

## **Title: Evolutionary-Scale Enzymology Enables Biochemical Constant Prediction Across a Multi-Peaked Catalytic Landscape**

**Authors: Duncan F. Muir<sup>1,2</sup>, Garrison P. R. Asper<sup>1</sup>, Pascal Notin<sup>3,4</sup>, Jacob A. Posner<sup>1,5</sup>, Debora S. Marks<sup>3,6</sup>, Michael J. Keiser<sup>7,8,9</sup>, Margaux M. Pinney<sup>1,10</sup>**

### **Affiliations:**

1. Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA, USA
2. Program in Biophysics, University of California, San Francisco, San Francisco, CA, USA
3. Department of Systems Biology, Harvard Medical School, Boston, MA, USA
4. Department of Computer Science, University of Oxford, Oxford, UK
5. Department of Biology, San Francisco State University, San Francisco, CA, USA
6. Broad Institute of Harvard and MIT, Cambridge, MA, USA
7. Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA
8. Institute for Neurodegenerative Diseases, University of California, San Francisco, San Francisco, CA, USA
9. Bakar Computational Health Sciences Institute, University of California, San Francisco, San Francisco, CA, USA
10. Valhalla Fellow, University of California San Francisco, San Francisco, CA, USA

### **Abstract:**

Quantitatively mapping enzyme sequence-catalysis landscapes remains a critical challenge in understanding enzyme function, evolution, and design. Here, we expand an emerging microfluidic platform to measure catalytic constants— $k_{\text{cat}}$  and  $K_M$ —for hundreds of diverse naturally occurring sequences and mutants of the model enzyme Adenylate Kinase (ADK). This enables us to dissect the sequence-catalysis landscape’s topology, navigability, and mechanistic underpinnings, revealing distinct catalytic peaks organized by structural motifs. These results challenge long-standing hypotheses in enzyme adaptation, demonstrating that thermophilic enzymes are not slower than their mesophilic counterparts. Combining the rich representations of protein sequences provided by deep-learning models with our custom high-throughput kinetic data yields semi-supervised models that significantly outperform existing models at predicting catalytic parameters of naturally occurring ADK sequences. Our work demonstrates a promising strategy for dissecting sequence-catalysis landscapes across enzymatic evolution and building family-specific models capable of accurately predicting catalytic constants, opening new avenues for enzyme engineering and functional prediction.

### **Main Text:**

#### **Introduction.**

Natural selection has shaped the catalytic parameters of enzymatic reactions, determining the rates and specificities that govern nearly all biological processes. The relationship between enzyme sequence and catalytic parameters is often conceptualized as a landscape, as first described by Fisher (1) and Wright (2), that is traversed through mutational “walks” (Fig. 1A). The topologies of sequence-catalysis landscapes, their responses to environmental changes, and their underlying molecular mechanisms are central to understanding how enzymes evolved to perform catalysis in distinct physicochemical environments (3–7). Furthermore, a quantitative mapping of these enzyme sequence-catalysis relationships can aid in developing predictive models for enzyme function and guide enzyme optimization.

This inherent complexity of protein sequence space has often restricted sequence-function landscapes to simplified illustrations to conceptualize—but not quantify—evolutionary processes (e.g., Fig. 1A) (7). Current experimental explorations of enzyme sequence-catalysis landscapes are limited to narrow regions of the underlying protein sequence space. Combinatorial mutational studies have uncovered the role of intramolecular epistasis in shaping the evolutionary paths available to enzymes (8–12) but typically explore fewer than five sites (9, 10, 13–20). Alternatively, deep mutational scanning (DMS) studies measure a large number of sequence-fitness relationships (21–23). However, this approach explores

a dense, local region of sequence space—all single mutations from wild-type—missing alternative peaks separated by multiple mutations. Furthermore, most mutational effects in DMS explorations are neutral or deleterious (24–29), limiting the potential to discover enzymes with improved catalytic parameters. In contrast, genomic sequencing has shown that orthologous enzymes exhibit significantly more sequence variation than is typically explored experimentally (30), potentially yielding substantial differences in catalytic properties from adaptation to diverse environments (e.g., temperature, [salt], pH). Nevertheless, catalytic activities are typically measured for few (typically <10) orthologs, leaving the topology of sequence-catalysis landscapes at the scale of naturally occurring sequences biochemically underexplored and underdetermined.

Existing approaches to model sequence-catalysis relationships are limited by a dearth of catalytic parameters for a given enzyme and narrow explorations of sequence space. In the absence of large-scale catalytic datasets, unsupervised deep-learning methods can learn complex distributions that may approximate these high-dimensional landscapes (31). Protein Language Models (PLMs), trained on the millions to billions of publicly available naturally occurring protein sequences, can often predict protein structure (32, 33). Beyond structure, functional properties of enzymes, such as thermodynamic stability, folding kinetics, solubility, substrate specificity, and catalytic rate are presumably encoded in sequence, suggesting that sequence-only models may also learn sequence-function relationships. Furthermore, early PLM work demonstrated that the models learned an underlying representation of protein sequence that organized amino acids by biochemical property (34), and recent work has suggested their utility in optimizing protein function directly (35, 36). Nevertheless, while PLM likelihoods (37–40), supervision on top of PLMs (41), and other unsupervised sequence-only models (42, 43) can predict single mutation fitness effects, evaluating these representations of sequence-function landscapes remains challenging due to fundamental data challenges. Namely, these models are often trained or evaluated on DMS datasets, which typically use a fitness-linked proxy to estimate enzyme function. Various molecular properties can influence these fitness values, such as catalytic activity, stability, and inhibition, and these multiple low-level biochemical properties may buffer or amplify each other (44). This reliance on complex, indirect labels for protein function prediction may underlie the “yawning chasm” in our ability to accurately predict the relationships between sequence and specific enzyme properties (6). Consequently, there is a pressing need for quantitative datasets of gold-standard catalytic parameters collected under consistent conditions to test whether sequence-only models can accurately predict enzyme properties across diverse evolutionary and environmental contexts.

Here, we quantitatively map and dissect the topology of an evolutionary-scale sequence-catalysis landscape. We leverage an emerging microfluidics platform, High-Throughput Microfluidic Enzyme Kinetics (HT-MEK) (45), to measure the Michaelis-Menten parameters  $k_{cat}$  (catalytic constant),  $K_M$  (Michaelis-Menten constant), and  $k_{cat}/K_M$  (catalytic efficiency) for hundreds of diverse naturally occurring and mutant sequences of the model enzyme adenylate kinase (ADK) (Fig. 1B, Fig. 1C). We dissect the topology and navigability of this sequence-catalysis landscape, showing that it is rugged, with at least three global peaks organized by structural motifs, yet this landscape remains navigable over long evolutionary timescales through path-dependent mechanisms. We address long-standing evolutionary hypotheses, for example, demonstrating that environmental temperature is not a primary determinant of enzyme catalytic rates and that thermophilic enzymes are not slower than their mesophilic counterparts. Finally, we show that a state-of-the-art unsupervised PLM organizes ADK sequence space by structure but not catalytic activity. Supervised and semi-supervised deep-learning models trained on the dataset newly collected herein to predict  $k_{cat}$  outperform prior models trained on only public databases, with performance improving as the training dataset size grows. We provide a strategy to develop models that directly learn from family-specific sequence-catalysis landscapes and accurately predict fundamental biochemical constants of interest.

## Results and Discussion.

### *High-throughput measurement of ADK catalytic parameters across the bacterial and archaeal tree of life*

The enzyme ADK is essential and ubiquitous across the tree of life. ADK regulates the cellular balance of adenylate currencies (ATP, ADP, and AMP), reversibly catalyzing the phosphoryl transfer between two ADP molecules (or ATP and AMP) using a  $Mg^{2+}$  cofactor that acts as an electrostatic “pivot” during the chemical step (Fig. 1B) (46). ADK consists of

three domains: CORE, AMP-binding (AMP), and LID (Fig. 1D). The CORE domain forms most of the nucleotide-binding site and provides many catalytic residues needed for phosphoryl transfer. Prior NMR experiments with ADKs from diverse bacteria and archaea have revealed that the rate-limiting step of ADK catalysis is a large conformational change that opens two domains, LID and AMP, enabling product release (47–49). These conformational dynamics have made ADK a valuable model for studying the relationship between enzyme dynamics and catalytic turnover (Fig. 1B) (46–48, 50–53).

To map the naturally occurring sequence-catalysis landscape of ADK, we gathered orthologous ADK sequences from bacteria and archaea from protein sequence databases (see Methods). We selected 193 orthologs based on sequence diversity and environmental adaptations. Specifically, we included ADK sequences from organisms adapted to divergent temperatures, as temperature adaptation has been suggested to drive differences in ADK catalytic parameters (47, 54–58) (see Methods, Table S1). The resulting ADK sequence library has an average pairwise sequence identity of 42% and median optimal growth temperatures ( $T_{\text{Growth}}$ ) ranging from 9 to 96 °C (Fig. 1C, Fig. S1). Despite low average sequence identity, most substrate-contacting residues are >90% conserved, including the key catalytic arginine residues R36, R88, R123, R156, and R167 (*E. coli* ADK [ecADK] numbering used throughout) (59) (Fig. S2). Since most of our orthologs lack experimentally derived structures, we used AlphaFold2 (60) to predict their structures (Fig. 1D). Most ortholog structural differences were found in the predicted conformation of the LID domain, which varied in its openness. However, none of the predictions were as open as the fully open conformation of ecADK (61) (Fig. 1B, D).

Measuring Michaelis-Menten kinetics for  $>10^2$  enzyme sequences is intractable with traditional bench-top biochemistry methods. To overcome this challenge, we redeveloped an emerging microfluidic technology, HT-MEK (45), to assay all orthologous ADK enzyme kinetics in parallel under identical conditions. Briefly, we recombinantly expressed, purified, and assayed all 193 ADK orthologs on a single HT-MEK device (Fig. 2A). ADK enzymatic activity was monitored “on-chip” by coupling the formation of ATP to the production of NADPH (62), which was detected using time-lapse microscopy (Fig. 2B). All ADK orthologs were tagged at the C-terminus with a flexible Ser-Gly linker and eGFP, which enabled the purification of each ADK ortholog and subsequent quantification of per-chamber ADK concentrations (Fig. 2B). Observed initial rates collected over a range of ADP concentrations were normalized by ADK concentration and fit to the Michaelis-Menten equation, obtaining  $k_{\text{cat}}$  and  $K_M$  values for each ADK ortholog (Fig. 2C-D). Across the 1792 chambers in our device, we obtain an average of 6-7 biological replicates for each ADK ortholog per on-chip experiment. Out of the 193 ADK orthologs, 181 expressed and displayed catalytic activity above background, and of those, 175 had bounded  $K_M$  values under our assay conditions (Table S2, Methods). For ADK orthologs in our library that have been previously characterized in the literature under common conditions and assays, we observe a strong correlation between those  $k_{\text{cat}}$  measurements and our on-chip measurements ( $r^2 = 0.96$ , Fig. 2E). Multiple controls provide confidence that these naturally occurring ADKs are natively-folded on-chip (Supplementary Text S1, Fig. S3-6).

Remarkably, while the  $k_{\text{cat}}$  values of all natural enzymes vary by  $\sim 5$  orders of magnitude (63), the  $k_{\text{cat}}$  values within the ADK family alone span at least three orders of magnitude, ranging from 1–803  $\text{s}^{-1}$  (Fig. 3A, B, Fig. S7). Thus, while the slowest measurable ADK in our library had a  $k_{\text{cat}}$  of 1  $\text{s}^{-1}$ , certain sequence combinations found in nature achieve catalytic turnover at least two orders of magnitude higher, approaching the activities of some of the fastest known enzymes (63). Thus, naturally occurring sequences performing analogous functions across different organisms can exhibit catalytic activities spanning orders of magnitude despite having superimposable structures, experimental and predicted, and nearly identical active sites (Fig. 1D, Fig. S2). This finding underscores the challenge of predicting catalytic function from sequence alone.

Given the wide variation in catalytic rates, we next explored whether ADK variants close in evolutionary space exhibit similar activities. Mapping the catalytic parameters  $k_{\text{cat}}$ ,  $K_M$ , and  $k_{\text{cat}}/K_M$  to the organismal taxonomic tree shows very little visual organization (Fig. 3A). To quantify this relationship between phylogenetic distance and catalytic rate, we computed the phylogenetic signal, which quantifies the autocorrelation between distance on a phylogenetic tree and continuous traits (Methods) (Fig. 3C-D, Fig. S8) (64). We quantified this relationship for  $k_{\text{cat}}$ ,  $K_M$ ,  $k_{\text{cat}}/K_M$ , and  $T_{\text{Growth}}$  and observed that  $k_{\text{cat}}$  values show a significant positive correlation over short phylogenetic distance (Fig. 3D), albeit weaker than  $T_{\text{Growth}}$  phylogenetic signal (Fig 3D). However, across medium to long phylogenetic distances,  $k_{\text{cat}}$  is decorrelated with phylogeny

(Fig 3C), with high  $k_{\text{cat}}$  values often interspersed among comparatively low ones. Similar behaviors are observed for  $K_M$  and  $k_{\text{cat}}/K_M$  (Fig. S8). This result suggests that high catalytic activity has independently evolved multiple times during ADK evolution along distinct lineages. Thus, computational models for predicting catalytic activity must encompass multiple evolutionary—and potentially mechanistic—routes to high activity.

***Many thermophilic orthologs remain highly active at mesophilic temperatures, and psychrophilic ADKs are not catalytically superior***

Next, we explored whether the distinct environmental niches and selection pressures of each organism were responsible for the variation and evolutionary distribution we observed in ADK catalytic rates. We focused on temperature, a pervasive environmental factor reported to drive adaptive changes in enzyme catalytic rates (54, 62, 65–73). Discussions of activity-stability trade-offs propose that stabilizing mutations in thermophilic enzymes increase rigidity during natural evolution, suppressing activity-promoting dynamics (65–68, 71, 74–78). Accordingly, psychrophilic enzymes are frequently reported to have higher  $k_{\text{cat}}$  values relative to their mesophilic and thermophilic counterparts (65–68, 76). However, these findings typically rely on comparisons of just a pair of sequences, and recent analyses of enzyme kinetic data in BRENDA (79) question the generality of this model (80).

We systematically tested this model across a wide range of growth temperatures under consistent conditions. When  $k_{\text{cat}}$  is plotted against the optimal growth temperature of the corresponding organisms, we do not observe the expected negative correlation but rather a weak positive correlation (Fig. 3E). Similarly weak trends are observed for  $K_M$  versus growth temperature, providing evidence against related hypotheses that enzymatic  $K_M$  values are increased in cold-adapted enzymes (Fig. S9) (65, 68, 71, 81). Since growth temperature correlates with melting temperature and thermodynamic stability, our data show that ADK activity and stability *do not* universally trade off during natural evolution. Consequently, psychrophilic ADKs are not universally catalytically superior, and thermophilic ADKs are not catalytically limited. In fact, some of the fastest ADKs come from thermophilic organisms, potentially because high ADK activity is needed at increased temperatures to regenerate the ATP pool to combat its thermal lability (82). These findings are particularly intriguing for ADK, which has been central to the long-standing debate on the trade-off between thermodynamic stability and catalytic activity, and show that natural evolution can jointly optimize stability and activity (47, 54, 62, 66, 83–85). The joint optimization of both properties is perhaps facilitated by the independent folding of the CORE and LID domains, allowing the CORE to evolve high thermodynamic stability while the LID retains the necessary flexibility for domain opening along the reaction path (86). Our findings underscore the importance of examining sequence-catalysis landscapes on a broad scale. While small, localized studies may suggest the existence of an activity-stability trade-off (65, 66, 68, 71, 74–78), the larger sequence space reveals that this trend does not hold universally for natural ADK sequences. Furthermore, our results strongly indicate that enzymatic catalytic rates are not under strong selection during temperature adaptation. Instead, recent bioinformatic trends observed in enzyme sequences across a wide range of  $T_{\text{Growth}}$  values are likely driven primarily by changes in stability rather than activity (87).

***The evolutionary-scale ADK sequence-catalysis landscape is rugged, with multiple structural peaks achieving distinct activity levels***

While protein “sequence space” is often discussed in conceptual terms, its high dimensionality (20 x sequence length) makes its concrete visualization challenging. To further dissect the topology of this landscape, we visualized the sequence space as a graph, where each node corresponds to an orthologous ADK sequence, and edges connect these nodes, weighted by the number of edits in a multiple sequence alignment (MSA). Using our 175 observed sequences, we traversed this subset of ADK sequence space by minimizing the total edit distance, forming a Minimum Spanning Tree (MST). This MST visually represents the rugged sequence-catalysis landscape of naturally occurring ADKs (Fig. 4A).

Since the ADK MST shows little organization by catalytic activity, we next explored whether ADK structural differences might organize it. The LID of ADK exists in three general forms: (1) “Lidless” ADKs, characterized by unstructured loops of varying lengths, present across all domains of life (Fig. 4B, top); (2) “Zn<sup>2+</sup>-binding” lids, which feature a structural Zn<sup>2+</sup>-binding site typically formed by four Cys residues, found in bacteria and archaea (Fig. 4B, middle); and (3) “H-Bond” lids,

which possess an extended hydrogen bond network and are found almost exclusively in bacteria (Fig. 4B, bottom) (88–90). Visualizing the ADK MST reveals that these three lid types—Zn<sup>2+</sup>-binding, H-Bond, and Lidless—each cluster into distinct neighborhoods, or "peaks" (Fig. 4A). A closer examination of these structural peaks reveals that while the two lidded peaks span similar dynamic ranges of catalytic activity, encompassing both the fastest and slowest ADKs, the Lidless sequences are significantly slower (Fig. 4C). Given that Lidless ADKs span all domains of life, while Zn<sup>2+</sup>-binding and H-Bond lids are limited to bacteria and archaea or bacteria alone, respectively (90), lid domains may have been acquired later in ADK evolution, potentially as a mechanism to enhance catalytic rates.

ADKs adapted to different environmental temperatures inhabit different structural peaks within the landscape. ADKs from psychrophilic organisms ( $T_{\text{Growth}} < 25$  °C) have mostly H-Bond lid types, whereas ADKs from thermophilic organisms ( $T_{\text{Growth}} > 50$  °C) have nearly exclusively Zn<sup>2+</sup>-binding lids (Fig. 4D, Fig. S10). The Zn<sup>2+</sup>-binding lid may be favored in thermophilic ADKs because it provides extra stability to the domain at high temperatures, consistent with prior studies that show an increase in melting temperature ( $T_m$ ) by 15 °C when installing a Zn<sup>2+</sup> binding site in the H-Bond lid of *E. coli* ADK (91). In contrast, Zn<sup>2+</sup> can be limiting in marine environments where many psychrophilic organisms live, potentially favoring the H-Bond over the Zn<sup>2+</sup>-binding lid (92). Intriguingly, thermophilic ADKs are predominantly located at internal nodes within the MST, indicating they share more sequence similarity with other ADKs compared to their mesophilic and psychrophilic counterparts (Fig. S10). This result aligns with previous observations of highly thermostable consensus sequences in enzyme families (93–98). Thus, while the temperature of an organism's environment does not appear to directly drive changes in catalytic properties (Fig. 3E), it does influence the distribution of organisms across this evolutionary-scale sequence-catalysis landscape, confining psychrophilic and thermophilic ADKs to specific neighborhoods of distinct lid structures (Fig. 4D, Fig. S10).

#### ***Navigating between peaks: mutational walks and extra-dimensional bypasses***

Given that all three lid types appear in nature, these peaks must have been accessible to one another throughout evolution. To explore the viability of transitioning between these lid-type peaks in our landscape, we first focused on the Zn<sup>2+</sup>-binding and H-Bond peaks. We generated combinatorial mutations to swap the four key residues that constitute the Zn<sup>2+</sup>-binding motif—Cys130, Cys133, Cys150, and Cys153 (CCCC) in *G. stearothermophilus* (gsADK)—with the corresponding residues in the H-Bond motif—His126, Ser129, Asp146, and Thr149 (HSDT) in ecADK (Fig. 4B, Fig. S11). We measured  $k_{\text{cat}}$  and  $K_M$  for all possible mutations that interconvert the Zn<sup>2+</sup>-binding and H-Bond motifs in gsADK and ecADK backgrounds in a single high-throughput experiment. When installing a Zn<sup>2+</sup>-binding motif in the *E. coli* ADK LID, the final Zn<sup>2+</sup>-binding motif (CCCC) retains 74% of the activity of the wild-type H-Bond motif (HSDT), consistent with previous studies (91), and there are mutational pathways without highly deleterious intermediates (Fig. 5A). Given the feasibility of a mutational walk from the H-Bond to the Zn<sup>2+</sup>-binding lid, we anticipated that the reciprocal walk in gsADK—starting with the Zn<sup>2+</sup>-binding motif (CCCC) and ending with the H-Bond motif (HSDT)—should also be possible. Surprisingly, nearly all mutational trajectories in this direction encounter an unfavorable intermediate or "pit," leading to a large loss of activity or an inexpressible variant (Fig. 5B, Table S3). This non-reciprocal behavior aligns with differences in the three-dimensional context of the two structural motifs. While the CCCC motif coordinates the Zn<sup>2+</sup> ion independently of neighboring residues, the H-Bond motif is more complex, involving additional interactions with second-shell residues, presumably contributing to the observed epistasis of the HSDT motif (Fig. S12). Supporting this model, the H-Bond lid occupies a narrower region of sequence space compared to the Zn<sup>2+</sup>-binding lid (79% vs. 65% average sequence identity, respectively, Fig. S13). This indicates that the sequence contexts available for the lid beyond the HSDT/CCCC motifs are more constrained in H-Bond lids than in Zn<sup>2+</sup>-binding lids.

We next explored whether navigation between these peaks could be achieved through "extra-dimensional bypasses" (i.e., whole LID swaps), thereby avoiding unfavorable valleys (5). Substituting the ecADK H-Bond LID into gsADK results in a highly active ADK that is even faster than wild-type gsADK (Fig. 5B). Interestingly, the reverse experiment, where the entire LID of ecADK is swapped into gsADK, leads to a dramatic decrease in activity, dropping below the level of any of the incremental cysteine mutants. These results suggest that to avoid inactive intermediates, whole LID swaps are necessary

to convert a Zn<sup>2+</sup>-binding lid into an H-Bond lid, whereas, in the reverse direction, it is more favorable to convert an H-Bond lid to a Zn<sup>2+</sup>-binding lid through single mutational steps.

Considering the existence of Lidless ADKs, which may represent the ancestral form of all ADKs given their presence across all domains of life (90), we explored navigating our landscapes through the insertion or deletion of the LID. Since we found Lidless ADKs to be significantly slower than their lidded counterparts (Fig. 4C), we asked whether their activity could be “rescued” by inserting a LID. Given the variability in the degree of “lidlessness”, we selected two Lidless ADKs for our study: *O. acuminata* ADK (ocADK), which has a seven amino-acid loop and expressed well but showed activity below our detection limit in the initial assay, and *D. deserti* ADK (ddADK), which has a longer 17 amino-acid loop and displayed above-average activity among Lidless ADKs. We generated chimeras incorporating the LIDs from gsADK, bsADK, ecADK, and vcADK into both ocADK and ddADK backgrounds. For ocADK, we found that three out of four LID chimeras successfully rescued catalytic activity above our detection limit (Fig. 5C). In contrast, for ddADK, which already had modest activity, we instead observed a decrease in  $k_{\text{cat}}$  with a commensurate decrease in  $K_M$  for measurable lid insertions (Fig. 5E). Therefore, LID insertion does not always increase  $k_{\text{cat}}$ , complicating the prediction of ADK activity, which depends on LID dynamics but not solely on the LID sequence due to potential functional coupling between these domains (55).

### ***Contributions from multiple peaks catalytically impair the ADK consensus sequence***

Consensus sequences have garnered interest for engineering proteins with enhanced stability while preserving catalytic activity, though the retention of activity depends on the specific natural enzyme used as a benchmark (93–98). For ADK, we anticipated that a consensus sequence incorporating elements from all three peaks would produce suboptimal amino acid combinations. Indeed, a previously constructed ADK consensus sequence that spans the landscape has a  $k_{\text{cat}}$  value of 5 s<sup>-1</sup>, making it slower than 91% of the naturally occurring sequences we measured (Fig. 3A) (94). To further test this hypothesis, we independently constructed consensus sequences from each of the three peaks. We observed that the internal consensus sequences for ADKs with H-Bond and Zn<sup>2+</sup>-binding LIDs were 47- and 1.6-fold faster in  $k_{\text{cat}}$  compared to the consensus sequence for the entire family (Fig. 5E). The larger increase in  $k_{\text{cat}}$  for the H-Bond lid consensus sequence aligns with the narrower sequence space for this lid type (62% vs. 45% internal pairwise sequence identity for H-Bond and Zn<sup>2+</sup>-binding LID orthologs, respectively, Fig. S14). The internal consensus sequence for the Lidless ADKs exhibited activity below our detection limit, perhaps due to the variability in loop length among Lidless ADKs (Fig. 5E). These results emphasize the importance of functional coupling between residue positions and the role of intramolecular epistasis in shaping sequence-catalysis landscapes (8–12).

### ***Changes in ADK dynamics tune activity across billions of years of evolution***

Since high ADK activity has evolved multiple times in distinct structural contexts, we investigated the molecular mechanisms underlying high activity in each peak to determine if they are general or unique. Prior work in ecADK has shown that interconversion between the closed and open states involves the local unfolding of the LID and that this conformational equilibrium can be tuned with osmolytes and mutations (Fig. 1B) (48, 50, 90, 99, 100). In particular, urea, which can stabilize the unfolded state of proteins, favors the more expanded, open forms of ecADK, increasing  $k_{\text{cat}}$  by 1.7-fold at 2 M urea (99). We hypothesized that if changes in conformational dynamics were driving the differences in ADK catalysis, we would observe this capacity for conformational tuning across most naturally occurring ADK sequences. A closer examination of ADK activity under gradually increasing concentrations of urea (0.0–2.0 M) supports this hypothesis: most naturally occurring ADKs exhibited activation by urea, including for all three lid types, suggesting that the capacity for conformational tuning has been conserved in multiple structural contexts (Fig. 6A, 6B, Fig. S6). Previous reports of activation of thermophilic enzymes by denaturants were interpreted as increasing motions of these presumably overly rigid enzymes (101–103). Here, we systematically demonstrate that low concentrations of urea activate most naturally occurring ADKs, regardless of their parent organisms' growth temperature, challenging the model of rigidity-activity trade-offs during temperature adaptation (65, 66, 68, 71, 74–78).

To identify mutations that influence ADK conformational dynamics throughout evolution, we turned to orthologs with high sequence similarity but large differences in catalytic rates. We focused on ADKs from *V. cholerae* (vcADK) and *P.*

*profundum* (ppADK), as they exhibited the largest ratio of fold-change in  $k_{cat}$  to edit-distance—a >50-fold difference in activity over 23 sequence differences—indicating the steepest activity “cliff” in our landscape (Fig. 4A, Fig. 6C). We hypothesized that functionally-relevant sequence positions would be in the LID. There are two key residue differences in the LID of ppADK and vcADK: Ala/Pro at position 128 and Asn/Val at position 135 (Fig. 6C). Indeed, the double mutation Ala128Pro/Asn135Val increases ppADK  $k_{cat}$  by 21%, while the reverse mutation, Pro128Ala/Val135Asn, decreases vcADK  $k_{cat}$  by 11% (Fig. 6D, Methods). An additional CORE mutation, Ala22Glu, was also selected because its proximity to the LID led us to hypothesize that the increased bulkiness and charge of the Glu side chain could destabilize the enzyme's closed state by electrostatic repulsion (Fig. 6C). The Ala22Glu mutation increased ppADK  $k_{cat}$  by 28%, while the reverse Glu22Ala mutation decreased activity in vcADK by a reciprocal 29% (Fig. 6D). Together, the Ala22Glu/Pro128Ala/Val135Asn mutations account for 49% of the activity difference between ppADK and vcADK (Fig. 6D).

Next, we assessed the generality of these activation mechanisms across the ADK landscape. Pro128 is frequently found in ADKs and shows a statistically significant association with high  $k_{cat}$  across diverse ADK sequences (Fig. 6E). We selected 20 H-Bond and Zn<sup>2+</sup>-binding ADKs and either mutated the residue at position 128 to proline or, in cases where position 128 was already proline, we mutated P128 to alanine. In the Xaa-to-Pro mutations, we primarily observed activating effects in ADKs with H-Bond lids (Fig. 6F). In Zn<sup>2+</sup>-binding ADKs, the activating effect of Pro was less pronounced, although most still showed increased activity (Fig. 6F). This smaller effect may be due to Pro128's position between Cys130 and Cys133 (*B. subtilis* ADK numbering), which coordinate the structural Zn<sup>2+</sup> (Fig. S12). In contrast, mutational effects at position 22 do not generalize to other orthologs (Fig. S15). Thus, conformational tuning can regulate ADKs separated by billions of years of evolution, whether through mutations or small molecule solutes.

### ***PLMs organize ADK sequence space by structure but not by catalytic activity***

We next explored the landscape of our naturally occurring ADK sequences as learned by the state-of-the-art model ESM-2 (32) to evaluate the sequence-catalysis relationship of PLMs. We fed representative ADK sequences from sequence databases (~5,000 sequences, see Methods) into the pre-trained 650-million-parameter ESM-2 model to obtain fixed-length embeddings for each ADK ortholog after mean-pooling (Methods). The continuous nature of these embeddings allows us to visualize the traditional concept of landscape using dimensionality reduction techniques like UMAP (104). We observed that the ADKs measured in our library broadly cover the landscape generated from ADK ESM-2 embeddings, with the various lid types forming distinct visual clusters (Fig. 7A). This structural organization aligns with our MST-derived sequence-catalysis landscape, with distinct lid-type peaks (Fig. 4A). Consistent with our MSA, H-Bond ADKs had the highest internal sequence identity among the major lid types, and occupy a narrower distribution in this dimensionality reduction (Fig. 7A, Fig. S14). Nevertheless, UMAP notoriously distorts distances between true clusters during dimensionality reduction (105), so we sought to further characterize the organization of ADK ESM-2 embeddings using quantitative metrics. To quantify the lid-type organization, we hierarchically clustered the embeddings based on euclidean distance (Fig. S16) and computed the adjusted mutual information score (AMI) by lid-type label (Fig. 7B). We found that while ESM-2 embedding clusters exhibit higher lid-type AMI than random, they retain less information than a one-hot encoded MSA (One-Hot MSA) clustered by Hamming distance (Fig. 7C, Fig. S17).

We next investigated whether ESM-2 had instead learned a representation that reflects similarity in enzymatic activity. We find that the ESM-2 landscape appears visually rugged when colored by  $k_{cat}$  (Fig. 7D). Furthermore, quantifying  $k_{cat}$  organization using the continuous manifold metric of trustworthiness (Methods) reveals that both ESM-2 and One-Hot MSA exhibit essentially random organization in their nearest five neighbors with respect to  $k_{cat}$  (0.52 and 0.50 respectively) (Fig. 7F). Consequently, naturally occurring sequences close in the ESM-2 landscape often have vastly different  $k_{cat}$  values. Similar relationships apply to  $K_M$  and  $k_{cat}/K_M$  (Fig. S18). The trustworthiness of ESM-2 embeddings relative to One-Hot MSA remains largely unchanged when the number of neighbors increases (Fig. S19). Given that ESM-2 has retained some lid-type organization, we also computed trustworthiness within lid types and found that this improves trustworthiness slightly (0.59 for Lidless and H-Bond) when not considering the global landscape (Fig. 7F). In a principal component analysis (PCA) on the ESM-2 embeddings (Fig. S20), we observed weak explained variance between the first principal component and  $k_{cat}$  ( $r^2 = 0.097$ ) (Fig. S21). PC1 may reflect lid-type differences, as removing the significantly slower Lidless

ADKs (Fig. 4C) weakens the regression (Fig. S21). Thus, while ESM-2 captures some high-level structural organization of ADKs, it fails to meaningfully encode the complex relationship between sequence and catalytic activity, perhaps because the co-evolutionary relationships learned are stronger for structure than for catalysis. These results highlight the limitations of current PLMs in predicting enzyme function from naturally occurring sequences alone.

### ***ADK $k_{cat}$ and $K_M$ prediction improve with increasing experimental data and supervision on top of PLMs***

Supervised models built on top of pre-trained PLMs can enhance the prediction of mutational fitness effects compared to a zero-shot regime (106–108). However, fitness labels represent an aggregate of multiple underlying biochemical properties, such as catalytic activity and thermodynamic stability, and many models are specifically trained to predict the effect of a single mutation (37). Recently, there have been attempts to model sequence-catalysis relationships directly, with models like DLKcat (109) and TurNuP (110) trained on catalytic turnover rates collated from published literature in the BRENDA database (79). Our dataset of naturally occurring ADK sequences with measured  $k_{cat}$  values provides a valuable test set to evaluate the extent to which models like DLKcat have learned the sequence- $k_{cat}$  relationship. When predicting  $k_{cat}$  from our sequences and substrate information (Methods), DLKcat performs poorly (spearman rho = -0.09, Fig. S22). We speculate that this poor performance may arise from challenges inherent to the BRENDA training data, collected under many experimental conditions, requiring scaling to standard temperature and pH. Furthermore, BRENDA has suffered data consistency issues from mis-annotations (111).

Apart from dataset inconsistencies, predicting  $k_{cat}$  for any naturally occurring enzyme sequence is inherently challenging, given the vast range of enzymatic activities (e.g., phosphorylation, hydrolysis, oxidation) and mechanisms. Instead, we considered whether we could use our newly collected experimental data to build an ADK-specific model that would be more predictive than a large model like DLKcat. We trained classic lightweight machine learning models—Random Forest (RF), Support Vector Regressor (SVR)—on increasing sub-samples of our dataset (20, 40, 60... 140 examples) and evaluated performance on a held-out 20% of our dataset for  $k_{cat}$  and  $K_M$  (Methods). For  $k_{cat}$  prediction, both SVR and RF outperform DLKcat with as few as 20 data points (mean RF spearman rho=0.14, mean SVR spearman rho=0.23, across 30 bootstrapped training set samplings). Performance steadily improved with increasing training data (Fig. 8A).

Having demonstrated that our newly collected experimental data enables lightweight ML models to outperform an existing deep-learning model trained on large literature databases for predicting  $k_{cat}$ , we next evaluated state-of-the-art semi-supervised approaches that combine pLMs with our experimental labels to further improve predictions. Specifically, we trained several variants of ProteinNPT (PNPT) (107), a pseudo-generative model that learns a joint representation of protein sequences and property annotations, as well as a convolutional linear regression (CLR) baseline (107). We experimented with different underlying pLM embeddings—ESM2 (32), Tranception (112), MSA Transformer (113)—and prediction targets— $k_{cat}$  only,  $K_M$  only, or both plus growth temperature and lid type. ProteinNPT outperformed the lightweight regressors in all scenarios, particularly when leveraging Tranception embeddings, and its performance also improved when trained on a greater number of experimental labels (Fig. 8A-B). Additionally, the model variant predicting all targets simultaneously generally outperformed single-target models for  $k_{cat}$  and  $K_M$  (Fig. S23), demonstrating its ability to leverage relationships between different properties, as  $k_{cat}$  and  $K_M$  are correlated for naturally occurring ADKs (Fig. S24), a trend previously observed for mutants of ecADK (48).

### **Conclusions and implications**

We redeveloped an emerging microfluidic method, HT-MEK, to measure the catalytic constants  $k_{cat}$  and  $K_M$  for hundreds of naturally occurring and mutant ADK sequences. This dataset offered a unique view of the sequence-catalysis landscape on an evolutionary scale, revealing highly diverse catalytic parameters for extant ADKs. Analysis of this data uncovered a rugged topology with at least three global peaks, which remain navigable over long evolutionary timescales through distinct, path-dependent mechanisms, including single mutations and extra-dimensional bypasses. Our findings underscore the importance of examining sequence-catalysis landscapes on a broad scale. While small-scale comparisons have led to the longstanding hypothesis that adaptation to environmental temperature drives differences in catalytic rates, we show that



thermophilic enzymes are not universally slower than their psychrophilic and mesophilic counterparts. A similar approach could be applied to assess the generality of other mechanistic and evolutionary hypotheses in enzymology.

PLMs predict protein structure by training on millions of protein sequences. Nevertheless, we show that enzymes with highly similar structures and analogous biochemical functions can exhibit vastly different catalytic parameters, hindering machine and deep-learning models that predict enzyme catalytic function from sequence alone. While current PLMs capture structural distinctions between naturally occurring variants of the same enzyme, they fail to accurately represent catalytic activity, perhaps because distinct structural solutions to high ADK activity have evolved over billions of years. This work underscores the need to integrate experimental data across a sequence-catalysis landscape with machine and deep-learning models to improve the prediction of catalytic parameters.

Supervised models trained on this study's kinetic dataset outperform models using broad but sparse existing catalytic databases. Strategically selecting sequences to experimentally characterize catalytically for model training with Bayesian optimization or active learning approaches could improve the trade-off between model performance and training set size. Furthermore, we envision training semi-supervised generative models of sequence and function on data from many experimental assays at once to learn diverse enzyme sequence-catalysis landscapes. These models could generate novel sequences that potentially surpass the limits of enzyme activity that evolution has explored thus far.

### **Acknowledgments:**

We thank members of the Pinney and Keiser laboratories, Siyuan Du, Stephanie Crilly, and Tony Capra for discussion and comments on the manuscript. We thank members of the Fordyce and Herschlag laboratories for helpful discussions and experimental advice. We acknowledge P. Suzuki for the design of the PS1.8K devices. D.F.M. thanks T. Bangalter, G.M. de Homem-Christo, D. Smith, and J. Vernon for additional support. **Funding:** This work was funded by a National Institutes of Health (NIH) grant (DP5OD033413), support from the Valhalla Foundation, and CZI grant DAF2018-191905 (DOI 10.37921/550142lkcjzw; M.J.K.) from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation (DOI 10.13039/100014989). D.F.M. was supported by a UCSF Discovery Fellowship. P.N. was supported by a Chan Zuckerberg Initiative Award (Neurodegeneration Challenge Network, CZI2018-191853). D.S.M. holds a Ben Barres Early Career Award from the Chan Zuckerberg Initiative as part of the Neurodegeneration Challenge Network (CZI2018-191853) and is supported by a NIH Transformational Research Award (TR01 1R01CA260415). **Author contributions:** D.F.M. M.M.P. and M.J.K. designed the study. D.F.M. G.P.R.A. and J.A.P. cloned the ADK library. D.F.M. and G.P.R.A. performed *in vitro* experiments. D.F.M. processed and analyzed kinetic data, and performed downstream computational, phylogenetic and statistical analysis. D.F.M. performed analyses of ESM-2 embeddings and machine learning experiments. P.N. performed deep-learning experiments with ProteinNPT with help from D.S.M. D.F.M. and M.M.P. wrote the manuscript with input from all authors. **Competing interests:** D.S.M. is an advisor for Dyno Therapeutics, Octant, Jura Bio, Tectonic Therapeutic and Genentech and a cofounder of Seismic. **Data and materials availability:** Summary tables of all measured kinetic parameters for each ADK variant are provided in the Supplementary Materials. All kinetic data generated in this study will be available in a registered Open Science Foundation repository upon publication. Code used for image processing and fitting kinetic parameters, as well as computational analyses will be made available in a public GitHub repository upon publication. Raw images from time-lapse microscopy will be made available upon request.

### **List of Supplementary Materials:**

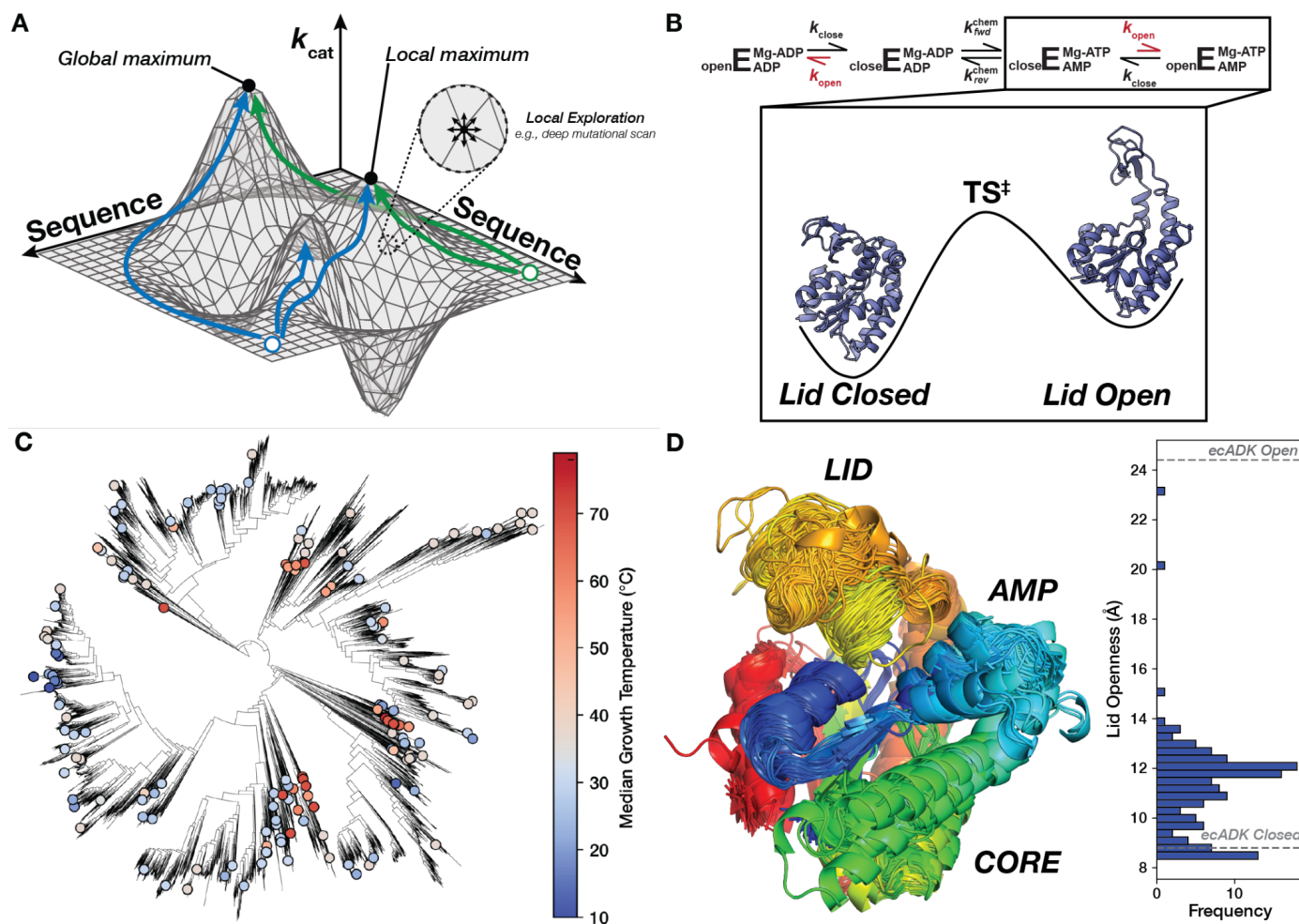
Materials and Methods

Supplementary Text

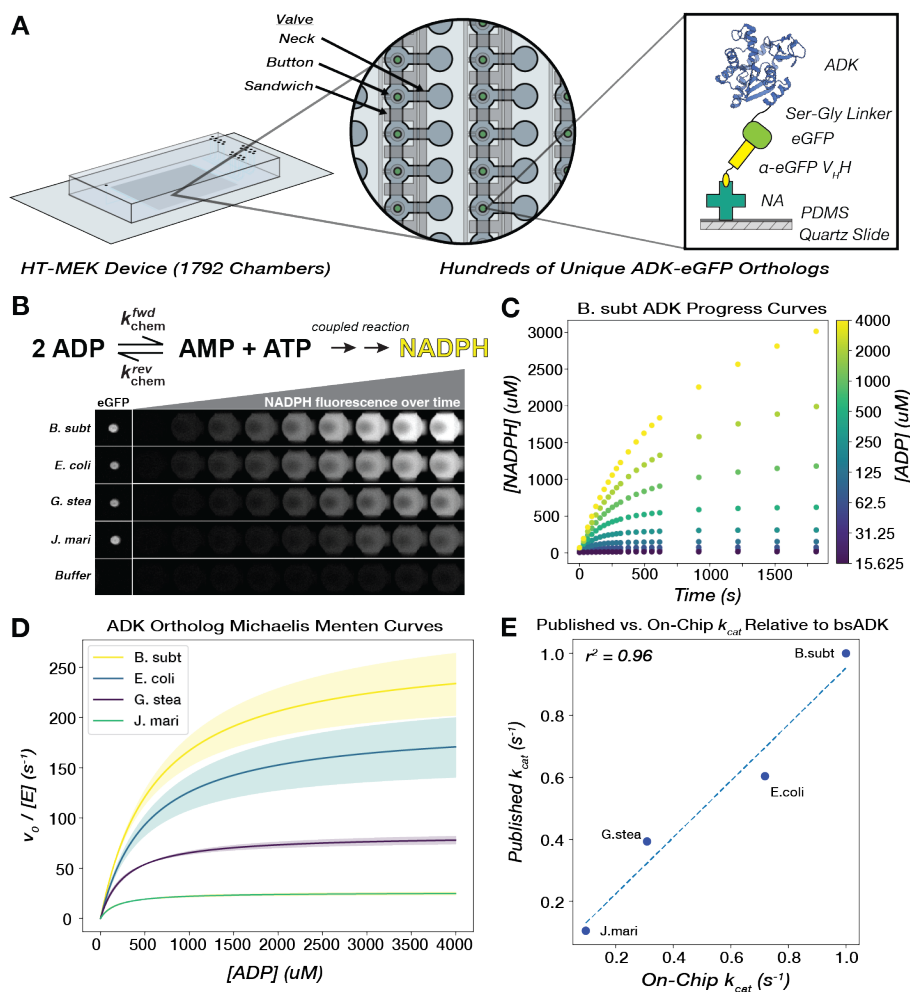
Tables S1 to S3

Figs. S1 to S28

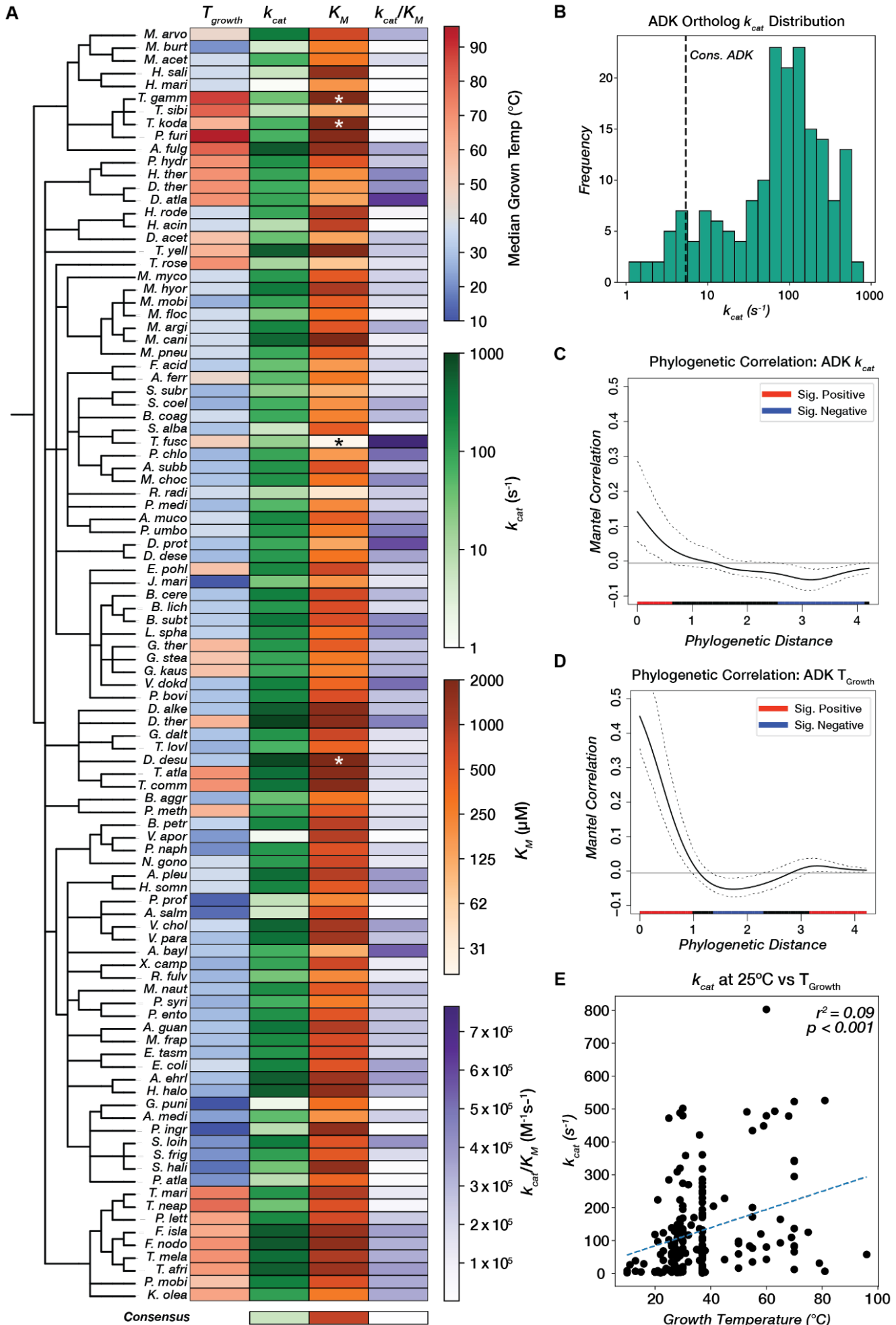
## Figures:



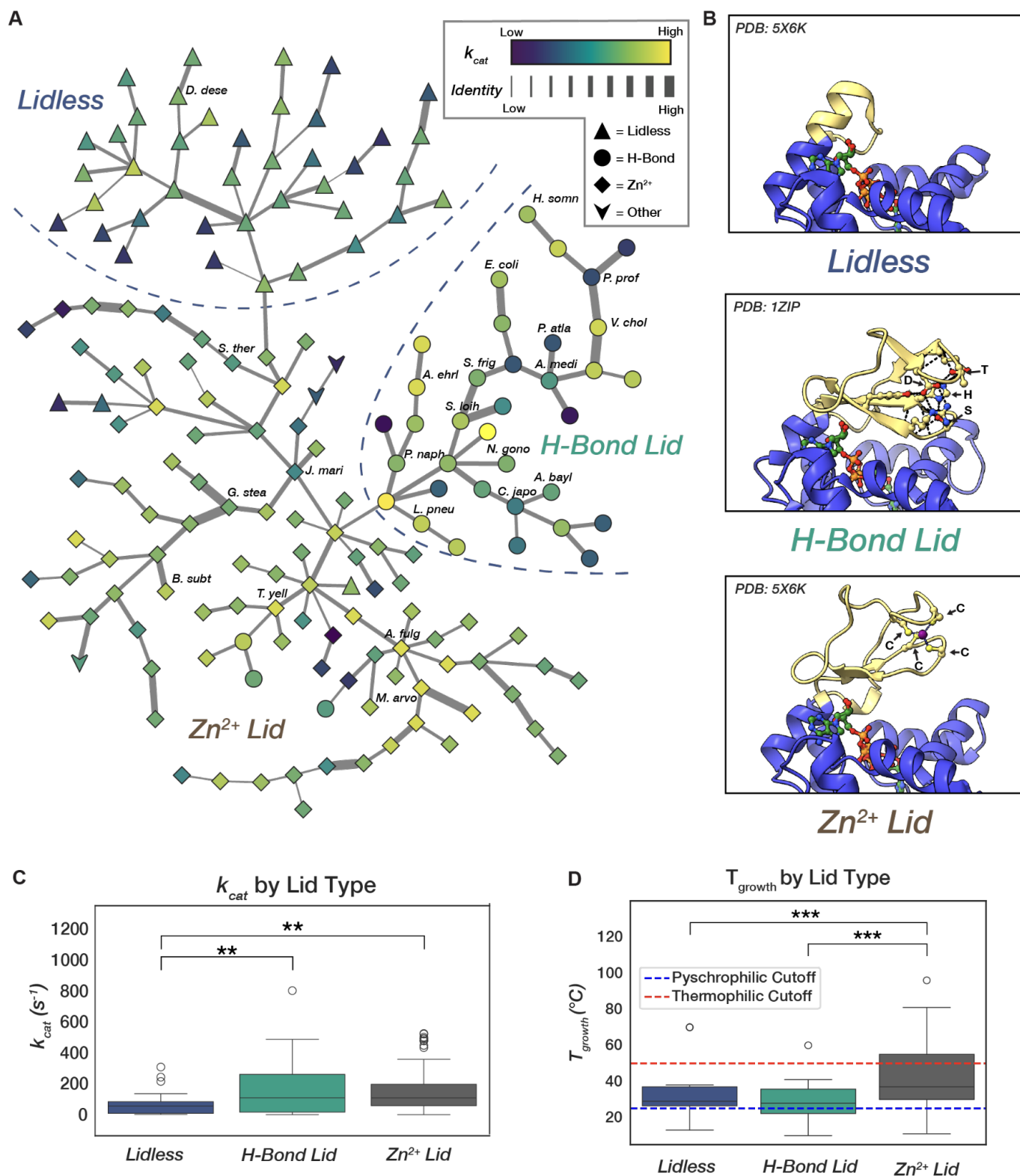
**Figure 1. Mapping the sequence-catalysis of Adenylate Kinase across evolution.** (A) Similar to a physical landscape, individual positions along the surface correspond to different enzyme sequences, with the height at each position representing their respective catalytic parameters. Sequences that occupy the highest “peaks” have the highest catalytic activities. The sequence space is effectively continuous for many sequences, allowing us to conceptualize the adaptive landscape as a two-dimensional surface in a three-dimensional space (2, 5, 7). While Deep-Mutational Scanning exercises assay many sequences, they explore a narrow region of sequence space (one mutation from wild-type) and are liable to be stuck in local optima. Green and blue paths show possible paths explorable by evolution that reach unique optima. (B) Schematic of the ADK reaction, where E represents the ADK enzyme, with the rate-limiting step,  $k_{open}$ , shown in red (46). Experimental structures of closed (PDB: 4AKE) and open ecADK (PDB: 1AKE) are depicted on a simplified energy landscape for this reaction step. (C) The ADK sequence library characterized herein (dots) spans the bacterial tree of life, with organisms adapted to optimal growth temperatures ranging from the coldest to the hottest environments on Earth (114). (D) Superimposition of AlphaFold2 predictions of the ADK orthologs, with sequences trimmed to align with ecADK in an MSA. Measurements between the c-alpha of positions 30 and 130 are histogrammed to display the range of observed LID “openness”, with measurements from experimental structures of ecADK in the open (PDB: 1AKE) and closed (PDB: 4AKE) conformations (61, 115).



**Figure 2. Michaelis-Menten parameters for hundreds of naturally occurring ADK sequences can be measured in parallel via high-throughput microfluidic devices.** (A) A single High-Throughput Microfluidic Enzyme Kinetics (HT-MEK) device enabled the expression and purification of up to 1792 enzyme variants in a single experiment. All enzyme variants are tagged with a C-terminal eGFP construct to facilitate capture on a functionalized "pedestal." This pedestal consists of neutravidin proteins (NA) non-specifically bound to a PDMS-coated quartz slide, which in turn binds a biotinylated anti-eGFP VHH nanobody, pulling down the eGFP-tagged ADK in each chamber. (B) ADK activity in the direction of ATP formation is monitored on-chip through coupled production of NADPH, and product formation is measured over the course of the assay with time-lapse inverted fluorescent microscopy. Four chambers containing exemplary orthologs are highlighted, as well as a control chamber that did not contain an ADK-encoding plasmid and thus did not show any expressed enzyme or detectable catalysis. (C) Scatter plot of progress curves for bsADK across multiple substrate concentrations (encoded by color). (D) Mean fits of initial rates to the Michaelis-Menten equation for four ADK orthologs. Shaded regions represent the standard deviation of  $k_{cat}$  across biological replicates for each ortholog. (E) On-chip catalytic measurements correlate with previously published values for orthologs, relative to *bsADK*, measured in the same reaction direction under comparable conditions (Methods).

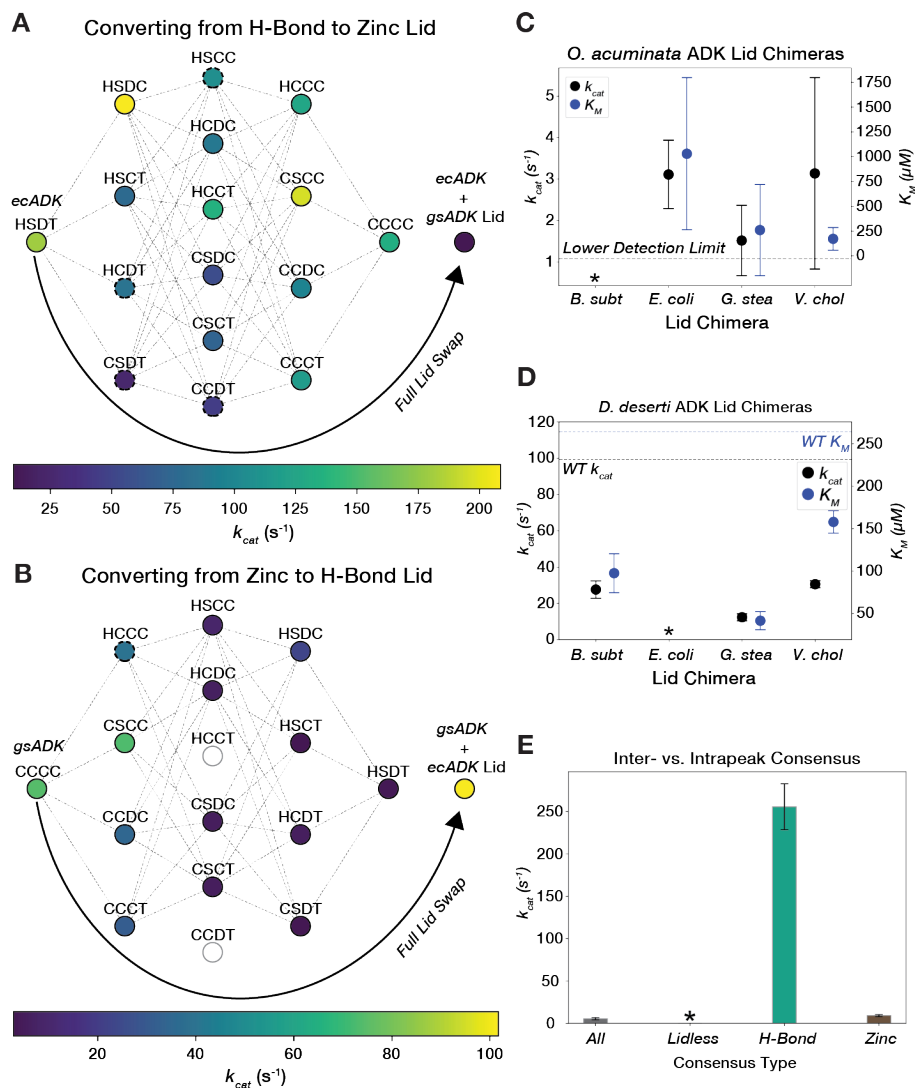


**Figure 3. Adenylate kinase catalytic parameters span three orders of magnitude and correlate weakly with phylogeny and environmental conditions.** (A) Catalytic parameters and  $T_{Growth}$  values for 100/181 ADK orthologs (including those with  $K_M$  outside the range of the assay) are displayed as a heatmap across a taxonomic tree (Methods).  $k_{cat}$  and  $K_M$  are colored on a log scale. Orthologs with  $K_M$  outside the range of the assay are labeled with asterisks (\*). See Fig. S7 for a heatmap of all 181 ADK orthologs. The consensus ADK sequence (94) is plotted at the bottom ( $k_{cat} = 5.4 s^{-1}$ ,  $K_M = 804 \mu M$ ). (B) Measured  $k_{cat}$  values for 175 orthologs span three orders of magnitude. The vertical dashed line represents the  $k_{cat}$  for the ADK consensus sequence (94). (C,D) Phylogenetic signal analysis of (C)  $k_{cat}$  and (D)  $T_{Growth}$ . Moran's I index of autocorrelation is plotted as a solid black line, with the 95% confidence interval outlined by dashed black lines. The colored bar at the bottom encodes the significance of autocorrelation, with red and blue representing positive and negative significant autocorrelation, respectively, and black representing nonsignificant autocorrelation. (E) Linear regression analysis between optimal growth temperature and  $k_{cat}$ . The regression line is plotted as a dashed blue line ( $r^2 = 0.09$ ,  $p < 0.001$ ).



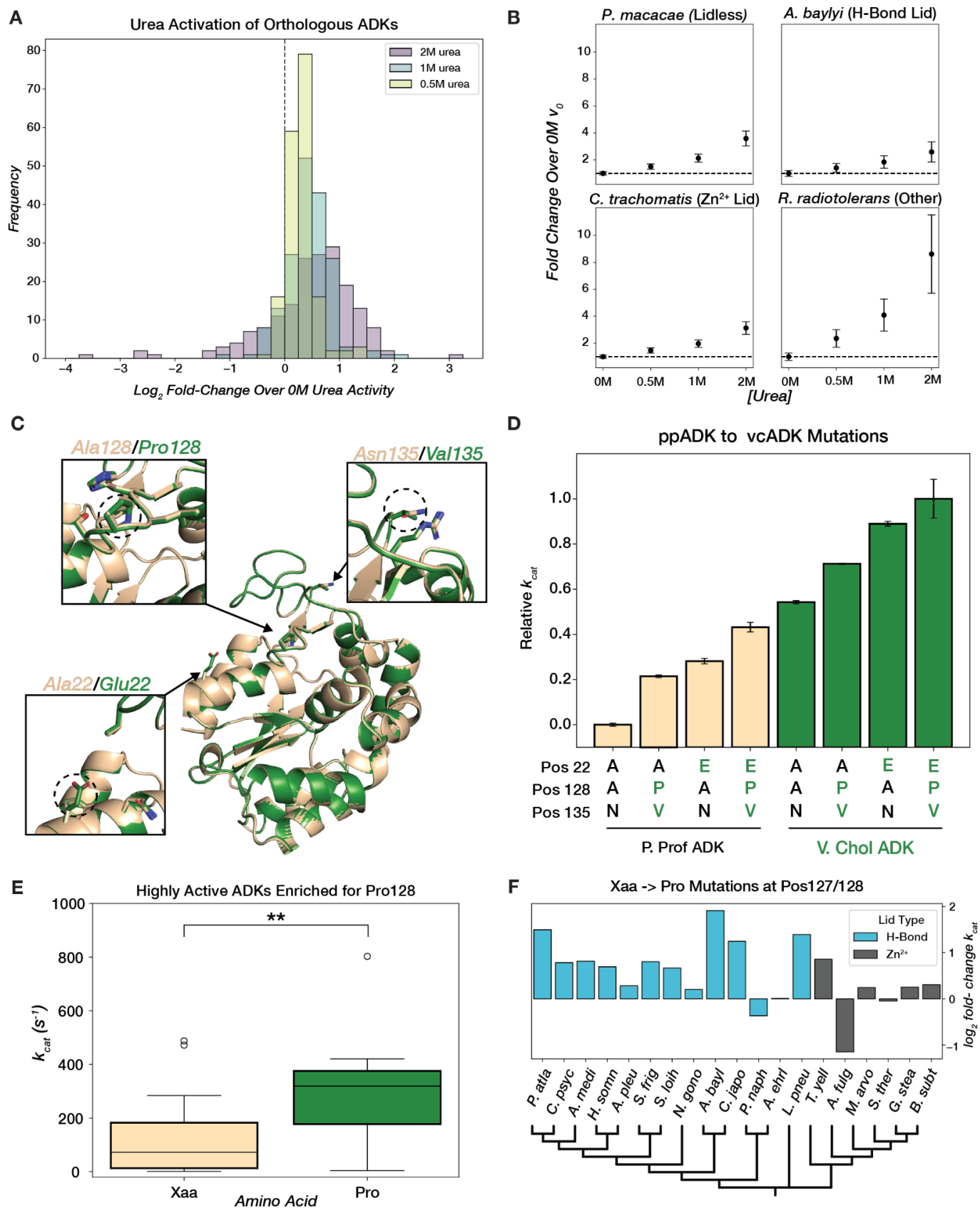
**Figure 4. Adenylate Kinase sequence space features multiple “peaks” linked to different lid types, each associated with distinct growth temperatures and reaching varying “heights” in activity. (A)** A Minimum-Spanning Tree (MST) from the all-by-all graph of ADK sequences with edges weighted by edit distance in an MSA. Edge thickness in the MST encodes increasing edit distance on a log scale. Node color encodes measured  $k_{cat}$  value in log scale, and node shape corresponds to lid type. Rough partitions between lid-type neighborhoods are labeled and outlined with dashed lines. Selected orthologs discussed in depth in this study are labeled. **(B)** The three major types of LID domains found in ADKs: “Lidless” containing a short loop that still provides catalytic Arg residues, and two lidded variations, one containing hydrogen-bonding network consisting of a conserved His-Ser-Asp-Thr tetrad (H-Bond Lid) and the other containing a

cysteine tetrad chelating a  $Zn^{2+}$  ion. **(C)** H-Bond and  $Zn^{2+}$  Binding Lid ADKs are significantly faster than lidless counterparts but have similar activity distributions to one another (ANOVA,  $F=6.92$ ,  $p\text{-val}=0.002$ ; Tukey-HSD,  $Zn^{2+}$  vs. Lidless  $p\text{-adj}=0.002$ , H-Bond vs. Lidless  $p\text{-adj}=0.006$ ,  $Zn^{2+}$  vs. H-Bond  $p\text{-adj}=0.963$ ). **(D)**  $Zn^{2+}$  Binding Lid ADKs have a significantly higher associated growth temperature than the other lid-types (ANOVA,  $F=16.39$ ,  $p\text{-val}=3.11e-7$ ; Tukey-HSD,  $Zn^{2+}$  vs. Lidless  $p\text{-adj}=0.0005$ , H-Bond vs. Lidless  $p\text{-adj}=0.370$ ,  $Zn^{2+}$  vs. H-Bond  $p\text{-adj}<1.0e-16$ ). Cutoffs for psychrophilicity and thermophilicity are shown as dashed lines (103).



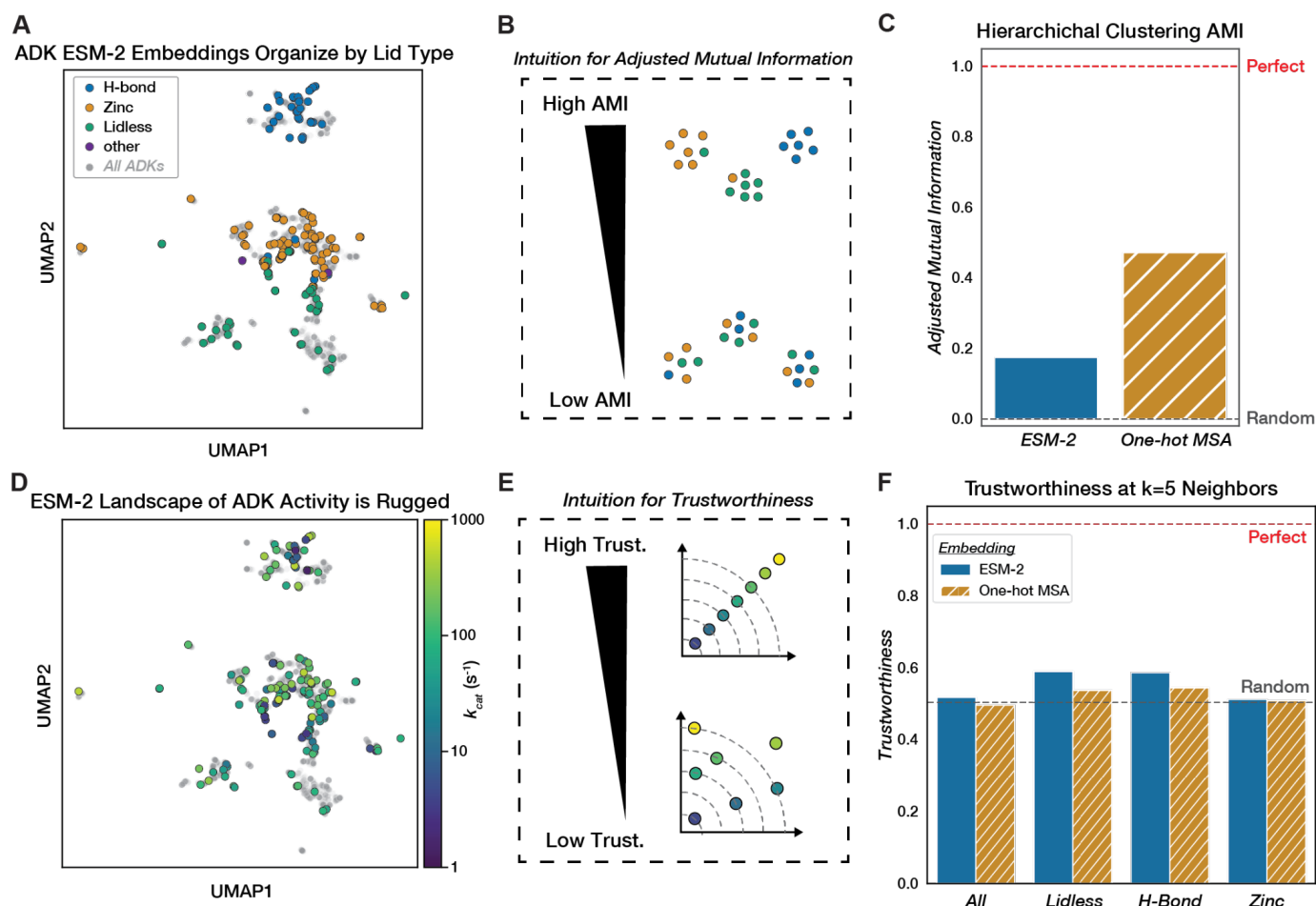
**Figure 5. Traversing the ADK sequence-catalysis landscape through mutational walks and “extra-dimensional bypasses”.** (A) A graph showing mutational pathways of swapping the chelating cysteine tetrad into the H-Bond lid of ecADK. Nodes represent variants along the pathway, with their  $k_{cat}$  value encoded by color. Dashed lines connect variants that are one mutation away from each other. Dashed circles represent variants with  $K_M$  fit outside of the bounds of the assay. Fully swapping the ecADK H-Bond LID for gsADK’s  $Zn^{2+}$ -binding LID is shown as an arrow. (B) A graph showing mutational pathways of swapping the hydrogen-bond motif into the  $Zn^{2+}$  lid of gsADK. Nodes represent variants along the pathway, with their  $k_{cat}$  value encoded by color. Dashed lines connect variants that are one mutation away from each other. Dashed circles represent variants with  $K_M$  fit outside of the bounds of the assay. Fully swapping the gsADK  $Zn^{2+}$ -binding LID for ecADK’s H-Bond LID is shown as an arrow. Empty circles represent variants that expressed poorly or displayed activity below the limit of detections and are not connected by edges in the graph. (C)  $k_{cat}$  and  $K_M$  are plotted for LID chimeras of ocADK. Error bars represent standard deviation across biological replicates. The lower detection limit for  $k_{cat}$  is plotted as a black dashed line. The bsADK LID chimera displayed activity below the lower limit of detection and is marked with an asterisk. (D)  $k_{cat}$  and  $K_M$  are plotted for LID chimeras of ddADK. Error bars represent standard deviation across biological replicates. The lower detection limit for  $k_{cat}$  is plotted as a black dashed line. The ecADK LID chimera exhibited a  $K_M$  below the lower bound of the assay and is marked with an asterisk. (E) Barplot of  $k_{cat}$  for consensus sequences, with error bars representing standard deviation across biological replicates. The lidless consensus sequence did not exhibit activity above background.





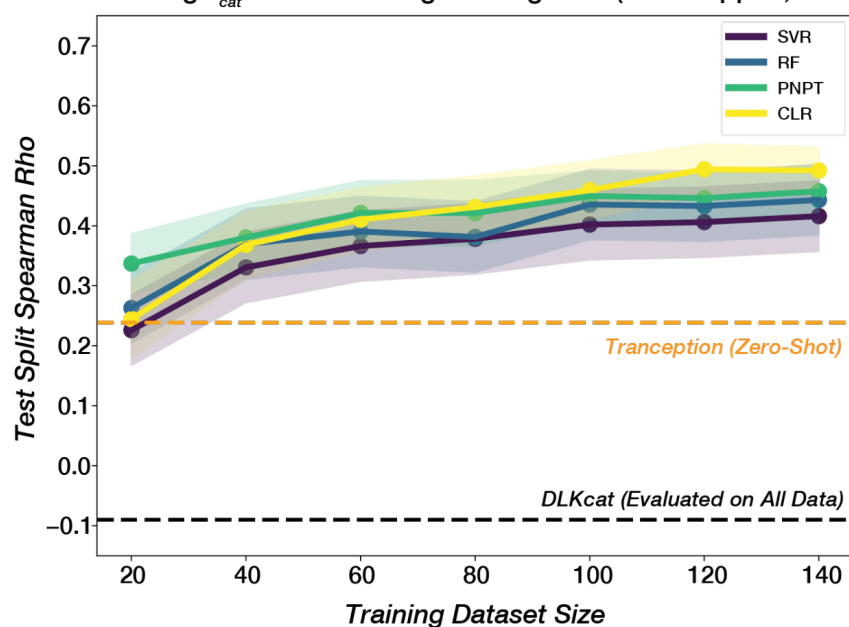
**Figure 6. ADK conformational tuning with osmolytes and mutations across evolution.** (A) Distribution of  $\log_2$  fold-change in initial reaction rate at saturating substrate concentration relative to 0M urea for 0.5M, 1M, and 2M urea. (B) Mean fold-change in initial rate over 0M urea at 4mM [substrate]. Error bars represent standard deviation across replicates. The black dashed line represents no change in initial rate. (C) Superimposed AF2 models of ppADK (tan) and vcADK (green) with positions 22, 128, and 135 highlighted. (D) Barplot of mean relative catalytic effects (technical replicates, n=2) of

mutations at key positions that differ between ppADK and vcADK. Error bars represent standard deviation across replicates. Variants in ppADK background are plotted in tan and vcADK background in green. Amino acid identity at positions 22, 128, and 135 are displayed below each bar.  $k_{cat}$  values for ppADK and vcADK mutants were collected off-chip (see Methods). **(E)** Boxplot of  $k_{cat}$  for H-Bond LID ADKs that have a proline (green) or different amino acid (tan) at position 128.  $t(37)=-3.162$ ,  $p=0.003$ . **(F)** Barplot of  $\log_2$  fold-change in  $k_{cat}$  for Xaa→Pro mutations in selected ADKs with either an H-Bond or  $Zn^{2+}$  LID, organized by a phylogenetic tree (plotted with arbitrary branch lengths).

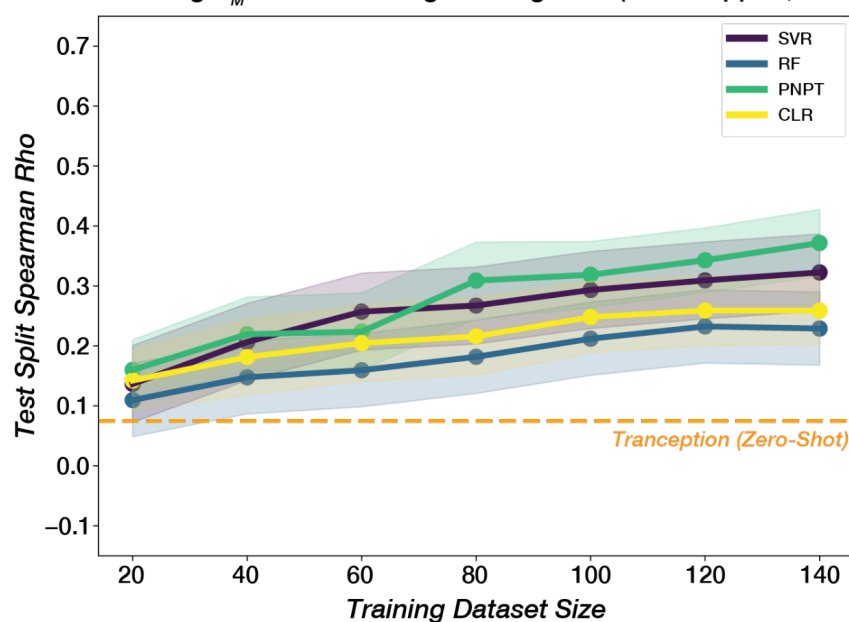


**Figure 7. Evaluating the structural and catalytic organization of ADK sequence space learned by a protein language model.** (A) UMAP of ADK ESM-2 embeddings, with all representative ADK sequences (~5,000, Methods) plotted (gray). Sequences with measured catalytic parameters are encoded with color by lid type. (B) AMI reflects the agreement of a clustering method with respect to another label. In this case, low AMI would suggest poor clustering by Lid-type and vice versa. (C) Barplot comparing the AMI of hierarchical clustering on ESM-2 embeddings or a one-hot encoded MSA using euclidean and hamming distance, respectively, with respect to lid type. Perfect AMI (1.0) is plotted as a red dashed line, and random AMI (0.0) is plotted as a gray dashed line. (D) UMAP of ADK ESM-2 embeddings with  $k_{cat}$  values encoded by color. Grey points represent all representative ADK sequences (~5,000, Methods). (E) Trustworthiness quantifies the level of organization retained between two embeddings: in this case, it quantifies how similar neighboring sequences in ESM-2 space are in “ $k_{cat}$  space”. (F) Grouped barplot of trustworthiness of ESM-2 ADK embeddings (by euclidean distance) and one-hot encoded MSA (by hamming distance) with respect to  $k_{cat}$  computed at k=5 neighbors. Trustworthiness is computed for all sequences, as well as by lid type. Perfect trustworthiness is plotted as a red dashed line ( $k_{cat}$  vs.  $k_{cat}$ ) and random as a gray dashed line ( $k_{cat}$  vs. an average of 30 shuffles of corresponding embeddings).

**A** Predicting  $k_{cat}$  with Increasing Training Data (Boostrapped, N=30)



**B** Predicting  $K_M$  with Increasing Training Data (Boostrapped, N=30)



**Figure 8. Improving  $k_{cat}$  and  $K_M$  prediction for ADK sequences with semi-supervised learning.** The mean test spearman rho for different models across 30 samplings is plotted against training dataset size for (A)  $k_{cat}$  and (B)  $K_M$ . Models include Random Forest (RF), Support Vector Regression (SVR), ProteinNPT (PNPT) (107), and Convolutional Linear Regression (CLR) (107). For embeddings, SVR used a One-hot encoded MSA, RF used ESM-2 embeddings, and CLR and PNPT used Tranception(112) embeddings. Embeddings and other model hyperparameters were selected based on aggregate (mean) performance for both  $k_{cat}$  and  $K_M$  prediction. Shaded regions represent 95% confidence intervals across 30 training/test set samplings at each dataset size (Methods). A zero-shot evaluation of the Tranception PLM (112) is plotted as a dashed orange line. DLKcat (109) performance evaluated on all 175 sequences is plotted in (A) as a dashed black line.

## References and Notes

1. R. A. Fisher, *The Genetical Theory of Natural Selection* (Oxford: Oxford University Press., 1930).
2. S. Wright, “J., 1932. The roles of mutation, inbreeding, cross-breeding and selection in evolution” in *Proceedings of the Sixth International Congress of Genetics, New York, Ithaca, I* (1932), pp. 356–366.
3. I. Fragata, A. Blanckaert, M. A. Dias Louro, D. A. Liberles, C. Bank, Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82 (2019).
4. J. A. G. M. de Visser, J. Krug, Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490 (2014).
5. M. Conrad, The geometry of evolution. *Biosystems.* **24**, 61–81 (1990).
6. X. Yi, A. M. Dean, Adaptive Landscapes in the Age of Synthetic Biology. *Mol. Biol. Evol.* **36**, 890–907 (2019).
7. D. M. McCandlish, Visualizing fitness landscapes. *Evolution* **65**, 1544–1558 (2011).
8. M. Lunzer, G. B. Golding, A. M. Dean, Pervasive cryptic epistasis in molecular evolution. *PLoS Genet.* **6**, e1001162 (2010).
9. G. Yang, D. W. Anderson, F. Baier, E. Dohmen, N. Hong, P. D. Carr, S. C. L. Kamerlin, C. J. Jackson, E. Bornberg-Bauer, N. Tokuriki, Higher-order epistasis shapes the fitness landscape of a xenobiotic-degrading enzyme. *Nat. Chem. Biol.* **15**, 1120–1128 (2019).
10. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114 (2006).
11. C. G. Acevedo-Rocha, A. Li, L. D’Amore, S. Hoebenreich, J. Sanchis, P. Lubrano, M. P. Ferla, M. Garcia-Borràs, S. Osuna, M. T. Reetz, Pervasive cooperative mutational effects on multiple catalytic enzyme traits emerge via long-range conformational dynamics. *Nat. Commun.* **12**, 1621 (2021).
12. M. S. Breen, C. Kemena, P. K. Vlasov, C. Notredame, F. A. Kondrashov, Epistasis as the primary factor in molecular evolution. *Nature* **490**, 535–538 (2012).
13. A. Papkou, L. Garcia-Pastor, J. A. Escudero, A. Wagner, A rugged yet easily navigable fitness landscape. *Science* **382**, eadh3860 (2023).
14. B. G. Hall, The EBG system of *E. coli*: origin and evolution of a novel beta-galactosidase for the metabolism of lactose. *Genetica* **118**, 143–156 (2003).
15. S. P. Miller, M. Lunzer, A. M. Dean, Direct demonstration of an adaptive constraint. *Science* **314**, 458–461 (2006).
16. M. Lunzer, S. P. Miller, R. Felsheim, A. M. Dean, The biochemical architecture of an ancient adaptive landscape. *Science* **310**, 499–501 (2005).
17. S. Noor, M. C. Taylor, R. J. Russell, L. S. Jermini, C. J. Jackson, J. G. Oakeshott, C. Scott, Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS One* **7**, e39822 (2012).
18. P. E. O’Maille, A. Malone, N. Dellas, B. Andes Hess Jr, L. Smentek, I. Sheehan, B. T. Greenhagen, J. Chappell, G. Manning, J. P. Noel, Quantitative exploration of the catalytic landscape separating divergent plant sesquiterpene synthases. *Nat. Chem. Biol.* **4**, 617–623 (2008).
19. E. R. Lozovsky, T. Chookajorn, K. M. Brown, M. Imwong, P. J. Shaw, S. Kamchonwongpaisan, D. E. Neafsey, D. M. Weinreich, D. L. Hartl, Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proceedings of the National Academy of Sciences* **106**, 12025–12030 (2009).
20. M.-R. Meini, P. E. Tomatis, D. M. Weinreich, A. J. Vila, Quantitative Description of a Protein Fitness Landscape

- Based on Molecular Features. *Mol. Biol. Evol.* **32**, 1774–1787 (2015).
21. N. S. Persky, D. Hernandez, M. Do Carmo, L. Brenan, O. Cohen, S. Kitajima, U. Nayar, A. Walker, S. Pantel, Y. Lee, J. Cordova, M. Sathappa, C. Zhu, T. K. Hayes, P. Ram, P. Pancholi, T. S. Mikkelsen, D. A. Barbie, X. Yang, R. Haq, F. Piccioni, D. E. Root, C. M. Johannessen, Defining the landscape of ATP-competitive inhibitor resistance residues in protein kinases. *Nat. Struct. Mol. Biol.* **27**, 92–104 (2020).
  22. E. M. Jones, N. B. Lubock, A. J. Venkatakrishnan, J. Wang, A. M. Tseng, J. M. Paggi, N. R. Latorraca, D. Cancilla, M. Satyadi, J. E. Davis, M. Madan Babu, R. O. Dror, S. Kosuri, Structural and functional characterization of G protein-coupled receptors with deep mutational scanning. *Elife* **9** (2020).
  23. W. Coyote-Maestas, D. Nedrud, A. Suma, Y. He, K. A. Matreyek, D. M. Fowler, V. Carnevale, C. L. Myers, D. Schmidt, Probing ion channel functional architecture and domain recombination compatibility by massively parallel domain insertion profiling. *Nat. Commun.* **12**, 1–16 (2021).
  24. G. Yang, C. M. Miton, N. Tokuriki, A mechanistic view of enzyme evolution. *Protein Sci.* **29**, 1724–1747 (2020).
  25. M. A. Stiffler, D. R. Hekstra, R. Ranganathan, Evolvability as a function of purifying selection in TEM-1  $\beta$ -lactamase. *Cell* **160**, 882–892 (2015).
  26. 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).
  27. A. Melnikov, P. Rogov, L. Wang, A. Gnirke, T. S. Mikkelsen, Comprehensive mutational scanning of a kinase in vivo reveals substrate-dependent fitness landscapes. *Nucleic Acids Res.* **42**, e112 (2014).
  28. J.-Y. van der Meer, H. Poddar, B.-J. Baas, Y. Miao, M. Rahimi, A. Kunzendorf, R. van Merkerk, P. G. Tepper, E. M. Geertsema, A.-M. W. H. Thunnissen, W. J. Quax, G. J. Poelarends, Using mutability landscapes of a promiscuous tautomerase to guide the engineering of enantioselective Michaelases. *Nat. Commun.* **7**, 10911 (2016).
  29. J. Z. Chen, D. M. Fowler, N. Tokuriki, Comprehensive exploration of the translocation, stability and substrate recognition requirements in VIM-2 lactamase. *Elife* **9** (2020).
  30. J. N. Copp, E. Akiva, P. C. Babbitt, N. Tokuriki, Revealing Unexplored Sequence-Function Space Using Sequence Similarity Networks. *Biochemistry* **57**, 4651–4662 (2018).
  31. P. Notin, N. Rollins, Y. Gal, C. Sander, D. Marks, Machine learning for functional protein design. *Nat. Biotechnol.* **42**, 216–228 (2024).
  32. Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y. Shmueli, A. Dos Santos Costa, M. Fazel-Zarandi, T. Sercu, S. Candido, A. Rives, Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).
  33. R. Wu, F. Ding, R. Wang, R. Shen, X. Zhang, S. Luo, C. Su, Z. Wu, Q. Xie, B. Berger, J. Ma, J. Peng, High-resolution *de novo* structure prediction from primary sequence, *bioRxiv* (2022)p. 2022.07.21.500999.
  34. A. Rives, J. Meier, T. Sercu, S. Goyal, Z. Lin, J. Liu, D. Guo, M. Ott, C. Lawrence Zitnick, J. Ma, R. Fergus, Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences, *bioRxiv* (2020)p. 622803.
  35. J. A. Ruffolo, A. Madani, Designing proteins with language models. *Nat. Biotechnol.* **42**, 200–202 (2024).
  36. A. Madani, B. Krause, E. R. Greene, S. Subramanian, B. P. Mohr, J. M. Holton, J. L. Olmos Jr, C. Xiong, Z. Z. Sun, R. Socher, J. S. Fraser, N. Naik, Large language models generate functional protein sequences across diverse families. *Nat. Biotechnol.*, doi: 10.1038/s41587-022-01618-2 (2023).
  37. P. Notin, A. W. Kollasch, D. Ritter, L. van Niekerk, S. Paul, H. Spinner, N. Rollins, A. Shaw, R. Weitzman, J. Frazer, M. Dias, D. Franceschi, R. Orenbuch, Y. Gal, D. S. Marks, ProteinGym: Large-Scale Benchmarks for Protein Design and Fitness Prediction. *bioRxiv*, doi: 10.1101/2023.12.07.570727 (2023).

38. B. Hie, E. D. Zhong, B. Berger, B. Bryson, Learning the language of viral evolution and escape. *Science* **371**, 284–288 (2021).
39. J. Meier, R. Rao, R. Verkuil, J. Liu, T. Sercu, A. Rives, Language models enable zero-shot prediction of the effects of mutations on protein function, *bioRxiv* (2021)p. 2021.07.09.450648.
40. B. L. Hie, K. K. Yang, P. S. Kim, Evolutionary velocity with protein language models predicts evolutionary dynamics of diverse proteins. *Cell Syst* **13**, 274–285.e6 (2022).
41. Y. Luo, G. Jiang, T. Yu, Y. Liu, L. Vo, H. Ding, Y. Su, W. W. Qian, H. Zhao, J. Peng, ECNet is an evolutionary context-integrated deep learning framework for protein engineering. *Nat. Commun.* **12**, 1–14 (2021).
42. A. J. Riesselman, J. B. Ingraham, D. S. Marks, Deep generative models of genetic variation capture the effects of mutations. *Nat. Methods* **15**, 816–822 (2018).
43. T. A. Hopf, J. B. Ingraham, F. J. Poelwijk, C. P. I. Schärfe, M. Springer, C. Sander, D. S. Marks, Mutation effects predicted from sequence co-variation. *Nat. Biotechnol.* **35**, 128–135 (2017).
44. J. I. Boucher, D. N. A. Bolon, D. S. Tawfik, Quantifying and understanding the fitness effects of protein mutations: Laboratory versus nature. *Protein Sci.* **25**, 1219–1226 (2016).
45. C. J. Markin, D. A. Mokhtari, F. Sunden, M. J. Appel, E. Akiva, S. A. Longwell, C. Sabatti, D. Herschlag, P. M. Fordyce, Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics. *Science* **373** (2021).
46. S. J. Kerns, R. V. Agafonov, Y.-J. J. Cho, F. Pontiggia, R. Otten, D. V. Pachov, S. Kutter, L. A. Phung, P. N. Murphy, V. Thai, T. Alber, M. F. Hagan, D. Kern, The energy landscape of adenylate kinase during catalysis. *Nat. Struct. Mol. Biol.* **22**, 124–131 (2015).
47. M. Wolf-Watz, V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E. Z. Eisenmesser, D. Kern, Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struct. Mol. Biol.* **11**, 945–949 (2004).
48. J. Ådén, A. Verma, A. Schug, M. Wolf-Watz, Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. *J. Am. Chem. Soc.* **134**, 16562–16570 (2012).
49. J. A. Hanson, K. Duderstadt, L. P. Watkins, S. Bhattacharyya, J. Brokaw, J.-W. Chu, H. Yang, Illuminating the mechanistic roles of enzyme conformational dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18055–18060 (2007).
50. U. Olsson, M. Wolf-Watz, Overlap between folding and functional energy landscapes for adenylate kinase conformational change. *Nat. Commun.* **1**, 111 (2010).
51. J. B. Stiller, S. J. Kerns, M. Hoemberger, Y.-J. Cho, R. Otten, M. F. Hagan, D. Kern, Probing the transition state in enzyme catalysis by high-pressure NMR dynamics. *Nature Catalysis*, 1–9 (2019).
52. M. Kovermann, J. Ådén, C. Grundström, A. E. Sauer-Eriksson, U. H. Sauer, M. Wolf-Watz, Structural basis for catalytically restrictive dynamics of a high-energy enzyme state. *Nat. Commun.* **6**, 7644 (2015).
53. B. Dulko-Smith, P. Ojeda-May, J. Ådén, M. Wolf-Watz, K. Nam, Mechanistic Basis for a Connection between the Catalytic Step and Slow Opening Dynamics of Adenylate Kinase. *J. Chem. Inf. Model.* **63**, 1556–1569 (2023).
54. V. Nguyen, C. Wilson, M. Hoemberger, J. B. Stiller, R. V. Agafonov, S. Kutter, J. English, D. L. Theobald, D. Kern, Evolutionary drivers of thermostability in enzyme catalysis. *Science* **355**, 289–294 (2017).
55. E. Bae, G. N. Phillips Jr, Roles of static and dynamic domains in stability and catalysis of adenylate kinase. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2132–2137 (2006).
56. S. Moon, D. Jung, G. N. Phillips, E. Bae, An integrated approach for thermal stabilization of a mesophilic adenylate kinase. *Proteins: Struct. Funct. Bioinf.* **82**, 1947–1959 (2014).

57. R. Couñago, C. J. Wilson, M. I. Peña, P. Wittung-Stafshede, Y. Shamoo, An adaptive mutation in adenylate kinase that increases organismal fitness is linked to stability–activity trade-offs. *Protein Eng. Des. Sel.* **21**, 19–27 (2008).
58. S. Moon, R. M. Bannen, T. J. Rutkoski, G. N. Phillips, E. Bae, Effectiveness and limitations of local structural entropy optimization in the thermal stabilization of mesophilic and thermophilic adenylate kinases. *Proteins: Struct. Funct. Bioinf.* **82**, 2631–2642 (2014).
59. P. Ojeda-May, A. U. Mushtaq, P. Rogne, A. Verma, V. Ovchinnikov, C. Grundström, B. Dulko-Smith, U. H. Sauer, M. Wolf-Watz, K. Nam, Dynamic Connection between Enzymatic Catalysis and Collective Protein Motions. *Biochemistry* **60**, 2246–2258 (2021).
60. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
61. C. W. Müller, G. J. Schlauderer, J. Reinstein, G. E. Schulz, Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure* **4**, 147–156 (1996).
62. H. G. Saavedra, J. O. Wrabl, J. A. Anderson, J. Li, V. J. Hilser, Dynamic allostery can drive cold adaptation in enzymes. *Nature* **558**, 324–328 (2018).
63. A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfik, R. Milo, The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry* **50**, 4402–4410 (2011).
64. F. Keck, F. Rimet, A. Bouchez, A. Franc, phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol. Evol.* **6**, 2774–2780 (2016).
65. K. S. Siddiqui, R. Cavicchioli, Cold-adapted enzymes. *Annu. Rev. Biochem.* **75**, 403–433 (2006).
66. W. Wang, J. Swartz, Rational engineering of enzymes for enhanced cold activity. *ACS Catal.* **14**, 12518–12528 (2024).
67. D. Georlette, V. Blaise, T. Collins, S. D’Amico, E. Gratia, A. Hoyoux, J.-C. Marx, G. Sonan, G. Feller, C. Gerday, Some like it cold: biocatalysis at low temperatures. *FEMS Microbiol. Rev.* **28**, 25–42 (2004).
68. G. Feller, C. Gerday, Psychrophilic enzymes: hot topics in cold adaptation. *Nat. Rev. Microbiol.* **1**, 200–208 (2003).
69. S. Bjelic, B. O. Brandsdal, J. Aqvist, Cold adaptation of enzyme reaction rates. *Biochemistry* **47**, 10049–10057 (2008).
70. M. Elias, G. Wieczorek, S. Rosenne, D. S. Tawfik, The universality of enzymatic rate-temperature dependency. *Trends Biochem. Sci.* **39**, 1–7 (2014).
71. P. A. Fields, Y. Dong, X. Meng, G. N. Somero, Adaptations of protein structure and function to temperature: there is more than one way to “skin a cat.” *J. Exp. Biol.* **218**, 1801–1811 (2015).
72. J. Åqvist, G. V. Isaksen, B. O. Brandsdal, Computation of enzyme cold adaptation. *Nat. Rev. Chem.* **1**, 0051 (2017).
73. M.-L. Liao, G. N. Somero, Y.-W. Dong, Comparing mutagenesis and simulations as tools for identifying functionally important sequence changes for protein thermal adaptation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 679–688 (2019).
74. P. Závodszy, J. Kardos, Svingor, G. A. Petsko, Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7406–7411 (1998).
75. R. Jaenicke, Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* **202**, 715–728 (1991).



76. S. D'Amico, P. Claverie, T. Collins, D. Georgette, E. Gratia, A. Hoyoux, M.-A. Meuwis, G. Feller, C. Gerday, Molecular basis of cold adaptation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**, 917–925 (2002).
77. P. A. Fields, Review: Protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 417–431 (2001).
78. R. Jaenicke, Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *Proc. Natl. Acad. Sci. U. S. A.* **97**, 2962–2964 (2000).
79. A. Chang, L. Jeske, S. Ulbrich, J. Hofmann, J. Koblitz, I. Schomburg, M. Neumann-Schaal, D. Jahn, D. Schomburg, BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Res.* **49**, D498–D508 (2021).
80. C. Stark, T. Bautista-Leung, J. Siegfried, D. Herschlag, Systematic investigation of the link between enzyme catalysis and cold adaptation. *Elife* **11** (2022).
81. P. A. Fields, G. N. Somero, Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 11476–11481 (1998).
82. W. Bains, Y. Xiao, C. Yu, Prediction of the maximum temperature for life based on the stability of metabolites to decomposition in water. *Life (Basel)* **5**, 1054–1100 (2015).
83. M. D. Daily, G. N. Phillips Jr, Q. Cui, Interconversion of functional motions between mesophilic and thermophilic adenylate kinases. *PLoS Comput. Biol.* **7**, e1002103 (2011).
84. S. Moon, J. Kim, E. Bae, Structural analyses of adenylate kinases from Antarctic and tropical fishes for understanding cold adaptation of enzymes. *Sci. Rep.* **7**, 16027 (2017).
85. H. Bönisch, J. Backmann, T. Kath, D. Naumann, G. Schäfer, Adenylate kinase from *Sulfolobus acidocaldarius*: expression in *Escherichia coli* and characterization by Fourier transform infrared spectroscopy. *Arch. Biochem. Biophys.* **333**, 75–84 (1996).
86. L. Rundqvist, J. Adén, T. Sparrman, M. Wallgren, U. Olsson, M. Wolf-Watz, Noncooperative folding of subdomains in adenylate kinase. *Biochemistry* **48**, 1911–1927 (2009).
87. M. M. Pinney, D. A. Mokhtari, E. Akiva, F. Yabukarski, D. M. Sanchez, R. Liang, T. Doukov, T. J. Martinez, P. C. Babbitt, D. Herschlag, Parallel molecular mechanisms for enzyme temperature adaptation. *Science* **371** (2021).
88. G. E. Schulz, E. Schiltz, A. G. Tomasselli, R. Frank, M. Brune, A. Wittinghofer, R. H. Schirmer, Structural relationships in the adenylate kinase family. *Eur. J. Biochem.* **161**, 127–132 (1986).
89. H. Munier-Lehmann, S. Burlacu-Miron, C. T. Craescu, H. H. Mantsch, C. P. Schultz, A new subfamily of short bacterial adenylate kinases with the *Mycobacterium tuberculosis* enzyme as a model: A predictive and experimental study. *Proteins: Struct. Funct. Bioinf.* **36**, 238–248 (1999).
90. T. P. Schrank, J. O. Wrabl, V. J. Hilser, Conformational heterogeneity within the LID domain mediates substrate binding to *Escherichia coli* adenylate kinase: function follows fluctuations. *Top. Curr. Chem.* **337**, 95–121 (2013).
91. V. Perrier, S. Burlacu-Miron, S. Bourgeois, W. K. Surewicz, A.-M. Gilles, Genetically Engineered Zinc-chelating Adenylate Kinase from *Escherichia coli* with Enhanced Thermal Stability \*. *J. Biol. Chem.* **273**, 19097–19101 (1998).
92. K. W. Bruland, Oceanographic distributions of cadmium, zinc, nickel, and copper in the North Pacific. *Earth Planet. Sci. Lett.* **47**, 176–198 (1980).
93. B. T. Porebski, A. M. Buckle, Consensus protein design. *Protein Eng. Des. Sel.* **29**, 245–251 (2016).
94. M. Sternke, K. W. Tripp, D. Barrick, Consensus sequence design as a general strategy to create hyperstable,

- biologically active proteins. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11275–11284 (2019).
95. D. Aerts, T. Verhaeghe, H.-J. Joosten, G. Vriend, W. Soetaert, T. Desmet, Consensus engineering of sucrose phosphorylase: the outcome reflects the sequence input: Consensus Engineering of Enzymes. *Biotechnol. Bioeng.* **110**, 2563–2572 (2013).
  96. B. J. Sullivan, V. Durani, T. J. Magliery, Triosephosphate isomerase by consensus design: dramatic differences in physical properties and activity of related variants. *J. Mol. Biol.* **413**, 195–208 (2011).
  97. M. Lehmann, C. Loch, A. Middendorf, D. Studer, S. F. Lassen, L. Pasamontes, A. P. G. M. van Loon, M. Wyss, The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng.* **15**, 403–411 (2002).
  98. K. M. Polizzi, J. F. Chaparro-Riggers, E. Vazquez-Figueroa, A. S. Bommarius, Structure-guided consensus approach to create a more thermostable penicillin G acylase. *Biotechnol. J.* **1**, 531–536 (2006).
  99. P. Rogne, M. Wolf-Watz, Urea-dependent adenylate kinase activation following redistribution of structural states. *Biophys. J.* **111**, 1385–1395 (2016).
  100. T. P. Schrank, D. W. Bolen, V. J. Hilser, Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16984–16989 (2009).
  101. C. Kujo, T. Ohshima, Enzymological characteristics of the hyperthermostable NAD-dependent glutamate dehydrogenase from the archaeon *Pyrobaculum islandicum* and effects of denaturants and organic solvents. *Appl. Environ. Microbiol.* **64**, 2152–2157 (1998).
  102. V. Wilquet, J. A. Gaspar, M. van de Lande, M. Van de Castele, C. Legrain, E. M. Meiering, N. Glansdorff, Purification and characterization of recombinant *Thermotoga maritima* dihydrofolate reductase. *Eur. J. Biochem.* **255**, 628–637 (1998).
  103. C. Vieille, G. J. Zeikus, Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43 (2001).
  104. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, *arXiv [stat.ML]* (2018). <http://arxiv.org/abs/1802.03426>.
  105. T. Chari, L. Pachter, The specious art of single-cell genomics. *PLoS Comput. Biol.* **19**, e1011288 (2023).
  106. C. Hsu, H. Nisonoff, C. Fannjiang, J. Listgarten, Learning protein fitness models from evolutionary and assay-labeled data. *Nat. Biotechnol.*, 1–9 (2022).
  107. P. Notin, R. Weitzman, D. S. Marks, Y. Gal, ProteinNPT: Improving protein property prediction and design with non-parametric transformers. *bioRxiv*org, 2023.12.06.570473 (2023).
  108. S. Biswas, G. Khimulya, E. C. Alley, K. M. Esvelt, G. M. Church, Low-N protein engineering with data-efficient deep learning. *Nat. Methods* **18**, 389–396 (2021).
  109. F. Li, L. Yuan, H. Lu, G. Li, Y. Chen, M. K. M. Engqvist, E. J. Kerkhoven, J. Nielsen, Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. *Nature Catalysis*, 1–11 (2022).
  110. A. Kroll, Y. Rousset, X.-P. Hu, N. A. Liebrand, M. J. Lercher, Turnover number predictions for kinetically uncharacterized enzymes using machine and deep learning. *Nat. Commun.* **14**, 4139 (2023).
  111. E. Rembeza, M. K. M. Engqvist, Experimental and computational investigation of enzyme functional annotations uncovers misannotation in the EC 1.1.3.15 enzyme class. *PLoS Comput. Biol.* **17**, e1009446 (2021).
  112. P. Notin, M. Dias, J. Frazer, J. Marchena-Hurtado, A. Gomez, D. S. Marks, Y. Gal, Tranception: protein fitness prediction with autoregressive transformers and inference-time retrieval, *arXiv [cs.LG]* (2022).

<http://arxiv.org/abs/2205.13760>.

113. R. M. Rao, J. Liu, R. Verkuil, J. Meier, J. Canny, P. Abbeel, T. Sercu, A. Rives, “MSA Transformer” in *Proceedings of the 38th International Conference on Machine Learning*, M. Meila, T. Zhang, Eds. (PMLR, 18--24 Jul 2021; <https://proceedings.mlr.press/v139/rao21a.html>) vol. 139 of *Proceedings of Machine Learning Research*, pp. 8844–8856.
114. M. K. M. Engqvist, Correlating enzyme annotations with a large set of microbial growth temperatures reveals metabolic adaptations to growth at diverse temperatures. *BMC Microbiol.* **18**, 177 (2018).
115. C. W. Müller, G. E. Schulz, Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap5A refined at 1.9 Å resolution: A model for a catalytic transition state. *J. Mol. Biol.* **224**, 159–177 (1992).
116. D. D. Le, T. C. Shimko, A. K. Aditham, A. M. Keys, S. A. Longwell, Y. Orenstein, P. M. Fordyce, Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3702–E3711 (2018).
117. P. M. Fordyce, D. Pincus, P. Kimmig, C. S. Nelson, H. El-Samad, P. Walter, J. L. DeRisi, 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**, E3084–93 (2012).
118. P. M. Fordyce, R. Hastings, A. Aditham, N. DelRosso, P. Suzuki, High-throughput thermodynamic and kinetic measurements of transcription factor/DNA mutations reveal how conformational heterogeneity can shape motif selectivity, *bioRxiv* (2023)p. 2023.11.13.566946.
119. I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.*, doi: 10.1093/nar/gkae268 (2024).
120. M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
121. K. Tamura, G. Stecher, S. Kumar, MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
122. A. Jarzab, N. Kurzawa, T. Hopf, M. Moerch, J. Zecha, N. Leijten, Y. Bian, E. Musiol, M. Maschberger, G. Stoehr, I. Becher, C. Daly, P. Samaras, J. Mergner, B. Spanier, A. Angelov, T. Werner, M. Bantscheff, M. Wilhelm, M. Klingenspor, S. Lemeer, W. Liebl, H. Hahne, M. M. Savitski, B. Kuster, Meltome atlas-thermal proteome stability across the tree of life. *Nat. Methods* **17**, 495–503 (2020).
123. C. Gerday, M. Aittaleb, J. L. Arpigny, E. Baise, J. P. Chessa, G. Garsoux, I. Petrescu, G. Feller, Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta* **1342**, 119–131 (1997).
124. S. D’Amico, T. Collins, J.-C. Marx, G. Feller, C. Gerday, Psychrophilic microorganisms: challenges for life. *EMBO Rep.* **7**, 385–389 (2006).
125. S. Davail, G. Feller, E. Narinx, C. Gerday, Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile *Bacillus TA41*. *J. Biol. Chem.* **269**, 17448–17453 (1994).
126. M. Davlieva, Y. Shamoo, Structure and biochemical characterization of an adenylate kinase originating from the psychrophilic organism *Marinibacillus marinus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 751–756 (2009).
127. R. Stroek, Y. Ge, P. D. Talbot, M. K. Glok, K. E. Bernás, C. M. Thomson, E. R. Gould, M. S. Alphey, H. Liu, G. J. Florence, J. H. Naismith, R. G. da Silva, Kinetics and structure of a cold-adapted hetero-octameric ATP phosphoribosyltransferase. *Biochemistry* **56**, 793–803 (2017).