RESEARCH ARTICLE SUMMARY

ENZYME DESIGN

Computational design of serine hydrolases

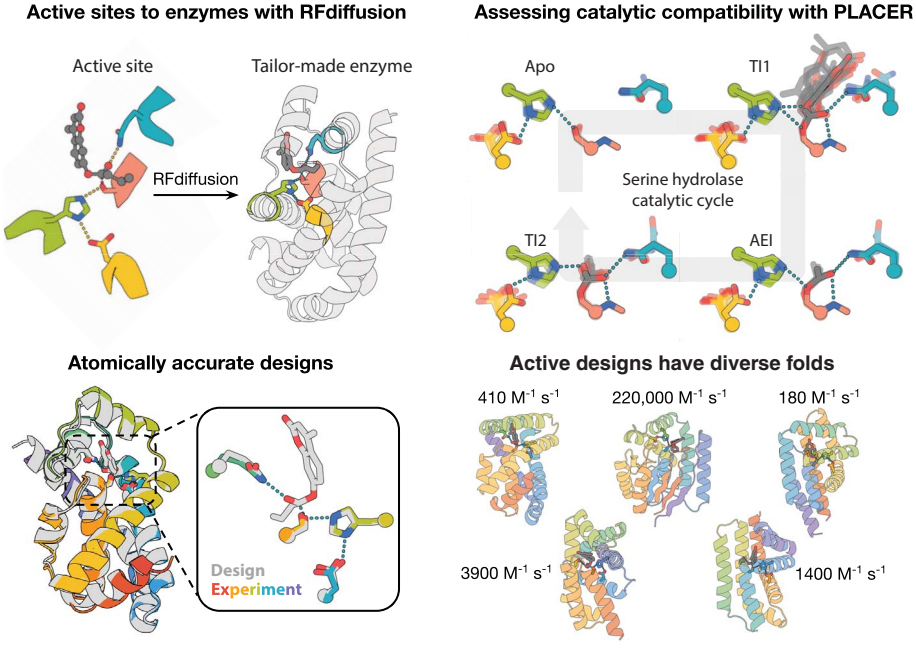
Anna Lauko†, Samuel J. Pellock*†, Kiera H. Sumida†, Ivan Anishchenko, David Juergens, Woody Ahern, Jihun Jeung, Alexander F. Shida, Andrew Hunt, Indrek Kalvet, Christoffer Norn, Ian R. Humphreys, Cooper Jamieson, Rohith Krishna, Yakov Kipnis, Alex Kang, Evans Brackenbrough, Asim K. Bera, Banumathi Sankaran, K. N. Houk, David Baker*

INTRODUCTION: Enzymes efficiently accelerate the chemical reactions that sustain all living systems using complex and atomically precise molecular architectures that are extremely challenging to design. Previous efforts to design enzymes have largely focused on finding geometric matches between model active sites and preexisting protein structures, an approach akin to buying a suit from a thrift store; it is unlikely the fit will be perfect. Most previous enzyme design efforts have also primarily focused on simplified active sites designed for a single state, even for enzymes that use multistep reaction mechanisms. These design limitations have limited the geometric precision and preorganization of the active site, key features of highly efficient enzymes. New approaches that generate protein structures tailored to a complex enzyme active site and assess the structural compatibility of the resulting design with each state along the reaction coordinate are needed to improve enzyme design.

RATIONALE: We reasoned that rather than searching for geometric matches between designed active sites and fixed protein backbones, we could instead leverage deep learning–based generative models to construct proteins that are tailor-made for the desired active site. Once generated, we could use deep learning–based ensemble modeling to structurally assess the compatibility of a designed enzyme with each step in a catalytic cycle. We reasoned that ester hydrolysis by serine hydrolases would serve as an excellent model reaction because the reaction mechanism requires a complex active site made up of intricate hydrogen bond networks to activate the serine nucleophile that can accommodate multiple transition states and intermediates throughout the reaction cycle, including the acyl-enzyme intermediate (AEI) and two tetrahedral intermediates (TI1, TI2).

RESULTS: We used RFdiffusion denoising trajectories starting from completely random residue distributions to generate protein backbones housing active sites that span a range of catalytic residue compositions and geometries and assigned amino acid sequences to these backbones using LigandMPNN. Experimental characterization of the resulting designs revealed multiple serine hydrolases with high structural accuracy to the predicted models (Cα root mean square deviations <1 Å between design and crystal structures), catalytic efficiencies as high as 2.2×10^5 M^{-1} s^{-1}, and folds distinct from natural hydrolases. Designs filtered for structural compatibility and preorganization across the catalytic cycle using a newly developed ensemble modeling method, PLACER, had considerably improved success rates and catalytic activities. Characterization of designs with increasingly complex active sites revealed structural and compositional features essential for catalysis: minimal active sites composed of Ser-His dyads yielded proteins containing reactive serines capable of covalent labeling with phosphonate probes and ester substrates, but the full Ser-His-Asp triad and two-component oxyanion hole are necessary for catalytic turnover. Rather than performing rounds of directed evolution, we were able to improve catalysis through explicit optimization of catalytic geometries. The RFdiffusion calculations sample a considerable range of folds and structures, enabling an analysis of the geometric determinants of catalysis that complements structural studies of native serine hydrolases.

CONCLUSION: We demonstrate that RFdiffusion can scaffold complex active sites with higher accuracies, success rates, and catalytic activities than previous computational methods. Deep learning–based assessment of active site preorganization at each step in a multistep reaction can considerably increase design success rates for enzymes with complex catalytic mechanisms. The enzyme structure generation and active site evaluation methods described here should be broadly applicable to the design of multistep enzymes for a wide variety of chemical reactions. ■



Computational design of serine hydrolases. RFdiffusion generates enzymes tailor-made for a given active site (top left). Assessing design compatibility across the catalytic cycle with the PLACER neural net. Conformational ensembles for each step of the reaction are shown (top right). Comparison with crystal structures shows that designs made with RFdiffusion and filtered with PLACER are atomically accurate (bottom left). Designed serine hydrolases expand the fold space of this ancient enzyme family (bottom right, catalytic efficiencies listed).

The list of author affiliations is available in the full article online.
*Corresponding author. Email: spellock@uw.edu (S.J.P.); dabaker@uw.edu (D.B.)
†These authors contributed equally to this work.
Cite this article as A. Lauko *et al.*, *Science* **388**, eadu2454 (2025). DOI: 10.1126/science.adu2454

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adu2454

RESEARCH ARTICLE

ENZYME DESIGN

Computational design of serine hydrolases

Anna Lauko[1,2,3]†, Samuel J. Pellock[1]*†, Kiera H. Sumida[1,2,4]†, Ivan Anishchenko[1,2], David Juergens[1,2,5], Woody Ahern[1,2,6], Jihun Jeung[1,2,5], Alexander F. Shida[1,2,5,7], Andrew Hunt[1,2], Indrek Kalvet[1,2,8], Christoffer Norn[1,2], Ian R. Humphreys[1,2], Cooper Jamieson[9], Rohith Krishna[1,2], Yakov Kipnis[1,2], Alex Kang[1,2], Evans Brackenbrough[1,21], Asim K. Bera[1,2], Banumathi Sankaran[1,2], K. N. Houk[9], David Baker[1,2,8]*

The design of enzymes with complex active sites that mediate multistep reactions remains an outstanding challenge. With serine hydrolases as a model system, we combined the generative capabilities of RFdiffusion with an ensemble generation method for assessing active site preorganization at each step in the reaction to design enzymes starting from minimal active site descriptions. Experimental characterization revealed catalytic efficiencies (k_{cat}/K_m) up to 2.2 × 10^5 M^{-1} s^{-1} and crystal structures that closely match the design models (Cα root mean square deviations <1 angstrom). Selection for structural compatibility across the reaction coordinate enabled identification of new catalysts remove with five different folds distinct from those of natural serine hydrolases. Our de novo approach provides insight into the geometric basis of catalysis and a roadmap for designing enzymes that catalyze multistep transformations.

Enzymes are powerful catalysts that substantially accelerate reaction rates in mild aqueous conditions. The ability to construct enzymes that catalyze arbitrary chemical reactions would have enormous utility across a wide range of applications, and hence, enzyme design has been a long-standing goal of computational protein design (*1*). De novo enzyme design has generally started from a specification of arrangements of catalytic residues around the reaction transition state (a theozyme) and sought to identify placements of this active site in preexisting scaffolds (*2*–*7*). Using fixed backbones restricts how accurately the catalytic geometry can be realized and has likely limited the activities of many designed enzymes to date prior to optimization by laboratory evolution, as recent studies of designed Kemp eliminases demonstrate (*8*–*10*). A further challenge of enzyme design is the preorganization of the active site such that the catalytic functional groups are accurately positioned relative to the transition state. Achieving preorganization is especially difficult for multistep reaction mechanisms because the enzyme must preferentially stabilize multiple transition states and intermediates, and current methods to evaluate design preorganization in silico are limited by low accuracy or computational cost (*7*, *11*–*14*). To enable the accurate design of multistep enzymes, new methods are needed for both the generation of proteins housing a given active site and the assessment of their structural compatibility with each step in the reaction.

Ester hydrolysis has served as a model reaction for computational enzyme design for decades (*15*–*20*), and justifiably so: Numerous mechanisms can be used for ester hydrolysis, enabling a range of distinct design approaches to target this reaction; activity is easily monitored by absorbance and fluorescence with reporter substrates; and esterases are highly valuable in industrial processes, most recently for their application in plastic recycling (*21*–*23*). The textbook example of enzymatic ester hydrolysis is the double-displacement reaction mechanism used by serine hydrolases, in which a serine nucleophile undergoes acylation to form the acyl-enzyme intermediate (AEI) that is subsequently hydrolyzed by an activated water. Despite extensive structural, mutational, and computational characterization of the mechanism of serine hydrolases found in nature (*24*–*35*), de novo design efforts attempting to use this machinery have been unsuccessful, and to our knowledge, no previous efforts have successfully constructed a serine hydrolase that extends beyond the fold space found in nature.

A major challenge in designing serine hydrolases is overcoming the stability of the AEI, the resolution of which is typically rate-limiting when activated esters are used. Numerous previously designed enzymes and peptide-based systems inactivate or substantially slow down after acylation (*6*, *15*–*17*). In addition to this chemical challenge, constructing the serine hydrolase active site combines some of the most difficult current challenges in protein design: (i) The catalytic site is very complex, requiring the scaffolding of at least four individual residues with atomic precision, a task that state-of-the-art design tools struggle to achieve (*36*); (ii) the serine nucleophile requires activation by construction of intricate hydrogen bond networks; and (iii) the active site must undergo subtle conformational changes throughout the multistep catalytic cycle, and although there has been recent progress in multistate design (*37*, *38*), it remains challenging, particularly when the energetic difference between desired states is small.

Previous efforts to design esterases have circumvented the challenges presented by serine hydrolases by using simpler, more easily designable active sites, leveraging nucleophiles more activated than serine, and by targeting reaction mechanisms that do not require the formation of stable covalent intermediates. For example, previously designed metallohydrolases skip the AEI by activating water to cleave esters in a single step (*18*, *39*); the noncanonical amino acid N_δ-methylhistidine has been employed to make the AEI less stable (*17*); and cysteine has been used in place of serine because of its greater nucleophilicity (*6*, *15*). Structural analysis of the resulting cysteine esterases indicated key interactions between the cysteine nucleophile and histidine base of the desired dyad or triad were not formed (*6*, *15*), suggesting that the inherent chemical reactivity of the residues employed, not their coordinated effort, may have been responsible for the observed steady-state rate enhancements. Even with these chemical interventions, the efficiency of the initial computational designs remain far below the range observed for natural enzymes.

One hypothesis for the lack of designed serine hydrolases to date is a potential geometric incompatibility between the complex hydrolase active site and the sets of fixed protein scaffold libraries previously used (*6*). We investigated whether increasing scaffold diversity could help identify backbones that more accurately reconstruct the desired active site and carried out a preliminary design campaign searching for placements of a serine hydrolase active site in a large library of scaffolds based on the nuclear transport factor 2 (NTF2) fold (*40*) (fig. S1 and Materials and methods, "NTF2 design campaign"). As in previous studies (*7*), experimental characterization of the resulting designs revealed activated serines but no catalytic turnover on ester substrates, despite a close match between the experimental and designed structures (fig. S2). We suspected that an inability to install

[1]Department of Biochemistry, University of Washington, Seattle, WA, USA. [2]Institute for Protein Design, University of Washington, Seattle, WA, USA. [3]Graduate Program in Biological Physics, Structure and Design, University of Washington, Seattle, WA, USA. [4]Department of Chemistry, University of Washington, Seattle, WA, USA. [5]Graduate Program in Molecular Engineering, University of Washington, Seattle, WA, USA. [6]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA. [7]Medical Scientist Training Program, University of Washington, Seattle, WA, USA. [8]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. [9]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA, USA.
*Corresponding author. Email: spellock@uw.edu (S.J.P.); dabaker@uw.edu (D.B.)
†These authors contributed equally to this work.

key catalytic features into NTF2s, such as the backbone oxyanion hole contact common to all serine hydrolases, limited the function of these designs.

We reasoned that advances in deep learning for protein design could enable the design of proteins de novo to directly scaffold the serine hydrolase active site and assess design compatibility for the entire multistep catalytic cycle. Recent advances in scaffolding functional sites with RFdiffusion have yielded improved in silico and experimental success rates across a range of design tasks (*36*, *41*, *42*); we aimed to use the same approach to generate serine hydrolases starting from geometric descriptions of an active site (Fig. 1A). To assess preorganization and functional interactions in each step of the catalytic cycle, we sought to

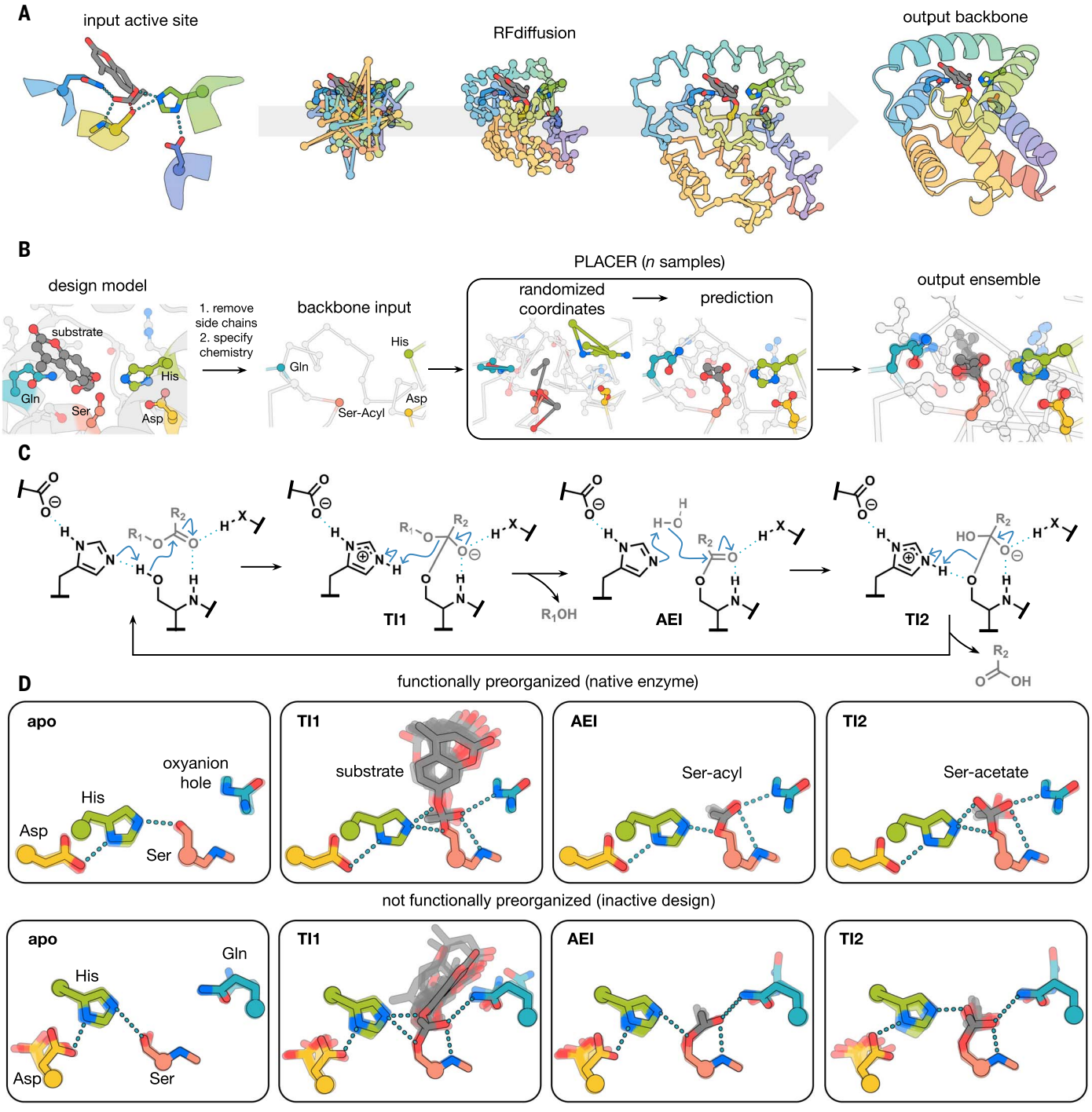


Fig. 1. Design methods. (**A**) Active site-specific backbone generation with RFdiffusion. Given the geometry of a possible active site configuration, RFdiffusion denoising trajectories generate backbone coordinates, which scaffold the site. (**B**) Generation of active site ensembles with PLACER. The coordinates of the side chains around the active site and any bound small molecule for the step in the reaction being considered are randomized, and *n* samples are carried out to generate an ensemble of predictions. (**C**) Mechanism of ester hydrolysis by serine hydrolases. (**D**) PLACER ensembles for distinct states along the reaction coordinate for hydrolysis of 4MU-Ac for a native serine hydrolase (top, PDB: 1IVY) and an inactive designed serine hydrolase from round 3 (bottom, josie).

leverage advances in deep learning–based prediction of protein–small molecule complexes by modeling structural ensembles of catalytic intermediates (Fig. 1B).

Assessing reaction path compatibility with PLACER

We set out to understand why previously designed serine hydrolases failed to appreciably catalyze ester hydrolysis and hypothesized that modeling each step of the reaction could be critical for assessing the ability of a design to achieve catalytic turnover. To model the extent to which a designed enzyme can stabilize each of the key states along the reaction coordinate and to assess the preorganization of the active site residues in the desired catalytic geometries, we developed a deep neural network that, given (i) the backbone coordinates of a small-molecule binding pocket or active site, (ii) the identities of the amino acid residues at each position, and (iii) the chemical connectivity of bound small molecules (but not their positions), generates the full atomic coordinates of the binding site, comprising both protein side chains and small molecules. We trained this network, called PLACER (protein-ligand atomistic conformational ensemble resolver) (*43*), on protein–small molecule complexes in the Protein Data Bank (PDB) by randomizing the atomic coordinates of side chains and small molecules within spherical regions with up to 600 heavy atoms, and seeking to minimize a loss function assessing the recapitulation of the atomic coordinates within the region. In benchmark tests, PLACER predicted regions within native structures with an average root mean square deviation (RMSD) of 1.1 Å. PLACER is stochastic, and repeated runs from different random seeds yield an ensemble of models for the predicted region (Fig. 1B).

We used PLACER to generate structural ensembles for each step of the catalytic cycle for a set of native and previously designed serine hydrolases. The catalytic cycle of serine hydrolases can be divided into four steps (Fig. 1C). First, the substrate binds to the apoenzyme (apo) and the catalytic serine, deprotonated by the catalytic histidine, attacks the carbonyl carbon of the ester to form the first tetrahedral intermediate (TI1). Second, the catalytic histidine protonates the leaving-group oxygen promoting its departure, leaving the active site serine covalently linked to the acyl group of the substrate (the AEI mentioned above). Third, the histidine deprotonates a water molecule, which attacks the AEI to generate a second tetrahedral intermediate (TI2). Finally, this intermediate is resolved by histidine-mediated protonation of serine and release of the acyl group, reconstituting the free enzyme and completing the catalytic cycle. Throughout, negatively charged transition states and intermediates are stabilized by at least two hydrogen bond donors that constitute the oxyanion hole. Perturbation of the histidine pK_a, which tunes its acid–base function, is mediated by interaction with aspartate or glutamate, the final residue in the triad (*44*–*46*).

Modeling this catalytic cycle with PLACER showed that native serine hydrolases are more preorganized than previous designed systems (Fig. 1D and fig. S3). At each step in the reaction, the catalytic residues sample the key hydrogen bonds essential for catalysis more often in native than previously designed serine hydrolases (fig. S3). Because the reaction rate should be proportional to the fraction of the enzyme in the active state, poor preorganization of the designed active sites is expected to compromise catalysis. To quantify the extent of active site formation in PLACER ensembles, we compute the frequency of formation of key interactions between the catalytic functional groups and reaction intermediates over each step of the reaction (see Materials and methods, "Filtering" section) and use this metric to assess new designs described in the following sections.

Design and characterization of serine hydrolases

We next set out to design proteins with active sites of increasing complexity, using RFdiffusion to scaffold serine hydrolase active site motifs and PLACER to assess their preorganization in each step of the reaction (Fig. 2, A and B). We designed catalysts for the hydrolysis of 4-methylumbelliferone (4MU) esters (Fig. 2C) that fluoresce upon hydrolysis. To generate active site motifs, we sampled positions of the catalytic side chains based on a density functional theory (DFT)–optimized transition state (see Materials and methods, "Motif generation") and an analysis of natural hydrolases (*33*), and enumerated α-helix and β-strand backbone conformations for each catalytic residue, keeping the interactions with the transition state fixed in space. For each combination of the backbone N, Cα, and C atoms for each of the catalytic residues, we used RFdiffusion to build up backbones starting from random noise that have coordinates that nearly exactly match the input catalytic residue backbone positions (average all-atom RMSD ~0.1 Å) and form a binding pocket for the substrate (see Materials and methods, "Motif generation" and "Backbone generation" sections). To drive folding to the designed state and to make favorable interactions with the substrate and active site residues, LigandMPNN (*47*) was used to design the sequence. Rosetta FastRelax (*48*) was used to refine the protein backbone and ligand pose, and sequence design with LigandMPNN was repeated with the new backbone as input (*49*). After three cycles of LigandMPNN and FastRelax, the structures of the designs were predicted with AlphaFold2 (AF2) (*50*), and designs for which all catalytic residue Cα atoms were positioned within 1.0 Å of the design models were selected for experimental characterization (see Materials and methods, "Sequence design" and "Filtering" sections for details).

In the first two rounds of design, we built relatively simple active sites consisting of Ser-His dyads with a single oxyanion hole contact from the backbone amide of the serine (Fig. 2, A and B), and explicitly evaluated the utility of PLACER to select designs for experimental characterization. Round 1 designs were filtered with AF2 alone, whereas round 2 designs that passed the AF2 filter were selected for experimental screening if PLACER ensembles of the apo state indicated that the key Ser-His hydrogen bond was formed (see Materials and methods, "Filtering" section; only 1.6% of round 2 designs that passed the AF2 RMSD filter were predicted to be preorganized by PLACER). For experimental testing, we obtained synthetic genes encoding 129 and 192 designs for rounds 1 and 2, respectively, for *Escherichia coli* overexpression and screening.

We used a fluorophosphonate (FP) activity-based probe and fluorescent 4MU-acetate (4MU-Ac) and 4MU-butyrate (4MU-Bu) ester substrates to identify designs with activated serines and esterase activity, respectively (Fig. 2C). The fraction of designs labeled by the FP probe in *E. coli* lysate increased fivefold from 3 to 17% from round 1 to round 2 (Fig. 2B and fig. S4). Designs that reacted with the FP probe were purified and incubated with 4MU esters, and two round 1 designs (1.6%) and 10 round 2 designs (5.2%) showed 4MU ester hydrolysis. Retrospective PLACER analysis of the round 1 designs revealed that the Ser-His H-bonds in the two catalytically active designs were predicted to be among the most preorganized (fig. S5). PLACER filtering of round 2 designs on the extent of formation of the key Ser-His H-bond not only increased the fraction of designs exhibiting FP probe labeling and 4MU hydrolysis, but also resulted in higher activities (Fig. 2, E and F). The progress curves for these round 1 and 2 designs plateau after approximately one enzyme equivalent of fluorescent product is formed (Fig. 2E), suggesting that the serine acylates but that the resulting AEI fails to hydrolyze (*32*). When incubated with substrate, mass spectra of these designs revealed a mass shift corresponding to acylation, further supporting protein inactivation after formation of the acylated intermediate (fig. S6).

We hypothesized that incorporating a histidine-stabilizing catalytic acid and a second oxyanion hole H-bond donor in a third round of designs (round 3) and filtering for PLACER preorganization in both the apo and AEI states could generate designs capable of catalytic turnover through hydrolysis of the AEI. For round 3 designs, we required all catalytic triad and oxyanion hole H-bonds to be highly preorganized in PLACER ensembles of both the apo and AEI

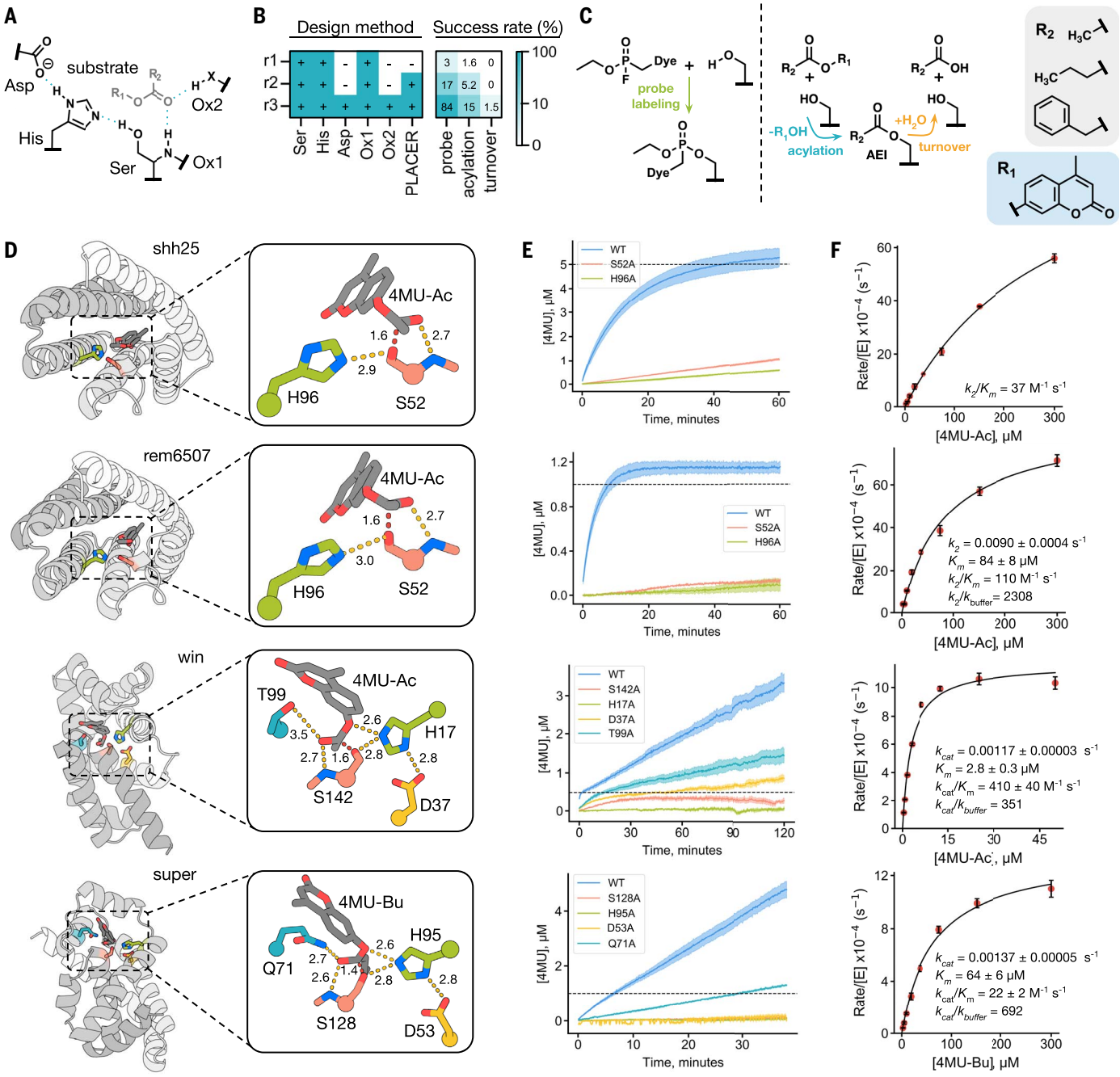


Fig. 2. Functional characterization of designed serine hydrolases. (**A**) Chemical schematic of a serine hydrolase active site. (**B**) Summary of design method and experimental success rate for probe labeling, single-turnover acylation, and catalytic turnover for each design round. (**C**) Chemical schematic depicting probe labeling, acylation, and catalytic turnover. (**D**) Fold (left) and active site (right) of serine hydrolase design models. (**E**) Reaction progress curves for the parent design and catalytic residue knockouts. Dashed line represents the enzyme concentration and shaded areas represent standard deviation of three technical replicates. (**F**) Michaelis-Menten plots derived from initial (shh25, rem6507) or steady-state velocities (win, super). Error bars represent standard deviation of three technical replicates.

states. Of 132 round 3 designs, 111 (84%) displayed FP probe labeling, 20 hydrolyzed 4MU substrates (15%), and two designs (1.5%) displayed catalytic turnover (Fig. 2, B and E). Active designs from all three rounds showed reduced activity upon mutation of any one of the catalytic residues (Ser, His, Asp/Glu, and oxyanion side-chain contact) (Fig. 2E), showing that the observed activities are dependent on the designed active site. To determine the kinetic parameters of the active designs, initial or steady-state rates were measured to determine k_2/K_m or k_{cat}/K_m for single-turnover and multiple-turnover designs, respectively (Fig. 2F and fig. S7). For the two designs that displayed catalytic turnover, called "super" and "win," k_{cat}/K_m values were 22 M^{-1} s^{-1} (k_{cat} = 0.00137 ± 0.00005 s^{-1}, K_m = 64 ± 6 μM) and 410 M^{-1} s^{-1} (k_{cat} = 0.00117 ± 0.00003 s^{-1}, K_m = 2.8 ± 0.3 μM), respectively, for the more preferred of the two 4MU substrates (win and super preferentially hydrolyzed 4MU-Ac and 4MU-Bu, respectively (fig. S8)). Despite the low K_m observed for win, we were unable to reach saturation of the initial

burst phase of the reaction by increasing substrate concentration up to 100 μM (fig. S9), suggesting that K_s >> K_m and that the low apparent K_m observed for win is a result of rapid acylation and not tight substrate binding.

Structural characterization of designed serine hydrolases

We pursued x-ray crystallography to determine the accuracy with which super and win were designed. We were able to solve crystal structures of both super and win, and found that they had very low Cα RMSDs of 0.8 Å over 165 residues and 0.83 Å over 160 residues (Fig. 3, A and D), respectively, to the design models. The design accuracy extends to the geometry of the active site: the side-chain conformations of the catalytic residues are in atomic agreement for super (all-atom RMSD = 0.38 Å over 22 atoms) and for win (all-atom RMSD = 0.86 Å over 20 atoms) except for a rotamer shift in the side-chain oxyanion contact, Thr^{99} (Fig. 3, B and E). In the active site of super, a water molecule sits above the nucleophilic serine and forms hydrogen bonds with the oxyanion hole contacts, which likely mimics the positioning of the carbonyl oxygen of the ester substrate (Fig. 3B). Similarly, in win, an acetate molecule is positioned at the catalytic center and hydrogen bonds to the catalytic serine (Ser^{142}), the side-chain oxyanion hole (Thr^{99}), and the histidine acid-base residue (His^{17}) (Fig. 3E).

Although the structures were solved in the absence of bound small-molecule substrate or transition state analog, overlay of the design model and crystal structure of super reveals high shape complementarity to the butyrate acyl group of its preferred substrate (Fig. 3C and fig. S8). At the same time, the 4MU moiety is largely exposed, corroborating the selectivity of super for 4MU-Bu over 4MU-Ac and suggesting that substrate binding, in this case, is largely driven by binding to the acyl group. For win, a rotamer shift in Phe^{98} in the crystal structure would clash with the butyrate moiety, and indeed, win is selective for the smaller substrate 4MU-Ac that avoids this clash (Fig. 3F and fig. S8).

The structures of super and win are very different from known structures; the closest matches found from Foldseek searches against all databases have template modeling (TM) scores of 0.52 and 0.46 for super and win, respectively (near or below the 0.5 cutoff below which structures are considered to have different topological folds), are proteins of unknown function, and have no similarity to known hydrolases at the fold or active site level (fig. S10, A and B), demonstrating that the design method used here yields structural solutions for serine hydrolase activity that extend well beyond those found in nature, expanding the structural space of this ancient enzyme family.

Filtering for preorganization across the reaction coordinate improves catalysis

We next sought to generate and compare designs filtered explicitly with PLACER for preorganization over two states (apo and AEI) or over all four states of the reaction path by carrying out additional iterations of LigandMPNN and FastRelax starting from the active design win (fixing only the identities of the four catalytic residues) (Fig. 4A and fig. S1). We obtained genes encoding 45 two-state filtered designs for experimental characterization, all of which were diverse in sequence compared to the original designs (mean sequence identity to the parent design of 58% and 61% within the active site), and found 38 (84%) labeled with FP probe (fig. S11A), and 9 (20%) displayed activity over background in a lysate screen (fig. S11C). Three of these, win1, win11, and

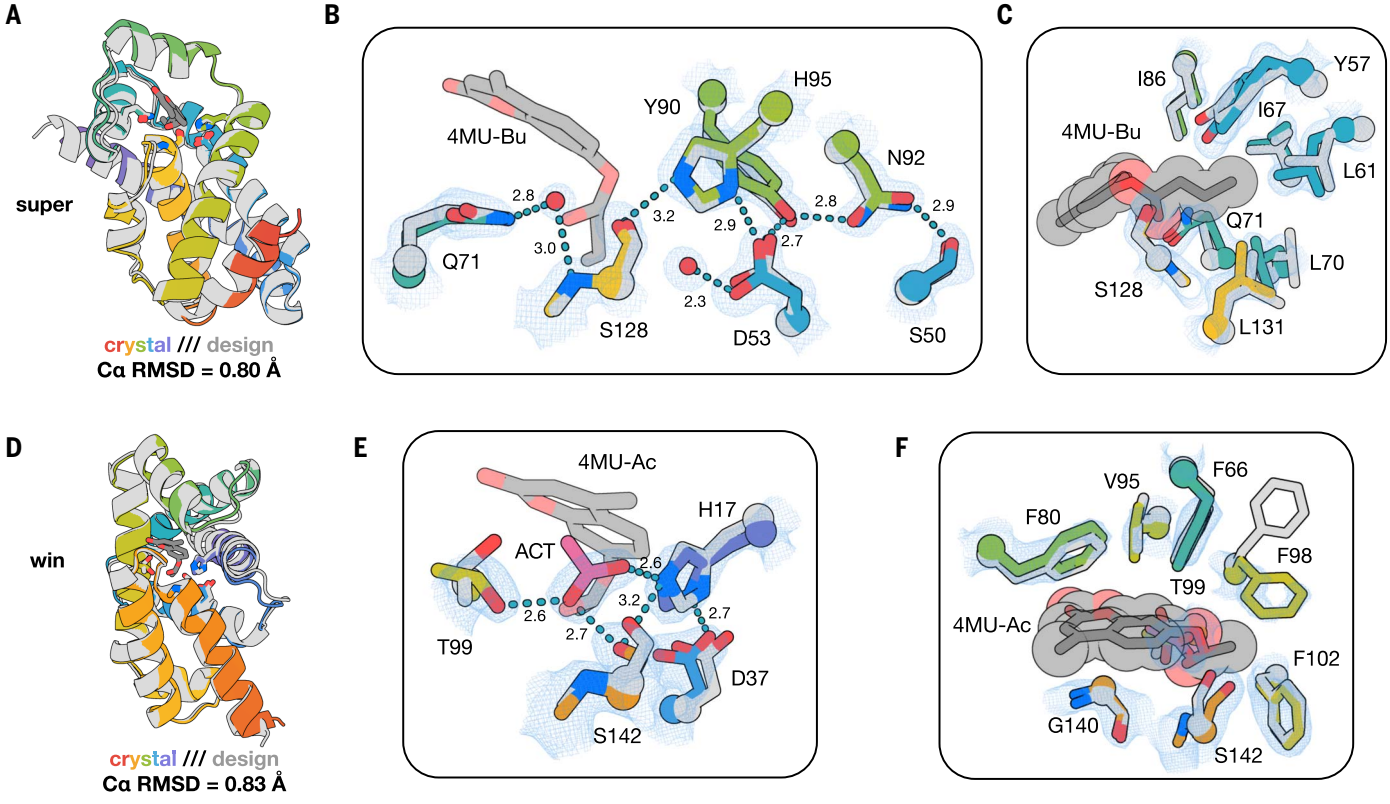


Fig. 3. Structural characterization of designed serine hydrolases. (**A** and **D**) Structural superposition of design models (gray) and crystal structures (rainbow) for super (A) and win (D). (**B** and **E**) Active site overlays of design models (gray) and crystal structures (rainbow) of super (B) and win (E) with 2*Fo* − *Fc* map shown at 1σ (blue mesh). (**C** and **F**) Superposition of substrate binding sites of the design models (gray) and crystal structures (rainbow) of super (C) and win (F) with 2*Fo* − *Fc* map shown at 1 σ (blue mesh). Distances shown in angstroms. Single-letter abbreviations for the amino acid residues are as follows: D, Asp; F, Phe; G, Gly; H, His; I, Ile; L, Leu; N, Asn; Q, Gln; S, Ser; T, Thr; V, Val; and Y, Tyr.

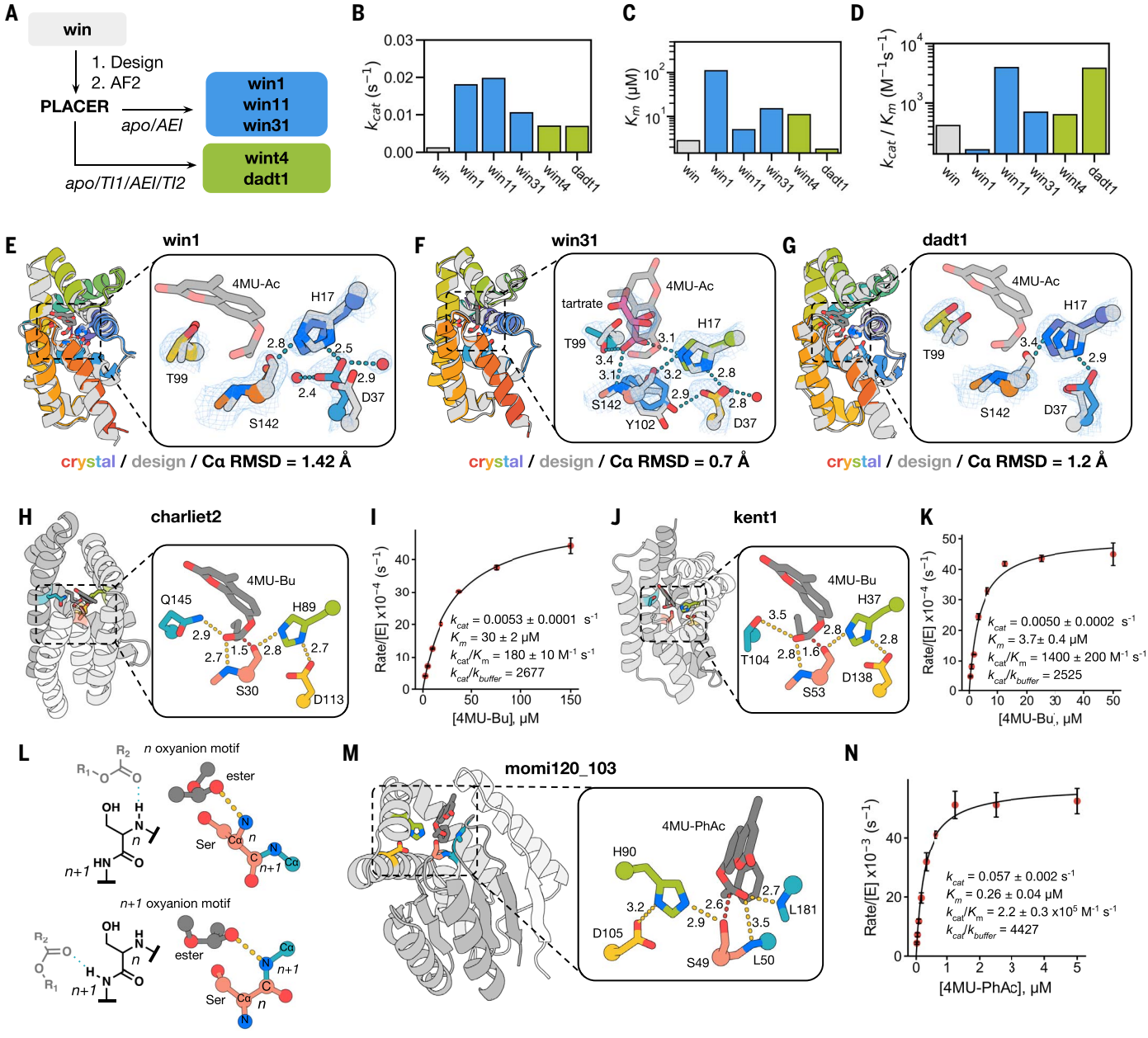


Fig. 4. Computational redesign and more complex folds improve catalysis. (**A**) Computational pipeline for redesign of win. (**B** to **D**) k_{cat} (B), K_m (C), and k_{cat}/K_m (D) of parent win compared to computational redesigns. (**E** to **G**) Structural superposition of design model and crystal structure of win1 (E), win31 (F), and (G) dadt1 with 2*Fo* − *Fc* map shown at 1σ. (**H** to **K**) Design models (H and J) and Michaelis-Menten plots (I and K) for active designs with distinct folds. (**L**) Chemical and structural comparison of *n* and *n+1* oxyanion hole motifs. (**M**) Chai-1 prediction of momi120_103 in complex with 4MU-PhAc. (**N**) Michaelis-Menten plot for momi120_103 with 4MU-PhAc. Error bars represent standard deviation of three technical replicates.

win31, displayed higher catalytic turnover compared to the starting design: win has a k_{cat} of 0.00117 s^{-1}, which increases 15-fold in win1 (0.018 s^{-1}), 17-fold in win11 (0.0197 s^{-1}), and 9-fold in win31 (0.0105 s^{-1}) (Fig. 4B and fig. S7). Of the 11 four-state filtered designs tested, 10 (91%) labeled with FP-probe (fig. S11B) and 8 (73%) displayed activity (fig. S11D). Two of these, dadt1 and wint4, displayed higher catalytic efficiencies than win, with k_{cat}/K_m values of 3800 and 640 M^{-1} s^{-1}, driven by increases to k_{cat} and decreases in K_m relative to win (Fig. 4, B to D, and fig. S7). Catalytic triad residue knockouts for all designs showed reductions in activity, and for win11 and win31, mutation of stabilizing residues in the second shell of the active site that H-bond to the catalytic aspartate also reduced activity (fig. S12). The two redesigns with the highest k_{cat} values (win1 and win11) do not display burst phase kinetics, suggesting that deacylation is no longer rate-limiting (fig. S7).

We determined the crystal structures of win1, win31, and dadt1, and comparison to the design models revealed Cα RMSDs of 1.42, 0.7, and 1.2 Å, respectively (Fig. 4, E to G). For win1, the active site closely matches the designed architecture (mean all-atom RMSD = 0.54 Å) (Fig. 4E), and Thr^{99}, the oxyanion hole contact, occupies the designed rotamer, which may account for the 15-fold increase in k_{cat} compared to win, in which Thr^{99} is rotated relative to the designed rotamer (Fig. 3E). In chain B of the win1 structure, the catalytic serine partially occupies a second conformer with an occupancy of 0.23 (fig. S13A). For win31,

five chains are present in the asymmetric unit, all of which closely match the design model (average Cα RMSD = 0.7 Å) at the backbone level (Fig. 4F and fig. S13B). Analysis of the active site across all chains in the asymmetric unit revealed mobility in the catalytic serine, side-chain oxyanion threonine, and a second shell tyrosine (fig. S13C), but overall a very close match to the design model active site with a mean all-atom RMSD of 0.7 Å. Tartrate, derived from the crystallization solution, fit the electron density present in the active site of all five chains and forms hydrogen bonds with the serine, histidine, and oxyanion hole contacts (Fig. 4F), likely mimicking key contacts used throughout the catalytic cycle. For dadt1, the active site closely matches the design model with a mean all-atom RMSD of 0.95 Å, and the T99 side-chain oxyanion residue occupies the designed conformation.

We next explored whether stringent PLACER filtering for optimal catalytic geometry and preorganization across the reaction coordinate could generate active esterases with new backbone topologies and active site geometries. We performed sequence design and PLACER filtering for the complete reaction coordinate on round 3 backbones excluding win (fig. S1), and of 20 designs tested, two (charliet2 and kent1) displayed esterase activity, with catalytic efficiencies of 180 and 1400 M^{-1} s^{-1} (Fig. 4, H to K), suggesting that structural variability in intermediate states of the reaction coordinate may have limited otherwise functional designs. We also used sequence design combined with PLACER filtering to modify the substrate selectivity of win1, converting it from accepting only the small acyl group of 4MU-Ac to processing the larger 4MU-phenylacetate (4MU-PhAc) substrate (fig. S14).

To test the generality of RFdiffusion combined with PLACER filtering, we applied it to a different active site configuration in which the oxyanion hole consists of two backbone amides, rather than a backbone amide and a side-chain H-bond donor, and where the first backbone amide of the oxyanion hole is the residue following the catalytic serine (*N+1*) rather than the catalytic serine itself (*N*) as in the previous designs (Fig. 4L, the N+1 position is more widely used than the N position in naturally occurring esterase oxyanion holes). We used the RFdiffusion and LigandMPNN/FastRelax design pipeline to generate 66 designs for this new catalytic site and the larger 4MU-PhAc substrate (fig. S1). The most active of these, momi, displayed a k_{cat}/K_m of 1240 M^{-1} s^{-1} and a k_{cat} of 0.1 s^{-1}, a fivefold faster rate than win11, the previous best design in terms of turnover number. The distribution of folds generated by RFdiffusion for this active site geometry differed from that of the original geometry, with more α/β fold solutions (as in the case of momi), showing how the RFdiffusion buildup approach crafts overall protein structure topology to the specific active site of interest. Natural esterases, to our knowledge, exclusively use the momi *N+1* oxyanion hole motif, suggesting that it is particularly well suited for ester hydrolysis. The high activity achieved without any prior experimental characterization for this new catalytic site shows that filtering for preorganization across the reaction cycle can yield new catalysts in one shot.

Several experimental results identify areas to address for improved function. First, kent1 inactivates after roughly 10 turnovers, and mass spectra of the catalyst and the serine knockout incubated with substrate reveal stable acylated species (fig. S15), indicating that designs that hydrolyze the AEI are still susceptible to inactivation, potentially from off-mechanism acylation events in the active site or acylation-induced conformational changes. Second, mutation of the side-chain oxyanion hole residue had variable effects on activity. In three designs (dadt1, charliet2, kent1) from design rounds 4 and 5 that underwent stringent PLACER filtering, mutation of the side-chain oxyanion hole residue had a modest effect on activity, suggesting limited contribution to catalysis (fig. S12). Analysis of the oxyanion hole geometries in these designs and others in earlier design rounds reveal in-plane hydrogen bonds to the oxygen of the substrate carbonyl (fig. S16 and supplementary text), in contrast to those found in nature, which are perpendicular to the plane of the carbonyl, where they likely stabilize the oxyanion transition state over the sp^2 carbonyl ground state (*34*, *51*, *52*).

We next explored whether existing designs could be improved by rebuilding suboptimal regions using RFdiffusion. Using the momi backbone as input to RFdiffusion, we built out the N terminus to further stabilize the active site but made no changes to the parent backbone or sequence (figs. S1 and S17). Of 65 designs tested, all showed activity, and one design, momi120, displayed a catalytic efficiency of 4300 M^{-1} s^{-1}, 3.5-fold greater than momi, driven by a twofold increase in k_{cat} and 1.5-fold decrease in K_m (fig. S17). We also used RFdiffusion to improve the suboptimal in-plane (with respect to the substrate carbonyl) oxyanion hole H-bond formed by Gln^{71} in super. The serine protease subtilisin uses a chemically similar side-chain oxyanion hole, Asn^{155}, with the amide positioned perpendicular to the plane of the substrate carbonyl (fig. S16A). Using the subtilisin oxyanion hole geometry as a guide, we mutated Gln^{71} to Asn in super, and repositioned it to form an analogous out-of-plane H-bond to the substrate carbonyl, then rebuilt the surrounding backbone of the protein with RFdiffusion to accommodate this change (fig. S18). Of the 150 designs screened, the two most active designs, superfast and supercool, showed 8-fold and 7-fold improvements in k_{cat} over the parent design super (k_{cat} = 0.00137 s^{-1}), and 19-fold and 13-fold improvements in k_{cat}/K_m, respectively (fig. S18). These results highlight productive design interventions made possible by RFdiffusion that are not easily accessible with traditional engineering tools like rational mutagenesis and directed evolution, where the sequence can be readily changed but not easily augmented with new structural features.

We redesigned momi120 for the hydrolysis of polyethylene terephthalate (PET) and screened 85 designs for activity on the sterically similar 4MU-PhAc substrate. All 85 designs displayed activity above background in a lysate screen, and two of the most active designs were further kinetically characterized and found to have $k_{cat}/K_m > 10^4$ M^{-1} s^{-1} (fig. S19). The most efficient design, momi120-103, has a k_{cat} for 4MU-PhAc of 0.057 s^{-1}, K_m of 0.26 μM, and a k_{cat}/K_m of 2.2×10^5 M^{-1} s^{-1} (Fig. 4N). PLACER and Chai-1 (*53*) predictions suggest that 4MU-PhAc fits with high shape complementarity into the redesigned pocket; the substitutions lining the binding pocket, particularly F76G (fig. S19), appear to provide a deeper pocket that may be the structural basis of the submicromolar K_m.

Structural determinants of catalysis

The high structural conservation of catalytic geometry in native serine hydrolases suggests that it is close to optimal for catalysis (*33*, *54*), but it is difficult to assess how activity depends on the detailed geometry of the interactions of the transition states with the catalytic serine, histidine, and oxyanion hole functional groups because although the identities of the catalytic residues can be readily changed by mutation, it is not straightforward to systematically vary side-chain or backbone geometries geometry. By contrast, our de novo buildup approach samples a wide range of catalytic geometries. To investigate how active site geometry and preorganization influence catalytic activity, we generated PLACER ensembles of all 812 experimentally characterized designs, categorized as inactive, FP probe labeling, acylation, and catalytic turnover, for each reaction step in the hydrolysis of 4MU-Ac (including design rounds 1 to 3 and previous NTF2-based designs). We summarize the strongest trends in the following paragraphs.

Increased preorganization and bending of the Ser-His H-bond were associated with higher rates of probe labeling, acylation, and turnover. All designs capable of catalyzing turnover displayed highly preorganized Ser-His H-bonds across all four states, whereas inactive designs often displayed rotamer shifts, causing loss of the interaction (Fig. 5, A and B). Designs that catalyzed turnover had Ser(Oγ):His(Nε-Cε) bond angles that were more acute (median, all states = 94°) than inactive designs (median, all states =

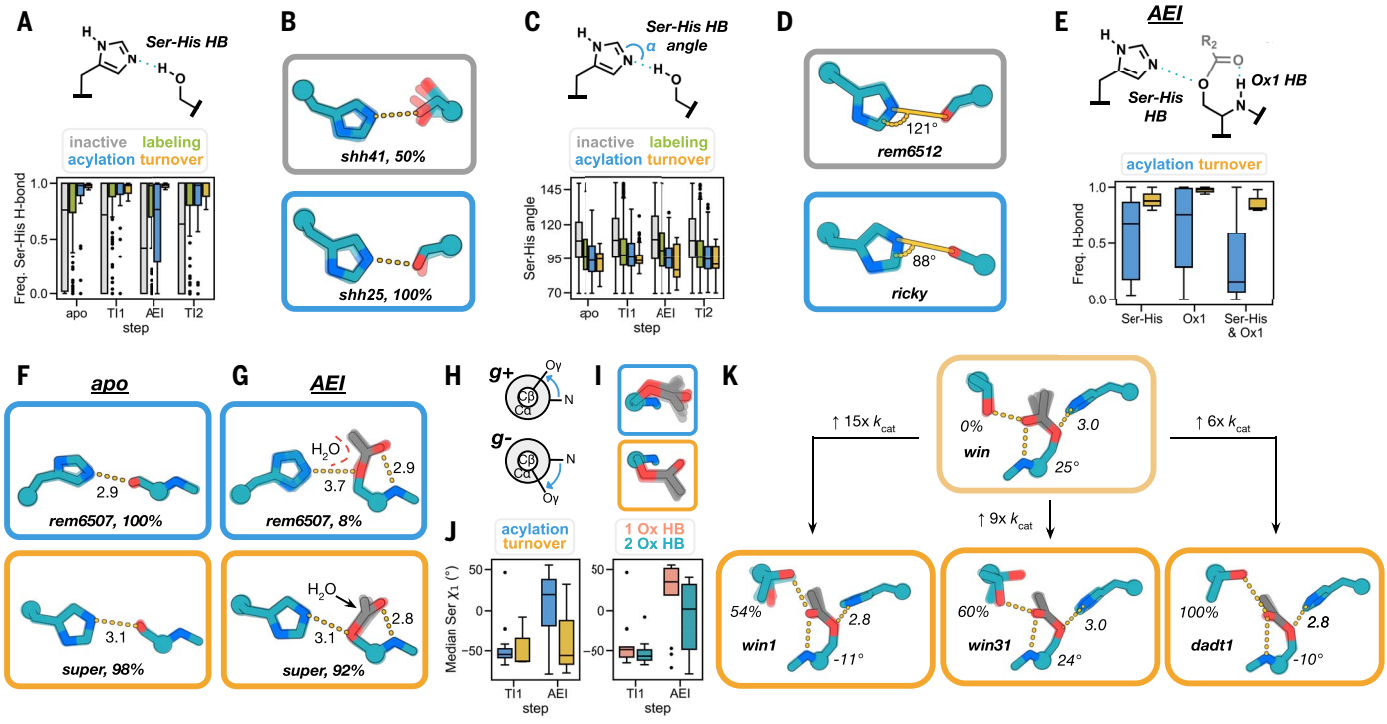


Fig. 5. PLACER ensembles reveal geometric determinants of catalysis. (**A**) Frequencies of catalytic Ser-His H-bond formation in PLACER ensembles for each reaction step, grouped by experimental outcome. (**B**) Apo PLACER ensembles of representative inactive (top) and acylating (bottom) designs. (**C**) Median angle (α) between serine Oγ, histidine Nϵ and Cϵ across PLACER ensembles for each reaction step, grouped by experimental outcome. (**D**) Apo PLACER ensembles of representative inactive (top) and acylating (bottom) designs; angle indicates median α. (**E**) AEI PLACER ensemble H-bond frequencies for designs that undergo acylation or full turnover. (**F**) PLACER ensembles of the apo state for an acylating (top) and multiple turnover design (bottom). (**G**) PLACER ensembles of the AEI state for a representative design that undergoes acylation (top) and a design that catalyzes turnover (bottom). Measurements shown represent median distances (angstroms) of key H-bonds indicated for each ensemble and percentages represent frequency of H-bond formation across all PLACER trajectories. (**H**) Newman projections of serine *g*+ and *g*- rotameric states (left). (**I**) PLACER ensembles of an acylating design (top) and a design that catalyzes turnover (bottom). (**J**) Median serine χ_1 angle across TI1 and AEI state PLACER ensembles for designs that catalyze acylation or turnover (left) and for the same designs grouped by number of oxyanion hole H-bonds. (**K**) AEI state PLACER ensembles for win, win1, win31, and dadt1, with percent of frames with correct oxyanion hole rotamer, Ser $\chi1$ angle, and catalytic Ser-His H-bond distance shown. Boxplots represent median, upper, and lower quartiles; whiskers extend 1.5 × IQR (interquartile range) above and below the upper and lower quartiles (respectively). Observations falling outside these ranges plotted as outliers.

108°), which were more similar to serine-histidine hydrogen bonds across the PDB (~125°) (*34*) (Fig. 5C). This acute H-bond is consistent with the reaction mechanism because this geometry allows histidine to participate, without changing conformation, in all of the necessary proton transfers involving serine, the leaving-group oxygen in TI1, and the hydrolytic water (*35*, *55*). This compromise in positioning is observed not only in our active designs but also in natural esterases (*34*, *55*, *56*).

The geometry of the serine rotamer throughout the catalytic cycle was also strongly correlated with experimental outcome. For designs that display acylation or turnover, we found that serine largely occupies the active *g*- rotamer (*54*) in the apo state. Designs that display turnover retain the *g*- serine conformer upon formation of the AEI, but designs that irreversibly acylate switch to the *g*+ rotamer in the AEI (Fig. 5, H to J). The *g*+ serine rotamer is catalytically incompetent in these designs because it leads to an acyl group conformation that occludes interaction of the hydrolytic water with histidine (Fig. 5G), increases the median Ser-His H-bond distance (Fig. 5G), and reduces the frequency with which the Ser-His and oxyanion hole-acyl group H-bonds form (Fig. 5E). The same retention of the *g*- rotamer in the AEI is observed in native crystal structures (*35*). PLACER analysis also revealed that the presence of a second oxyanion hole residue favors the active *g*- serine rotamer: those designs with only one oxyanion hole H-bond (from the backbone amide of the serine nucleophile) shift from *g*- to *g*+ upon acylation, and designs with two oxyanion hole H-bonds predominantly occupy *g*- Ser rotamers (Fig. 5J, right). The second oxyanion hole contact in serine hydrolases thus not only stabilizes the transition state but likely helps orient intermediates in catalytically productive conformations.

Differential preorganization may also explain activity trends in the win, win1, win31, and dadt1 series. PLACER analysis of the crystal structures of these designs revealed that in the AEI state, the more active redesigns win1, win31, and dadt1 sample the designed T99 oxyanion hole rotamer in 56, 60, and 100% of predictions, respectively, but the less active win never adopts this rotamer (Fig. 5K). Although both observed rotamers place T99 Oγ within hydrogen bonding distance of the oxyanion, the designed rotamer-oxyanion dihedral angle (91°) adopted by the redesigns much more closely matches the angles observed in native serine hydrolases, suggesting it is likely more optimal for selective transition state stabilization (*34*, *51*, *52*). We also observed differences in the serine rotameric state and the preorganization of the acyl group in the AEI state. Both win and win31 occupy the catalytically unfavorable *g*+ rotamer across the entire AEI ensemble, but win1 and dadt1 both display a less pronounced rotameric shift, which leads to shorter Ser-His H-bond distances (mean H-bond distance of 2.8 Å in win1 and dadt1 compared to 3.1 Å in win and win31). Overall, the acyl groups of win1 and especially win31 and dadt1 display less conformational

heterogeneity than that of win, which may increase the likelihood of histidine-mediated water attack (Fig. 5K).

Conclusion

The substantial catalytic efficiencies, the complexity of the active sites, and the atomic accuracy of the designs described here represent major advances in computational enzyme design. The serine catalytic triad plus oxyanion hole mechanism involves complex machinery that is challenging to scaffold [compared to, for example, the Kemp eliminase, which requires only a general base in a hydrophobic environment (*2*)], necessitates chemical activation of serine, and proceeds through a complex multistep mechanism that traverses a chemically stable AEI. The designed serine hydrolases described here have efficiencies up to 2.2×10^5 M^{-1} s^{-1}, a major improvement in function for computationally designed enzymes. For example, the previously designed esterase OE1 has a k_{cat}/K_m = 210 M^{-1} s^{-1} and reached an efficiency of 3190 M^{-1} s^{-1} after four rounds of directed evolution and screening over 12,000 clones, despite the use of a more activated N_δ-methylhistidine nucleophile (*17*). The closest comparable de novo design in terms of mechanism, in which a cysteine-based catalytic triad was mutated into a peptide-based helical barrel that proceeds via a more activated thioester intermediate (*15*), has a k_{cat}/K_m of 3.7 M^{-1} s^{-1} and k_{cat} of 0.0005 s^{-1}, 60,000 times less efficient and 400 times slower than the most efficient (momi120-103) and highest turnover design (momi120) described here, respectively. The ability to accelerate the hydrolysis of a chemically stable acyl-enzyme intermediate has been a decades-old challenge in enzyme design. To approximate the deacylation rate enhancement, we compared the uncatalyzed rate of hydrolysis of ethyl acetate (2.5 to 5.0) × 10^{-10} s^{-1} (*57*)) to the lower limit of the deacylation rate constant of momi (k_{cat}, 0.076 s^{-1}, pH 7.0, 25°C), yielding an estimated rate enhancement of over 10^8. Taken together, the design of serine hydrolases spanning five folds not represented in natural esterases, the considerable improvement in activity over previously designed esterases, and the acceleration of deacylation represent key advances in enzyme design.

The designs described here are not as efficient as native serine hydrolases with their cognate substrates (e.g., the k_{cat}/K_m of acetylcholinesterase with acetylcholine is >10^8 M^{-1} s^{-1}) (*58*), but they have efficiencies comparable to or better than natural proteases for activated esters (α-chymotrypsin with *p*-nitrophenyl acetate k_{cat}/K_m: 3530 M^{-1} s^{-1}, k_{cat}: 0.0053 s^{-1}; subtilisin with *p*-nitrophenyl acetate k_{cat}/K_m: 610 $M^{-1}s^{-1}$, k_{cat}: 0.23 s^{-1}) (*59*, *60*) and are within the distribution of catalytic efficiencies observed in nature (*58*). Higher k_{cat} values could likely be achieved through optimization of the catalytic geometry, further preorganization of the active site (*8*, *9*), and increasing active site complexity. Acetylcholinesterase employs three backbone amide hydrogen bonds to the oxyanion and an additional network of hydrogen bonds to stabilize the catalytic aspartate (*61*, *62*). The current designs do not use this machinery, and comparison of catalytic triad and oxyanion hole geometries to those found in highly efficient native serine hydrolases highlights differences that could be responsible for the remaining activity gap (see supplementary text). Our de novo buildup approach using RFdiffusion coupled with PLACER ensemble analysis to ensure design accuracy and preorganization should allow us to test these hypotheses by direct construction, which should complement more traditional approaches based on structural examination, computational analysis, and optimization by experimental approaches like directed evolution.

Previous efforts to design catalytic triad-based designs have failed to achieve multiple turnover; in some cases, such as our preliminary NTF2-based designs, a backbone amide oxyanion hole was impossible to achieve owing to scaffold limitations, whereas in others based on native scaffolds, the histidine geometry was difficult to control, which likely limited activation of the leaving groups and water (fig. S20) (*7*). De novo backbone generation building outward from a specified active site with RFdiffusion, described here for serine hydrolases and also recently used to generate retroaldolases (*42*), overcomes these limitations by enabling generation of almost any desired catalytic geometry. We further show that the deep neural network PLACER can rapidly generate ensembles for a series of reaction intermediates to predict preorganization, and provide insights that would otherwise require labor-intensive structural studies. For example, PLACER revealed pervasive off-target conformational changes in the acyl-enzyme intermediate, providing feedback on design flaws that would go unnoticed when considering only a single state in the catalytic cycle. The value of this approach is evident in the substantial improvement in experimental success rate upon filtering with PLACER, suggesting that such ensemble generation will be useful for enzyme design moving forward. Although the designs described here do use a known mechanism, the geometries sampled and the folds that scaffold them are distinct from those found in native proteins, and the insights provided by PLACER for these geometries suggests that the approach should prove valuable for assessing catalytic geometries for which no native precedent exists. We anticipate that the ability to precisely position multiple catalytic groups using RFdiffusion, and to assess active site organization throughout a complex reaction cycle using PLACER should enable the design of a wide variety of new catalysts, such as PETases, amidases, and ligases, in the near future.

Materials and methods

NTF2 design campaign

Catalytic geometries from a previous analysis of native serine hydrolases (*33*) were used to generate constraint files for use in the RosettaMatch algorithm (*63*). The scaffold set used for matching was a set of idealized nuclear transport factor 2 (NTF2) fold proteins generated with trRosetta (*40*). After matching, sequence design was performed using LigandMPNN and FastRelax and designs were filtered using AlphaFold2 as described below. An additional filter was used requiring that all catalytic hydrogen bonds in the active site be formed in the AlphaFold2 prediction.

Computational design of serine hydrolases

Motif generation

Motifs were built in an iterative process. First, a substrate rotamer in a transition state geometry (either 4MU-Bu or 4MU-Ac) was placed in accordance with geometries in (*33*) in relation to a three-residue stub of the serine and local oxyanion hole from one of two natural serine hydrolase crystal structures, in which all residues other than serine were mutated to alanine (*N* oxyanion hole: 1scn, residues 220-222; *N+1* oxyanion hole: 1lns, residues 347-349). The transition state geometry of the substrate ester group was determined by density functional theory geometry optimization (B3LYP-D3(BJ)/6-31G(d)). Next, positions and rotamers of histidine on three-residue helical or strand stubs flanked by alanine were sampled around the catalytic serine and filtered for those structures in which the histidine simultaneously formed hydrogen bonds with the catalytic serine and the substrate leaving group oxygen. This process resulted in 108 unique motifs for design rounds 1 and 2. For the round 3 motifs, initially the aspartate or glutamate residue and second oxyanion hole hydrogen bond were added in a similar manner using geometric sampling of hydrogen-bonding conformations and rotamers. However, backbones produced from these motifs had exceedingly low AF2 success rates, presumably due to the generation of incompatible combinations of backbone conformations. To ensure that the remaining catalytic residue stubs were placed in physically plausible geometries, we generated 10,000 backbones with RFdiffusion using the simple substrate-Ser-His motifs as input, and then searched these backbones using Rosetta for positions on secondary structure that could accommodate the aspartate or glutamate triad residue to hydrogen bond to histidine. These stubs were then extracted, and in a final step, the same process was repeated to generate stubs for the second oxyanion hole, considering all

hydrogen bond donating sidechains, ultimately producing 2238 unique round 3 motifs with Ser-His-Asp/Glu catalytic triads, and Ser/Thr/Tyr/His/Trp oxyanion holes.

Backbone generation

See supplemental text for a detailed description of CA diffusion, which was employed to generate backbones to scaffold the generated active sites.

Sequence design

We performed three cycles of LigandMPNN (*47*) and Rosetta FastRelax (*64*) to design sequences for backbones generated from RFdiffusion. To encourage formation of hydrogen bond contacts to the catalytic histidine (for round 1 motifs) and to the catalytic aspartate/glutamate (round 3 motifs), the log probabilities used by LigandMPNN to select residues were biased toward polar amino acids for all residues with Cα within 8 Å of the active site. Catalytic residues were kept fixed and Rosetta enzyme constraints (*63*, *65*) were applied during the relax steps to maintain the catalytic geometry during each LigandMPNN/FastRelax cycle. Constraints were defined for each hydrogen bonding interaction between the catalytic dyad, backbone oxyanion hole, and substrate using the starting motif geometry with tolerances of 0.1 Å for distances and 5° for angles and dihedrals. For designs with catalytic triads, the His-Asp interaction was constrained.

Filtering

After sequence design, designs were filtered on the recapitulation of the motif catalytic geometry after FastRelax and the shape complementarity of the binding site to the substrate using Rosetta. Passing designs were used as input to AF2 (*50*) for single sequence structure prediction. AF2 was run using model 4 with three recycles. Designs were filtered for a global Cα RMSD < 1.5 Å, pLDDT > 75, and catalytic residue Cα RMSD < 1.0 Å. In the case of final round N+1 oxyanion hole designs, a modified version of Initial Guess AF2 was used to predict designs with sparse template information provided (see supplementary text).

Designs that passed AF2 filters were subsequently analyzed using PLACER. PLACER is a denoising neural network trained on x-ray and electron microscopy structures from the PDB to recapitulate the correct atom positions from partially corrupted input structures provided the atom type and bond connectivity is known. PLACER predictions were done for a spatial crop of 600 atoms closest to the active site. The inputs to the network included the protein backbone coordinates within the crop and the amino acid sequence with side-chain coordinates randomly initialized around the respective Cα atoms. For proteins without a crystal structure, the AF2 model was used. For every designed protein, we modeled 5 reaction states representing the chemical modifications the catalytic serine undergoes in the course of the reaction: 1) apo, 2) substrate bound, 3) tetrahedral intermediate 1 (TI1), 4) acylenzyme intermediate (AEI), and 5) tetrahedral intermediate 2 (TI2). We used 50 different seeds to generate an ensemble of 50 PLACER models for each reaction state (apo, substrate bound, TI1, AEI, and TI2). For each of the 50 models in a given ensemble, the presence and geometry of key hydrogen bonds in each individual model (see supplementary text) were determined. To analyze native hydrolases with PLACER, a set of native crystal structures was collected (*34*) (PDB IDs: 1ACB_E, 1C5L_H, 1H2W_A, 1IC6_A, 1IVY_A, 1PFQ_A, 1QNJ_A, 1QTR_A, 1ST2_A, 2H5C_A, 2QAA_A, 3MI4_A, 5JXG_A), the active site locations identified, and the aforementioned process applied.

Backbone resampling for momi and super redesign campaigns

The design model of momi was provided as input to RFdiffusion and the entirety of the protein was fixed while a region of secondary structure was diffused at the N terminus. The length of this region was randomly sampled from a range of 20 to 50 amino acids for 1000 independent diffusion trajectories. The contigs flag for RFdiffusion was as follows: contigs: {region_length},A1-160. For each backbone, the sequence of the original momi input was kept fixed while the newly diffused region at the N terminus was designed as described previously with LigandMPNN and FastRelax, with 10 sequences generated per backbone.

To generate designs in complex with the PET substrate, momi120 was redesigned around a 2-mer of the PET polymer. The PET 2-mer was aligned into the active site based on the geometry of the original momi120 design in complex with 4MU-PhAc substrate. Two regions of secondary structure which clashed with the aligned PET substrate, region 1 (residues 66-87) which flanks the lower cleft of the active site and region 2 (residues 94-104) which sits above the catalytic histidine, were subsequently remodeled with RFdiffusion. The lengths of region 1 and 2 were randomly sampled from a range of 18 to 28 amino acids and 7 to 17 amino acids, respectively, for 1000 independent diffusion trajectories. The contigs flag for RFdiffusion was formatted as follows: contigs:A1-65,{region1_length},A88-93,{region2_length},A105-194. The sequence of the entire structure was designed as described above. Twenty sequences were generated per backbone and designs were filtered as previously described with AF2 and PLACER. For 74 backbones that passed AF2 and PLACER filters, sequences were designed again as described above with 1000 sequences generated per backbone and subsequently filtered for confidence and self-consistency by single sequence AF2 prediction.

To generate a version of super with an optimized oxyanion hole sidechain geometry, we started by superimposing the active sites of super and subtilisin (PDB: 1scn) by alignment of the catalytic serine backbone atoms. Residues 56-91 that flank the oxyanion hole residue Gln^{71} in super were removed and Asn^{155} that was aligned from subtilisin was copied into the structure. We used RFdiffusion2 (*66*), a backbone generation model capable of scaffolding individual atoms or functional groups, to reconstruct the removed region of super and scaffold the newly placed amide group of Asn. We sampled lengths between 48-58 residues to generate 10,000 unique backbones, which were then designed and filtered as described above.

In-gel fluorescence screening with activity-based probes

DNA encoding the designed proteins was ordered from IDT as eblocks and the GoldenGate method was used to clone them into vector LM627 (addgene), which contains a C-terminal SNAC tag followed by a hexahistidine-tag. Resulting plasmid was transformed into BL21(DE3) cells and grown overnight in 1 ml of LB supplemented with 50 μg/ml kanamycin. For expression, 100 μl of overnight culture was used to inoculate 1 ml of LB media and grown for 1.5 hours at 37°C on a Heidolph shaker at 1300 rpm and then 10 μl of 100 mM IPTG was added and cultures were incubated at 37°C with shaking for an additional 3 hours. Cultures were centrifuged at 4000*g* for 10 min and supernatant removed. Cell pellets were resuspended in 200 μl of 20 mM HEPES (pH 7.4), containing 50 mM NaCl, 0.1 mg/ml lysozyme, and 0.01 mg/ml deoxyribonuclease I (DNase I). After 15 min, lysates were frozen in liquid nitrogen and subsequently thawed at room temperature. For labeling, 10 μl of lysate was incubated with 1 μM FP-TAMRA probe (10 μl of 2 μM stock in lysis buffer) for 1 hour at room temperature before quenching with 2× Laemmli sample buffer. Labeled samples were heated at 95°C for 5 min and 10 μl of each sample was separated on a BioRad AnykD Criterion precast gel and fluorescence imaging performed using a LI-COR Odyssey M imager. Gels were subsequently stained with Coomassie blue and imaged again.

Lysate screening

DNA encoding the designed proteins was ordered from IDT as eblocks and cloned by the GoldenGate method into vector pCOOL1 which contains a C-terminal mScarlet-i3 fusion to enable normalization of activity in lysate by enzyme concentration. Resulting plasmid was transformed into BL21(DE3) cells and cultures were grown overnight at 1 ml scale in 2 ml deep-well 96-well round bottom plates on a Heidolph shaker at 1300 rpm and 37°C. For expression, 50 μl of the overnight cultures were

used to inoculate 1 ml of autoinduction media in 2 ml deep-well 96-well round bottom plates and incubated at 1300 rpm and 37°C for approximately 24 hours. Cultures were centrifuged at 4000*g* for 10 min and supernatant decanted, washed with buffer (20 mM HEPES, 50 mM NaCl, pH 7.4), and incubated on a Heidolph shaker at 1300 rpm at room temp for 5 min to resuspend. Plates were centrifuged again at 4000*g* for 10 min and supernatant decanted. For lysis, cell pellets were resuspended with 500 µl of lysis buffer (20 mM HEPES, 50 mM NaCl, 0.01 mg/ml DNAse I, 0.01 mg/ml lysozyme, 1 mM EDTA, 0.1% triton X-100) and incubated for 2 hours on a Heidolph shaker (1300 rpm, 37°C). Plates were centrifuged at 4300*g* for 30 min and supernatant collected for screening. For activity screening, 4 or 6 µl of lysate was aliquoted into microtiter plates and reactions initiated by addition of 36 or 54 µl of buffer containing 111.1 µM 4MU-Ac or 4MU-Bu, 20 mM HEPES, 50 mM NaCl, pH 7.4, 5% dimethyl sulfoxide (DMSO). Volume sizes were modified depending on plate type used, where half-area plates were used for 40 µl reaction volume and full-area plates were used with 60 µl reaction volume. Upon addition of substrate, microtiter plates were measured once for mScarlet-i3 signal and then subsequently monitored continuously for the generation of 4MU (ex: 365 nm, em: 445 nm) on a Neo2 plate reader.

Protein expression and purification

Genes encoding the designed proteins were ordered from IDT as eblocks and cloned via the Golden Gate method into vector LM627 as previously described (*67*). Resulting plasmid was transformed into BL21(DE3) cells and grown overnight in 1 ml of LB supplemented with 50 µg/ml kanamycin, after which 500 µl of overnight was used to inoculate 50 ml of autoinduction media (*68*), which was grown 4-6 hours at 37°C and then overnight at 18°C. Cultures were spun down at 4000*g* for 15 min, and supernatant decanted. Cell pellets were resuspended in 25 ml of cold wash buffer (40 mM imidazole, 500 mM NaCl, 50 mM sodium phosphate, pH 7.4) with 1 mg/ml lysozyme and 0.1 mg/ml DNAse I. Cell slurries were sonicated on ice for 2.5 min at 80% amplitude, 10 s on 10 s off. The resulting lysate was centrifuged at 14000*g* for 30 min and the supernatant was applied to 1 ml of Ni-NTA resin equilibrated with wash buffer. The resin was subsequently washed with 15 ml of wash buffer 3 times and once with 400 µl of elution buffer (400 mM imidazole, 500 mM NaCl, 50 mM sodium phosphate, pH 7.4) followed by elution with 1.3 ml elution buffer. The eluate was purified by size-exclusion chromatography (SEC) on a Superdex 75 Increase 10/300 GL with running buffer of 20 mM HEPES, 50 mM NaCl, pH 7.4. Samples were either used immediately in downstream experiments or snap frozen in liquid nitrogen and stored at −80°C. Protein molecular weight was confirmed by liquid chromatography–mass spectrometry (LC-MS).

Kinetic analysis

To characterize hits identified from in-gel fluorescence and lysate screens for catalytic turnover, we incubated purified protein samples with fluorogenic substrates 4MU-Ac, 4MU-Bu, and 4MU-PhAc. Kinetic screens were either performed in 40 µl reaction volumes in 96-well half area plates or 60 µl reaction volume in 96-well full-area plates. Protein and substrate were prepared fresh in 20 mM HEPES, 50 mM NaCl, pH 7.4, 5% DMSO. Either 4 or 6 µl of enzyme was added to microtiter plates and the reactions were initiated by addition of substrate (36 or 54 µl). Generation of the fluorogenic product 4MU was monitored continuously (excitation 365 nm, emission 445 nm) on a Neo2 plate reader with incubation at 30°C. Analysis of the resulting data was carried out using custom scripts (see computational methods). In cases where single-turnover activity was observed, initial velocities were used to determine k_2/K_m. For those designs that displayed a clear burst phase followed by a slower steady-state rate, straight-line fits of the steady-state velocities were used to determine Michaelis-Menten catalytic parameters. To determine the uncatalyzed reaction rate in assay buffer (20 mM HEPES, 50 mM NaCl, pH 7.4, 5% DMSO), substrate was diluted in buffer alone and rates determined at multiple substrate concentrations, after which the rate was determined from fitting [S] versus rate with an equation of the form rate = k_{buffer}[S].

Crystallography

Proteins for crystallography were prepared as described above, but SEC was done with SNAC tag cleavage buffer (*69*). After SEC, protein eluate was incubated with 500 mM guanidinium hydrochloride and 2 mM $NiCl_2$ overnight at room temperature to remove the C-terminal His tag. The SNAC cleavage reaction was applied to a nickel column equilibrated with wash buffer to remove any uncleaved product and resulting eluate applied to a Superdex 75 Increase 10/300 GL column with 20 mM HEPES, 50 mM NaCl, pH 7.4 as the running buffer. Samples were concentrated and stored at −80°C or immediately used for crystallization. Crystallization screening was performed using a Mosquito LCP by STP Labtech and resulting crystals were harvested directly from the screening plate. Crystallization conditions for each design were as follows: n8 (15 mg/ml) in 0.1 M Bis-Tris pH 5.5, 25% (w/v) PEG 3350, super (50 mg/ml) in 0.2 M potassium fluoride, 20% (w/v) PEG 3350, win (42 mg/ml) in 0.1 M sodium acetate pH 4.6, 8% (w/v) PEG 4000, win1 (54 mg/ml) in 60% v/v Tacsimate pH 7.0, win31 (60 mg/ml) in 0.2 M diammonium tartrate and 20% (w/v) PEG 3350, and dadt1 (27 mg/ml) in 0.1 M potassium chloride, 0.02 M Tris pH 7.0, and 20% PEG4000. Data were processed with XDS (*70*), phased and refined with Phenix (*71*), and model building performed with COOT (*72*). Percent Ramachandran favored, allowed, and outliers for each structure are as follows: n8 (98.21, 1.79, 0.00), super (99.37, 0.63, 0.00), win (97.99, 2.01, 0.00), win1 (99.68, 0.32, 0.00), win31 (99.36, 0.64, 0.00), and dadt1 (100, 0, 0). Coordinates are deposited in the PDB with PDB IDs of 9DED (n8), 9DEE (super), 9DEF (win), 9DEG (win1), 9DEH (win31), and 9MRB (dadt1).

Mass spectrometry

Intact mass spectra of protein samples were obtained by reverse-phase LC-MS on an Agilent G6230B TOF after desalting using an Advance-Bio RP-Desalting column. Deconvolution using a total entropy algorithm was performed using Bioconfirm. In some cases, protein samples (1 mg/ml) were incubated overnight with substrate (300 µM) in SEC running buffer at room temperature prior to mass spectrometry analysis.

Structural similarity search of the PDB and AFDB

To assess the structural novelty of our designed enzymes, we used FoldSeek (*73*) to compare our crystal structures and select design models against all available databases. Searches were performed in TM-align mode and the highest TM-score hit was used for structural comparison.

REFERENCES AND NOTES

1. S. L. Lovelock *et al*., The road to fully programmable protein catalysis. *Nature* **606**, 49–58 (2022). doi: 10.1038/s41586-022-04456-z; pmid: 35650353
2. D. Röthlisberger *et al*., Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008). doi: 10.1038/nature06879; pmid: 18354394
3. L. Jiang *et al*., De novo computational design of retro-aldol enzymes. *Science* **319**, 1387–1391 (2008). doi: 10.1126/science.1152692; pmid: 18323453
4. J. B. Siegel *et al*., Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313 (2010). doi: 10.1126/science.1190239; pmid: 20647463
5. H. K. Privett *et al*., Iterative approach to computational enzyme design. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3790–3795 (2012). doi: 10.1073/pnas.1118082108; pmid: 22357762
6. F. Richter *et al*., Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J. Am. Chem. Soc.* **134**, 16197–16206 (2012). doi: 10.1021/ja3037367; pmid: 22871159
7. S. Rajagopalan *et al*., Design of activated serine-containing catalytic triads with atomic-level accuracy. *Nat. Chem. Biol.* **10**, 386–391 (2014). doi: 10.1038/nchembio.1498; pmid: 24705591
8. R. Otten *et al*., How directed evolution reshapes the energy landscape in an enzyme to boost catalysis. *Science* **370**, 1442–1446 (2020). doi: 10.1126/science.abd3623; pmid: 33214289
9. R. V. Rakotoharisoa *et al*., Design of efficient artificial enzymes using crystallographically enhanced conformational sampling. *J. Am. Chem. Soc.* **146**, 10001–10013 (2024). doi: 10.1021/jacs.4c00677; pmid: 38532610
10. A. Broom *et al*., Ensemble-based enzyme design can recapitulate the effects of laboratory directed evolution in silico. *Nat. Commun.* **11**, 4808 (2020). doi: 10.1038/s41467-020-18619-x; pmid: 32968058
11. G. Kiss, D. Röthlisberger, D. Baker, K. N. Houk, Evaluation and ranking of enzyme designs. *Protein Sci.* **19**, 1760–1773 (2010). doi: 10.1002/pro.462; pmid: 20665693

12. S. J. Fleishman, S. D. Khare, N. Koga, D. Baker, Restricted sidechain plasticity in the structures of native proteins and complexes. *Protein Sci.* **20**, 753–757 (2011). doi: 10.1002/pro.604; pmid: 21432939
13. H. A. Bunzel *et al*., Evolution of dynamical networks enhances catalysis in a designer enzyme. *Nat. Chem.* **13**, 1017–1022 (2021). doi: 10.1038/s41557-021-00763-6; pmid: 34413499
14. M. P. Frushicheva, J. Cao, Z. T. Chu, A. Warshel, Exploring challenges in rational enzyme design by simulating the catalysis in artificial kemp eliminase. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16869–16874 (2010). doi: 10.1073/pnas.1010381107; pmid: 20829491
15. A. J. Burton, A. R. Thomson, W. M. Dawson, R. L. Brady, D. N. Woolfson, Installing hydrolytic activity into a completely de novo protein framework. *Nat. Chem.* **8**, 837–844 (2016). doi: 10.1038/nchem.2555; pmid: 27554410
16. D. N. Bolon, S. L. Mayo, Enzyme-like proteins by computational design. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14274–14279 (2001). doi: 10.1073/pnas.251555398; pmid: 11724958
17. A. J. Burke *et al*., Design and evolution of an enzyme with a non-canonical organocatalytic mechanism. *Nature* **570**, 219–223 (2019). doi: 10.1038/s41586-019-1262-8; pmid: 31132786
18. S. Studer *et al*., Evolution of a highly active and enantiospecific metalloenzyme from short peptides. *Science* **362**, 1285–1288 (2018). doi: 10.1126/science.aau3744; pmid: 30545884
19. B. S. Der, D. R. Edwards, B. Kuhlman, Catalysis by a de novo zinc-mediated protein interface: Implications for natural enzyme evolution and rational enzyme engineering. *Biochemistry* **51**, 3933–3940 (2012). doi: 10.1021/bi201881p; pmid: 22510088
20. Y. S. Moroz *et al*., New Tricks for Old Proteins: Single Mutations in a Nonenzymatic Protein Give Rise to Various Enzymatic Activities. *J. Am. Chem. Soc.* **137**, 14905–14911 (2015). doi: 10.1021/jacs.5b07812; pmid: 26555770
21. V. Tournier *et al*., An engineered PET depolymerase to break down and recycle plastic bottles. *Nature* **580**, 216–219 (2020). doi: 10.1038/s41586-020-2149-4; pmid: 32269349
22. S. Yoshida *et al*., A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **351**, 1196–1199 (2016). doi: 10.1126/science.aad6359; pmid: 26965627
23. E. L. Bell *et al*., Directed evolution of an efficient and thermostable PET depolymerase. *Nat. Catal.* **5**, 673–681 (2022). doi: 10.1038/s41929-022-00821-3
24. D. M. Blow, Structure and mechanism of chymotrypsin. *Acc. Chem. Res.* **9**, 145–152 (1976). doi: 10.1021/ar50100a004
25. P. Carter, J. A. Wells, Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564–568 (1988). doi: 10.1038/332564a0; pmid: 3282170
26. P. Carter, J. A. Wells, Functional interaction among catalytic residues in subtilisin BPN'. *Proteins* **7**, 335–342 (1990). doi: 10.1002/prot.340070405; pmid: 2199971
27. L. Polgár, The catalytic triad of serine peptidases. *Cell. Mol. Life Sci.* **62**, 2161–2172 (2005). doi: 10.1007/s00018-005-5160-x; pmid: 16003488
28. P. Bryan, M. W. Pantoliano, S. G. Quill, H. Y. Hsiao, T. Poulos, Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3743–3745 (1986). doi: 10.1073/pnas.83.11.3743; pmid: 3520553
29. D. R. Corey, C. S. Craik, An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J. Am. Chem. Soc.* **114**, 1784–1790 (1992). doi: 10.1021/ja00031a037
30. B. Zerner, R. P. M. Bond, M. L. Bender, Kinetic Evidence for the Formation of Acyl-Enzyme Intermediates in the α-Chymotrypsin-Catalyzed Hydrolyses of Specific Substrates. *J. Am. Chem. Soc.* **86**, 3674–3679 (1964). doi: 10.1021/ja01072a016
31. J. Kraut, Serine proteases: Structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46**, 331–358 (1977). doi: 10.1146/annurev.bi.46.070177.001555; pmid: 332063
32. L. Hedstrom, Serine protease mechanism and specificity. *Chem. Rev.* **102**, 4501–4524 (2002). doi: 10.1021/cr000033x; pmid: 12475199
33. A. J. T. Smith *et al*., Structural reorganization and preorganization in enzyme active sites: Comparisons of experimental and theoretically ideal active site geometries in the multistep serine esterase reaction cycle. *J. Am. Chem. Soc.* **130**, 15361–15373 (2008). doi: 10.1021/ja803213p; pmid: 18939839
34. S. Du *et al*., Conformational Ensembles Reveal the Origins of Serine Protease Catalysis, *bioRxiv* 2024.02.28.582624 [Preprint] (2024); https://doi.org/10.1101/2024.02.28.582624
35. E. S. Radisky, J. M. Lee, C.-J. K. Lu, D. E. Koshland Jr., Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6835–6840 (2006). doi: 10.1073/pnas.0601910103; pmid: 16636277
36. J. L. Watson *et al*., De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089–1100 (2023). doi: 10.1038/s41586-023-06415-8; pmid: 37433327
37. F. Praetorius *et al*., Design of stimulus-responsive two-state hinge proteins. *Science* **381**, 754–760 (2023). doi: 10.1126/science.adg7731; pmid: 37590357
38. A. Pillai *et al*., De novo design of allosterically switchable protein assemblies. *Nature* **632**, 911–920 (2024). doi: 10.1038/s41586-024-07813-2; pmid: 39143214
39. M. L. Zastrow, A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro, Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.* **4**, 118–123 (2011). doi: 10.1038/nchem.1201; pmid: 22270627
40. A. H.-W. Yeh *et al*., De novo design of luciferases using deep learning. *Nature* **614**, 774–780 (2023). doi: 10.1038/s41586-023-05696-3; pmid: 36813896
41. R. Krishna *et al*., Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* **384**, eadl2528 (2024). doi: 10.1126/science.adl2528; pmid: 38452047
42. M. Braun *et al*., Computational design of highly active de novo enzymes. bioRxiv 2024.08.02.606416 [Preprint] (2024); https://doi.org/10.1101/2024.08.02.606416.
43. I. Anishchenko *et al*., Modeling protein-small molecule conformational ensembles with ChemNet. bioRxivorg 2024.09.25.614868 [Preprint] (2024); https://doi.org/10.1101/2024.09.25.614868.
44. P. A. Frey, A. D. Hegeman, *Enzymatic Reaction Mechanisms* (Oxford Univ. Press, 2007).doi: 10.1093/oso/9780195122589.001.0001
45. C. Walsh, *Enzymatic Reaction Mechanisms* (Freeman, 1979).
46. W. P. Jencks, *Catalysis in Chemistry and Enzymology* (Courier Corporation, 1987).
47. J. Dauparas *et al*., Atomic context-conditioned protein sequence design using LigandMPNN. bioRxiv 2023.12.22.573103 [Preprint] (2023); https://doi.org/10.1101/2023.12.22.573103.
48. R. Das, D. Baker, Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* **77**, 363–382 (2008). doi: 10.1146/annurev.biochem.77.062906.171838; pmid: 18410248
49. N. R. Bennett *et al*., Improving de novo protein binder design with deep learning. *Nat. Commun.* **14**, 2625 (2023). doi: 10.1038/s41467-023-38328-5; pmid: 37149653
50. J. Jumper *et al*., Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). doi: 10.1038/s41586-021-03819-2; pmid: 34265844
51. L. Simón, J. M. Goodman, Hydrogen-bond stabilization in oxyanion holes: Grand jeté to three dimensions. *Org. Biomol. Chem.* **10**, 1905–1913 (2012). doi: 10.1039/c2ob06717j; pmid: 22273994
52. L. Simón, J. M. Goodman, Enzyme catalysis by hydrogen bonds: The balance between transition state binding and substrate binding in oxyanion holes. *J. Org. Chem.* **75**, 1831–1840 (2010). doi: 10.1021/jo901503d; pmid: 20039621
53. J. Chai Discovery *et al*., Chai-1: Decoding the molecular interactions of life. bioRxiv 2024.10.10.615955 [Preprint] (2024); https://doi.org/10.1101/2024.10.10.615955
54. A. R. Buller, C. A. Townsend, Intrinsic evolutionary constraints on protease structure, enzyme acylation, and the identity of the catalytic triad. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E653–E661 (2013). doi: 10.1073/pnas.1221050110; pmid: 23382230
55. E. Zakharova, M. P. Horvath, D. P. Goldenberg, Structure of a serine protease poised to resynthesize a peptide bond. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11034–11039 (2009). doi: 10.1073/pnas.0902463106; pmid: 19549826
56. G. Dodson, A. Wlodawer, Catalytic triads and their relatives. *Trends Biochem. Sci.* **23**, 347–352 (1998). doi: 10.1016/S0968-0004(98)01254-7; pmid: 9787641
57. R. Wolfenden, Y. Yuan, The "neutral" hydrolysis of simple carboxylic esters in water and the rate enhancements produced by acetylcholinesterase and other carboxylic acid esterases. *J. Am. Chem. Soc.* **133**, 13821–13823 (2011). doi: 10.1021/ja204116a; pmid: 21793525
58. A. Bar-Even *et al*., The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011). doi: 10.1021/bi2002289; pmid: 21506553
59. F. J. Kezdy, M. L. Bender, The kinetics of the α-chymotrypsin-catalyzed hydrolysis of p-nitrophenyl acetate. *Biochemistry* **1**, 1097–1106 (1962). doi: 10.1021/bi00912a021; pmid: 14032227
60. L. Polgar, M. L. Bender, The reactivity of thiol-subtilisin, an enzyme containing a synthetic functional group. *Biochemistry* **6**, 610–620 (1967). doi: 10.1021/bi00854a032; pmid: 6047645
61. J. L. Sussman *et al*., Atomic structure of acetylcholinesterase from Torpedo californica: A prototypic acetylcholine-binding protein. *Science* **253**, 872–879 (1991). doi: 10.1126/science.1678899; pmid: 1678899
62. A. Warshel, S. Russell, Theoretical correlation of structure and energetics in the catalytic reaction of trypsin. *J. Am. Chem. Soc.* **108**, 6569–6579 (1986). doi: 10.1021/ja00281a021
63. A. Zanghellini *et al*., New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci.* **15**, 2785–2794 (2006). doi: 10.1110/ps.062353106; pmid: 17132862
64. M. D. Tyka *et al*., Alternate states of proteins revealed by detailed energy landscape mapping. *J. Mol. Biol.* **405**, 607–618 (2011). doi: 10.1016/j.jmb.2010.11.008; pmid: 21073878
65. F. Richter, A. Leaver-Fay, S. D. Khare, S. Bjelic, D. Baker, De novo enzyme design using Rosetta3. *PLOS ONE* **6**, e19230 (2011). doi: 10.1371/journal.pone.0019230; pmid: 21603656
66. D. Kim *et al*., Computational Design of Metallohydrolases, bioRxiv 2024.11.13.623507 [Preprint] (2024); https://doi.org/10.1101/2024.11.13.623507.
67. B. I. M. Wicky *et al*., Hallucinating symmetric protein assemblies. *Science* **378**, 56–61 (2022). doi: 10.1126/science.add1964; pmid: 36108048
68. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005). doi: 10.1016/j.pep.2005.01.016; pmid: 15915565
69. B. Dang *et al*., SNAC-tag for sequence-specific chemical protein cleavage. *Nat. Methods* **16**, 319–322 (2019). doi: 10.1038/s41592-019-0357-3; pmid: 30923372
70. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010). doi: 10.1107/S0907444909047337; pmid: 20124692
71. D. Liebschner *et al*., Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in *Phenix*. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019). doi: 10.1107/S2059798319011471; pmid: 31588918
72. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi: 10.1107/S0907444904019158; pmid: 15572765
73. M. van Kempen *et al*., Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024). doi: 10.1038/s41587-023-01773-0; pmid: 37156916
74. Code for Computational design of serine hydrolases, Zenodo (2025); doi: 10.5281/zenodo.14642126

ACKNOWLEDGMENTS

We thank L. Goldschmidt and K. Van Wormer for maintaining the computational and wet lab resources at the Institute for Protein Design. We thank A. P. Green, F. J. Hardy, F. Richter, and D. Hilvert for helpful advice during the development of this project. We thank B. D. Weitzner for providing code scripts. We thank F. J. Hardy and M. A. Kennedy for reading and editing drafts of the manuscript. **Funding:** Howard Hughes Medical Institute (HHMI) (I.K., A.K.B., D.B.). The Open Philanthropy Project Improving Protein Design Fund (S.J.P., K.H.S., I.K., E.B., A.K.B.). The Washington Research Foundation (S.J.P.). National Institutes of Health (NIH) and/or the National Institute of General Medical Sciences (NIGMS) Award (T32GM008268) (A.L.). A gift from Microsoft (I.A., D.J., D.B.). The Defense Threat Reduction Agency Grant (HDTRA1-19-1-0003) (S.J.P., A.L.). The Audacious Project at the Institute for Protein Design (A.K., E.B., A.K.B., A.L.). The Novo Nordisk Foundation (C.N., NNF180C0030446). The Schmidt Science Fellows, in partnership with the Rhodes Trust (A.H.). The Schmidt Futures foundation (S.J.P., A.L., K.H.S.) Crystallographic data were collected at the Advanced Light Source (ALS), which is supported by the Director, Office of Science, Office of 20 Basic Energy Sciences, and US Department of Energy under contract number DE-AC02-05CH11231. **Author contributions:** Conceptualization: A.L., S.J.P., D.B. Methodology: A.L., S.J.P., K.H.S., I.A., D.J., W.A., I.K., C.N., I.R.H., C.J., R.K., Y.K. Investigation: A.L., S.J.P., K.H.S., J.J., A.F.S., A.H., A.K., E.B., A.K.B., B.S. Visualization: A.L., S.J.P. Funding acquisition: D.B. Supervision: D.B., K.N.H. Writing – original draft: A.L., S.J.P., D.B. Writing – review and editing: A.L., S.J.P., I.A., K.H.S., D.J., A.H., I.K., C.J., I.R.H., A.K.B., K.N.H., D.B. **Competing interests:** A provisional patent (application number 63/535,404) covering the PLACER network presented in this paper has been filed by the University of Washington. D.B. and I.A. are inventors on this patent. D.B. is a cofounder and shareholder of Vilya, an early-stage biotechnology company that has licensed the provisional patent. **Data and materials availability:** Coordinates and structure factors are available in the Protein Data Bank with PDB IDs 9DED (n8), 9DEE (super), 9DEF (win), 9DEG (win1), 9DEH (win31), and 9MRB (dadt1). Code and example execution commands for each step of the design pipeline are available at https://github.com/laukoag/

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adu2454
Supplementary Text
Figs. S1 to S20
Tables S1 to S6
References (75–77)
MDAR Reproducibility Checklist

Submitted 30 October 2024; accepted 3 February 2025
Published online 13 February 2025
10.1126/science.adu2454