RESEARCH ARTICLE SUMMARY

ENZYMOLOGY

Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics

C. J. Markin†, D. A. Mokhtari†, F. Sunden, M. J. Appel, E. Akiva, S. A. Longwell, C. Sabatti, D. Herschlag*, P. M. Fordyce*

INTRODUCTION: Enzymes possess extraordinary catalytic proficiency and specificity. These properties ultimately derive from interactions not just between the active-site residues and the substrate but from functional interactions throughout a folded enzyme. Therefore, understanding the origins of catalytic proficiency and specificity will require the ability to make mutations throughout the protein. Traditionally, enzyme active sites have been characterized by means of site-directed mutagenesis (SDM), revealing much about the catalytic functions of these residues; nevertheless, SDM is low-throughput, costly, and labor intensive. By contrast, recently developed high-throughput mutational scanning techniques assay large numbers of sequences but provide only coarse estimates of function, such as the amount of product generated at a particular time under a particular set of conditions or overall organismal fitness.

RATIONALE: New technologies are needed to overcome the limitations of current approaches and allow deep characterization of many enzyme variants in a cost- and time-efficient manner. To meet this challenge, we developed a high-throughput microfluidic platform that allows the simultaneous expression and purification of more than 1500 rationally chosen enzyme mutants in hours and allows their quantitative functional characterization in days. HT-MEK (High-Throughput Microfluidic Enzyme Kinetics) can be used with any enzyme system that can be tagged and expressed in vitro and has a direct or coupled fluorogenic assay.

RESULTS: As a first application of HT-MEK, we functionally characterized 1036 single-site mutants that contain either a glycine or valine substitution at each position within PafA (phosphate-irrepressible alkaline phosphatase

of Flavobacterium), a well-studied enzyme from the alkaline phosphatase superfamily. For each mutant, we measured Michaelis-Menten kinetics [apparent unimolecular rate constant (k_{cat}), Michaelis constant (K_m), and k_{cat}/K_m] for multiple substrates, inhibition constants, and effects on folding, obtaining more than 5000 kinetic and thermodynamic constants from more than 670,000 total reactions. We found that most mutations (702 of 1036) yielded statistically significant effects on some aspect of catalysis. By systematically and independently varying expression and assay conditions, we determined that 232 of these mutations reduced catalysis by promoting the formation of a long-lived, catalytically inactive misfolded state, whereas none did so through equilibrium unfolding under our assay conditions. Combining these functional measurements with prior mechanistic knowledge allowed us to systematically assess the effect of each mutation. Different groups of residues affected different aspects of function, with residues that affect a particular function forming large, spatially contiguous regions that spanned from the active site up to 20 Å from the active site and to the enzyme surface.

CONCLUSION: HT-MEK has allowed us to uncover functional effects throughout PafA and to identify the catalytic features affected by different groups of residues. Some of these effects are readily rationalized through inspection of structural interconnections to the active site, whereas others were nonobvious, including large distal and surface effects and the discovery of a long-lived misfolded state. These results underscore the need to measure the effects of mutations on multiple kinetic and thermodynamic parameters across multiple reaction conditions and thus the need for this new technology. Because HT-MEK is applicable to any enzyme with a direct or coupled fluorescent readout and provides an in-depth and quantitative analysis of mutant space rapidly and at modest cost, it may be the method of choice to characterize new enzymes. In future applications, HT-MEK can be used to dissect potential evolutionary trajectories, determine the functional consequences of human disease-associated allelic variants, identify surfaces with nascent allosteric potential for rational control of catalysis, and direct the adaptation of natural and designed enzymes for new functions and roles. ∎
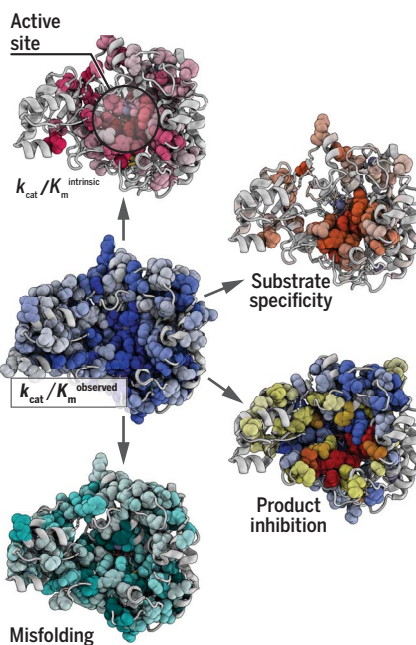
High-throughput quantitative enzymology

1568 chamber microfluidic device

Fluorescence activity assays
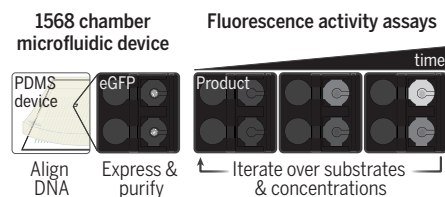
Kinetic and thermodynamic parameters for 1036 PafA phosphatase variants
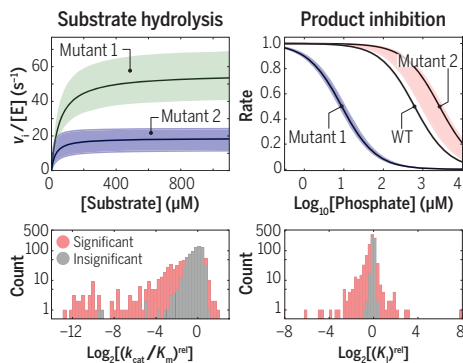
Systematic, multidimensional measurements revealing PafA functional architecture

Simultaneous expression, purification, and biochemical characterization of enzyme variants in a microfluidic device makes it possible to measure Michaelis-Menten parameters and inhibition constants for more than 1500 variants in days. Mutational effects across multiple assays reveal an extensive functional architecture in which physically contiguous residue regions extending to the enzyme surface control or alter particular aspects of catalysis.

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf8761

ENZYMOLOGY

Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics

C. J. Markin[1]†, D. A. Mokhtari[1]†, F. Sunden[1], M. J. Appel[1], E. Akiva[2], S. A. Longwell[3], C. Sabatti[4,5], D. Herschlag[1,6,7]*, P. M. Fordyce[3,7,8,9]*

Systematic and extensive investigation of enzymes is needed to understand their extraordinary efficiency and meet current challenges in medicine and engineering. We present HT-MEK (High-Throughput Microfluidic Enzyme Kinetics), a microfluidic platform for high-throughput expression, purification, and characterization of more than 1500 enzyme variants per experiment. For 1036 mutants of the alkaline phosphatase PafA (phosphate-irrepressible alkaline phosphatase of Flavobacterium), we performed more than 670,000 reactions and determined more than 5000 kinetic and physical constants for multiple substrates and inhibitors. We uncovered extensive kinetic partitioning to a misfolded state and isolated catalytic effects, revealing spatially contiguous regions of residues linked to particular aspects of function. Regions included active-site proximal residues but extended to the enzyme surface, providing a map of underlying architecture not possible to derive from existing approaches. HT-MEK has applications that range from understanding molecular mechanisms to medicine, engineering, and design.

Understanding how sequence encodes function remains a fundamental challenge in biology. Linear chains of amino acids fold into three-dimensional protein structures that carry out the physical and chemical tasks needed for life, such as highly efficient and specific catalysis. Sequence variations across organisms and individuals confer beneficial and deleterious effects: Variation throughout evolution creates proteins with improved or new functions, but variation among individuals can also compromise function and cause disease (1–3). An enhanced predictive understanding of the sequence-function landscape could have profound impacts across biology, from enabling efficient protein design to improving detection of rare allelic variants that drive disease (4–7), but new approaches and data are needed to attain this goal.

Understanding sequence-function relationships within enzymes poses a particular challenge. Structural and biochemical studies of enzymes have revealed the sites of substrate binding and catalytic transformation, the residues directly involved in catalysis, and roles

for these residues. Nevertheless, residues outside the active site are needed for the active site to assemble and function and for control of function by allosteric ligands and covalent modifications (8–10). Despite their importance, the roles played by residues outside the active site, which comprise the majority of amino acids in an enzyme, remain largely unexplored.

This dearth of knowledge stems from the nature of experimental approaches currently available. Site-directed mutagenesis (SDM) has traditionally been used to assess function by means of in-depth biochemical assays that yield kinetic and thermodynamic constants. However, SDM is time-, resource-, and labor-intensive, limiting investigation to a small number of residues. By contrast, deep mutational scanning (DMS) provides the ability to assay the effects of all 20 amino acids at every position within an enzyme (5, 11, 12). However, DMS lacks the depth and dimensionality of traditional SDM studies, typically providing a scalar readout with an uncertain relationship to the multiple fundamental physical constants needed to describe an enzyme's function.

Marrying the strengths of traditional SDM and emerging DMS is needed to usher in a new era of mechanistic enzymology. Here, we present HT-MEK (High-Throughput Microfluidic Enzyme Kinetics), a platform capable of simultaneously expressing, purifying, and characterizing more than 1500 rationally chosen enzyme mutants in parallel with the depth and precision of traditional SDM. Each HT-MEK experiment provides 1000s of measurements and multiple kinetic and thermodynamic constants [such as the apparent unimolecular rate constant (k_{cat}), Michaelis constant (K_m), k_{cat}/

K_m, and inhibition constant (K_i)] in days and at low cost.

To guide HT-MEK development and demonstrate its capabilities, we carried out a comprehensive mechanistic investigation of the effects of mutations to every residue within the alkaline phosphatase superfamily member PafA (phosphate-irrepressible alkaline phosphatase of Flavobacterium) (Fig. 1A and fig. S1). PafA and related phosphomonoesterases are among the most prodigious catalysts known, with rate enhancements of up to ~10^{27}-fold, providing a large dynamic range to explore (13). We also anticipated that PafA, a secreted enzyme, would be highly stable, potentially allowing us to more deeply probe catalysis without obfuscation from global unfolding. We found that 702 of the 1036 mutants investigated have substantial functional consequences, with none arising from equilibrium unfolding. Additional experiments revealed that many mutations promote the formation of a long-lived, catalytically incompetent misfolded state both in vitro and in cells. The multidimensional measurements provided by HT-MEK allowed us to decouple this misfolding from catalytic effects and quantify mutational effects on particular aspects of catalysis and mechanism using an approach we call Functional Component Analysis. This approach revealed spatially contiguous regions of residues that extend from the active site to the enzyme surface responsible for optimizing particular catalytic strategies. Surface residues with functional effects upon mutation may compose candidate allosteric regions useful for engineering rational control of catalytic activity. The HT-MEK platform and the quantitative multidimensional datasets it can provide will have broad utility for future efforts to understand catalytic mechanisms, natural variation, and evolutionary trajectories and to design enzymes with new functions.

Results

HT-MEK device and experimental pipeline

HT-MEK is built around a two-layer polydimethylsiloxane (PDMS) microfluidic device with 1568 chambers and integrated pneumatic valves (Fig. 1B and fig. S2) (14, 15). Each chamber is composed of two compartments (DNA and Reaction) separated by a valve (Neck), with adjacent chambers isolated from one another by a second valve (Sandwich). A third valve (Button) reversibly excludes or exposes a circular patch of the reaction compartment surface, enabling surface patterning for on-chip protein immobilization and purification (Fig. 1C) and subsequent simultaneous initiation of successive on-chip reactions across the device (Fig. 1D). Each DNA compartment of each chamber is programmed with a specified enzyme variant by aligning the device to a spotted array of DNA plasmids that encode for the expression of C-terminally enhanced green

[1]Department of Biochemistry, Stanford University, Stanford, CA 94305, USA. [2]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA 94158, USA. [3]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [4]Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA. [5]Department of Statistics, Stanford University, Stanford, CA 94305, USA. [6]Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA. [7]ChEM-H Institute, Stanford University, Stanford, CA 94305, USA. [8]Department of Genetics, Stanford University, Stanford, CA 94305, USA. [9]Chan Zuckerberg Biohub; San Francisco, CA 94110, USA.
*Corresponding author. Email: herschla@stanford.edu (D.H.); pfordyce@stanford.edu (P.M.F.)
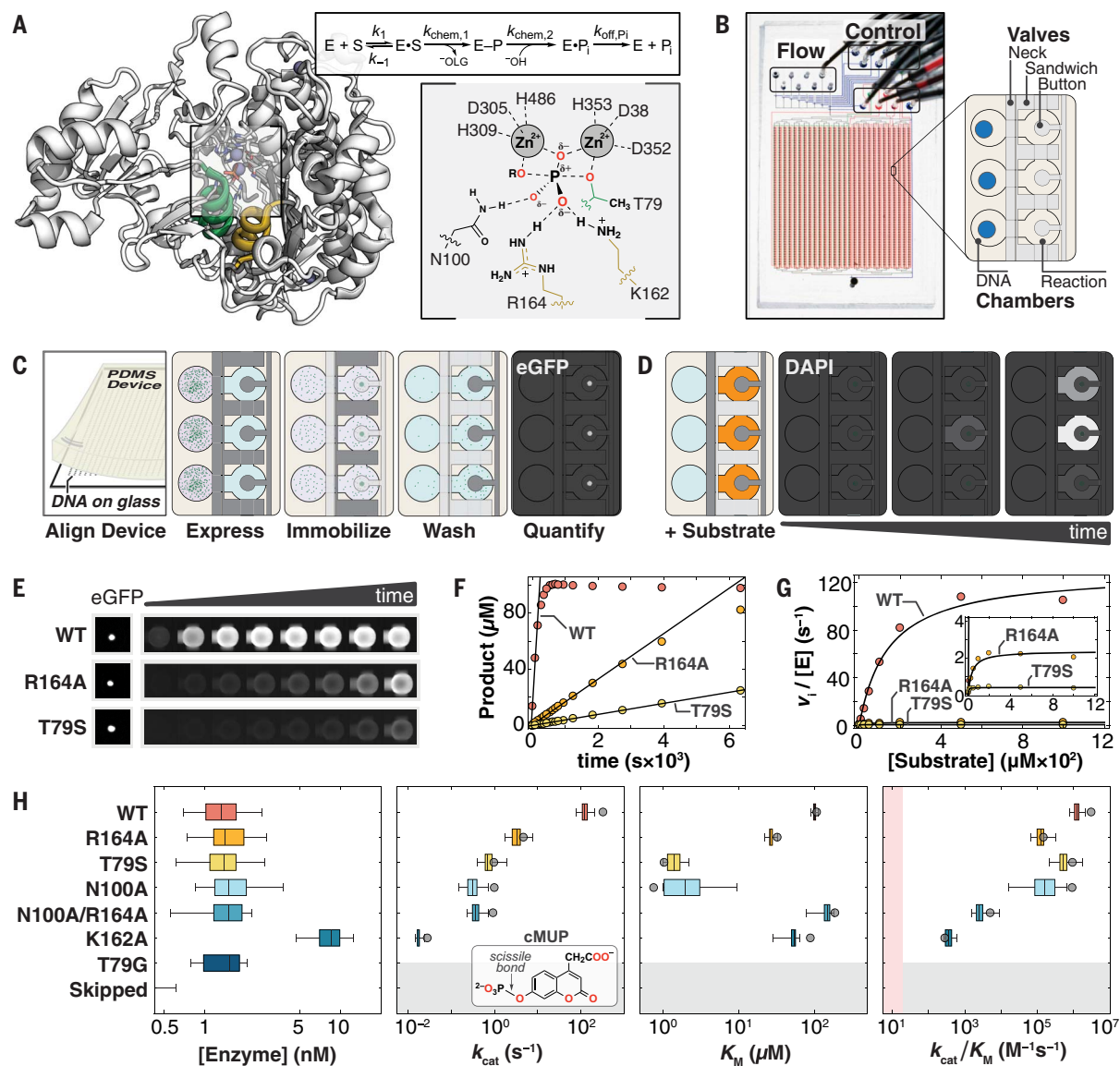†These authors contributed equally to this work.

Fig. 1. Overview of HT-MEK and PafA. (**A**) (Left) Crystal structure of WT PafA [Protein Data Bank (PDB) ID: 5TJ3] highlighting "nucleophile" (residues 77 to 89, green) and "monoesterase" (residues 161 to 171, yellow) helices. (Right) PafA catalytic cycle (S, phosphate monoester substrate; E–P, covalent phospho-threonine intermediate; ⁻OLG, alkoxide leaving group) and active site in the transition state of monoester hydrolysis. (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example, T79S indicates that threonine at position 79 was replaced by serine.) (**B**) HT-MEK microfluidic device image and schematic showing solution (Flow) and pneumatic manifold (Control) input ports, device valves, and chambers. (**C**) Schematic of on-chip enzyme expression pipeline. Dark and light gray valves are pressurized (closed) and depressurized (open), respectively. (**D**) Schematic of on-chip activity assays by using fluorogenic substrate in reaction chambers (orange). (**E**) Sample images of (left) immobilized enzyme and (right) fluorogenic product over time for WT PafA and two active-site mutants (R164A and T79S) with the substrate cMUP. (**F**) Example cMUP progress curves for chambers containing wild type and two active-site mutants and initial rate fits to these data. (**G**) Michaelis-Menten fits to initial rates yield k_{cat}, K_m, and k_{cat}/K_m for cMUP. (**H**) (Left) On-chip expressed concentrations for WT PafA and six active-site mutants calculated by using eGFP calibration curves. (Middle left to right) Comparisons of on-chip (box plots) and off-chip (gray circle) values of k_{cat}, K_m, and k_{cat}/K_m values (log₁₀ axis) for cMUP for seven PafA variants (table S1) and (inset) cMUP structure. K162A was expressed at higher concentration in a later experimental tier (with higher [E]); "skipped" refers to chambers without plasmid DNA; pink shaded region denotes 10-fold above the median apparent second-order rate constant for T79G control chambers (supplementary materials, materials and methods).

fluorescent protein (eGFP)–tagged variants (fig. S3). After alignment, device surfaces are patterned with antibodies to eGFP beneath the Button valve and passivated with bovine serum albumin elsewhere. All enzymes are then expressed in parallel through the introduction of an *Escherichia coli* in vitro transcription-translation system and purified by means of capture with surface-patterned immobilized antibody and washing (fig. S4). Production of up to 1568 different purified enzymes takes ~10 hours, with most steps automated. Enzymes are immobilized under the Button valves that protect against flow-induced loss of enzyme during solution exchange and allow repeated synchronous initiation of reactions.

To obtain catalytic rate parameters, we quantify (i) the concentration of immobilized

enzyme in each chamber, using an eGFP calibration curve (fig. S5), and (ii) the amount of product formed as a function of reaction time, using a chamber-specific product calibration curve (fig. S6). We then fit reaction progress curves in each chamber to obtain initial rates (v_i) for each substrate concentration using a custom image-processing pipeline and convert observed rates (v_i) to enzyme-normalized rate constants according to the eGFP intensity in each chamber and calibration curve (Fig. 1, E and F, and figs. S7 and S8). This process—repeated on a single device for multiple substrate concentrations, multiple substrates, and multiple inhibitors—provides the data necessary to obtain Michaelis-Menten parameters and other kinetic and thermodynamic constants (Fig. 1G and fig. S7).

HT-MEK reproduces kinetic constants previously measured with traditional assays

To demonstrate the technical capabilities of HT-MEK, we applied it to study seven previously characterized PafA variants: wild type, five active site mutants (T79S, N100A, R164A, K162A, and N100A/R164A), and one mutant lacking detectable activity (T79G, negative control) (16). Activities of wild-type (WT) PafA and the six mutants span a broad range in k_{cat} (>10^4-fold), k_{cat}/K_m (>10^4-fold), and K_m (>10^2-fold) for aryl phosphate monoester hydrolysis, providing a stringent initial test of HT-MEK dynamic range (table S1). Nearly all DNA-containing chambers expressed enzyme (>90%), and all mutants expressed at similar levels as determined from eGFP fluorescence (K162A was deliberately expressed at higher concentrations in a later experimental tier, described below) (Fig. 1H and fig. S9).

Although fluorogenic phosphate ester substrates permit kinetic assays of phosphatase activity with a high dynamic range, microfluidic assays that use the commercial 4-methylumbelliferyl phosphate ester (MUP) were complicated by partitioning of the hydrophobic fluorescent product into the hydrophobic PDMS, which increases background and distorts kinetic measurements. To address this limitation, we synthesized MUP derivatives of similar reactivity (cMUP and a corresponding methyl phosphodiester, MecMUP) that bear a charged moiety on the leaving group to eliminate PDMS absorption (figs. S10 to S12 and table S1).

Accurately resolving enzymatic rates spanning many orders of magnitude poses technical challenges because different acquisition times are needed at catalytic extremes (fig. S13), and even a small concentration of contaminating fast enzyme introduced into nearby chambers during fluid exchanges can obscure the true rates for the most catalytically compromised mutants. To address the first challenge, we expressed enzymes at two concentrations: ~1.5 nM (for accurate measurement of fast enzymes) and ~15 nM (to speed reactions for efficient detection of slow enzymes) (Fig. 1H and fig. S14). To identify regions of the device with contaminating enzymes from other chambers, we interspersed chambers that were empty or contained the inactive T79G mutant and also measured their apparent activity (figs. S15 and S16 and table S2). Per-device normalizations (0.4 and 1.4-fold) were used to account for small variations in apparent activity owing to nonspecific adsorption of mutant enzymes to chamber walls (supplementary materials, materials and methods). This normalization increased precision across replicates but did not affect conclusions (fig. S17). HT-MEK assays recapitulated cMUP kinetic parameters (k_{cat}, K_m, and k_{cat}/K_m) accurately and over a wide dynamic range (>10^4-fold in k_{cat}/K_m), indicating that the device and added eGFP tag did not alter activity (Fig. 1H and table S3).

Many mutations throughout PafA affect phosphate monoester hydrolysis

To explore functional effects of mutations throughout PafA, we created mutant libraries in which we introduced two residues with widely differing side-chain properties at each position: (i) glycine, to ablate side chain interactions and increase backbone flexibility, and (ii) valine, to introduce a branched chain hydrophobe of average volume. Native valine and glycine residues were mutated to alanine. Nearly all of these 1052 possible mutants (1036; 98%) were successfully cloned, sequenced, expressed, and assayed with HT-MEK (fig. S18).

We first measured the catalytic effect of each substitution on the steady-state kinetic parameters for cMUP hydrolysis (k_{cat}, K_m, and k_{cat}/K_m) (Fig. 2A). To facilitate efficient measurement, we performed experiments in three tiers based on reaction rates (table S2). In tier 1, we assayed devices that contained all variants at low enzyme concentration ([E] ~ 1.5 nM), with enzymes printed in duplicate and the above-noted active site mutants distributed throughout as fiducial controls. Tier 2 and 3 measurements, using higher enzyme concentration ([E] ≈ 15 nM), focused successively on the slowest variants with increasing assay times and increasing numbers of replicates to provide high-precision measurements of these slower mutants. Each device was used to measure tens of cMUP progress curves, and all expressed variants were stable over days, facilitating high-throughput data collection. Per-experiment data reports contain all data collected for each chamber, including initial rate plots and fit Michaelis-Menten curves (an example is provided in fig. S19, and full data sets are available as data files S1 and S2); per-mutant summaries combine data from all experiments and include estimates of statistical significance (Fig. 2, B and C, and fig. S20). In total, we acquired a median of nine and seven replicates for valine and glycine mutants, respectively, over 16 experiments (figs. S21 to 24 and table S2). The wealth and precision of these data allowed us to resolve differences across ranges of 10^4, 10^2, and 10^5-fold for k_{cat}, K_m, and k_{cat}/K_m, respectively.

As expected, mutations of active-site residues and catalytic Zn^{2+} ligands were highly deleterious, and positional effects varied for valine and glycine (Fig. 2, D to F). Nevertheless, a surprisingly large number of mutations throughout the enzyme were deleterious, with decreases in k_{cat}/K_m observed for 267 of the 1036 mutants ($P < 0.01$). We also observed 35 mutants with increased activity (fig. S21). These measurements provide a quantitative survey of Michaelis-Menten kinetic constants for mutations throughout a large enzyme but do not tell us why so many mutations alter activity.

The most obvious explanation for these widespread effects would be destabilization leading to a significant fraction of unfolded enzyme. Beyond destabilization, mutations can have other repercussions, altering the catalytic effectiveness of particular active-site residues, reducing Zn^{2+} affinity at the bimetallo active site or altering enzymatic protonation states. The ability to efficiently measure catalytic activity for all mutants under different assay and expression conditions as afforded by HT-MEK allowed us to test each of these possibilities.

Widespread mutational effects do not arise from equilibrium unfolding

Reflecting its role as a secreted phosphatase designed to function in harsh and variable environments, WT PafA is highly stable and remains folded, as inferred from circular dichroism (CD) spectra, even after exposure to 4 M urea for 14 days (fig. S25). This stability suggests that any individual mutation is unlikely to substantially unfold the enzyme. To directly test this expectation, we measured cMUP activities in the presence of increasing concentrations of urea. If a variant were already partially unfolded in the absence of urea, then even low concentrations of added urea would cause substantial additional unfolding (17), proportionally lowering activity (Fig. 2G and supplementary text S1). By contrast, we observed only minor rate effects for all mutants (six ± onefold decrease at 3 M compared with 0 M urea, ±SEM), which is considerably less than the >10^9-fold decrease predicted for an unfolding effect and consistent with inhibition by urea, with K_i = 2.6 M for all variants (Fig. 2H and fig. S26).

Equilibrium unfolding would also predict a Zn^{2+} concentration dependence for the observed rate because added or removed Zn^{2+} would pull the equilibrium between folded (and Zn^{2+}-bound) and unfolded states (Fig. 2G). Consistent with an absence of unfolding
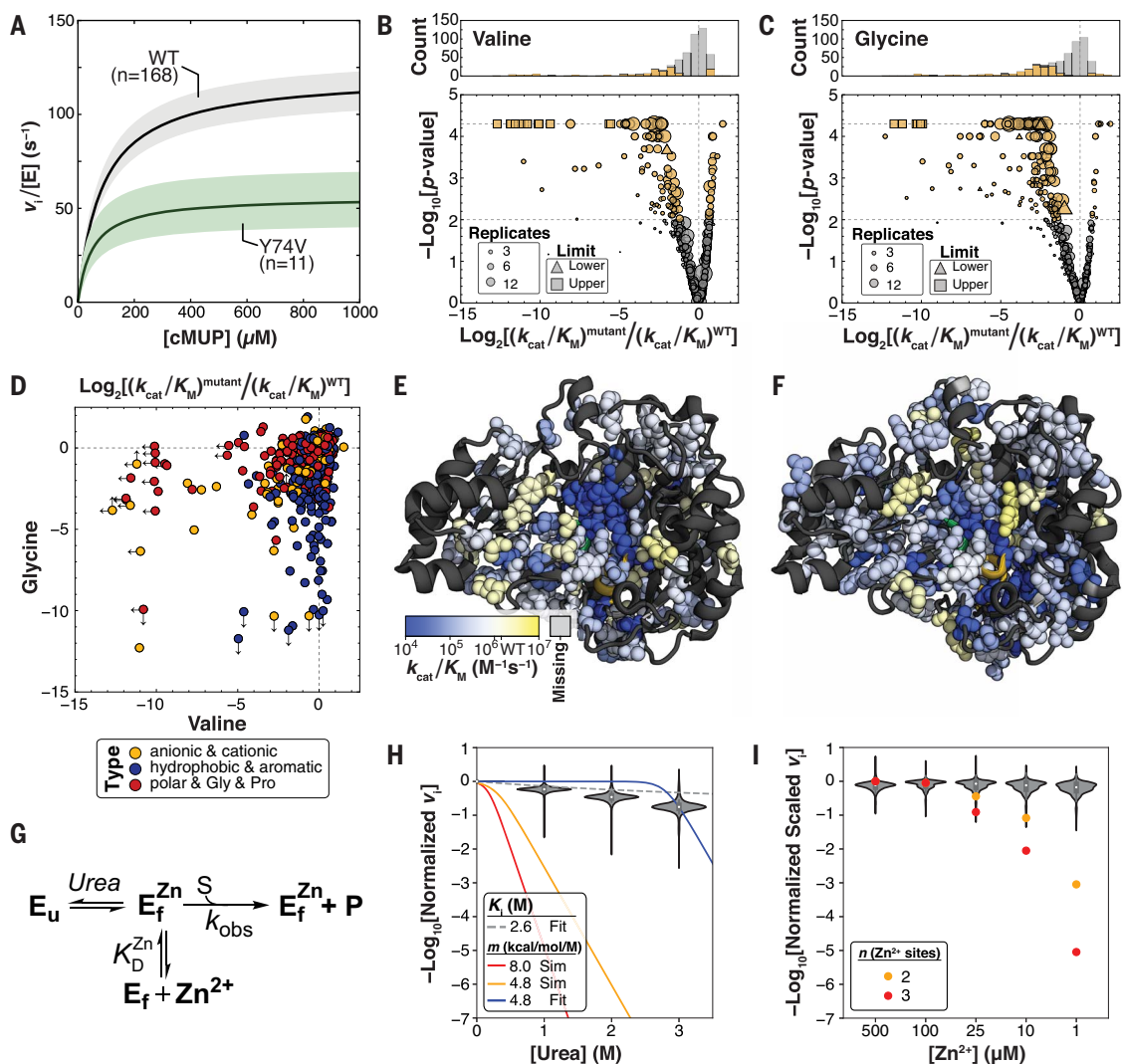
Fig. 2. HT-MEK measurements of aryl phosphate monoester (cMUP) hydrolysis for valine and glycine scans of PafA. (**A**) Median Michaelis-Menten curves for WT PafA and an example mutant with k_{cat} and K_m effects. Colored regions denote 99% confidence intervals (CIs) on medians of parameters from replicates. (**B**) Valine substitutions at 126 positions alter cMUP k_{cat}/K_m at $P < 0.01$ (105 slower, 21 faster, gold markers; gray, $P \geq 0.01$). (**C**) Glycine substitutions at 176 positions alter cMUP k_{cat}/K_m at $P < 0.01$ (162 slower, 14 faster, gold; gray, $P \geq 0.01$). (**D**) Effect of glycine versus valine substitution on cMUP k_{cat}/K_m at each position, colored by identity type of the native residue. Arrows denote the presence and direction of measurement limits. (**E** and **F**) Valine (E) and glycine (F) substitution effects on cMUP hydrolysis on the PafA structure. $P < 0.01$ sites shown as spheres, $P \geq 0.01$ and missing sites shown as ribbons. Alternate views are provided in fig. S21. (**G**) Model of equilibrium unfolding with varying urea and Zn^{2+}. (**H**) Activity of mutants relative to wild type at 50 μM cMUP as a function of urea. Red and yellow lines show predicted dependencies if mutants were 10% unfolded in the absence of urea, assuming m values of 4.8 and 8 kcal/mol/M (supplementary text S1). The blue curve is the best fit of the data for m = 4.8 kcal/mol/M with fraction unfolded (f_u) at 0 M [urea] as a free parameter (fit $f_u = 1 \times 10^{-10}$), and the dashed gray curve is the fit competitive inhibition of WT PafA by urea (fig. S26). (**I**) Normalized activity of mutants relative to wild type as a function of $[Zn^{2+}]$. Orange and red points show expected activities for Zn^{2+}-dependent unfolding events if mutants were 10% unfolded at 100 μM Zn^{2+} assuming two or three Zn^{2+} binding events, respectively.

effects, we observed no dependence on Zn^{2+} concentration over a ~10^3-fold range (Fig. 2I and supplementary text S1). This observation also establishes that observed rate effects for mutants with measurable rates do not arise from loss of bound Zn^{2+} owing to lowered Zn^{2+} affinity. Last, mutants were unaltered in their pH dependencies (fig. S27 and supplementary text S1), ruling out altered protonation states as responsible for observed kinetic effects.

A general high-throughput assay for phosphate release and additional mutational effects

Although fluorogenic probes provide a sensitive and convenient method for directly visualizing enzyme activity in kinetic assays, many reactions lack a direct fluorogenic readout. To allow future application of HT-MEK to a much broader range of enzymes, we developed an on-chip coupled assay in which inorganic phosphate (P_i) is detected through fluorescence em-

itted upon binding to a modified phosphate binding protein (PBP) (Fig. 3A) (18). Calibration curves for P_i and PBP and control measurements using the PafA substrate methyl phosphate (MeP) established that this coupled on-chip assay can detect submicromolar P_i formation and accurately reproduce off-chip kinetic constants (figs. S28 to S30).

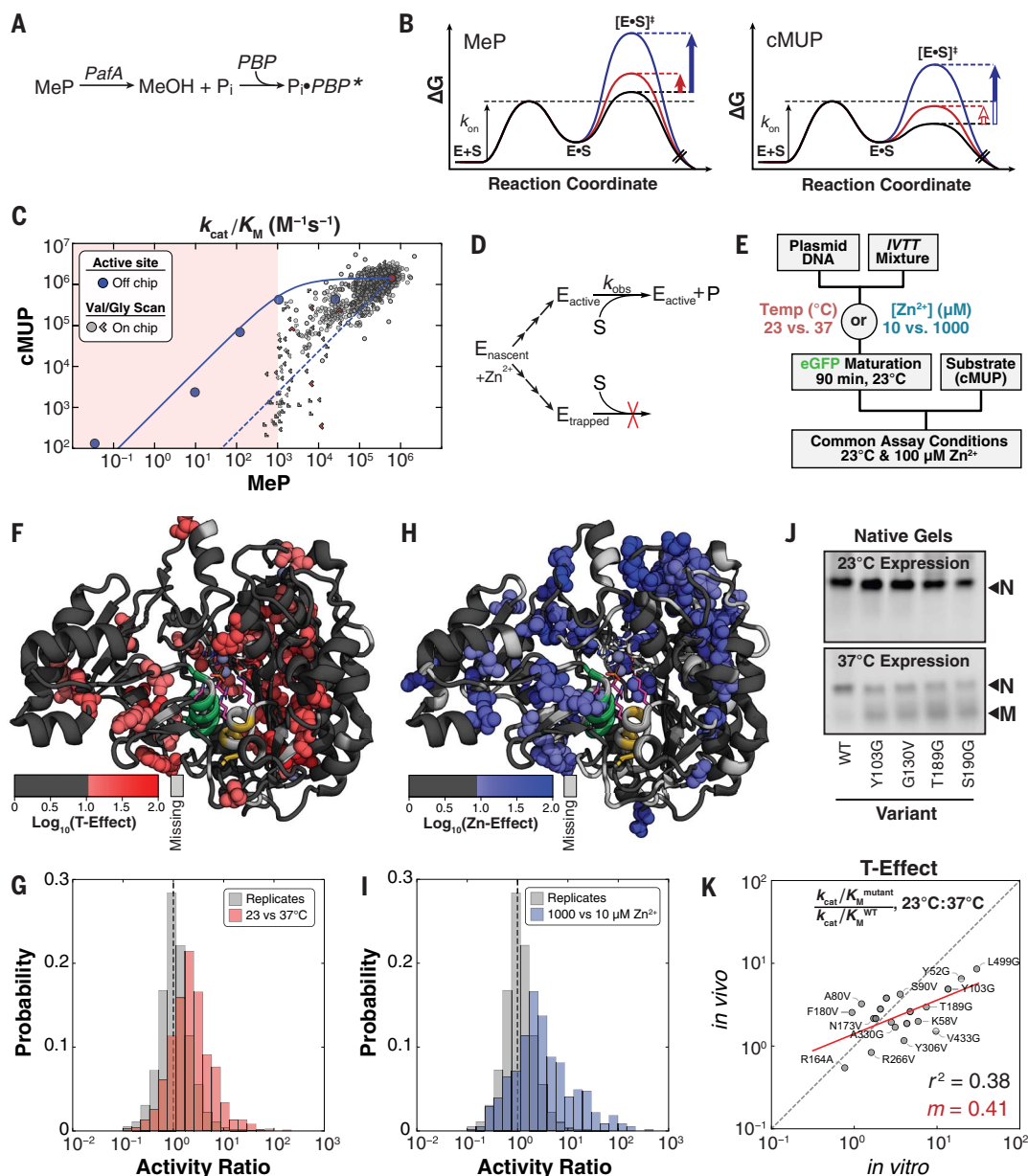Beyond their limited availability, fluorogenic substrates are often more reactive than

Fig. 3. Substrates with different intrinsic reactivity reveal kinetic folding traps. (**A**) Schematic illustrating real-time detection of P_i produced by MeP hydrolysis on-chip using a fluorophore-conjugated PBP (*18*). (**B**) Reaction coordinate diagrams schematically illustrating substrates with either chemistry (MeP) or binding (cMUP) as the rate-limiting step. Solid arrows indicate observable mutational effects on the chemical step, and open arrows indicate effects that are not observable. (**C**) Measured k_{cat}/K_m for MeP versus k_{cat}/K_m for cMUP for active-site mutants off-chip (blue points) and for valine and glycine mutants. Limits in one or both directions are indicated with chevrons pointing in the direction or quadrant of the limit. The red shaded region denotes lower limit of detection for MeP, the solid blue line indicates predicted relationship between k_{cat}/K_m for cMUP and MeP hydrolysis, and the dashed blue line shows predicted relationship for enzymes with WT activity but a varying fraction of inactive enzyme (supplementary text S2). (**D**) Scheme showing kinetic partitioning between folded (active) and misfolded (kinetically trapped) enzyme during expression. (**E**) Pulse-chase experiment to test for expression-dependent misfolding by varying expression conditions while keeping reaction conditions constant. (**F**) Residues with temperature-dependent activity changes shown

as spheres on the PafA structure, with active-site residues N100, K162, R164, and T79 shown as purple sticks. (**G**) Histograms of the ratio of measured activities after expression at 23° and 37°C [cMUP $(k_{cat}/K_m^{mutant})/(k_{cat}/K_m^{WT})$, 23°C/37°C] (red) compared with the ratios of randomly paired replicates expressed under standard conditions (37°C and 100 μM Zn^{2+}; gray). (**H**) Residues with expression $[Zn^{2+}]$–dependent activity changes shown as spheres on the PafA structure. (**I**) Histograms of the ratio of measured activities across Zn^{2+} concentrations [cMUP $(k_{cat}/K_m^{mutant})/(k_{cat}/K_m^{WT})$, 1000 μM/10 μM Zn^{2+}] (blue) compared with the ratios of randomly paired replicates expressed under standard conditions (37°C and 100 μM Zn^{2+}; gray). (**J**) (Top) Native gels for enzymes expressed by means of in vitro transcription-translation off-chip at 23°C. All constructs appear as a single band corresponding to the natively folded species ("N"). (Bottom) At 37°C, non-WT variants have an additional band corresponding to the putatively misfolded ("M") species. (**K**) Expression temperature effects on-chip (in vitro) versus in *E. coli* (in vivo). The red line indicates linear regression of log-transformed data with the indicated slope (values provided in table S5). In vivo protein levels were determined from eGFP fluorescence to estimate total expressed protein concentration.

naturally occurring substrates, potentially rendering binding rate–limiting and obscuring mutational effects on the chemical step of catalysis, and there is evidence for this behavior with PafA (*16*). On-chip MeP hydrolysis assays, therefore, complement the wide dynamic range of cMUP assays by identifying modest effects on the chemical step. Off-chip measurements of several active-site mutants revealed a decrease in observed MeP activity (owing to effects on the chemical step) without a concomitant change in cMUP activity, as expected for rate-limiting cMUP binding (Fig. 3B, red vertical arrows). Once transition-state destabilization was sufficiently large, MeP and cMUP reactions both slowed (Fig. 3B, blue vertical arrows). The solid blue line in Fig. 3C is a fit to the rate model derived from the free energy–reaction profiles in Fig. 3B for a series of active-site mutants (Fig. 3C, blue points; figs. S31; and supplementary text S2) (*19*) and predicts the kinetic behavior expected for the PafA glycine- and valine-scanning library mutants.

HT-MEK kinetic measurements revealed k_{cat}/K_m effects for almost half of the mutants (498 of 1035 with $P < 0.01$) (fig. S32), but few exhibited the predicted behavior (Fig. 3C, solid blue line versus gray symbols). Instead, the mutants tended to fall between the predicted line and a diagonal line representing equally deleterious effects on the reactions of both substrates (Fig. 3C, blue solid and dashed lines, respectively). Equally deleterious effects are expected for enzymes with WT activity but only a fraction of the enzyme in the active configuration, with less correctly folded variants further down the diagonal. Thus, observed intermediate effects could represent combinations of effects on the chemical step and on the fraction of the mutant population that is active.

Many mutations reduce catalysis by altering folding

The urea, Zn^{2+}, and pH data presented above provided strong evidence against equilibrium unfolding for any of the variants (Fig. 2, H and I). We therefore considered and tested an alternate model in which inactive enzyme resulted from a nonequilibrium process: the formation of long-lived misfolded proteins during expression (Fig. 3D) (*20*). Because temperature is known to affect folding efficiency (*21*, *22*), and because PafA binds multiple Zn^{2+} ions during folding, we varied the expression temperature (23° and 37°C) and Zn^{2+} concentration (10 and 1000 μM) from our standard expression conditions (37°C and 100 μM Zn^{2+}). We then measured reaction rates under identical assay conditions (23°C and 100 μM Zn^{2+}, our standard assay conditions), so that any observed rate changes must arise from differences during folding that persisted over time (Fig. 3E). Many mutations had differential ef-

fects on observed catalytic activity when expressed at 23° versus 37°C ("T-Effect") or with different concentrations of Zn^{2+} ("Zn-Effect"). Whereas T-Effects were found predominantly in the enzyme core, Zn-Effects were concentrated around the distal Zn^{2+} (Fig. 3, F to I, and figs. S33 and S34). These results strongly support the presence of persistent nonequilibrium folding effects, which have been anecdotally observed in other systems (*20*, *21*, *23–25*). A second prediction of the misfolding model was also met: Variants with T- and Zn-Effects were not systematically altered in K_m under varied expression conditions (fig. S35), as expected under a model of nonequilibrating states. PafA folding thus apparently involves one or more branchpoints sensitive to temperature and Zn^{2+} that lead to active PafA or one or more long-lived inactive states (Fig. 3D).

Altered folding pathways promote a long-lived inactive state in vitro and in vivo

Misfolding could be an artifact of high-throughput on-chip expression or could also arise during standard expression in vitro and possibly in vivo. To test for chip-induced misfolding effects, we selected 19 variants with varying predicted amounts of misfolding for off-chip expression by means of in vitro transcription-translation and kinetic characterization (fig. S36 and table S4). Activities were similar off- and on-chip (fig. S37), suggesting that the chip is not responsible for the observed misfolding.

Native gels and kinetic assays provided additional support for the misfolded state. Mutants predicted to misfold had an additional band of distinct mobility when expressed at high temperature that was not present or diminished when expressed at lower temperature (Fig. 3J, misfolded state "M"). Transient treatment with thermolysin, a protease that cleaves within exposed hydrophobic regions that occur in unfolded or misfolded proteins, resulted in loss of M but not the native state (native state "N") (fig. S38A) (*26*). Nevertheless, despite degradation of the majority of the protein (present as M), the total observed enzyme activity for each mutant was unchanged before and after degradation (fig. S38B), indicating that M lacked substantial activity and that N and M did not equilibrate over the hours taken to carry out these experiments. The observation that different temperature-sensitive mutants gave the same mobility on a native gel and similar protease sensitivity suggests the formation of a single misfolded state or set of states with similar properties.

Our observation of a long-lived inactive state raised the question of whether analogous misfolding occurs in cells, where cellular machinery can assist folding. We recombinantly expressed 21 variants in *E. coli* that did and did not undergo temperature-dependent misfold-

ing in vitro (fig. S36 and table S5). Expression in vivo was also temperature dependent, with changes in apparent k_{cat}/K_m values for the in vivo–expressed PafA mutants that correlated with the change observed in vitro [coefficient of determination (R^2) = 0.38] (Fig. 3K). Temperature effects were generally smaller for in vivo–expressed mutants, suggesting that cellular factors may partially rescue or preferentially degrade the misfolded state, analogous to cellular degradation of thermodynamically destabilized mutants (*27*).

To test whether misfolding was a consequence of the eGFP tag, we overexpressed a non–eGFP-tagged variant (Y103G). We found wild type–like activity when it was expressed at 23°C but 260-fold less activity when it was expressed at 37°C (table S6 and fig. S39), suggesting that >99% of the 37°C-expressed enzyme was in the M state. CD spectra of the purified Y103G mutant provided independent evidence for a structural alteration that accompanied misfolding in vivo. WT PafA exhibits identical CD spectra when expressed at 37° or 23°C; by contrast, the CD spectrum for Y103G matches that of wild type when the mutant is expressed at low temperature (23°C) but not at higher temperature (37°C) (fig. S40), with the observed difference at 37°C suggesting loss of about one-third of PafA's α-helical character in M (table S7).

Together, these results suggest that cellular folding conditions and chaperones are insufficient to prevent mutations from causing PafA to misfold and form long-lived inactive states in an *E. coli* expression strain [BL21 (DE3)]. However, the degree to which misfolding is rescued likely varies between strains and is sensitive to growth conditions and expression levels (*28*). A tendency to form kinetically stable misfolded states may therefore exert a selective pressure and influence the fitness landscape of proteins in cells (*29–33*).

Dissecting the origins of observed catalytic effects

HT-MEK assays allow us to quantify and dissect the degree to which observed changes in activity arise from changes in the amount of expressed protein, the amount that is correctly folded, and the catalytic efficiency of the correctly folded enzyme. Below, we isolate the catalytic effects for our PafA variants. We then take advantage of HT-MEK's ability to provide quantitative kinetic and thermodynamic constants for multiple substrates and inhibitors and use these data to probe PafA's functional architecture at a global level.

To remove folding effects, we compared k_{cat}/K_m values measured for substrates with different rate-limiting steps (cMUP and MeP) (Fig. 3C and supplementary text S3 and S4), represented the datapoint for each PafA variant in Fig. 3C as a superposition of a catalytic

effect (Fig. 3C, blue solid line, defined by PafA active site variants that do not have misfolding effects; fig. S41) and a misfolding effect (Fig. 3C, diagonal blue dashed line), and solved for both (Fig. 4A and fig. S42). With this approach, we were able to quantify catalytic effects (k_{cat}/K_m^{chem}) for 946 of the 1036 variants and obtain upper limits for an additional 65 (fig. S43). Deleterious catalytic effects were found for mutations at 161 of PafA's 526 positions (Fig. 4B, figs. S44 and S45, and table S8). Mutations at an even larger number of positions, 232, gave folding effects. The magnitude of folding effects did not correlate with observed eGFP fluorescence intensities, establishing that enzyme misfolding does not influence eGFP folding (fig. S46).

The largest catalytic effects cluster in and around the active site, with the fraction of residues exhibiting effects diminishing with distance from the active site (Fig. 4C and table S8). Although positions that give catalytic effects tend to cluster, the pattern is asymmetric

and complex (Fig. 4D). Many large distal effects are found surrounding the bound Zn^{2+} that is >15 Å from the active site ("distal Zn^{2+}"), suggesting that its coordination may restrain the conformational mobility of surrounding residues and propagate to the active site. Nevertheless, it is difficult to rationalize catalytic effects, especially more distal effects. To better understand these patterns and to relate effects to the specific mechanisms used in catalysis, we defined functional components (FCs) from prior mechanistic studies of PafA and other alkaline phosphatase (AP) superfamily members (16, 34, 35) and systematically assessed the effect of each mutation on each FC.

FC1: Mutations disrupting O2 phosphoryl oxygen atom interactions

Our first FC (FC1) is derived from the observation that removal of two active-site side chains, K162 and R164, reduces monoester hydrolysis by 10^5-fold but has no effect on diester hydrolysis, rendering PafA an equally

potent phosphate mono- and diesterase (fig. S47) (16). Because K162 and R164 interact with the phosphoryl oxygen that is anionic in monoesters but esterified in diesters (Fig. 5A, O2) (34), we expect other residues that support the formation of these interactions to similarly disrupt monoester but not diester hydrolysis. We define these effects as FC1 = $\Delta^{diester}/\Delta^{monoester}$, where $\Delta = (k_{cat}/K_m)^{mutant}/(k_{cat}/K_m)^{WT}$. Although the simplest expectation is that mutations to residues neighboring K162 and R164 will have FC1 effects, we cannot predict how large and how varied these effects are, how far they extend, or whether there are remote regions or surface sites that have large effects. In addition, we cannot predict whether residues that affect FC1 also contribute to other catalytic mechanisms, represented as other FCs below.

To measure diester activity on-chip and determine FC1 effects, we synthesized a fluorogenic diester substrate suitable for HT-MEK (MecMUP) and measured k_{cat}/K_m for the PafA mutant libraries (high K_m values for the noncognate diesterase activities preclude estimation of k_{cat} and K_m separately) (fig. S48) (16). We obtained k_{cat}/K_m values for 857 of the 1036 mutants and upper limits for an additional 178 (fig. S49, A and B), and on-chip diester rate constants matched off-chip measurements (fig. S49C). Because folding affects hydrolysis of both substrates equally, we compared measured rates directly, without correcting for folding effects (fig. S50A), allowing stronger statistical inference (fig. S50B).

Many mutants had FC1 effects: 88 Val and 93 Gly mutations (fig. S50C), corresponding to 156 of the 494 measurable positions (of a total of 526 positions) (Fig. 5, B to D). Seven of the 10 measurable non–active-site residues directly contacting K162 or R164 (second-shell residues) exhibited FC1 effects (Fig. 5E; fig. S51, A and B; and table S9A), which is consistent with frequent second-shell effects observed in directed evolution experiments (36–38). Of the three active-site Zn^{2+} ligands with measurable effects upon mutation, we observed an FC1 effect for D38G (which accepts a hydrogen bond from K162) but no FC1 effect for D352G or H353V (which do not interact with K162 or R164). The active-site variants T79S and N100G also had FC1 effects (tables S9B and S10), which is consistent with coupling between active-site residues attributable to shared contacts with K162 and R164 (Fig. 5E).

Although the largest effects were observed for active-site residues and next-largest for the second-shell residues, there was no additional drop in effect size after the third shell; the majority of residues with >10-fold FC1 effects were found in the third shell and beyond (15 of 23) (fig. S51 and table S11). Four of these lie at the enzyme surface (Fig. 5D and fig. S51), which is consistent with the hypothesis that
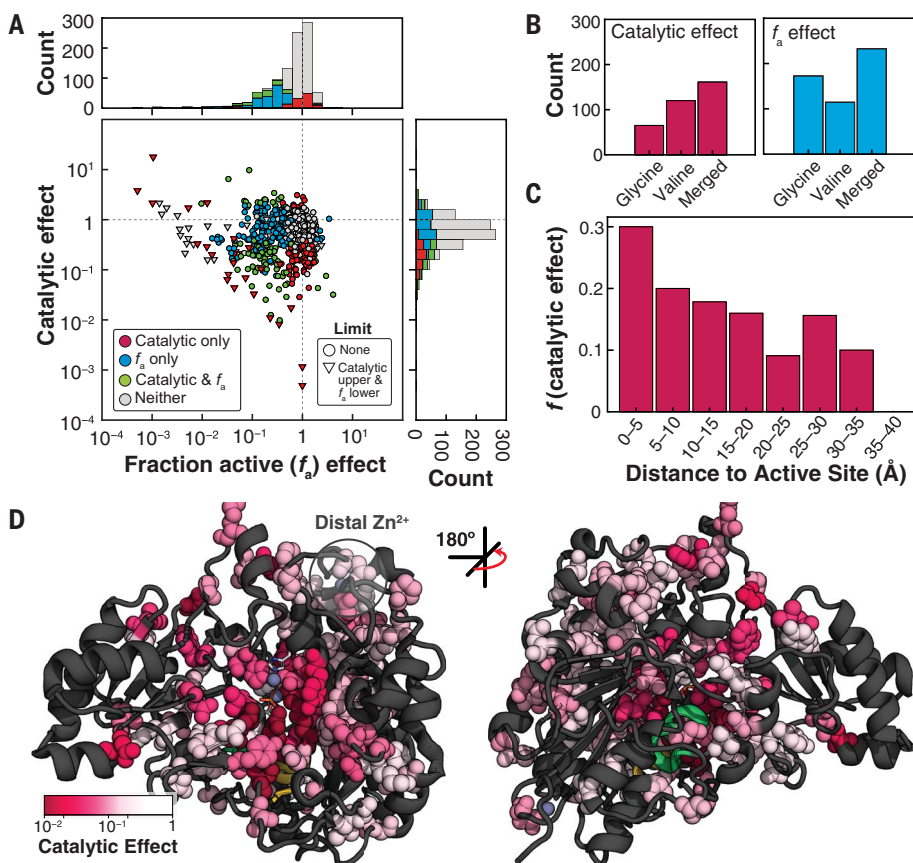

Fig. 4. Separating mutational effects on catalysis from changes in the fraction of active enzyme (f_a). (**A**) Catalytic and f_a effects (relative to wild type) for all variants colored by effect. Distributions of effects are projected along axes (stacked). (**B**) Counts of mutants with catalytic ($P < 0.05$) and f_a effects ($P < 0.01$) for the valine and glycine libraries, and total number of positions with effects (out of 525, "Merged") (fig. S45). (**C**) Fraction of non–active-site mutants within each distance bin with deleterious catalytic effects. $P < 0.05$; 5-Å bins (table S8). (**D**) Positions with deleterious catalytic effects on the PafA structure [161; "Merged" set in (B)]. Residues with deleterious effects (greater than fivefold down from wild type) are colored according to the most deleterious effect and shown as spheres.

enzymes possess a reservoir of allosteric potential and suggests that HT-MEK can be used to identify regions that are potential sites for allosteric inhibitors and drugs (39–43).

PafA has three of four nonterminal auxiliary domains (ADs) found within the AP superfamily, referred to as ADs 2 to 4 (supplementary text S5), that sit around the universally conserved Rossmann fold (Fig. 5F and fig. S52). ADs 2 and 4 are present in both AP superfamily phosphate mono- and diesterases, whereas AD 3 contains K162 and R164 and is considerably

more extensive in the monoesterases (figs. S52 and S53 and table S12). Despite these apparent functional and evolutionary differences, FC1 effects are found in all three ADs to a similar extent and with similar magnitudes (fig. S54 and table S13). The largest FC1 effect outside of the active site or second shell comes from a solvent-exposed surface residue, D473, within AD 4 (Fig. 5F, yellow); a dramatically larger effect for the valine substitution (>60-fold versus less than twofold for D473G) suggests that a change in local folding may allosteri-

cally disrupt the O2 site 20 Å away. Two nearby residues (L434 and I451) also exhibit >10-fold FC1 effects, and all three lie along a three-way interface between ADs 3 and 4 and the Rossmann core domain (Fig. 5F). Future multimutant cycle experiments (44–46) using HT-MEK may allow testing of whether mutations at interfaces result in more extensive conformational rearrangements than mutations within a single domain and dissect the underlying structural properties accounting for long-range effects.
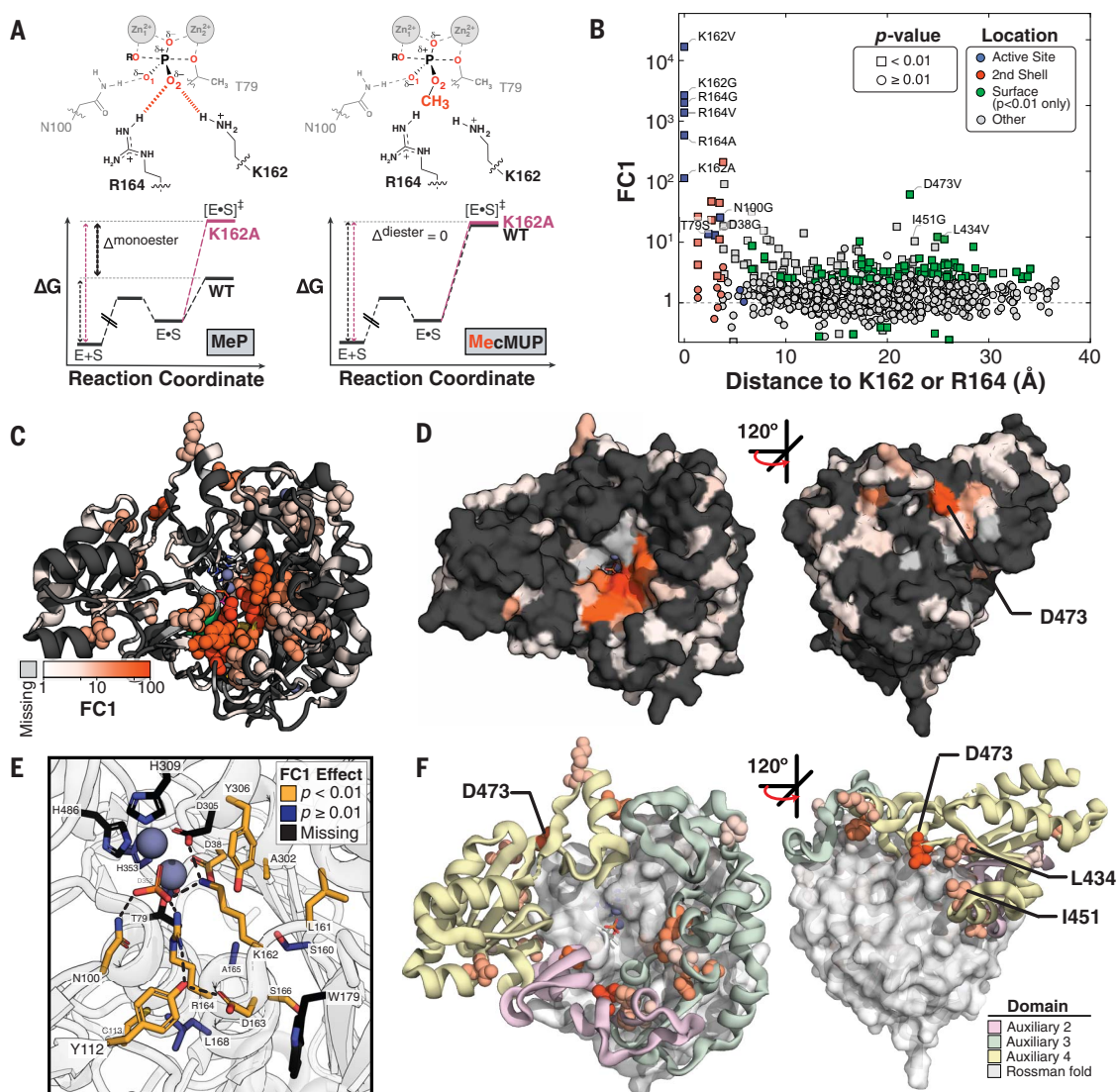
Fig. 5. FC1: Catalytic effects through the O_2 phosphoryl oxygen atom.
(**A**) Schematic of the transition states for reaction of (left) MeP and (right) MecMUP highlighting the methyl group on the O_2 phosphoryl atom of MecMUP (orange) and the O_2 interactions with K162 and R164. Thin black and pink arrows indicate the reaction barrier for wild type and a mutant PafA, respectively, for each reaction; the thick black arrow denotes their difference for the two reactions ($\Delta^{monoester}$ and $\Delta^{diester}$). The ratio of k_{cat}/K_m effects (the difference between $\Delta^{monoester}$ and $\Delta^{diester}$ in energy space) gives FC1. (**B**) FC1 values as a function of the minimum distance to K162 or R164, with active site, second shell, and significantly affected surface residues colored. (**C**) PafA positions with FC1

effects ($P < 0.01$) when mutated to valine or glycine, colored by FC1 magnitude. Effects greater than fivefold are shown as spheres, with ribbon coloring for positions with effects less than or equal to fivefold. (**D**) PafA surface representation with FC1 effects colored as in (C). (**E**) FC1 effects of active site and second-shell residues indicating residues with significant FC1 effects (yellow), without FC1 effects (blue), and without measurements (black) (table S9). (**F**) Distribution of FC1 effects within the Rossmann core and Auxiliary Domains. Distal (greater than or equal to third shell) positions with effects greater than fivefold shown as spheres colored corresponding to their auxiliary domain. Only effects in auxiliary domains are shown.

FC2 and FC3: Effects on phosphate affinity

To provide catalysis, enzymes must bind their transition states more strongly than their substrates; otherwise, the energetic barrier for the reaction and reaction rate would remain the same as in solution (47, 48). Enzymes must also limit the binding of substrates and products to allow sufficient turnover in the presence of higher substrate and product concentrations (49, 50). Ground-state destabilization has been considered a possible mechanism for achieving this, and there is evidence for ground-state destabilization by PafA and other AP superfamily members through electrostatic repulsion between the anionic nucleophile, T79 in

PafA, and the negatively charged phosphoryl oxygens (Fig. 6A) (16, 35, 51). For PafA, mutating T79 to serine increases the affinity for P_i, the reaction's product and a ground-state analog, by 100-fold, whereas in E. coli AP, the nucleophile S102G mutation increases affinity more than 1000-fold (16, 35); the S102G mutation ablates the nucleophile entirely, and the chemically conservative T79S substitution presumably allows greater mobility, reducing electrostatic repulsion (16). We therefore expect that mutations that perturb this mechanism will increase P_i affinity and define the second FC, FC2, as strengthened P_i binding ($K_i^{mutant}/K_i^{WT} < 1$).

Conversely, active-site residues typically make both ground-state and transition-state interactions, so their removal weakens binding and diminishes catalysis, in some instances to a similar extent ("uniform binding") and in others preferentially destabilizing the transition state (52–55). As expected, mutations to the PafA active-site residues that interact with the phosphoryl O1 and O2 oxygen atoms (N100, K162, and R164) weaken P_i binding, but N100 and K162 mutations have much larger effects on catalysis, indicating preferential transition-state stabilization (fig. S55) (16). We therefore define the third FC (FC3) as weakened P_i binding ($K_i^{mutant}/K_i^{WT} > 1$).
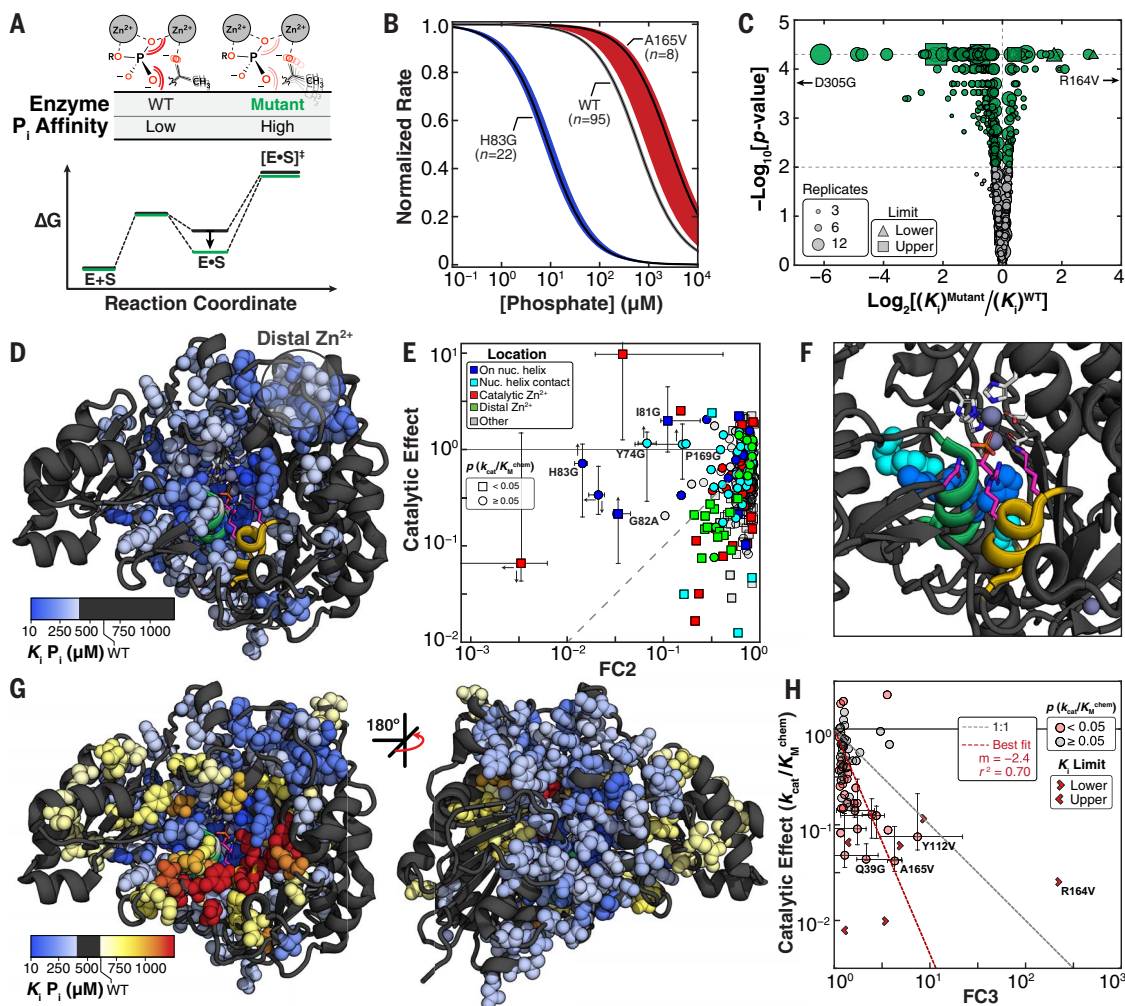
Fig. 6. FC2 and FC3: Mutational effects on affinity for P_i. (**A**) Reaction coordinate diagram for hypothetical mutant (green) with diminished destabilization relative to wild type (black) and schematic illustrating a model for an FC2 effect through increased flexibility of the T79 nucleophile. (**B**) Median P_i inhibition curves for wild type and two mutants with FC2 (H83G) and FC3 (A165V) effects. Colored regions denote 99% CIs on the medians of replicate measurements. (**C**) Volcano plot of K_i P_i effects for glycine and valine scan mutants. $P < 0.01$, green markers; $P \geq 0.01$, gray markers. (**D**) PafA structure with positions for which K_i P_i is >1.5-fold tighter than that of wild type (at $P < 0.01$) shown as spheres colored by the magnitude of the larger effect. The "nucleophile" helix (green) is largely obscured. (**E**) Scatter plot of catalytic and FC2 effects ($P < 0.01$) colored by location. Error bars correspond

to 95% CIs determined from bootstrapping; up and down arrows indicate lower and upper k_{cat}/K_m^{chem} limits, respectively; left and right arrows indicate upper and lower K_i P_i limits. (**F**) PafA active-site (pink residues) and the five mutants possessing large FC2 effects without k_{cat}/K_m^{chem} effects shown as spheres, corresponding to the labeled mutants in (E). (**G**) Front and back views of FC2 (blue spheres) and FC3 effects (red spheres). FC2 effects are those shown in (D), and FC3 effects are shown for $P < 0.01$. (**H**) Scatter plot of catalytic (k_{cat}/K_m^{chem}) and FC3 effects for mutants with FC3 effects ($P < 0.01$). Gray and red points indicate significant and insignificant k_{cat}/K_m^{chem} effects, respectively. Error bars correspond to 95% CIs, as in (E); red dashed line indicates the best fit line to mutants with significant FC3 and catalytic (k_{cat}/K_m^{chem}) effects, excluding active-site mutants.

To measure inhibition constants, we quantified rates of cMUP hydrolysis as a function of P_i concentration and fit observed initial rates to a competitive inhibition model (Fig. 6B and fig. S56). HT-MEK–determined inhibition constants agreed with previous off-chip measurements (fig. S57) and are of higher precision than kinetic constants because they are not sensitive to uncertainties in total enzyme concentration or active fraction (Fig. 6C; reports of inhibition measurements are available in data files S1 and S2).

We uncovered 331 mutants that increased P_i affinity (FC2) and 73 that decreased P_i affinity (FC3) (Fig. 6C and fig. S58). Thus, about one-third of all mutants measurably altered affinity, and four times as many mutations enhanced binding as weakened it. Because it is highly unusual to enhance function through random variation, this observation suggests that residues at many positions are evolutionarily selected to prevent tight P_i binding. Mutations with ground-state destabilization effects (FC2) were located in an extended yet spatially contiguous region that included the helix containing the T79 oxyanion ("nucleophile helix"), the catalytic Zn^{2+} ions, and the distal Zn^{2+} site (Fig. 6D and tables S14 and S15).

Catalytic effects (k_{cat}/K_m^{chem}) for mutants with FC2 effects ranged from insignificant to 100-fold reductions (Fig. 6E and fig. S59). For most, the catalytic effect was greater than the FC2 effect, which is consistent with functional roles beyond the prevention of too-tight binding. Nevertheless, five mutants of residues that form a spatially contiguous subregion on and adjacent to the nucleophile helix have the largest FC2 effects but little or no catalytic effect (Fig. 6E and F), suggesting that these mutations allow rearrangements to reduce electrostatic repulsion in the presence of bound substrate.

Mutants that weakened P_i binding (FC3 effects) were also located in a contiguous region (Fig. 6G). The largest FC3 effects were near active-site residues K162 and R164 and had larger effects on catalysis (k_{cat}/K_m^{chem}) than P_i binding (Fig. 6H, red, and figs. S60 to S62), presumably reflecting the greater constraints present in the transition state than in the ground state (56, 57).

Comparisons of the mutational effects across FC1 to FC3 reveal that many mutations outside the active site preferentially affect either FC1 or FC2. Further, several preferentially alter FC2 without dramatically altering FC1, reducing the fraction of active enzyme, or reducing overall catalysis (fig. S63). A small number of residues at the interface of FC2 and FC3 effects yield either an FC2 or FC3 effect depending on the substitution (fig. S64). The ability to selectively alter particular properties through specific mutations provides a potential starting point for attempts to engineer enzymes with

desired kinetic and thermodynamic constants and behaviors.

FC4: Rates of phosphoenzyme hydrolysis

Linking observed mutational rate effects to their physical and chemical origins requires knowledge of the step that is being observed: the rate-limiting step. Pre–steady-state approaches (such as stopped flow and rapid quench) are the gold standard for determining rates of individual reaction steps but do not readily scale to large libraries. For PafA, extensive prior mechanistic knowledge allowed us to determine mutational effects on individual PafA reaction steps for 992 PafA variants without requiring pre–steady-state measurements.

The steady-state kinetic constant k_{cat}/K_m can be limited by substrate binding or chemical cleavage of the substrate to form the covalent enzyme-phosphate (E–P) species (Fig. 7A, k_1 and $k_{chem,1}$ steps), and k_{cat} can be limited by hydrolysis of E–P ($k_{chem,2}$) or dissociation of product P_i ($k_{off,Pi}$) (Fig. 7A). We deconvoluted mutational effects on $k_{chem,2}$ using measurements of k_{cat} for cMUP and K_i for P_i, and these $k_{chem,2}$ effects define FC4 (supplementary text S6). Seven mutants changed the rate-limiting step from E–P hydrolysis to P_i release ($k_{off,Pi} < k_{chem,2}$) (fig. S65), which is consistent with the observation of naturally occurring alkaline phosphatases of the AP superfamily with either of these steps rate limiting (58).

Overall, we found 18 Val and 36 Gly mutants that reduce $k_{chem,2}$ (Fig. 7B and fig. S66), and these overlap substantially with those affecting k_{cat}/K_m^{chem} (table S16). In the simplest scenario, mutations reducing $k_{chem,2}$ would be a subset of those reducing k_{cat}/K_m^{chem} because k_{cat}/K_m^{chem} includes both a phosphoryl transfer chemical step and an additional step (binding). Consistent with this expectation, mutations affect either k_{cat}/K_m^{chem} alone or k_{cat}/K_m^{chem} and $k_{chem,2}$ to a similar or smaller extent (Fig. 7C).

Discussion

HT-MEK uses automatic valved microfluidics to carry out high-throughput expression, purification, and comprehensive biochemical characterization of enzymes at unprecedented scale and mechanistic depth. In the future, HT-MEK can be applied to the vast number of enzymes whose activity can be monitored through fluorescence directly or with a coupled assay. In particular, the ability to sensitively detect P_i (18) renders HT-MEK immediately applicable to adenosine triphosphatases, guanosine triphosphatases, helicases, protein chaperones, polymerases (with pyrophosphatase present), and many others.

Deleterious effects throughout the enzyme and at the surface, with many of the largest effects distant from the active site, mirror results obtained from DMS studies of other enzymes that often reveal distal effects on product
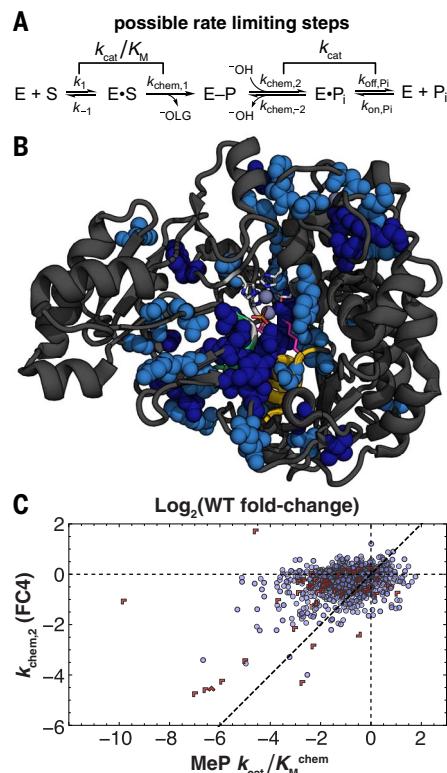
Fig. 7. FC4: Mutational effects on phosphoenzyme intermediate hydrolysis. (**A**) Schematic of PafA catalytic cycle with possible rate-limiting steps under saturating (k_{cat}) and subsaturating (k_{cat}/K_m) conditions. (**B**) Structure showing positions with decreased $k_{chem,2}$ ($P < 0.05$ and $P < 0.1$ in dark and light blue, respectively) upon mutation (spheres). Positions lacking an estimate of $k_{chem,2}$ for both substitutions are light gray ribbons, and positions without significant effects are black ribbons. (**C**) Scatter plot comparing mutational effects on phosphoenzyme hydrolysis ($k_{chem,2}$, or FC4) and k_{cat}/K_m^{chem}. Limits are indicated with red chevrons pointing in limit directions (supplementary text S6).

formation or organismal fitness (12, 59–61). However, the ability of HT-MEK to provide quantitative measurements of a battery of kinetic and thermodynamic parameters yields an atlas of PafA functional "architecture" with unprecedented detail (table S17). Regions of residues with similar catalytic signatures extending throughout the enzyme affirm that the enzyme beyond the active site is not a passive, monolithic scaffold but rather contributes to function in multiple ways.

Comparing patterns of mutational effects between FCs reveals that these architectural solutions are idiosyncratic. For example, K162/R164, and the T79 nucleophile, sit at the N-termini of their respective helices, but the largest FC2 effects were on and around the nucleophile helix, whereas FC1 effects were not similarly distributed around the monoesterase helix. This

difference perhaps reflects a need for more interactions to secure the nucleophile helix against ground-state electrostatic repulsive forces. Twice as many glycine as valine mutations increased P_i binding (fig. S58), potentially because side-chain ablation allows structural rearrangements to reduce electrostatic repulsion without disrupting favorable binding interactions.

The need for interactions throughout the enzyme for optimal catalysis highlights the challenges inherent in de novo enzyme design because it suggests that computational efforts must consider in detail the interactions of a prohibitively large number of residues. The detailed architectural maps provided by HT-MEK can focus computational and experimental mutagenesis by distinguishing catalytic from folding effects and by identifying residues and regions that affect particular aspects of catalysis.

The presence of a long-lived inactive population of many PafA variants emphasizes how different underlying (and unexpected) biophysical mechanisms can compromise an enzyme's activity and thus the need to explicitly decouple effects on folding and catalysis to understand function. Analogous behavior observed for these mutants in *E. coli* is consistent with selective pressure to avoid misfolding in vivo and adds to growing evidence that kinetic factors affect stable protein expression in cells (*20*, *30*, *62–64*). We speculate that highly stable proteins such as PafA and other secreted enzymes may be more prone to forming kinetically trapped states.

Although PafA functional regions have superficial similarities to sectors and other measures of evolutionary covariation and coconservation (*65–69*), our data report on sequence-function relationships in ways that sequence analyses alone cannot. On the basis of a metagenomic alignment of 14,505 PafA-like AP superfamily sequences (fig. S67 and supplementary text S7), we found that observed k_{cat}/K_m values correlate only modestly with conservation at a given position (figs. S68 and S69). Positions at which mutations promote misfolding tend to be most conserved, suggesting that the largest selective pressure at the residue level may be related to ensuring correct folding, whereas correlations between individual FCs and conservation are weak, with many large effects at nonconserved residues (figs. S69 and S70). For PafA, selective pressures will differ with the available P_i at the organism's physical location, whether available P_i varies temporally, and whether there is competition for P_i from other organisms in the same ecological niche. Residues with critical roles in responding to these adaptive pressures would be poorly conserved, changing frequently despite being tightly linked to survival. These results underscore the need for in-depth functional studies to uncover the relationships between the evolution of particular functions and conservation.

Enzymes are the targets of many therapeutics, are altered in genetic diseases, serve as tools for molecular biology, and play critical roles in industrial processes. Its rapidity and low cost make HT-MEK a powerful tool for future applications across all these areas. In basic research, the large, highly quantitative datasets provided by HT-MEK can greatly extend and even supplant traditional SDM approaches for the initial characterization of new enzymes and for in-depth mechanistic investigation. Combined with recent advances in gene synthesis, HT-MEK can rapidly functionally characterize metagenomic variants, providing a critically needed dimension to phylogenetic analyses. In medicine, we anticipate that HT-MEK will rapidly determine the functional effects of human enzyme allelic variants of unknown relevance identified from sequence data and systematically identify candidate allosteric surfaces within currently "undruggable" therapeutic target enzymes. We anticipate HT-MEK contributing to these and still more areas of basic and applied biology, medicine, and engineering.

Materials and methods summary

A full description of materials and methods is provided in the supplementary materials. Briefly, we fabricated microfluidic devices and aligned them to plasmid DNA arrays as previously described (*70–72*), then connected devices to a custom pneumatic manifold (*73*) and imaged using a fully automated fluorescence microscope. The plasmid DNAs coded for a set of specified PafA variants, each with a C-terminal eGFP tag. The PafA-eGFP variants were expressed in vitro on-chip and immobilized to Button valves by means of antibodies to eGFP (*72*), in parallel, before quantification of the enzyme concentration within each chamber by using an eGFP standard curve. This entire process takes ~10 hours to complete. Before measuring kinetics of substrate hydrolysis for cMUP, MeP, and MecMUP, we determined product standard curves for each chamber by introducing a series of product concentrations [cMU for cMUP and MecMUP hydrolysis and P_i in the presence of 30 μM of modified PBP (*18*) for MeP hydrolysis], imaging the device after each addition. To measure rates of hydrolysis, we flowed a single concentration of substrate into the device with the Button valves closed and then started reactions by (i) closing the Sandwich valves to isolate individual reaction chambers and (ii) opening the Button valves to expose the immobilized enzyme within each chamber to substrate. We quantified initial rates of product formation for each chamber by imaging over time across the entire device and converting measured intensities to product concentrations using standard curves. To determine k_{cat}, K_m, and k_{cat}/K_m for cMUP hydrolysis and k_{cat}/K_m for MeP and MecMUP hydrolysis, we quantified initial rates of product formation for a series of increasing substrate concentrations, scaled rates by measured enzyme concentration, and fit to a simple Michaelis-Menten model. For the nonfluorogenic MeP substrate, we detected formation of P_i by performing reactions in the presence of 30 μM PBP. To determine the K_i for P_i inhibition, we assayed rates of cMUP hydrolysis at 10 or 50 μM cMUP in the presence of increasing P_i concentration, fit initial rates as described above, and then fit these rates to a competitive inhibition model. For each assay type, we carried out multiple replicate experiments for each parameter for each mutant, allowing us to perform extensive quality control measurements and determine the statistical significance of effects.

REFERENCES AND NOTES

1. Q. Wang *et al.*, Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. *Nat. Commun.* **11**, 2539 (2020). doi: 10.1038/s41467-019-12438-5; pmid: 32461613
2. M. Claussnitzer *et al.*, A brief history of human disease genetics. *Nature* **577**, 179–189 (2020). doi: 10.1038/s41586-019-1879-7; pmid: 31915397
3. M. Lek *et al.*, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016). doi: 10.1038/nature19057; pmid: 27535533
4. A. Stein, D. M. Fowler, R. Hartmann-Petersen, K. Lindorff-Larsen, Biophysical and mechanistic models for disease-causing protein variants. *Trends Biochem. Sci.* **44**, 575–588 (2019). doi: 10.1016/j.tibs.2019.01.003; pmid: 30712981
5. J. B. Kinney, D. M. McCandlish, Massively parallel assays and quantitative sequence-function relationships. *Annu. Rev. Genomics Hum. Genet.* **20**, 99–127 (2019). doi: 10.1146/annurev-genom-083118-014845; pmid: 31091417
6. C. Blanco, E. Janzen, A. Pressman, R. Saha, I. A. Chen, Molecular fitness landscapes from high-coverage sequence profiling. *Annu. Rev. Biophys.* **48**, 1–18 (2019). doi: 10.1146/annurev-biophys-052118-115333; pmid: 30601678
7. J. Weile, F. P. Roth, Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. *Hum. Genet.* **137**, 665–678 (2018). doi: 10.1007/s00439-018-1916-x; pmid: 30073413
8. G. P. Lisi, J. P. Loria, Allostery in enzyme catalysis. *Curr. Opin. Struct. Biol.* **47**, 123–130 (2017). doi: 10.1016/j.sbi.2017.08.002; pmid: 28865247
9. K. A. Reynolds, R. N. McLaughlin, R. Ranganathan, Hot spots for allosteric regulation on protein surfaces. *Cell* **147**, 1564–1575 (2011). doi: 10.1016/j.cell.2011.10.049; pmid: 22196731
10. D. A. Keedy *et al.*, An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**, e36307 (2018). doi: 10.7554/eLife.36307; pmid: 29877794
11. D. M. Fowler *et al.*, High-resolution mapping of protein sequence-function relationships. *Nat. Methods* **7**, 741–746 (2010). doi: 10.1038/nmeth.1492; pmid: 20711194
12. C. L. Araya, D. M. Fowler, Deep mutational scanning: Assessing protein function on a massive scale. *Trends Biotechnol.* **29**, 435–442 (2011). doi: 10.1016/j.tibtech.2011.04.003; pmid: 21561674
13. J. K. Lassila, J. G. Zalatan, D. Herschlag, Biological phosphoryl-transfer reactions: Understanding mechanism and catalysis. *Annu. Rev. Biochem.* **80**, 669–702 (2011). doi: 10.1146/annurev-biochem-060409-092741; pmid: 21513457
14. P. M. Fordyce *et al.*, De novo identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. *Nat. Biotechnol.* **28**, 970–975 (2010). doi: 10.1038/nbt.1675; pmid: 20802496

15. S. J. Maerkl, S. R. Quake, A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315**, 233–237 (2007). doi: 10.1126/science.1131007; pmid: 17218526

16. F. Sunden *et al.*, Mechanistic and evolutionary insights from comparative enzymology of phosphomonoesterases and phosphodiesterases across the alkaline phosphatase superfamily. *J. Am. Chem. Soc.* **138**, 14273–14287 (2016). doi: 10.1021/jacs.6b06186; pmid: 27670607

17. M. M. Santoro, D. W. Bolen, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068 (1988). doi: 10.1021/bi00421a014; pmid: 3233195

18. M. Brune, J. L. Hunter, J. E. T. Corrie, M. R. Webb, Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase. *Biochemistry* **33**, 8262–8271 (1994). doi: 10.1021/bi00193a013; pmid: 8031761

19. P. J. O'Brien, D. Herschlag, Alkaline phosphatase revisited: Hydrolysis of alkyl phosphates. *Biochemistry* **41**, 3207–3225 (2002). doi: 10.1021/bi012116y; pmid: 11863460

20. R. M. Addabbo *et al.*, Complementary role of co- and post-translational events in de novo protein biogenesis. *J. Phys. Chem. B* **124**, 6488–6507 (2020). doi: 10.1021/acs.jpcb.0c03039; pmid: 32456434

21. J. M. Sturtevant, M. H. Yu, C. Haase-Pettingell, J. King, Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein. *J. Biol. Chem.* **264**, 10693–10698 (1989). doi: 10.1016/S0021-9258(18)81678-7; pmid: 2525128

22. M. Guo, Y. Xu, M. Gruebele, Temperature dependence of protein folding kinetics in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17863–17867 (2012). doi: 10.1073/pnas.1201797109; pmid: 22665776

23. D. H. Smith, J. King, Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein. III. Intensive polypeptide chains synthesized at 39°C. *J. Mol. Biol.* **145**, 653–676 (1981). doi: 10.1016/0022-2836(81)90308-9; pmid: 7265218

24. D. Baker, J. L. Sohl, D. A. Agard, A protein-folding reaction under kinetic control. *Nature* **356**, 263–265 (1992). doi: 10.1038/356263a0; pmid: 1552947

25. H. Im, M.-S. Woo, K. Y. Hwang, M.-H. Yu, Interactions causing the kinetic trap in serpin protein folding. *J. Biol. Chem.* **277**, 46347–46354 (2002). doi: 10.1074/jbc.M207682200; pmid: 12244055

26. C. Park, S. Marqusee, Probing the high energy states in proteins by proteolysis. *J. Mol. Biol.* **343**, 1467–1476 (2004). doi: 10.1016/j.jmb.2004.08.085; pmid: 15491624

27. S. Thompson, Y. Zhang, C. Ingle, K. A. Reynolds, T. Kortemme, Altered expression of a quality control protease in *E. coli* reshapes the in vivo mutational landscape of a model enzyme. *eLife* **9**, e53476 (2020). doi: 10.7554/eLife.53476; pmid: 32701056

28. T. Gidalevitz, V. Prahlad, R. I. Morimoto, The stress of protein misfolding: From single cells to multicellular organisms. *Cold Spring Harb. Perspect. Biol.* **3**, a009704–a009704 (2011). doi: 10.1101/cshperspect.a009704; pmid: 21536706

29. J. L. Sohl, S. S. Jaswal, D. A. Agard, Unfolded conformations of α-lytic protease are more stable than its native state. *Nature* **395**, 817–819 (1998). doi: 10.1038/27470; pmid: 9796818

30. S. S. Jaswal, J. L. Sohl, J. H. Davis, D. A. Agard, Energetic landscape of α-lytic protease optimizes longevity through kinetic stability. *Nature* **415**, 343–346 (2002). doi: 10.1038/415343a; pmid: 11797014

31. A. L. Watters *et al.*, The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell* **128**, 613–624 (2007). doi: 10.1016/j.cell.2006.12.042; pmid: 17289578

32. F. Bemporad *et al.*, Biological function in a non-native partially folded state of a protein. *EMBO J.* **27**, 1525–1535 (2008). doi: 10.1038/emboj.2008.82; pmid: 18451804

33. R. E. Jefferson, T. M. Blois, J. U. Bowie, Membrane proteins can have high kinetic stability. *J. Am. Chem. Soc.* **135**, 15183–15190 (2013). doi: 10.1021/ja407232b; pmid: 24032628

34. J. G. Zalatan, T. D. Fenn, D. Herschlag, Comparative enzymology in the alkaline phosphatase superfamily to determine the catalytic role of an active-site metal ion. *J. Mol. Biol.* **384**, 1174–1189 (2008). doi: 10.1016/j.jmb.2008.09.059; pmid: 18851975

35. L. D. Andrews, T. D. Fenn, D. Herschlag, Ground state destabilization by anionic nucleophiles contributes to the activity of phosphoryl transfer enzymes. *PLOS Biol.* **11**, e1001599 (2013). doi: 10.1371/journal.pbio.1001599; pmid: 23843744

36. D. Hilvert, Design of protein catalysts. *Annu. Rev. Biochem.* **82**, 447–470 (2013). doi: 10.1146/annurev-biochem-072611-101825; pmid: 23746259

37. M. I. Freiberger, A. B. Guzovsky, P. G. Wolynes, R. G. Parra, D. U. Ferreiro, Local frustration around enzyme active sites. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4037–4043 (2019). doi: 10.1073/pnas.1819859116; pmid: 30765513

38. K. L. Morley, R. J. Kazlauskas, Improving enzyme properties: When are closer mutations better? *Trends Biotechnol.* **23**, 231–237 (2005). doi: 10.1016/j.tibtech.2005.03.005; pmid: 15866000

39. O. Dagliyan, N. V. Dokholyan, K. M. Hahn, Engineering proteins for allosteric control by light or ligands. *Nat. Protoc.* **14**, 1863–1883 (2019). doi: 10.1038/s41596-019-0165-3; pmid: 31076662

40. J. W. McCormick, D. Pincus, O. Resnekov, K. A. Reynolds, Strategies for engineering and rewiring kinase regulation. *Trends Biochem. Sci.* **45**, 259–271 (2020). doi: 10.1016/j.tibs.2019.11.005; pmid: 31866305

41. K. Gunasekaran, B. Ma, R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? *Proteins* **57**, 433–443 (2004). doi: 10.1002/prot.20232; pmid: 15382234

42. J. A. Hardy, J. A. Wells, Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* **14**, 706–715 (2004). doi: 10.1016/j.sbi.2004.10.009; pmid: 15582395

43. G. R. Bowman, E. R. Bolin, K. M. Hart, B. C. Maguire, S. Marqusee, Discovery of multiple hidden allosteric sites by combining Markov state models and experiments. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2734–2739 (2015). doi: 10.1073/pnas.1417811112; pmid: 25730859

44. F. Sunden, A. Peck, J. Salzman, S. Ressl, D. Herschlag, Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site. *eLife* **4**, e06181 (2015). doi: 10.7554/eLife.06181; pmid: 25902402

45. P. J. Carter, G. Winter, A. J. Wilkinson, A. R. Fersht, The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (*Bacillus stearothermophilus*). *Cell* **38**, 835–840 (1984). doi: 10.1016/0092-8674(84)90278-2; pmid: 6488318

46. P. Carter, J. A. Wells, Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564–568 (1988). doi: 10.1038/332564a0; pmid: 3282170

47. D. Herschlag, A. Natarajan, Fundamental challenges in mechanistic enzymology: Progress toward understanding the rate enhancements of enzymes. *Biochemistry* **52**, 2050–2067 (2013). doi: 10.1021/bi4000113; pmid: 23488725

48. W. P. Jencks, Mechanism of enzyme action. *Annu. Rev. Biochem.* **32**, 639–676 (1963). doi: 10.1146/annurev.bi.32.070163.003231; pmid: 14140708

49. A. R. Fersht, *Enzyme Structure and Mechanism* (W.H. Freeman and Co., ed. 2, 1985).

50. W. P. Jencks, Binding energy, specificity, and enzymic catalysis: The Circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.* **43**, 219–410 (1975). doi: 10.1002/9780470122884.ch4; pmid: 892

51. L. D. Andrews, H. Deng, D. Herschlag, Isotope-edited FTIR of alkaline phosphatase resolves paradoxical ligand binding properties and suggests a role for ground-state destabilization. *J. Am. Chem. Soc.* **133**, 11621–11631 (2011). doi: 10.1021/ja203370b; pmid: 21692505

52. W. J. Albery, J. R. Knowles, Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry* **15**, 5631–5640 (1976). doi: 10.1021/bi00670a032; pmid: 999839

53. F. M. Menger, Analysis of ground-state and transition-state effects in enzyme catalysis. *Biochemistry* **31**, 5368–5373 (1992). doi: 10.1021/bi00138a018; pmid: 1606161

54. J. D. McCarter, M. J. Adam, S. G. Withers, Binding energy and catalysis. Fluorinated and deoxygenated glycosides as mechanistic probes of Escherichia coli (lacZ) β-galactosidase. *Biochem. J.* **286**, 721–727 (1992). doi: 10.1042/bj2860721; pmid: 1417731

55. S. A. Moore, W. P. Jencks, Formation of active site thiol esters of CoA transferase and the dependence of catalysis on specific binding interactions. *J. Biol. Chem.* **257**, 10893–10907 (1982). doi: 10.1016/S0021-9258(18)33908-5; pmid: 6955308

56. A. Whitty, C. A. Fierke, W. P. Jencks, Role of binding energy with coenzyme A in catalysis by 3-oxoacid coenzyme A transferase. *Biochemistry* **34**, 11678–11689 (1995). doi: 10.1021/bi00037a005; pmid: 7547900

57. Y.-L. Zhang *et al.*, Impaired transition state complementarity in the hydrolysis of O-arylphosphorothioates by protein-tyrosine phosphatases. *Biochemistry* **38**, 12111–12123 (1999). doi: 10.1021/bi990836i; pmid: 10508416

58. W. E. Hull, S. E. Halford, H. Gutfreund, B. D. Sykes, Phosphorus-31 nuclear magnetic resonance study of alkaline phosphatase: The role of inorganic phosphate in limiting the enzyme turnover rate at alkaline pH. *Biochemistry* **15**, 1547–1561 (1976). doi: 10.1021/bi00652a028; pmid: 4092

59. E. E. Wrenbeck, L. R. Azouz, T. A. Whitehead, Single-mutation fitness landscapes for an enzyme on multiple substrates reveal specificity is globally encoded. *Nat. Commun.* **8**, 15695 (2017). doi: 10.1038/ncomms15695; pmid: 28585537

60. J. R. Klesmith, J.-P. Bacik, E. E. Wrenbeck, R. Michalczyk, T. A. Whitehead, Trade-offs between enzyme fitness and solubility illuminated by deep mutational scanning. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2265–2270 (2017). doi: 10.1073/pnas.1614437114; pmid: 28196882

61. P. Bandaru *et al.*, Deconstruction of the Ras switching cycle through saturation mutagenesis. *eLife* **6**, e27810 (2017). doi: 10.7554/eLife.27810; pmid: 28686159

62. C. M. Dobson, Protein folding and misfolding. *Nature* **426**, 884–890 (2003). doi: 10.1038/nature02261; pmid: 14685248

63. M. S. Faber, E. E. Wrenbeck, L. R. Azouz, P. J. Steiner, T. A. Whitehead, Impact of in vivo protein folding probability on local fitness landscapes. *Mol. Biol. Evol.* **36**, 2764–2777 (2019). doi: 10.1093/molbev/msz184; pmid: 31400199

64. V. Zhao, W. M. Jacobs, E. I. Shakhnovich, Effect of protein structure on evolution of cotranslational folding. *Biophys. J.* **119**, 1123–1134 (2020). doi: 10.1016/j.bpj.2020.06.037; pmid: 32857962

65. N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, Protein sectors: Evolutionary units of three-dimensional structure. *Cell* **138**, 774–786 (2009). doi: 10.1016/j.cell.2009.07.038; pmid: 19703402

66. O. Rivoire, K. A. Reynolds, R. Ranganathan, Evolution-based functional decomposition of proteins. *PLOS Comput. Biol.* **12**, e1004817 (2016). doi: 10.1371/journal.pcbi.1004817; pmid: 27254668

67. V. H. Salinas, R. Ranganathan, Coevolution-based inference of amino acid interactions underlying protein function. *eLife* **7**, e34300 (2018). doi: 10.7554/eLife.34300; pmid: 30024376

68. T. Teşileanu, L. J. Colwell, S. Leibler, Protein sectors: Statistical coupling analysis versus conservation. *PLOS Comput. Biol.* **11**, e1004091 (2015). doi: 10.1371/journal.pcbi.1004091; pmid: 25723535

69. P. Creixell *et al.*, Hierarchical organization endows the kinase domain with regulatory plasticity. *Cell Syst.* **7**, 371–383.e4 (2018). doi: 10.1016/j.cels.2018.08.008; pmid: 30243563

70. P. M. Fordyce *et al.*, Basic leucine zipper transcription factor Hac1 binds DNA in two distinct modes as revealed by microfluidic analyses. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E3084–E3093 (2012). doi: 10.1073/pnas.1212457109; pmid: 23054834

71. D. D. Le *et al.*, Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3702–E3711 (2018). doi: 10.1073/pnas.1715888115; pmid: 29588420

72. A. K. Aditham, C. J. Markin, D. A. Mokhtari, N. DelRosso, P. M. Fordyce, High-throughput affinity measurements of transcription factor and DNA mutations reveal affinity and specificity determinants. *Cell Syst.* **12**, 112–127.e11 (2021). doi: 10.1016/j.cels.2020.11.012; pmid: 33340452

73. K. Brower *et al.*, An open-source, programmable pneumatic setup for operation and automated control of single- and multi-layer microfluidic devices. *HardwareX* **3**, 117–134 (2018). doi: 10.1016/j.ohx.2017.10.001; pmid: 30221210

74. C. J. Markin *et al.*, Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics. Open Science Repository (2021); doi: 10.17605/OSF.IO/QRN3C

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

Science

Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics

C. J. Markin, D. A. Mokhtari, F. Sunden, M. J. Appel, E. Akiva, S. A. Longwell, C. Sabatti, D. Herschlag and P. M. Fordyce

Go big or you'll get lost

Rational mutagenesis is a common approach to investigating or engineering enzyme function in vitro, but the ease with which one can manipulate protein sequences belies many pitfalls in connecting sparse activity data to an enzyme's true functional landscape. Using a high-throughput platform, Markin *et al.* expressed, purified, and performed an array of kinetic measurements on a target esterase, collecting data from >1000 mutations spanning the entire protein (see the Perspective by Baumer and Whitehead). Protein misfolding into an inactive state, rather than decreased equilibrium stability, was a crucial factor in negatively affected variants spread throughout the protein. When combined with prior mechanistic understanding and structures, four "functional components" help to rationalize the otherwise complex spatial pattern of effects of mutations on different aspects of enzyme function, all of which would be invisible from mutagenesis of just a few residues.

Science, abf8761, this issue p. eabf8761; see also abj8346, p. 391

ARTICLE TOOLS	http://science.sciencemag.org/content/373/6553/eabf8761
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2021/07/21/373.6553.eabf8761.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/373/6553/391.full
REFERENCES	This article cites 115 articles, 16 of which you can access for free http://science.sciencemag.org/content/373/6553/eabf8761#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service