

Article

Rational Engineering of (S)-Norcoclaurine Synthase for Efficient Benzyloquinoline Alkaloids Biosynthesis

João P. M. De Sousa , Nuno C. S. A. Oliveira  and Pedro A. Fernandes 

LAQV/REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal

* Correspondence: pafernan@fc.up.pt

Abstract: (S)-Norcoclaurine is synthesized in vivo through a metabolic pathway that ends with (S)-norcoclaurine synthase (NCS). The former constitutes the scaffold for the biosynthesis of all benzyloquinoline alkaloids (BIAs), including many drugs such as the opiates morphine and codeine and the semi-synthetic opioids oxycodone, hydrocodone, and hydromorphone. Unfortunately, the only source of complex BIAs is the opium poppy, leaving the drug supply dependent on poppy crops. Therefore, the bioproduction of (S)-norcoclaurine in heterologous hosts, such as bacteria or yeast, is an intense area of research nowadays. The efficiency of (S)-norcoclaurine biosynthesis is strongly dependent on the catalytic efficiency of NCS. Therefore, we identified vital NCS rate-enhancing mutations through the rational transition-state macrodipole stabilization method at the Quantum Mechanics/Molecular Mechanics (QM/MM) level. The results are a step forward for obtaining NCS variants able to biosynthesize (S)-norcoclaurine on a large scale.

Keywords: enzyme catalysis; protein engineering; alkaloid biosynthesis; benzyloquinoline alkaloids; (S)-norcoclaurine



Citation: De Sousa, J.P.M.; Oliveira, N.C.S.A.; Fernandes, P.A. Rational Engineering of (S)-Norcoclaurine Synthase for Efficient Benzyloquinoline Alkaloids Biosynthesis. *Molecules* **2023**, *28*, 4265. <https://doi.org/10.3390/molecules28114265>

Academic Editors: Elisa Nuti, Filipe Menezes and Grzegorz Popowicz

Received: 23 April 2023

Revised: 11 May 2023

Accepted: 20 May 2023

Published: 23 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

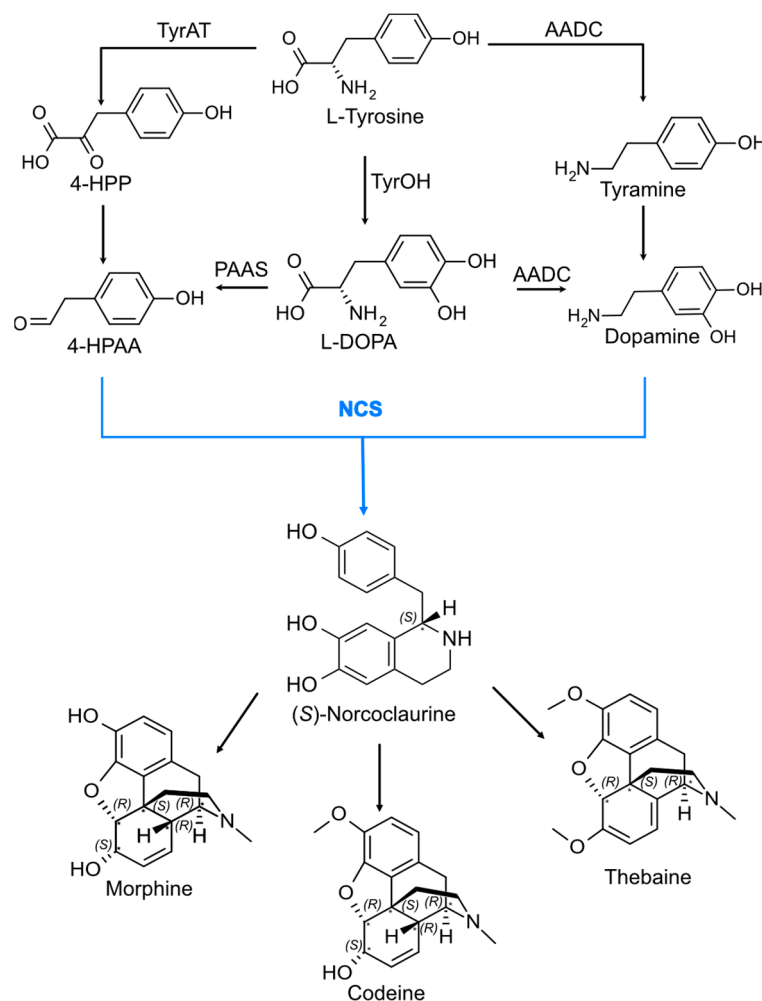
1. Introduction

The BIAs are one of the most relevant drug groups for chronic pain management, whose most well-known members include the opiates morphine, codeine, and thebaine and the semi-synthetic opioids oxycodone, hydrocodone, and hydromorphone [1]. These drugs are listed as essential medicines by the World Health Organization for their utility in managing chronic pain and palliative care [2]. Despite this, more than one-half of the world's population has low to nonexistent access to treatment for moderate or severe pain [3]. The only source of these drugs or their synthetic precursors is the opium poppy (*Papaver somniferum*) harvest, which depends on the highly-susceptible opium poppy crops, introducing a risk factor in their worldwide supply [4,5]. Biotechnology-based forms of BIAs production arising from the poppy's enzymes usage in solution or heterologous hosts can be one solution for diversifying BIAs production methods [6–10].

Currently, the enzyme-based alternatives for diversifying production are mostly focused on using either *Escherichia coli* or *Saccharomyces cerevisiae* as microbial hosts for the synthesis of BIAs [11]. The most prominent development of BIAs production was achieved by Pyne and coworkers in *S. cerevisiae* using a strain capable of synthesizing (S)-reticuline to titers up to the 4.6 g L^{−1} [8], an improvement of 57,000-fold compared to previous efforts [10]. The pathway engineering was performed by manipulating both endogenous and exogenous genes to maximize carbon flux into BIAs biosynthesis and minimize the formation of fusel products.

The early stages of naturally occurring opioids in the opium poppy are shown in Scheme 1. BIAs downstream products have a high degree of three-dimensional structure complexity and might have up to five stereocenters, as in the case of morphine or codeine. Nonetheless, these complex molecules arise from the condensation of two equivalents of

L-tyrosine derivatives, dopamine, and 4-hydroxyphenylacetaldehyde (4-HPAA). Endogenous L-tyrosine can be first hydroxylated to L-DOPA and then *P. somniferum*'s aromatic amino acid decarboxylase (AADC) decarboxylates L-DOPA to synthesize dopamine. Alternatively, the *P. somniferum* AADC can also decarboxylate L-tyrosine to tyramine, which can be further hydroxylated to dopamine. Another L-tyrosine equivalent is necessary to synthesize 4-HPAA; an L-tyrosine aminotransferase (TyrAT) converts the L-tyrosine residue to 4-hydroxyphenylpyruvate (4-HPP), which is then decarboxylated to yield 4-HPAA. Besides the naturally occurring opioids, (S)-norcoclaurine is also the precursor to other bioactive BIAs such as magnoflorine, sanguinarine, and noscapine [1].



Scheme 1. Early stages of the biosynthesis of naturally occurring opioids; molecular representation of (S)-norcoclaurine and some of the most widely used downstream opioids with potent analgesic properties; TyrOH-L-tyrosine hydroxylase; AADC—aromatic amino acid decarboxylase; TyrAT-L-tyrosine aminotransferase; PAAS-phenylacetaldehyde synthase; NCS-(S)-norcoclaurine synthase. The stereocenters are highlighted with an “*”. A more detailed pathway can be found in Ref. [1].

Next to the derivatization of the two L-tyrosine equivalents, NCS comes into play, one of the most relevant enzymes in the BIAs biosynthesis pathway. This enzyme is responsible for the enantioselective condensation of dopamine and 4-HPAA (Scheme 1). Since NCS has two substrates, the order of substrate binding to the active site pocket influences the outcome of the reaction. X-ray structural evidence deposited under the PDB accession code 5NON led to a general acceptance that dopamine is the first substrate to enter deep within the active site of NCS, followed by 4-HPAA which lodges at the active site entrance. This substrate binding order is corroborated by NCS's high substrate

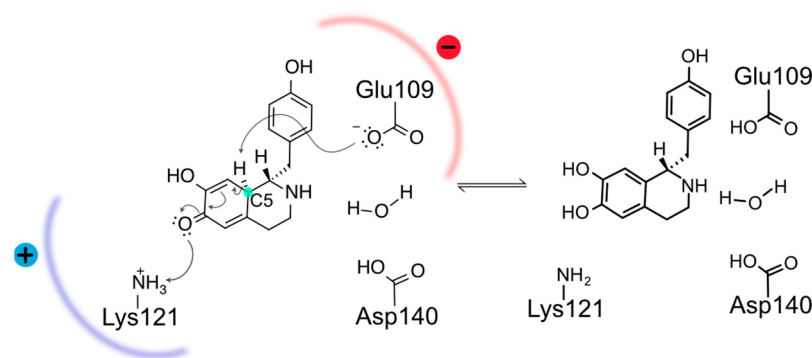
promiscuity towards the aldehyde substrate and the fact that a catalytic lysine residue establishes an interaction with the catechol moiety of the bound mimic, demonstrating this residue's crucial activity as a proton acceptor during NCS's reaction mechanism [12]. Other substrate binding poses might exist within the enzyme active site, at certain time periods, as previous crystallographic structures have demonstrated (PDB accession code: 2VQ5), but these alternative poses might lead to unreactive enzyme:substrate complexes and therefore dissociate [13].

The demonstrated promiscuity of NCS opens the possibility of engineering NCS variants with enhanced activity towards synthesizing (S)-norcoclaurine [14–16].

Furthermore, NCS is considered as a key bottleneck in BIAs biosynthesis once a drop in titer was reported from dopamine to (S)-norcoclaurine synthesis (from mg L^{-1} to $\mu\text{g L}^{-1}$) [10].

The viability of heterologous production of BIAs on an industrial scale is dependent on the improvement of the catalytic efficiency of the pathway enzymes to achieve viable product concentrations, space–time yields, and catalyst productivity, which in turn determine the commercial viability of the method [17]. Being NCS the initial branchpoint and a significant bottleneck in the BIAs biosynthesis pathway, it is one of the most appealing targets for biocatalyst improvement [18,19].

Previous structural and computational studies have dissected the intricacies of *Thalictrum flavum* NCS (TfNCS) substrate binding, reaction mechanism, and enantioselectivity [12,20]. The complete cycle involves five transition states, from which the last and rate limiting is the rearomatization triggered by the deprotonation of the quinonoid intermediate (QIN) at the C5 position, which leads to the protonation of the quinone moiety by the catalytic Lys residue (Scheme 2).



Scheme 2. Rearomatization step of NCS reaction mechanism. The negative Glu109 and positive Lys121 sidechains form a dipole within the enzyme active site at the rate-determining intermediate state that disappears in the product of this step.

In this work, we have used the ‘own *n*-layered integrated molecular orbital and molecular mechanics’ (ONIOM) QM/MM method [21] to calculate the contribution of the residues outside the active site for the activation energy of the rearomatization step, as was previously performed for other enzymes with biotechnological applications [22,23]. Our approach began with simulating the quinonoid intermediate (QIN) rearomatization step for the wild-type *Papaver somniferum* NCS (*PsNCS*) through a linear transit scan of the QIN-H•••OOC-Glu109 coordinate. We then fully optimized the reactant (R) and transition state (TS) stationary points for this specific step and calculated the energy barrier for the wild-type *PsNCS* (ΔE_{wt}). All the details concerning the modeling of *PsNCS* and building the QM/MM model can be found in the methods section. Next, a point deletion of each residue from the MM layer in both R and TS was performed, and the associated energy barrier (ΔE_{m}) was calculated. The magnitude of each residue's influence in the energy barrier ($\Delta\Delta E$) is given by the difference in energy between the ΔE_{m} and ΔE_{wt} ($\Delta\Delta E_{\text{m-wt}} = \Delta E_{\text{m}} - \Delta E_{\text{wt}}$). We propose that the experimental procedure to design NCS variants with enhanced first-

order reaction rates should focus on residues that raise the wild-type barrier, which can be identified by the ones that retrieve a negative $\Delta\Delta E_{m-wt}$ value.

2. Results and Discussion

The generated conformations from the MD simulations reveal an average RMSD of 1.70 Å for the protein backbone and 0.69 Å for the active site (see Figure S4). Therefore, we can infer that QIN and the modeled catalytic water molecule do not move away from their positions in the active site and that the overall structure of *PsNCS* is stable. This is not only evidence of a well-constructed enzyme model but also of a realistic docking pose for reaction to happen.

We have also monitored the variations in the length of the two most important interactions for the reaction: Lys121-NH3⁺ •••O=C-QIN and Glu109-COO[−] •••HC-QIN. The collected data indicate that these are well conserved through the course of the MD simulation (Figures 1 and 2). Regarding the interaction with Glu109, it is interesting to note that its carboxylate side chain has no conformational freedom to switch around the oxygen atoms, as depicted by the graphic. We see that one of them is always closer to QIN's leaving proton than the other, with an average distance of 2.99 Å against one of 4.80 Å (Figure 1). This is probably due to the hydrogen bonding network established in the active site, which may induce it to adopt specific conformations to favor the course of the reaction. As for the interaction with Lys121, the almost superimposable graphics lead us to deduce that there is no specificity in this protonation. The amino group side chain of Lys121 has the torsional freedom to flip, so all the protons are in range for hydrogen bond formation with the carbonyl group of QIN, with average distances between 2.98 and 3.16 Å (Figure 2).

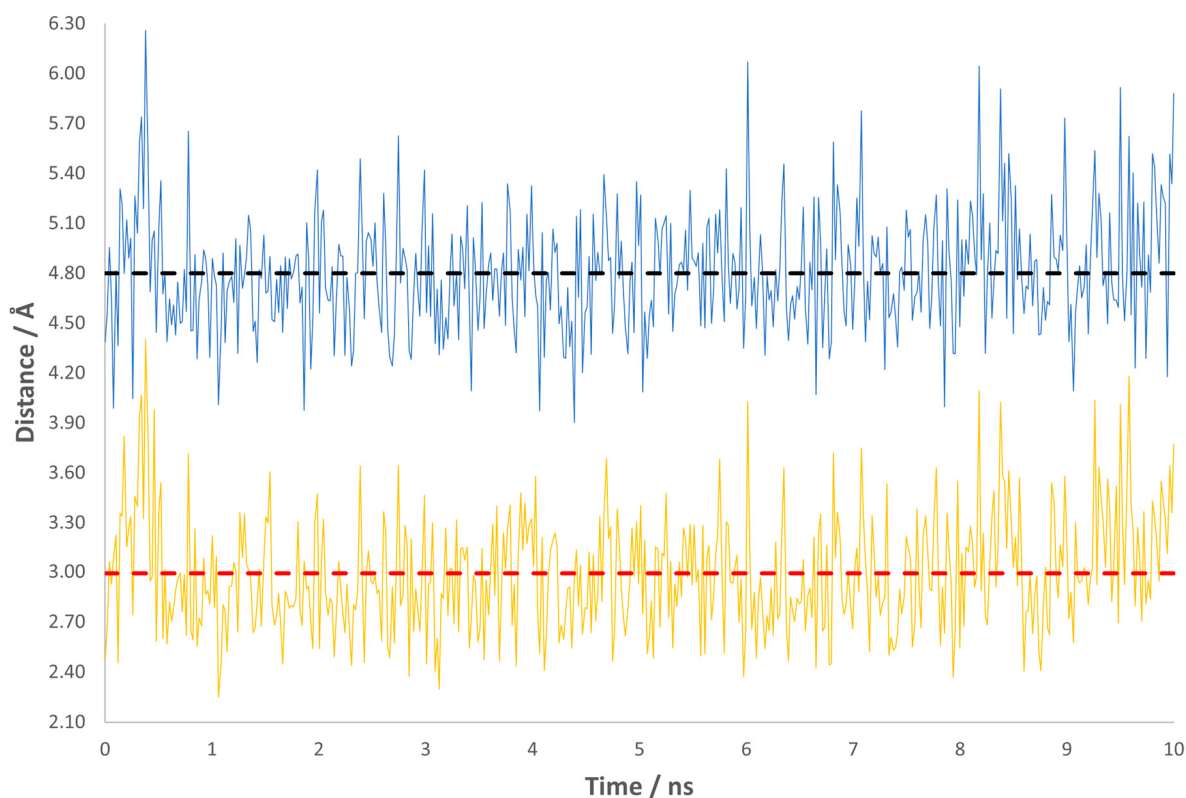


Figure 1. Variations in the distance for the interaction Glu109-COO[−] •••HC-QIN throughout the production stage of the MD simulation. The yellow and blue lines represent each oxygen atom from the carboxylate side chain of Glu109. The dashed lines show the average distance for each interaction: the red dashed line marks 2.99 Å and the black one 4.80 Å.

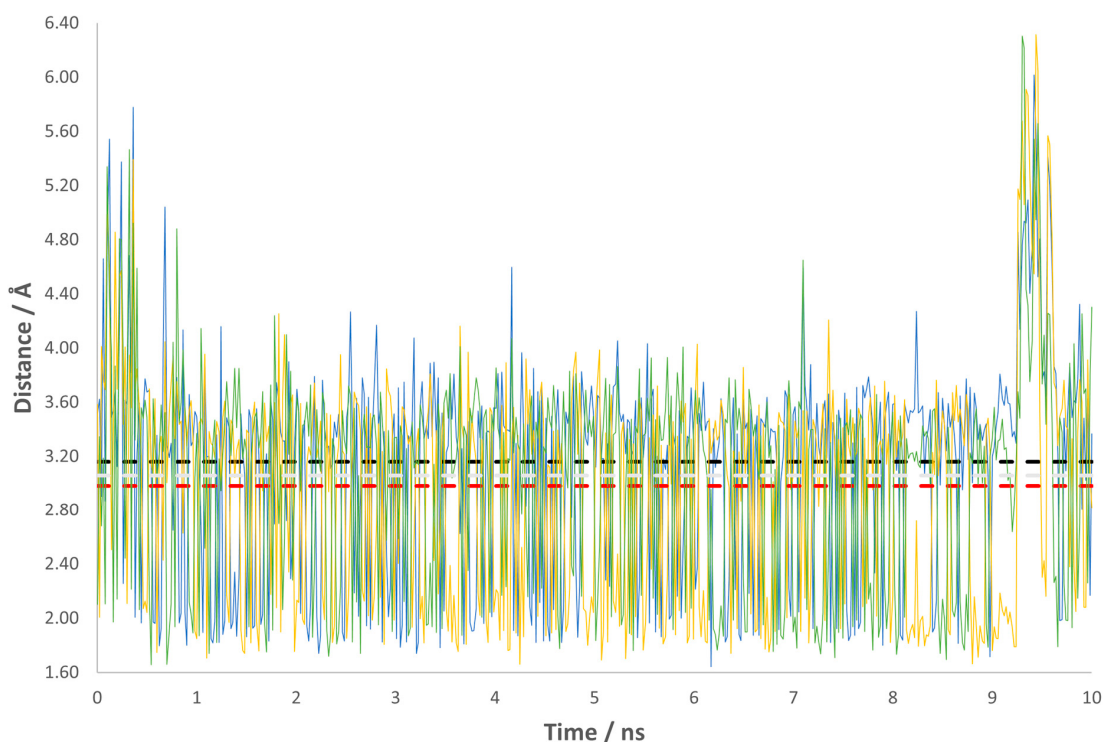


Figure 2. Variations in the distance for the interaction $\text{Lys121-NH}_3^+ \cdots \text{O}=\text{C-QIN}$ throughout the production stage of the MD simulation. The yellow, green and blue lines represent each proton from the amino group side chain of Lys121. The dashed lines mark the average distance for each interaction: red dashed line $\rightarrow$ 2.98 Å; light gray dashed line $\rightarrow$ 3.06 Å; black dashed line $\rightarrow$ 3.16 Å.

As the MD simulations show that our model maintains the distances involved in the QIN reactivity, we selected the structure obtained after the phased minimization protocol to proceed to the simulation of the rate limiting step through QM/MM.

The overall orientation of the active site was maintained after initial QM/MM geometry optimizations, even though a slight approximation of QIN towards Lys121 was noticed (Figure 3).

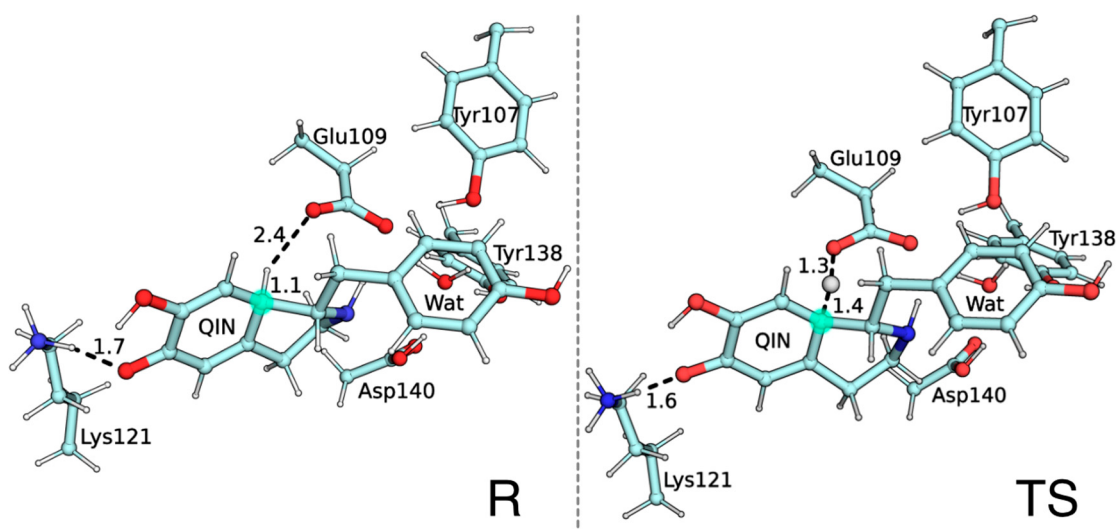


Figure 3. R and TS structures for the QM layer atoms. Distances between atoms participating in the rearomatization step are displayed in Å. Carbon C5 is marked in light green.

As expected, the TS of the rate-limiting rearomatization step is characterized by the concerted and asynchronous proton transfer triggered by QIN deprotonation at the C5 position.

A closer look at the TS geometry confirms the stepwise fashion of the proton transfers occurring in QIN as the C5 proton is halfway through the QIN-H $\cdots$ $\cdots$ ⁻OOC-Glu109 coordinate. In contrast, the bond between the Lys121 amino N and its proton is not yet broken (Figure 3).

The single-point energy calculations employing the ONIOM(B3LYP-D3/6-311+G(2d,2p): amber theoretical level retrieved a ΔE_{wt} for PsNCS of 10.6 kcal mol⁻¹ in very good agreement (~2 kcal mol⁻¹ lower) with the TfNCS value predicted by Himo and Sheng [20].

The influence of the deletion of each residue on the barrier is shown in Figure 4. As the transformation of the R into the TS involves a partial neutralization from the Glu109 and the development of a negative charge at the QIN oxygen atom in the *para* position (QIN_{OX}; Lys121 is still positive at the TS), all residues that stabilize the negative charge at Glu109 more extensively than at QIN_{OX} will over stabilize the R in relation to the TS, generating an anticatalytic effect. Therefore, these shall be the principal target residues for rate-enhancing mutations. Thus, negative residues closer to QIN_{OX} than to Glu109 or positive residues closer to Glu109 than to QIN_{OX} are the ideal targets for mutation. This principle is confirmed by the obtained results, highlighted in Figure 4.

In summary, the charged residues that raise the barrier for more than 1 kcal mol⁻¹ and thus are the ideal candidates to test for experimental mutagenesis are Asp64, Asp71, Arg105, Lys110, Glu117, and Lys163. The large barrier-lowering effect induced by the deletion of Tyr60 is due to the loss of an ionic interaction with the Lys121 amino group and Tyr60 backbone carbonyl, which destabilizes the reactant. However, for mutation purposes, Tyr60 is not an ideal target because it is an important residue for the previous steps of the NCS mechanism.

The results in Figure 4 hold if the enzyme does not undergo significant structural rearrangements after introducing the mutations. However, experience shows that many mutations cause structural rearrangements. Thus, ideal mutations cancel the side chains charge but retain most of the other characteristics of the residues, such as Glu/Gln, Asp/Asn, or Lys/Gln.

In general, mutating a charged residue for a polar residue is not highly disruptive because the environment around the charged residue is highly polar, responding well to the mutation. Conversely, mutation of a charged/polar residue by a hydrophobic residue, and vice versa, is highly disruptive because the surrounding environment is not tuned for the mutation.

However, in NCS, the charged residues identified as favorable for mutation, although not far from the active center, are located on the enzyme's surface. Moreover, their charged groups are exposed to the solvent and free of hydrogen bridges with other residues. This is an exceptionally favorable situation for mutation because of the low disruptive propensity of the mutations. The function of these residues likely is to confer adequate solubility to the NCS, so their mutation by polar or oppositely charged residues seems amenable from a structural perspective.

In sum, the proposed mutations are not disruptive nor are significant changes in the protein scaffold expected.

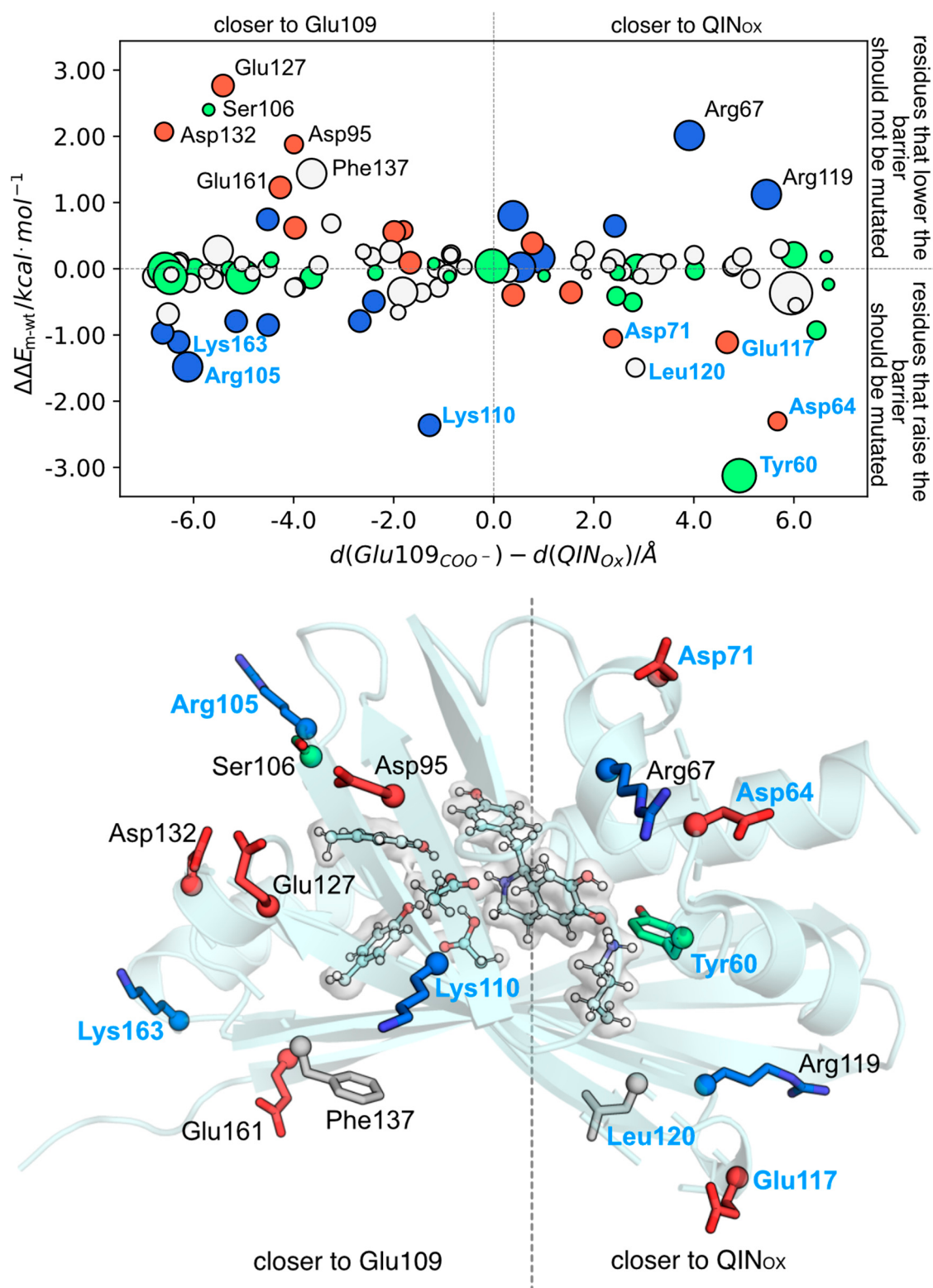


Figure 4. (Top) $\Delta\Delta E_{m-wt}$ values for each NCS mutant as a function of the difference of the distances from each residue's: (1) charged center (for Glu/Asp/Lys/Arg) or (2) side chain center of mass, to either QIN_{ox} or Glu109_{COO}[−]. White circles- hydrophobic residues; green circles- polar residues; blue circles- positive residues; red circles- negative residues. Circle size is correlated with residue size; (Bottom) Cartoon representation of NCS with the residues whose deletion induced a higher $|\Delta\Delta E|$ in sticks. Residues labeled in blue correspond to the ones whose mutation lowers the barrier.

3. Methods

The computational protocol in this work can be described as follows: (a) construction of the PsNCS:QIN complex through homology modeling and molecular docking; (b) MD simulations to stabilize the modeled structure (50 + 100 ps) and to briefly sample the conformational space of the system (10 ns); (c) QM/MM calculation of the reactant, transition state, and product of the last, rate-limiting, step of the reaction; and (d) calculation and rationalization of the influence of the residues surrounding the active site on the activation energy and prediction of specific mutations that increase the reaction rate.

3.1. Construction of the Psncs:Qin Complex

To the date of this work, there were no solved crystallographic structures of PsNCS available at the Protein Data Bank. To overcome this, we built a PsNCS structure through homology modeling, using SWISS-MODEL [24–28]. After choosing a PsNCS target sequence (UniProtKB accession number Q4QTJ1), the SWISS-MODEL template library (SMTL version 05/05/2021, PDB release 30/04/2021) was searched with BLAST [29] and Hhblits [30,31] for evolutionary related structures matching the target sequence. From the results gathered by this sequence alignment-based search, we chose a recent truncated *Tf*NCS crystallographic structure (PDB ID 6RP3) as the template [32]. It was obtained by X-ray crystallography with a resolution of 1.81 Å and displayed a sequence identity of 53.70%. These are already signs of a potentially precise model, confirmed with our subsequent structure validation tests (results shown in SI). For internal structure validation, we created a Ramachandran plot of our model using MolProbity (Figure S1) [33–35]. For external structure validation, we evaluated its tertiary structure (folding) by comparing its structural motifs with the ones found in native proteins of identical size, using the estimators QMEAN, from SWISS-MODEL, and ProSaII, from ProSa-Web (Figure S2) [36,37]. We added the hydrogen atoms to our structure and solvated the enzyme with a cubic box of TIP3P water molecules [38], using the xLEaP program from Amber18 [39]. The system was then subjected to a four-step minimization protocol. In the first step, the water molecules were minimized, while harmonic potentials restrained the other atoms of the system. Afterward, we sequentially allowed the hydrogen atoms, the residue's side chains, and the whole system to move freely. In each minimization, the restraint force constant applied to the fixed atoms was 50.0 kcal mol^{−1} Å^{−2}.

The template structure had a reaction intermediate mimic from which we built the QIN ligand with the GaussView software [40]. For the protein–ligand docking, the solvation box had to be removed from the system. The technique was performed with vsLab [41], a VMD [42] plug-in for molecular docking procedures using the AutoDock software [43–45]. For the search stage, we set up a box comprising the area of the active site and defined the number of possible results to 15, using the default options for the genetic algorithm. The scoring stage was carried out with the AutoDock scoring function. From the resulting four poses (A to D, see Figure S3), we chose to proceed with complex B because the position and interactions of QIN were the most appropriate for the reaction to happen.

The parameters for QIN were determined from an optimized structure at the HF/6-31G(d) level in vacuum. We used the antechamber program, from AmberTools18 [39], with the general AMBER force field (GAFF) [46] to derive the intramolecular and Lennard–Jones parameters, and the restrained electrostatic surface potential (RESP) method [47] to derive atomic charges at the HF/6-31G(d) level of theory. We manually added a water molecule to the active site and protonated the carboxylate side chain of the residue Asp140 because these characteristics are considered relevant for the reaction's mechanism [20]. The AMBER *ff*14SB force field [48] was used to parametrize the remaining atoms of the enzyme. To prepare the system for the simulations, a new cubic box of TIP3P water molecules [38] was added to solvate the enzyme within 10 Å of the solute, as well as 2 Na⁺ and 1 Cl[−] counterions to neutralize the system's charge. These procedures were carried out with xLEaP, from Amber18. This system was subjected to a five-step minimization protocol, with the same restraint force constant for the fixed atoms. In the first step, only the manually added water molecule was allowed to move. The remaining water molecules, the hydrogen

atoms, the residue's side chains, and the whole system were sequentially minimized afterward. Both minimizations were carried out with the *sander* module of Amber18.

3.2. MD Simulations

The simulated system had 36,079 atoms. The SHAKE algorithm [49] was used in all stages of the simulation to constrain all bonds involving hydrogen atoms, allowing the usage of an integration time step of 2 fs. The particle mesh Ewald (PME) method was used to calculate the long-range interactions, whereas short-range interactions are described with Lennard–Jones potentials, using a 10 Å cutoff.

A 20 ps progressive heating was performed linearly from 0 to 310.15 K, followed by 30 ps at 310.15 K in the canonical ensemble (NVT). The system was equilibrated at that temperature for 100 ps in the isothermal-isobaric ensemble (NPT, 310.15 K, 1 bar), using the Langevin thermostat with a collision frequency of 2 ps^{−1} and isotropic position scaling. A production stage was then run in the NPT ensemble for 10 ns to analyze the conformations of the active site and the stability of the protein backbone, in the same regulation conditions. Heating and equilibration stages were run with the *sander* module of Amber18, and PMEMD.CUDA [50–52] was used for the production stage. The presented root-mean-square deviation (RMSd) values were calculated with CPPTRAJ [53] from AmberTools18. Data on the variations in the length of interactions were acquired from VMD.

3.3. Oniom Model Details

The QM/MM model was defined from the minimized structure obtained after the docking procedure. The counterions and most of the solvation box were removed, keeping a solvation shell with all water molecules within 7 Å of the enzyme, ensuring that all the outer residue's side chains were solvated. The final model had a total of 6706 atoms and an overall charge of −1 au. The QM layer, treated at the density functional theory level, comprised 107 atoms, including the QIN substrate, the catalytic water molecule, the side chains of the proposed catalytic residues (Tyr107, Glu109, Lys121, and Asp140), and the side chain of the nearby interacting residue Tyr138, as depicted by Figure 3. The remaining 6599 atoms of the system were described by the AMBER *ff10* force field [54] and the MM layer waters were kept fixed.

All QM/MM calculations were carried out within the ONIOM methodology [55,56] using the Gaussian 09 program. Preparation of the input files and output files analysis was facilitated using molUP [57] a VMD plug-in providing a graphical interface for preparation and analysis of QM/MM calculations. The truncated bonds in the QM layer's frontier were treated with the link-atom approach using hydrogen atoms [58]. The system's geometry was optimized twice: a first optimization with the mechanical embedding scheme was followed by a second one employing the electrostatic embedding scheme.

In order to study the reaction path, we scanned this optimized structure's reaction coordinate (i.e., the distance between the oxygen of the carboxylate side chain of Glu109 and the hydrogen of QIN). The structure corresponding to the energy maximum in the scan was taken as a guess for a subsequent full transition state (TS) optimization, and its vibrational frequencies were calculated to confirm its TS nature (a single imaginary frequency). We followed the intrinsic reaction coordinate (IRC) [59] of the TS to obtain the reactants and products of the reaction, which were further optimized and also subjected to a frequency calculation to verify the absence of imaginary frequencies.

Zero-point, thermal, and entropic corrections were also calculated, at 298.15 K, to obtain free energies of activation. In order to assess the influence of the MM layer residues on the activation energy, we performed a series of single-point energy calculations removing every protein residue within 10 Å of the active site region, other than Gly or Pro, employing an approach proven successful in previous work carried out within our group [22,60,61]. The difference in activation energy calculated with a structure with a deleted residue and the complete structure gives the QM/MM energy contribution of that residue for the wild-type barrier.

All QM/MM calculations were performed using the B3LYP density functional [62,63]. The 6-31G(d) basis set [64] was used for the scans, optimizations, and frequency calculations, whereas the larger 6-311+G(2d,2p) basis set, along with the (D3) dispersion correction [65], was used for the single-point calculation of the three optimized stationary states (Figure S5).

3.4. Transition-State Macrodipole Stabilization

In order to assess the influence of the MM layer residues on the activation energy, we performed a series of single-point energy calculations removing every protein residue within 10 Å of the active site region, in both R and TS stationary structures, other than Gly or Pro, employing an approach proven successful in previous work carried out within our group [22,60,61]. The difference in activation energy calculated with a structure with a deleted residue and the complete structure gives the QM/MM energy contribution of that residue for the wild-type barrier.

A total of 105 individual residue deletions in the MM layer residues of each stationary structure were performed, adding up to 210 single-point calculations for this experiment. It is important to note that, as a conceptual experiment, the values obtained for each deletion are simply a quantitative estimate of the stabilization or destabilization effect that a specific residue might have on the reaction barrier and are not meant to be a theoretical calculation of the activation free energy associated to an NCS mutant.

In the conceptual deletion experiment, the values analyzed for the energetic differences are the ones obtained from the single-point calculations, without zero-point, thermal, or entropic corrections. The interaction between the layers was treated with the electrostatic embedding formalism.

4. Conclusions

The essence of catalysis is the stabilization of the transition state with respect to the reactant. It is useless to predict mutations that decrease the activation energy by looking at a crystallographic structure since this only represents one piece of the puzzle, the reactant state. While crystallographic structures co-crystallized with transition state analogs can be helpful, the fine detail they provide is not sufficient. For example, the protonation of Glu109 might antecede, precede, or be synchronous with the deprotonation of Lys121. Such fine details change the associated transition state's nature and charge distribution and are intangible by looking at transition state analogs.

The only way to know the transition state with sufficient accuracy is through computer simulations, such as those performed here.

In this work, we rationally identified mutations that increase the efficiency of the NCS, enabling its optimization for heterologous, or enzymatic, biosynthesis of the precursors of first-line analgesics. The proposed mutations are not disruptive from a structural perspective. The results point the way to the rational optimization of the crop-independent biosynthesis of BIAs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules28114265/s1>. Reference [66] is cited in the supplementary materials.

Author Contributions: Conceptualization, J.P.M.D.S., N.C.S.A.O. and P.A.F.; Methodology, J.P.M.D.S. and N.C.S.A.O.; Software, J.P.M.D.S. and N.C.S.A.O.; Validation, P.A.F.; Formal analysis, J.P.M.D.S., N.C.S.A.O. and P.A.F.; Investigation, J.P.M.D.S. and N.C.S.A.O.; Resources, P.A.F.; Data curation, J.P.M.D.S.; Writing—original draft, J.P.M.D.S. and N.C.S.A.O.; Writing—review & editing, P.A.F.; Visualization, J.P.M.D.S., N.C.S.A.O. and P.A.F.; Supervision, P.A.F.; Funding acquisition, J.P.M.D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financed by National Funds through FCT/MCTES- Portuguese Foundation for Science and Technology within the scope of the project UIDB/50006/2020. J.P.M.S. thanks FCT for the grant SFRH/BD/144205/2019.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study can be found in the article or in the associated Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Not applicable.

References

1. Singh, A.; Menéndez-Perdomo, I.M.; Facchini, P.J. Benzyloquinoline Alkaloid Biosynthesis in Opium Poppy: An Update. *Phytochem. Rev.* **2019**, *18*, 1457–1482. [\[CrossRef\]](#)
2. World Health Organization. *WHO Model List of Essential Medicines—22nd List, 2021*; WHO: Geneva, Switzerland, 2021.
3. Lohman, D.; Schleifer, R.; Amon, J.J. Access to Pain Treatment as a Human Right. *BMC Med.* **2010**, *8*, 8. [\[CrossRef\]](#)
4. Atanasov, A.G.; Waltenberger, B.; Pferschy-Wenzig, E.-M.; Linder, T.; Wawrosch, C.; Uhrin, P.; Temml, V.; Wang, L.; Schwaiger, S.; Heiss, E.H. Discovery and Resupply of Pharmacologically Active Plant-Derived Natural Products: A Review. *Biotechnol. Adv.* **2015**, *33*, 1582–1614. [\[CrossRef\]](#) [\[PubMed\]](#)
5. David, B.; Wolfender, J.-L.; Dias, D.A. The Pharmaceutical Industry and Natural Products: Historical Status and New Trends. *Phytochem. Rev.* **2015**, *14*, 299–315. [\[CrossRef\]](#)
6. Sousa, J.P.; Ramos, M.J.; Fernandes, P.A. Modern Strategies for the Diversification of the Supply of Natural Compounds: The Case of Alkaloid Painkillers. *ChemBioChem* **2022**, *23*, e202100623. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Galanie, S.; Thodey, K.; Trenchard, I.J.; Filsinger Interrante, M.; Smolke, C.D. Complete Biosynthesis of Opioids in Yeast. *Science* **2015**, *349*, 1095–1100. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Pyne, M.E.; Kevvai, K.; Grewal, P.S.; Narcross, L.; Choi, B.; Bourgeois, L.; Dueber, J.E.; Martin, V.J. A Yeast Platform for High-Level Synthesis of Tetrahydroisoquinoline Alkaloids. *Nat. Commun.* **2020**, *11*, 3337. [\[CrossRef\]](#)
9. Hawkins, K.M.; Smolke, C.D. Production of Benzyloquinoline Alkaloids in *Saccharomyces Cerevisiae*. *Nat. Chem. Biol.* **2008**, *4*, 564–573. [\[CrossRef\]](#)
10. DeLoache, W.C.; Russ, Z.N.; Narcross, L.; Gonzales, A.M.; Martin, V.J.; Dueber, J.E. An Enzyme-Coupled Biosensor Enables (S)-Reticuline Production in Yeast from Glucose. *Nat. Chem. Biol.* **2015**, *11*, 465–471. [\[CrossRef\]](#)
11. Narcross, L.; Fossati, E.; Bourgeois, L.; Dueber, J.E.; Martin, V.J. Microbial Factories for the Production of Benzyloquinoline Alkaloids. *Trends Biotechnol.* **2016**, *34*, 228–241. [\[CrossRef\]](#)
12. Lichman, B.R.; Sula, A.; Pesnot, T.; Hailes, H.C.; Ward, J.M.; Keep, N.H. Structural Evidence for the Dopamine-First Mechanism of Norcoclaurine Synthase. *Biochemistry* **2017**, *56*, 5274–5277. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Ilari, A.; Franceschini, S.; Bonamore, A.; Arengi, F.; Botta, B.; Macone, A.; Pasquo, A.; Bellucci, L.; Boffi, A. Structural Basis of Enzymatic (S)-Norcoclaurine Biosynthesis. *J. Biol. Chem.* **2009**, *284*, 897–904. [\[CrossRef\]](#)
14. Roddan, R.; Ward, J.M.; Keep, N.H.; Hailes, H.C. Pictet–Spenglerases in Alkaloid Biosynthesis: Future Applications in Biocatalysis. *Curr. Opin. Chem. Biol.* **2020**, *55*, 69–76. [\[CrossRef\]](#)
15. Lichman, B.R.; Zhao, J.; Hailes, H.C.; Ward, J.M. Enzyme Catalysed Pictet–Spengler Formation of Chiral 1, 1'-Disubstituted-and Spiro-Tetrahydroisoquinolines. *Nat. Commun.* **2017**, *8*, 14883. [\[CrossRef\]](#)
16. Vaissier Welborn, V.; Head-Gordon, T. Computational Design of Synthetic Enzymes. *Chem. Rev.* **2018**, *119*, 6613–6630. [\[CrossRef\]](#)
17. Sheldon, R.A.; Brady, D. Streamlining Design, Engineering, and Applications of Enzymes for Sustainable Biocatalysis. *ACS Sustain. Chem. Eng.* **2021**, *9*, 8032–8052. [\[CrossRef\]](#)
18. Samanani, N.; Liscombe, D.K.; Facchini, P.J. Molecular Cloning and Characterization of Norcoclaurine Synthase, an Enzyme Catalyzing the First Committed Step in Benzyloquinoline Alkaloid Biosynthesis. *Plant J.* **2004**, *40*, 302–313. [\[CrossRef\]](#)
19. Minami, H.; Dubouzet, E.; Iwasa, K.; Sato, F. Functional Analysis of Norcoclaurine Synthase in *Coptis Japonica*. *J. Biol. Chem.* **2007**, *282*, 6274–6282. [\[CrossRef\]](#)
20. Sheng, X.; Himo, F. Enzymatic Pictet–Spengler Reaction: Computational Study of the Mechanism and Enantioselectivity of Norcoclaurine Synthase. *J. Am. Chem. Soc.* **2019**, *141*, 11230–11238. [\[CrossRef\]](#)
21. Svensson, M.; Humbel, S.; Froese, R.D.; Matsubara, T.; Sieber, S.; Morokuma, K. Oniom: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and Pt (P(*t*-Bu)₃)₂ + H₂ Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357–19363. [\[CrossRef\]](#)
22. Sousa, J.P.; Neves, R.P.; Sousa, S.F.; Ramos, M.J.; Fernandes, P.A. Reaction Mechanism and Determinants for Efficient Catalysis by DszB, a Key Enzyme for Crude Oil Bio-Desulfurization. *ACS Catal.* **2020**, *10*, 9545–9554. [\[CrossRef\]](#)
23. Viegas, M.F.; Neves, R.P.; Ramos, M.J.; Fernandes, P.A. Qm/Mm Study of the Reaction Mechanism of Thermophilic Glucuronoyl Esterase for Biomass Treatment. *ChemPhysChem* **2022**, *23*, e202200269. [\[CrossRef\]](#)
24. Bertoni, M.; Kiefer, F.; Biasini, M.; Bordoli, L.; Schwede, T. Modeling Protein Quaternary Structure of Homo- and Hetero-Oligomers Beyond Binary Interactions by Homology. *Sci. Rep.* **2017**, *7*, 10480. [\[CrossRef\]](#)
25. Studer, G.; Rempfer, C.; Waterhouse, A.M.; Gumieny, R.; Haas, J.; Schwede, T. Qmeandisco—Distance Constraints Applied on Model Quality Estimation. *Bioinformatics* **2020**, *36*, 1765–1771. [\[CrossRef\]](#)

26. Studer, G.; Tauriello, G.; Bienert, S.; Biasini, M.; Johner, N.; Schwede, T. Promod3—A Versatile Homology Modelling Toolbox. *PLoS Comput. Biol.* **2021**, *17*, e1008667. [CrossRef]
27. Bienert, S.; Waterhouse, A.; De Beer, T.A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The Swiss-Model Repository—New Features and Functionality. *Nucleic Acids Res.* **2017**, *45*, D313–D319. [CrossRef]
28. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L. Swiss-Model: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [CrossRef]
29. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. Blast+: Architecture and Applications. *BMC Bioinform.* **2009**, *10*, 421. [CrossRef]
30. Remmert, M.; Biegert, A.; Hauser, A.; Söding, J. Hhblits: Lightning-Fast Iterative Protein Sequence Searching by Hm-Hm Alignment. *Nat. Methods* **2012**, *9*, 173–175. [CrossRef]
31. Steinegger, M.; Meier, M.; Mirdita, M.; Vöhringer, H.; Haunsberger, S.J.; Söding, J. Hh-Suite3 for Fast Remote Homology Detection and Deep Protein Annotation. *BMC Bioinform.* **2019**, *20*, 473. [CrossRef]
32. Roddan, R.; Gygli, G.; Sula, A.; Méndez-Sánchez, D.; Pleiss, J.; Ward, J.M.; Keep, N.H.; Hailes, H.C. Acceptance and Kinetic Resolution of A-Methyl-Substituted Aldehydes by Norcoclaurine Synthases. *ACS Catal.* **2019**, *9*, 9640–9649. [CrossRef]
33. Davis, I.W.; Leaver-Fay, A.; Chen, V.B.; Block, J.N.; Kapral, G.J.; Wang, X.; Murray, L.W.; Arendall, W.B., III; Snoeyink, J.; Richardson, J.S. Molprobability: All-Atom Contacts and Structure Validation for Proteins and Nucleic Acids. *Nucleic Acids Res.* **2007**, *35*, W375–W383. [CrossRef] [PubMed]
34. Chen, V.B.; Arendall, W.B.; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. Molprobability: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 12–21. [CrossRef] [PubMed]
35. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B. Molprobability: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci.* **2018**, *27*, 293–315. [CrossRef]
36. Wiederstein, M.; Sippl, M.J. Prosa-Web: Interactive Web Service for the Recognition of Errors in Three-Dimensional Structures of Proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410. [CrossRef]
37. Sippl, M.J. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins Struct. Funct. Bioinform.* **1993**, *17*, 355–362. [CrossRef]
38. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
39. Case, D.A.; Belfon, H.M.A.K.; Ben-Shalom, I.Y.; Berryman, J.T.; Brozell, S.R.; Cerutti, D.S.; Cheatham, G.A., III; Cisneros, V.W.D.; Cruzeiro, T.A.; Darden, R.E.; et al. *Amber 2018*; University of California: San Francisco, CA, USA, 2018.
40. Frisch, M.J.; Schlegel, G.W.T.H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; Li, X.; et al. Gaussian, Inc.: Wallingford, CT, USA, 2016. Available online: [https://www.scirp.org/\(S\(lz5mqp453ed%20snp55rrgct55\)\)/reference/referencespapers.aspx?referenceid=2418053](https://www.scirp.org/(S(lz5mqp453ed%20snp55rrgct55))/reference/referencespapers.aspx?referenceid=2418053) (accessed on 13 July 2022).
41. Cerqueira, N.; Ribeiro, J.; Fernandes, P.; Ramos, M. Vslab—An Implementation for Virtual High-Throughput Screening Using Autodock and Vmd. *Int. J. Quantum Chem.* **2011**, *111*, 1208–1212. [CrossRef]
42. Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]
43. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. Autodock4 and Autodocktools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef]
44. Huey, R.; Morris, G.M.; Olson, A.J.; Goodsell, D.S. A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152. [CrossRef] [PubMed]
45. Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639–1662. [CrossRef]
46. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef] [PubMed]
47. Bayly, C.I.; Cieplak, P.; Cornell, W.; Kollman, P.A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The Resp Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280. [CrossRef]
48. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. Ff14sb: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99sb. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef]
49. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [CrossRef]
50. Le Grand, S.; Götz, A.W.; Walker, R.C. Spfp: Speed without Compromise—A Mixed Precision Model for Gpu Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374–380. [CrossRef]
51. Gotz, A.W.; Williamson, M.J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555. [CrossRef]
52. Salomon-Ferrer, R.; Gotz, A.W.; Poole, D.; Le Grand, S.; Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888. [CrossRef]
53. Roe, D.R.; Cheatham III, T.E. Ptraaj and Cpptraaj: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [CrossRef]

54. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. [[CrossRef](#)]
55. Chung, L.W.; Sameera, W.; Ramozzi, R.; Page, A.J.; Hatanaka, M.; Petrova, G.P.; Harris, T.V.; Li, X.; Ke, Z.; Liu, F. The Oniom Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796. [[CrossRef](#)]
56. Chung, L.W.; Hirao, H.; Li, X.; Morokuma, K. The Oniom Method: Its Foundation and Applications to Metalloenzymes and Photobiology. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 327–350. [[CrossRef](#)]
57. Fernandes, H.S.; Ramos, M.J.; MFSA Cerqueira, N. *Molup: A Vmd Plugin to Handle Qm and Oniom Calculations Using the Gaussian Software*; Wiley Online Library: Hoboken, NJ, USA, 2018.
58. Sousa, S.F.; Ribeiro, A.J.; Neves, R.P.; Brás, N.F.; Cerqueira, N.M.; Fernandes, P.A.; Ramos, M.J. Application of Quantum Mechanics/Molecular Mechanics Methods in the Study of Enzymatic Reaction Mechanisms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1281. [[CrossRef](#)]
59. Fukui, K. The Path of Chemical Reactions—the Irc Approach. *Acc. Chem. Res.* **1981**, *14*, 363–368. [[CrossRef](#)]
60. Ribeiro, A.J.; Santos-Martins, D.; Russo, N.; Ramos, M.J.; Fernandes, P.A. Enzymatic Flexibility and Reaction Rate: A Qm/Mm Study of Hiv-1 Protease. *Acs Catal.* **2015**, *5*, 5617–5626. [[CrossRef](#)]
61. Medina, F.E.; Neves, R.P.; Ramos, M.J.; Fernandes, P.A. Qm/Mm Study of the Reaction Mechanism of the Dehydratase Domain from Mammalian Fatty Acid Synthase. *ACS Catal.* **2018**, *8*, 10267–10278. [[CrossRef](#)]
62. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785. [[CrossRef](#)]
63. Becke, A.D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098. [[CrossRef](#)]
64. Ditchfield, R.; Hehre, W.J.; Pople, J.A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728. [[CrossRef](#)]
65. 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. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)] [[PubMed](#)]
66. Benkert, P.; Tosatto, S.C.; Schomburg, D. Qmean: A Comprehensive Scoring Function for Model Quality Assessment. *Proteins Struct. Funct. Bioinform.* **2008**, *71*, 261–277. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.