Adapted DFTB3 repulsive potentials reach DFT accuracy for hydride transfer reactions in enzymes

José Luís Velázquez-Libera,[*,†,‡] Rodrigo Recabarren,[*,¶] David Adrian Saez,[§] Carlos Castillo,[¶] J. Javier Ruiz-Pernía,[†] Iñaki Tuñón,[†] and Esteban Vöhringer-Martinez[*,¶]

†*Departamento de Química Física, Universitat de Valencia, Valencia 46100, Spain*

‡*Departamento de Bioinformática, Centro de Bioinformática, Simulación y Modelado (CBSM), Facultad de Ingeniería, Universidad de Talca, Talca 3460000, Chile*

¶*Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción, Chile*

§*Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío-Bío, Chillán, Chile*

E-mail: josevlibera2010@gmail.com; rrecabarren@udec.cl; evohringer@udec.cl

Abstract

Enzymatic hydride transfer reactions play a crucial role in numerous metabolic pathways, yet their accurate computational modeling remains challenging due to the trade-off between accuracy and computational efficiency. Ideally, molecular dynamics simulations should sample all enzyme configurations along the reaction path using post Hartree-Fock or DFT QM/MM electrostatic embedding methods, but these are computationally expensive. Here, we introduce a simple approach to improve the third-order density functional tight binding (DFTB3) semi-empirical method to model hydride transfer reactions in enzymes. We identified deficiencies in DFTB3's description

of the potential energy surface for the hydride transfer step in Crotonyl-CoA Carboxylase/Reductase (Ccr) and developed a systematic methodology to address these limitations. Our approach involves modifying DFTB3's repulsive potential functions using linear combinations of harmonic functions, guided by analysis of C-H and C-C distance distributions along the reaction path. The optimized DFTB3 Hamiltonian significantly improved the description of the hydride transfer reaction in Ccr, reproducing the reference DFT activation barrier within 0.1 kcal/mol. We also addressed the transferability of our method by applying it to another hydride transfer reaction bearing the 1,4-dihydropyridine motif but exhibiting distinct structural features of the reactant , as well as the hydride transfer reaction in Dihydrofolate Reductase (DHFR). In both cases our adapted DFTB3 Hamiltonian correctly reproduced the DFT reference and experimentally observed activation barriers. The low computational cost and transferability of our method will enable more accurate and efficient QM/MM molecular dynamics simulations of hydride transfer reactions, potentially accelerating research in enzyme engineering and drug design.

Introduction

Enzymatic reactions play a fundamental role in biochemical processes, and the study of their mechanisms is of significant interest in biochemistry and computational biology. Understanding the intricate details of these reactions is crucial for gaining insights into enzymatic activity and for developing applications in biotechnology and drug design. Among enzymatic processes, hydride transfer reactions are of particular importance due to their involvement in numerous metabolic pathways. These reactions, often catalyzed by enzymes like NAD(P)H-dependent oxidoreductases,[1,2] play a vital role in maintaining cellular energy balance and enabling complex biochemical transformations. Their mechanisms have inspired the design of new biomimetic compounds with potential applications in green chemistry and drug manufacturing.[3–5] Possibly the most clear example of exploiting hydride transfer reac-

tions in biocatalysis is the use of ene-reductases for asymmetric reduction of activated C=C bonds in both synthetic and industrial applications.[6–8]

Quantum mechanics / molecular mechanics (QM/MM) methods are widely employed for modeling enzymatic reaction mechanisms,[9,10] including hydride transfer reactions.[11–17] The dynamic nature of enzymes, characterized by flexibility and numerous degrees of freedom, necessitates thorough exploration of the conformational space, presenting challenges in navigating complex potential energy surfaces.[18] To address this, molecular dynamics simulations (MD) employing a hybrid QM/MM Hamiltonian are commonly utilized. To enhance computational efficiency, these simulations often rely on semiempirical methods, which can be significantly faster—up to three orders of magnitude—than their DFT counterparts.[19] However, the use of semiempirical methods introduces a potential drawback, as they may exhibit inaccuracies in describing diverse chemical processes. A common approach to overcoming this challenge is to perform a careful parametrization of the semiempirical method for the reaction of interest, where neglect of diatomic differential overlap (NDDO)-based methods with specific reaction parameters are common examples.[20,21] One example is the reparametrization of the AM1 semiempirical method to describe the hydride transfer reaction in the enzyme dihydrofolate reductase (DHFR).[12] However, problems with its transferability to other hydride transfer reactions have been observed.[22]

Among the available semiempirical Hamiltonians, the self-consistent charge density functional tight-binding (SCC-DFTB) method[23] has demonstrated good performance in enzyme-catalyzed reaction benchmarking studies, outperforming methods like AM1 and PM3 in overall reaction energies; however, there have also been notable cases where SCC-DFTB falls short.[24] As a significant improvement, the third-order density functional tight-binding method (DFTB3) was developed.[25] DFTB3 has proven to be more robust as an alternative to its predecessor, providing reliable energetics and geometries for a wide range of chemical systems.[26] Aiming to improve the performance of DFTB3 for specific reactions, element-specific parameterizations for magnesium, zinc, copper, and others have been reported,[27–29]

with a particular focus on their use in enzymatic systems. Addressing DFTB3 limitations often involves re-parametrizing the repulsive potentials (RPs). In the density functional tight binding (DFTB) theory, the RPs hold a critical role, encompassing necessary corrections to faithfully reproduce target DFT results.[30] Serving as a key adjustable parameter, RPs have been generated using several methods to better reproduce chemical properties,[29,31–34] with recent advancements incorporating machine learning techniques.[35–37]

In this study, our goal was to create a new parameterization of RPs tailored for describing hydride transfer reactions, with a potential application to NAD(P)H-dependent oxidoreductases. To the best of our knowledge, there is no systematic study showing the performance of DFTB3 for hydride transfer reactions in enzymatic systems. We selected Crotonyl-CoA Carboxylase/Reductase (Ccr),[38] a unique enzyme class within the medium chain dehydrogenase/reductase superfamily, as our model system. Ccr is one of the most efficient CO_2-fixing enzymes and has already been included in synthetic CO_2 fixation pathways.[39,40] Ccr catalyzes the reductive carboxylation of the substrate crotonyl-CoA to form the carboxylated product (2*S*)-ethylmalonyl-CoA as shown in Figure 1. Upon absence of CO_2 in the active site, a reduced side product can be formed which would cause reduction in catalytic efficiency. After the initial hydride transfer step and formation of an enolate intermediate, the enzyme can promote the formation of a covalent adduct (C2 adduct). The latter has been proposed to enhance catalytic efficiency by "storing" the reactive enolate in a more stable intermediate.[41]

This study aims to improve the accuracy and applicability of DFTB3 for investigating enzymatic hydride transfer reactions, with a focus on Ccr's catalytic process. The proposed modifications hold the potential to be applied to a wide range of hydride transfer reactions. We also propose a rather simple approach to find relevant regions where the RPs can be modified. Due to the low computational cost of semiempirical methods, an improved description of hydride transfer reactions in biochemical contexts can advance our understanding of enzymatic mechanisms and contribute in areas of enzyme engineering, protein and drug

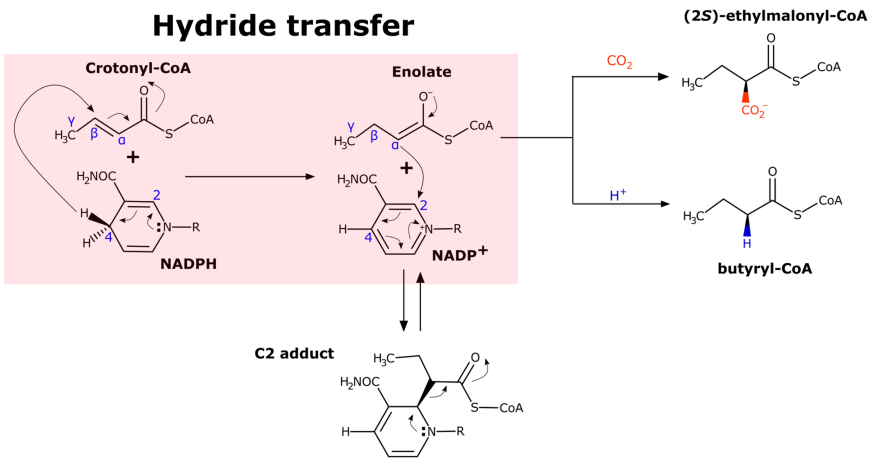

Figure 1: Ccr's characterized catalytic pathways. Highlighted in a red square is the hydride transfer step creating the reactive enolate intermediate.

design.

Methodology

Improvements of the DFTB3 Hamiltonian

In this subsection, we delve into the application of DFTB methods, particularly DFTB3, to model the intricate details of the system under investigation. We begin by laying out a concise overview of its theoretical foundation and computational principles, with a special focus on the *Repulsive Potential (RP)* terms.

Theoretical Basis

DFTB3 belongs to the family of tight-binding (TB) methods, which approximate the electronic structure of a system by using limited atomic orbital (AO) basis-set representation.[30] However, unlike conventional TB methods, DFTB3 integrates principles from *Density Functional Theory (DFT)* through a self-consistent field cycle, allowing a more accurate description of various chemical phenomena. The energy functional encompasses terms of electronic

contributions (E^{H0}, E^{γ}, and E^{Γ}) and the element specific pairwise RPs (E^{rep}):

$$\begin{aligned} E^{\mathrm{DFTB3}} =& E^{\mathrm{H0}} + E^{\gamma} + E^{\Gamma} + E^{\mathrm{rep}} \\ =& \sum_{iab} \sum_{\mu \in a} \sum_{\nu \in b} n_i c_{\mu i} c_{\nu i} H^0_{\mu\nu} + \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab} \\ &+ \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab} + \frac{1}{2} \sum_{ab} V^{\mathrm{rep}}_{ab}(r_{ab}) \end{aligned} \tag{1}$$

where: $E^{\mathrm{H0}} = \sum_{iab} \sum_{\mu \in a} \sum_{\nu \in b} n_i c_{\mu i} c_{\nu i} H^0_{\mu\nu}$ is the electronic term of first order and represents the energy contribution from an atomic orbital Hamiltonian depending only on the reference density. The electronic term of second order $E^{\gamma} = \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab}$ models the coulombic interactions between the net charges at atoms a (Δq_a) and b (Δq_b), being γ_{ab} a function that yields the Hubbard parameter when $a = b$ ($\gamma_{aa} = U_a$, the on-site self-repulsion) and $1/r_{ab}$ for large distances r_{ab} (pure coulombic interactions between the charges Δq_a and Δq_b). The electronic term of third order $E^{\Gamma} = \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab}$ accounts for the derivative of the γ function with respect to charge ($\Gamma_{ab} = \frac{\partial \gamma_{ab}}{\partial q_a}\Big|_{q_a^0}$), which significantly improves the description of charged systems by allowing DFTB3 to capture dynamic charge responses and enhance electrostatics. Finally, the repulsive energy contribution is approximated as the sum of the short-ranged two-center potentials $E^{rep} = \frac{1}{2} \sum_{ab} V^{\mathrm{rep}}_{ab}(r_{ab})$, only depending on the interatomic distances r_{ab}.

In DFTB parameterization, the repulsive potential function (RPF) for a specific pair of atoms, denoted as a and b, represented as $V^{\mathrm{rep}}_{ab}(r_{ab})$, is characterized by a differentiable curve. This curve is partitioned into segments through the application of an exponential decay function, subsequently followed by the incorporation of fourth-order polynomials (spline).

The RPFs are formally represented as:

$$V^{\text{rep}}(r) = \begin{cases} S_0(r) = \alpha e^{\beta r + \gamma} & \forall\ r \in [0, r_1), \\ S_i(r) = \sum_{k=0}^{4} a_{k,i}(r - r_i)^k & \forall\ r \in [r_i, r_{i+1}), \quad i = 1, 2, \ldots, n \\ 0 & \forall\ r \in [r_{n+1}, \infty), \end{cases} \tag{2}$$

where: the coefficients α, β, γ, and $a_{k,i}$, are the parameters to be fitted for a given set of spline curve interval boundaries r_i. For interatomic distances close to 0Å, the RPF (as $S_0(r)$) would raise, considerably increasing the potential energy of the system, while for distances larger or equal than a certain cutoff ($r \geq r_{n+1}$) it will be zero.

The large number of parameters in the RPFs offers a mechanism for fine-tuning DFTB3's performance, enabling it to more accurately capture specific physical processes that might be deficient in its available parameterizations. However, caution is warranted when utilizing this flexibility. Over-fitting to a specific dataset can compromise transferability to other systems or conditions, potentially limiting the model's broader applicability. Additionally, any modifications to the RPFs should align with established physical principles to maintain the model's reliability. To ensure the efficacy and robustness of fitted RPFs, it is crucial to conduct thorough validation against independent data, thereby assessing both their accuracy and predictive power.

Improving DFTB3

While DFTB3 offers a valuable tool for simulating diverse chemical systems, its inherent approximations sometimes limit its accuracy. This subsection details our proposed methodology to refine the repulsive potential terms within DFTB3, specifically targeting the hydride transfer step in Ccr's catalytic mechanism proposed in the literature.

Identifying deficiencies in the potential energy surface (PES) description. Our initial step involved analyzing the deficiencies in DFTB3's description of the PES for the hy-

dride transfer step (see Figure 1 and further discussion in "Results and discussions" section). This process involved the optimization of the transition state (TS) at the M06-2X/def2-TZVP level theory in vacuum, followed by IRC calculations at the same level of theory to map the reaction pathway and identify discrepancies with the DFTB3-predicted energy profile on DFT geometries.

C-H and C-C Distance Distributions. By examining the distributions of C-H and C-C distances around the reactants, TS and products geometries, we gained insights into the regions where the repulsive potential terms exerted the most significant influence on the total energy. This analysis allowed to tailor these terms within specific intervals relevant to correct C-H and C-C RPFs to approximate DFTB3's potential energy profiles on the reaction mechanism to DFT energies at the M06-2X/def2-TZVP level of theory.

Modifying the RPFs. Considering the observed distance distributions, we opted to modify the C-H and C-C RPFs represented by cubic splines. Our approach consisted in adding linear combinations of harmonic functions (LCHFs) bounded within the target intervals and with amplitudes of our choice to the original RPFs. Similarly, we also tested linear combinations of Gaussian functions (LCGFs) as alternative (see the definition in the SI), similar to previous studies,[42,43] (see Figure 2). However, LCHFs strategy showed more accurate fit. The LCHFs are highly versatile, allowing for precise adjustments even in overlapping intervals. This enabled us to fine-tune the RPs without introducing abrupt changes in their shape. On the other hand, the LCHFs can be readily confined to specific intervals with precise control over their boundaries, ensuring targeted modifications aligned with the C-H and C-C distance distributions. Unlike LCGFs, which tend to be highly localized around their central points, LCHFs exhibit a broader spatial distribution, facilitating smoother adjustments across the desired intervals.

Formally, the harmonic functions (HF) proposed in this work are defined as:

$$HF(\mathbf{r}, lb, hb, \Delta E) = \begin{cases} 0 & \forall\ \mathbf{r} \in (-\infty, lb] \cup [hb, \infty), \\ \frac{\Delta E}{2}\left(\sin\left(\frac{2\pi}{(hb-lb)}\mathbf{r} - \frac{(hb+3lb)\pi}{2(hb-lb)}\right) + 1.0\right) & \forall\ \mathbf{r} \in (lb, hb), \end{cases} \tag{3}$$

where: ΔE is the maximum value at the center of the objective interval (lb, hb). For every objective interval, we construct a LCHFs by adding two HFs with ΔE values of opposite sign, where the one with the lower absolute value has its boundaries equidistant from the outer boundaries of the other by a small magnitude (~ 0.05Å). While the HFs have the absolute values of their maximum at the center of the inner and outer intervals, the maximum of their sum will be the difference between their respective ΔE values. The same approximation was applied to obtain the tested LCGFs (see the GFs' definition in the SI). As an example, for the objective interval (1.20Å, 2.60Å) and ΔE of 2.0 kcal/mol, the curves plotted in Figure 2 are defined as:

$$HF(\mathbf{r}) = HF(\mathbf{r}, \text{lb}{=}1.20, \text{hb}{=}2.60, \Delta\text{E}{=}2.0) \tag{4}$$

$$LCHF(\mathbf{r}) = HF(\mathbf{r}, 1.20, 2.60, 6.0) + HF(\mathbf{r}, 1.25, 2.55, -4.0) \tag{5}$$

$$GF(\mathbf{r}) = GF(\mathbf{r}, \text{lb}{=}1.20, \text{hb}{=}2.60, \Delta\text{E}{=}2.0, \text{w}{=}1E^{-5}) \tag{6}$$

$$LCGF(\mathbf{r}) = GF(\mathbf{r}, 1.20, 2.60, 6.0, 10^{-5}) + GF(\mathbf{r}, 1.25, 2.55, -4.0, 10^{-5}) \tag{7}$$

Optimization of LCHFs' parameters. By determining the target intervals and taking into account the properties of the LCHFs, we iteratively refined the parameter values until the resulting potential energy profiles closely matched the high-level DFT reference data. This simple optimization procedure minimizes the need for complex algorithms, making it readily applicable to diverse systems.

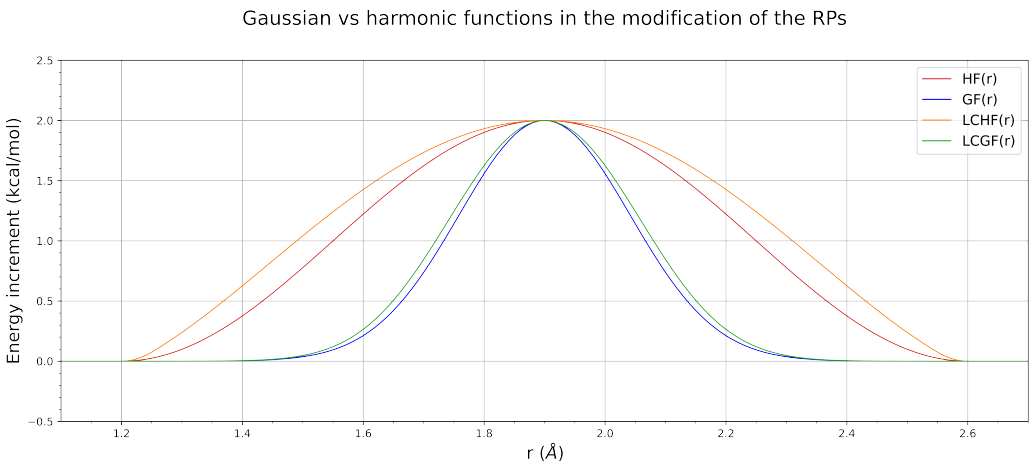

Figure 2: Visualization of LCHFs versus LCGFs. The plots depict four curves defined between 1.20 and 2.60 Å with a maximum of 2 kcal/mol at 1.9 Å, showing how LCHFs (orange line) offer greater flexibility and control of the dispersion of the curves over the objective interval compared to LCGFs (green line). LCGFs decrease immediately, being more localized around the center of the interval than LCHFs. We also plot a Gaussian function (line in blue) and a harmonic (line in red) as baselines to assess the magnitude of the corrections made by using the linear combinations instead of simply the basic functions in each case, respectively.

Computational Methods

Benchmark of density functionals and basis sets

A model system was defined consisting of the ribose-nicotinamide portion of NADP(H), and the thioester fraction of the crotonyl-CoA substrate. A preliminary transition structure was sought using the autodE software[44] at the PBE0/def2-SVP level of theory. This transition structure (TS) was optimized *in vacuo* at the M06-2X/def2-TZVP level using the software Gaussian 16.[45] The presence of an imaginary frequency corresponding to the hydride transfer was verified. The reaction coordinate was followed *in vacuo* at the same level of theory, leading to configurations of the reactant (reduced nicotinamide configuration) and product (enolate-type configuration). The geometries of these endpoints plus the TS were extracted, and their energy evaluated using different methods and basis sets. For the density functionals, calculations with the D3 dispersion correction[46] were also included. To evaluate the performance of these DFT methods, explicitly correlated and local wave-

function methods were employed using the software Molpro,[47–49] along with cc-pVXZ-F12 (X=D,T) correlation-consistent basis sets to determine the energy of each of the optimized states.[50] Density fitting was utilized for all energy computations. Highly accurate energies were obtained using the composite method:

$$\Delta E_{comp} = \Delta E(\text{MP2-F12/VTZ-F12}) + \Delta E_{CC}(\text{VDZ-F12}) \tag{8}$$

where:

$$\Delta E_{CC}(\text{VDZ}) \ = E(\text{LCCSD(T)-F12/VDZ-F12}) - E(\text{LMP2/VDZ-F12}) \tag{9}$$

The calculated values served as reference energies to evaluate the performance of multiple functionals/basis sets combinations.

System preparation and reaction setup

The computational models of Ccr were constructed based on the crystal structure of the ternary complex of *K. setae* Ccr (*Ks*Ccr, PDB ID: 6NA4) and representative structures from previous work.[41] This complex displays a tetrameric arrangement, organized as a dimer of dimers. Each dimer comprises one closed and one open subunit (designated as subunits A/C and B/D), with the closed subunits representing the catalytically active form of the enzyme (subunits A and B). Each subunit includes a NADPH cofactor, with the closed subunits also containing the substrate analogue butyryl-CoA. Details regarding system preparation, parametrization, and initial equilibration of the system are documented in previous work.[41] Representative structures of reactants and the enolate intermediate from a converged string simulation, describing the direct mechanism, were used to model the hydride transfer reaction. These two states served as the start and end points for the subsequent reaction modeling studies discussed below.

QM/MM MD simulations and minimum free energy path calculations

Ccr's QM region was reduced compared to our previous study[41] as the carboxylation reaction was not studied, making it comparable to the model used in the initial benchmark calculations (see Figure 3). The link atom approach, as implemented in the *sander* program of the AmberTools18 suite,[51] was employed to saturate missing valences at the QM/MM interface, summing to a total of 48 QM atoms (two link atoms). Initial QM/MM MD simulations were performed in *sander* using the DFTB3 semiempirical method's (3OB-3-1 parameter set),[26,29] while the MM region was treated with the CHARMM36 force field and parameters derived for NADPH cofactor and crotonyl-CoA substrate.[41]

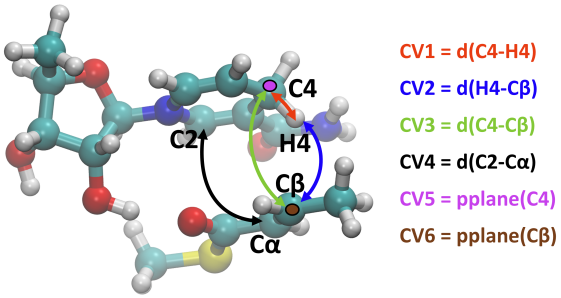

Figure 3: Representation of the QM region and the employed collective variables used in the ASM method to describe the hydride transfer reaction in Ccr.

The adaptive string method (ASM)[52] was used to study the hydride transfer reaction in Ccr. Our implementation of the ASM approach in *sander* has been used extensively for the study of several enzymatic reactions.[41,53–56] This allows finding the minimum free energy path (MFEP) between two states, in this case, the reactants and the enolate intermediate, within a space of collective variables (see Figure 3) and generating the final free energy profile for the reaction under study. An initial path was constructed by linear interpolation of collective variable (CV) values between structures of the reactants and the enolate. This path was discretized into 60 nodes, with each node undergoing an independent QM/MM MD simulation. Node positions are continuously reparametrized to keep them equidistant, evolving through regions of lower free energy until converging to the MFEP. The obtained

MFEP is used to generate a path CV (s reaction coordinate) on which all relevant configurational space is sampled by means of Umbrella Sampling (US) simulations.[57] Final PMFs were generated using the Weighted Histogram Analysis Method (WHAM).[58]

Selection and validation of DFTB3's modified parameter sets

During the optimization of the LCHFs' parameters (lb, hb, and ΔE), we determined the objective intervals and the respective values of ΔE with the highest impact on correcting DFTB3's deficiencies on describing the PES along the IRC profile obtained at the M06-2X/def2-TZVP level of theory. For each DFTB3 parameter set including a modified version of C-C and/or C-H RPs, we performed single-point calculations on the IRC geometries at DFTB3 level and evaluated the performance of the corrections taking as reference the energies at the M06-2X/def2-TZVP level of theory. Once we determined a few groups of modified parameters significantly improving DFTB3 Hamiltonian for describing the IRC reaction path, we carried out tests through QM/MM restrained geometry optimizations along the antisymmetric coordinate for the hydride transfer reaction at the DFTB3/CHARMM36 level of theory. To evaluate and select the best parameter sets, we computed single-point calculations on DFTB3's optimized geometries at the M06-2X/def2-TZVP/CHARMM36 level of theory, in this way we could compare the semiempirical and DFT energy profiles.

The validation of the selected sets of DFTB3 parameters was finally completed through free energy calculations using the ASM approach. In the case of the potential of mean force (PMF) profiles obtained at the DFTB3/CHARMM36 level with modified parameters, corrections were made at the M06-2X/def2-TZVP/CHARMM36 level, as we report in our previous work,[59] to evaluate the new DFTB3 parameter sets' performance in the predictions of the free energy surface (FES). Full free energy calculations at the M06-2X/def2-SVP/CHARMM36 level were also carried out and served as energetic and structural reference to the semiempirical calculations described above.

Results and discussion

In our previous study of the hydride transfer step in Ccr,[41] we observed a severe underestimation of the free energy barrier in the reductive carboxylation mechanism when employing QM/MM MD simulations with the DFTB3 method and the 3OB-3-1 parameters set, which required correction at a higher level of theory. In this study, we recalculated the free energy profile for this step using a smaller QM region (see Figure 4), observing a similar behavior, which we also corrected using single-point calculations at the DFT level. To address these shortcomings and enhance the predictive accuracy of DFTB3 for studying Ccr's hydride transfer without relying on *a posteriori* corrections, we developed an approach based on the analysis of structural changes along the reaction path to guide the fitting process of the RPs. Our strategy involves the introduction of harmonic functions to modify the C-H and C-C RPs in DFTB3. This simple yet effective modification aims to improve the representation of bond-breaking and bond-forming events during the hydride transfer, thereby improving the energetic description of the reaction.

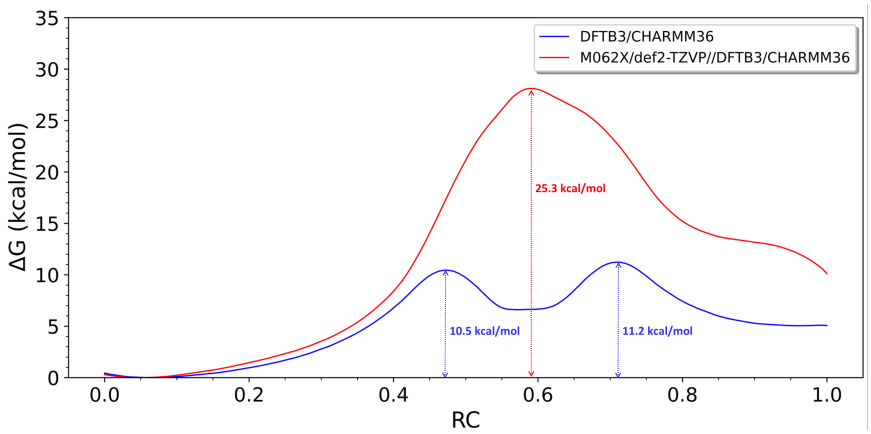

Figure 4: Free energy profile along the reaction coordinate obtained from the adaptive string method (ASM) computed at DFTB3/CHARMM36 level of theory (in blue) and its corrected version at the M06-2X/def2-TZVP/CHARMM36 level of theory (in red).

We validated the effectiveness of our proposed modifications by performing QM/MM molecular dynamics simulations at the M06-2X/def2-SVP/CHARMM36 level of theory. These calculations served as a reference for the accuracy and reliability of our modified

DFTB3 RPs. Additionally, we applied the modified RPs to the well-known Dihydrofolate Reductase (DHFR) enzyme where they significantly improved the predicted free energy barrier compared to the original DFTB3 parameters. As a final test, the new parametrization was tested in a small system bearing the 1,4-dihydropyridine core motif of NAD(P)H and using benzyne as a substrate, where it corrected the energy profile provided by the original DFTB3 Hamiltonian, showing the transferability of the new parameters.

Analysis of DFTB3 deficiencies in describing the hydride transfer step in Ccr

To disclose the origin of the deficiencies in correctly describing the hydride transfer reaction by the DFTB3 Hamiltonian, we optimized the transition state at the M06-2X/def2-TZVP level of theory in vacuum. Then, we calculated the activation barrier using different DFT functionals and post-HF methods to validate an accurate and computationally efficient electronic structure method (see Table S1 in the Supporting information). The benchmark evidences that the M06-2X functional in combination with the def2-TZVP basis set describe the potential energy difference between reactants and transition state accurately as well as the potential energy difference to the enolate, the product of the hydride transfer reaction. With the validated reference electronic structure method, we followed the intrinsic reaction path from the transition state to the reactant and the enolate as shown in Figure 5A (curve in blue). When we calculated the potential energy with the DFTB3 Hamiltonian on the same structures along the reaction path, significant deviations appeared at the transition state (see the orange curve at Figure 5A). Moreover, the DFTB3 Hamiltonian presented a local minimum on the potential energy surface close to the transition state instead of a maximum.

At the transition state of the hydride transfer reaction studied at the M06-2X/def2-TZVP level of theory, the two carbon atoms approach each other to 2.70 Å leaving the hydrogen atom at 1.47 Å from the donor and 1.27 Å from the acceptor. To analyze the contributions

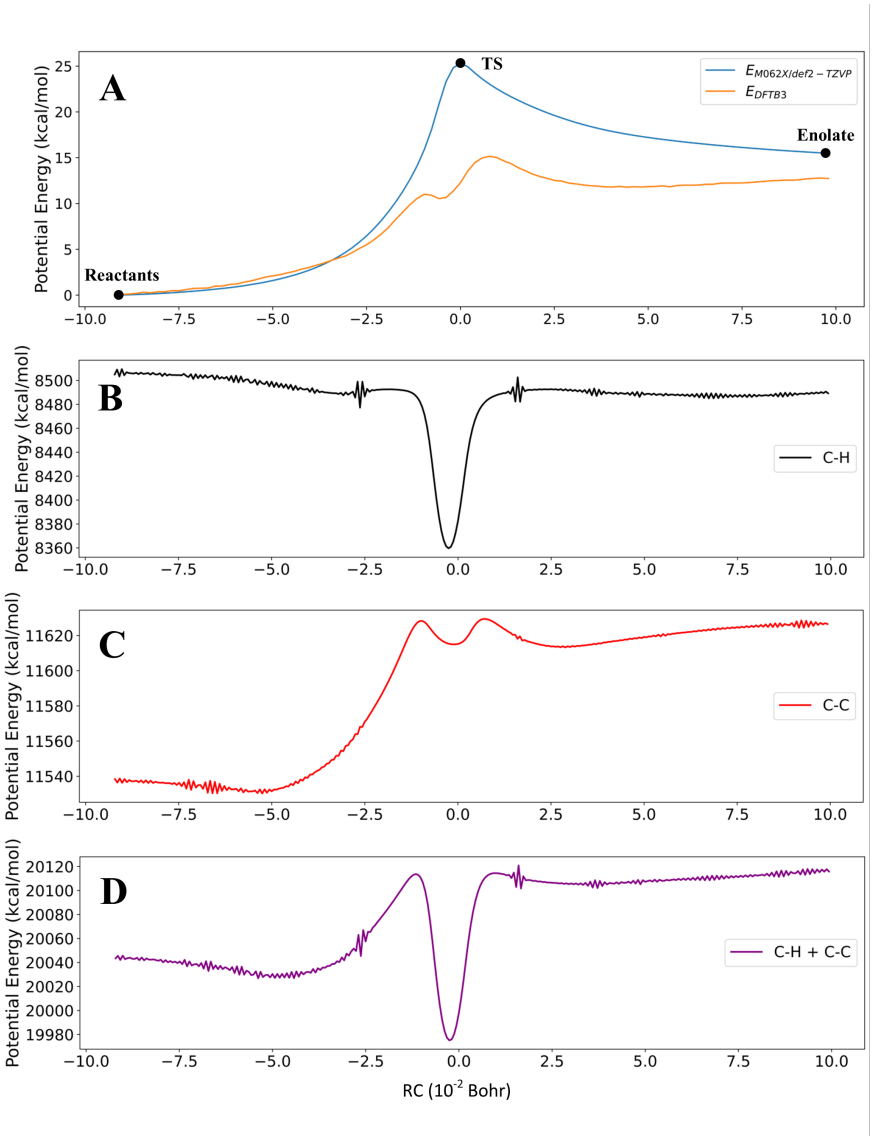

Figure 5: **A)** Comparison between the minimum energy path along the intrinsic reaction coordinate obtained at the M06-2X/def2-TZVP level (curve in blue) of theory and the DFTB3 potential energy for the same molecular geometries (curve in orange). The reactants, TS and enolate states are highlighted as a reference. The local minimum of DFTB3 close to the transition state is explained by the contribution of the RPs to the total DFTB3 Hamiltonian: **B)** the C-H RPs, **C)** the C-C RPs, and **D)** the sum of C-H and C-C RPs.

of the RPFs' along the reaction path, we show their individual contributions to the potential energy provided by all C-H and C-C atom pairs, respectively, as well as the sum of both for each structure along the IRC path in Figure 5B-D. Interestingly, the observed local minimum close to the TS in DFTB3's PES in Figure 5A coincides with the minima observed in the C-H and C-C total potential energies as well as their sum. We concluded that the DFTB3 Hamiltonian's deficiency to describe the hydride transfer reaction relies mostly in both RPs. We hypothesized that this anomalous behavior comes from not considering structures with stretched C-H bonds in the original fitting process of the 3OB-3-1 parameter set derivation. [26]

To elucidate the origin of the total potential energy provided by the repulsive potential functions (RPFs) between C-H or C-C pairs, we show the individual RPFs for each C-H and C-C atom pairs as implemented in the DFTB3 Hamiltonian (3OB-3-1 set of parameters) in Figures 6A and 6B, respectively. Although both potentials are repulsive at short distances, they present attractive regions at distances larger than 1.95 Å for C-H and 2.95 Å for C-C, reaching zero around the 3.5 Å and 4.5 Å, respectively. To examine the changes in these distances along the IRC path, we compared the C-C and C-H distance distributions of all C-C and C-H pair atoms for the reactants and TS in Figures 6C and 6D. This comparison allowed us to determine which distance intervals are more frequently occupied at the TS compared to the reactants. As the reaction progresses to the transition state, the NADPH ring moves closer to the substrate to transfer the hydride from C_4 to C_β (see Figure 6E). This reduces the C-H and C-C distances, shifting populations towards the attractive regions of the RPFs, specifically within 2.1-2.7 Å and 2.9-3.2 Å for C-H, and 2.9-3.4 Å for C-C pairs. To address DFTB3's shortcomings in modeling hydride transfer reactions, one could consider adjusting the attractive regions of the C-H and C-C RPFs to be more repulsive. Meanwhile, the C-H distances ranging from 1.0-2.0 Å displayed the anticipated changes due to the hydride transfer from C_4 to C_β (refer to Figure 6C). Altering the RPFs for the lower end of this C-H distance range might impact the characterization of all C-H bonds in the

system, potentially introducing large errors in the description of C-H interactions.

We also analyzed changes in distance distributions between reactant and the enolate (see Figure S1). The changes in the attractive region of the C-H distance distributions were similar. However, the population increase in the range between 2.9-3.2 Å was more moderate (see Figure S1B), which allowed us to modify the RPFs differentiating the corrections in the TS and the enolate region. Furthermore, the C-C distance distributions in the range 2.3-2.7 Å showed different patterns among the reactants, TS and enolate geometries (see Figures 6D and S1D), making this a target region for adjusting the relative energies among these stationary states.

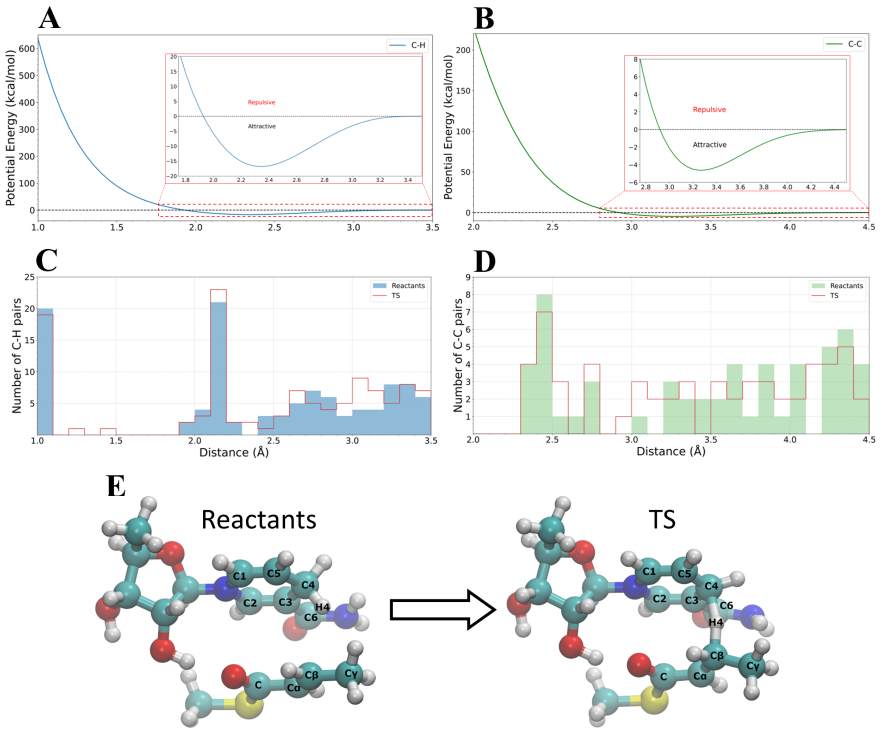

Figure 6: Analysis of DFTB3's C-H and C-C RPs: The repulsive potential functions (RPFs) are depicted for: **A)** C-H and **B)** C-C atom pairs. In **C)** and **D)** the C-H and C-C atom pair distance distributions for the reactants are compared to the TS geometry for Ccr's hydride transfer reaction. To facilitate the comprehension of the distance distributions represented in **C)** and **D)**, in **E)** the structures of the reactants and TS are shown, with labels at the reactive H4 hydrogen (white), the donor C4 and acceptor Cβ carbons (cyan) (oxygen atoms are shown in red, nitrogen blue and sulfur in yellow) .

Adjusting DFTB3's C-H and C-C RPs significantly improves the energetics of the hydride transfer reaction in Ccr

Based on the previous analysis of the C-H and C-C distance distributions of the reactant, transition state and enolate intermediate, we added local combination of harmonic functions (LCHFs) to the C-H and C-C RPFs contained in the 3OB-3-1 parametrization of the DFTB3 Hamiltonian to correct its deficiencies.

Our first objective was to make the attractive region of the C-H RPF's more repulsive, specifically the intervals between 2.1-2.7 Å and 2.9-3.2 Å shown in Figure 7A. For each tested configuration of the LCHFs, labeled as CH_X (with X=1 to 9; see Figure 7A and Table S2 for more details), we recalculated the potential energy along the reaction path for the modified DFTB3 Hamiltonian. As shown in Figure 7B, the addition of harmonic functions eliminates the local minimum at the transition state, approximating the activation barrier to the reference DFT method.

Although the activation barrier was significantly improved with the added harmonic functions to the C-H RPF's, the energy difference between the enolate intermediate and the reactant still deviated significantly from the reference M06-2X value (Figure 7B). Thus, we turned our focus to adjusting the C-C repulsive potential to achieve this effect. We added LCHFs on the C-C pair RPFs testing different combination of LCHFs for C-H RPF's that best reproduced the activation barrier. Figure 7C shows the C-C RPF (curve in green) and its step-wise modified versions as CC_X (with X=1,3,5,7,9; see Table S2 for more details). Following the analysis on C-C pairs distance distributions in the above section, we increased the repulsion in the C-C RPF's attractive region in a range including the interval 2.9-3.4 Å, combined with the reduction in the interval 2.3-2.7 Å. The modifications to the C-H and C-C RPFs are not independent, as they both influence the description of the transition state. While reducing the C-C RPF in the range 2.3-2.7 Å could potentially decrease the repulsion between $C4$ and $C\beta$ near the TS, this effect is moderated by the modifications made to

the C-H RPFs. Therefore, we focused on finding the optimal combination of C-H and C-C RPF modifications that would improve the overall energetics of the reaction, that is the relative energies between the reactant, transition state, and enolate intermediate. Among the possible combinations of the modified RPFs for C-H and C-C pairwise distances, we found that the configurations combining CH_4 with CC_1 to CC_9 were the ones showing more agreement in the potential energy profiles for the IRC geometries between the corrected DFTB3 and M06-2X/def2-TZVP Hamiltonians (see Figure 7D).

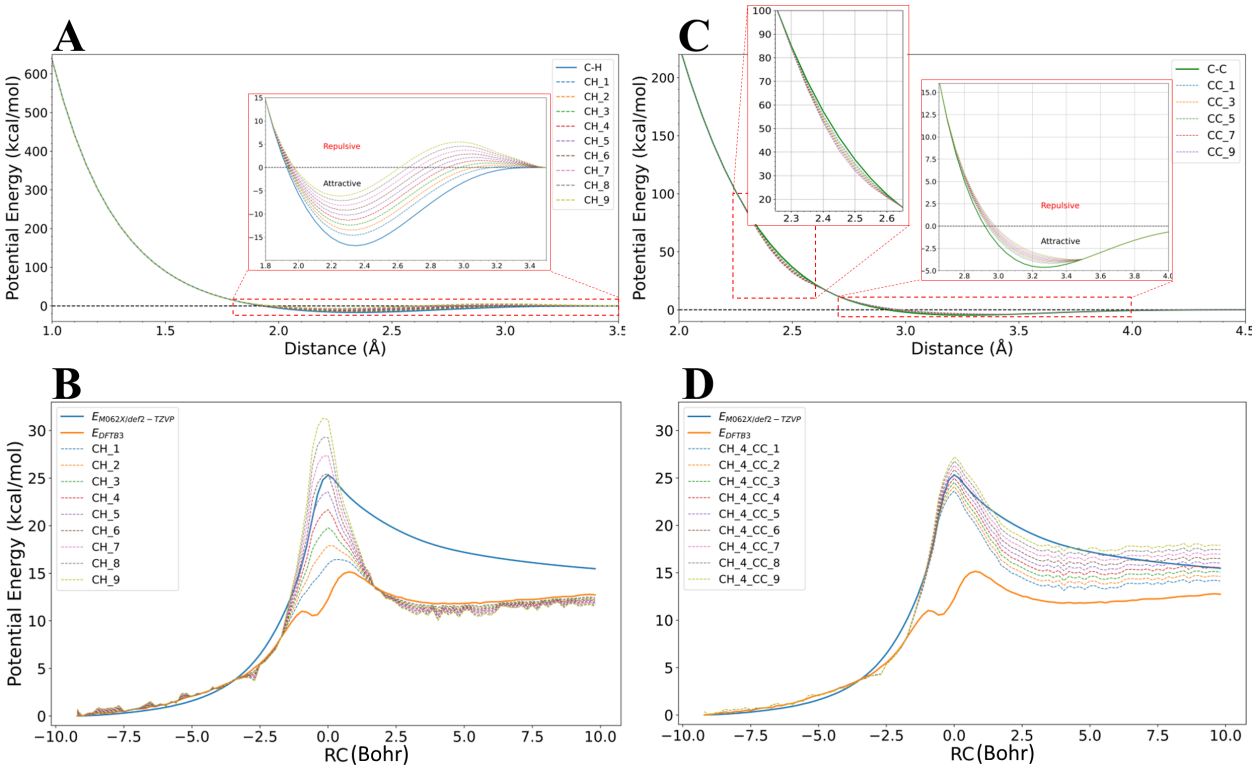

Figure 7: Step-wise corrections to DFTB3's RPFs for **A)** C-H and **C)** C-C pairs. The representations of the corrections to DFTB3's Hamiltonian along the IRC's reaction path are shown for the addition of LCHFs to C-H and C-C RPFs. In **B)** are represented the corrected DFTB3 profiles for the modified C-H RPFs. In **D)** are represented the corrected DFTB3 profiles obtained from the combination of one configuration of C-H RPF's modification from C) and the addition of LCHFs to C-C RPFs. As a reference in C) and D), is represented the IRC profile obtained at the M06-2X/def2-TZVP level of theory (curve in blue).

Validation of the improved DFTB3 Hamiltonian using QM/MM simulations in Ccr

To evaluate the improved DFTB3 Hamiltonian in a more realistic model of the enzymatic environment, we conducted constrained QM/MM geometry optimizations along the reaction coordinate for the hydride transfer reaction using a complete solvated model of the Ccr enzyme. We compared the results of various parameterizations at the DFTB3/CHARMM36 level with single-point calculations at the M06-2X/def2-TZVP/CHARMM36 level. Our calculations showed that, while the different C-H parameterizations agree in addressing the severe energy underestimation at the TS region in the QM/MM model, the CH_4 modification provided a better estimate of the activation energy compared to the DFT-corrected energy profile (see Figure S2). As expected, modifying the C-H RPFs does not correct the underestimation in the reaction energy as evidenced before by the IRC calculations. We then tested different combinations of the CH_4 parametrization with CC RPF's modifications, and while we were not able to fully correct the reaction energy compared to the reference calculations, we found that the combination of LCHFs CH_4 with CC_3 provided the most significant improvement in describing the activation barrier (see Figure S3), as discussed earlier for the IRC profile.

With the optimized C-H and C-C pair potentials we then calculated the free energy profile of the hydride transfer reaction in the Ccr enzyme using the modified DFTB3/CHARMM36 Hamiltonian. As reference, we carried out QM/MM molecular dynamics simulations with the adaptive string method at the M06-2X/def2-SVP/CHARMM36 level of theory. Figure 8 depicts the comparison between both free energy profiles (black and blue curves), as well as the corrected one by means of single point calculations at the M06-2X/def2-TZVP/CHARMM36 level of theory (curve in red). The new DFTB3 Hamiltonian (curve in blue) reproduces the reference activation barrier (curve in black) within 0.1 kcal/mol, anticipating the transition state along the reaction path. This result is validated by the DFT-based corrections on the

optimized DFTB3 profile, with a barrier of only 1.5 kcal/mol higher (see Figure 8), while at the enolate region the energies were up to 8.8 kcal/mol higher. Here, it is worth to mention that these deviations are similar to the ones in the QM/MM potential energy profiles (see Figure S3). On the other hand, the corrected profile in red showed good agreement with the profile obtained at the M06-2X/def2-SVP/CHARMM36 level of theory. The disagreements in these results could be explained by slight differences in geometries. As shown in Figure S4, although the geometries at the transition state were improved by the modifications made to the original DFTB3's RPFs (see the green curve in Figure S4), they still reproduce slightly shorter donor-acceptor C-C distances (CV3, Figure 3) compared to reference M06-2X/def2-SVP geometries.

As mentioned above, the main energy difference between the modified DFTB3 Hamiltonian and the DFT reference is observed for the formed enolate after passing the transition state. It is possible that the nature of the DFTB3 Hamiltonian is not able to describe the $NADP^+$ and enolate correctly, as they involve considerably charge separation. This has been reported to be a challenge for DFTB3 due to the use of a minimal basis sets, which significantly underestimates polarizability.[25] Modifications to include a larger basis set are beyond the scope of this work. However, this error can be easily corrected by adding a correction accounting for the difference in the electronic energy at higher DFT level.

The modified DFTB3 parameters are transferable to other hydride transfer reactions

A desired quality of a system-specific parametrization when using semiempirical Hamiltonians is that it can be transferable to other systems to describe similar chemical reactions. Thus, we evaluated the new CH_4_CC_3 parametrization on a small model system used to study hydride transfer reactions in organic chemistry, where the substrate exhibits distinct structural features compared to Ccr's substrate (see Figure S5A in the SI). This model

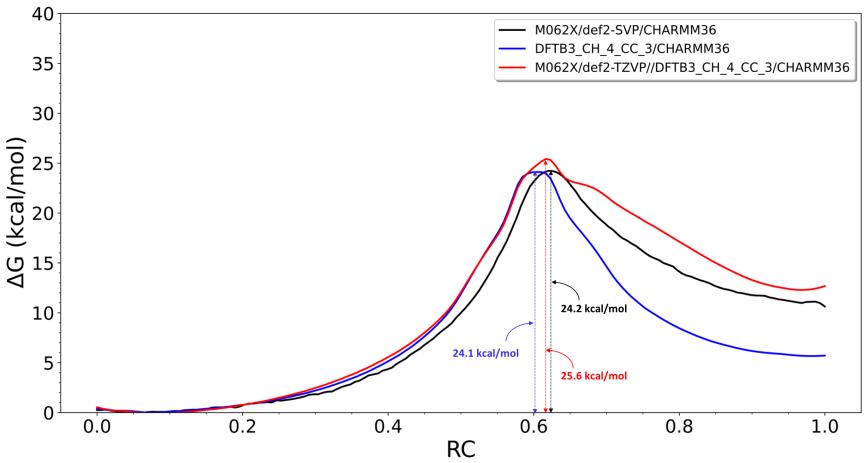

Figure 8: Comparison of the free energy profiles and activation barriers, obtained at the DFTB3_CH_4_CC_3/CHARMM36 (curve in blue) and M06-2X/def2-SVP/CHARMM36 (curve in black) levels of theory, for the hydride transfer reaction in the Ccr enzyme. In red is represented the corrected profile for the DFTB3_CH_4_CC_3/CHARMM36 Hamiltonian at the M06-2X/def2-TZVP/CHARMM36 level of theory.

involves a reaction between a derivative of 1,4-dihydropyridine, specifically diethyl-1,2,6-trimethyl-1,4-dihydropyridine-3,5-dicarboxylate (a member of the Hantzsch ester family), and the substrate benzyne.[60] The system was chosen for two main reasons: first, due to its structural similarities to NAD(P)H, where the 1,4-dihydropyridine ring represents the core of the molecule; and second, because it features distinct structural elements, such as an unsaturated substrate (benzyne) that is sufficiently different from crotonyl-CoA in Ccr. Our calculations reveal that the original DFTB3 parametrization also yields a severe underestimation of the barrier and the creation of a fictitious minimum near the transition state (TS). The new parametrization of the RPs resolves this issue, yielding a barrier height that closely matches DFT results (see Figure S5B and the Supporting Information for further details).

As a final test, we evaluated the transferability of our improved DFTB3 Hamiltonian to the hydride transfer reaction catalyzed by the Dihydrofolate Reductase (DHFR) enzyme. DHFR has become a model system for enzyme-catalyzed reactions,[61–70] making it an excellent choice for evaluating the new parametrization. QM/MM molecular dynamics simulations were carried out using the ASM at the DFTB3/CHARMM36 level of theory for the

original and optimized DFTB3 Hamiltonian. Figure 9 shows details on the reaction and ASM setup, as well as the free energy profiles obtained at the DFTB3/CHARMM36 level of theory (curves in blue) and the correction made using the Hamiltonian M06-2X/def2-TZVP/CHARMM36 on DFTB3 geometries (curve in red in Figure 9B). Similar to the case of Ccr, the free energy profile generated using DFTB3 with original 3OB-3-1 parameters severely underestimated the activation barrier and showed a minimum at the TS region, as evidenced by the corrections at higher level (see the blue and red curves in Figure 9B). On the other hand, our improved DFTB3 Hamiltonian (curve in blue in Figure 9C) provided an activation free energy of 16.8 kcal/mol that was in perfect agreement with previous theoretical reports for DHFR from *Escherichia coli*,[61,70] and experimental data (14.2-14.3 kcal/mol).[71,72] It is worth-noting that our calculations do not account for nuclear quantum effects, which are expected to lower the activation barrier and improve alignment with experimental results.[70]

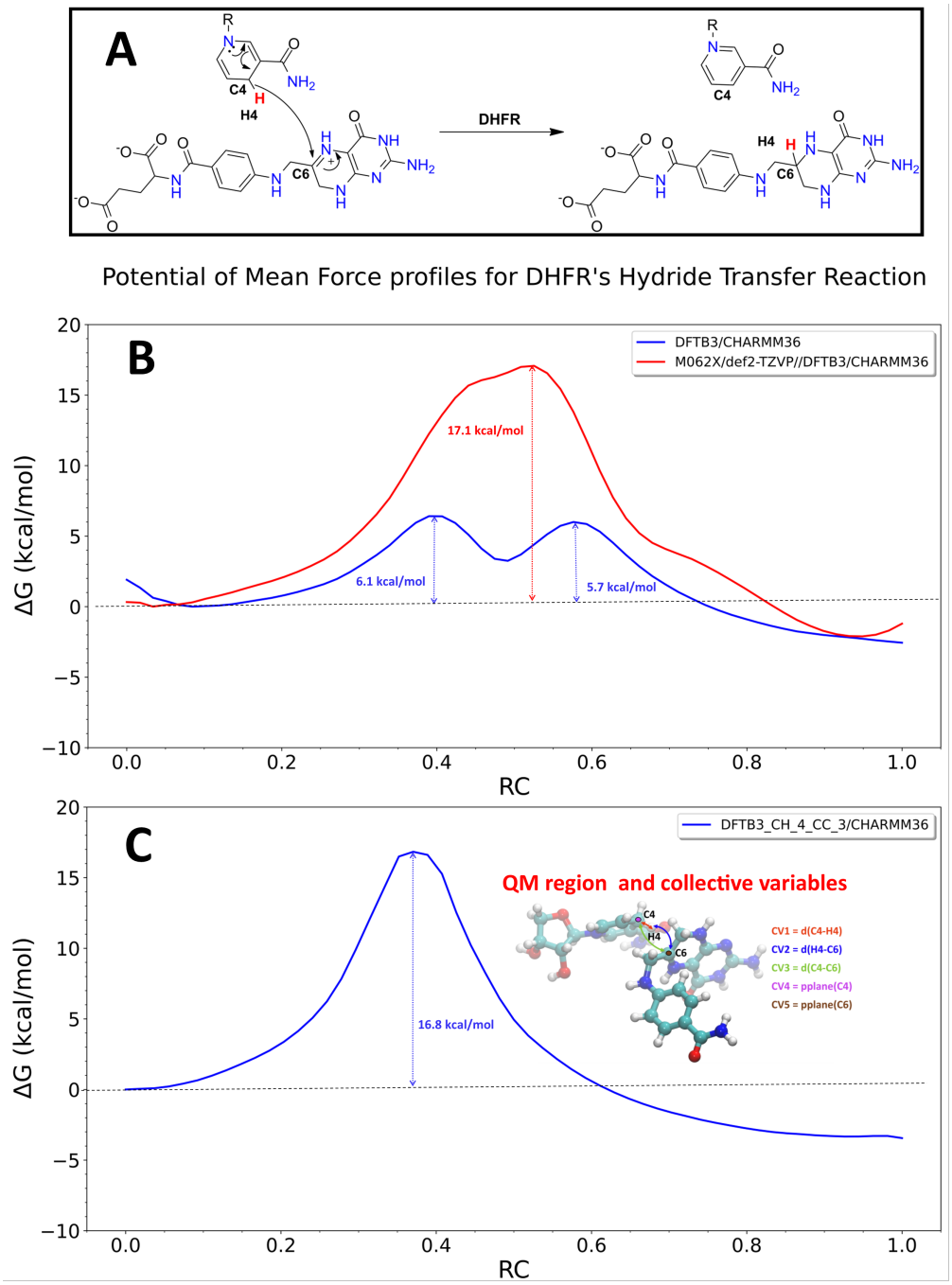

Figure 9: The hydride transfer reaction catalyzed by the DHFR enzyme: **A)** Simplified representation of the reaction mechanism. **B)** QM region's structural representation for the reactants, and the free energy profile and activation barriers computed at the DFTB3/CHARMM36 level of theory using DFTB3's original parameterization (in blue) and its corrected version at the M062X/def2-TZVP/CHARMM36 level of theory (in red). **C)** Free energy profile and activation barrier, obtained at the DFTB3/CHARMM36 level of theory using DFTB3's optimized parameter set (curve in blue)

Conclusions

Although DFTB3 is a widely used semiempirical method for studying (bio)organic systems and often serves as a viable alternative to DFT methods for QM/MM molecular dynamics simulations, it has limitations in accurately describing the energetics of hydride transfers. This limitation was demonstrated in the enzyme Ccr and subsequently verified in DHFR. To address this issue, we introduced a straightforward, generalizable approach based on analyzing C-H and C-C distance distributions, allowing for modifications to the CH and CC RPs within specific regions by using linear combinations of harmonic functions (LCHFs). This method effectively corrected the erroneous energetic descriptions. To our knowledge, this is the first systematic study to examine the effects of modifying C-H and C-C RPs to enhance the current DFTB3 parametrization (3OB-3-1 set of parameters) for hydride transfer reactions. The results of the new parametrization were validated against reference simulations at the DFT level.

The modified DFTB3 Hamiltonian significantly improved the description of the hydride transfer reaction in Ccr, reproducing the reference DFT activation barrier within 0.1 kcal/mol and providing a qualitatively correct free energy profile. Moreover, we demonstrated the transferability of our approach by applying the new DFTB3 parameters to the hydride transfer reaction in DHFR, where it correctly reproduced the activation free energy barrier without further adjustments, and a small model system bearing the 1,4-dihydropyridine motif but with distinct structural features. This suggests that these parameters are likely transferable, particularly within the family of NAD(P)H-dependent oxidoreductases. Our work represents a significant step forward in the accurate and efficient modeling of hydride transfer reactions in enzymatic systems using more efficient semiempirical methods for extended sampling. We expect that the improved DFTB3 parametrization will enable more reliable QM/MM molecular dynamics simulations, potentially accelerating research in enzyme engineering, drug design, and the development of bio-inspired catalysts.

Acknowledgement

J.L.V.-L. thanks the financial support of the project: "ANID/CONICYT FONDECYT Postdoctorado No.3240216". R.R. thanks the financial support of the project: "ANID/CONICYT FONDECYT Postdoctorado No.3210695". J.L.V.-L. and R.R. thanks NLHPC, Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02). E.V.M. acknowledges financial support by ANID FONDECYT REGULAR 1240345 and the Max-Planck Society. I.T. and J. J. R-P acknowledge financial support from grant PID2021-123332OB-C22 funded by MCIN/AEI/10.13039/501100011033/ and by "ERDF A way of making Europe". C.C. acknowledges support from ANID-Subdirección de Capital Humano/Doctorado Nacional/2021 Folio 21212208. The authors acknowledge the Computational center at the University of Valencia for computational facilities, including Tirant supercomputer. R.R. and J.L.V.-L thank Mirolad Andelkovic for his support in running string simulations.

Supporting Information Available

The Supporting Information includes detailed mathematical formulations of the GFs used in parameter optimization, extended analysis of DFTB3's deficiencies in describing Ccr's hydride transfer step, and thorough validation studies using QM/MM simulations represented in six supporting figures showing detailed analyses of RPs, comparative energy profiles, and reaction path analyses, along with two tables presenting activation energies for different functionals and basis sets, and numerical values for the modified parameters combining harmonic functions. Additionally, we shared in a GitHub repository (`https://github.com/josevlibera2010/DFTB3-CH-CC-RPs_For_HydrideTransfers`) the modified DFTB3 parameter sets for C-H and C-C RPs, along with the IRC output file and details on the ASM calculations performed.

References

(1) Sellés Vidal, L.; Kelly, C. L.; Mordaka, P. M.; Heap, J. T. Review of NAD(P)H-dependent oxidoreductases: Properties, engineering and application. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2018**, *1866*, 327–347.

(2) Gębicki, J.; Marcinek, A.; Zielonka, J. Transient Species in the Stepwise Interconversion of NADH and NAD+. *Acc. Chem. Res.* **2004**, *37*, 379–386, Publisher: American Chemical Society.

(3) Knaus, T.; Paul, C. E.; Levy, C. W.; de Vries, S.; Mutti, F. G.; Hollmann, F.; Scrutton, N. S. Better than Nature: Nicotinamide Biomimetics That Outperform Natural Coenzymes. *J. Am. Chem. Soc.* **2016**, *138*, 1033–1039, Publisher: American Chemical Society.

(4) Paul, C. E.; Gargiulo, S.; Opperman, D. J.; Lavandera, I.; Gotor-Fernández, V.; Gotor, V.; Taglieber, A.; Arends, I. W. C. E.; Hollmann, F. Mimicking Nature: Synthetic Nicotinamide Cofactors for C=C Bioreduction Using Enoate Reductases. *Org. Lett.* **2013**, *15*, 180–183, Publisher: American Chemical Society.

(5) Banerjee, S.; J. Sadler, P. Transfer hydrogenation catalysis in cells. *RSC Chemical Biology* **2021**, *2*, 12–29, Publisher: Royal Society of Chemistry.

(6) Winkler, C. K.; Faber, K.; Hall, M. Biocatalytic reduction of activated CC-bonds and beyond: emerging trends. *Current Opinion in Chemical Biology* **2018**, *43*, 97–105.

(7) Kumar Roy, T.; Sreedharan, R.; Ghosh, P.; Gandhi, T.; Maiti, D. Ene-Reductase: A Multifaceted Biocatalyst in Organic Synthesis. *Chemistry – A European Journal* **2022**, *28*, e202103949, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/chem.202103949.

(8) Toogood, H. S.; Scrutton, N. S. Discovery, Characterization, Engineering, and Applications of Ene-Reductases for Industrial Biocatalysis. *ACS Catal.* **2018**, *8*, 3532–3549, Publisher: American Chemical Society.

(9) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728, Publisher: American Chemical Society.

(10) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition* **2009**, *48*, 1198–1229, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.200802019.

(11) Dhoke, G. V.; Davari, M. D.; Schwaneberg, U.; Bocola, M. QM/MM Calculations Revealing the Resting and Catalytic States in Zinc-Dependent Medium-Chain Dehydrogenases/Reductases. *ACS Catal.* **2015**, *5*, 3207–3215, Publisher: American Chemical Society.

(12) Doron, D.; Major, D. T.; Kohen, A.; Thiel, W.; Wu, X. Hybrid Quantum and Classical Simulations of the Dihydrofolate Reductase Catalyzed Hydride Transfer Reaction on an Accurate Semi-Empirical Potential Energy Surface. *J. Chem. Theory Comput.* **2011**, *7*, 3420–3437, Publisher: American Chemical Society.

(13) Hernández-Ortega, A.; Ferreira, P.; Merino, P.; Medina, M.; Guallar, V.; Martínez, A. T. Stereoselective Hydride Transfer by Aryl-Alcohol Oxidase, a Member of the GMC Superfamily. *ChemBioChem* **2012**, *13*, 427–435, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cbic.201100709.

(14) Delgado, M.; Görlich, S.; Longbotham, J. E.; Scrutton, N. S.; Hay, S.; Moliner, V.; Tuñón, I. Convergence of Theory and Experiment on the Role of Preorganization, Quantum Tunneling, and Enzyme Motions into Flavoenzyme-Catalyzed Hydride Transfer. *ACS Catal.* **2017**, *7*, 3190–3198, Publisher: American Chemical Society.

(15) Gillet, N.; Ruiz-Pernía, J. J.; Lande, A. d. l.; Lévy, B.; Lederer, F.; Demachy, I.; Moliner, V. QM/MM study of L-lactate oxidation by flavocytochrome b2. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15609–15618, Publisher: The Royal Society of Chemistry.

(16) Ribeiro, A. J. M.; Yang, L.; Ramos, M. J.; Fernandes, P. A.; Liang, Z.-X.; Hirao, H. Insight into Enzymatic Nitrile Reduction: QM/MM Study of the Catalytic Mechanism of QueF Nitrile Reductase. *ACS Catal.* **2015**, *5*, 3740–3751, Publisher: American Chemical Society.

(17) Serapian, S. A.; Crosby, J.; Crump, M. P.; van der Kamp, M. W. Path to Actinorhodin: Regio- and Stereoselective Ketone Reduction by a Type II Polyketide Ketoreductase Revealed in Atomistic Detail. *JACS Au* **2022**, *2*, 972–984, Publisher: American Chemical Society.

(18) Cui, Q.; Pal, T.; Xie, L. Biomolecular QM/MM Simulations: What Are Some of the "Burning Issues"? *J. Phys. Chem. B* **2021**, *125*, 689–702, Publisher: American Chemical Society.

(19) Thiel, W. Semiempirical quantum–chemical methods. *WIREs Computational Molecular Science* **2014**, *4*, 145–157, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1161.

(20) Rossi, I.; Truhlar, D. G. Parameterization of NDDO wavefunctions using genetic algorithms. An evolutionary approach to parameterizing potential energy surfaces and direct dynamics calculations for organic reactions. *Chemical Physics Letters* **1995**, *233*, 231–236.

(21) Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. Direct dynamics calculations with NDDO (neglect of diatomic differential overlap) molecular orbital theory with specific reaction parameters. *J. Phys. Chem.* **1991**, *95*, 4618–4627, Publisher: American Chemical Society.

(22) Świderek, K.; Arafet, K.; Kohen, A.; Moliner, V. Benchmarking Quantum Mechanics/Molecular Mechanics (QM/MM) Methods on the Thymidylate Synthase-Catalyzed Hydride Transfer. *J. Chem. Theory Comput.* **2017**, *13*, 1375–1388, Publisher: American Chemical Society.

(23) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **1998**, *58*, 7260–7268, Publisher: American Physical Society.

(24) Sirirak, J.; Lawan, N.; Kamp, M. W. V. d.; Harvey, J. N.; Mulholland, A. J. Benchmarking quantum mechanical methods for calculating reaction energies of reactions catalyzed by enzymes. *PeerJ Phy. Chem.* **2020**, *2*, e8, Publisher: PeerJ Inc.

(25) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *7*, 931–948, Publisher: American Chemical Society.

(26) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *9*, 338–354, Publisher: American Chemical Society.

(27) Lu, X.; Gaus, M.; Elstner, M.; Cui, Q. Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. *J. Phys. Chem. B* **2015**, *119*, 1062–1082, Publisher: American Chemical Society.

(28) Gaus, M.; Jin, H.; Demapan, D.; Christensen, A. S.; Goyal, P.; Elstner, M.; Cui, Q. DFTB3 Parametrization for Copper: The Importance of Orbital Angular Momentum Dependence of Hubbard Parameters. *J. Chem. Theory Comput.* **2015**, *11*, 4205–4219, Publisher: American Chemical Society.

(29) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur

and Phosphorus for Chemical and Biological Applications. *J. Chem. Theory Comput.* **2014**, *10*, 1518–1537, Publisher: American Chemical Society.

(30) Elstner, M.; Seifert, G. Density functional tight binding. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2014**, *372*, 20120483, Publisher: Royal Society.

(31) Doemer, M.; Liberatore, E.; Knaup, J. M.; Tavernelli, I.; Rothlisberger, U. In situ parameterisation of SCC-DFTB repulsive potentials by iterative Boltzmann inversion. *Molecular Physics* **2013**, *111*, 3595–3607, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00268976.2013.842011.

(32) Ammothum Kandy, A. K.; Wadbro, E.; Aradi, B.; Broqvist, P.; Kullgren, J. Curvature Constrained Splines for DFTB Repulsive Potential Parametrization. *J. Chem. Theory Comput.* **2021**, *17*, 1771–1781, Publisher: American Chemical Society.

(33) Bodrog, Z.; Aradi, B.; Frauenheim, T. Automated Repulsive Parametrization for the DFTB Method. *J. Chem. Theory Comput.* **2011**, *7*, 2654–2664, Publisher: American Chemical Society.

(34) Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. Automatized Parametrization of SCC-DFTB Repulsive Potentials: Application to Hydrocarbons. *J. Phys. Chem. A* **2009**, *113*, 11866–11881, Publisher: American Chemical Society.

(35) Panosetti, C.; Engelmann, A.; Nemec, L.; Reuter, K.; Margraf, J. T. Learning to Use the Force: Fitting Repulsive Potentials in Density-Functional Tight-Binding with Gaussian Process Regression. *J. Chem. Theory Comput.* **2020**, *16*, 2181–2191, Publisher: American Chemical Society.

(36) Kranz, J. J.; Kubillus, M.; Ramakrishnan, R.; von Lilienfeld, O. A.; Elstner, M. Generalized Density-Functional Tight-Binding Repulsive Potentials from Unsupervised Ma-

chine Learning. *J. Chem. Theory Comput.* **2018**, *14*, 2341–2352, Publisher: American Chemical Society.

(37) Stöhr, M.; Medrano Sandonas, L.; Tkatchenko, A. Accurate Many-Body Repulsive Potentials for Density-Functional Tight Binding from Deep Tensor Neural Networks. *J. Phys. Chem. Lett.* **2020**, *11*, 6835–6843, Publisher: American Chemical Society.

(38) Erb, T. J.; Berg, I. A.; Brecht, V.; Müller, M.; Fuchs, G.; Alber, B. E. Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. *Proceedings of the National Academy of Sciences* **2007**, *104*, 10631–10636, Publisher: Proceedings of the National Academy of Sciences.

(39) Stoffel, G. M. M.; Saez, D. A.; DeMirci, H.; Vögeli, B.; Rao, Y.; Zarzycki, J.; Yoshikuni, Y.; Wakatsuki, S.; Vöhringer-Martinez, E.; Erb, T. J. Four amino acids define the CO2 binding pocket of enoyl-CoA carboxylases/reductases. *Proceedings of the National Academy of Sciences* **2019**, *116*, 13964–13969, Publisher: Proceedings of the National Academy of Sciences.

(40) Schwander, T.; Schada von Borzyskowski, L.; Burgener, S.; Cortina, N. S.; Erb, T. J. A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **2016**, *354*, 900–904, Publisher: American Association for the Advancement of Science.

(41) Recabarren, R.; Tinzl, M.; Saez, D. A.; Gomez, A.; Erb, T. J.; Vöhringer-Martinez, E. Covalent Adduct Formation as a Strategy for Efficient CO2 Fixation in Crotonyl-CoA Carboxylases/Reductases. *ACS Catal.* **2023**, *13*, 6230–6241, Publisher: American Chemical Society.

(42) Goyal, P.; Elstner, M.; Cui, Q. Application of the SCC-DFTB Method to Neutral and Protonated Water Clusters and Bulk Water. *J. Phys. Chem. B* **2011**, *115*, 6790–6805, Publisher: American Chemical Society.

(43) Riccardi, D.; Schaefer, P.; Yang,; Yu, H.; Ghosh, N.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. Development of Effective Quantum Mechanical/Molecular Mechanical (QM/MM) Methods for Complex Biological Processes. *J. Phys. Chem. B* **2006**, *110*, 6458–6469, Publisher: American Chemical Society.

(44) Young, T. A.; Silcock, J. J.; Sterling, A. J.; Duarte, F. autodE: Automated Calculation of Reaction Energy Profiles— Application to Organic and Organometallic Reactions. *60*, 4266–4274, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202011941.

(45) Frisch, M. J. et al. Gaussian 16 Revision B.01.

(46) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *132*, 154104.

(47) Werner, H.-J.; Knowles, P. J., et al. MOLPRO, version , a package of ab initio programs.

(48) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *WIREs Computational Molecular Science* **2012**, *2*, 242–253.

(49) Werner, H.-J. et al. The Molpro quantum chemistry package. *The Journal of Chemical Physics* **2020**, *152*, 144107.

(50) Saez, D. A.; Vogt-Geisse, S.; Vöhringer-Martinez, E. Ene-adducts from 1,4-dihydropyridines and α,β-unsaturated nitriles: asynchronous transition states displaying aromatic features. *Org. Biomol. Chem.* **2022**, *20*, 8662–8671, Publisher: The Royal Society of Chemistry.

(51) Case, D. A. et al. Amber 18. *University of California, San Francisco.* **2018**, *2018*.

(52) Zinovjev, K.; Tuñón, I. Adaptive Finite Temperature String Method in Collective Variables. *121*, 9764–9772, Publisher: American Chemical Society.

(53) Recabarren, R.; Zinovjev, K.; Tuñón, I.; Alzate-Morales, J. How a Second Mg2+ Ion Affects the Phosphoryl-Transfer Mechanism in a Protein Kinase: A Computational Study. *ACS Catal.* **2021**, *11*, 169–183, Publisher: American Chemical Society.

(54) Velázquez-Libera, J. L.; Caballero, J.; Alzate-Morales, J.; Ruiz-Pernía, J. J.; Tuñón, I. Insights into the Structural and Energetic Descriptions of Ubiquitin Specific Protease 7 (USP7) Catalytic Mechanisms by Hybrid QM/MM Simulations. *ChemCatChem* **2023**, *15*, e202300344, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cctc.202300344.

(55) Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón, I. Multiscale Simulations of SARS-CoV-2 3CL Protease Inhibition with Aldehyde Derivatives. Role of Protein and Inhibitor Conformational Changes in the Reaction Mechanism. *ACS Catal.* **2021**, *11*, 4157–4168, Publisher: American Chemical Society.

(56) A. Ramos-Guzmán, C.; Andjelkovic, M.; Zinovjev, K.; Javier Ruiz-Pernía, J.; Tuñón, I. The impact of SARS-CoV-2 3CL protease mutations on nirmatrelvir inhibitory efficiency. Computational insights into potential resistance mechanisms. *Chemical Science* **2023**, *14*, 2686–2697, Publisher: Royal Society of Chemistry.

(57) Kästner, J. Umbrella sampling. *WIREs Computational Molecular Science* **2011**, *1*, 932–942, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.66.

(58) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **1992**, *13*, 1011–1021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.540130812.

(59) Ramos-Guzmán, C. A.; Velázquez-Libera, J. L.; Ruiz-Pernía, J. J.; Tuñón, I. Testing Affordable Strategies for the Computational Study of Reactivity in Cysteine Proteases: The Case of SARS-CoV-2 3CL Protease Inhibition. *18*, 4005–4013, Publisher: American Chemical Society.

(60) Trinchera, P.; Sun, W.; Smith, J. E.; Palomas, D.; Crespo-Otero, R.; Jones, C. R. Intermolecular Aryne Ene Reaction of Hantzsch Esters: Stable Covalent Ene Adducts from a 1,4-Dihydropyridine Reaction. *Org. Lett.* **2017**, *19*, 4644–4647, Publisher: American Chemical Society.

(61) Luk, L. Y. P.; Javier Ruiz-Pernía, J.; Dawson, W. M.; Roca, M.; Loveridge, E. J.; Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J.; Tuñón, I.; Moliner, V.; Allemann, R. K. Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. *Proceedings of the National Academy of Sciences* **2013**, *110*, 16344–16349, Publisher: Proceedings of the National Academy of Sciences.

(62) Ruiz-Pernía, J. J.; Behiry, E.; Luk, L. Y. P.; Loveridge, E. J.; Tuñón, I.; Moliner, V.; Allemann, R. K. Minimization of dynamic effects in the evolution of dihydrofolate reductase. *Chem. Sci.* **2016**, *7*, 3248–3255, Publisher: The Royal Society of Chemistry.

(63) Luk, L. Y. P.; Loveridge, E. J.; Allemann, R. K. Protein motions and dynamic effects in enzyme catalysis. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30817–30827, Publisher: The Royal Society of Chemistry.

(64) Liu, C. T.; Francis, K.; Layfield, J. P.; Huang, X.; Hammes-Schiffer, S.; Kohen, A.; Benkovic, S. J. Escherichia coli dihydrofolate reductase catalyzed proton and hydride transfers: Temporal order and the roles of Asp27 and Tyr100. *Proceedings of the National Academy of Sciences* **2014**, *111*, 18231–18236, Publisher: Proceedings of the National Academy of Sciences.

(65) Hammes-Schiffer, S.; Watney, J. B. Hydride transfer catalysed by Escherichia coli and

Bacillus subtilis dihydrofolate reductase: coupled motions and distal mutations. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2006**, *361*, 1365–1373, Publisher: Royal Society.

(66) Ruiz-Pernía, J. J.; Tuñón, I.; Moliner, V.; Allemann, R. K. Why Are Some Enzymes Dimers? Flexibility and Catalysis in Thermotoga maritima Dihydrofolate Reductase. *ACS Catal.* **2019**, *9*, 5902–5911, Publisher: American Chemical Society.

(67) Angelastro, A.; Ruiz-Pernía, J. J.; Tuñón, I.; Moliner, V.; Luk, L. Y. P.; Allemann, R. K. Loss of Hyperconjugative Effects Drives Hydride Transfer during Dihydrofolate Reductase Catalysis. *ACS Catal.* **2019**, *9*, 10343–10349, Publisher: American Chemical Society.

(68) Roston, D.; Kohen, A.; Doron, D.; Major, D. T. Simulations of remote mutants of dihydrofolate reductase reveal the nature of a network of residues coupled to hydride transfer. *Journal of Computational Chemistry* **2014**, *35*, 1411–1417, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.23629.

(69) Mhashal, A. R.; Vardi-Kilshtain, A.; Kohen, A.; Major, D. T. The role of the Met20 loop in the hydride transfer in Escherichia coli dihydrofolate reductase. *Journal of Biological Chemistry* **2017**, *292*, 14229–14239, Publisher: Elsevier.

(70) Doron, D.; Stojković, V.; Gakhar, L.; Vardi-Kilshtain, A.; Kohen, A.; Major, D. T. Free Energy Simulations of Active-Site Mutants of Dihydrofolate Reductase. *J. Phys. Chem. B* **2015**, *119*, 906–916, Publisher: American Chemical Society.

(71) Cameron, C. E.; Benkovic, S. J. Evidence for a Functional Role of the Dynamics of Glycine-121 of Escherichia coli Dihydrofolate Reductase Obtained from Kinetic Analysis of a Site-Directed Mutant. *Biochemistry* **1997**, *36*, 15792–15800, Publisher: American Chemical Society.

(72) Swanwick, R. S.; Maglia, G.; Tey, L.-h.; Allemann, R. K. Coupling of protein motions and hydrogen transfer during catalysis by Escherichia coli dihydrofolate reductase. *Biochemical Journal* **2006**, *394*, 259–265.

TOC Graphic