Predicting Effects of Site-Directed Mutagenesis on Enzyme Kinetics by QM/MM and QM Calculations: A Case of Glutamate Carboxypeptidase II

Daniel Bím,[1,2] Michal Navrátil,[1] Ondrej Gutten,[1] Jan Konvalinka,[1] Zsofia Kutil,[3] Martin Culka,[1] Václav Navrátil,[1] Anastassia N. Alexandrova,[2] Cyril Bařinka,[3,] Lubomír Rulíšek[1,]**

[1] Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo náměstí 2, 166 10 Praha 6, Czech Republic; [2] Department of Chemistry and Biochemistry, University of California, Los Angeles, California, U. S. A. 90095-1569; [3] Institute of Biotechnology of the Czech Academy of Sciences, Průmyslová 595, 252 50 Vestec, Czech Republic.

Corresponding authors: cyril.barinka@ibt.cas.cz; rulisek@uochb.cas.cz

ABSTRACT. Quantum and molecular mechanics (QM/MM) and QM-only (cluster model) modeling techniques represent the two workhorses in mechanistic understanding of enzyme catalysis. One of the stringent tests for QM/MM and/or QM approaches is to provide quantitative answers to real-world biochemical questions, such as the effect of single-point mutations on enzyme kinetics, of an immediate relevance to enzyme design. This translates into predicting the *relative* activation energies to 1-2 kcal.mol^{-1} accuracy. Herein, we employ glutamate carboxypeptidase II (GCPII), a dizinc metallopeptidase, as a model system. Its reaction mechanism is quite understood: the reaction coordinate comprises 'tetrahedral intermediate' and two transition states and experimental activation Gibbs free energies of ~17.5 kcal.mol^{-1} can be inferred for the known $k_{cat} \approx 1$ s^{-1}. We correlate experimental kinetic data (including E424H variant, newly characterized in this work) for various GCPII mutants ($k_{cat} = 8.6 \cdot 10^{-5}$ s^{-1} - 2.7 s^{-1}) with the energy profiles calculated by QM/MM and QM-only (cluster model) approaches. We show that the near-quantitative agreement between the experimental values and the calculated activation energies ($\Delta H^{\ddagger}$) can be obtained and recommend the combination of the two protocols: QM/MM optimized structures and cluster model (QM) energetics. The trend in *relative* activation energies is mostly independent of the QM method (DFT functional) used. Last but not least, a satisfactory correlation between experimental and theoretical data allows us to provide qualitative and fairly simple explanations of the observed kinetic effects which are thus based on a rigorous footing.

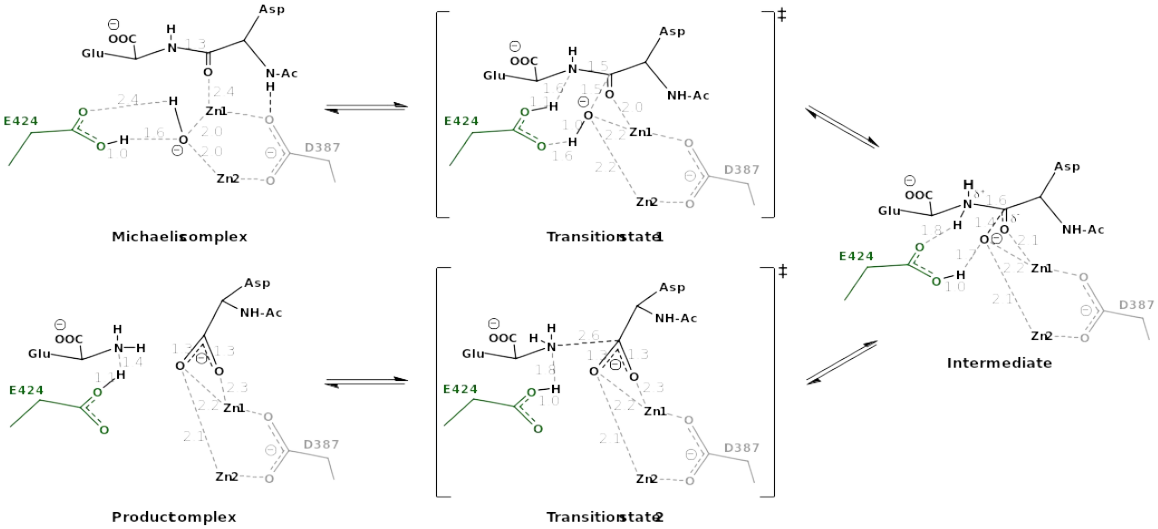
1. Introduction

Computer modeling and design of enzyme catalysis has become an integral part of biochemical research.[1–3] Over more than two decades of methods development, supported by the available (and growing) computer power, there is ample expertise to characterize in detail reaction coordinates of most of (metallo)enzymes.[4–17] With due respect to the semiempirical quantum chemical methods (exemplified by the empirical valence bond approach, EVB)[18] and force-field methods, the progress has been mostly made by employing quantum mechanics (QM). QM methods are used either as standalone (sometimes denoted 'cluster modeling', where a cluster is mostly equivalent to enzyme's active site)[7,19–21] or in combination with molecular mechanical methods (MM), either in a more static approach: QM/MM[12–15] or including dynamics (QM/MM MD).[16,17,22]

To date, reaction mechanisms were described for several hundreds, if not thousands of (metallo)enzymes (see, e.g., refs. [23–41] for selected examples from various research groups). The last critical and exhaustive survey has been made by Senn and Thiel[42] whereas the very recent commentary of Cui et al.[43] discusses remaining burning issues in the QM/MM field. It has been argued that accuracy on the order of 5 kcal.mol^{-1} in calculated activation (free) energies still enables one to elucidate the correct reaction coordinate for a particular enzyme.[44] Somewhat more practical test for the validity and computational robustness of a particular reaction mechanism, although scarcely done, is the effect of single-point mutations on the enzymes kinetics. Strategic mutagenesis (as opposed to laboratory evolution) can be part of enzyme design protocol. This requires accuracy in computed *relative* activation energies in the order of ~1-2 kcal.mol^{-1} (corresponding to a roughly one order of magnitude in experimental k_{cat}). Considering subtleties of "computational enzymology", such as conformational freedom within an enzyme-substrate complex that adds

to the inherent inaccuracy of the QM or QM/MM methods, this goal seems rather challenging.[6–8,45,46]

In our opinion, a well-established general protocol for handling these intricate issues is still missing. Ongoing discussions are mostly trying to contrast (or appease) two most commonly used models: cluster model (QM-only approach) or/and/vs. QM/MM.[7,10,12,16,19,42] At the same time, more advanced



Scheme 1. wt-GCPII reaction mechanism, inferred from QM/MM modeling, X-ray crystallography, and site-directed mutagenesis. Note that the reaction mechanism was established in ref. [50], whereas the distances are corrected based on our calculations.

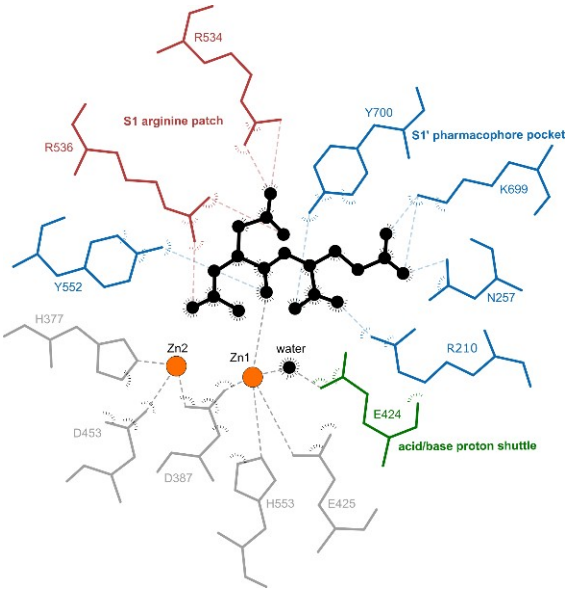


Figure 1. Schematic view of the wt-GCPII active site.

approaches to compute free energy changes along the reaction coordinate more rigorously, such as QM/MM free-energy perturbation methods (QM/MM-FEP), are advocated for.[47–49] In our opinion, huge computational demands make the FEP-based QM methods somewhat impractical, and perhaps not always necessary for obtaining fairly reliable contours and energy profiles of enzymatic reactions.

In this work, we want to add few arguments into the methodological discussions, by employing well-calibrated model of the prototypic, but in many aspects non-trivial, metalloenzyme, dizinc glutamate carboxypeptidase II (GCPII).

Human glutamate carboxypeptidase II (GCPII; E.C. 3.4.17.21) is a dinuclear zinc-dependent peptidase that catalyzes, *inter alia*, the hydrolysis of *N*-acetyl-L-aspartyl-L-glutamate (NAAG) to *N*-acetyl aspartate (NAA) and glutamate.[51] GCPII is expressed in various tissues of the human body, including the prostate, nervous system, small intestine, and kidney.[52,53] The overexpression of GCPII in the neovasculature of a variety of solid tumors and malignant prostate tissue makes GCPII the prime target in imaging, diagnosis, and treatment of a majority of solid cancers.[54]

The active site of human GCPII contains two Zn^{2+} ions that are indispensable for its hydrolytic function, and play both functional and structural roles. The catalytic Zn1, coordinated by the side chains of H553, E425, and bridging D387, polarizes the carbonyl group of the NAAG peptide bond and stabilizes the anionic tetrahedral intermediate in the reaction mechanism (**Scheme 1**). The co-catalytic Zn2 (coordinated by

H377, D453, and bridging D387) plays a role in the substrate binding and assists in the orientation of the hydroxide nucleophile and stabilization of the transition state (**Scheme 1**).[50,55] The GCPII substrate-binding pocket comprises two differentiable sites – S1' and S1.[56] The amphipathic S1' site is optimized for glutamate binding by fixating the α-carboxylate group through a salt bridge with the R210 side chain, and hydrogen bonds to Y552 and Y700 side chains; and γ-carboxylate by hydrogen bonds to N257 and K699 (**Figure 1**).[57] The S1 site has a funnel-shaped architecture with a characteristic polar 'arginine patch' region near the Zn2 active site (R534, R536, and R463) that is responsible for specificity towards acidic residues at the P1 position of substrates. Previously, the arginine patch was also proposed to draw the resulting product of the hydrolysis (NAA) from the active site into the exit channel as a part of the completion of the GCPII catalytic cycle.[50,58]

The wild-type GCPII (wt-GCPII) exhibits a strong preference for the NAAG binding and hydrolysis with reported K_m = 1.15 μM and k_{cat} = 1.1 s^{-1}, respectively, and single amino acid substitutions in the S1' and S1 sites typically resulted in the decreased catalytic efficiency of the mutants. Past mutagenesis studies revealed the role of each of the targeted residues in the protein stability, substrate binding, and/or reaction mechanism. For example, the E424A mutation abolished GCPII enzymatic activity, which is in line with E424 role as a proton shuttle required for NAAG activation and regeneration of the catalytic cycle. R210 in the S1' pocket has been shown to be a critical residue affecting both the substrate binding affinity and the catalytic cycle as it directly interacts with the α-carboxylate of the NAAG glutamate moiety and helps in positioning the Y552 residue side chain. The Y552 is, in turn, crucial for the stabilization of the transition state, and both Y552 and Y700 hydroxyl groups are required for the polarization of the NAAG scissile peptide bond carbonyl group, increasing its electrophilicity for the OH^- attack. As an example of non-polar (hydrophobic) interactions between GCPII and NAAG, the side chains of F209 and L428 are particularly important for substrate binding in the S1' site. Finally, by mutating residues in the S1 pocket, it has been demonstrated that the S1 architecture is responsible for GCPII specificity toward glutamate and aspartate functionalities in the P1 position of the *N*-acetylated substrates.[50,59,60]

Herein, we used the kinetic data from several mutagenetic experiments (carried out previously) to assess the applicability and limits of the multiscale modeling techniques to accurately describe the subtle effects of single-point mutations on the turnover number (k_{cat}). More specifically, on a precisely defined model, we want to define and discuss key parameters of the QM and QM/MM modeling techniques that should be considered in calculations. Last but not least, by presenting new experimental data (E424H mutation which was suggested by calculations), we give an example of how the experiment and theory might work synergistically to advance our understanding of biomolecular systems.

2. Computational Details

2.1. Protein setup. The computational protocols and the general QM/MM setup that was previously successfully applied in mechanistic and structural studies of wt-GCPII[50,61–63] have been employed in this work. This also concerns the initial protein structure used in the modeling (the full structures are deposited in the supporting information, in the PDB format). The model structure is the fully optimized all-atom QM/MM model[61-63] originating in the 1.71 Å resolution X-ray structure of the E424A mutant with the NAAG ligand (PDB code 3BXM). This included addition of a few missing residues in the 3BXM crystal structure, added solvation sphere of ~12,700 water molecules (radius of 50 Å from the center of the protein), and two Na^+ ions to neutralize the overall charge of the protein. Apart from utilizing the previous (non-trivial) protein preparation and equilibration,[61] we believe that the adopted approach also warrants the reliable and experimentally consistent position of the ligand in the active site. The quantum system of wt-GCPII comprised 324 atoms (i.e., side chains or backbone atoms of the following residues: F209, R210, N257, H377, D387, E424, E425, L428, N451, D453, S517, G518, N519, D520, F521, R534, R536, Y552, H553, R580, K699, Y700, Zn ions, hydroxide, chloride ion, and 10 water molecules) whereas the part of the protein that was allowed to relax in the QM/MM calculations comprised additional 28 residues and 9 water molecules in the vicinity of the QM system (i.e., any residue/water with an atom with R < 2.0 Å from any atom of the QM region was included as a whole). This definition of the QM system warrants that all residues subjected to mutagenesis were included in optimization, allowing to keep the same distribution of residues between individual QM/MM regions and study all effects of mutagenesis quantum mechanically (i.e., from the first principles). Different GCPII variants were prepared by

a direct replacement of the residues in the reactant complex (RC) of wt-GCPII followed by QM/MM optimization. In some cases, this required addition of water molecules in the void space generated by the mutations (**Table 1**).

Table 1. Sizes and differences in QM systems used for the studied mutants in QM/MM calculations.

	QM atoms	QM charge	added waters
wt-GCPII	324	0	-
E424A	321	0	1
E424H	326	1	0
R210A	322	-1	3
R210K	322	0	0
Y552I	325	0	1
Y700F	323	0	0
F209W	329	0	0

2.2. QM/MM and QM Calculations. The ComQum software was used for all QM/MM calculations. A detailed description of the contributions to the total QM/MM energy is summarized in the Computational Details section in the Supporting Information, while more technical details can be found in refs. 64–67. In brief, the ComQum package combines Turbomole and Amber programs for the QM/MM optimizations. It implements electrostatic embedding, the hydrogen link-atom scheme, and a microiterative approach; i.e., the MM system is fully relaxed after each optimization step in the QM region. The total (QM/MM) energy is then calculated as:

$$E_{\text{QM/MM}} = E_{\text{QM-pchg}} + E_{\text{MM123}} - E_{\text{MM1}},$$

where $E_{\text{QM-pchg}}$ corresponds to the QM energy of System 1 embedded in a set of point charges of the Systems 2 and 3 (MM part; the self-interaction of point charges not included in the $E_{\text{QM/MM}}$ term), E_{MM123} is the MM energy of the entire system (MM charges of the System 1 zeroed), and E_{MM1} is the MM energy of System 1, again with the MM charges of the System 1 zeroed. Note that MM part thus formally consists of Systems 2 and 3, which are (energetically) treated on the same footing. However, System 3 atoms (outer part of the protein) are fixed at their original crystallographic positions.

All DFT calculations (QM region) were performed using the Turbomole 6.6 program.[68] Geometry optimizations were carried out using the TPSS functional[69] with the def2-SV(P) basis set,[70] including the empirical zero-damping dispersion correction (D3),[71] and expedited by the RI-J approximation.[72] The transition states were obtained as one- or two-dimensional scans of QM/MM potential energy surfaces (PESs). In the first transition state structure, this required a simultaneous variation of the distance between (i) Zn-bound hydroxide nucleophile and the peptide-bond carbonyl of the NAAG substrate, and (ii) hydrogen atom of the protonated-E424 carboxyl group and the NAAG amide nitrogen. In the second transition state, the variation of the peptide bond C−N distance was coupled with the hydrogen atom transfer from E424 shuttle to the emerging amine functional group (*cf.* reaction mechanism in **Scheme 1**). The MM calculations were performed employing the Amber ff14SB force field.[73]

For the QM(TPSS-D3/def2-SV(P))/MM equilibrium geometries, the single-point energies were evaluated with TPSS,[69] TPSSh,[69,74] and B3LYP[75] methods and the def2-TZVP basis set *(i)* embedded in the set of point charges from the MM region to obtain more accurate $E_{\text{QM-pchg}}$ term above, and *(ii)* in the homogeneous environment represented by the conductor-like screening model (COSMO)[76] with dielectric constant $\varepsilon = 4$ to evaluate approach analogous to the one adopted in cluster modeling. We consider the QM/MM optimized geometries as the more naturally 'restrained' model as compared to the standard cluster model approach which requires fixing some of the active-site atomic coordinates at the crystal structure positions.

2.3 Kinetic model. The experimental free energy of activation ($\Delta G^{\ddagger}_{\text{expt}}$) for each GCPII variant was calculated from the measured rate constants (k_{cat}) using the Eyring equation:

$$k_{cat} = \frac{k_B T}{h} e^{\frac{-\Delta G^{\neq}_{expt}}{RT}}$$

In the calculations, the reaction is approached by the kinetic model: $RC \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathfrak{I} \xrightarrow{k_2} PC$, under steady-state approximation (where RC, IM, and PC stand for reactant complex, intermediate, and product complex). The steady-state conditions assume the intermediate concentration is low and constant during the course of the reaction. This is achieved in the GCPII due to the relatively high energy of the IM, so that $k_{-1} + k_2 >> k_1$.

Under steady-state approximation:

$$\frac{d[\mathfrak{I}]}{dt} = k_1[RC] - (k_{-1} + k_2)[\mathfrak{I}] \approx 0$$

$$[\mathfrak{I}] \approx \frac{k_1[RC]}{k_{-1} + k_2}$$

$$\frac{d[PC]}{dt} \approx \frac{k_1 k_2}{k_{-1}+k_2}[RC] = k_{obs}[RC]$$

The k_1, k_{-1}, k_2 were calculated from the energies of TS1, IM and TS2 and substituted into Eq. (4) to estimate the calculated k_{obs} and $\Delta H^{\ddagger}_{obs}$ using the Eyring equation. Note that $\Delta H^{\ddagger}_{obs}$ obtained from the kinetic model is equivalent to the energy of the higher reaction barrier (TS1 or TS2), unless the two barriers are very close in energy.

3. Experimental

3.1 Cloning and purification of GCPII mutants

Plasmids encoding GCPII mutants were generated by the standard QuikChange site-directed mutagenesis approach. For the GCPII(E424H) and GCPII(F209W) mutants, plasmids encoding wt-GCPII flanked by the *N*-terminal Strep-II and Avi purification tag, respectively, were used as templates.[77,78] The identity of the resulting mutants were verified by Sanger DNA sequencing. Expression and purification of the mutants were carried out according to the protocol reported for the wild-type proteins.[77,78] Briefly, the mutants were heterologously overexpressed in Schneider's S2 cells and purified to homogeneity by Strep-Tactin (GCPII(E424H)) or Streptavidin Mutein (GCPII(F209W)) affinity chromatography, followed by size-exclusion chromatography. The final protein preparation was >95% pure as determined by reducing SDS-PAGE.

3.2. Kinetic measurements

The enzymatic activity of GCPII(E424H) (and wt-GCPII as a control) was determined using a protocol described in ref.[79] with ^{3}H-labeled *N*-acetyl aspartyl glutamate (3H-NAAG) as a substrate. The enzymes were preincubated in reaction buffer (50 mM Bis-Tris, 150 mM NaCl, 0.001% $E_{12}C_8$, pH 7.5) in a total volume of 80 µl at 37 °C for 15 min. For the GCPII(F209W) mutant (and wt-GCPII as a control), 25 mM Tris-HCl, 0.1 % Triton X-100, pH 7.5, was used as the assay buffer. The reaction was initiated by the addition of 40 µl of a mixture of 10 µM NAAG and 50 nM ^{3}H-NAAG (49.1 Ci/mmol in 20 mM Tris buffer, Perkin Elmer) to a total reaction volume of 120 µl. After 240 min, the reaction was terminated by the addition of 120 µl of 200 mM potassium phosphate, 50 mM EDTA, 2 mM β-mercaptoethanol, pH 7.4. The released glutamate was separated from the reaction mixture by ion-exchange chromatography and quantified by liquid scintillation. Kinetic parameters for remaining mutants (and a control wild-type reaction) were taken from ref. [55].

4. Results

4.1. Influence of single-point mutations on GCPII kinetics: Experimental data. To determine kinetic parameters for wt-GCPII and its mutants, we expressed the GCPII variants in insect S2 cells and purified them to homogeneity using the combination of affinity and size exclusion chromatography.[77,78] The purified proteins were then used to determine their catalytic rate constants using NAAG as a substrate (**Table 2**).

Table 2. Catalytic rate constants of GCPII variants with NAAG as a substrate.

Mutant	k_{cat} (s^{-1})	wt/mut
Wild type*,$,#	0.65 – 2.2	1
E424A, E424M#	N.D.	N.D.
E424S#,@	< 10^{-5}	>100,000
E424H#	8.6x10^{-5} ± 0.2x10^{-5}	7,325
F209W$	2.7 ± 0.2	0.88
R210A*	0.023 ± 0.001	47.8
R210K*	0.130 ± 0.020	8.5
Y552I*	0.014 ± 0.001	78.6
Y700F*	0.075 ± 0.003	14.7

N.D. – activity below the detection limit of the assay
* - values from ref. [59]
$,# - k_{cat} values of corresponding wt-GCPII were determined to be 2.2 ± 0.2$ and 0.63 ± 0.08# likely reflecting the different composition of the assay buffers
@E424S – activity not measurable using biochemical assays but can be inferred from X-ray structures where hydrolyzed product (glutamate) is observed in the active site. This is in contrast with crystallographic studies using E424A and E424M, where the intact substrates are observed.

The (measurable) k_{cat} values span the range of $8.6 \cdot 10^{-5}$ s^{-1} to 2.7 s^{-1}, with extremes represented by the GCPII(E424H) and GCPII(F209W) mutants, respectively. In addition, NAAG-hydrolyzing activities of mutants of the E424 proton shuttle (E424A, E424M, and E424S) were below the detection limit of 10^{-5} s^{-1} of the assay. Our structural data on GCPII/NAAG complexes further revealed that while E424A and E424M mutants are completely inactive (we can observe intact NAAG in the GCPII active site; ref. 50 and *unpublished data*), the E424S mutant likely retains a residual hydrolytic activity. This can be inferred from the fact that glutamate, the NAAG hydrolysis product, is observed in crystals from GCPII(E424S)/NAAG droplets (data not shown), reflecting our observations for the wt-GCPII/NAAG crystallization setups.[55]

Notwithstanding the fact that the kinetic data were determined using different constructs, purification

schemes, and reaction conditions in different laboratories, the comparison of k_{cat} values for the wild-type enzymes (0.65 – 2.2 s^{-1}) clearly show surprising consistency in obtained enzymatic parameters. This observation gives us confidence that within limits of experimental variability, our kinetic data are robust and fairly reproducible, and thus well-suited as experimental basis for corresponding *in silico* calculations.

4.2. Influence of single-point mutations on GCPII kinetics: QM/MM and QM Calculations.

4.2.1 wt-GCPII

For the wt-GCPII with the large QM/MM active site (consisting of approximately 320 atoms) presented herein, we have found a reaction mechanism consistent with our findings published previously, which is discussed in detail in the Introduction section. Specifically, the hydrolytic cleavage of the peptide bond is accomplished through two transition states and one intermediate (**Figures 2-3**). The two transition states were found to be relatively close in energy, with the formation of the second transition state (TS2) being the rate-determining step of the reaction (ca. 5 kcal.mol^{-1} higher in energy; irrespective of the QM method used in the QM/MM scheme – comparing TPSS-D3, B3LYP-D3, and TPSSh-D3, with varying amount of exact exchange). The overall activation enthalpy, $\Delta H^{\ddagger}$ = 29.9 kcal.mol^{-1} (QM(B3LYP-D3)/MM level), appears to be significantly overestimated as compared to the experiment (k_{cat} = 1.1 s^{-1}, corresponding to $\Delta G^{\ddagger}_{expt}$ = 17.4 kcal.mol^{-1}). This is not surprising, as such overestimation of the $\Delta H^{\ddagger}$ was previously described in a benchmark study on a smaller GCPII model system.[60] Similarly, non-hybrid QM(TPSS-D3)/MM yields results with the $\Delta H^{\ddagger}$ lowered by ca. 3 kcal.mol^{-1}, still quite high as

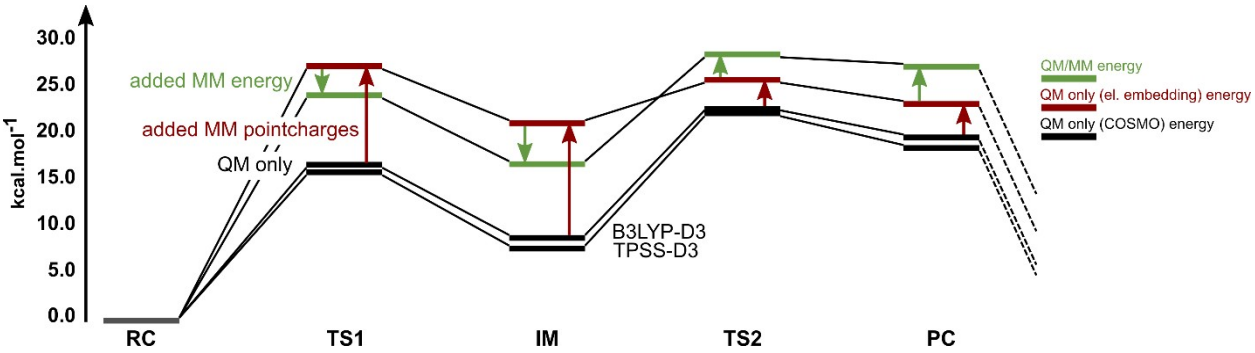


Figure 2. Energetic profile of the wt-GCPII reaction mechanism at different levels of theory. Note that energies of all reaction intermediates are evaluated using the identical QM regions from the QM/MM optimized geometries; however, the overall energy is composed of different energetic terms. In QM/MM profile (green), the QM energy obtained at the DFT (TPSS-D3) level is complemented by the MM energy of the 'System 2' (see Computational Details) and the QM-region is embedded in the set of MM point charges. In red, the MM energy of system 2 is omitted, but the QM-region is affected by the electrostatics from the MM point charges. In black (QM only; COSMO), all of the components from the MM region are neglected and QM system is calculated in the homogeneous dielectric continuum with ε = 4 at the fixed coordinates from the QM/MM equilibration.

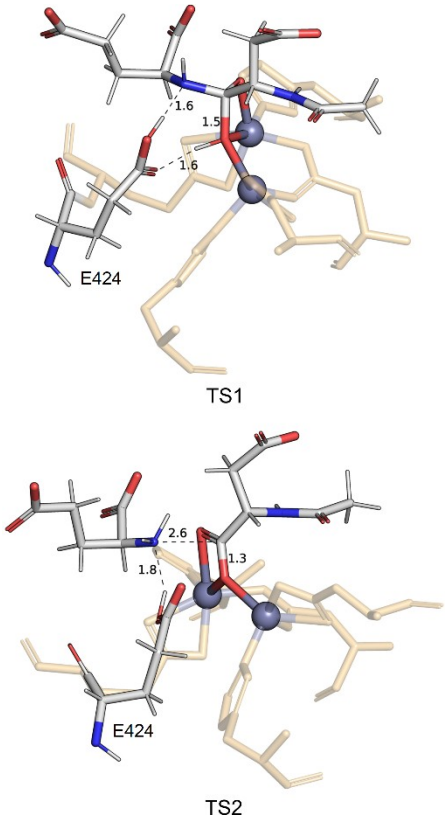


Figure 3. The structure of TS1 and TS2 from the QM/MM calculations of the wt-GCPII. The E424 residue possess a dual role in the catalysis. In TS1, the E424 increases the nucleophilicity of the hydroxyl to assist the attack on the NAAG's carbonyl. In TS2, the proton donation by E424 to NAAG's nitrogen facilitates the C−N bond cleavage.

compared to the experiment. In ref. 60, various popular density functionals were calibrated for the in vacuo peptide bond hydrolysis energetics employing the reference MP2 and CCSD(T) calculations. We have demonstrated that TPSS functional with D3 dispersion correction performed relatively well for the larger model systems and was in a very good agreement with the CCSD(T) reference energies. Also, we have demonstrated a satisfactory convergence of the calculated barriers with the increasing size of the quantum system (converged at ~300 atoms). These initial benchmark calculations led us to select the model system and the TPSS-D3 method for the study presented herein. The zero-point vibrational energy (ZPVE) corrections, obtained from the frequency calculations on the full QM-model system has relatively small effect on the TS1 and TS2 energies, *cf.* **Table S1**.

We consider such overestimation of activation energies to originate from the improper description of the MM part of the QM/MM model system (a standard hydrogen-link atom approach is used). In particular, we have observed an imbalance between the energetic terms originating in the MM point charges acting on the QM region and the MM energy of the protein outside the QM region. On one hand, the MM point charges lead to a destabilization of the IM by as much as ~13 kcal.mol^{-1}, along with the increase of $\Delta H^{\ddagger}$ (TS1) and $\Delta H^{\ddagger}$ (TS2) by 11 kcal.mol^{-1} and 3 kcal.mol^{-1}, respectively. In contrast, such a strong effect of the point charges is *back-corrected* by the large energetic changes in the MM energies, stabilizing the TS1 and IM by ~ 3 and 5 kcal.mol^{-1} and destabilizing the TS2 by 3 kcal.mol^{-1} (*cf.* **Figure 2**). Such dramatic changes originating from the MM construct based on force-field methods appear to be quite unreliable. We propose that better energetics might be obtained by using the QM/MM method for obtaining reliable and realistic geometries (in the context of the enzyme), but disregard the MM-originated energetic terms and point charges. Note that similar approach was advocated for by Sumner et al., recommending QM/MM methodology for geometry optimizations with energies evaluated through the 'big-QM approach'.[80] The presented approximation can be supported also by the recently reported relative "near-sightedness" of the electrostatic interactions and electric fields exerted on the active sites of enzymes by the rest of the protein macromolecule.[81–83]

In fact, the QM-only (COSMO; ε = 4) energies evaluated as single-points at the equilibrium QM/MM geometries (denoted as TPSS-D3//QM/MM) demonstrate a significantly remedied behavior with $\Delta H^{\ddagger}$(TS2) = 23.2 kcal.mol^{-1}. The energetic profile is presented in **Figure 2**. Note that more advanced COSMO for real solvents (COSMO-RS) method[84,85] was previously suggested to provide better energetics in enzymatic catalysis, when the enzyme was approximated with the σ profile of octanol.[63] Herein, we report that COSMO (ε = 4) provides comparable results as COSMO-RS (octanol) within ~0.5 kcal.mol^{-1} (**Table S2**). In addition, for COSMO-RS, the TS1 energy is increased from 15.5 to 16.5 kcal.mol^{-1} when the environment is more polar (hexane → octanol → water), while the effect on TS2 is opposite going from 23.9 to 21.7 kcal.mol^{-1}.

4.2.2 E424 proton shuttle mutations

***E424A*.** The mutation of the E424 proton shuttle to alanine is associated with the total loss of the GCPII hydrolytic activity. The E424 is involved, as the proton shuttle, in both activation steps of the NAAG

hydrolysis (*cf.* **Figure 3**). It both increases the nucleophilicity of the Zn-bound hydroxyl and stabilizes the *N*-terminal ammonium group of the emerging glutamate product in the course of the NAAG peptide bond hydrolysis. Therefore, the loss of catalytic activity is not surprising. Herein, we show that the NAAG hydrolysis is prevented due to substantially higher TS2 ($\Delta H^{\ddagger}$ = 42.6 kcal.mol^{-1} calculated at the TPSS-D3//QM/MM level), whereas the IM might still be accessible via a TS1 of $\Delta H^{\ddagger}$ = 28.3 kcal.mol^{-1} (roughly corresponding to k_1 in the range of 10^{-8} s^{-1}).

E424H. Consistent with the function of E424 as a proton shuttle, the E424H variant also exhibits a significant decrease in the activity. In contrast to the E424A mutation, the histidine residue can, in principle, function in the NAAG hydrolysis as the proton shuttle and thus take over the role of E424 in transferring a proton from the water/hydroxyl nucleophile to the NAAG. Qualitatively, two trends are expected to counterbalance each other: higher pK_a (which should facilitate the second step of the reaction) and different geometry of the side chain that may presumably interfere with the reaction mechanism. Computations suggest that H424 can deprotonate the Zn-bound water in the reactant complex and generate the activated OH^- nucleophile, analogously to the wild-type E424 residue. However, the computed activation enthalpy corresponding to the TS1 is 11 kcal.mol^{-1} higher than that of the wild-type enzyme (and 4.8 kcal.mol^{-1} higher than the rate-determining TS2 of the wt-GCPII). Note that TS1 energy is calculated to be 28.0 kcal.mol^{-1} at the TPSS-D3//QM/MM level of theory. An increase in the TS1 energy can be attributed to the lower nucleophilicity of the OH^-, because of the missing H-bonding interaction with the E424, and the energetically demanding distortion/movement of the H424 toward the TS1. In other words, while the E424 seems to be optimized to facilitate the NAAG hydrolysis, the misfit of the H424 in the active site requires a larger reorganization (energy). Noticeably, the experimentally observed turnover number (rate constant) is in accord with the computational predictions; significantly lower *wrt* the wt-GCPII.

Interestingly, when the TS1 is overcome, the reaction can be completed through a low-lying TS2, only 2.5 kcal.mol^{-1} above the energy of the intermediate. This is in stark contrast with E424A, for which the high TS2 prohibits the reaction to occur. However, the TS2 is also the rate-determining step in the wt-GCPII. The improvement in the second step of the hydrolysis using the E424H variant can be attributed to a worse stabilization of the reaction intermediate and the higher pK_a of the H424 (as compared to E424), which would facilitate the product formation *via* easier proton transfer to the glutamate of the NAAG hydrolysis product.

4.2.3 Mutations of the S1' pharmacophore pocket residues

The key role in the stabilization of the 'tetrahedral intermediate' (IM) of the NAAG hydrolysis is performed by the R210, Y552, and Y700 residues. Along the course of the reaction, these residues are utilized to pull NAAG apart by forming a salt bridge and the hydrogen bond between the NAAG glutamate α-carboxylate and the guanidinium group of R210 and the Y700 hydroxyl group, respectively. Furthermore, there is a hydrogen bond between the NAA fragment and the hydroxyl group of Y552 (**Figure S3**) Since the R210 and Y552 residues are coupled and positioned next to each other, a mutation of either of them results in a disruption of the both modes of the NAAG stabilization. Consistent with this picture, the calculated energy of the IM is higher for the Y552I, R210A, and R210K variants (4.5-5.0 kcal.mol^{-1} destabilization of the IM *wrt* the wt-GCPII) than for Y700F (3.5 kcal.mol^{-1} destabilization of the IM). Similarly, the energy of TS1 is increased by all of these mutations by ~3.5-6.0 kcal.mol^{-1}. The rate-determining TS2 is also energetically higher as compared to the wt-GCPII by ~1-2 kcal.mol^{-1} in R210K and Y552I and by ~4-4.5 kcal.mol^{-1} in R210A and Y700F. This is consistent with the decrease in the turnover number of these GCPII variants, ranging from k_{cat} = 0.014 s^{-1} to k_{cat} = 0.13 s^{-1} (*cf.* wt-GCPII k_{cat} = 1.1 s^{-1}).

The F209W mutation leads to a slightly higher catalytic activity of GCPII than observed in the wild-type enzyme ($\Delta\Delta G^{\ddagger}_{expt}$ ~ −0.5 kcal.mol^{-1}). From the QM/MM optimizations, no significant changes are observed along the reaction coordinate. Also, the TPSS-D3//QM/MM energy profile is almost undistinguishable from the wt-GCPII, and the energies of the TS1, IM, and TS2 are within ca. 0.5 kcal.mol^{-1} from wt-GCPII. The rate-limiting TS2 is 0.5 kcal.mol^{-1} higher in energy, which is way below the expected error of available computational methods. Therefore, we cannot conclude with certainty whether the higher turnover number originates from the effects directly associated with the reaction energetics in the active site, or whether some other (distal) effects take place, such as the conformational flexibility of the protein, or the subtle changes in the substrate binding and the dynamics of the nearby residues not accounted for in our QM/MM model.

4.3. Methodological Issues.

An extensive QM system consisting of ca. 320 atoms brings multiple advantages, such as the possibility to discern the effects of single-point mutations directly on the Zn site within the same level of theory. This allows us to reveal even minor effects of how the 'new' residues interact with the Zn site and evaluate the impact of specific mutations along the full reaction coordinate. In the previous section, we were thus able to qualitatively discuss the effect of each mutation to the turnover number from the comparison of the geometric and electronic differences the mutations imposed on the substrate, the E424 proton shuttle, the reacting hydroxyl nucleophile, or both Zn ions' coordination spheres.

On the other hand, such a large QM region also raises several methodological issues that should be mentioned. Large computational expenses are associated with evaluating even a single reaction coordinate (e.g., for the wt-GCPII). Herein, we compared eight different variants of GCPII (wt-GCPII and E424A, E424H, R210A, R210K, Y552I, Y700F, and F209W mutants). This was possible because of the almost identical reaction mechanism, differing only by small deviations in the reaction intermediates' geometries and energies. Still, the completion of the project required several rounds of optimizations of each of the reaction intermediates due to the considerable flexibility of the QM region and the associated chemical 'hysteresis' effect. The forward and backward reaction was repeated until convergence of the equilibrium geometry and energy of the intermediates. The very same flexibility of the QM region also led us not to include two other initially tested GCPII variants, namely R534L and R536L. The arginine residues mutations required another 3 water molecules to be added to fill the space occupied by the arginines. However, this greatly added to the system's complexity which became intractable by the QM/MM methodology. The backward and forward reactions always provided intermediates with largely reordered water molecules, thus differing in geometries and energies. Substantial sampling, by computationally intensive QTCP or QM/MM-FEP methods might be required, and that is beyond the scope of our work.

Additionally, the large quantum region has led to ZPVE corrections (and also entropic corrections *via* the standard calculation of the vibrational, rotational, and translational contributions to the overall partition function) of different magnitudes based on the methodology that was used for frequency calculations. We have tested two approaches: (i) numerical calculation of frequencies in the QM region embedded in the MM pointcharges (i.e., direct QM/MM-like frequency calculations), and (ii) reoptimizing the QM region in the gas phase, followed by the analytical harmonic frequency calculations. In these, we have computed the frequencies on the complete QM systems, whereas the QM/MM link atoms were fixed at the crystal positions and excluded from the frequency analysis by assigning the infinite mass. Besides the large deviations in energy corrections, both approaches provided a large number of imaginary frequencies, mainly originating from the numerical instability of frequency calculations. Therefore, the transition states structures were instead confirmed by performing one- or two-dimensional scans on the potential energy surfaces. And, the unreliable thermal enthalpic and entropic contributions to the free energy surface were neglected in the calculations. Here, we thus rely on comparing the differences in the reaction coordinates for various GCPII variants on the ΔH level. Considering the reliability of the two approaches, the first one appears to be more robust, providing the ZPVE corrections that only rarely exceed the threshold of ~1 kcal.mol^{-1} in the relative TS1 and TS2 energies. While the latter approach was repeatedly utilized in smaller model systems, the gas-phase optimization seems to affect the large QM model significantly, leading to more notable contributions to the free energies that are violating the trends observed in the experimental turnover numbers and this approach is not recommended.

Finally, we note that different methods (i.e., choice of the functional or the basis set) might lead to slightly different conclusions. However, we found that GCPII system is quite robust *wrt* the functional selection, and the shift in energies appears to be systematic between the different methods. Therefore, the differences (not absolute values) in the activation enthalpies as compared in **Figure 4** are only slightly affected when comparing non-hybrid (TPSS) and hybrid (B3LYP, TPSSh) functionals (*cf.* **Table S3**). The selection of a smaller basis set appears to be more dramatic and leads to a severe underestimation of the TS1 and IM energy by ca. 10 kcal mol^{-1} in all of the cases. On the other hand, the TS2 remains unaffected and is described within ca. 2 kcal mol^{-1} accuracy (*cf.* **Table S4**).

5. Discussion

The primary role of theoretical modeling in understanding the enzyme catalysis is to elucidate the reaction mechanism of a particular enzyme that is to describe the reaction coordinate corresponding to lowest (free) energy profile. A practical test of a particular computational methodology is its usefulness

for the 'real-world' biochemical problems, such as influence of

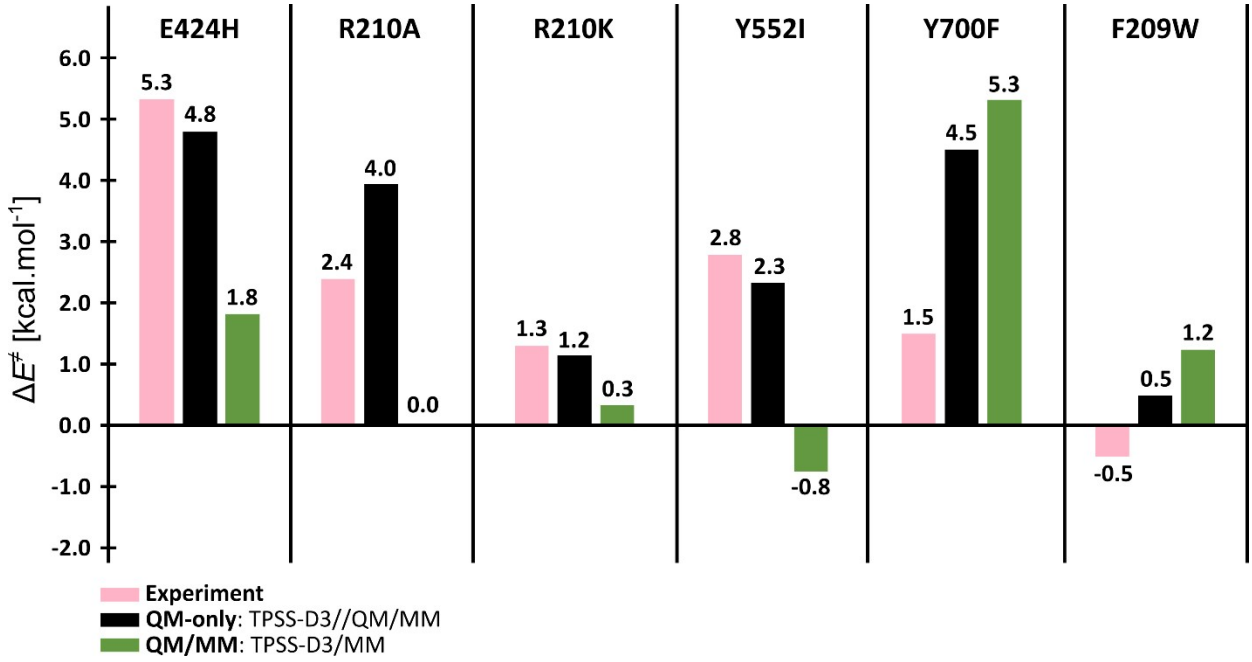


Figure 4. The relative activation energies for the rate determining step of the NAAG hydrolysis computed by the QM/MM and QM-only (on top of the QM/MM geometries). For all of the cases, the QM-only approach provides relative $\Delta H^{\ddagger}$ that is closer to the value derived from the experimental turnover number. Note that E424A (not shown) is predicted to be inactive in both computational schemes, which is consistent with the experiment. Raw energetic data are presented in **Table S5**.

single-point mutations on enzyme activity (whether natural or artificial). The ongoing discussion revolves around comparing (or contrasting) the two most commonly used approaches: QM/MM and cluster (QM-only) modeling. While the accuracy of the former suffers from the (in principle) less accurate MM part and the QM-MM coupling, the latter will be always associated with errors related to the incompleteness of the system (such as flexibility of the protein framework, which is excluded from typical cluster model approaches, or long-range effects from the protein bulk). In this study, employing GCPII as the test system, we argue that it is the combination of the two that leads to the most consistent results. This is in accordance with our previous QM/MM and QM studies on protein-ligand binding and metalloenzymatic reactions mechanisms.[25,63,86,87] The QM/MM optimization allows to obtain a correct "protein geometry" and in our opinion also yields a smoother potential energy surface for the particular enzymatic reaction. On the contrary, some transition states obtained in cluster models (QM-only) may suffer from (artificial) constraints on the QM boundary and may in fact disappear within the context of the relaxed protein surrounding. This can be seen by comparing our reaction profiles to the similar profiles obtained for aminopeptidases by Himo and coworkers[88] which are characterized by a greater number of transition states. These are often not relevant for the overall (highest) activation energies, but unnecessarily complicate the mechanistic description of the reaction mechanism.

Using the protocol, we show that the agreement between computed and experimental values of the activation energies (rate constants) is satisfactory (semi-quantitative). We mostly predicted experimental observations to within 1-2 kcal.mol^{-1} accuracy (**Figure 4**). The absolute values are somewhat higher, but this has been thoroughly discussed in Ref. 60 and might be partially attributed to an overestimation of the activation barriers computed by the DFT. The only exception has been the Y700F mutant where the calculations predicted a decrease in the turnover number approximately by more than three orders of magnitude whereas experimentally the reaction is slower by only one order. For most of the mutants, calculations provided a plausible hypothesis concerning the origin of the

variations in the observed turnover numbers which may in general lead to an optimization of the catalytic efficiency of the particular enzyme, saving a decent amount of experimental work.

The combined protocol presented herein avoids all the subtleties of the QM/MM procedure (mechanical vs. electronic embedding, zeroing, and 'displacing' the charges on the atoms neighboring with the QM region, treating of the junction atoms, etc.) that may severely hamper the reaction energy profiles.[13,66] In contrast, equilibrium geometries are thought to be satisfactorily predicted using various QM/MM approaches, and when followed by the technically more straightforward QM-only calculations, the approach might provide more reliable reaction energetics. Simultaneously, we must argue that the results presented herein might not represent a general setback of the QM/MM methodology, but might instead be characteristic for GCPII and other related proteins. We speculated that the higher (inaccurate) activation energies originate from the QM region in the field of surrounding point-charges entering into the QM calculations via one-electron Hamiltonian. However, this might be associated with a significant charge surplus in the active site in GCPII (note that NAAG substrate has a total charge of -3). In **Note S1** in the Supporting Information, we demonstrate that electrostatic energy is indeed significantly lower in the RC of wt-GCPII, overall increasing the energies of TS1 and IM.

6. Conclusions

In this work, we demonstrated the applicability of the combined QM/MM and QM protocol to (semi-)quantitatively predict the effect of single-point mutations in an enzyme (in our case GCPII) on the rate constant (turnover number, k_{cat}). We argued that the former (QM/MM) is recommended for providing the optimal geometries of the reactant (Michaelis complex), transition states, intermediates, and products, whereas the energy profile is preferably computed by QM-only approach in a dielectric continuum. Thus, one avoids the inaccuracies in the MM part and in the QM and MM coupling. Such an approach might be recommended provided that a sufficiently large QM system is used (in our case ~320 atoms). In addition to the Zn_2 active site, and the second-sphere residues, we have also included all mutated residues. Once the (semi-)quantitative agreement between experimental and theoretical kinetic data has been obtained, a qualitative explanation (based on the computed numbers) has been provided for most of the mutants, and that leads to a better understanding of the peptide bond hydrolysis by GCPII. Based on the calculated data the new E424H mutation was conceived and the GCPII(E424H) mutant was experimentally characterized. It not only demonstates the mutually beneficial coupling between theory and experiment, but also addresses one of the most important issues in hydrolytic cleavage: the role of the general acid/base in the catalytic process. In GCPII, it thus appears that there is no substitute for the wt E424 residue.

ASSOCIATED CONTENT

Supporting Information. Equilibrium geometries of all studied systems (QM regions in QM/MM simulations and full PDBs); Additional Computational Details; Supplementary Tables S1-S8, comparing results obtained with different computational methods; Supplementary Figure S3; and Note S1 commenting on the origin of the electrostatic imbalance along the GCPII reaction coordinate. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

E-mail: cyril.barinka@ibt.cas.cz;
rulisek@uochb.cas.cz

Notes

The authors declare no competing financial interest.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

We thank Barbora Havlinova for her excellent technical assistance and Dr. Tibor Andras Rokob for the help with the protein preparation/setup and technical subtleties of QM/MM calculations. The financial support of the Ministry of Education, Youth and Sports (project LTAUSA19148 to L.R.), the Grant Agency of the Czech Republic (grants 18-14167S to C.B. and 19-22269Y to Z.K.), the Czech Academy of Sciences (RVO: 86652036), and National Science Foundation (NSF CHE-1903808 to. A.N.A.) is gratefully acknowledged. Computational

resources were in part provided by the IT4Innovations National Supercomputing Center, under the project LM2015070 (Large Infrastructures for Research, Experimental Development and Innovations), UCLA-IDRE cluster Hoffman2, and Extreme Science and Engineering Discovery Environment (XSEDE).

REFERENCES

(1) Thiel, W. Computational Catalysis—Past, Present, and Future. *Angew. Chemie Int. Ed.* **2014**, *53*, 8605–8613.

(2) Harvey, J. N.; Himo, F.; Maseras, F.; Perrin, L. Scope and Challenge of Computational Methods for Studying Mechanism and Reactivity in Homogeneous Catalysis. *ACS Catal.* **2019**, *9*, 6803–6813.

(3) Świderek, K.; Tuñón, I.; Moliner, V. Predicting Enzymatic Reactivity: From Theory to Design. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 407–421.

(4) Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J. Protocol for Computational Enzymatic Reactivity Based on Geometry Optimisation. *ChemPhysChem* **2018**, *19*, 669–689.

(5) Ahmadi, S.; Barrios Herrera, L.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R. Multiscale Modeling of Enzymes: QM-Cluster, QM/MM, and QM/MM/MD: A Tutorial Review. *Int. J. Quantum Chem.* **2018**, *118*, e25558.

(6) **Saura, P.; Röpke, M.; Gamiz-Hernandez, A. P.; Kaila, V. R. I. Quantum Chemical and QM/MM Models in Biochemistry. In *Biomolecular Simulations: Methods and Protocols*; Bonomi, M., Camilloni, C.; Springer New York: New York, NY, 2019, 75–104.**

(7) **Stiebritz, M. T.; Hu, Y. Computational Methods for Modeling Metalloproteins. In *Metalloproteins: Methods and Protocols*; Hu, Y.; Springer New York: New York, NY, 2019, 245–266.**

(8) Magalhães, R. P.; Fernandes, H. S.; Sousa, S. F. Modelling Enzymatic Mechanisms with QM/MM Approaches: Current Status and Future Challenges. *Isr. J. Chem.* **2020**, *60*, 655–666.

(9) **Ryde, U. QM/MM Calculations on Proteins. In *Computational Approaches for Studying Enzyme Mechanism Part A*; Voth, G. A.; Ed.; Academic Press, 2016, *577*, 119–158.**

(10) Quesne, M. G.; Borowski, T.; de Visser, S. P. Quantum Mechanics/Molecular Mechanics Modeling of Enzymatic Processes: Caveats and Breakthroughs. *Chem. - Eur. J.* **2016**, *22*, 2562–2581.

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

(12) Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; 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.

(13) Rovira, C. The Description of Electronic Processes inside Proteins from Car-Parrinello Molecular Dynamics: Chemical Transformations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 393–407.

(14) Wei, W.-J.; Qian, H.-X.; Wang, W.-J.; Liao, R.-Z. Computational Understanding of the Selectivities in Metalloenzymes. *Front. Chem.* **2018**, *6*, 638.

(15) Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780–6786.

(16) Sheng, X.; Kazemi, M.; Planas, F.; Himo, F. Modeling Enzymatic Enantioselectivity Using Quantum Chemical Methodology. *ACS Catal.* **2020**, *10*, 6430–6449.

(17) Nechay, M. R.; Valdez, C. E.; Alexandrova, A. N. Computational Treatment of Metalloproteins. *J. Phys. Chem. B* **2015**, *119*, 5945–5956.

(18) Kamerlin, S. C. L.; Warshel, A. The Empirical Valence Bond Model: Theory and Applications. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 30–45.

(19) Siegbahn, P. E. M.; Himo, F. The Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 323–336.

(20) Blomberg, M. R. A. How Quantum Chemistry Can Solve Fundamental Problems in Bioenergetics. *Int. J. Quantum Chem.* **2015**, *115*, 1197–1201.

(21) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum

Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601–3658.
(22) Gallup, N. M.; Alexandrova, A. N. Use of QM/DMD as a Multiscale Approach to Modeling Metalloenzymes. In *Computational Approaches for Studying Enzyme Mechanism Part A*; Voth, G. A.; Ed.; Academic Press, **2016**, *577*, 319–339.
(23) Karasulu, B.; Patil, M.; Thiel, W. Amine Oxidation Mediated by Lysine-Specific Demethylase 1: Quantum Mechanics/Molecular Mechanics Insights into Mechanism and Role of Lysine 661. *J. Am. Chem. Soc.* **2013**, *135*, 13400–13413.
(24) Dutta, D.; Mishra, S. Active Site Dynamics in Substrate Hydrolysis Catalyzed by DapE Enzyme and Its Mutants from Hybrid QM/MM-Molecular Dynamics Simulation. *J. Phys. Chem. B* **2017**, *121*, 7075–7085.
(25) Serapian, S. A.; van der Kamp, M. W. Unpicking the Cause of Stereoselectivity in Actinorhodin Ketoreductase Variants with Atomistic Simulations. *ACS Catal.* **2019**, *9*, 2381–2394.
(26) Lawan, N.; Chasing, P.; Santatiwongchai, J.; Muangpil, S. QM/MM Molecular Modelling on Mutation Effect of Chorismate Synthase Enzyme Catalysis. *J. Mol. Graph. Model.* **2019**, *87*, 250–256.
(27) Suñer-Rubio, A.; Cebrián-Prats, A.; González-Lafont, À.; Lluch, J. M. Unraveling How the Gly526Ser Mutation Arrests Prostaglandin Formation from Arachidonic Acid Catalyzed by Cyclooxygenase-2: A Combined Molecular Dynamics and QM/MM Study. *RSC Adv.* **2020**, *10*, 986–997.
(28) Blomberg, M. R. A. Can Reduction of NO to N 2 O in Cytochrome c Dependent Nitric Oxide Reductase Proceed through a Trans-Mechanism? *Biochemistry* **2017**, *56*, 120–131.
(29) Maršavelski, A.; Vianello, R. What a Difference a Methyl Group Makes: The Selectivity of Monoamine Oxidase B Towards Histamine and N -Methylhistamine. *Chem. - Eur. J.* **2017**, *23*, 2915–2925.
(30) Wojdyła, Z.; Borowski, T. DFT Study of the Mechanism of Manganese Quercetin 2,3-Dioxygenase: Quest for Origins of Enzyme Unique Nitroxygenase Activity and Regioselectivity. *J. Biol. Inorg. Chem.* **2016**, *21*, 475–489.
(31) Siegbahn, P. E. M. Model Calculations Suggest That the Central Carbon in the FeMo-Cofactor of Nitrogenase Becomes Protonated in the Process of Nitrogen Fixation. *J. Am. Chem. Soc.* **2016**, *138*, 10485–10495.
(32) Lan, C.-L.; Chen, S.-L. The Decarboxylation of α,β-Unsaturated Acid Catalyzed by Prenylated FMN-Dependent Ferulic Acid Decarboxylase and the Enzyme Inhibition. *J. Org. Chem.* **2016**, *81*, 9289–9295.
(33) Bím, D.; Chalupský, J.; Culka, M.; Solomon, E. I.; Rulíšek, L.; Srnec, M. Proton–Electron Transfer to the Active Site Is Essential for the Reaction Mechanism of Soluble Δ^9-Desaturase. *J. Am. Chem. Soc.* **2020**, *142*, 10412–10423.
(34) Srnec, M.; Solomon, E. I. Frontier Molecular Orbital Contributions to Chlorination versus Hydroxylation Selectivity in the Non-Heme Iron Halogenase SyrB2. *J. Am. Chem. Soc.* **2017**, *139*, 2396–2407.
(35) Zapata-Torres, G.; Fierro, A.; Barriga-González, G.; Salgado, J. C.; Celis-Barros, C. Revealing Monoamine Oxidase B Catalytic Mechanisms by Means of the Quantum Chemical Cluster Approach. *J. Chem. Inf. Model.* **2015**, *55*, 1349–1360.
(36) Bykov, D.; Neese, F. Six-Electron Reduction of Nitrite to Ammonia by Cytochrome c Nitrite Reductase: Insights from Density Functional Theory Studies. *Inorg. Chem.* **2015**, *54*, 9303–9316.
(37) Cassimjee, K. E.; Manta, B.; Himo, F. A Quantum Chemical Study of the ω-Transaminase Reaction Mechanism. *Org. Biomol. Chem.* **2015**, *13*, 8453–8464.
(38) Piazzetta, P.; Marino, T.; Russo, N. The Working Mechanism of the β-Carbonic Anhydrase Degrading Carbonyl Sulphide (COSase): A Theoretical Study. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14843–14848.
(39) Hernández-Ortega, A.; Quesne, M. G.; Bui, S.; Heyes, D. J.; Steiner, R. A.; Scrutton, N. S.; de Visser, S. P. Catalytic Mechanism of Cofactor-Free Dioxygenases and How They Circumvent Spin-Forbidden Oxygenation of Their Substrates. *J. Am. Chem. Soc.* **2015**, *137*, 7474–7487.
(40) Hopmann, K. H. Full Reaction Mechanism of Nitrile Hydratase: A Cyclic Intermediate and an Unexpected Disulfide Switch. *Inorg.*

Chem. **2014**, *53*, 2760–2762.

(41) Reilley, D. J.; Popov, K. I.; Dokholyan, N. V.; Alexandrova, A. N. Uncovered Dynamic Coupling Resolves the Ambiguous Mechanism of Phenylalanine Hydroxylase Oxygen Binding. *J. Phys. Chem. B* **2019**, *123*, 4534–4539.

(42) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chemie Int. Ed.* **2009**, *48*, 1198–1229.

(43) 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.

(44) Siegbahn, P. E. M. A Quantum Chemical Approach for the Mechanisms of Redox-Active Metalloenzymes. *RSC Adv.* **2021**, *11*, 3495–3508.

(45) Ribeiro, A. J. M.; 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.

(46) Kazemi, M.; Himo, F.; Åqvist, J. Enzyme Catalysis by Entropy without Circe Effect. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2406–2411.

(47) Klusák, V.; Bařinka, C.; Plechanovová, A.; Mlčochová, P.; Konvalinka, J.; Rulíšek, L.; Lubkowski, J. Reaction Mechanism of Glutamate Carboxypeptidase II Revealed by Mutagenesis, X-Ray Crystallography, and Computational Methods. *Biochemistry* **2009**, *48*, 4126–4138.

(48) Duarte, F.; Amrein, B. A.; Blaha-Nelson, D.; Kamerlin, S. C. L. Recent Advances in QM/MM Free Energy Calculations Using Reference Potentials. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850*, 954–965.

(49) Huggins, D. J.; Biggin, P. C.; Dämgen, M. A.; Essex, J. W.; Harris, S. A.; Henchman, R. H.; Khalid, S.; Kuzmanic, A.; Laughton, C. A.; Michel, J.; Mulholland, A. J.; Rosta, E.; Sansom, M. S. P.;| van der Kamp, M. W. Biomolecular Simulations: From Dynamics and Mechanisms to Computational Assays of Biological Activity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2019**, *9*, e1393.

(50) Boulanger, E.; Harvey, J. N. QM/MM Methods for Free Energies and Photochemistry. *Curr. Opin. Struct. Biol.* **2018**, *49*, 72–76.

(51) Robinson, M. B.; Blakely, R. D.; Couto, R.; Coyle, J. T. Hydrolysis of the Brain Dipeptide N-Acetyl-L-Aspartyl-L-Glutamate. Identification and Characterization of a Novel N-Acetylated Alpha-Linked Acidic Dipeptidase Activity from Rat Brain. *J. Biol. Chem.* **1987**, *262*, 14498–14506.

(52) Rovenská, M.; Hlouchová, K.; Šácha, P.; Mlčochová, P.; Horák, V.; Zámečník, J.; Bařinka, C.; Konvalinka, J. Tissue Expression and Enzymologic Characterization of Human Prostate Specific Membrane Antigen and Its Rat and Pig Orthologs. *Prostate* **2008**, *68*, 171–182.

(53) Sokoloff, R. L.; Norton, K. C.; Gasior, C. L.; Marker, K. M.; Grauer, L. S. A Dual-Monoclonal Sandwich Assay for Prostate-Specific Membrane Antigen: Levels in Tissues, Seminal Fluid and Urine. *Prostate* **2000**, *43*, 150–157.

(54) Hupe, M. C.; Philippi, C.; Roth, D.; Kümpers, C.; Ribbat-Idel, J.; Becker, F.; Joerg, V.; Duensing, S.; Lubczyk, V. H.; Kirfel, J.; Sailer, V.; Kuefer, R.; Merseburger, A. S.; Perner, S.; Offermann, A. Expression of Prostate-Specific Membrane Antigen (PSMA) on Biopsies Is an Independent Risk Stratifier of Prostate Cancer Patients at Time of Initial Diagnosis. *Front. Oncol.* **2018**, *8*, 623.

(55) Mesters, J. R.; Barinka, C.; Li, W.; Tsukamoto, T.; Majer, P.; Slusher, B. S.; Konvalinka, J.; Hilgenfeld, R. Structure of Glutamate Carboxypeptidase II, a Drug Target in Neuronal Damage and Prostate Cancer. *EMBO J.* **2006**, *25*, 1375–1384.

(56) Pavlicek, J.; Ptacek, J.; Barinka, C. Glutamate Carboxypeptidase II: An Overview of Structural Studies and Their Importance for Structure-Based Drug Design and Deciphering the Reaction Mechanism of the Enzyme. *Curr. Med. Chem.* **2012**, *19*, 1300–1309.

(57) Bařinka, C.; Rovenská, M.; Mlčochová, P.; Hlouchová, K.; Plechanovová, A.; Majer, P.; Tsukamoto, T.; Slusher, B. S.; Konvalinka, J.; Lubkowski, J. Structural Insight into the Pharmacophore Pocket of Human Glutamate Carboxypeptidase II. *J. Med. Chem.* **2007**, *50*, 3267–3273.

(58) Barinka, C.; Hlouchova, K.; Rovenska, M.; Majer, P.; Dauter, M.; Hin, N.; Ko, Y.-S.; Tsukamoto, T.; Slusher, B. S.; Konvalinka, J.; Lubkowski, J. Structural Basis of

Interactions between Human Glutamate Carboxypeptidase II and Its Substrate Analogs. *J. Mol. Biol.* **2008**, *376*, 1438–1450.
(59) Mlčochová, P.; Plechanovová, A.; Bařinka, C.; Mahadevan, D.; Saldanha, J. W.; Rulíšek, L.; Konvalinka, J. Mapping of the Active Site of Glutamate Carboxypeptidase II by Site-Directed Mutagenesis. *FEBS J.* **2007**, *274*, 4731–4741.
(60) Navrátil, V.; Klusák, V.; Rulíšek, L. Theoretical Aspects of Hydrolysis of Peptide Bonds by Zinc Metalloenzymes. *Chem. Eur. J.* **2013**, *19*, 16634–16645.
(61) Navrátil, M.; Tykvart, J.; Schimer, J.; Pachl, P.; Navrátil, V.; Rokob, T. A.; Hlouchová, K.; Rulíšek, L.; Konvalinka, J. Comparison of Human Glutamate Carboxypeptidases II and III Reveals Their Divergent Substrate Specificities. *FEBS J.* **2016**, *283*, 2528–2545.
(62) Plechanovová, A.; Byun, Y.; Alquicer, G.; Škultétyová, L.; Mlčochová, P.; Němcová, A.; Kim, H.-J.; Navrátil, M.; Mease, R.; Lubkowski, J.; Pomper, M.; Konvalinka, J.; Rulíšek, L.; Bařinka, C. Novel Substrate-Based Inhibitors of Human Glutamate Carboxypeptidase II with Enhanced Lipophilicity. *J. Med. Chem.* **2011**, *54*, 7535–7546.
(63) Barinka, C.; Novakova, Z.; Hin, N.; Bím, D.; Ferraris, D. V.; Duvall, B.; Kabarriti, G.; Tsukamoto, R.; Budesinsky, M.; Motlova, L.; Rojas, C.; Slusher, B. S.; Rokob, T. A.; Rulíšek, L.; Tsukamoto, T. Structural and Computational Basis for Potent Inhibition of Glutamate Carboxypeptidase II by Carbamate-Based Inhibitors. *Bioorg. Med. Chem.* **2019**, *27*, 255–264.
(64) Ryde, U. The Coordination of the Catalytic Zinc Ion in Alcohol Dehydrogenase Studied by Combined Quantum-Chemical and Molecular Mechanics Calculations. *J. Comput.-Aided. Mol. Des.* **1996**, *10*, 153–164.
(65) Ryde, U.; Olsson, M. H. M. Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. *Int. J. Quantum Chem.* **2001**, *81*, 335–347.
(66) Cao, L.; Ryde, U. On the Difference Between Additive and Subtractive QM/MM Calculations. *Front. Chem.* **2018**, *6*, 89.
(67) **Rokob, T. A.;** Rulíšek, L. Curvature correction for microiterative optimizations with QM/MM electronic embedding. *J. Comput. Chem.* **2012**, *33*, 1197-1206.
(68) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
(69) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta--Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
(70) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
(71) 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.
(72) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240*, 283–290.
(73) 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.
(74) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137.
(75) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(76) Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc. Perkin Trans. 2* **1993**, 799–805.
(77) Barinka, C.; Ptacek, J.; Richter, A.; Novakova, Z.; Morath, V.; Skerra, A. Selection and Characterization of Anticalins Targeting Human Prostate-Specific

Membrane Antigen (PSMA). *Protein Eng., Des. Sel.* **2016**, *29*, 105–115.

(78) Tykvart, J.; Šácha, P.; Bařinka, C.; Knedlík, T.; Starková, J.; Lubkowski, J.; Konvalinka, J. Efficient and Versatile One-Step Affinity Purification of in Vivo Biotinylated Proteins: Expression, Characterization and Structure Analysis of Recombinant Human Glutamate Carboxypeptidase II. *Protein Expr. Purif.* **2012**, *82*, 106–115.

(79) Barinka, C.; Rinnová, M.; Šácha, P.; Rojas, C.; Majer, P.; Slusher, B. S.; Konvalinka, J. Substrate Specificity, Inhibition and Enzymological Analysis of Recombinant Human Glutamate Carboxypeptidase II. *J. Neurochem.* **2002**, *80*, 477–487.

(80) **Sumner, S.; Söderhjelm, P.; Ryde, U. Effect of Geometry Optimizations on QM-Cluster and QM/MM Studies of Reaction Energies in Proteins. *J. Chem. Theory Comput.* 2013, *9*, 4205–4214.**

(81) Morgenstern, A.; Jaszai, M.; Eberhart, M. E.; Alexandrova, A. N. Quantified Electrostatic Preorganization in Enzymes Using the Geometry of the Electron Charge Density. *Chem. Sci.* **2017**, *8*, 5010–5018.

(82) Fuller, J.; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N. Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *J. Chem. Inf. Model.* **2019**, *59*, 2367–2373.

(83) Wang, L.; Fried, S. D.; Markland, T. E. Proton Network Flexibility Enables Robustness and Large Electric Fields in the Ketosteroid Isomerase Active Site. *J. Phys. Chem. B* **2017**, *121*, 9807–9815.

(84) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235.

(85) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.

(86) Vavřina, Z.; Gutten, O.; Smola, M.; Zavřel, M.; Aliakbar Tehrani, Z.; Charvát, V.; Kožíšek, M.; Boura, E.; Birkuš, G.; Rulíšek, L. Protein–Ligand Interactions in the STING Binding Site Probed by Rationally Designed Single-Point Mutations: Experiment and Theory. *Biochemistry* **2021**, *60*, 607–620.

(87) Smola, M.; Gutten, O.; Dejmek, M.; Kožíšek, M.; Evangelidis, T.; Aliakbar Tehrani, Z.; Novotná, B.; Nencka, R.; Birkuš, G.; Rulíšek, L.; Boura, E. Ligand Strain and Its Conformational Complexity Is a Major Factor in the Binding of Cyclic Dinucleotides to STING Protein. *Angew. Chemie Int. Ed.* **2021**, *60*, 10172–10178.

(88) Chen, S.-L.; Marino, T.; Fang, W.-H.; Russo, N.; Himo, F. Peptide Hydrolysis by the Binuclear Zinc Enzyme Aminopeptidase from Aeromonas Proteolytica: A Density Functional Theory Study. *J. Phys. Chem. B* **2008**, *112*, 2494–2500.

TOC GRAPHICS