bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Computational remodeling of an enzyme conformational landscape for altered substrate selectivity

Antony D. St-Jacques,[1,2] Joshua M. Rodriguez,[3] Matthew G. Eason,[1,2] Scott M. Foster, [1,2] Safwat T. Khan,[1,2] Adam M. Damry,[1,2] Natalie K. Goto,[1,2] Michael C. Thompson,[3] Roberto A. Chica[1,2,*]

[1] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5

[2] Center for Catalysis Research and Innovation, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5

[3] Department of Chemistry and Biochemistry, University of California, Merced, Merced, California 95343, United States

[*] To whom correspondence should be addressed:

Roberto A. Chica, Email: rchica@uottawa.ca

Keywords

Computational protein design; multistate design; enzyme engineering; aminotransferase; transaminase; biocatalysis; substrate selectivity

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Structural plasticity of enzymes dictates their function. Yet, our ability to rationally remodel enzyme conformational landscapes to tailor catalytic properties remains limited. Here, we report a computational procedure for tuning conformational landscapes that is based on multistate design. Using this method, we redesigned the conformational landscape of a natural aminotransferase to preferentially stabilize a less populated but reactive conformation, and thereby increase catalytic efficiency with a non-native substrate to alter substrate selectivity. Steady-state kinetics of designed variants revealed selectivity switches of up to 1900-fold, and structural analyses by room-temperature X-ray crystallography and multitemperature nuclear magnetic resonance spectroscopy confirmed that conformational equilibria favoured the target conformation. Our computational approach opens the door to the fine-tuning of enzyme conformational landscapes to create designer biocatalysts with tailored functionality.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Main text

Enzymes are flexible macromolecules that sample multiple structural states, described by a conformational energy landscape (*1*). It is the relative stability of these conformational states, and the ability of enzymes to transition between them, that ultimately dictates enzymatic function (*2-4*). Analyses of directed evolution trajectories have shown that evolution can reshape enzyme conformational landscapes by enriching catalytically productive states and depopulating non-productive ones (*5-7*), leading to enhanced catalytic activity. Similar mechanisms contribute to the evolution of substrate selectivity, where transient conformations responsible for activity on non-cognate substrates become enriched (*8-10*). Thus, it should be possible to harness the pre-existing conformational plasticity of enzymes to rationally design biocatalysts with tailored catalytic properties by fine-tuning their conformational landscapes. However, predicting the effect of mutations on these energy landscapes remains challenging, and current computational enzyme design protocols, which focus on a single structural state, are poorly optimized for this task. New design methodologies are therefore required for the targeted alteration of subtle conformational states and equilibria, which would in turn facilitate the design of biocatalysts with customized activity and selectivity.

Here, we report a computational procedure for rationally tuning enzyme conformational landscapes that is based on multistate computational protein design, a methodology that allows protein sequences to be optimized on multiple structural states (*11*). As a case study, we remodelled the conformational landscape of aspartate aminotransferase, an enzyme that switches between open and closed conformations via hinge movement, which involves the rotation of a protein domain relative to another around an axis between two planes. Using our approach, we enriched the less populated but catalytically active closed conformation in order to increase

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

catalytic efficiency (k_{cat}/K_M) with the non-native substrate L-phenylalanine, leading to altered substrate selectivity. Steady-state kinetics revealed k_{cat}/K_M increases of up to 100-fold towards this aromatic amino acid, resulting in a selectivity switch of up to 1900-fold, and structural analyses by room-temperature X-ray crystallography and multitemperature nuclear magnetic resonance (NMR) spectroscopy confirmed that the conformational landscape was remodelled to favour the target state. Our methodology for conformational landscape fine-tuning could be incorporated into *de novo* enzyme design pipelines, opening the door to the creation of more complex and efficient designer biocatalysts than previously possible.

Results

Computational remodelling of conformational landscape

E. coli aspartate aminotransferase (AAT) is a pyridoxal phosphate (PLP)-dependent enzyme that catalyses the reversible transamination of L-aspartate with α-ketoglutarate, yielding oxaloacetate and L-glutamate (Supplementary Figure 1a). During its catalytic cycle, AAT undergoes hinge movement to switch between an open conformation, in the ligand-free form, and a closed conformation upon association with substrates or inhibitors (Figure 1a,b) (*12*). Previously, a hexamutant (HEX) of AAT that is approximately two orders of magnitude more catalytically efficient than the wild type (WT) with the aromatic amino acid L-phenylalanine (Table 1, Supplementary Figure 1b), and similarly efficient for transamination of L-aspartate, was engineered by replacing six of the 19 residues that are strictly conserved in AAT enzymes by those found at corresponding positions in the homologous *E. coli* tyrosine aminotransferase (*13*). Unexpectedly, it was found that unlike WT, HEX was closed in its ligand-free form (Figure 1c) (*14*), demonstrating that those six mutations shifted its conformational equilibrium to favour the

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

closed state (Figure 1d), which is the active conformation of the enzyme (*15*). This equilibrium shift was accompanied by enhanced L-phenylalanine transamination activity resulting from a larger increase in affinity for this non-native substrate than for the native L-aspartate substrate (*13*). Thus, we postulated that we could rationally remodel the conformational landscape of AAT by using multistate computational protein design (*11*) to identify novel mutation combinations that can preferentially stabilize the closed conformation over the open conformation, and in doing so, increase substrate selectivity for L-phenylalanine.

To test this hypothesis, we implemented a computational strategy (Figure 2) that proceeds in five steps: (1) identification of hinge-bending residues involved in transition between open and closed conformations; (2) generation of structural ensembles approximating backbone flexibility to model open and closed conformational states; (3) optimization of side-chain rotamers for all allowed amino-acid combinations at key hinge-bending residues and neighboring positions, on each ensemble; (4) calculation of energy differences between open and closed states to predict preferred conformation, and (5) combinatorial library design using computed energy differences to select mutant sequences for experimental testing.

To identify hinge-bending residues for the open/closed conformational transition, crystal structures of WT AAT in its open and closed forms (PDB ID: 1ARS and 1ART, respectively (*12*)) were used as input for hinge movement analysis with DynDom, a program that identifies domains, hinge axes and hinge bending residues in proteins for which two conformations are available (*16*). DynDom analysis (Supplementary Table 1) revealed a small moving domain (Figure 1b) that rotates by 7.1 degrees about the hinge axis from the larger fixed domain, and identified 25 hinge-bending residues. We selected two of these residues, Val35 and Lys37 (numbering based on Uniprot sequence P00509), for design because they are found on the flexible loop connecting the

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

moving and fixed domains (Figure 2). We also selected for design residues Thr43 and Asn64, which are not part of the hinge, but whose side chains form tight packing interactions with those of Val35 and Lys37. Interestingly, these residues comprise four of the six positions that were mutated in HEX (Table 1), demonstrating that our analysis using only WT structures led to the identification of positions that contribute to controlling the open/closed conformational equilibrium in AAT.

Next, we generated backbone ensembles from the open- and closed-state crystal structures to approximate the intrinsic flexibility of these two conformational states using the PertMin algorithm (*17*), which we previously showed to result in improved accuracy of protein stability predictions when used as templates in multistate design (*18*). Using the protein design software Phoenix (*19, 20*), we optimized rotamers for all combinations of proteinogenic amino acids with the exception of proline at the four designed positions on each backbone ensemble, yielding Boltzmann-weighted average energies for 130,321 (19^4) AAT sequences that reflect their predicted stability on each conformational state. To identify mutant sequences that preferentially stabilized the closed conformation, we computed the energy difference between closed and open state ensembles ($\Delta E = E_{closed} - E_{open}$) for each sequence. As a final step, we used these ΔE values as input to the CLEARSS library design algorithm (*19*) to generate a 24-member combinatorial library of AAT mutants predicted to favour the closed state (Closed library, Supplementary Table 2) with a range of values (–86.0 to –9.5 kcal mol^{-1}) encompassing that of HEX (–76.8 kcal mol^{-1}, Table 2). As controls, we also generated two libraries of sequences predicted to favour the open conformation (Supplementary Table 2): the $Open_{Low}$ library, which contains 18 sequences predicted to stabilize the open state with ΔE values (0.8–16.4 kcal mol^{-1}) comparable to that of the WT (14.2 kcal mol^{-1}, Table 2), and the $Open_{High}$ library, which contains 24 sequences predicted to

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

more strongly favour the open state due to substantial destabilization of the closed state by >120 kcal mol^{-1}. While we postulated that the $Open_{Low}$ library would yield mutants with wild-type-like conformational landscapes and therefore similar catalytic efficiency and substrate selectivity, we hypothesized that $Open_{High}$ library mutants would be less efficient than WT with both native and non-native substrates due to their strong destabilization of the closed conformation, which is the active form of the enzyme (*15*). Thus, experimental characterization of these three mutant libraries, which comprise non-overlapping sequences (Figure 2), allowed us to assess the ability of the ΔE metric to predict sequences with conformational landscapes favouring the open or closed states.

Kinetic analysis of designs

We screened the three mutant libraries for transamination activity with the non-native substrate L-phenylalanine (Methods) and selected the most active mutants from each library for kinetic analysis. All selected mutants catalyzed transamination of L-phenylalanine or L-aspartate with α-ketoglutarate, and displayed substrate inhibition with this acceptor substrate, as is the case for WT (Table 1, Supplementary Tables 3–4, Supplementary Figures 2–5). All Closed library mutants displayed catalytic efficiencies towards L-phenylalanine that were improved by approximately two orders of magnitude relative to WT (Table 1), comparable to HEX, and all were similarly or more active with this non-native substrate than with L-aspartate, in stark contrast with WT AAT, which prefers the native substrate by a factor of 100 (Supplementary Figure 6). By contrast, $Open_{Low}$ and $Open_{High}$ mutants favoured the native over the non-native substrate by up to 50-fold, similar to WT. Surprisingly, $Open_{Low}$ mutants had lower K_M values and were more catalytically efficient than WT with both L-phenylalanine and L-aspartate, which could be due to the fact that these mutants have ΔE values similar to the WT but stabilize both the open and closed

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

149 conformations by >10 kcal mol^{-1} (Table 2, Supplementary Table 2). This is not the case for
150 Open$_{High}$ mutants, which have similar K_M values but are less catalytically efficient with the native
151 substrate than WT, in agreement with our hypothesis that strong destabilization of the catalytically
152 active closed conformation would result in less efficient catalysis. Overall, these kinetic results
153 support the hypothesis that the change in substrate selectivity is linked to the conformational
154 equilibrium previously suggested by the data from the HEX mutant (*13, 14*).
155
156 *Structural analysis of designs*
157 To provide structural information on the conformations adopted by AAT mutants, we
158 turned to room-temperature X-ray crystallography, which provides insight into enzyme
159 conformational ensembles under conditions that are relevant to catalysis (*21*) and free of potential
160 distortions or conformational bias introduced by sample cryocooling (*22*). We crystallized WT,
161 HEX, and select variants from the Closed (VFIT and VFIY), Open$_{Low}$ (VFCS) and Open$_{High}$
162 (AIFS) libraries. All six enzymes yielded crystals under similar conditions (Supplementary Table
163 5), which could only be obtained in the presence of maleate (Supplementary Figure 1c), an
164 inhibitor that stabilizes the closed conformation when bound by the enzyme (*23*). To obtain
165 structures in the absence of maleate, we applied a rigorous crystal soaking method to serially dilute
166 and extract the inhibitor from the crystallized enzymes (Methods). We collected X-ray diffraction
167 data at room-temperature (278 K) for all variants with the exception of AIFS, which could only be
168 measured at cryogenic temperature (100 K) because we could only obtain small crystals that were
169 not robust to radiation damage at non-cryogenic temperatures. We applied statistical criteria
170 (Supplementary Table 6) to assign high-resolution cut-offs of 1.37–2.31 Å for our data sets, and
171 all structures were subsequently determined by molecular replacement in space group P63 with an

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

enzyme homodimer in the asymmetric unit. In the inhibitor-bound state, all structures were closed as expected (Supplementary Figure 7) due to electrostatic interactions between maleate and the side chains of Arg280 and Arg374 (Supplementary Figure 8). Upon soaking crystals of the WT enzyme to remove the bound inhibitor, we observed that one subunit (chain A) within the enzyme homodimer was in the open conformation (Figure 3), confirming that the soaking procedure was able to remove the bound maleate, and that the crystal lattice could accommodate the domain rotation required for opening and closing of this subunit. In the maleate-free structures, a sulfate ion and one or more ordered water molecules occupy the inhibitor binding site (Supplementary Figure 9), as was previously observed in WT (*24*) and HEX (*14*) structures at cryogenic temperatures.

Superposition of bound and unbound structures for each variant confirmed that closed library mutants VFIT and VFIY remained in the closed conformation in the inhibitor-free form, similar to HEX, while $Open_{Low}$ variant VFCS and $Open_{High}$ variant AIFS adopted the open conformation, similar to WT (Figure 3). AIFS is unique in that the electron density is especially weak in the regions corresponding to the helix formed by Pro12–Leu19 and the loop connecting moving and fixed domains (Leu31–Thr43), even at cryogenic temperatures, demonstrating that these structural segments are disordered when the enzyme adopts the open conformation (Supplementary Figure 10). This result suggests that the four mutations of AIFS, three of which are found within the Leu31–Thr43 loop, contribute to destabilize the open conformation, consistent with the calculated E_{open} value of this variant (Table 2) being >7 kcal mol^{-1} higher than that of the other variants that favour the open state (WT and VFCS). Interestingly, the amplitude of the open/closed conformational transition that occurs in open variants upon maleate binding (Figure 3) correlated with their computed ΔE values (VFCS < WT < AIFS), suggesting that this

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

metric can be used to fine-tune this enzyme conformational landscape. Furthermore, DynDom analyses of all variants confirmed that only WT, VFCS, and AIFS undergo hinge motion upon maleate binding (Supplementary Table 7), which rotates the moving domain relative to the fixed domain by 4.6, 2.6, and 5.9 degrees, respectively.

NMR analysis of conformational landscapes

Having demonstrated crystallographically that our designed variants adopted the target conformation in the absence of ligand, we turned to NMR spectroscopy to gain insights into their conformational equilibria in solution, and compared results against those obtained for WT and HEX. We first measured ^{1}H-^{15}N HSQC spectra for WT in the presence and absence of L-aspartate (Supplementary Figure 11a). As expected for a protein of this size, the spectrum consisted of a large number of peaks that were broad with a high degree of overlap. It was nonetheless possible to assign the unique chemical shifts of peaks from the indole NH group for three native Trp residues by comparison of these spectra with those acquired with single Trp mutants (Supplementary Figure 11b,c). This allowed assignment of a peak that was significantly broadened in spectra of both HEX and Closed library mutant VFIY to the indole NH from Trp307 (Supplementary Figure 11d), whose side chain is closest to the designed hinge residues. Moreover, the Trp307 indole peak was no longer detectable when the L-aspartate substrate was present for both WT and mutant enzymes, most likely being broadened beyond detection. Since all conditions that favour the closed state (i.e., HEX and Closed library mutations and/or L-aspartate binding) broaden the Trp307 indole resonance, this exchange appears to be associated with the closed state, potentially due to conformational dynamics around the hinge.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In order to characterize the thermodynamics of the exchange processes around the hinge region, we labelled AAT variants at a single site with ^{19}F using site-specific incorporation of the noncanonical amino acid 4-trifluoromethyl-L-phenylalanine (*25*). Phe217 was chosen as the incorporation site since it is proximal to hinge residues but is not in direct contact with the substrate (Figure 4a). The ^{19}F spectrum of WT at 278 K showed a single peak centred at approximately –60 ppm that shifts downfield as temperature is increased (Figure 4b), similar to what is observed for free 4-trifluoromethyl-L-phenylalanine in solution (Supplementary Figure 12). By contrast, the HEX spectrum at 278 K showed a large broad peak centred at around –61.6 ppm, with another peak of substantially lower intensity also appearing at a similar shift to that observed in the WT spectrum (–60.4 ppm). The relative intensity of these 2 peaks changed as the temperature was increased, with the low intensity peak increasing as the major peak decreased. This is characteristic of two-state exchange, with the equilibrium between the two states being shifted by the temperature change. Given that the crystal structure of HEX in its ligand-free form showed a closed conformation (Figure 3), it is likely that the major peak reflects a local chemical environment created by the closed state, with a small population in an open state similar to that seen in the WT spectrum.

Using peak deconvolution and integration, it was possible to calculate relative populations for each species for variants undergoing the observed two-state exchange, along with the free energy difference between states (Supplementary Figure 13, Supplementary Table 8). We calculated ΔG at 278 K for HEX to be –1.72 kcal mol^{-1} (Table 2), a small difference that would be compatible with the interconversion between these two states to be part of the catalytic cycle (*26*). These peak volumes could also be calculated over the entire temperature range tested, giving rise to a van't Hoff relationship with a small degree of curvature (Supplementary Figure 14).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Deviations from linearity can occur when there are differences in heat capacity between the two states, as would be expected for a process involving a change in conformational states over the temperature range tested (*27*). By contrast, the presence of a single peak in WT spectra suggests that the closed state is not significantly populated under these conditions, as expected from its open-state crystal structures obtained under both cryogenic (*12, 24*) and room-temperature conditions.

We next analyzed two mutants from the Closed library (VFIT and VFIY). ^{19}F NMR spectra at 278 K for both mutants showed a narrow peak centred at approximately –60.2 ppm that is similar to that of WT (Figure 4). Spectra of these variants also showed another broad peak centred at approximately –61.2 ppm that is similar to the HEX peak characteristic of the closed conformation. The relative intensity of the two peaks showed similar temperature dependence, with an increase in the relative intensity of the WT-like peak as temperature was increased to 308 K. Peak deconvolution and integration (Supplementary Figure 13–14, Supplementary Table 8) were used to evaluate the population of the two states, and confirmed that at low temperatures both VFIT and VFIY favour the state resembling that adopted by the closed HEX mutant ($T < 288$ K or 283 K, respectively). However, unlike HEX, at higher temperatures the alternate conformation with the WT-like peak becomes the favoured state (Table 2). Interestingly, ΔE values for these mutants were smaller in magnitude than that calculated for HEX, supporting the predictive nature of the calculated energy differences between open and closed states for these sequences.

To determine if these differences in the temperature dependence of exchange could be observed in the crystal state, we also solved the maleate-free structures of WT, HEX, and VFIT at 303 K (Supplementary Table 9, Supplementary Figure 9), and calculated isomorphous difference density maps by subtracting electron density at 278 K (Supplementary Figure 15). Comparing

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

263 VFIT data obtained at 278 K and 303 K results in substantial difference density throughout chain
264 A, the chain that opens when WT crystals are soaked to remove maleate. By contrast, similar
265 comparisons for WT and HEX showed relatively little difference density. This analysis confirms
266 that VFIT undergoes larger local conformational changes than either HEX or WT when the
267 temperature of the crystal is increased to 303 K, in agreement with our van't Hoff analysis of NMR
268 data (Table 2, Supplementary Table 8). The agreement between temperature-dependent X-ray
269 crystallography and NMR data provides strong evidence that the conformational exchange
270 detected in the NMR experiments reflects a dynamic equilibrium between closed and open
271 conformations, with mutations that favour the closed conformation also increasing selectivity
272 toward L-phenylalanine.

273 We next analyzed two mutants designed to favour the open conformation in their ligand-
274 free forms. $Open_{Low}$ mutant VFCS showed ^{19}F NMR spectra that were very similar to those of WT
275 within the tested temperature range (Figure 4b, Supplementary Figure 13), consistent with its
276 preference for the open conformation (Figure 3, Table 2). However, $Open_{High}$ mutant AIFS gave
277 rise to spectra that were distinct from those of all other variants (Figure 4b), but could be
278 deconvoluted to two exchanging peaks (Supplementary Figure 13), to allow estimation of
279 populations (Supplementary Table 8, Supplementary Figure 14). We postulate that those peaks
280 correspond to alternate open conformations distinct from the one sampled by the other variants.
281 This hypothesis is supported by our observations that ligand-free AIFS is open at low temperature
282 (Figure 3) but contains disordered segments around the Pro12–Leu19 helix and the loop containing
283 three of the four designed positions (Supplementary Figure 10), which are located close to the
284 Phe217 position where the ^{19}F label was introduced. The observed heterogeneity could therefore
285 correspond to a mixture of these alternate open conformations. Additional support that the

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

286 exchange detected in AIFS differed from that of HEX and closed mutants was provided by the
287 van't Hoff analysis, which showed no curvature for AIFS unlike for the closed mutants
288 (Supplementary Figure 14), suggesting that the $Open_{High}$ variant does not undergo the open/closed
289 conformational transition within this temperature range.
290
291 **Discussion**
292 Here, we successfully remodeled the conformational landscape of an enzyme via targeted
293 alterations to the equilibrium between two distinct conformational states related by a hinge-
294 bending motion. The resulting equilibrium shift promoted activity towards a non-native substrate,
295 leading to a selectivity switch of up to 1900-fold. As many enzymes undergo hinge-mediated
296 domain motions during their catalytic cycles (*28*), the multistate design approach presented here
297 should be straightforward to implement for such enzymes. Given the ability of our design
298 procedure to distinguish between closed and open states whose free energy difference is on the
299 order of a single hydrogen bond, our approach could, in principle, also be applied to preferentially
300 stabilize catalytically competent substates involving more subtle structural changes, such as
301 backbone carbonyl flips (*29*) or side-chain rotations (*30*). This methodology could therefore help
302 to tailor catalytic efficiency or substrate selectivity by mimicking, *in silico,* the processes of
303 evolution that harness altered conformational equilibria to tune function (*7, 31, 32*).
304 The predictive capacity of our multistate design framework could only be achieved by
305 evaluating the energy of sequences on multiple conformational states. For example, the VFCS
306 variant that prefers the open state is predicted to be more stable on the closed state than Closed
307 library mutants VFIT and VFIY (Table 2), and more stable on the Open state than AIFS even
308 though it is less open than this variant (Figure 3). Furthermore, there were no obvious trends in

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

designed mutations that could explain their effect on the conformational landscape, as none of these introduced bulkier or smaller amino acids at all or specific residue positions to cause or alleviate steric clashes in the open or closed conformations so as to shift the equilibrium towards one of these states, which has been the approach others have used to shift conformational equilibria (*33*). Thus, subtle effects of mutation combinations on the relative stability of each conformational state were likely responsible for the observed preference of mutants for the open or closed conformations.

Our results demonstrate the utility of multistate design, with ΔE values calculated from ensemble energies of open and closed states, for the targeted alteration of subtle conformational equilibria, an approach that represents a useful alternative to heuristic methods that others have used to tune the relative stability of protein conformational states (*34*). Extending this concept, we envision that de novo design of artificial enzymes with native-like catalytic efficiency and selectivity for complex multistep chemical transformations will require a holistic approach where every conformational state and/or substate required to stabilize reaction intermediates and transition states are explicitly modelled, and their relative energies optimized. The multistate design method for conformational landscape remodelling presented here could therefore be incorporated into *de novo* enzyme design pipelines, helping to bridge the gap between the carving of an active site for transition-state stabilization, and the modulation of conformational dynamics required for efficient passage along the reaction coordinate.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. S. J. Benkovic, G. G. Hammes, S. Hammes-Schiffer, Free-energy landscape of enzyme catalysis. *Biochemistry* **47**, 3317-3321 (2008).
2. S. J. Kerns, R. V. Agafonov, Y. 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).
3. D. D. Boehr, D. McElheny, H. J. Dyson, P. E. Wright, The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638-1642 (2006).
4. K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins. *Nature* **450**, 964-972 (2007).
5. E. Campbell, M. Kaltenbach, G. J. Correy, P. D. Carr, B. T. Porebski, E. K. Livingstone, L. Afriat-Jurnou, A. M. Buckle, M. Weik, F. Hollfelder, N. Tokuriki, C. J. Jackson, The role of protein dynamics in the evolution of new enzyme function. *Nature chemical biology* **12**, 944-950 (2016).
6. J. A. Kaczmarski, M. C. Mahawaththa, A. Feintuch, B. E. Clifton, L. A. Adams, D. Goldfarb, G. Otting, C. J. Jackson, Altered conformational sampling along an evolutionary trajectory changes the catalytic activity of an enzyme. *Nature communications* **11**, 5945 (2020).
7. M. A. Maria-Solano, J. Iglesias-Fernandez, S. Osuna, Deciphering the Allosterically Driven Conformational Ensemble in Tryptophan Synthase Evolution. *J Am Chem Soc* **141**, 13049-13056 (2019).
8. N. Tokuriki, D. S. Tawfik, Protein dynamism and evolvability. *Science* **324**, 203-207 (2009).
9. M. A. Maria-Solano, E. Serrano-Hervas, A. Romero-Rivera, J. Iglesias-Fernandez, S. Osuna, Role of conformational dynamics in the evolution of novel enzyme function. *Chem Commun (Camb)* **54**, 6622-6634 (2018).
10. B. E. Clifton, C. J. Jackson, Ancestral Protein Reconstruction Yields Insights into Adaptive Evolution of Binding Specificity in Solute-Binding Proteins. *Cell chemical biology* **23**, 236-245 (2016).
11. J. A. Davey, R. A. Chica, Multistate approaches in computational protein design. *Protein Sci* **21**, 1241-1252 (2012).
12. A. Okamoto, T. Higuchi, K. Hirotsu, S. Kuramitsu, H. Kagamiyama, X-ray crystallographic study of pyridoxal 5'-phosphate-type aspartate aminotransferases from Escherichia coli in open and closed form. *J Biochem* **116**, 95-107 (1994).
13. J. J. Onuffer, J. F. Kirsch, Redesign of the substrate specificity of Escherichia coli aspartate aminotransferase to that of Escherichia coli tyrosine aminotransferase by homology modeling and site-directed mutagenesis. *Protein Sci* **4**, 1750-1757 (1995).
14. V. N. Malashkevich, J. J. Onuffer, J. F. Kirsch, J. N. Jansonius, Alternating arginine-modulated substrate specificity in an engineered tyrosine aminotransferase. *Nat Struct Biol* **2**, 548-553 (1995).
15. M. M. Islam, M. Goto, I. Miyahara, H. Ikushiro, K. Hirotsu, H. Hayashi, Binding of C5-dicarboxylic substrate to aspartate aminotransferase: implications for the conformational change at the transaldimination step. *Biochemistry* **44**, 8218-8229 (2005).
16. S. Hayward, R. A. Lee, Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *Journal of Molecular Graphics and Modelling* **21**, 181-183 (2002).
17. J. A. Davey, R. A. Chica, Multistate Computational Protein Design with Backbone Ensembles. *Methods Mol Biol* **1529**, 161-179 (2017).
18. J. A. Davey, R. A. Chica, Improving the accuracy of protein stability predictions with multistate design using a variety of backbone ensembles. *Proteins* **82**, 771-784 (2014).
19. B. D. Allen, A. Nisthal, S. L. Mayo, Experimental library screening demonstrates the successful application of computational protein design to large structural ensembles. *Proc Natl Acad Sci U S A* **107**, 19838-19843 (2010).
20. R. A. Chica, M. M. Moore, B. D. Allen, S. L. Mayo, Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc Natl Acad Sci U S A* **107**, 20257-20262 (2010).
21. J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, T. Alber, Hidden alternative structures of proline isomerase essential for catalysis. *Nature* **462**, 669-673 (2009).
22. D. A. Keedy, H. van den Bedem, D. A. Sivak, G. A. Petsko, D. Ringe, M. A. Wilson, J. S. Fraser, Crystal cryocooling distorts conformational heterogeneity in a model Michaelis complex of DHFR. *Structure* **22**, 899-910 (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

23. I. Miyahara, K. Hirotsu, H. Hayashi, H. Kagamiyama, X-ray crystallographic study of pyridoxamine 5'-phosphate-type aspartate aminotransferases from Escherichia coli in three forms. *J Biochem* **116**, 1001-1012 (1994).
24. J. Jäger, M. Moser, U. Sauder, J. N. Jansonius, Crystal Structures of Escherichia coli Aspartate Aminotransferase in Two Conformations: Comparison of an Unliganded Open and Two Liganded Closed Forms. *Journal of Molecular Biology* **239**, 285-305 (1994).
25. J. T. Hammill, S. Miyake-Stoner, J. L. Hazen, J. C. Jackson, R. A. Mehl, Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat Protoc* **2**, 2601-2607 (2007).
26. S. Kuramitsu, K. Hiromi, H. Hayashi, Y. Morino, H. Kagamiyama, Pre-steady-state kinetics of Escherichia coli aspartate aminotransferase catalyzed reactions and thermodynamic aspects of its substrate specificity. *Biochemistry* **29**, 5469-5476 (1990).
27. G. Sharma, E. A. First, Thermodynamic analysis reveals a temperature-dependent change in the catalytic mechanism of bacillus stearothermophilus tyrosyl-tRNA synthetase. *J Biol Chem* **284**, 4179-4190 (2009).
28. P. M. Khade, A. Kumar, R. L. Jernigan, Characterizing and Predicting Protein Hinges for Mechanistic Insight. *Journal of Molecular Biology* **432**, 508-522 (2020).
29. R. Otten, R. A. P. Padua, H. A. Bunzel, V. Nguyen, W. Pitsawong, M. Patterson, S. Sui, S. L. Perry, A. E. Cohen, D. Hilvert, D. Kern, How directed evolution reshapes the energy landscape in an enzyme to boost catalysis. *Science* **370**, 1442-1446 (2020).
30. N. S. Hong, D. Petrovic, R. Lee, G. Gryn'ova, M. Purg, J. Saunders, P. Bauer, P. D. Carr, C. Y. Lin, P. D. Mabbitt, W. Zhang, T. Altamore, C. Easton, M. L. Coote, S. C. L. Kamerlin, C. J. Jackson, The evolution of multiple active site configurations in a designed enzyme. *Nature communications* **9**, 3900 (2018).
31. B. E. Clifton, J. A. Kaczmarski, P. D. Carr, M. L. Gerth, N. Tokuriki, C. J. Jackson, Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein. *Nature chemical biology* **14**, 542-547 (2018).
32. M. Kaltenbach, J. R. Burke, M. Dindo, A. Pabis, F. S. Munsberg, A. Rabin, S. C. L. Kamerlin, J. P. Noel, D. S. Tawfik, Evolution of chalcone isomerase from a noncatalytic ancestor. *Nature chemical biology* **14**, 548-555 (2018).
33. J. S. Marvin, H. W. Hellinga, Manipulation of ligand binding affinity by exploitation of conformational coupling. *Nature Structural Biology* **8**, 795-798 (2001).
34. K. Y. Wei, D. Moschidi, M. J. Bick, S. Nerli, A. C. McShan, L. P. Carter, P.-S. Huang, D. A. Fletcher, N. G. Sgourakis, S. E. Boyken, D. Baker, Computational design of closely related proteins that adopt two well-defined but structurally divergent folds. *Proceedings of the National Academy of Sciences* **117**, 7208-7215 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Apparent kinetic parameters of *E. coli* AAT and its mutants for transamination of various amino-acid donors with α-ketoglutarate as acceptor

Enzyme[a]	Mutations				L-Aspartate[b]			L-Phenylalanine[b]			Selectivity[c]
	V35	K37	T43	N64	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})	
Controls											
WT	-	-	-	-	0.21 ± 0.03	7.8 ± 0.3	37000 ± 5000	N.D.[e]	N.D.[e]	400 ± 100	0.01
HEX[d]	L	Y	I	L	0.059 ± 0.007	1.32 ± 0.03	22000 ± 3000	0.27 ± 0.03	9.0 ± 0.2	33000 ± 4000	1.5
Closed Library											
IYIT	I	Y	I	T	0.014 ± 0.002	0.155 ± 0.005	11000 ± 2000	0.34 ± 0.03	10.1 ± 0.2	30000 ± 3000	2.7
VFIT	-	F	I	T	0.027 ± 0.003	0.90 ± 0.01	33000 ± 4000	0.90 ± 0.06	22.8 ± 0.4	25000 ± 2000	0.8
VFIY	-	F	I	Y	0.12 ± 0.01	0.263 ± 0.005	2200 ± 200	0.42 ± 0.04	17.3 ± 0.4	41000 ± 4000	19
VYIT	-	Y	I	T	0.08 ± 0.01	0.37 ± 0.01	4600 ± 600	1.11 ± 0.07	41.8 ± 0.6	38000 ± 2000	8
VYIY	-	Y	I	Y	0.09 ± 0.02	0.244 ± 0.006	2700 ± 600	0.58 ± 0.02	20.9 ± 0.2	36000 ± 1000	13
$Open_{Low}$ Library											
IFCA	I	F	C	A	0.031 ± 0.004	2.63 ± 0.06	80000 ± 10000	2.3 ± 0.2	47 ± 1	20000 ± 2000	0.25
MFCA	M	F	C	A	0.018 ± 0.003	1.21 ± 0.02	70000 ± 10000	1.08 ± 0.08	35.9 ± 0.6	33000 ± 3000	0.47
VFCA	-	F	C	A	0.050 ± 0.006	4.4 ± 0.1	90000 ± 10000	4.2 ± 0.2	47 ± 1	11200 ± 600	0.12
VFCS	-	F	C	S	0.068 ± 0.007	6.0 ± 0.1	88000 ± 9000	12 ± 1	87 ± 3	7200 ± 700	0.08
$Open_{High}$ Library											
AIFS	A	I	F	S	0.32 ± 0.03	2.42 ± 0.04	7600 ± 700	2.6 ± 0.2	0.54 ± 0.01	210 ± 20	0.03
CIFC	C	I	F	C	0.13 ± 0.01	2.62 ± 0.04	20000 ± 2000	5.9 ± 0.4	3.89 ± 0.08	660 ± 50	0.03
CIFS	C	I	F	S	0.28 ± 0.02	0.65 ± 0.01	2300 ± 200	5.4 ± 0.4	0.257 ± 0.006	48 ± 4	0.02
SIFH	S	I	F	H	0.33 ± 0.03	1.11 ± 0.02	3400 ± 300	3.7 ± 0.3	0.245 ± 0.008	66 ± 6	0.02
SIFS	S	I	F	S	0.35 ± 0.04	1.04 ± 0.02	3000 ± 300	3.9 ± 0.2	0.294 ± 0.006	75 ± 4	0.03

[a] Mutants are named on the basis of the amino-acid identity at the four designed positions. For example, the VFIT mutant from the Closed Library contains Val, Phe, Ile, and Thr residues at positions 35, 37, 43, and 64, respectively. WT and HEX refer to wild-type AAT and the previously published hexamutant (*13*), respectively.

[b] All experiments were performed in triplicate using a single enzyme batch. Errors of regression fitting, which represent the absolute measure of the typical distance that each data point falls from the regression line, are provided. Concentrations of α-ketoglutarate used to determine apparent kinetic parameters of donor substrates are reported on Supplementary Table 3.

[c] Selectivity is defined as (k_{cat}/K_M L-phenylalanine) / (k_{cat}/K_M L-aspartate).

[d] This variant also contains the T104S and N285S active-site mutations.

[e] Individual parameters K_M and k_{cat} could not be determined accurately because saturation was not possible at the maximum substrate concentration tested (40 mM, Supplementary Figure 2), which is the substrate's solubility limit. Catalytic efficiency (k_{cat}/K_M) was therefore calculated from the slope of the linear portion ([S] << K_M) of the Michaelis-Menten model ($v_0 = (k_{cat}/K_M)[E_0][S]$).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Conformational equilibrium of AAT variants

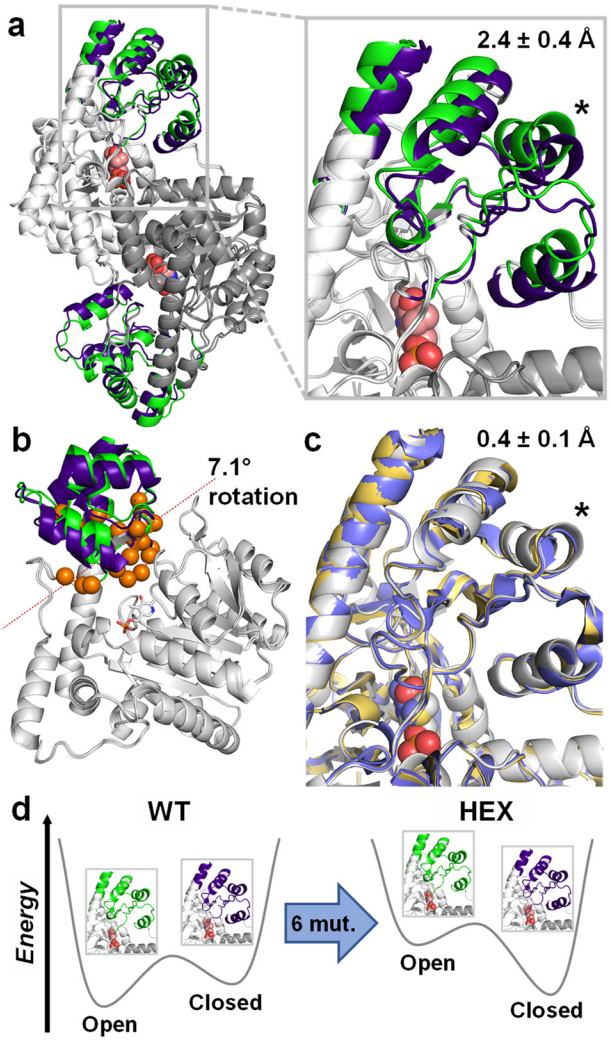
Enzyme	Computational[a]			Experimental[b]				
	E_{closed} (kcal mol^{-1})	E_{open} (kcal mol^{-1})	ΔE (kcal mol^{-1})	ΔH (kcal mol^{-1})	ΔS (kcal mol^{-1} K^{-1})	ΔC_p (kcal mol^{-1} K^{-1})	ΔG, 278 K (kcal mol^{-1})	ΔG, 303 K (kcal mol^{-1})
WT	–358.0	–372.2	14.2	N.D.	N.D.	N.D.	N.D.	N.D.
HEX	–350.2	–273.4	–76.8	–9.9	–0.032	0.972	–1.72	–0.31
VFIT	–366.1	–320.6	–45.5	–10.7	–0.038	0.667	–0.59	0.78
VFIY	–357.8	–300.0	–57.8	–19.6	–0.070	0.369	–0.38	1.59
VFCS	–385.6	–389.7	4.2	N.D.	N.D.	N.D.	N.D.	N.D.
AIFS	–237.2	–364.4	127.2	N.D.	N.D.	N.D.	N.D.	N.D.

[a] Boltzmann-weighted average potential energies (T = 300 K) for the open and closed state ensembles were computed using the Phoenix energy function (Methods). The energy difference reported corresponds to $E_{closed} - E_{open}$.

[b] Differences in enthalpy (ΔH), entropy (ΔS), heat capacity (ΔC_p) and Gibbs free energy (ΔG) are given in the direction of enzyme closing (e.g., $G_{closed} - G_{open}$), and were calculated using a reference temperature of 298 K. N.D. indicates "not determined".

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

441 **Figures**



442

443 **Figure 1. AAT conformational landscape.** (a) *E. coli* AAT is a 90 kDa homodimer that undergoes a conformational
444 change from an open (green, PDB ID: 1ARS) to closed (dark blue, PDB ID: 1ART) state upon substrate binding. This
445 conformational transition involves rotation of a small moving domain (colored) relative to a fixed domain (white and
446 grey for chains A and B, respectively), which causes a 2.4 ± 0.4 Å displacement (mean C_α distance ± s.d.) of the helix
447 formed by residues K355–F365 (indicated by an asterisk). The PLP cofactor bound at the active site is shown as
448 spheres (salmon). (b) Hinge movement analysis of chain A reveals a 7.1-degree rotation of the moving domain relative
449 to the fixed domain along an axis between two planes (dotted line). Hinge-bending residues and PLP are shown as
450 orange spheres and white sticks, respectively. (c) Superposition of HEX structures in the absence (yellow, PDB ID:
451 1AHE) and presence (blue, PDB ID: 1AHY) of bound inhibitor with that of the WT closed state (white, PDB ID:
452 1ART) show that this mutant is closed in both cases. C_α displacement (mean ± s.d.) of residues K355–F365 (asterisk)
453 is indicated. (d) These results demonstrate that the six mutations (mut.) of HEX remodel its conformational landscape
454 to favour the closed conformation.

455

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

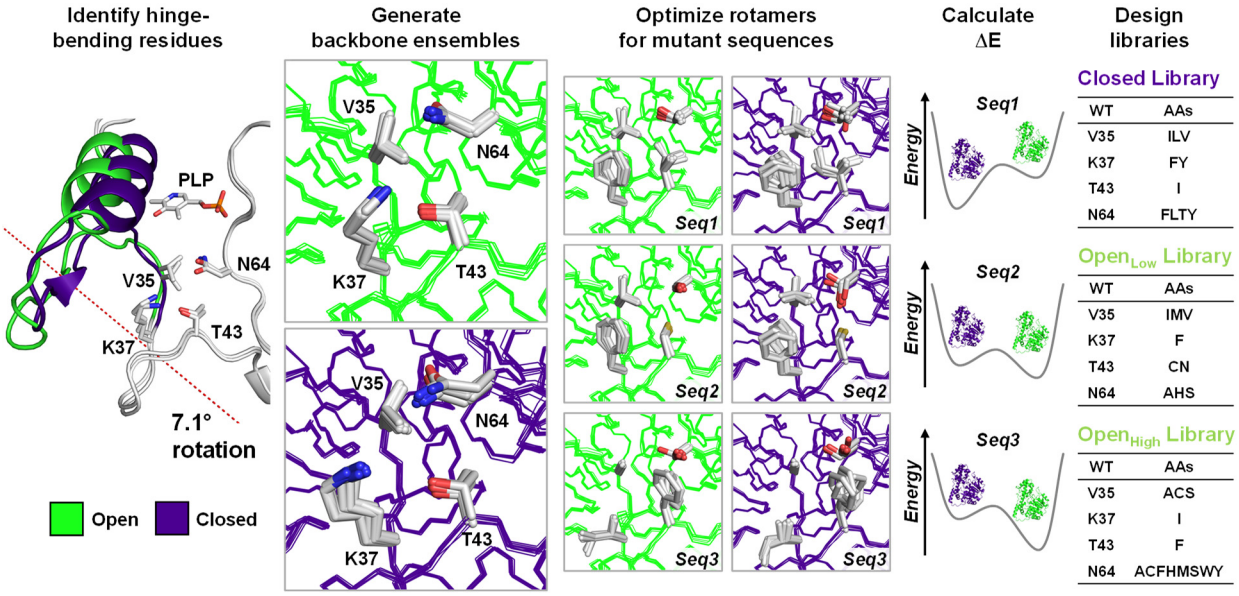


Figure 2. Computational remodelling of AAT conformational landscape by multistate design. To remodel the AAT conformational landscape, we followed a 5-step process: (1) identification of hinge-bending residues involved in transition between open (green) and closed (dark blue) conformational states; (2) generation of structural ensembles approximating backbone flexibility to model open and closed states; (3) optimization of rotamers for mutant sequences on both open- and closed-state ensembles; (4) calculation of energy differences between conformational states ($\Delta E = E_{closed} - E_{open}$) to predict equilibrium of each mutant, and (5) combinatorial library design using ΔE values to generate Closed, $Open_{Low}$ and $Open_{High}$ libraries for experimental testing.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

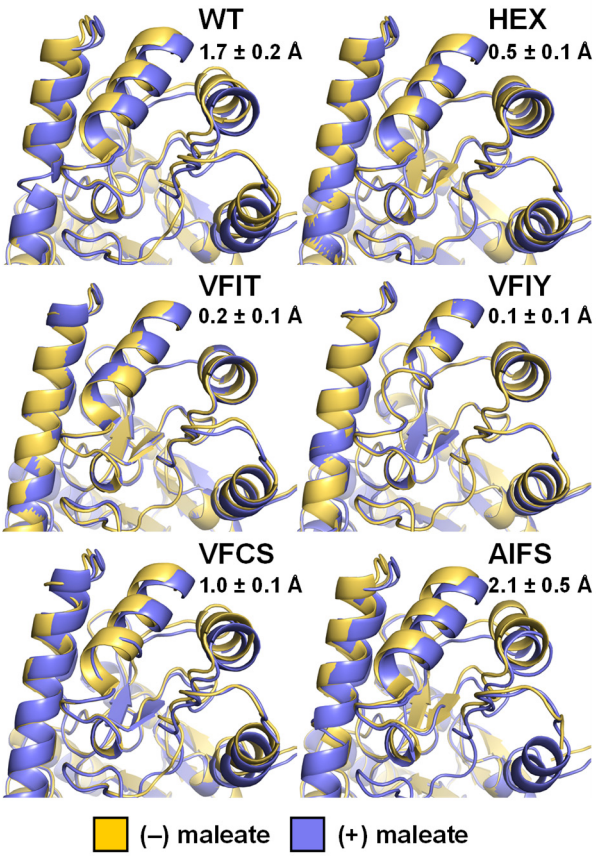


Figure 3. Crystal structures. Overlay of AAT structures (Chain A) in the presence and absence of maleate show transition from open to closed states upon inhibitor binding for WT, VFCS, and AIFS, but not for HEX, VFIT, and VFIY, which already adopt the closed conformation in the absence of bound inhibitor. Average displacements of helix formed by residues K355–F365 upon maleate binding are reported as the average pairwise distance of corresponding C_α atoms for the 11 residues comprising this helix (mean ± s.d.).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.16.508321; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

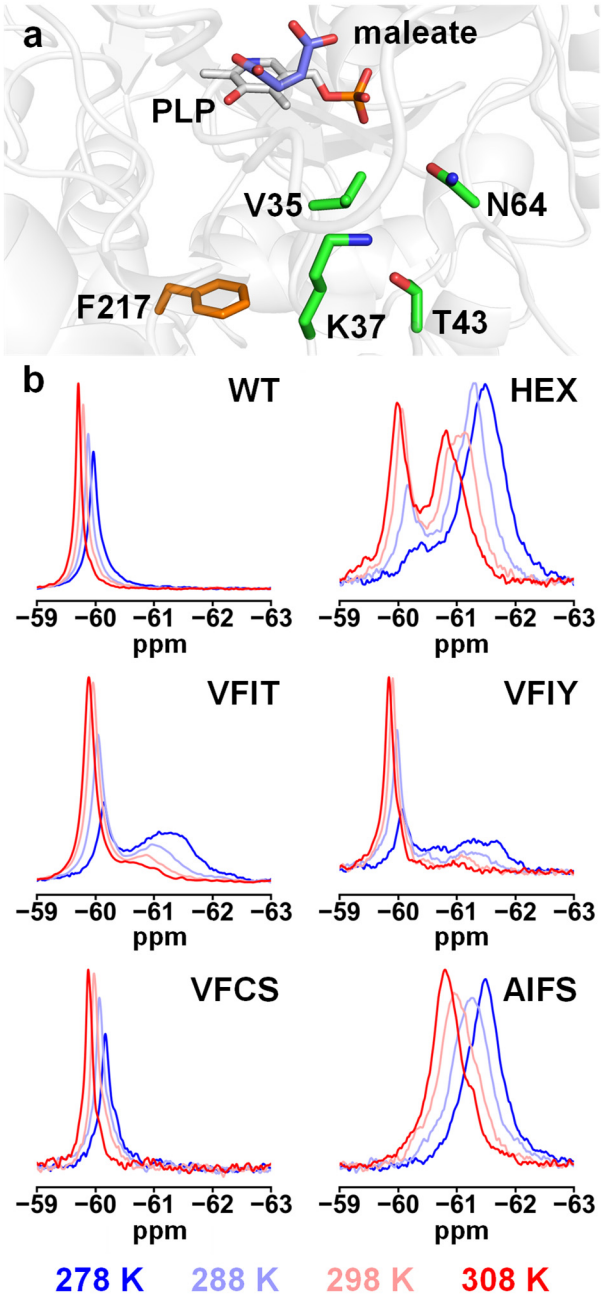


Figure 4. Conformational landscape analysis by NMR. (a) To evaluate conformational equilibrium of AAT variants, we introduced the fluorinated amino acid 4-trifluoromethyl-L-phenylalanine at position F217 (orange), which is located closer to hinge-bending and designed residues (green) than to the bound maleate inhibitor (blue). Crystal structure shown is that of wild-type (WT) AAT at 278 K (PDB ID: 8E9K). (b) ^{19}F NMR spectra of AAT variants in the absence of ligand show dynamic equilibrium between 278 K and 308 K for HEX, VFIT, VFIY, and AIFS, confirming that these proteins are undergoing exchange. This is not the case for WT and $Open_{Low}$ library mutant VFCS, who both adopt predominantly the open conformation within this temperature range. For HEX and Closed library mutants VFIT and VFIY, the open conformation is enriched as temperature increases. For AIFS, spectra suggest that this $Open_{High}$ library mutant samples conformations distinct from those sampled by the other variants.