

Genetically Tunable Enzymatic C–H Amidation for Lactam Synthesis

Inha Cho[†], Zhi-Jun Jia[†], Frances H. Arnold*

Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, MC 210-41, Pasadena, California 91125, USA.

*Correspondence to: Email: frances@cheme.caltech.edu

[†]These authors contributed equally to this work.

Abstract:

A major challenge in carbon–hydrogen (C–H) bond functionalization is to have the catalyst control precisely where a reaction takes place. Here we report engineered cytochrome P450 enzymes that perform unprecedented enantioselective C–H amidation reactions and control the site selectivity to divergently construct β -, γ - and δ -lactams, completely overruling the inherent reactivities of the C–H bonds. The enzymes, expressed in *Escherichia coli* cells, accomplish this abiological carbon–nitrogen (C–N) bond formation via reactive iron-bound carbonyl nitrenes generated from nature-inspired acyl-protected hydroxamate precursors. This transformation is exceptionally efficient (up to 1,020,000 total turnovers) and selective (up to 25:1 regioselectivity and 96% enantiomeric excess), and can be performed easily on preparative scale.

One Sentence Summary:

Directed evolution of cytochrome P450 enzymes fine-tunes site selectivity of new-to-nature C–H amidation for modular, sustainable and scalable preparation of enantio-enriched β -, γ - and δ -lactams.

Main Text:

Functionalization of unactivated carbon–hydrogen (C–H) bonds represents one of the most sought-after strategies to construct and diversify molecules, dramatically reshaping the logic of synthetic chemistry (1-4). Since organic molecules often contain multiple, similar C–H bonds, a major challenge is to control the site at which a reaction takes place. Current approaches to target specific C–H bonds rely mainly on substrate control, wherein the substrate guides the catalyst with directing groups (5, 6) or bears one C–H bond that is inherently more reactive (7-11). The former strategy necessitates extra synthetic steps to manipulate directing groups, while both strategies significantly limit the breadth of potential substrates. An appealing strategy is to use a modular catalyst platform that could be tuned to deliver different site selectivities, thereby enabling regiodivergent C–H functionalization of a broad range of substrates. However, such strategies remain elusive (12, 13).

Enzymes, the catalytic machinery of the biological world, exert exquisite control over selectivity in biochemical transformations (14). This is exemplified by cytochrome P450 enzymes, which can hydroxylate a specific C–H bond within a complex molecule (15, 16). Furthermore, this site selectivity can be tuned by protein engineering techniques like directed evolution (17-19). In

the past few years, our group and others expanded the boundaries of biocatalytic C–H functionalization by repurposing native cytochrome P450s to perform new-to-nature carbene and nitrene transfer reactions (20-22). These enzymes have not yet been engineered to control the site selectivity of the non-natural reactions comparable to what has been achieved with P450-catalyzed hydroxylation.

It was recently shown that C–H insertion of sulfonyl nitrenes can be catalyzed by variants of cytochrome P450_{BM3} having the axial heme-ligating cysteine amino acid residue substituted with serine (so-called cytochrome P411s) (23, 24). Carbonyl nitrenes are synthetically more useful than sulfonyl nitrenes, but are generally considered ineffective for C–H functionalization due to their low stability and facile decomposition to isocyanates through a well-known Curtius-type rearrangement (25). Recently, Chang and co-workers reported that an iridium-based catalytic system could harness carbonyl nitrenes for intramolecular C–H amidation; however, these catalysts were not enantioselective and only formed the thermodynamically favored products (26). We hypothesized that cytochrome P450s and other hemoproteins could be engineered to generate reactive carbonyl nitrenes, direct them to C–H amidation over competing decomposition pathways, and precisely control the enantioselectivity and site selectivity of the reaction.

Rather than using the carbonyl analogues of the sulfonyl azides to generate enzyme-bound nitrenes, we sought a more easily accessible and biocompatible nitrene precursor. A recent report by Tsutsumi and coworkers identified a naturally occurring nitrene-transfer reaction catalyzed by a cytochrome P450 enzyme in which the nitrene is generated from an acyl-protected hydroxylamine (27). We reasoned that acyl-protected hydroxamates **1**, which are readily prepared from carboxylic acids in one step (**Fig. 1A**), could act as suitable nitrene sources in new enzymatic reactions.

Motivated by the extensive pharmaceutical applications of β -lactams (28), we challenged hemoproteins to construct these strained, four-membered rings through intramolecular C–H amidation, an unprecedented reaction (29-31). We chose as the model substrate hydroxamate **1a** with two sets of reactive C–H bonds, potentially affording β -lactam **2a** through C(sp^3)–H amidation or δ -lactam **3a** through C(sp^2)–H amidation (**Fig. 1B**). We first tested whether free heme could catalyze C–H amidation in biocompatible conditions and observed alkyl amine **4** as the sole product, presumably from Curtius-type rearrangement of the carbonyl nitrene followed by hydrolysis (**Table S1**). We then tested a panel of hemoprotein variants, including P450s, P411s, cytochromes *c*, and globins. Most evaluated hemoproteins, including the P411 variants previously engineered for C–H insertion of sulfonyl nitrenes (23, 24), only generated amide **5**, presumably through reduction of the carbonyl nitrene intermediate. A few hemoproteins of various lineages, however, produced small amounts of the lactam products (**Tables S1 and S2**). In particular, P411variant **E10-V78F S438A (E10FA)**, which was originally engineered to catalyze carbene transfer to alkynes (32), was able to transform substrate **1a** to the desired β -lactam (**2a**), giving only 1.5% yield but 90% enantiomeric excess (*ee*). The δ -lactam **3a** was also detected, with 1.0% yield. In stark contrast to the reaction catalyzed by free heme, amine **4** was not observed in the reaction using this hemoprotein, supporting our hypothesis that an enzyme could divert highly unstable carbonyl nitrenes to the desired C–H amidation and bypass the Curtius-type rearrangement.

We chose E10FA as the parent for directed evolution of *lactam synthases* (LS). To improve catalytic activity as well as site- and enantio-selectivity of C–H amidation for the synthesis of β -lactam **2a**, amino acid residues located near the heme iron in a closely related crystal structure

(P411 **E10**) were targeted for sequential rounds of site-saturation mutagenesis and screening (**Fig. 1C**). In each round of mutagenesis, the libraries of enzymes were expressed and screened as whole-cell catalysts in 96-well plates for increased product formation and/or enantioselectivity. Two rounds of site-saturation mutagenesis and screening introduced mutations F78L and A264G, leading to a 28-fold improvement in total turnover number (TTN) to 52,000 for β -lactam **2a** while almost eliminating formation of δ -lactam **3a** (**Fig. 1D**). Double site-saturation mutagenesis and screening subsequently added mutations T327R and V328M, which further increased TTN to 71,000 and *ee* to 96%. Another round of double site-saturation led to the discovery of the final variant LS_{sp3} containing two more mutations L263N and A268F, which affords the desired β -lactam **2a** with 223,000 TTN (96% yield) and 96% *ee*. Adding both the *E. coli* expressing LS_{sp3} and precursor **1a** portion-wise to a reaction mixture boosted the TTN for β -lactam **2a** to 1,020,000 (see supplementary materials for details).

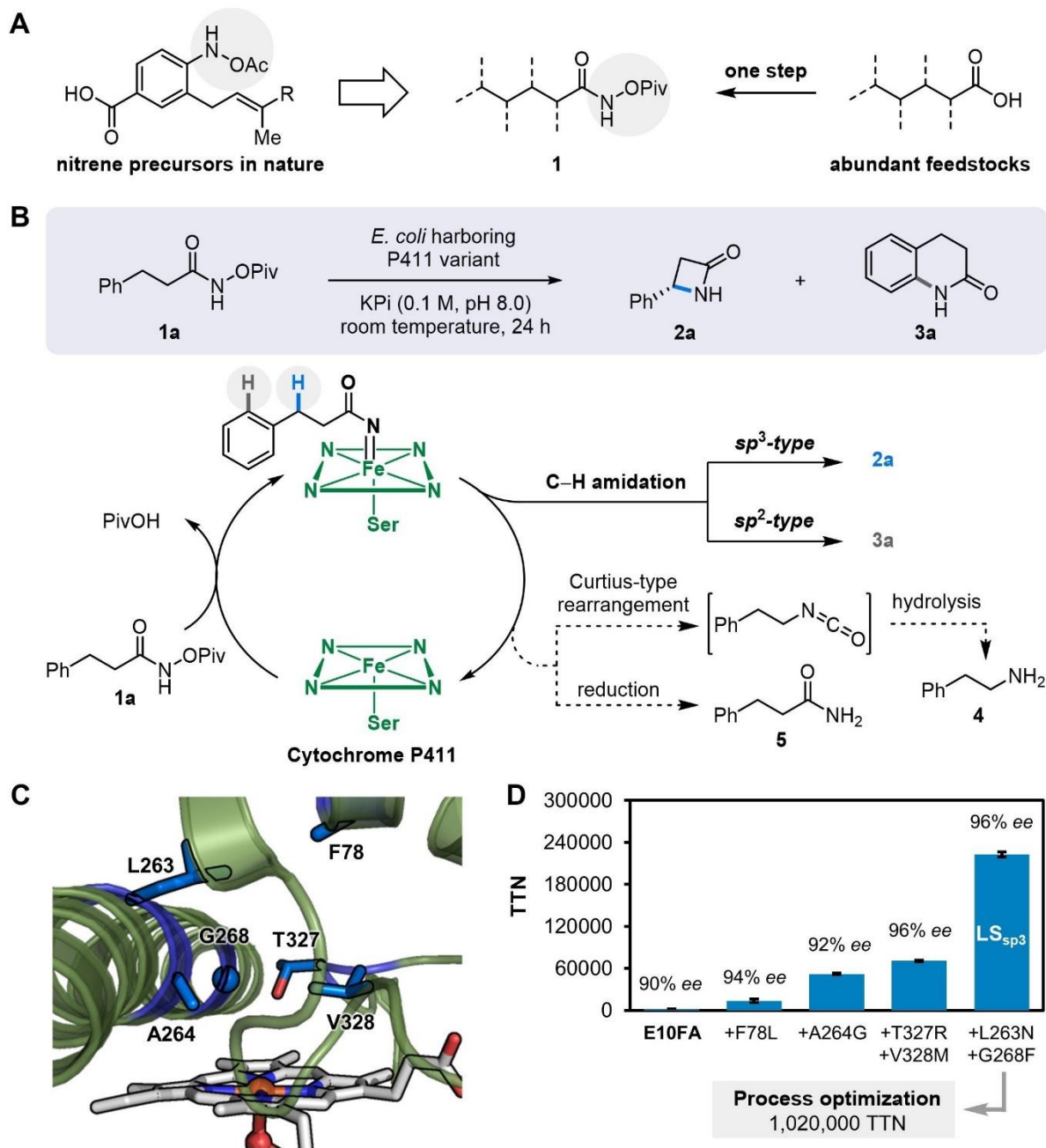


Fig. 1. Hemeprotein-catalyzed intramolecular C–H amidation. (A) Design of acyl-protected hydroxamate **1** as nitrene precursor, inspired by naturally occurring nitrene precursors in the biosynthesis of benzastatins (**27**) (Ac, acetyl; Piv, pivaloyl). (B) Reaction scheme and proposed catalytic cycle of intramolecular C–H amidation catalyzed by P411 variant. (C) Crystal structure of a variant closely related to P411 **E10** (PDB ID: 5UCW), with mutated residues marked in blue. (D) Directed evolution of **E10FA** to **LS_{sp3}** for the synthesis of β -lactam **2a** and further improvement of total turnover number (TTN) via process optimization. Experiments were typically performed at analytical scale using suspensions of *E. coli* cells expressing **E10FA** variants [optical density at 600 nm (OD_{600}) = 2.5] in KPi (0.1 M, pH 8.0) buffer, 50 mM substrate **1a**, 5 vol % organic cosolvents, and 400 μ L reaction volume at room temperature under anaerobic conditions for 36

hours. See supplementary materials for further details. Single-letter amino acid abbreviations (here or in Fig. 3): A, Ala; E, Glu; F, Phe; G, Gly; L, Leu; M, Met; N, Asn; P, Pro; R, Arg; T, Thr; Y, Tyr.

With variant LS_{sp3} in hand, we next evaluated a collection of hydroxamates for enantioselective β -lactam synthesis (**Fig 2A**). Various aryl substituents were accepted to furnish β -lactams **2b–2g** with excellent TTN (up to 180,000) and *ee* (up to 94%); only trace amount of δ -lactams through C(*sp*²)-H amidation were detected. Introducing two methyl groups adjacent to the carbonyl group did not hamper the reaction, leading to the formation of product **2h** in 191,000 TTN. Olefin-containing substrates could also undergo highly enantioselective C–H amidation (**2i, 2j**) with no competing aziridination observed, highlighting the enzyme's exquisite chemoselectivity. An α -alkynyl C–H bond was also readily amidated (**2k**). Substrates with heteroatom substituents, such as amide and silyl groups, reacted well to give desired products **2l** and **2m**.

The reactive C–H bonds in all the preceding substrates were weakened by adjacent substituents. This is not required, however, as evidenced by the ester-substituted β -lactam **2n**, which was formed with 111,000 turnovers. Substrates with secondary and tertiary aliphatic C–H bonds were also accepted, leading to products **2o–2p**. The pristine site selectivity of the enzyme was further highlighted by the formation of product **2q** from a substrate with multiple similar aliphatic C–H bonds. Even aliphatic C–H bonds adjacent to a trifluoromethyl group were amenable to C–H amidation, affording product **2r** with 58,000 turnovers and 88% *ee*. The transformation could be applied to preparative-scale synthesis without sacrificing yield or enantioselectivity (**Fig. 2B**). Isolation of most products was also straightforward: product **2s**, for example, was simply filtered from the aqueous reaction mixture, resulting in 86% yield (1.62 g, 92% HPLC purity) and 92% *ee*.

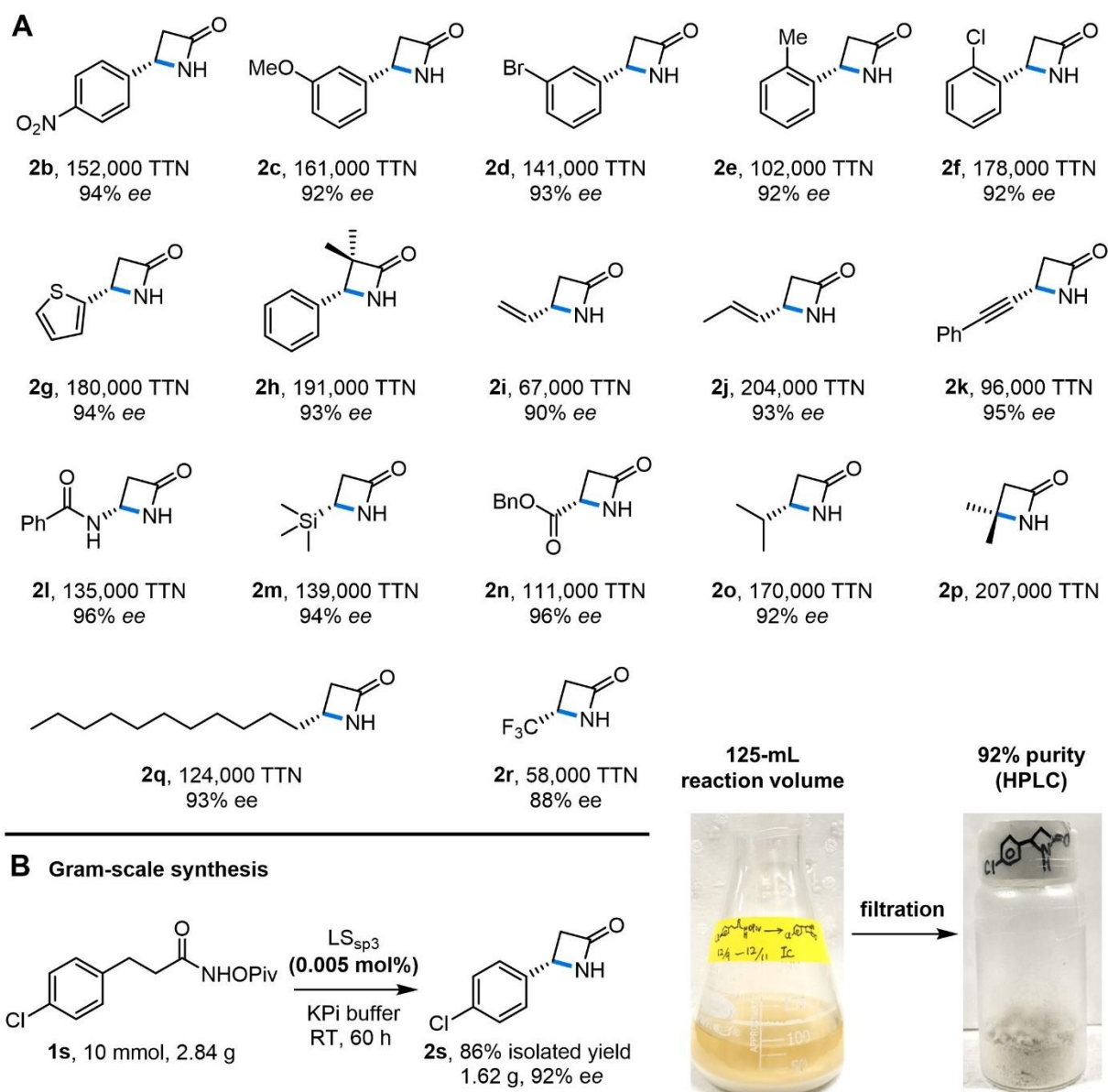


Fig. 2. Scope of β -lactam products. (A) Synthesis of β -lactams from corresponding hydroxamate substrates using $\text{LS}_{\text{sp}3}$. Experiments were typically performed at analytical scale using $\text{LS}_{\text{sp}3}$ -expressing *E. coli* cells resuspended to $\text{OD}_{600} = 2.5$ or 3.0 in KPi (0.1 M, pH 8.0) buffer, 25 mM or 50 mM substrate **1**, 5 vol % organic cosolvents, and 400 μL reaction volume at room temperature under anaerobic conditions for 36 hours. (B) Gram-scale synthesis of β -lactam **2s** using cell lysate of *E. coli* expressing $\text{LS}_{\text{sp}3}$. See supplementary materials for further details. In all cases, regioisomeric ratio (*r.r.*) > 20:1.

Having achieved β -lactam synthesis by intramolecular $\text{C}(\text{sp}^3)\text{-H}$ amidation, we next asked whether other types of C-H bonds could be amidated selectively. The parent enzyme E10FA showed trace reactivity towards $\text{C}(\text{sp}^2)\text{-H}$ bonds, but only two rounds of site-saturation mutagenesis and screening generated variant $\text{LS}_{\text{sp}2}$, with mutations F78A and T327P, that exclusively functionalized $\text{C}(\text{sp}^2)\text{-H}$ bonds (**Fig. 3A**). Instead of forming the corresponding β -

lactams (**2a**, **2s–2u**), LS_{sp2} diverted the substrates to the corresponding δ -lactams (**3a**, **3s–3u**) with excellent TTNs (**Fig. 3B**).

To further demonstrate the tunability of enzyme-catalyzed C–H amidation, we synthesized precursor **1v** with two sets of C(*sp*³)–H bonds, potentially affording β -lactam **2v** or γ -lactam **6v** (**Fig. 3C**). Variant LS_{sp3} gave β -lactam **2v** as the major product, together with **6v** in 2.2:1 ratio. Sequential site-saturation mutagenesis on residues R327, P329 and A330 and screening delivered variant LS _{β} , which produces β -lactam **2v** with *r.r.* up to 25:1. Bearing mutations R327P, P329L, and A330V from LS_{sp3}, LS _{β} provided 180,000 TTN and 96% *ee* for **2v**. As shown in the synthesis of β -lactams **2w–2y**, this variant generally enforces high selectivity for the β -lactam products, despite the presence of weaker benzylic C–H bonds and the option to form less-strained γ -lactams. Meanwhile, variant LS _{γ} was also revealed, which has mutations P329V A330M and can selectively synthesize γ -lactam **6v** (**Fig. 3C**). The site selectivity of LS _{γ} was largely insensitive to aromatic substitution, such that γ -lactams **6w–6y** were afforded with up to 25:1 *r.r.*, 209,000 TTN, and 97% *ee* (**Fig. 3C**).

Finally, we challenged the enzymes with nitrene precursor **1z**, which has three sets of reactive C(*sp*³)–H bonds. Variants LS _{β} , LS _{γ} , and LS_{sp3} selectively formed β -lactam **2z**, γ -lactam **6z** and δ -lactam **3z**, respectively. The excellent site selectivity shown in this example further showcases the powerful regiocontrol of the enzymes that is tunable by directed evolution (**Fig. 3D**).

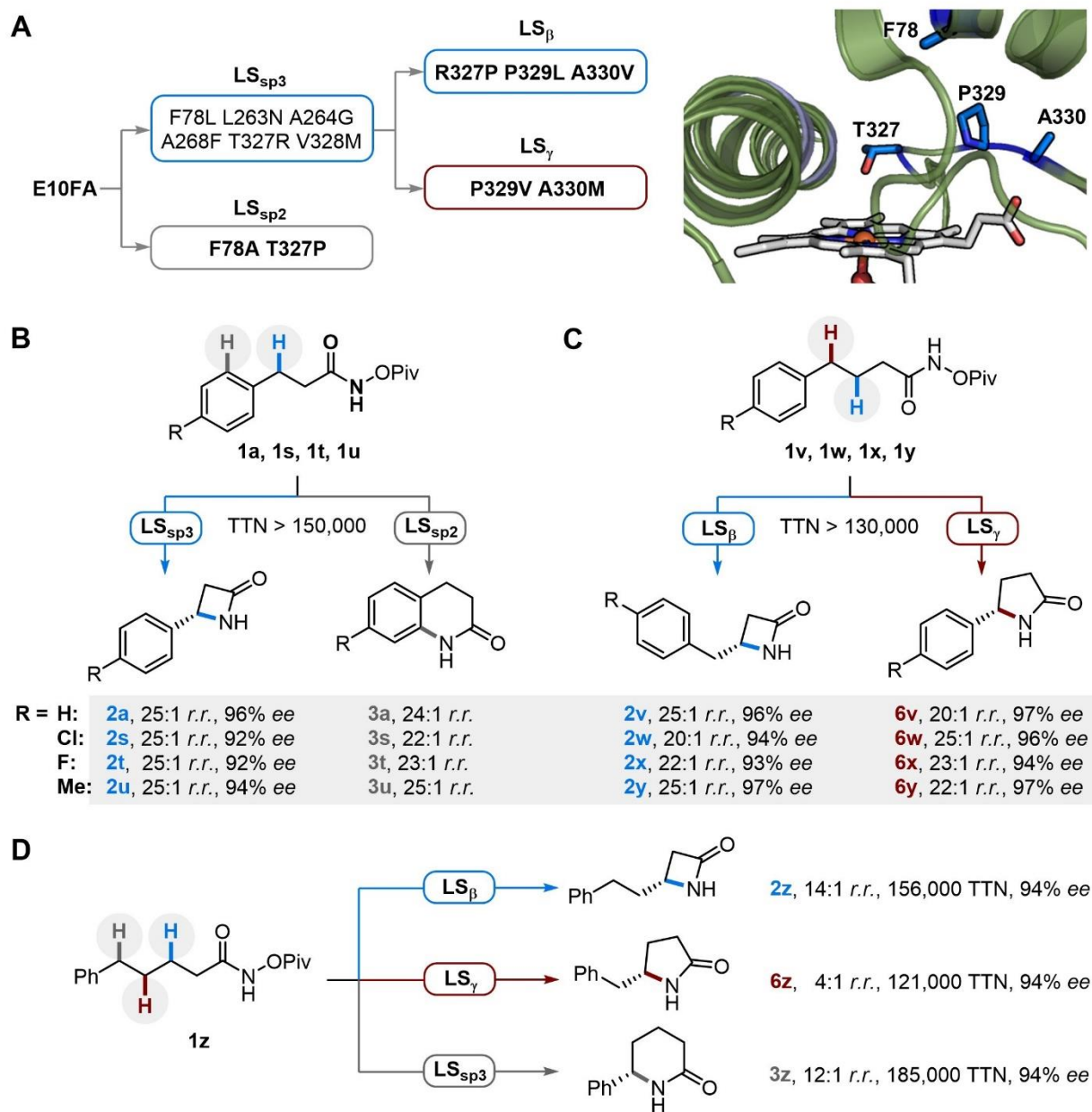


Fig. 3. Engineering lactam synthases for regiodivergent intramolecular C–H amidation. (A) Evolutionary trajectory of variants LS_{sp2}, LS_{sp3}, LS_β, and LS_γ. (B) Selectivity and scope of LS_{sp2} and LS_{sp3}-catalyzed intramolecular C–H amidation. (C) Selectivity and scope of LS_β and LS_γ. (D) Regiodivergent amidation of aliphatic, homobenzylic, and benzylic C(*sp*³)-H bonds catalyzed by LS_β, LS_γ, and LS_{sp3}. Regioisomeric ratio (*r.r.*) indicates the mole ratio of major product to combined minor regioisomers. See supplementary materials for details.

We have engineered a new ‘lactam synthase’ that can be tuned by directed evolution to convert substrates into different lactams through a catalyst-controlled C–H amidation process. Reactivity trends due to bond strength, inductive effects, steric accessibility, or ring strain could be completely overturned in this catalyst-controlled process. Our findings suggest that genetically

tunable enzymatic catalysis may provide a general strategy to address the challenge of site selectivity in C–H functionalization.

References and Notes:

1. K. Godula, D. Sames, C–H Bond Functionalization in Complex Organic Synthesis. *Science* **312**, 67-72 (2006).
2. W. R. Gutekunst, P. S. Baran, C–H Functionalization Logic in Total Synthesis. *Chem. Soc. Rev.* **40**, 1976-1991 (2011).
3. J. Yamaguchi, A. D. Yamaguchi, K. Itami, C–H Bond Functionalization: Emerging Synthetic Tools for Natural Products and Pharmaceuticals. *Angew. Chem. Int. Ed.* **51**, 8960-9009 (2012).
4. H. M. L. Davies, D. Morton, Recent Advances in C–H Functionalization. *J. Org. Chem.* **81**, 343-350 (2016).
5. K. M. Engle, T.-S. Mei, M. Wasa, J.-Q. Yu, Weak Coordination as a Powerful Means for Developing Broadly Useful C–H Functionalization Reactions. *Acc. Chem. Res.* **45**, 788-802 (2012).
6. F. Zhang, D. R. Spring, Arene C–H Functionalisation Using a Removable/Modifiable or a Traceless Directing Group Strategy. *Chem. Soc. Rev.* **43**, 6906-6919 (2014).
7. M. P. Doyle, R. Duffy, M. Ratnikov, L. Zhou, Catalytic Carbene Insertion into C–H Bonds. *Chem. Rev.* **110**, 704-724 (2010).
8. E. M. Simmons, J. F. Hartwig, Catalytic Functionalization of Unactivated Primary C–H Bonds Directed by an Alcohol. *Nature* **483**, 70-73 (2012).
9. J. D. Cuthbertson, D. W. C. MacMillan, The Direct Arylation of Allylic sp^3 C–H Bonds via Organic and Photoredox Catalysis. *Nature* **519**, 74-77 (2015).
10. G. J. Choi, Q. Zhu, D. C. Miller, C. J. Gu, R. R. Knowles, Catalytic Alkylation of Remote C–H Bonds Enabled by Proton-Coupled Electron Transfer. *Nature* **539**, 268-271 (2016).
11. T. Newhouse, P. S. Baran, If C–H Bonds Could Talk: Selective C–H Bond Oxidation. *Angew. Chem. Int. Ed.* **50**, 3362-3374 (2011).
12. K. Liao *et al.*, Design of Catalysts for Site-Selective and Enantioselective Functionalization of Non-Activated Primary C–H Bonds. *Nat. Chem.* **10**, 1048-1055 (2018).
13. L. Ping, D. S. Chung, J. Bouffard, S.-g. Lee, Transition Metal-Catalyzed Site- and Regio-Divergent C–H Bond Functionalization. *Chem. Soc. Rev.* **46**, 4299-4328 (2017).
14. S. J. Benkovic, S. Hammes-Schiffer, A Perspective on Enzyme Catalysis. *Science* **301**, 1196-1202 (2003).
15. P. R. Ortiz de Montellano, Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chem. Rev.* **110**, 932-948 (2010).
16. X. Huang, J. T. Groves, Oxygen Activation and Radical Transformations in Heme Proteins and Metalloporphyrins. *Chem. Rev.* **118**, 2491-2553 (2018).
17. R. Fasan, Tuning P450 Enzymes as Oxidation Catalysts. *ACS Catalysis* **2**, 647-666 (2012).

18. A. R. H. Narayan *et al.*, Enzymatic Hydroxylation of an Unactivated Methylene C–H Bond Guided by Molecular Dynamics Simulations. *Nat. Chem.* **7**, 653-660 (2015).
19. J.-b. Wang, G. Li, M. T. Reetz, Enzymatic Site-Selectivity Enabled by Structure-Guided Directed Evolution. *Chem. Comm.* **53**, 3916-3928 (2017).
20. O. F. Brandenburg, R. Fasan, F. H. Arnold, Exploiting and Engineering Hemoproteins for Abiological Carbene and Nitrene Transfer Reactions. *Curr. Opin. Biotechnol.* **47**, 102-111 (2017).
21. R. K. Zhang, X. Huang, F. H. Arnold, Selective C–H Bond Functionalization with Engineered Heme Proteins: New Tools to Generate Complexity. *Curr. Opin. Chem. Biol.* **49**, 67-75 (2019).
22. R. K. Zhang *et al.*, Enzymatic Assembly of Carbon–Carbon Bonds via Iron-Catalysed sp^3 C–H Functionalization. *Nature* **565**, 67-72 (2019).
23. T. K. Hyster, C. C. Farwell, A. R. Buller, J. A. McIntosh, F. H. Arnold, Enzyme-Controlled Nitrogen-Atom Transfer Enables Regiodivergent C–H Amination. *J. Am. Chem. Soc.* **136**, 15505-15508 (2014).
24. C. K. Prier, R. K. Zhang, A. R. Buller, S. Brinkmann-Chen, F. H. Arnold, Enantioselective, Intermolecular Benzylic C–H Amination Catalysed by an Engineered Iron-Haem Enzyme. *Nat. Chem.* **9**, 629-634 (2017).
25. S. Bräse, C. Gil, K. Knepper, V. Zimmermann, Organic Azides: An Exploding Diversity of a Unique Class of Compounds. *Angew. Chem. Int. Ed.* **44**, 5188-5240 (2005).
26. S. Y. Hong *et al.*, Selective Formation of γ -Lactams via C–H Amidation Enabled by Tailored Iridium Catalysts. *Science* **359**, 1016-1021 (2018).
27. H. Tsutsumi *et al.*, Unprecedented Cyclization Catalyzed by a Cytochrome P450 in Benzastatin Biosynthesis. *J. Am. Chem. Soc.* **140**, 6631-6639 (2018).
28. R. Fernandes, P. Amador, C. Prudêncio, β -Lactams: Chemical Structure, Mode of Action and Mechanisms of Resistance. *Rev. Med. Microbiol.* **24**, 7-17 (2013).
29. M. Cherest, X. Lusinchi, A Novel Electrophilic *N*-Amidation via Electron Deficient Complexes: Action of Ferric Chloride on *N*-Acetyloxyamides. *Tetrahedron Lett.* **30**, 715-718 (1989).
30. S. Hosseyni, A. Jarrahpour, Recent Advances in β -Lactam Synthesis. *Org. Biomol. Chem.* **16**, 6840-6852 (2018).
31. C. R. Pitts, T. Lectka, Chemical Synthesis of β -Lactams: Asymmetric Catalysis and Other Recent Advances. *Chem. Rev.* **114**, 7930-7953 (2014).
32. K. Chen, X. Huang, S. B. J. Kan, R. K. Zhang, F. H. Arnold, Enzymatic Construction of Highly Strained Carbocycles. *Science* **360**, 71-75 (2018).

Acknowledgments: We thank K. Chen, S. Brinkmann-Chen, D. C. Miller and D. K. Romney for comments on the manuscript; X. Huang for helpful discussions; K. Chen for the parent variant E10FA; The Caltech Center for Catalysis and Chemical Synthesis and the Caltech Mass Spectrometry Laboratory for analytical support. **Funding:** Supported by NSF Division of Molecular and Cellular Biosciences grant MCB-1513007; Joseph J. Jacobs Institute for Molecular

Engineering for Medicine (I.C.); and Deutsche Forschungsgemeinschaft (JI 289/1-1 to Z.-J.J.). **Author contributions:** I.C. and Z.-J.J. designed the research and performed the experiments with guidance from F.H.A.. I.C., Z.-J.J. and F.H.A. wrote the manuscript. **Competing interests:** Caltech has filed a provisional patent application (CIT-8187-P); and **Data and materials availability:** All data are available in the main text or the supplementary materials.