# Ground-State Destabilization Controls the Selectivity of a Cofactor-Free Decarboxylase

Michal Biler,[a] Anna K. Schweiger,[b] Robert Kourist[b,*] and Shina Caroline Lynn Kamerlin[a,*]

[a] Department of Chemistry - BMC, Uppsala University, BMC Box 576- S-751 23 Uppsala, Sweden

[b] Institute of Molecular Biotechnology, Graz University of Technology, NAWI Graz, BioTechMed, Petersgasse 14, 8010 Graz, Austria.

Corresponding author email addresses: kourist@tugraz.at, lynn.kamerlin@kemi.uu.se

# Abstract

Bacterial arylmalonate decarboxylase (AMDase) and evolved variants have become a valuable tool with which to access both enantiomers of a broad range of chiral arylaliphatic acids with high optical purity. Yet, the molecular principles responsible for the substrate scope, activity and selectivity of this enzyme are only poorly understood to this day, greatly hampering the predictability and design of improved enzyme variants for specific applications. In this work, empirical valence bond simulations were performed on wild-type AMDase and variants thereof, to obtain a better understanding of the underlying molecular processes determining reaction outcome. Our results clearly reproduce the experimentally observed substrate scope, and support a mechanism driven by ground-state destabilization of the carboxylate group being cleaved by the enzyme. In addition, our results indicate that, in the case of the non-converted or poorly-converted substrates studied in this work, increased solvent exposure of the active site upon binding of these substrates can disturb the vulnerable network of interactions responsible for facilitating the AMDase-catalyzed cleavage of $CO_2$. Our results thus allow insight into the tight interaction network determining AMDase selectivity, which in turn provides guidance for the identification of target residues for future enzyme engineering.

# Introduction

Enzymatic catalysis of the formation and breaking of C-C bonds is currently receiving increasing attention.[1] In this context, enzymatic decarboxylation in particular has become highly attractive for the synthesis of optically pure building blocks[2] and the synthesis of alkenes[1,3–5] and alkanes from bio-based precursors.[6] The release of gaseous $CO_2$ renders decarboxylases quasi-irreversible, which has been exploited to drive numerous enzymatic cascade reactions.[7–11] In general, enzymatic decarboxylation can proceed in both an oxidative[4] and a non-oxidative[1] manner. Most non-oxidative decarboxylases employ organic cofactors such as pyridoxyl phosphate, thiamine diphosphate or an N-terminal pyruvyl group as electron sinks to accommodate the intermediary charge after cleavage of carbon dioxide. Interestingly, three cofactor-independent decarboxylases use substrate-assisted catalysis, and thus have the ability to cleave C-C bonds without an internal electron sink. With its highly unusual mechanism, orotidine-5′-phosphate decarboxylase has emerged as a model to study enzymes using ground-state destabilization as a catalytic principle.[12] Among several discussed mechanisms, one uses a so-called "Circe"-effect, in which binding of the phosphate group accommodates the substrate in a binding mode where unfavorable interactions lead to cleavage of a carboxylate group of the substrate. In this vein, the mechanism of phenolic acid decarboxylase (PAD) has been suggested to proceed *via* a quinone methide intermediate formed by protonation of the substrate double bond.[3] This strictly requires hydrogen bonding of the *p*-hydroxy group of the substrate with two tyrosine residues. In both cases, the involvement of functional groups of the substrate strictly limits the substrate scope. For instance, PAD decarboxylates differently substituted cinnamic acid derivatives, but all substrates must bear a *p*-hydroxy group.[1,13]

Bacterial arylmalonate decarboxylase from *Bordetella bronchiseptica* (AMDase, EC 4.1.1.76) was discovered in the early 1990s by the Ohta group, on the basis of a functional screen.[14,15] The enzyme catalyzes the stereospecific decarboxylation of α-disubstituted malonic acids, resulting in pure enantiomers of the respective monoacids (**Scheme 1**). While the acid-catalyzed decarboxylation of prochiral arylmalonates forms racemic product, AMDase catalyzes this reaction in a stereoselective manner. In addition, due to its outstanding stereoselectivity, AMDase has been utilized for the synthesis of a wide range of α-chiral carboxylic acids,[14] including among them several α-arylpropionates with pharmaceutical activity, such as naproxen[16,17] and flurbiprofen,[18–20] α-hydroxy and α-amino acids,[21] and α-heterocyclic[22] and α-alkenyl[23] propionates. Furthermore, combination with metal-catalyzed reduction allowed also for the synthesis of optically pure α-alkyl propionates.[9]
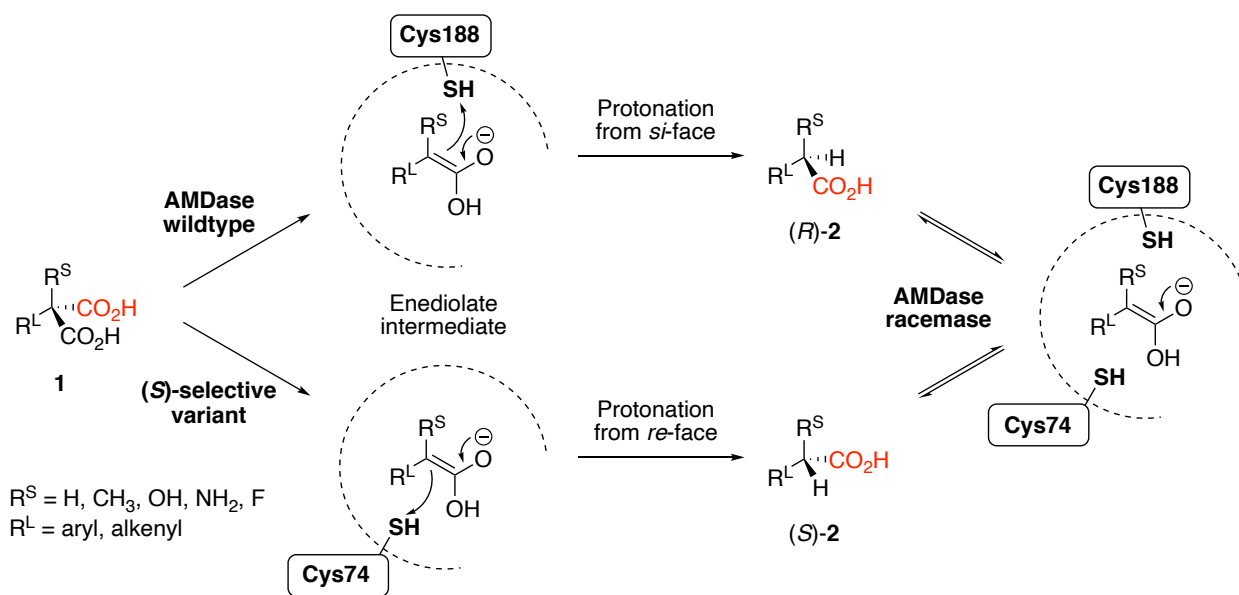


**Scheme 1.** Reaction mechanism of wild-type AMDase and its variants with inverted enantioselectivity (when introducing the G74C substitution, *i.e.* swapping the catalytic cysteine from position 188 to position 74) and promiscuous racemic activity (when introducing/maintaining cysteines at both positions 74 and 188 simultaneously). The pro-(*R*) carboxylate is shown in black, and the pro-(*S*) carboxylate in red.

Initial studies of this enzyme, performed in the absence of a crystal structure, showed that AMDase requires a substituent with a delocalized π-electron system,[15] which can be provided either by an aromatic group or an alkene. The smaller substituent can be a hydrogen atom, a methyl group, a fluorine atom or an amino or hydroxy group; larger substituents such as an ethyl group are not accepted.[2,15] Several bacterial arylmalonate decarboxylases from different bacteria have been isolated.[24–27] All show strict preference for the formation of the (R)-enantiomers. Using both enantiomers of pseudochiral $^{13}$C-labelled malonates, Ohta and coworkers could show that AMDase exclusively cleaves the pro-(R)-carboxylate (**Scheme 1**).[28]

Following from this, the elucidation of several structures of AMDase in both its unliganded and ligand-bound forms[23,29–31] revealed the presence of two binding pockets in the active site. While the first contains several hydrogen-bond donors, the second is mostly composed of hydrophobic residues. Micklefield and coworkers suggested a mechanism that proceeds in two steps: (**1**) Binding of the pro-(S)-carboxylate in the former pocket, stabilized by several H-bonds, pushes the pro-(R)-carboxylate into a configuration with very unfavorable interactions in the hydrophobic pocket, leading to facile cleavage of the C-C bond and the formation of a planar intermediate.[31] (**2**) The donation of a proton by cysteine 188 from one side explains the formation of the pure (R)-products. Ohta and coworkers shifted the position of the catalytic cysteine to the other side, resulting in the formation of pure (S)-enantiomers[32] (**Scheme 1**). While the stereoinversion led the G74C/C188S variant to lose its activity by 20,000-fold, iterative saturation mutagenesis of the hydrophobic pocket restored the activity.[33–35]

Decarboxylation of isotope-labelled malonates confirmed that the (S)-selective variants also cleave the pro-(R)-carboxylate.[33] A variant with both catalytic cysteines present (*i.e.* C188 intact and the artificial C74 introduced by the G74C substitution) has racemizing activity, which allows

4

for study of the second half-reaction of the mechanism.[36,37] Semi-empirical QM/MM calculations showed that the racemization proceeds in a stepwise fashion, through stepwise deprotonation and reprotonation of the planar intermediate shown in **Scheme 1**. The stabilization of the planar intermediate requires a delocalized $\pi$-electron system. The 3.5 kcal mol$^{-1}$ energy barrier to the deprotonation step was lower than that of the initial deprotonation of the cysteines (at 25 kcal mol$^{-1}$), which might explain the drastic pH-dependence of the G74C/C188G variant.

A quantum mechanical model of AMDase[38] confirmed that in the decarboxylation of methylphenyl malonate **1a**, the cleavage of the C-C bond is rate-determining. The authors also argued that the enantioselectivity of the reaction is determined already during substrate binding, as only one binding mode was found to be energetically viable. In the case of a smaller vinyl malonate substrate, it was argued that due to the energetic accessibility of multiple binding modes, both the binding step and the subsequent transition states contribute to the observed selectivity. We note that these calculations were however performed with truncated models of the system, and the results obtained were heavily dependent on model size. In a smaller 81 atom model comprising of only the substrate and residues forming the dioxyanion hole, the authors obtained a small energy difference of only 1.5 kcal mol$^{-1}$ between the cleavage of the pro-($R$) and the pro-($S$) carboxylate groups. However, upon extension of the model to include several other key residues (to a total of 223 atoms), this energy difference increased from 1.5 to 18.3 kcal mol$^{-1}$.

A more recent computational study[39] has studied AMDase with the same two cluster models as presented in ref. [38], but using soft harmonic confining potentials on the boundaries of the system, rather than the fixed atom model of ref. [38]. When using soft harmonic confining potentials, the authors obtained a smaller energy difference of 6.4 kcal mol$^{-1}$ with the larger cluster model, and were able to reproduce the enantioselectivity with the same model. These differences disclose the

complexities when modeling the system using truncated models. A full enzyme model would be useful in order to obtain a better overview of the molecular origins of the observed selectivity. This could be achieved by a complete electrostatic and dynamic treatment within either a QM/MM, empirical valence bond or related framework. In particular, the somewhat non-intuitive results obtained from iterative saturation mutagenesis require a model that takes into account at least the complete first coordination sphere. That is, the hypothetical mechanism for AMDase presented in ref. [38] explains the strict preference of AMDase for cleaving the pro-(R)-carboxylate, the inversion of stereopreference in the G74C/C188X variants, and the racemising activity of the G74C variant. The quantum mechanical model provided an energy profile for the reaction and indicated a plausible binding mode for the substrate. Yet, the predictability of the outcome of amino acid substitutions in the active site is very limited.

Saturation mutagenesis of (R)-selective[18,23] and (S)-selective[34,35] AMDase variants allowed for significant increases in AMDase activity through very conservative substitutions in the active site. So far, it is very difficult to rationalize why exchanges like L19I, L40V, V43I/L and V156L exert such a remarkable effect on AMDase activity. Moreover, the substrate selectivity of AMDase is very difficult to explain: that is, while AMDase catalyzes the decarboxylation of a large series of arylmalonates with a small second substituent (such as H, Me, F), α-ethyl arylmalonates are not converted.[2,15,35] In addition, while the second substituent might be quite large, AMDase does accept p-isobutylphenyl malonate (which would lead to optically pure ibuprofen) only with very poor catalytic efficiency.[35] In both poorly or non-converted substrates, the inductive effect of the alkyl substituents might impede the stabilization of the planar, charged dienoate intermediate, or their size might lead to steric hindrance.
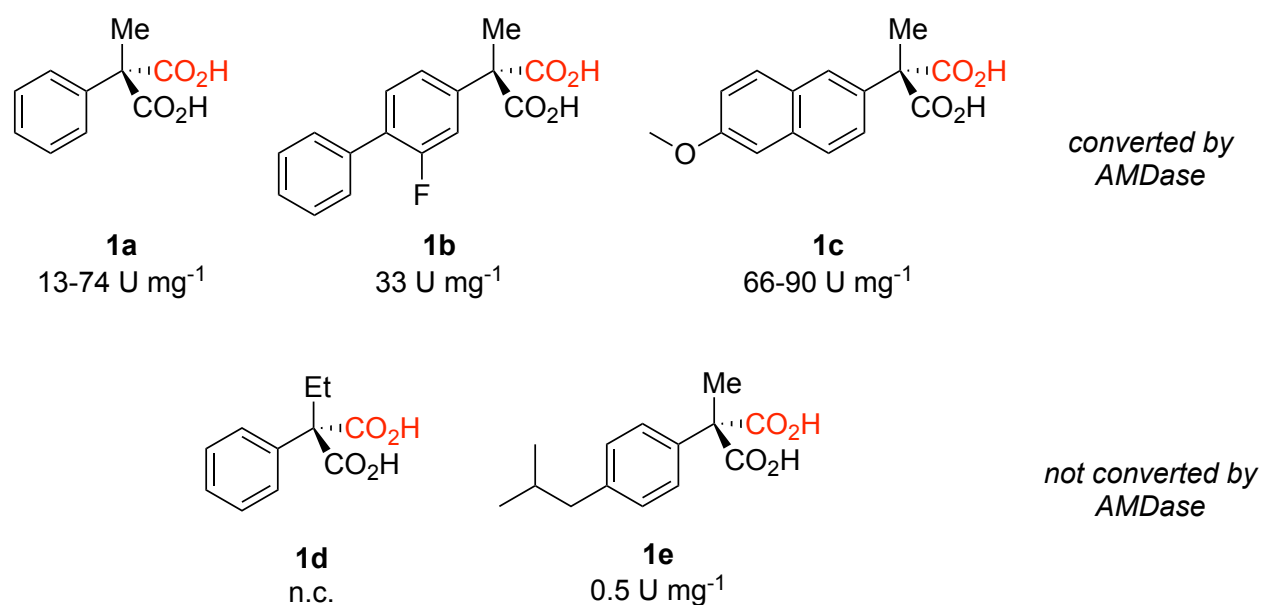
**Scheme 2.** Model compounds used in this study and their experimentally observed acceptance by wild-type AMDase. The pro-($R$) carboxylate is shown in black, and the pro-($S$) carboxylate is shown in red. Shown here are also the specific activities for each compound (U mg$^{-1}$), based on data presented in refs. [15,18,34,35]. We note that **1d** is fully not converted by AMDase, wherease **1e** is converted, but with very low conversion efficiency as shown in **Table 1**.

Obviously, the activity and selectivity of AMDase can be determined by very subtle interactions in the active site. In order to obtain a dynamic model of the decarboxylation, and to obtain insights into the factors determining substrate acceptance and activity of active-site variants, we investigated the rate-determining first half-reaction (the decarboxylation step) of the decarboxylation of substrates shown in **Scheme 2** as catalyzed by wild-type enzyme and substituted variants of AMDase, using the empirical valence bond (EVB) approach.[40] We have considered the cleavage of both the pro-($R$) and pro-($S$) carboxylate groups for each substrate and enzyme variant considered in this work, taking into account multiple potential binding modes of each substrate. Our calculations produce convincing reaction pathways in agreement with experimental observables, pointing to a strongly favored binding mode leading to production of

the (*R*)-enantiomer in wild-type AMDase and to the (*S*)-enantiomer in variants with the catalytic cysteine transferred to the opposite side of the active site. They rationalize the origins of the tremendous catalytic efficiency of this enzyme, as well as of mutational effects on this activity. Finally (and importantly), our EVB simulations are able to both reproduce and provide a rationale for the unusual substrate acceptance of this enzyme, laying the groundwork for future protein engineering effort on this enzyme.

## Methodology

The empirical valence bond (EVB) approach[40] is our methodology of choice in this study, based on the previous successes of both ourselves and others in using this approach to describe enzyme selectivity.[41–44] In this work, we have performed EVB simulations of the decarboxylation of compounds **1a** through **1e** (**Scheme 2**) by wild-type and mutant variants of AMDase, specifically by the CLG-IPL variant (which contains the G74C, V156L, C188G, V43I, A125P, and M159L substitutions) in the case of compounds **1a**, **1b**, **1c**, and **1e,** by the G74C/C188G and G74C/C188A variants in the case of compound **1b**, and by the G74C/C188G variant in the case of compound **1a** and **1c** (these variants were selected based on the availability of experimental data[18,20,23,34,35], with the exception of the G74C/C188A variant for which experimental data is not available). We have provided an in-depth description of our simulation protocol in the **Supporting Information**, and therefore we only provide here a brief summary of our methodology.

Our starting point for simulations of the wild-type enzyme was the structure of wild-type AMDase from *Bordetella bronchiseptica*, in complex with the potential mechanism-based inhibitor benzylphosphonate (PDB ID: 3IP8[23,45]). Due to the lack of structural data on the enzyme variants of interest to this work, all subsequent mutations were manually generated based on the

8

wild-type crystal structure using the Dunbrack and Cohen backbone-dependent rotamer library,[46] as implemented into the PyMOL Molecular Graphics System.[47] The specific side chain rotamers used in the simulations were chosen based on visual inspection for proximity to nearby side chains (to avoid steric clashes), as well as the calculated percentage probability of finding each side chain in a given rotameric state.

Substrates were placed into the active site by docking using AutoDock Vina v. 1.1.2,[48] which resulted in numerous binding poses. These can, in turn, be grouped into two representative highly ranked binding poses (**Figure S1**), the top ranking of which ('**Mode I**') has been the focus of simulations and analysis performed in this work, for reasons described in the *Supplementary Methodology* section of the **Supporting Information**. System setup was performed as described in the **Supporting Information**. Once system setup was complete, all enzyme-substrate complex variants of interest to this work were first equilibrated at the approximate EVB transition state ($\lambda$ = 0.5) for 30 ns, followed by EVB simulations performed on the endpoints of the equilibration runs and propagated from the approximate EVB transition states, using the valence bond states shown in **Figure S2**. Each EVB simulation was performed in 51 individual mapping windows per trajectory of 200 ps length each.

For each system, we performed two independent sets of equilibrations and EVB systems, taking account the cleavage of each of the pro-($R$) and pro-($S$) carboxylate groups per compound (the separate equilibrations were necessary as we are propagating from the transition states). Each set of simulations for the cleavage of each carboxylate group was performed in 30 individual replicates (60 per substrate), leading to total cumulative equilibration and EVB simulation timescales of 1.8 and 0.612 μs per enzyme-substrate complex, respectively.

All EVB simulations were performed using the *Q6* simulation package[49] and the OPLS-AA force field,[50] and all EVB parameters necessary to reproduce our work, as well as a detailed description of the computational methodology and subsequent simulation analysis can be found in the **Supporting Information**.

## Results and Discussion

*Empirical Valence Bond Simulations of AMDase Selectivity Towards Different Compounds*

In this work we study decarboxylation of five π-conjugated compounds (**Scheme 2**) differing in their degree of aromaticity and attached substituents, by both wild-type AMDase and its variants (CLG-IPL, G74C/C188G and G74C/C188A). The choice of the enzyme to study was led by the fact that wild-type AMDase from *B. bronchiseptica* converts compounds **1a-c** in an (*R*)-selective fashion,[15,18] whereas compounds **1d-e** are curiously either not converted at all (**1d**) or only very poorly converted (**1e**).[15,35] The CLG-IPL variant, which carries six amino acid substitutions, was studied here because of its shift to (*S*)-selectivity[18,35] and the doubly substituted variants were studied for their overall low activity levels after introducing the substitutions.[34,35] Moreover, it has been experimentally demonstrated that even a simple interchange to glycine or alanine at position 188 can have a crucial influence on the enzyme kinetics,[32,34] and therefore we considered variants with both glycine and alanine present at position 188.

The AMDase-catalyzed breakdown of compounds **1a** through **1e** to produce optically pure (*R*)- and/or (*S*)-products is a multi-step reaction, initiated through the rate-limiting cleavage of a carboxylic group to yield a sp²-hybridized planar intermediate. This is followed by proton transfer to the intermediate from a nearby amino acid side chain. Critically, it is unclear which carboxylic group of the substrate is preferentially cleaved during this process, as this is not seen in the

stereochemistry of the final product. Based on isotope-labeling experiments it would appear that, in both the wild-type enzyme[28,31] and the (*S*)-selective S36N/G74C/C188S variant of AMDase,[33] there is a strong preference for cleavage of the pro-(*R*) carboxylate group of the substrate. Moreover, it can be argued that multiple binding modes would lead to a mixture of the two enantiomers of the formed α-arylpropionates, while S36N/G74C/C188S and other variants with the motif G74C/C188S produce (*S*)-enantiomers. As described in the *Methodology* section, our docking simulations provided multiple possible binding modes in the active site for each substrate considered in this work, although only **Mode I**-like conformations such as that illustrated in **Figure S1** are catalytically productive.

In **Mode I**, the pro-(*S*) carboxylate of the substrate is closer to Cys188 and is stabilized by hydrogen bonding interactions from the diaoxyanion hole of AMDase, while the pro-(*R*) carboxylate of the substrate is partly located in the hydrophobic pocket. Upon equilibration (**Figure 1**), the substrate rotates slightly such that the pro-(*R*) carboxylate is fully in the hydrophobic pocket. In contrast, in **Mode II**, the substrate is rotated by 180° along the z-axis, such that the pro-(*R*) carboxylate group is instead closer to Cys188, and the pro-(*S*) carboxylate group is located in the hydrophobic pocket, in contrast to what would be expected from experimental studies.[28,31,33] In addition, EVB simulations of enzyme-substrate complexes with the substrate bound in **Mode II** provided very high activation free energies in the range of $24 - 41$ kcal mol$^{-1}$, further suggesting that this is not a catalytically viable binding mode in the AMDase active site, and therefore we have not considered this binding mode further for detailed analysis. Finally, for completeness, in the present work, we independently simulate the cleavage of each of the two carboxylate groups of the substrate, resulting in two different potential decarboxylation routes per

compound, allowing us to obtain computational predictions of the pro-($R$) *vs.* pro-($S$) preference of AMDase towards each compound studied here.
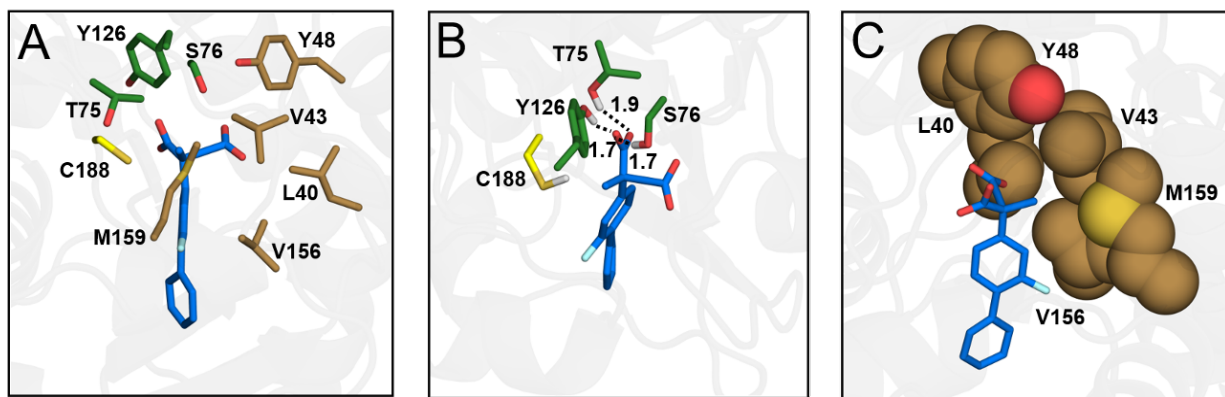


**Figure 1**. An illustration of the catalytically preferred binding mode of compound **1b**, 'Mode I', after molecular dynamics equilibration in preparation for EVB simulations. (**A**) An overview of the AMDase binding pocket. (**B**) A detailed overview of the interactions between the substrate and oxyanion hole. (**C**) A detailed overview of substrate positioning in the hydrophobic pocket. The corresponding amino acids main chains are for simplicity excluded from the figure. As can be seen from this figure, after initial equilibration, the substrate rotates slightly compared to the initial docking pose (**Figure S1**) such that the pro-($S$) carboxylate group of the substrate is stabilized by the dioxyanion hole, and the pro-($R$) carboxylate group points towards the hydrophobic pocket. The initial docking poses for both **Mode I** and **Mode II** prior to equilibration are shown in **Figure S1**. We note that compound **1b** is selected merely for illustration purposes, and similar binding modes were obtained for all compounds studied in this work.

As described in the *Methodology* section, due to the lack of experimental data on the non-enzymatic reaction, our EVB simulations were instead calibrated against DFT calculations of the decarboxylation of compounds **1a** to **1e** in implicit solvent, which means that our EVB calculations are only as quantitatively reliable as the performance of the DFT functional used (although it still allows us to make a qualitative assessment of relative trends between systems). The absolute and relative activation and reaction free energies obtained from our DFT calculations of each

12

compound are shown in **Table S1**, and the corresponding activation and reaction free energies obtained from our EVB calibration, as well as the associated EVB parameters, are shown in **Table S2**. All other bonded and non-bonded parameters necessary to reproduce these values are also provided as **Supporting Information**, and the optimized geometries of key stationary points are shown in **Figure 2**. As expected, in all cases, the DFT "product state" (corresponding to intermediate state in the case of the multi-step enzymatic reaction; **Table S1**) is very close in energy (within 2.7 kcal mol$^{-1}$) to the corresponding transition state. This indicates that this is a labile intermediate along the full reaction coordinate (which would include reprotonation of this intermediate by a suitable proton donor, similarly to the enzymatic reaction).
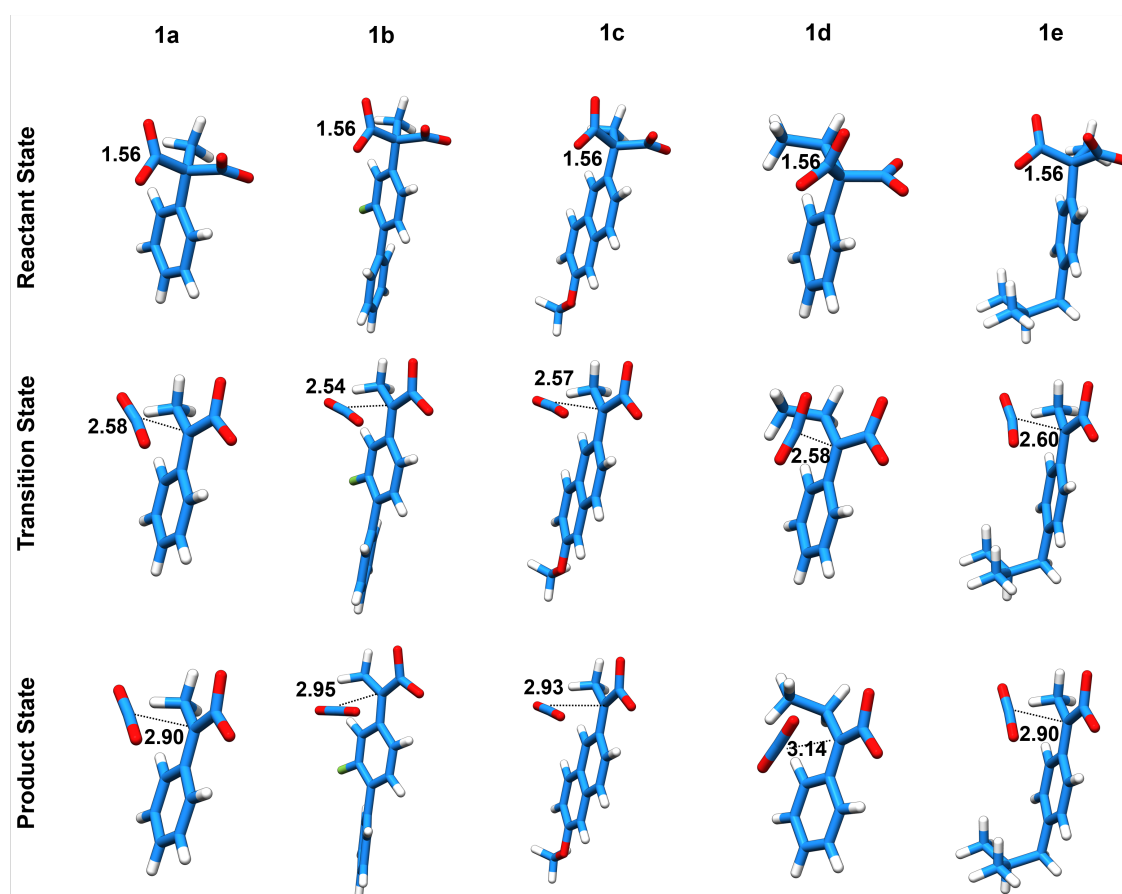


**Figure 2**. Optimized geometries of key stationary points, obtained at the M062X/6-311+G(d,p)/SMD level of theory. The optimized C-C distance for each reacting state is shown on this figure in Å.

Due to the high energy of the intermediate, we expect any subsequent proton transfer step to be facile,[37] and have not modelled it in the present work, as our focus is on the initial decarboxylation step. In addition, we note that the calculated activation free energies, $\Delta G^{\ddagger}$, for the decarboxylation of compounds **1a** and **1c** through **1e** are very similar, and the lower activation free energy calculated for the non-enzymatic decarboxylation of compound **1b** is likely due to the electron-withdrawing effect of the fluorine substituent. Also, curiously, despite the fact that the calculated activation free energies for the non-enzymatic decarboxylation of compounds **1d** and **1e** are only 0.7 and 1.6 kcal mol$^{-1}$ higher than that of compound **1a**, and all three compounds have similar conjugated $\pi$-systems, only compound **1a** is converted by AMDase, and compounds **1d** and **1e** are either not or only barely converted by the enzyme.[15,35] This suggests that it is less likely that AMDase reactivity depends simply on the electronic properties of the different substrates. In addition, we note that all five compounds are decarboxylated through similar transition states, with C-C distances of 2.54 - 2.60 Å at the bond being cleaved during the reaction. The corresponding EVB calibration for the non-enzymatic reaction is summarized in **Table S2**.

Following from this, we performed detailed EVB simulations of the decarboxylation of compounds **1a** through **1e** (**Scheme 2**) by wild-type and variants of AMDase, as described in the *Methodology* section, with the resulting data shown in **Table 1** and **Figure 3**. This table also shows the corresponding selectivities, kinetics ($k_{cat}$) and activation free energies estimated based on experimentally measured activities of each variant towards each compound studied here, where experimental data is available (see **Table 1**).[18,20,23,34,35] We note that we do not provide analysis of other substrate binding modes (see e.*g.* **Mode II** shown in **Figure S1**), due to the high calculated activation free energies when starting from these poses (as described in the *Supplementary Methodology* section), as well as the fact that binding in **Mode II** is in contradiction with

experimental observations.[33] The presence of only one preferred binding mode is also in agreement with prior DFT calculations of ref. [38] suggesting only one binding mode in AMDase for larger and bulky substrates.

From the data shown in **Table 1** and **Figure 3**, it can be seen that our EVB models only show turnover of compounds **1a**, **1b**, **1c** and **1e**, in good agreement with experimental observables,[18,20,23,28,34,35] whereas the activation free energies for compound **1d** is very high for the cleavage of both carboxylic groups, suggesting that this compound is not transformed by the enzyme. In cases where experimental data was available to allow for activation free energies to be estimated by experiment, we typically obtain activation within ~3 kcal mol$^{-1}$ of the experimental value for cleavage of the energetically preferred carboxylate group. We consider this acceptable due to the lack of experimental data on the reference reaction, necessitating our calculations to be calibrated to DFT calculations, thus introducing uncertainity. In addition, our calculations are able, with reasonable quantitative accuracy, to reproduce the experimentally observed loss of activity upon substitution of C188 to either glycine or alanine,[34,35] as observed in the calculated G74C/C188G and G74C/C188A variants, as well as the fact that the substitution to alanine is more detrimental to the activity of the enzyme than the substitution to glycine.[32]

In terms of selectivity, it is important to bear in mind that the preference for the cleavage of the bond to a given carboxylate group in the initial decarboxylation step (**Scheme 1** and **Table 1**), does not translate directly to the final product selectivity. That is, all reactions proceed through a common planar intermediate, with the selectivity being determined in the second step of the reaction upon reprotonation of the planar intermediate. This, in turn, is dependent on the binding pose of the substrate in the Michaelis complex, as illustrated in **Figure 4**. This figure shows three theoretical substrate binding poses to the wild-type AMDase active site.

**Table 1.** Calculated activation ($\Delta G^{\ddagger}$) and reaction free energies ($\Delta G^0$), obtained using the empirical valence bond approach, as well as relevant corresponding experimental observables, for the decarboxylation of compounds **1a** through **1e** by wild-type AMDase and variants.[a]

| System | | Pro-(*R*) | | Pro-(*S*) | | Experimental data | | |
|---|---|---|---|---|---|---|---|---|
| | | $\Delta G^{\ddagger}$ | $\Delta G^0$ | $\Delta G^{\ddagger}$ | $\Delta G^0$ | Selectivity | $k_{cat}$ | $\Delta G^{\ddagger}_{exp}$ |
| **1a** | **WT** | **15.6±0.4** | 14.0±0.6 | 26.6±0.6 | 24.9±0.6 | (*R*) | 279[23] | 14.1[23] |
| | **G74C/C188G** | **23.1±0.6** | 21.4±0.6 | 30.3±0.7 | 28.7±0.6 | (*S*) | 0.004[35] | 21.6[35] |
| | **CLG-IPL** | 26.8±0.7 | 24.6±0.7 | **18.1±0.4** | 17.2±0.4 | (*S*) | 3.8[35] | 17.4[35] |
| **1b** | **WT** | **15.9±0.7** | 12.9±0.9 | 20.2±0.7 | 18.5±0.8 | (*R*) | 15.1,[18] 31[20] | 16.1,[18] 15.4[20] |
| | **G74C/C188G** | **17.9±0.9** | 14.1±0.9 | 23.4±0.7 | 21.7±0.7 | (*S*) | -- | -- |
| | **G74C/C188A** | **20.7±0.9** | 17.2±1.0 | 21.2±0.7 | 18.0±0.9 | (*S*) | -- | -- |
| | **CLG-IPL** | 16.7±0.5 | 14.0±0.7 | **15.8±0.6** | 12.5±0.7 | (*S*) | 23.7,[18] 70[20] | 15.9,[18] 15.0[20] |
| **1c** | **WT** | **18.0±0.3** | 17.1±0.3 | 24.6±0.9 | 21.5±1.0 | (*R*) | 38.7[18] | 15.6[18] |
| | **G74C/C188G** | **22.7±0.4** | 20.6±0.5 | 26.9±0.6 | 25.3±0.7 | (*S*) | 0.077[34] | 19.0[34] |
| | **CLG-IPL** | 22.3±0.7 | 20.3±0.7 | **14.4±0.4** | 13.7±0.5 | (*S*) | 4.3[18] | 16.9[18] |
| **1d** | **WT** | **28.3±0.8** | 25.8±0.9 | 32.9±1.8 | 29.9±1.7 | -- | -- | -- |
| **1e** | **WT** | **18.0±0.4** | 16.3±0.5 | 35.4±0.7 | 33.7±0.7 | (*R*) | 0.23[35] | 19.1[35] |
| | **CLG-IPL** | 34.4±1.7 | 31.7±1.5 | **17.1±0.6** | 15.9±0.6 | (*S*) | 0.56[35] | 18.6[35] |

[a] All calculated values are averages and standard error of the mean over 30 individual EVB trajectories per system, as described in the *Methodology* section, and shown here are data obtained from modeling the decarboxylation of each compound through cleavage of either the pro-(*R*) or pro-(*S*) carboxylate groups. WT denotes the wild-type enzyme. Both experimental and calculated activation and reaction free energies are presented in kcal mol⁻¹. Shown here are also the experimentally observed selectivities for each compound, as well as the corresponding kinetics ($k_{cat}$, s⁻¹) and activation free energies ($\Delta G^{\ddagger}_{exp}$) derived from the experimentally observed activities towards each compound by each variant, as presented in refs. [18,20,23,34,35]. The $k_{cat}$ values were either taken directly from the literature, or were estimated by using the relationship $k_{cat}$ = (specific activity x molecular weight). The calculated activation free energies were obtained from the $k_{cat}$ values using transition state theory at temperature 30 °C (for ref. [18]), 37 °C (for ref. [35]), and 25 °C for the rest. Note that the specific activities were obtained from bar graphs provided in ref. [18] and therefore the

experimental kinetics and energetics are only approximate. '--' denotes that experimental data is not available for a given system.
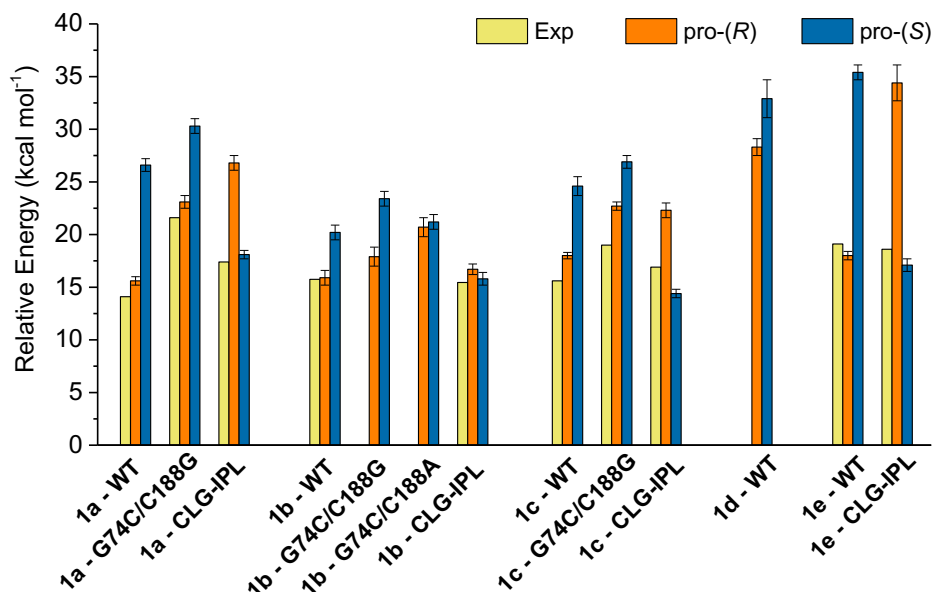


**Figure 3.** Calculated (pro-(R) and pro-(S)) and, where available, experimental (Exp) activation free energies ($\Delta G^{\ddagger}$, kcal mol$^{-1}$) for the decarboxylation of compounds **1a** through **1e** (**Scheme I**) by wild-type (WT) AMDase and its variants. All calculated values are averages and standard error of the mean over 30 individual EVB trajectories per system, as described in the *Supplementary Methodology* section. The raw data for this figure is provided in **Table 1**.

In **Poses A** and **B** (**Figure 4**), one of the pro-(S) and pro-(R) carboxylate groups is stabilized by the dioxyanion hole with cleavage of the other carboxylate group, respectively, resulting in the (R)-product after protonation of the planar intermediate. In **Pose C**, the substrate is rotated compared to **Pose A** such that it is now the pro-(R) carboxylate group that is stabilized by the dioxyanion hole, leading to cleavage of the pro-(S) carboxylate group and the (S)-product after reprotonation. We note that the product selectivity would be expected to be reversed in the G74C/C188X variants, where the cysteine protonating the leaving group is now located on the

opposite face of the planar intermediate, and therefore the same binding pose as in the wild-type enzyme will lead to the opposite product selectivity.
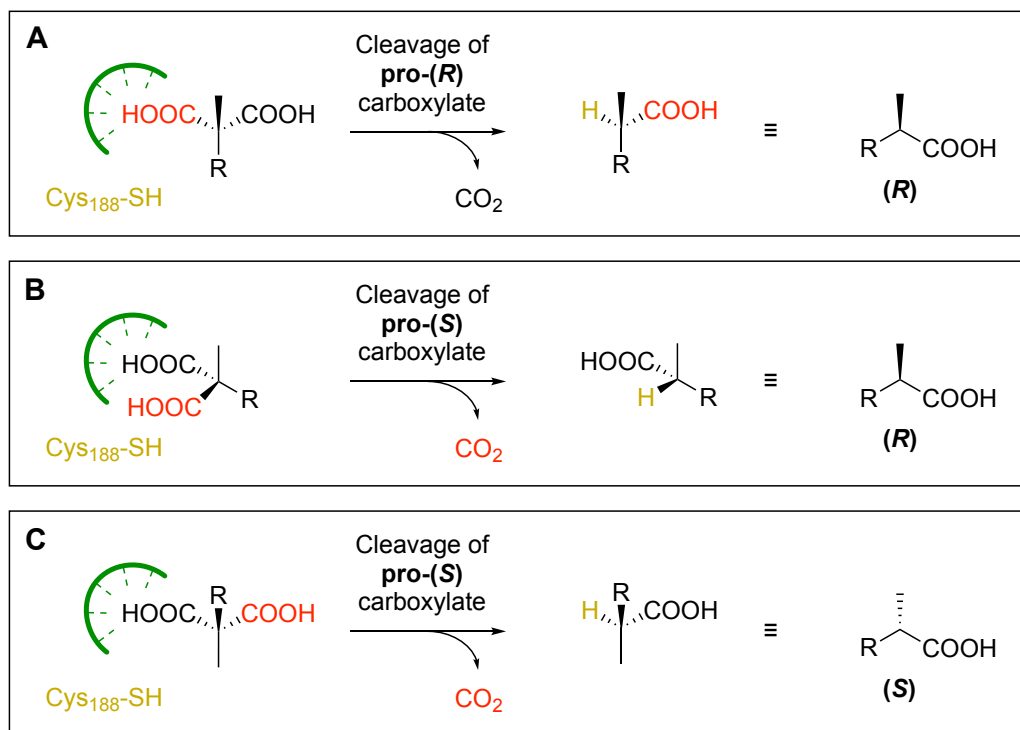


**Figure 4.** Relationship between different possible substrate binding poses in the wild-type AMDase active site and the resulting substrate selectivity. The pro-($R$) carboxylate group is shown in black, and the pro-($S$) carboxylate group is shown in red. The dioxynaion hole is shown in green, with interactions wih the carboxylated group shown schematically by dashed lines. Substituent R can be either an aromatic or alkenyl group. For clarity, we have not shown the hydrophobic pocket on this figure. These binding modes are denoted as **Poses A** to **C** in the discussion in the main text.

As can be seen from **Figures 5**, and **S3** to **S6**, we typically observe Michaelis complexes with the substrate in **Pose A** when we model cleavage of the pro-($R$) carboxylate group, and **Pose B** when we model cleavage of the pro-($S$) carboxylate group, consistent with the expected outcomes illustrated in **Figure 5**. We note that since we initiate our trajectories from the respective transition

states, as described in the *Methodology* section, it is possible for the trajectories to result in different substrate binding poses despite starting from the same initial binding mode. Here, we artificially distinguish between binding "modes" and "poses" simply to distinguish between the binding modes obtained from our docking simulations and the putative binding poses shown in **Figure 5**. However, this is a purely semantic distinction for clarity of the discussion. Similar binding poses are observed in the AMDase variants studied here. We note also that we do not observe **Pose C** in any of our simulations, consistent with the fact that the bulky substrates studied here (**Scheme 2**) are not able to bind to the active site in **Pose C**, however this does not rule out the possibility that this binding mode may hold for smaller substrates.
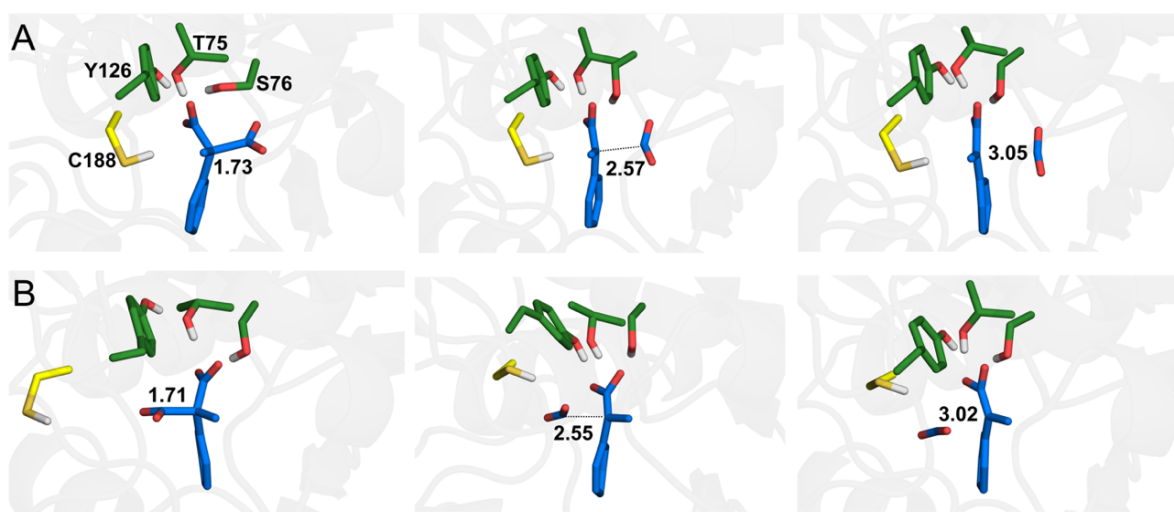


**Figure 5**. Representative structures of the Michaelis complexes, transition states, and intermediate states, for cleavage of (**A**) the pro-(*R*) and (**B**) the pro-(*S*) carboxylate groups of compound **1a** by wild-type AMDase, as obtained from EVB simulations of these reactions. The structures shown here are the centroids of the top ranked cluster obtained from clustering on RMSD, performed as described in the **Supporting Information**. The C-C distances labeled on this figure are averages at each stationary point over all trajectories (see **Table 2**). Corresponding representative structures of key stationary points during simulations of the wild-type AMDase catalyzed decarboxylation of compounds **1b** to **1e** can be found in **Figures S3** to **S6**. The color-coding of key residues follows **Figure 1A**.

For all compounds studied in this work (**Scheme 2** and **Table 1**), we observe preferential cleavage of the pro-(*R*) carboxylate by wild-type AMDase by 1.5 – 11 kcal mol$^{-1}$ depending on substrate, as is to be expected due to the destabilization of the pro-(*R*) carboxylate by unfavorable interactions in the so-called hydrophobic pocket[31] (**Figure 1**). We note that this preference is preserved also in the case of compounds **1d** and **1e**, which are either not experimentally observed to be converted by AMDase (**1d**) or only very poorly converted (**1e**). Based on the schema presented in **Figure 4** and the binding poses observed in **Figures 5** and **S3** to **S6**, this would be expected to lead to the (*R*)-product in all cases. This is in agreement with isotope-labeling experiments performed by two independent groups[28,31,33] on the (*R*)-selective wild-type and the (*S*)-selective variant S36N/G74C/C188S, which have shown that the preferred carboxylate to be cleaved is the pro-(*R*) carboxylate.

In the case of the G74C/C188G and G74C/C188A variants, these variants would be expected to result in the formation of pure (*S*)-enantiomers.[34,35] However, isotope labeling has confirmed that (*S*)-selective AMDase variants can also cleave the pro-(*R*) carboxylate.[33] Our simulations show preferential cleavage of the pro-(*R*) carboxylate group (**Table 1**) with the Michaelis complex bound in **Pose A** of **Figure 4**. However, as the cysteine side chain donating the proton to the intermediate in the second step of the reaction is now on the opposite face of the intermediate, cleavage of the pro-(*R*) carboxylate group would be expected to result in the formation of the (*S*)-product in an inversion of the expectation for the wild-type enzyme, and thus our results are once again consistent with experiment.

Finally, in the case of the CLG-IPL variant (which carries six amino acid substitutions: G74C/M159L/C188G/V43I/A125P/V156L), we observe preferential cleavage of the bond to the pro-(*S*) carboxylate group, although as with the G74C/C188X double mutants, this would still be

expected to lead to the (*S*)-product due to the Michaelis complex being bound in **Pose B** (**Figure 5**). We note that isotope labeling studies such as those performed in refs. [28,31,33] require the challenging synthesis of both pseudochiral malonate isomers in their pure form. Therefore, there are no corresponding isotope-labeling studies of the CLG-IPL variant. However, our modeled (*S*)-selectivity is in good agreement with the observation of the pure (*S*)-enantiomers found in experiments.[18,34] Capturing the selectivity of the CLG-IPL variant correctly is non-trivial from a simulation perspective as, due to the lack of a crystal structure for this enzyme, the six amino substitutions were manually inserted into the system, which may have created artificial perturbations to the structure. Furthermore, one of the six amino acid substitutions is to a proline in a loop (which is the consensus amino acid at this position in the related maleate *cis*/*trans* isomerases[51]), which may introduce conformational rigidity to the structure compared to substitution to other amino acids. This, in turn may influence the entire secondary structure, and that might not be captured in our computationally generated starting structure for our simulations. We noted that in addition to an activity increase in the decarboxylation of flurbiprofen malonate **1b**, this variant showed also remarkable differences in the relative activity towards differently substituted α-aryl propionates.[18] Nevertheless, despite the curious swap to a preference for cleavage of the pro-(*S*) carboxylate, our calculations reproduce both the expected formation of the (*S*)-enantiomer and the experimental activation free energies for the decarboxylation of compounds **1a** through **1c**, and **1e** by the CLG-IPL variant of AMDase with good quantitative accuracy compared to experiment[18,20,34,35] (**Table 1**). Furthermore, the switch in preference for pro-(*S*) carboxylate cleavage for the CLG-IPL variant is reproduced with all compounds studied, thus increasing our confidence in the reliability of our model. In addition, this is a particularly interesting AMDase variant, as each of the hydrophobic residues introduced into this variant (*i.e.*

proline, leucine, isoleucine) have each been shown to be very important determinants of AMDase activity.[18,34,35]

*Exploring the Molecular Origin of the Observed Effects*

While our EVB models for the reactions catalyzed by wild-type AMDase and its variants do not provide perfect quantitative agreement with experiment, due to the uncertainties involved in the energetics of the corresponding non-enzymatic reactions (which are used to calibrate our EVB potentials), they nevertheless appear to provide meaningful qualitative insights into both AMDase substrate preference, as well as selectivity towards cleavage of a given carboxylate group. In particular, our model only shows turnover of compounds **1a** through **1c** and **1e**, in good agreement with experiment, and very high activation barriers for compounds **1d**, in agreement with the fact that decarboxylation of this substrate is not experimentally observed. In addition, experimentally, the activity of AMDase towards substrate **1e** is significantly lower than towards other substrates **1a** through **1c**.[18,20,23,34,35] This could be due to the presence of sterically bulky and/or flexible ethyl and isobutyl groups, which would make compounds **1d** and **1e** challenging to accommodate in the hydrophobic pocket of the AMDase active site, resulting in non-productive binding modes.

In our simulations, we observe larger motions of these substrates (RMSD of up to 1.9 Å compared to the starting structure) compared to substrates such as **1a**, where the substrate RMSD over the course of the simulation is 1 Å or less compared to the starting structure. In addition to this, the ethyl and isobutyl groups of compounds **1d** and **1e**, respectively, are also highly 'floppy' and fluctuate extensively across the simulation time (**Figure 6**), making it more challenging for these compounds to settle into a productive binding mode in the AMDase active site. In conjunction with this, as will be discussed later in this section, in the case of compounds **1d** and **1e** we observe greater solvent penetration of the active site compared to the other compounds

studied in this work, which will counteract the destabilizing effect of the hydrophobic pocket. Finally, the inductive effect of the alkyl substituents would be expected to destabilize the charged intermediate formed upon cleavage of either carboxylate group, thus making the corresponding decarboxylation also energetically unfavorable through a Hammond effect. Indeed, our EVB simulations (**Table 1**) support this at least in the case of compound **1d**, as the reaction free energy for formation of this charged intermediate is significantly higher (by as high as 12.9 kcal mol$^{-1}$, in the case of cleavage of the bond to the pro-(*R*) carboxylate group) for the decarboxylation of this compound compared to the the other compounds studied in this work.
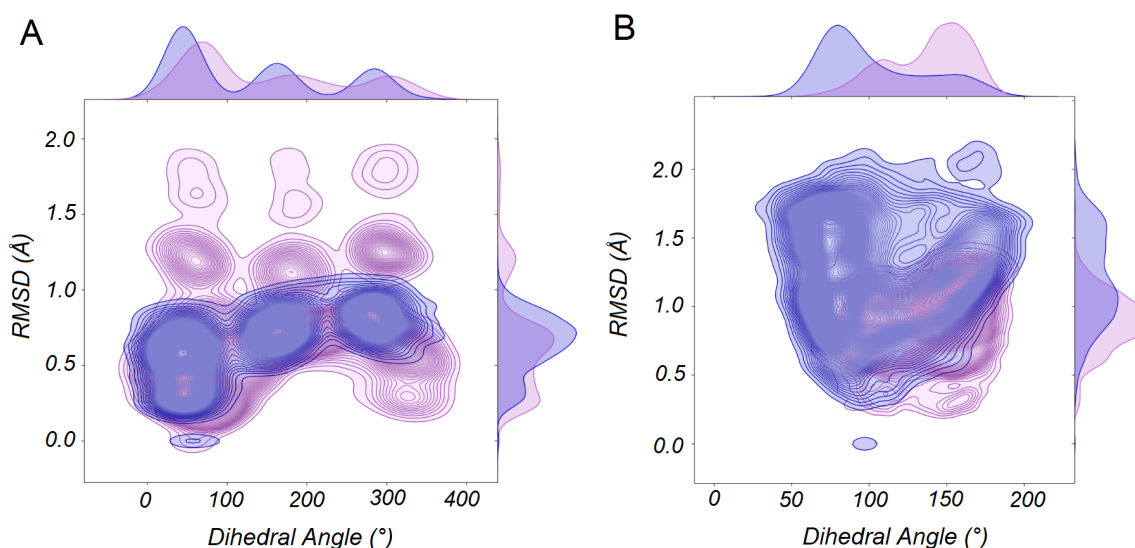


**Figure 6.** The joint distribution of the dihedral angles along the ethyl and isobutyl groups of compounds (**A**) **1d** and (**B**) **1e**, as well as the root mean square deviations of the substrate (RMSD), during 30 ns molecular dynamics simulations of each compound in complex with wild-type AMDase in preparation for subsequent EVB simulations. In the case of the dihedral angles, the C1-C2-C3-C4 and C1-C2-C3-H1 atoms of the ethyl group and of isobutyl group of **1d** and **1e**, respectively, were chosen for analysis in each case (see **Figure S2**). Snapshots were taken every 100 ps of the 30 ns simulations, and thus this analysis was performed on 9000 discrete data points per plot.

**Figure 5, S3** to **S6,** and **Table 2**, respectively, show representative structures of the Michaelis complexes, transition states and intermediate states for the decarboxylation of compounds **1a** through **1e** by wild-type AMDase obtained from our EVB simulations, as well as the average C-C distances for the reactions catalyzed by all AMDase variants studied in this work, and in the corresponding non-enzymatic reactions, as modelled by EVB. In terms of transition state geometries, in the case of the non-enzymatic reactions, all EVB transition states have similar C-C distances of ~2.57Å, similarly to the DFT calculations the EVB potentials were calibrated to (**Table S1**). Similar transition state geometries are observed for the AMDase-catalyzed reaction of compound **1a** and **1e**, irrespectively of which of the two carboxylate groups are being cleaved. In contrast, in the case of compounds **1b**, this distance contracts slightly to up to 2.49 Å in the CLG-IPL variant (cleavage of the pro-(*S*) carboxylate group), whereas in compounds **1c** and **1d**, this distance expands to up to 2.72 Å (compound **1c**, wild-type, cleavage of the pro-(*R*) carboxylate group). Overall, however, these distances are very similar, and indicate that irrespectively of compound, the C-C bond is mostly cleaved at the transition state (Pauling bond orders of ~0.14 to 0.2, using the relationship $r = r_e - 0.6\ln(n)$; where $r_e$ is taken to be 1.56 Å as defined in our EVB force field, as per the EVB parameters provided as **Supporting Information**). Therefore, it is clear that the differences in selectivity between the different substrates are unlikely to be due to intrinsic structural effects due to differences in the intrinsic nature of the transition states involved, as the different transition states are very structurally similar. Rather, they are driven by differences in the initial binding pose of the substrate, as discussed in the previous section.

**Table 2.** Calculated C-C distances at the Michaelis complexes (MC), transition states (TS) and intermediate states (IS) of the decarboxylation of compounds **1a** through **1e**, as catalyzed by wild-type (WT) AMDase and its variants, as well as the corresponding non-enzymatic reaction.[a]

| System | | pro-(*R*) | | | pro-(*S*) | | |
|---|---|---|---|---|---|---|---|
| | | MC | TS | IS | MC | TS | IS |
| **1a** | **Non-Enzymatic** | 1.71 ± 0.01 | 2.71 ± 0.01 | 3.32 ± 0.01 | 1.72 ± 0.01 | 2.69 ± 0.01 | 3.34 ± 0.02 |
| | **WT** | 1.73 ± 0.01 | **2.57 ± 0.01** | 3.05 ± 0.002 | 1.71 ± 0.01 | 2.55 ± 0.01 | 3.02 ± 0.02 |
| | **G74C/C188G** | 1.70 ± 0.01 | **2.63 ± 0.02** | 3.18 ± 0.02 | 1.71 ± 0.01 | 2.62 ± 0.02 | 3.11 ± 0.02 |
| | **CLG-IPL** | 1.71 ± 0.01 | 2.63 ± 0.02 | 3.14 ± 0.02 | 1.72 ± 0.01 | **2.62 ± 0.02** | 3.10 ± 0.02 |
| **1b** | **Non-Enzymatic** | 1.73 ± 0.01 | 2.77 ± 0.01 | 3.32 ± 0.02 | 1.73 ± 0.01 | 2.78 ± 0.01 | 3.34 ± 0.02 |
| | **WT** | 1.74 ± 0.01 | **2.52 ± 0.01** | 3.16 ± 0.02 | 1.71 ± 0.01 | 2.56 ± 0.01 | 3.05 ± 0.02 |
| | **G74C/C188G** | 1.72 ± 0.01 | **2.50 ± 0.01** | 3.15 ± 0.02 | 1.73 ± 0.01 | 2.56 ± 0.01 | 3.09 ± 0.02 |
| | **G74C/C188A** | 1.72 ± 0.01 | **2.51 ± 0.01** | 3.16 ± 0.02 | 1.71 ± 0.01 | 2.53 ± 0.01 | 3.19 ± 0.02 |
| | **CLG-IPL** | 1.74 ± 0.01 | 2.56 ± 0.01 | 3.11 ± 0.02 | 1.72 ± 0.01 | **2.49 ± 0.01** | 3.04 ± 0.01 |
| **1c** | **Non-Enzymatic** | 1.72 ± 0.01 | 2.79 ± 0.02 | 3.31 ± 0.02 | 1.72 ± 0.01 | 2.80 ± 0.02 | 3.28 ± 0.02 |
| | **WT** | 1.73 ± 0.01 | **2.68 ± 0.02** | 3.11 ± 0.02 | 1.73 ± 0.01 | 2.71 ± 0.02 | 3.07 ± 0.02 |
| | **G74C/C188G** | 1.71 ± 0.01 | **2.71 ± 0.02** | 3.12 ± 0.02 | 1.71 ± 0.01 | 2.69 ± 0.02 | 3.09 ± 0.02 |
| | **CLG-IPL** | 1.71 ± 0.01 | 2.63 ± 0.01 | 3.15 ± 0.003 | 1.73 ± 0.01 | **2.65 ± 0.01** | 3.00 ± 0.01 |
| **1d** | **Non-Enzymatic** | 1.71 ± 0.01 | 2.75 ± 0.02 | 3.38 ± 0.03 | 1.72 ± 0.01 | 2.77 ± 0.02 | 3.49 ± 0.03 |
| | **WT** | 1.71 ± 0.01 | **2.69 ± 0.01** | 3.22 ± 0.02 | 1.72 ± 0.01 | 2.69 ± 0.01 | 3.33 ± 0.03 |
| **1e** | **Non-Enzymatic** | 1.71 ± 0.01 | 2.70 ± 0.02 | 3.33 ± 0.02 | 1.72 ± 0.01 | 2.69 ± 0.01 | 3.34 ± 0.02 |
| | **WT** | 1.73 ± 0.01 | **2.59 ± 0.02** | 3.13 ± 0.02 | 1.72 ± 0.01 | 2.56 ± 0.02 | 3.15 ± 0.02 |
| | **CLG-IPL** | 1.73 ± 0.01 | 2.57 ± 0.02 | 3.18 ± 0.02 | 1.73 ± 0.01 | **2.61 ± 0.02** | 3.17 ± 0.03 |

[a] All values are average values and standard error of the mean (Å), over 30 independent EVB trajectories per system, obtained as described in the *Methodology* section, and shown here are data obtained from modeling the decarboxylation of each compound through cleavage of either the pro-(*R*) or pro-(*S*) carboxylate group. Shown here is also the corresponding data for the non-enzymatic reaction for reference. Note that the non-enzymatic transition states for the same compound are extremely similar in our EVB calculations. Data was collected every 10 ps of a 200 ps EVB mapping window corresponding to each stationary point per individual EVB trajectory, *i.e.* the data presented is averaged over 600 snapshots per system.

Similarly, in terms of structural effects, we also considered the impact of substrate binding on the active site volume of AMDase, calculated at the Michaelis complexes of wild-type AMDase

and its variants in complex of each of compounds **1a** through **1e**. These were calculated using Pocket VOlume MEasurer (POVME) 3.0,[52] as in our previous work.[53] As can be seen from **Figure 7** and **Table S3**, the calculated active site volumes largely follow substrate size. That is, the smallest active site volumes are observed in the case of compounds **1a** and **1d**, which differ only by substitutent (methyl for **1a**, ethyl for **1d**). This is followed by compound **1e**, which has an additional isopropyl group compared to compounds **1a**, and finally the multiring substrates **1b** and **1c**. The standard deviations on the calculated values also increase with increasing substrate size, but only slightly compared to the absolute volumes, suggesting the active site is flexible enough to accommodate also the bulky larger substrates, without being excessively "floppy".
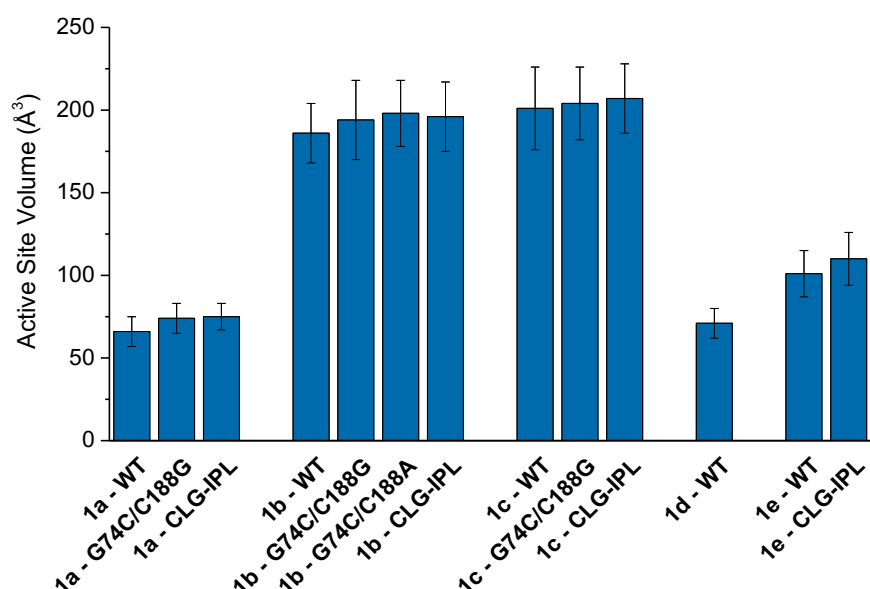


**Figure 7.** Average active site volumes during simulations of wild-type AMDase and its variants in complex with compounds **1b** to **1e**, calculated using Pocket VOlume MEasurer (POVME) 3.0.[52] Data is presented as average values and standard deviations over structures obtained at the Michaelis complexes of 30 independent EVB trajaectories, and analysis was performed on 600 snapshots per system (extracting data every 10 ps of the 200ps mapping window corresponding to the Michaelis complex of each individual EVB trajectory). The corresponding raw data for this figure is presented in **Table S3**.

We also considered the solvent-accessibility of the active site in our simulations, taking into account that one of the two carboxylate groups is stabilized by a dioxyanion hole while the other (more likely to be cleaved) carboxylate group is located in a hydrophobic pocket. As can be seen from **Figure 8** and **Table S4**, there is significant variety in the number of water molecules in close proximity (within 4 Å) of the carboxylate group being cleaved, with compounds that are turned over by AMDase typically having less than one water molecule close to the reacting group at the transition state, and with this number increasing to as many as four in the case of compounds **1d** and **1e** which either do not or are unlikely to react in the AMDase active site. This is likely due to the high flexibility of these substrates in the AMDase active site (**Figure 4**), which provides space for additional water molecules to enter the active site. We note that the number of water molecules for G74C/C188X variants is up to two, which may unfavorably also contribute to their low activity. The importance of sequestering the active site from solvent has been discussed in several prior experimental and computational studies by us and others,[54–57] and, in particular, from a computational perspective, we and others have clearly shown a correlation between activity loss and increased solvation of the active site for several enzymes.[54,56,58,59] Therefore, it is perhaps unsurprising to see yet again for AMDase increased solvent exposure of the active site in conjunction with the binding of compounds **1d** and **1e**, are either not turned over at all or only poorly converted by this enzyme, respectively, despite not being significantly structurally different to other compounds that are reactive (**Table 1** and **Scheme 2**).
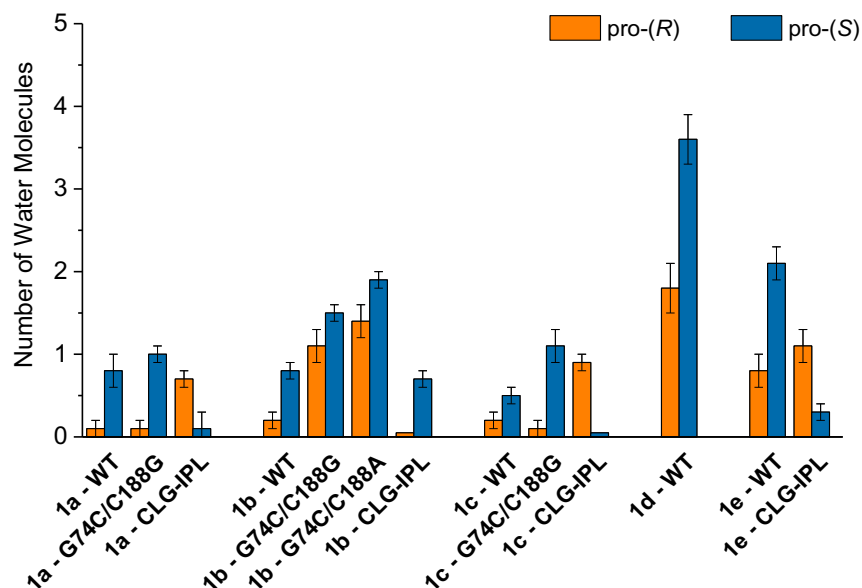
**Figure 8.** Average number of water molecules within 4 Å of the carboxylate group being cleaved (either pro-($R$) or pro-($S$), as relevant) during the last 25 ns of our 30 ns equilibration runs at the transition state for each reaction modeled in this work. Data is presented as average values and standard error of the mean over 30 individual trajectories per system, with data collected every 10 ps of simulation time. For the corresponding raw data associated with this figure, see **Table S4**.

Finally, although hydrophobic effects clearly dominate in determining the selectivity of AMDase (through destabilizing the pro-($R$) carboxylate and sequestering the active site from solvent), we have also considered the electrostatic contribution of individual amino acids to the calculated activation free energies (**Figure 9** and **Table S5**). This data was estimated by applying the linear response approximation (LRA)[60,61] to our EVB trajectories, as in previous work.[54,56,62] From this data, it can be seen that in the case of wild-type AMDase, where the preferred carboxylate group being cleaved is the pro-($R$) carboxylate, the T75 and Y126 side chains from the dioxyanion hole provide modest stabilizing contributions to the developing charge at the transition state, by stabilizing the pro-($S$) carboxylate group, although this contribution is offset by

a destabilizing contribution from the S76 side chain. In the case of cleavage of the pro-(*S*) carboxylate group (**Figure S7**), this is inversed with stabilizing contributions from T75 and S76, offset by a destabilizing contribution from Y126. Similarly, in the case of the side chains forming the hydrophobic pocket, contributions from all residues but M159 are destabilizing to the cleavage of the pro-(*R*) carboxylate group (**Figure 9**), whereas the inverse is observed for cleavage of the pro-(*S*) carboxylate (**Figure S7**) where the residues from the hydrophobic pocket make modest stabilizing contributions to the activation free energy for the decarboxylation reaction, and the side chain of M159 is destabilizing. These contributions are in conceptual agreement with how charge development is localized in the respective transition state. However, the fact that not all residues in the dioxyanion hole or hydrophobic pocket make stabilizing or destabilizing contributions for any given system also indicates that the residue contributions are more complex than than a simple model where one set of residues stabilizes and the other set of residues destabilizes the decarboxylation reaction.
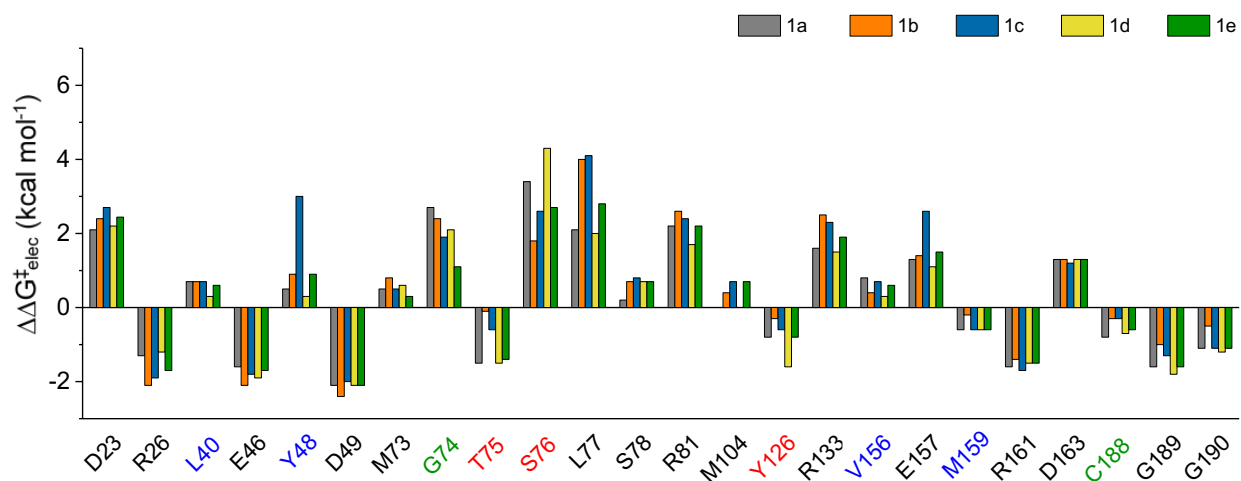


**Figure 9.** Electrostatic contributions of individual amino acids ($\Delta\Delta G^{\ddagger}_{elec}$, kcal mol$^{-1}$) to the calculated activation free energies for the decarboxylation of compounds **1a** to **1e** by wild-type AMDase. Data was obtained from our calculated EVB trajectories using the linear response approximation (LRA)[60,61] as in our

previous work,[54,56,62] and is presented as average values and standard error of the mean over 30 individual trajectories per system. The corresponding raw data and associated standard error of the mean for each value is shown in **Table S5.** Amino acids forming the oxyanion hole are highlighted in red, those forming the hydrophobic pocket in blue, and the catalytically important residues at positions 74 and 188 in green. Shown here is data corresponding to the energetically preferred cleavage of the pro-($R$) carboxylate group (**Table 1**). The corresponding figure and raw data for the cleavage of the pro-($S$) carboxylate group are shown in **Figure S7** and **Table S6**.

In summary, we have performed detailed EVB simulations of the decarboxylation of compounds **1a** through **1e** (**Scheme 2**) by wild-type AMDase and its variants, considering models for cleavage of both the pro-($R$) and pro-($S$) carboxylate groups. Our simulations can reproduce the substrate-preference of this enzyme, as well as providing a rationale for its selectivity. Specifically, our simulations clearly demonstrate a role for ground-state destabilization, through creating a hydrophobic cage for the carboxylate group being cleaved, with loss of activity in the case of compounds **1d** and **1e** being linked to increased stabilization of the carboxylate group being cleaved through greater solvent exposure of the active site, coupled with destabilization of the resulting cationic intermediate through inductive effects. In addition, from our electrostatic contributions, it can be seen that while the hydrophobic pocket mostly destabilizes the predominantly preferentially cleaved pro-($R$) carboxylate group, the actual interplay between stabilizing and destabilizing interactions within the hydrophobic pocket and dioxyanion hole is in fact quite complex. These simulations therefore provide clear insights into effects that can be easily manipulated in further engineering of this biocatalytically important enzyme.

# Conclusions

The unique capacity of co-factor-free decarboxylases to cleave C-C bonds under mild reaction conditions raises several questions regarding the destabilization of carbon-carbon bonds and the stabilization of an intermediary charge without the aid of an electron sink provided by an external cofactor. While its biological role has still not been clarified,[2] arylmalonate decarboxylase is a unique biocatalyst for the production of optically pure carboxylic acids from prochiral arylaliphatic malonic acids. However, despite increasing insight into the underlying molecular processes involved in the reaction mechanism,[2,23,30] several important questions remain unanswered. While a hydrophobic pocket in the active site was revealed to be a key determinant for AMDase activity,[30,31] the results of amino acid exchanges in this region have often been counterintuitive.[23,34,35] Similarly, restrictions on the substrate scope of this enzyme were hard to understand.[2,15,35] It remained unclear to which extent steric effects or the reactivity of the substrate control the acceptance of different substrates by AMDase. We were particularly curious to which degree a possible "ground-state destabilization"-driven mechanism or "Circe-effect" guides substrate acceptance, activity and selectivity. In line with this, a complete representation of this interaction network is challenging to decipher by truncated pure QM models, as the results obtained appear to be strongly dependent on the system size.[38,39]

Our empirical valence bond simulations help us obtain molecular-level insight into the drivers for the experimentally observed substrate acceptance of AMDase. We are able to reproduce activation free energies for the AMDase-catalyzed decarboxylation of compounds **1a** to **1c** and **1e** by all AMDase variants studied to within 3 kcal mol⁻¹ of the experimental value (where known). In all cases, our simulations are able to correctly predict both the product-selectivity and the substrate discrimination of AMDase. For all compounds, the preference for cleavage of the pro-

(*R*) *vs* the pro-(*S*) carboxylate group appears to be driven by substrate positioning in the Michaelis complex, with preferential cleavage of the carboxylate group that interacts most closely with the hydrophobic pocket.

In the case of compound **1d**, where no turnover is observed experimentally, our EVB calculations also yield activation free energies of >28 kcal mol$^{-1}$, depending on carboxylate group being cleaved. Further analysis of our simulations indicate, that this is due to a combination of inadequate substrate binding in the active site due to the presence of the bulky ethyl group, combined with greater solvent penetration into the active site, which is unfavorable for the decarboxylation reaction, and ties in with other work[54–58] emphasizing the importance of sequestering the active site from solvent. Similar observations are mde in the case of compound **1e** which is only poorly converted by AMDase. Finally, LRA calculations[60,61] performed on our EVB simulations allow us to decompose the contribution of individual amino acid side chains to the calculated activation free energies for each system, highlighting the complex interplay between individual stabilizing and destabilizing contributions of residues from the dioxyanion hole and hydrophobic pocket for cleavage of the pro-(*R*) and pro-(*S*) carboxylate groups, respectively.

This interplay between stabilizing and destabilizing contributions from the dioxyanion hole and hydrophobic pocket reflect the fact that, in turn, enzyme catalysis can, in theory, be facilitated by either stabilization of the transition state or destabilization of the ground-state, for example by placing the charged carboxylate group to be cleaved in a hydrophobic pocket as in the case of AMDase. While the latter may seem counter-intuitive, there nevertheless exist many examples of ground-state destabilization playing an important role both in catalysis by natural enzymes,[63–65] and by designed enzymes[66–70] where ground-state destabilization provides an easy design strategy (although there are limits to how far one can push design by affecting simply $K_M$). In particular,

the concept of ground-state destabilization in catalysis of decarboxylation reactions has been discussed extensively, for example in the case of the orotidine-5′-phosphate decarboxylase (OMPDC),[71–75] a decarboxylase with tremendous catalytic power (this enzyme which provides 31 kcal mol$^{-1}$ of transition state stabilization compared to the non-enzymatic reaction[76]). Like AMDase, OMPDC is one of the few cofactor-free decarboxylases, making the comparison particularly pertitent. In the case of OMPDC, evidence has been put forward that catalysis is not due to desolvation effects or ground-state destabilization, but rather due to electrostatic stabilization of the transition state for the decarboxylation reaction,[12,77,78] as well as the involvement of a ligand-gated conformational change that drives catalysis.[79,80] In the present case, our data indicate, that electrostatic interactions also play a clear role in stabilizing the *individual* transition states for the AMDase catalyzed decarboxylation reaction. However, both desolvation and ground-state destabilization clearly appear to be critical for determining the selectivity *between* different potential transition states, thus leading to the observed substrate- and product-selectivities in this enzyme. This is important both for being able to rationalize the effect of amino acid substitutions on AMDase selectivity, as well as for understanding the mechanistic principles in cofactor-free enzymes that have the capacity to cleave C-C bonds with the limited catalytic set of functional groups provided by the 20 canonical amino acids.

## Associated Content

Supporting Information: Additional simulation details, simulation analysis, and all EVB parameters necessary to reproduce this work.

## Acknowledgments

## References

(1)     Payer, S. E.; Faber, K.; Glueck, S. M. Non-Oxidative Enzymatic (De)Carboxylation of (Hetero)Aromatics and Acrylic Acid Derivatives. *Adv. Synth. Catal.* **2019**, *361*, 2402 – 2420.

(2)     Miyamoto, K.; Kourist, R. Arylmalonate Decarboxylase — A Highly Selective Bacterial Biocatalyst with Unknown Function. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8621–8631.

(3)     Frank, A.; Eborall, W.; Hyde, R.; Hart, S.; Turkenburg, J. P.; Grogan, G. Mutational Analysis of Phenolic Acid Decarboxylase from *Bacillus Subtilis* (BsPAD), Which Converts Bio-Derived Phenolic Acids to Styrene Derivatives. *Catal. Sci. Technol.* **2012**, *2*, 1568–1574.

(4)     Rude, M. A.; Baron, T. S.; Brubaker, S.; Alibhai, M.; Del Cardayre, S. B.; Schirmer, A. Terminal Olefin (1-Alkene) Biosynthesis by a Novel P450 Fatty Acid Decarboxylase from *Jeotgalicoccus* Species. *Appl. Environ. Microbiol.* **2011**, *77*, 1718–1727.

(5)    Rui, Z.; Harris, N. C.; Zhu, X.; Huang, W.; Zhang, W. Discovery of a Family of Desaturase-Like Enzymes for 1-Alkene Biosynthesis. *ACS Catal.* **2015**, *5*, 7091–7094.

(6)    Huijbers, M. M. E.; Zhang, W.; Tonin, F.; Hollmann, F. Light-Driven Enzymatic Decarboxylation of Fatty Acids. *Angew. Chem., Int. Ed.* **2018**, *57*, 13648–13651.

(7)    Steinreiber, J.; Schürmann, M.; Wolberg, M.; van Assema, F.; Reisinger, C.; Fesko, K.; Mink, D.; Griengl, H. Overcoming Thermodynamic and Kinetic Limitations of Aldolase-Catalyzed Reactions by Applying Multienzymatic Dynamic Kinetic Asymmetric Transformations. *Angew. Chem., Int. Ed.* **2007**, *46*, 1624–1626.

(8)    Gómez Baraibar, Á.; Reichert, D.; Mügge, C.; Seger, S.; Gröger, H.; Kourist, R. A One-Pot Cascade Reaction Combining an Encapsulated Decarboxylase with a Metathesis Catalyst for the Synthesis of Bio-Based Antioxidants. *Angew. Chem., Int. Ed.* **2016**, *55*, 14823–14827.

(9)    Enoki, J.; Mügge, C.; Tischler, D.; Miyamoto, K.; Kourist, R. Chemoenzymatic Cascade Synthesis of Optically Pure Alkanoic Acids by Using Engineered Arylmalonate Decarboxylase Variants. *Chem. Eur. J.* **2019**, *25*, 5071–5076.

(10)   Busto, E.; Simon, R. C.; Kroutil, W. Vinylation of Unprotected Phenols Using a Biocatalytic System. *Angew. Chem., Int. Ed.* **2015**, *54*, 10899–10902.

(11)   Baud, D.; Peruch, O.; Saaidi, P.-L.; Fossey, A.; Mariage, A.; Petit, J.-L.; Salanoubat, M.; Vergne-Vaxelaire, C.; de Berardinis, V.; Zaparucha, A. Biocatalytic Approaches towards the Synthesis of Chiral Amino Alcohols from Lysine: Cascade Reactions Combining Alpha-Keto Acid Oxygenase Hydroxylation with Pyridoxal Phosphate- Dependent Decarboxylation. *Adv. Synth. Catal.* **2017**, *359*, 1563–1569.

(12)   Richard, J. P.; Amyes, T. L.; Reyes, A. C. Orotidine 5′-Monophosphate Decarboxylase:

Probing the Limits of the Possible for Enzyme Catalysis. *Acc. Chem. Res.* **2018**, *51*, 960–969.

(13)  Schweiger, A. K.; Ríos-Lombardía, N.; Winkler, C. K.; Schmidt, S.; Morís, F.; Kroutil, W.; González-Sabín, J.; Kourist, R. Using Deep Eutectic Solvents to Overcome Limited Substrate Solubility in the Enzymatic Decarboxylation of Bio-Based Phenolic Acids. *ACS Sustainable Chem. Eng.* **2019**, *7*, 16364–16370.

(14)  Miyamoto, K.; Ohta, H. Enzyme-Mediated Asymmetric Decarboxylation of Disubstituted Malonic Acids. *J. Am. Chem. Soc.* **1990**, *112*, 4077–4078.

(15)  Miyamoto, K.; Ohta, H. Purification and Properties of a Novel Arylmalonate Decarboxylase from *Alcaligenes Bronchisepticus* KU 1201. *Eur. J. Biochem.* **1992**, *210*, 475–481.

(16)  Aßmann, M.; Stöbener, A.; Mügge, C.; Gaßmeyer, S. K.; Hilterhaus, L.; Kourist, R.; Liese, A.; Kara, S. Reaction Engineering of Biocatalytic (*S*)-Naproxen Synthesis Integrating in-Line Process Monitoring by Raman Spectroscopy. *React. Chem. Eng.* **2017**, *2*, 531–540.

(17)  Aßmann, M.; Mügge, C.; Gaßmeyer, S. K.; Enoki, J.; Hilterhaus, L.; Kourist, R.; Liese, A.; Kara, S. Improvement of the Process Stability of Arylmalonate Decarboxylase by Immobilization for Biocatalytic Profen Synthesis. *Front. Microbiol.* **2017**, *8*.

(18)  Gaßmeyer, S. K.; Wetzig, J.; Mügge, C.; Assmann, M.; Enoki, J.; Hilterhaus, L.; Zuhse, R.; Miyamoto, K.; Liese, A.; Kourist, R. Arylmalonate Decarboxylase-Catalyzed Asymmetric Synthesis of Both Enantiomers of Optically Pure Flurbiprofen. *ChemCatChem* **2016**, *8*, 916–921.

(19)  Terao, Y.; Ijima, Y.; Kakidani, H.; Ohta, H. Enzymatic Synthesis of (*R*)-Flurbiprofen. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 2395–2397.

(20)  Enoki, J.; Linhorst, M.; Busch, F.; Baraibar, Á. G.; Miyamoto, K.; Kourist, R.; Mügge, C.

Preparation of Optically Pure Flurbiprofen via an Integrated Chemo-Enzymatic Synthesis Pathway. *Mol. Catal.* **2019**, *467*, 135–142.

(21)   Tamura, K.; Terao, Y.; Miyamoto, K.; Ohta, H. Asymmetric Decarboxylation of α-Hydroxy- and α-Amino-α-Phenylmalonate Catalyzed by Arylmalonate Decarboxylase from *Alcaligenes Bronchisepticus*. *Biocatal. Biotransform.* **2008**, *26*, 253–257.

(22)   Lewin, R.; Goodall, M.; Thompson, M. L.; Leigh, J.; Breuer, M.; Baldenius, K.; Micklefield, J. Enzymatic Enantioselective Decarboxylative Protonation of Heteroaryl Malonates. *Chem. Eur. J.* **2015**, *21*, 6557–6563.

(23)   Okrasa, K.; Levy, C.; Wilding, M.; Goodall, M.; Baudendistel, N.; Hauer, B.; Leys, D.; Micklefield, J. Structure-Guided Directed Evolution of Alkenyl and Arylmalonate Decarboxylases. *Angew. Chem., Int. Ed.* **2009**, *48*, 7691–7694.

(24)   Yatake, Y.; Miyamoto, K.; Ohta, H. Screening, Cloning, Expression, and Purification of an Acidic Arylmalonate Decarboxylase from *Enterobacter Cloacae* KU1313. *Appl. Microbiol. Biotechnol.* **2008**, *78*, 793–799.

(25)   Miyamoto, K.; Yatake, Y.; Tamura, K.; Terao, Y.; Ohta, H. Purification and Characterization of Arylmalonate Decarboxylase from *Achromobacter* Sp. KU1311. *J. Biosci. Bioeng.* **2007**, *104*, 263–267.

(26)   Fernández-Álvaro, E.; Kourist, R.; Winter, J.; Böttcher, D.; Liebeton, K.; Naumer, C.; Eck, J.; Leggewie, C.; Jaeger, K.-E.; Streit, W.; Bornscheuer, U. T. Enantioselective Kinetic Resolution of Phenylalkyl Carboxylic Acids Using Metagenome-Derived Esterases. *Microb. Biotechnol.* **2010**, *3*, 59–64.

(27)   Maimanakos, J.; Chow, J.; Gaßmeyer, S. K.; Güllert, S.; Busch, F.; Kourist, R.; Streit, W. R. Sequence-Based Screening for Rare Enzymes: New Insights into the World of AMDases

Reveal a Conserved Motif and 58 Novel Enzymes Clustering in Eight Distinct Families. *Front. Microbiol.* **2016**, *7*.

(28) Miyamoto, K.; Tsuchiya, S.; Ohta, H. Stereochemistry of Enzyme-Catalyzed Decarboxylation of Alpha-Methyl-Alpha-Phenylmalonic Acid. *J. Am. Chem. Soc.* **1992**, *114*, 6256–6257.

(29) Obata, R.; Miyamoto, K.; Ohta, H.; Nakasako, M. Crystal Structures of Arylmalonate Decarboxylase – Implications for Enantioselective Reaction. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008**, *64*, C260–C260.

(30) Obata, R.; Nakasako, M. Structural Basis for Inverting the Enantioselectivity of Arylmalonate Decarboxylase Revealed by the Structural Analysis of the Gly74Cys/Cys188Ser Mutant in the Liganded Form. *Biochemistry* **2010**, *49*, 1963–1969.

(31) Okrasa, K.; Levy, C.; Hauer, B.; Baudendistel, N.; Leys, D.; Micklefield, J. Structure and Mechanism of an Unusual Malonate Decarboxylase and Related Racemases. *Chem. Eur. J.* **2008**, *14*, 6609–6613.

(32) Ijima, Y.; Matoishi, K.; Terao, Y.; Doi, N.; Yanagawa, H.; Ohta, H. Inversion of Enantioselectivity of Asymmetric Biocatalytic Decarboxylation by Site-Directed Mutagenesis Based on the Reaction Mechanism. *Chem. Commun.* **2005**, No. 7, 877–879.

(33) Miyamoto, K.; Tsutsumi, T.; Terao, Y.; Ohta, H. Stereochemistry of Decarboxylation of Arylmalonate Catalyzed by Mutant Enzymes. *Chem. Lett.* **2007**, *36*, 656–657.

(34) Miyauchi, Y.; Kourist, R.; Uemura, D.; Miyamoto, K. Dramatically Improved Catalytic Activity of an Artificial (S)-Selective Arylmalonate Decarboxylase by Structure-Guided Directed Evolution. *Chem. Commun.* **2011**, *47*, 7503.

(35) Yoshida, S.; Enoki, J.; Kourist, R.; Miyamoto, K. Engineered Hydrophobic Pocket of (S)-

Selective Arylmalonate Decarboxylase Variant by Simultaneous Saturation Mutagenesis to Improve Catalytic Performance. *Biosci., Biotechnol., Biochem.* **2015**, *79*, 1965–1971.

(36)   Terao, Y.; Miyamoto, K.; Ohta, H. Introduction of Single Mutation Changes Arylmalonate Decarboxylase to Racemase. *Chem. Commun.* **2006**, No. 34, 3600.

(37)   Busch, F.; Enoki, J.; Hülsemann, N.; Miyamoto, K.; Bocola, M.; Kourist, R. Semiempirical QM/MM Calculations Reveal a Step-Wise Proton Transfer and an Unusual Thiolate Pocket in the Mechanism of the Unique Arylpropionate Racemase AMDase G74C. *Catal. Sci. Technol.* **2016**, *6*, 4937–4944.

(38)   Lind, M. E. S.; Himo, F. Theoretical Study of Reaction Mechanism and Stereoselectivity of Arylmalonate Decarboxylase. *ACS Catal.* **2014**, *4*, 4153–4160.

(39)   Dasgupta, S.; Herbert, J. M. Using Atomic Confining Potentials for Geometry Optimization and Vibrational Frequency Calculations in Quantum-Chemical Models of Enzyme Active Sites. *J. Phys. Chem. B* **2020**, *124*, 1137–1147.

(40)   Warshel, A.; Weiss, R. M. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218–6226.

(41)   Frushicheva, M. P.; Warshel, A. Towards Quantitative Computer-Aided Studies of Enzymatic Enantioselectivity: The Case of *Candida Antarctica* Lipase A. *ChemBioChem* **2012**, *13*, 215–223.

(42)   Schopf, P.; Warshel, A. Validating Computer Simulations of Enantioselective Catalysis; Reproducing the Large Steric and Entropic Contributions in *Candida Antarctica* Lipase B. *Proteins: Struct., Funct., Bioinf.* **2014**, *82*, 1387–1399.

(43)   Bauer, P.; Carlsson, Å. J.; Amrein, B. A.; Dobritzsch, D.; Widersten, M.; Kamerlin, S. C. L. Conformational Diversity and Enantioconvergence in Potato Epoxide Hydrolase 1. *Org.*

*Biomol. Chem.* **2016**, *14*, 5639–5651.

(44)  Amrein, B. A.; Bauer, P.; Duarte, F.; Janfalk Carlsson, Å.; Naworyta, A.; Mowbray, S. L.; Widersten, M.; Kamerlin, S. C. L. Expanding the Catalytic Triad in Epoxide Hydrolases and Related Enzymes. *ACS Catal.* **2015**, *5*, 5702–5713.

(45)  Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(46)  Dunbrack, R. L.; Cohen, F. E. Bayesian Statistical Analysis of Protein Side-Chain Rotamer Preferences. *Protein Sci.* **1997**, *6*, 1661–1681.

(47)  The PyMOL Molecular Graphics System, Version 2.0. Schrödinger, LLC.

(48)  Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, *31*, 455–461.

(49)  Bauer, P.; Barrozo, A.; Purg, M.; Amrein, B. A.; Esguerra, M.; Wilson, P. B.; Major, D. T.; Åqvist, J.; Kamerlin, S. C. L. Q6: A Comprehensive Toolkit for Empirical Valence Bond and Related Free Energy Calculations. *SoftwareX* **2018**, *7*, 388–395.

(50)  Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(51)  Fisch, F.; Fleites, C. M.; Delenne, M.; Baudendistel, N.; Hauer, B.; Turkenburg, J. P.; Hart, S.; Bruce, N. C.; Grogan, G. A Covalent Succinylcysteine-like Intermediate in the Enzyme-Catalyzed Transformation of Maleate to Fumarate by Maleate Isomerase. *J. Am. Chem. Soc.* **2010**, *132*, 11455–11457.

(52)  Wagner, J. R.; Sørensen, J.; Hensley, N.; Wong, C.; Zhu, C.; Perison, T.; Amaro, R. E.

POVME 3.0: Software for Mapping Binding Pocket Flexibility. *J. Chem. Theory Comput.* **2017**, *13*, 4584–4592.

(53)     Calixto, A. R.; Moreira, C.; Pabis, A.; Kötting, C.; Gerwert, K.; Rudack, T.; Kamerlin, S. C. L. GTP Hydrolysis Without an Active Site Base: A Unifying Mechanism for Ras and Related GTPases. *J. Am. Chem. Soc.* **2019**, *141*, 10684–10701.

(54)     Kulkarni, Y. S.; Liao, Q.; Petrović, D.; Krüger, D. M.; Strodel, B.; Amyes, T. L.; Richard, J. P.; Kamerlin, S. C. L. Enzyme Architecture: Modeling the Operation of a Hydrophobic Clamp in Catalysis by Triosephosphate Isomerase. *J. Am. Chem. Soc.* **2017**, *139*, 10514–10525.

(55)     Richard, J. P.; Amyes, T. L.; Goryanova, B.; Zhai, X. Enzyme Architecture: On the Importance of Being in a Protein Cage. *Curr. Opin. Chem. Biol.* **2014**, *21*, 1–10.

(56)     Blaha-Nelson, D.; Krüger, D. M.; Szeler, K.; Ben-David, M.; Kamerlin, S. C. L. Active Site Hydrophobicity and the Convergent Evolution of Paraoxonase Activity in Structurally Divergent Enzymes: The Case of Serum Paraoxonase 1. *J. Am. Chem. Soc.* **2017**, *139*, 1155–1167.

(57)     Li, C.-H.; Tu, S.-C. Active Site Hydrophobicity Is Critical to the Bioluminescence Activity of *Vibrio Harveyi* Luciferase. *Biochemistry* **2005**, *44*, 12970–12977.

(58)     Liao, Q.; Kulkarni, Y.; Sengupta, U.; Petrović, D.; Mulholland, A. J.; van der Kamp, M. W.; Strodel, B.; Kamerlin, S. C. L. Loop Motion in Triosephosphate Isomerase is not a Simple Open and Shut Case. *J. Am. Chem. Soc.* **2018**, *140*, 15889–15903.

(59)     van der Kamp, M. W.; Chaudret, R.; Mulholland, A. J. QM/MM Modelling of Ketosteroid Isomerase Reactivity Indicates That Active Site Closure Is Integral to Catalysis. *FEBS J.* **2013**, *280*, 3120–3131.

(60) Lee, F. S.; Chu, Z.-T.; Bolger, M. B.; Warshel, A. Calculations of Antibody-Antigen Interactions: Microscopic and Semi-Microscopic Evaluation of the Free Energies of Binding of Phosphorylcholine Analogs to McPC603. *Protein Eng., Des. Sel.* **1992**, *5*, 215–228.

(61) Muegge, I.; Tao, H.; Warshel, A. A Fast Estimate of Electrostatic Group Contributions to the Free Energy of Protein-Inhibitor Binding. *Protein Eng., Des. Sel.* **1997**, *10*, 1363–1372.

(62) Purg, M.; Elias, M.; Kamerlin, S. C. L. Similar Active Sites and Mechanisms Do Not Lead to Cross-Promiscuity in Organophosphate Hydrolysis: Implications for Biotherapeutic Engineering. *J. Am. Chem. Soc.* **2017**, *139*, 17533–17546.

(63) Andrews, L. D.; Fenn, T. D.; Herschlag, D. Ground State Destabilization by Anionic Nucleophiles Contributes to the Activity of Phosphoryl Transfer Enzymes. *PLoS Biol.* **2013**, *11*, e1001599.

(64) Griswold, W. R.; Castro, J. N.; Fisher, A. J.; Toney, M. D. Ground-State Electronic Destabilization via Hyperconjugation in Aspartate Aminotransferase. *J. Am. Chem. Soc.* **2012**, *134*, 8436–8438.

(65) Zhang, Y.; Schramm, V. L. Ground-State Destabilization in Orotate Phosphoribosyltransferases by Binding Isotope Effects. *Biochemistry* **2011**, *50*, 4813–4818.

(66) Jencks, W. P. Destabilization Is as Important as Binding. *Philos. Trans. R. Soc., A* **1993**, *345*, 3–10.

(67) Eiben, C. B.; Siegel, J. B.; Bale, J. B.; Cooper, S.; Khatib, F.; Shen, B. W.; Players, F.; Stoddard, B. L.; Popovic, Z.; Baker, D. Increased Diels-Alderase Activity through Backbone Remodeling Guided by Foldit Players. *Nat. Biotechnol.* **2012**, *30*, 190–192.

(68) Mak, W. S.; Siegel, J. B. Computational Enzyme Design: Transitioning from Catalytic

Proteins to Enzymes. *Curr. Opin. Struct. Biol.* **2014**, *27*, 87–94.

(69) Jindal, G.; Ramachandran, B.; Bora, R. P.; Warshel, A. Exploring the Development of Ground-State Destabilization and Transition-State Stabilization in Two Directed Evolution Paths of Kemp Eliminases. *ACS Catal.* **2017**, *7*, 3301–3305.

(70) Hong, N.-S.; Petrović, D.; Lee, R.; Gryn'ova, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C.-Y.; Mabbitt, P. D.; et al. The Evolution of Multiple Active Site Configurations in a Designed Enzyme. *Nat. Commun.* **2018**, *9*, 3900.

(71) Lee, J. K.; Houk, K. N. A Proficient Enzyme Revisited: The Predicted Mechanism for Orotidine Monophosphate Decarboxylase. *Science* **1997**, *276*, 942–945.

(72) Feng, W. Y.; Austin, T. J.; Chew, F.; Gronert, S.; Wu, W. The Mechanism of Orotidine 5'-Monophosphate Decarboxylase: Catalysis by Destabilization of the Substrate. *Biochemistry* **2000**, *39*, 1778–1783.

(73) Appleby, T. C.; Kinsland, C.; Begley, T. P.; Ealick, S. E. The Crystal Structure and Mechanism of Orotidine 5'-Monophosphate Decarboxylase. *Proc. Natl. Acad. Sci.* **2000**, *97*, 2005–2010.

(74) Wu, N.; Mo, Y.; Gao, J.; Pai, E. F. Electrostatic Stress in Catalysis: Structure and Mechanism of the Enzyme Orotidine Monophosphate Decarboxylase. *Proc. Natl. Acad. Sci.* **2000**, *97*, 2017–2022.

(75) Miller, B. G.; Wolfenden, R. Catalytic Proficiency: The Unusual Case of OMP Decarboxylase. *Annu. Rev. Biochem.* **2002**, *71*, 847–885.

(76) Radzicka, A.; Wolfenden, R. A Proficient Enzyme. *Science* **1995**, *267*, 90–93.

(77) Warshel, A.; Štrajbl, M.; Villà, J.; Florián, J. Remarkable Rate Enhancement of Orotidine 5'-Monophosphate Decarboxylase Is Due to Transition-State Stabilization Rather Than to

Ground-State Destabilization. *Biochemistry* **2000**, *39*, 14728–14738.

(78)   Goryanova, B.; Amyes, T. L.; Richard, J. P. Role of the Carboxylate in Enzyme-Catalyzed Decarboxylation of Orotidine 5′-Monophosphate: Transition State Stabilization Dominates Over Ground State Destabilization. *J. Am. Chem. Soc.* **2019**, *141*, 13468–13478.

(79)   Amyes, T. L.; Richard, J. P.; Tait, J. J. Activation of Orotidine 5‘-Monophosphate Decarboxylase by Phosphite Dianion: The Whole Substrate Is the Sum of Two Parts. *J. Am. Chem. Soc.* **2005**, *127*, 15708–15709.

(80)   Moreira, C.; Calixto, A. R.; Richard, J. P.; Kamerlin, S. C. L. The Role of Ligand-Gated Conformational Changes in Enzyme Catalysis. *Biochem. Soc. Trans.* **2019**, *47*, 1449–1460.

# Table of Contents Graphic