Temperature Sensitive Contacts in Disordered Loops Tune Enzyme I Activity

Daniel Burns, Aayushi Singh, Vincenzo Venditti*, Davit A Potoyan*

Significance

Temperature affects the catalytic rates of all enzymes. The impact of temperature on the catalytic activity of an enzyme, however, is convoluted from contributions of protein sequence, structure and dynamics. As such, understanding and designing the molecular features of enzymes which tune catalytic rates at different temperatures remains a fundamental challenge in biophysics. In this work we have employed molecular simulations and mutagenesis experiments to reveal the temperature tuning mechanism of mesophilic and thermophilic homologues of the C domain of bacterial Enzyme I. We find that enzymes can be tuned to their physiological temperatures through a network of temperature sensitive residue contacts localized in the disordered loops. Furthermore, we find that among temperature sensitive contacts some exhibit linear and others parabolic dependence on temperature. These clues offer a promising physics-based approach for tuning enzyme activity.

Abstract

Homologous mesophilic and thermophilic enzymes with identical folds often exhibit different thermal and kinetic behaviors. Understanding how enzyme sequence encodes catalytic activity at functionally optimal temperatures is a fundamental problem in biophysics. Recently it was shown that the residues that tune catalytic activities of thermophilic/mesophilic variants of the C-terminal domain of bacterial Enzyme I (EIC) are largely localized within disordered loops, offering a model system with which to investigate this phenomenon. In this work, we employ molecular dynamics simulations and mutagenesis experiments to reveal a mechanism of sequence-dependent catalytic temperature tuning of EIC homologs.

We find that a network of contacts in the catalytic loops is particularly sensitive to changes in temperature, with some contacts exhibiting distinct linear or parabolic temperature-dependent trends. Moreover, these trends define structurally clustered dynamical modes and can distinguish regions that tend toward order or disorder at higher temperatures. Assaying several thermophilic EIC mutants, we show that complementary mesophilic mutations to the most temperature-sensitive positions exhibit the most enhanced activity while mutations to relatively temperature insensitive positions exhibit the least enhanced activities. These results provide a mechanistic explanation of sequence-dependent temperature tuning and offer a computational method for rational enzyme modification.

Introduction

Enzymes are sophisticated catalytic machines, designed by billions of years of natural selection for optimal activity at their physiological temperatures[1–3]. To fold rapidly and function reliably enzyme sequences have evolved to minimize strong energetic frustration, a demand met by forming rigid scaffolds[4,5]. Studies of energy landscapes of proteins however, have shown that some degree of frustration is beneficial for facilitating functional dynamics[6]. Enzyme active sites in particular have been found to be highly frustrated regardless of oligomeric state, overall topology, or catalytic mechanism[7,8]. Thus, enzymes are shaped to meet conflicting functional needs for folding and functional dynamics which manifests in having both rigid scaffolds and disordered regions. The fact that homologous enzymes can share an identical fold but exhibit drastically different rates of conformational exchange[9–11] suggests an important functional roles for disordered regions in tuning temperature-dependent dynamics. Identifying this temperature tuning mechanism is an important step in learning the key design principles of thermostable enzymes.

A homologous mesophilic-thermophilic enzyme pair is an ideal system for investigating this temperature tuning mechanism. To study the temperature tuning mechanisms of residue substitutions, we employed the well characterized Enzyme I C-terminal domain (EIC) (fig 1).

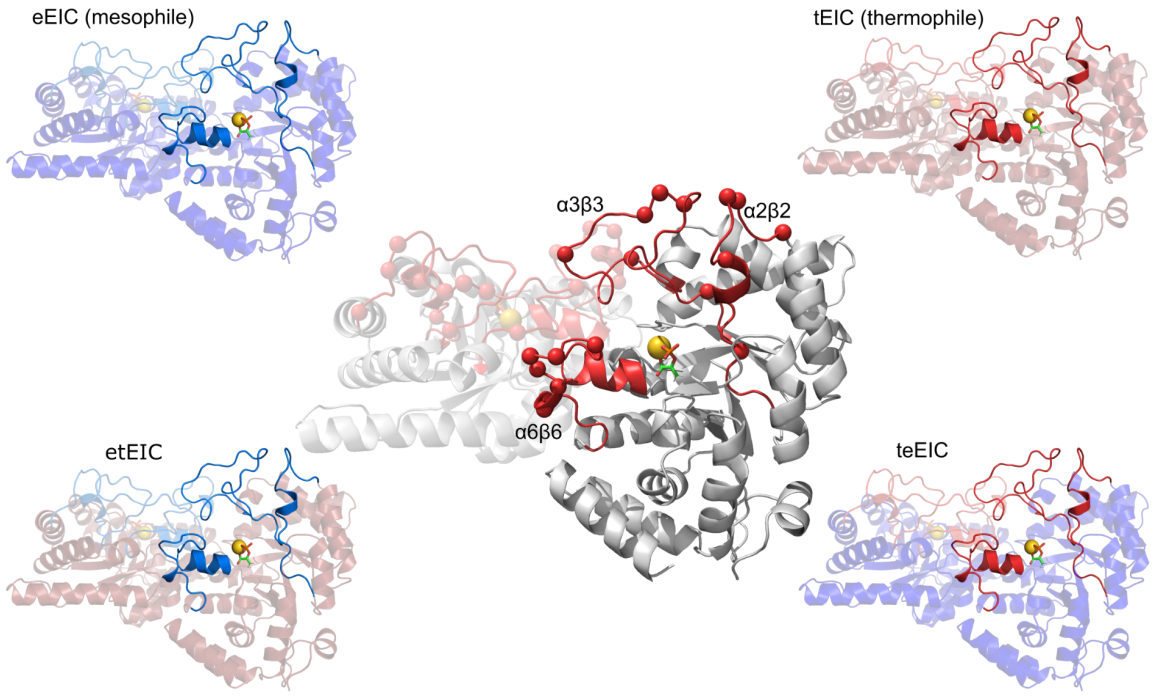

Figure 1. Top: EIC Hybrid construction from wild type homologs. In the center is the EIC structure with mutated residues highlighted as red spheres and names of loops show on top. In the upper left corner is the eEIC; a mesophilic homolog of EIC. In the upper right corner in the tEIC; a thermophilic homolog of EIC. In the bottom right corner is the etEIC; a hybrid formed with a thermophilic scaffold and mesophilic loop. In the bottom right is the teEIC; a hybrid formed with a mesophilic scaffold and thermophilic loop.

Here we used the mesophilic EIC from *Escherichia coli* (eEIC), with physiological temperature of 37°C (for the natural, full-length enzyme), its structurally-identical thermophilic counterpart from *Thermoanaerobacter tencongensis* (tEIC) with a physiological temperature of 65°C (full-length), and two hybrids composed of each wild types' scaffold and disordered catalytic loops swapped with one another (fig 1) . That is, teEIC contains the thermophile's loops on the mesophile's scaffold and etEIC contains the mesophile's loops on the thermophile's scaffold. [12]. There are 21 residue substitutions that distinguish the mesophilic and thermophilic loops but importantly, all of the catalytic residues are conserved and there are no deletions or insertions.

The C-terminal dimer catalyzes a phosphoryl transfer reaction from phosphoenolpyruvate (PEP) to water, a reaction that is dependent on the disordered loops assuming a compact conformation[13]. The loops can also assume a catalytically incompentent expanded conformation which involves a large magnitude motion of the α3β3 away from the active site. The differences in activity between the wild types can largely be explained by the populations of the expanded or compact conformation determined as a function of the distance between K340 (on the α3β3 loop) and PEP[12].

Previous work demonstrated that the catalytic loops of etEIC exhibit the same expanded to compact conformational exchange rate as the wild type mesophile at 37°C, indicating that this hybrid's loop dynamics are encoded by the mutations exclusively within the loops [12]. This implies that, in the presence of a stable scaffold, catalytically relevant dynamics can be tuned with a local subset of mutations. This is consistent with other studies demonstrating the amenability of thermophilic enzymes to modification[14]. In addition to shared dynamics, etEIC and eEIC have similar catalytic rates[12,15], implying that these dynamics underlie catalysis.

With the loop dynamics isolated, we can easily detect the effects of individual mutations on activity. We hypothesize that the most temperature sensitive inter-residue contacts serve to tune the enzyme's activity to its physiological temperature. To test our hypothesis we have sampled catalytic loop conformational ensembles across wide range of temperatures for each EIC homolog using Hamiltonian Replica Exchange Molecular Dynamics (HREMD) simulations. We then calculated the contact frequencies of all loop residue pairs at each temperature and applied Principal Component Analysis (PCA) to extract the most temperature sensitive contacts via their loading scores. Finally, we used these most sensitive contact pairs as a guide to design mutant tEIC enzymes based on the natural substitutions in eEIC. These mutants returned activities corresponding well with their PCA based temperature sensitivity rankings. These results strongly support our hypothesis and the utility of our analysis method which we expect to be a valuable research tool for enzyme design.

Results

Identifying the temperature sensitive contacts in enzyme loops. In order to identify the contacts in the tEIC catalytic loop conformational ensemble which are most sensitive to temperature changes, we simulated the enzyme across 20 temperatures ranging from 310K° to 510K°. Sampling is done by Hamiltonian Replica Exchange Molecular Dynamics (HREMD), a sampling scheme also known as REST2 (Replica exchange with solute scaling) [16]. The objective of HREMD simulations is to enhance the sampling of disordered loop conformations by reducing the strength of the non-bonded interactions involving protein atoms. While the entire

system is not heated, the effect of hamiltonian perturbation is similar and has been shown to reproduce melting curves and folding temperatures with high accuracy[17]. From the sampled MD trajectories we calculated all loop residue contacts with cutoff distances dependent on the chemical groups involved [18]. Contact distances collected from equilibrium sampling at different temperatures are converted to a contact frequency dataset. The contact frequency datasets have (NT, NP) shape where NT (number of rows) corresponds to temperatures sampled and NP (number of columns) corresponds to contacting residue pairs. An example dataset is available in supplementary figure (SI 1). To reduce the dimensionality of the dataset and to rank the temperature sensitivity of specific residue interactions we employed Principal Component Analysis (PCA).

Classifying temperature dependence of contacts. Here we used PCA on the contact frequency dataset described in the previous section to identify and rank residue-residue interactions which are most sensitive to the changes of temperature. Eigenvalues of PCs in this context quantify the degree of temperature-dependent variance of various collective residue contact making/breaking modes [Fig 2AB]. The contact pairs' loading scores (components of eigenvector) describe the individual contribution of temperature to the residue-residue contact frequency variance [Fig 2C]. In other words, loading scores provide a natural rank for contact's temperature sensitivity in terms of the dominant trend of the PC.
Our hypothesis predicts that contacts with very high sensitivity to temperature have the most influence on activity. We identify a highly temperature sensitive contact as occurring in the 98th percentile of the normalized loading score distributions. These include the top 7 contacts on each PC, entirely in the narrowest end of the loading score histogram tail [fig 2C].

Because our previous work demonstrated that the loop dynamics of tEIC are essentially uninfluenced by scaffold interactions (etEIC kex/ kcat) we directed our mutation experiments and focused our simulation analysis on tEIC. Other homologs display similar PC trends indicating that tEICs data is representative of the general behavior of the loops regardless of sequence composition (SI 3).

For tEIC, PC 1 describes nearly 85% of the variance and captures a linear change in contact frequencies with temperatures (fig 2). This represents the dominant effect of heating, increasing configuration entropy via progressive contact breaking toward denaturation. The next PCs identify the less obvious effects of heating. Since their eigenvalues decay exponentially, we focused on the first four PCs which accounted for over 98% of variance and disregarded the remaining PCs.

PC 2 describes less than 10% of the variation in the data, however it captures an important trend of contact frequencies that increase parabolically [fig 2B & D Blue line], becoming progressively more stable across a wide range of temperatures. These are likely to be highly relevant to function as they represent increased order at higher energy states, a fundamental feature of catalysis [19,20].
PC 3 captures a sharply increasing contact frequency within the physiological temperature range before rapidly decaying at higher temperatures. For tEIC this decay begins around its optimal physiological temperature. PC 4's trend is similar to that of PC 3 [fig 2B] but proved to be more predictive of activity enhancing natural substitutions ('Mutation Experiments' Section). The remaining PC trends are less obvious and begin to resemble noise after the 4th PC.

By applying PCA and loading score 'filtering' to the contact frequency dataset, we reduced the initially intractable web of contacts to only the contacts with the strongest response to changes in temperature (fig 2E).

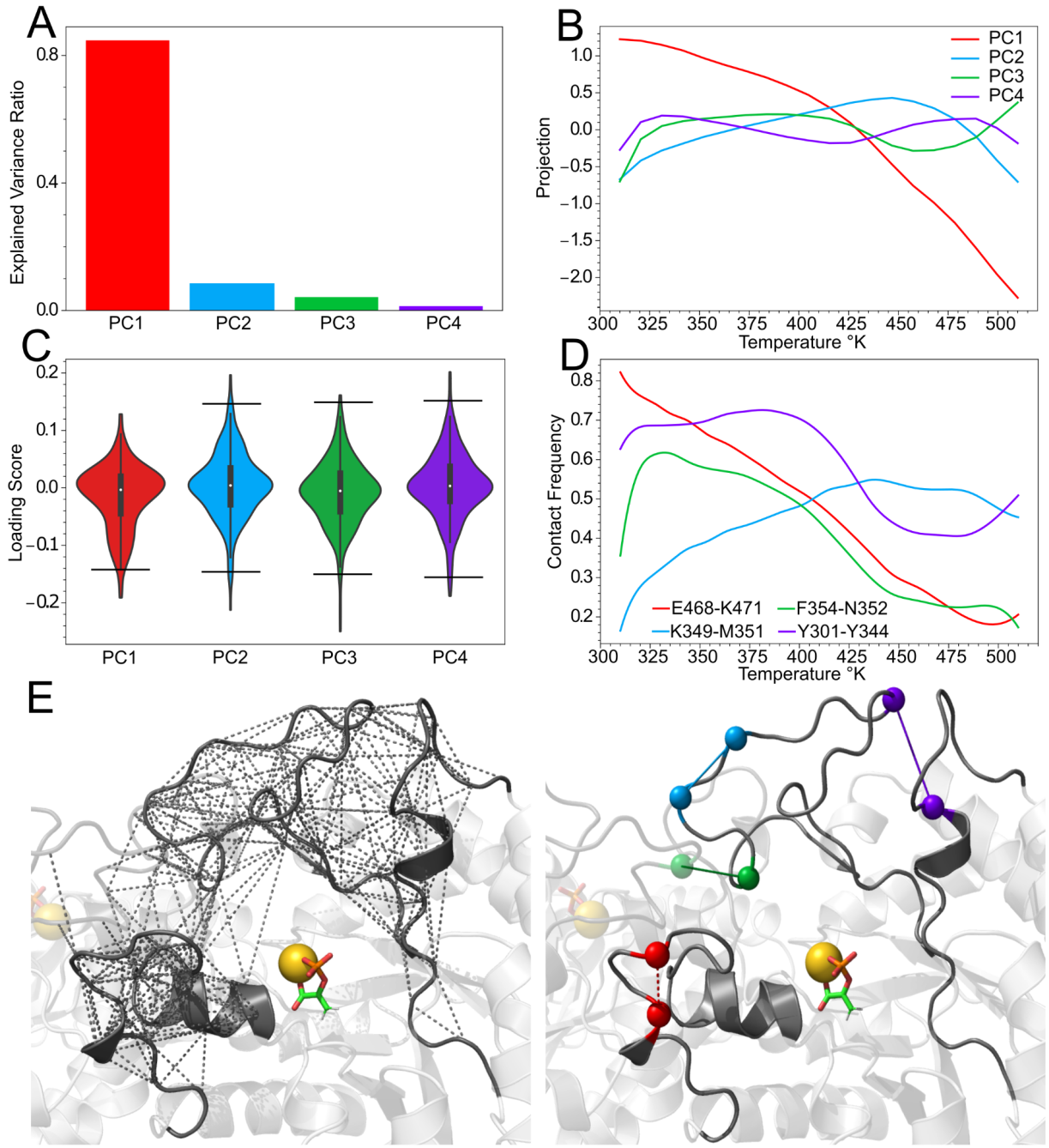

Figure 2. (A) tEIC contact frequency explained variance by PC. (B) Eigenvector components vs effective loop temperature. (C) Histograms of PC loading scores with bars indicating the cutoff for the 98th percentile of score absolute values constituting the highly temperature sensitive contacts. (D)Contact frequency vs temperature for the highest ranking naturally substituted temperature sensitive contacts from PCs 1-4.

Format of contact name is "subunit_id:residue name: residue index". (E) left: All the averaged contacts depicted on one subunit before PCA categorization and ranking. Right: The most temperature sensitive contacts distilled from the left image.

Mapping temperature sensitive contact networks. A global view of the contacts and their PC ranked temperature sensitivity is best visualized with contact maps (Fig 3). On PC1 the highly temperature sensitive contacts form an obvious cluster (Fig 3 'PC1') on the α6β6 loop, which is more apparent when viewed on the structure (fig 3). These high ranked PC1 contacts have a strong tendency to decrease across the full temperature and frequency range. Such obvious clustering of a dynamic mode suggests that nature has assigned a unique function to this loop. The fact that many of the most sensitive contacts in the α2β2 and α3β3 loops tend to increase in frequency in the physiological temperature range suggests that nature has segregated the dynamics of the different loops according to categories of order and disorder (SI 5). It may be that the α6β6 loop acts as a sort of 'entropy antennae', absorbing thermal energy and, through a specific set of interactions, powering the other loops' motions into more coordinated dynamics. Indeed, it has been shown that enzyme activity can be tuned simply by modifying the length of a disordered region which changes the amount of entropic force generated by its random motions and shifting the enzyme's population to a state of higher affinity toward an allosteric inhibitor[21].

Overall, the narrow tails of the loading score histograms and limited number of hot colors within the contact maps clearly indicate that highly temperature sensitive contacts are uncommon within tEIC's loops. This observation is consistent with our hypothesis, in that limited numbers of highly temperature sensitive contacts do indeed exist. Had all of the contact loading scores fallen within a narrow range, it would seem unlikely that any one contact could provide substantially more influence on the enzyme's dynamics than another. Since many of these highly temperature sensitive contacts involve natural substitutions, including some of the highest ranked contacts, it is reasonable that temperature dependent dynamics acts as a selection pressure. The notion that a small subset of residues could be responsible for tuning activity is also consistent with many mutation studies reporting single non-catalytic residue mutations that drastically alter or abolish function [22,23].

PC2 is the first PC to capture a number of strongly associating contacts with parabolic profiles. On PC2, the highest ranking contact involves M351. This natural substitution is in the middle of 9 conserved residues on the α3β3 loop and directly involved in the structurally and functionally vital dimer interface[24]. A glutamate in the mesophile, this residue represents the largest change in hydrophobicity among the loop substitutions suggesting it has a significant impact on function.

The top sensitive contacts on PCs 3 and 4 are entirely devoid of residues in the α6β6 loop further supporting the notion that the α6β6 loop is defined by a distinct PC 1 dynamic mode. PC 3 and 4's sensitive contacts share the same regions of the structure and display similar frequency trends however PC 4 appears shifted more towards the α2β2 loop indicating that it better defines this region. Additionally, their score plots (SI 6) show several of their high ranked contacts share extreme values on both PCs. As such, there is likely to be a gradient between these modes, perhaps including PC 2, with certain residues acting as junctions between them.

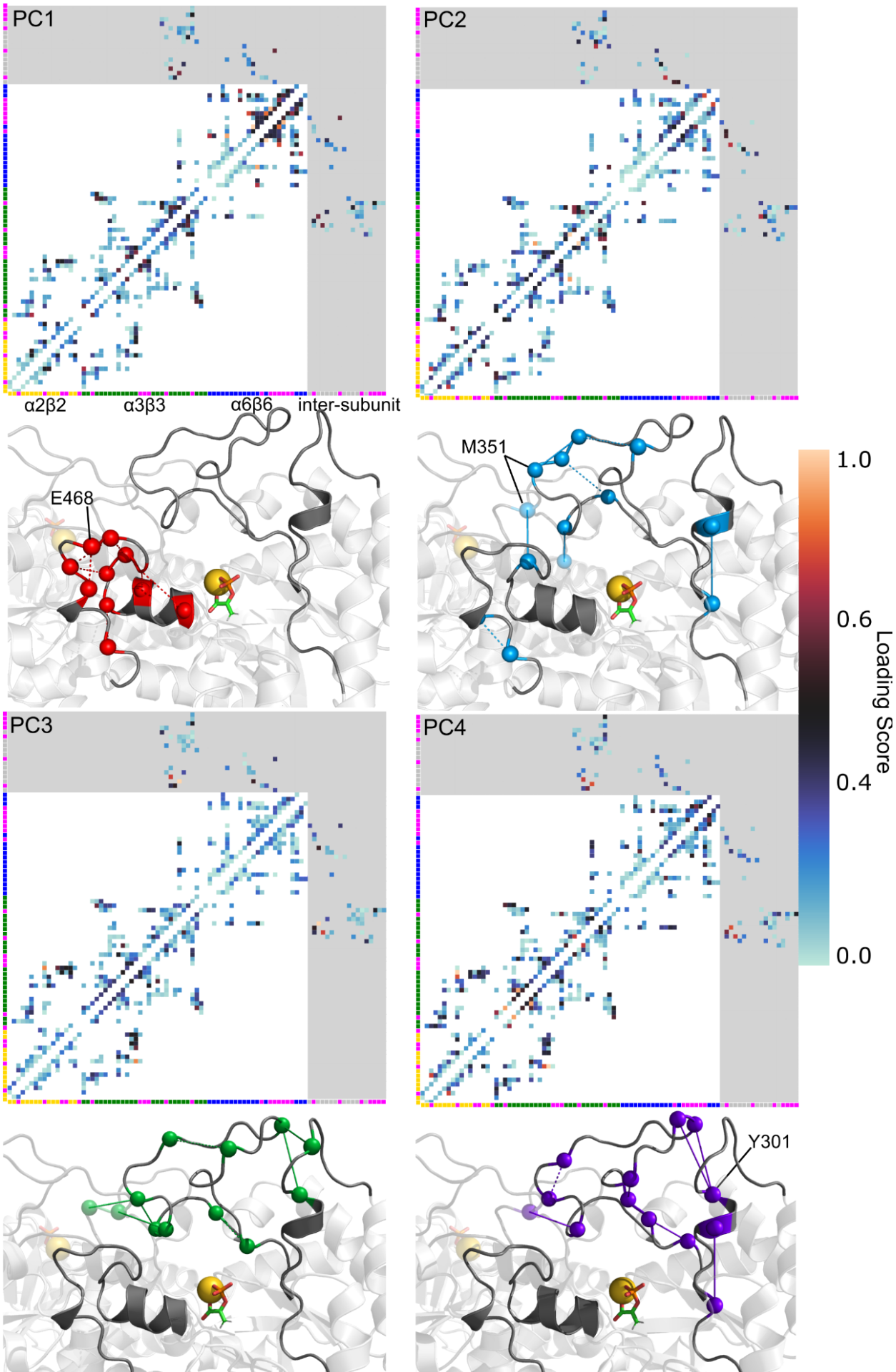
PC1
PC2
α2β2
α3β3
α6β6
inter-subunit
E468
M351
1.0
0.6
0.4
0.0
Loading Score
PC3
PC4
Y301

Figure 3. Contact maps showing normalized loading scores (temperature sensitivity) according to the colorbar (right). Contact maps correspond to the PC indicated by colors of the spheres on the structures below them. Red: PC1, Blue: PC2, Green: PC3, Purple: PC4. Solid lines indicate contacts that have a tendency in the physiological temperature range to increase in frequency with temperature, dashed lines tend to decrease in frequency in the same range. The tick marks on the axes of the contact maps indicate residues and are colored as follows: yellow: α2β2, green: α3β3, blue: α6β6 and magenta indicates a naturally substituted position.

Though we speculate on the selection pressures and functions of temperature sensitive contacts, experimental results are necessary to tell us whether they have any biological relevance.

Mutation Experiments. The EI homologs provide a known set of function-conserving mutations between the mesophile and the thermophile. If the residues that occur in highly temperature sensitive contacts are among the natural substitutions, it suggests that these positions have been selected based on their simultaneous tuning of loop dynamics and preservation of catalytic conformational space.

To investigate this, we made point mutations to several of tEIC's naturally substituted positions with the complementary residues of eEIC. We assayed these mutants for activity, by measuring the rate of PEP hydrolysis at 40° C, a temperature at which tEIC is nearly inactive, and compared their initial rates with etEIC (the hybrid containing the mesophilic loop residues), which provided the hypothetical maximum rate we would expect from a subset of the loop mutations to tEIC at 40° C[12]. In our model, the hydrolysis rate of PEP indicates how effectively the mutation has enhanced the disordered loop dynamics underlying catalysis. For the natural substitutions that do not appear in the most temperature sensitive contacts (the top 98th percentile of loading scores on the first four PCs), we expected their complementary point mutation to yield the least enhanced activities.

The assay results ultimately revealed the exceptional correspondence between tEIC's mutant activities and the individual contact rankings of its PCs 1 through 4 [fig 4B].

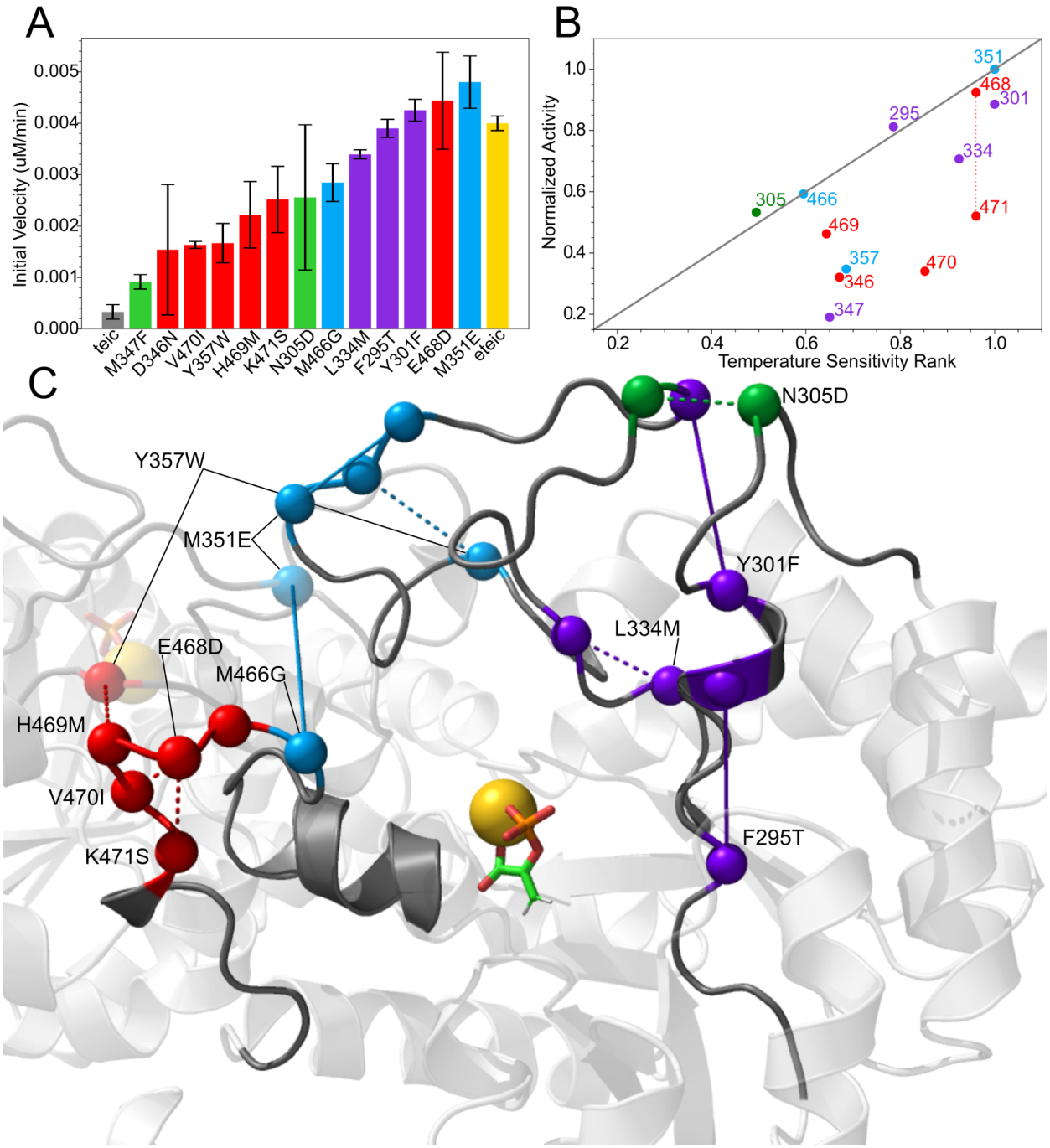

Figure 4.A: Assay results for tEIC mutations colored according to the PC in which they have the highest loading score contact (red: PC1, blue: PC2, green: PC3, purple: PC4). B: Mutant activity vs temperature sensitivity rank on their highest tEIC loading scoring contact PC. dotted line indicates that the loading score for 468 and 471 comes from the same contact C: Contacts involving residues from the mutant assays colored again according to the PC on which their highest loading score occurs (lowest activity enhancing mutations on 346 and 347 contacts not shown).

Remarkably, the mutations that returned the highest activities are indeed participants in the highest ranked substituted contacts and thus, the most temperature sensitive.

The top PC1 guided mutants (V470I, K471S, E468D) resulted in mixed activity enhancement (Fig 4A) with the overall top ranked substituted contact on PC 1 (2nd overall) involving residues 468 and 471. While the K471S mutant returned moderate activity enhancement, E468D returned the second highest enhancement of any mutation suggested by the 4 PCs. The residue M351 is involved in the highest ranked temperature sensitive tEIC contact on PC 2 and its mutant resulted in the highest activity enhancement among all of our mutants. Its contact with the conserved K349 exhibits a strong parabolic trend, increasing in frequency over a large temperature range. These increasing contact frequency trends are perhaps the most interesting and relevant to catalysis.

The third highest mutant activity was returned by Y301F. Y301's contact with residue Y344 ranks 2nd and 1st on PC 3 and 4 respectively.

The F295T mutant returned the 4th highest activity with its contact to the conserved residue 299 ranking 4th on PC 2 and in the top 6 of PC 4.

The fifth most active mutant, L334M, ranks 3rd on PC4, the only substituted residue outranking it on this PC being 301. The thermophilic leucine makes many contacts involving loop and scaffold residues. It makes a significant contact with Y301 and is thus involved in the same network that is responsible for the large rate enhancement seen with the Y301F mutant and likely that of F295 as all of these residues have high scoring contacts on PC4.

Depicting the experimentally mutated residue's highest ranking contacts such that they are colored according to the PC on which they have the highest loading scores results in a clear structural clustering [fig 4C]of the different modes defined by the PCs. The α6β6 and α2β2 loops are characterized by their own dynamic modes while the α3β3 exhibits a PC2 mode near the subunit interface and a PC4 mode near the α2β2 interface. Visualizing the contacts according this same color scheme for all of the top contacts on PCs 1-4 reveals that PC3 describes conserved residue contacts that interact at the interface of the 2 subunits' α3β3 loops (SI 6) with strong structural clustering maintained by PCs 1 and 4. A mutation that returned unexpected activity enhancement was N305D which did not appear in any of tEICs top contacts but does appear as the highest ranked PC 2 contact for eEIC data.

The PC data can differ significantly between the homologs and can be useful for interpreting outlier data such as N305D that might not serve the same role in each homolog or in the context of different neighboring residues. Nevertheless, the highest ranking (or most temperature sensitive) tEIC PC data corresponds very well with the tEIC mutant activity results. The PC1, 2, and 4 represent distinct modes and it is interesting to note that the top 3 enhancing mutations come from each of these PCs respectively [fig 4 A]. It seems that each mode has nearly equal influence on activity at least from the residue involved in its most sensitive contact.

Why nature selected so many mutations for lower temperature activity when a single mutation can accomplish the same enhancement may be for a number of reasons including stability of folding intermediates, the sequence that the natural mutations occurred in, modulation of enzymatic efficiency, altered activity in the presence of the host organism's micromolecular solutes, interactions with the removed N-domain or other biomolecules in the enzymatic pathway of EIC. Many of the loop mutations may be functionally silent except perhaps for effects on thermal-stability evidenced by the fact that the point mutations have less of an effect on thermal-stability than the full set of loop mutations in etEIC [SI 8]. Very little thermal-stability was sacrificed by the most activity enhancing mutation of M351E exemplifying

the value of rationally guided mutations in settings where desirable features (thermal-stability) need to be retained by minimal perturbations to the wild type sequence.

Among the top PCs, a larger eigenvalue does not appear to be more predictive of activity enhancing mutations, however the loading scores clearly are.

These temperature sensitivity rankings are also significant seeing that the conserved residues outnumber and could have excluded the substituted residues from the most temperature sensitive PC ranks. Not only do we see substituted residues in the highest ranked, most temperature sensitive positions, we also see residues with the highest rank having maximum effect on activity. This remarkable rank-activity relationship is lending additional support to our hypothesis that temperature sensitivity is a natural mechanism of tuning activity.

Discussion

Energy landscape theories of proteins have shown that frustration and disorder are evolutionarily conserved features facilitating functional dynamics of proteins[7,8]. The nature of interactions which balance the competing needs for having ordered catalytic sites and dynamical disorder, however, are currently not well understood. Here we have employed molecular dynamics simulations as well as in vitro assay experiments on the C domain of bacterial Enzyme I (EIC) to reveal a mechanism of temperature-tuning in enzymes.

We find that below a functionally optimal temperature for a thermophilic enzyme, critical residue contacts are not accessible in disordered loop motions. As temperature increases, certain contacts break, allowing other critical contacts to form. Eventually a temperature is reached where the optimal interactions are occurring to cycle the enzyme through catalytically relevant conformations at a rate tuned to the functional needs of the organism. This enhancement can involve a critical contact whose altered dynamics promote relevant conformational sampling. By mutating a specific residue involved in the critical contact, the enzyme can sample the catalytic conformations at a lower temperature.

Furthermore, we present a computational pipeline for generating and ranking temperature-sensitive contact networks which can identify residues critical for temperature sensitivity of enzymes. We find highly temperature-sensitive loop contacts displaying linear and parabolic temperature dependent frequency trends. These trends define structurally clustered dynamical modes and can distinguish contacts which tend toward order or disorder at higher temperatures. Assaying several thermophilic EIC mutants we show that complementary mesophilic mutations to the most temperature sensitive positions exhibit the most enhanced activity while mutations to relatively temperature insensitive positions exhibit the least enhanced activities.

We envision temperature sensitive contact analysis providing insights into temperature sensing mechanisms of systems such as ion channels, biomolecular condensates, and functionally relevant dynamics of biomolecular assemblies in general. Finally it is worth speculating that regions involving temperature sensitive contacts might make excellent drug targets. Dynamical mode centers important for activity might be disrupted or have their dynamics quenched by the binding of small molecules. Our future work will explore new protein systems and novel applications, complementary methods, and enhancements to temperature sensitive contact analysis.

Materials and Methods

Simulation Protocol. Crystal structure for tEIC was obtained from PDB entries 2BG5 (tEIC). Molecular dynamics simulations were carried out using Gromacs 2018.8 (The last 50ns of teEIC was done with 2019 on Expanse) utilizing the leapfrog integration method with a 2fs timesteps [25–28]. Short range electrostatic and Lennard-Jones interactions were calculated with a plain coulomb cutoff of 1.2 nm. The Particle mesh Ewald electrostatics scheme was employed for long range electrostatics with a grid spacing of 0.16 nm [29]. Bonds to hydrogen were constrained using the LINCS algorithm [30]. The system was solvated in a dodecahedral simulation box with TIP3P water molecules and neutralized with 100mM NaCl. The solvent and solute were separately coupled to a modified Berendsen thermostat (Velocity Rescale) with a reference temperature of 310K. Simulations included PEP to simulate the dynamics relevant to the rate limiting chemical step. The system was described according to the CHARMM 36 force field with the PEP ligand molecule parameterized with cgenff [31–33]. The system was energy minimized with the steepest descent minimization and equilibrated in the NVT and NPT ensembles for 100 ps.

Hamiltonian Replica Exchange Molecular Dynamics Simulations. To efficiently sample the temperature dependent conformational ensembles of Enzyme I disordered loops we employed Hamiltonian Replica Exchange Molecular Dynamics (HREMD) [34–36], specifically the REST2 solute tempering scheme. The choice of the Hamiltonian method was motivated by the size of the Enzyme which makes application of temperature replica exchange (TREMD) computationally costly. The residues 291-309, 332-360, and 454-477 within the active site loops of the constructs were selected for hamiltonian perturbation. Residues within these loops were identified as central to catalysis in previous work via NMR and structure visualization[12]. A total of 20 replicas with an effective temperature range of 310-510° K was chosen as it gave an exchange rate around 25%. Hamiltonian replica exchange simulations were performed with gromacs 2018.8 patched with Plumed 2.5.5 on Xsede's Comet and Expanse HPC systems [37,38][39]. Exchanges were attempted every 800 timesteps. Trajectory snapshots were extracted every 10ps with the exchange acceptance rate of approximately 25%. Simulations were conducted for 200ns which generated an effective 400ns sampling by combining contact data from the two subunits. A wall potential restricting the PEP molecule to within 6.5Å of cysteine 502, a residue centered beneath the binding site, was employed to prevent unbinding of PEP and ensure enhanced sampling of the relevant motions associated with the bound substrate. A weak 50 kJ/mol dihedral restraint was also applied to the CCOP dihedral angle of PEP to bias the hypothetically catalytically competent PEP conformation.

Contact analysis. We calculated the inter-atomic contacts and contact frequencies of the loop residues using GetContacts[40] and MDAnalysis python library [41]. Contacts between directly adjacent residues were discarded as well as contacts whose frequencies did not exceed 5% across the 20 replicas. The frequencies for both subunits were averaged and contacts not

occurring in both subunits were also discarded. All subsequent calculations were performed on the averaged values. The PCA loading scores were normalized [maximum - minimum absolute value/ maximum]. The 98th percentile of these scores was calculated to determine which contacts to consider highly temperature sensitive. Molecular visualizations were produced with pymol.

Mutations. Point mutations were performed according to the Agilent Quickchange II protocol.

Enzyme Expression. Plasmids for etEIC and tEIC constructs with His and EIN solubility tags were transformed into BL21 cells and cultured on LB media overnight. All culture media contained 100µg/ml ampicillin. Single colonies were grown in 20ml of LB media overnight at 37°C and used to inoculate 1L LB cultures. These were grown to an OD600 of 0.6 before cooling to 20°C and adding 0.238g of IPTG for overnight expression. Cultures were spun at 4,000xg for 30 minutes and resultant pellets frozen at -80°C.

Enzyme Purification. Pellets were thawed on ice and cells lysed and homogenized at (1200 psi) via Avestin Emulsiflex in 20ml of buffer containing 50mM Tris pH 8, 300mM NaCL, and 5mM imidazole. Lysate was centrifuged at 37,000 x g for 60 minutes followed by filtration through a 10kDa membrane and loaded on to a Bio-Rad Nuvia Ni-charged column. Protein was eluted with a 500mM Imidazole buffer applied as a gradient. Eluted protein was buffer exchanged with the original buffer and incubated with TEV protease to remove the solubility tag. Samples were filtered and again passed through the Ni column where the solubility tag was retained on the eluted EIC collected. The fractions were buffer exchanged with 20mM Tris pH 8, 1mM EDTA, and 2mM DTT before being applied to a Bio-Rad EnrichQ column and eluted with a 1M NaCL buffer gradient. The purified product was buffer exchanged with water and stored at -80°C.

Activity assays. (A280/ uvVis) values were used with molar extinction coefficients obtained from (expasy protparam) sequence analysis to determine protein concentrations. Enzymes were diluted to a concentration of 50 uM in an NMR buffer containing 3mM PEP and 1mM TSP. Reactions at 40°C were monitored by the decrease in PEP (vinyl hydrogen?) peak intensity at 5 minute intervals on a 700MHz Bruker NMR over the course of 2 ½ hours. PEP concentration was normalized against TSP peak intensity. Enzyme concentrations from the experimental samples were normalized against the etEIC experimental sample concentration determined by SDS PAGE band intensities with imageJ [42]. The average concentration (band intensities) from the two etEIC replicates was divided by the mutant's band intensity to produce a normalization coefficient for each corresponding rate.

Data availability. All study data are included in the article and/or SI.

Acknowledgements. This work was supported by funds from the National Institute of General Medical Sciences with grant no. R35GM133488 [to V. V.] and grant no. R35 GM138243 [to D.P.]

References

1. Graves JE, Somero GN. Electrophoretic and Functional Enzymic Evolution in Four Species of Eastern Pacific Barracudas from Different Thermal Environments. Evolution. 1982;36: 97–106.

2. Arcus VL, van der Kamp MW, Pudney CR, Mulholland AJ. Enzyme evolution and the temperature dependence of enzyme catalysis. Curr Opin Struct Biol. 2020;65: 96–101.

3. Nguyen V, Wilson C, Hoemberger M, Stiller JB, Agafonov RV, Kutter S, et al. Evolutionary drivers of thermoadaptation in enzyme catalysis. Science. 2017;355: 289–294.

4. Röder K, Wales DJ. Evolved Minimal Frustration in Multifunctional Biomolecules. J Phys Chem B. 2018;122: 10989–10995.

5. Chen M, Chen X, Schafer NP, Clementi C, Komives EA, Ferreiro DU, et al. Surveying biomolecular frustration at atomic resolution. Nat Commun. 2020;11: 5944.

6. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG. On the role of frustration in the energy landscapes of allosteric proteins. Proc Natl Acad Sci U S A. 2011;108: 3499–3503.

7. Freiberger MI, Guzovsky AB, Wolynes PG, Parra RG, Ferreiro DU. Local frustration around enzyme active sites. Proc Natl Acad Sci U S A. 2019;116: 4037–4043.

8. Gianni S, Freiberger MI, Jemth P, Ferreiro DU, Wolynes PG, Fuxreiter M. Fuzziness and Frustration in the Energy Landscape of Protein Folding, Function, and Assembly. Acc Chem Res. 2021;54: 1251–1259.

9. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nat Struct Mol Biol. 2004;11: 945–949.

10. Bhabha G, Ekiert DC, Jennewein M, Zmasek CM, Tuttle LM, Kroon G, et al. Divergent evolution of protein conformational dynamics in dihydrofolate reductase. Nat Struct Mol Biol. 2013;20: 1243–1249.

11. Callender R, Dyer RB. The dynamical nature of enzymatic catalysis. Acc Chem Res. 2015;48: 407–413.

12. Dotas RR, Nguyen TT, Stewart CE Jr, Ghirlando R, Potoyan DA, Venditti V. Hybrid Thermophilic/Mesophilic Enzymes Reveal a Role for Conformational Disorder in Regulation of Bacterial Enzyme I. J Mol Biol. 2020;432: 4481–4498.

13. Venditti V, Tugarinov V, Schwieters CD, Grishaev A, Clore GM. Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. Nat Commun. 2015;6: 5960.

14. Finch AJ, Kim JR. Thermophilic Proteins as Versatile Scaffolds for Protein Engineering. Microorganisms. 2018;6. doi:10.3390/microorganisms6040097

15. Venditti V, Clore GM. Conformational selection and substrate binding regulate the

monomer/dimer equilibrium of the C-terminal domain of Escherichia coli enzyme I. J Biol Chem. 2012;287: 26989–26998.

16. Wang L, Friesner RA, Berne BJ. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). J Phys Chem B. 2011;115: 9431–9438.

17. Stirnemann G, Sterpone F. Recovering Protein Thermal Stability Using All-Atom Hamiltonian Replica-Exchange Simulations in Explicit Solvent. J Chem Theory Comput. 2015;11: 5573–5577.

18. getcontacts: Library for computing dynamic non-covalent contact networks in proteins throughout MD Simulation. Github; Available: https://github.com/getcontacts/getcontacts

19. Kovermann M, Ådén J, Grundström C, Sauer-Eriksson AE, Sauer UH, Wolf-Watz M. Structural basis for catalytically restrictive dynamics of a high-energy enzyme state. Nat Commun. 2015;6: 7644.

20. Hammes GG, Benkovic SJ, Hammes-Schiffer S. Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. Biochemistry. 2011;50: 10422–10430.

21. Keul ND, Oruganty K, Schaper Bergman ET, Beattie NR, McDonald WE, Kadirvelraj R, et al. The entropic force generated by intrinsically disordered segments tunes protein function. Nature. 2018;563: 584–588.

22. Liu B, Qin F. Single-residue molecular switch for high-temperature dependence of vanilloid receptor TRPV3. Proc Natl Acad Sci U S A. 2017;114: 1589–1594.

23. Gnutt D, Timr S, Ahlers J, König B, Manderfeld E, Heyden M, et al. Stability Effect of Quinary Interactions Reversed by Single Point Mutations. J Am Chem Soc. 2019;141: 4660–4669.

24. Nguyen TT, Ghirlando R, Roche J, Venditti V. Structure elucidation of the elusive Enzyme I monomer reveals the molecular mechanisms linking oligomerization and enzymatic activity. Proc Natl Acad Sci U S A. 2021;118. doi: 10.1073/pnas.2100298118

25. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1-2: 19–25.

26. Páll S, Abraham MJ, Kutzner C, Hess B, Lindahl E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. Solving Software Challenges for Exascale. Springer International Publishing; 2015. pp. 3–27.

27. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013;29: 845–854.

28. Miyamoto S, Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comput Chem. 1992;13: 952–962.

29. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys. 1995;103: 8577–8593.

30. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem. 1997;18: 1463–1472.

31. Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem. 2013;34: 2135–2145.

32. Vanommeslaeghe K, MacKerell AD. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. Journal of Chemical Information and Modeling. 2012. pp. 3144–3154. doi:10.1021/ci300363c

33. Vanommeslaeghe K, Raman EP, MacKerell AD Jr. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. J Chem Inf Model. 2012;52: 3155–3168.

34. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett. 1999;314: 141–151.

35. Okabe T, Kawata M, Okamoto Y, Mikami M. Replica-exchange Monte Carlo method for the isobaric–isothermal ensemble. Chem Phys Lett. 2001;335: 435–439.

36. Fukunishi H, Watanabe O, Takada S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. J Chem Phys. 2002;116: 9058–9067.

37. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: New feathers for an old bird. Comput Phys Commun. 2014;185: 604–613.

38. Bussi G, Tribello GA. Analyzing and Biasing Simulations with PLUMED. Methods Mol Biol. 2019;2022: 529–578.

39. Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, et al. XSEDE: Accelerating Scientific Discovery. Computing in Science & Engineering. 2014. pp. 62–74. doi:10.1109/mcse.2014.80

40. Venkatakrishnan AJ, Fonseca R, Ma AK, Hollingsworth SA, Chemparathy A, Hilger D, et al. Uncovering patterns of atomic interactions in static and dynamic structures of proteins. bioRxiv. 2019. p. 840694. doi:10.1101/840694

41. Gowers RJ, Linke M, Barnoud J, Reddy TJE, Melo MN. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. 2019. Available: https://www.osti.gov/servlets/purl/1565806

42. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9: 676–682.