Received: 29 March 2021 | Revised: 16 June 2021 | Accepted: 30 June 2021

DOI: 10.1002/qua.26768

REVIEW

International Journal of Quantum Chemistry WILEY

The role of quantum chemistry in covalent inhibitor design

Levente M. Mihalovits | György G. Ferenczy | György M. Keserű

Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Budapest, Hungary

Correspondence
György G. Ferenczy and György M. Keserű, Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Magyar tudósok körútja 2, Budapest 1117, Hungary.
Email: ferenczy.gyorgy@ttk.hu and keseru.gyorgy@ttk.hu

Funding information
Nemzeti Kutatási Fejlesztési és Innovációs Hivatal, Grant/Award Numbers: SNN 125496, SNN 135335

Abstract

The recent ascent of targeted covalent inhibitors (TCI) in drug discovery brings new opportunities and challenges to quantum chemical reactivity calculations supporting discovery efforts. TCIs typically form a covalent bond with the targeted nucleophilic amino acid side chain. Their reactivity that can be both computed and experimentally measured is therefore one of the key factors in determining inhibitory potency. Calculation of relevant quantum chemical descriptors and corresponding reaction barriers of model reactions represent efficient ways to predict intrinsic reactivities of covalent ligands. A more comprehensive description of covalent ligand binding is offered by mixed quantum mechanical/molecular mechanical (QM/MM) potentials. Reaction mechanisms can be investigated by the exploration of the potential energy surface as a function of suitable reaction coordinates, and free energy surfaces can also be calculated with molecular dynamics based simulations. Here we review the methodological aspects and discuss applications with primary focus on high-end QM/MM simulations to illustrate the current status of quantum chemical support to covalent inhibitor design. Available QM approaches are suitable to identify likely reaction mechanisms and rate determining steps in the binding of covalent inhibitors. The efficient QM/MM prediction of ligand reactivities complemented with the computational description of the recognition step makes these computations highly useful in covalent drug discovery.

KEYWORDS
free energy, molecular mechanics/quantum mechanics, reaction barrier, reactivity, targeted covalent inhibitors

1 | INTRODUCTION

During the past decade, the design of targeted covalent inhibitors (TCIs) has gained increased attention. Formerly, covalent inhibitors were typically filtered out in drug discovery programs due to the risk of off-target activity attributed to their reactivity. Few compounds acting by covalent mechanism of action were discovered serendipitously. However, a paradigm change has occurred around the millennium owing to the recognition of distinct therapeutic advantages of covalent inhibition that include potentially full target occupancy and long-action, decoupling pharmacodynamics from pharmacokinetics. Nowadays TCIs are well represented in a wide range of therapies, treating various pathological conditions, such as cancer [1, 2], autoimmune diseases [3, 4], disorders of the central neural [5] and cardiovascular system [6], gastrointestinal related illnesses [7],

Dedicated to the memory of Professor István Mayer.

infections [8] and many others. Though the most abundant indication is oncology; 10 out of the 14 FDA-approved covalent drugs between 2011 and 2019 were anticancer medicines [9].

The binding of covalent inhibitors follows a two-step mechanism (Equation 1), namely the non-covalent binding, also known as the molecular recognition, and the covalent reaction in which the targeted nucleophilic sidechain of the enzyme reacts with the electrophilic center of the inhibitor molecule.

$$E + I \underset{}{\overset{K_i}{\rightleftharpoons}} E \cdot I \xrightarrow{k_{\text{inact}}} E - I \,(E: \text{enzyme}; I: \text{inhibitor}) \tag{1}$$

Typical targeted sidechains are serine, threonine, tyrosine, lysine and most importantly cysteine [10–12]. Due to the low occurrence of cysteine in the human proteome [13, 14] and its elevated nucleophilicity, cysteine targeting inhibitors possess lower chance of selectivity and toxicity issues [15], however, cysteine targeting is not an available or feasible option in some enzymes. Drug design attempts usually aims to improve both binding steps by optimizing the secondary non-covalent interactions between the molecule and the protein environment and fine tuning the warhead to achieve optimal covalent reactivity toward the targeted sidechain.

Computational methods traditionally support drug discovery campaigns primarily by predicting and/or evaluating non-covalent binding events and to guide optimization efforts. The exploding increase in computer performance and method developments have allowed researchers to perform more precise and computationally extensive calculations, which made computational studies highly effective in the field of non-covalent drug discovery [16–18]. Nevertheless, quantum mechanical (QM) approaches were not typical among the computational schemes used in non-covalent drug discovery programs owing to the size and complexity of the systems involved, and also to the primary importance of non-covalent interactions that are treated sufficiently at molecular mechanics level. However, the proper treatment of the chemical bond formation by covalent inhibitors requires the use of quantum chemical methods and the purpose of the present contribution is to review the diverse set of quantum chemical tools applied in the design of covalent inhibitors, in particular, the estimation of k_{inact} (Equation 1), the rate constant of the covalent bond formation. We first give a methodological overview of the various quantum chemistry approaches ranging from descriptor calculations in model systems to the evaluation of free energy changes of the chemical bond formation in extended systems. Then applications identifying covalent inhibition mechanisms and calculating rate constants for covalent bond formation are reviewed.

The comparison of computed and experimental potencies of inhibitors requires special considerations for covalent ligands. Reversible inhibitors are commonly characterized by K_i and IC_{50} values [19]. The K_i inhibition constant is the dissociation constant of the protein-ligand complex of competitive inhibitors, while IC_{50} expresses the ligand concentration that produces 50% inhibition of the biological activity [20]. However, in the case of irreversible covalent inhibition, K_i is associated with the first non-covalent step (cf. Equation 1) and IC_{50} values are time-dependent owing to the significant barrier of the chemical bond formation. Nevertheless, the dissociation constants of reversible covalent inhibitors are suitable measures of inhibitor potency as far as their slow binding equilibrium is considered. In contrast, irreversible covalent ligands are appropriately characterized by K_i and k_{inact} [21, 22]. In the forthcoming discussion we primarily consider k_{inact}, the rate constant of the chemical reaction, as the target of the computational investigations.

Studies of covalent inhibition relate to and benefit from investigations in two closely related fields. Toxicity predictions (see, e.g., References [23, 24]) often use descriptors, model reactions and experimental models that are relevant for cysteine TCIs. Theoretical studies of enzymatic reaction mechanisms [25, 26] use methodological tools common with computational covalent inhibitor studies, but they are not considered in the current review unless the mechanism of covalent inhibition is discussed.

2 | METHODS

2.1 | Descriptors

Comprehensive treatments of reactions in highly complex systems, like in solvated protein-ligand complexes are highly demanding, however, simplified models are useful in describing relevant aspects. Following the idea of ligand based drug design, reactivity studies can focus solely on the structure and properties of the inhibitor molecules, which can be investigated by quantum mechanics based molecular descriptors [27]. There is a large variety of descriptors, like HOMO-LUMO energies, net atomic charges, pK_a proton dissociation constants, NMR chemical shifts, polarizability, electron affinity, and so forth [28–31], which can be calculated by quantum mechanics and give information on molecular properties. Investigations on the relationship between QM descriptors and reactivity have a long history and are particularly widespread in the field of toxicology [32–34]. Typically, descriptors correlate with reactivity for compounds with a limited structural variability. Descriptor calculations are performed with effective quantum chemical methods (e.g., density functional theory [DFT] with hybrid functional and medium-sized basis set, or at semi-empirical level) [35], which allow obtaining the required quantities at a limited computational cost and with limited accuracy. This is, however,

most often a satisfactory compromise between reproducing modest differences in properties of similar compounds and the approximate nature of the relationship between descriptors and reactivities.

2.2 | Calculations on model system

A more demanding approach to estimate the reactivity of a ligand toward a protein sidechain is to perform QM calculations on model systems mimicking the labeling reaction. The selection of a model requires a judicious choice to ensure chemical relevance and computational tractability, and models typically include the ligand, or a truncated ligand containing the warhead, that is, the reactive electrophilic moiety, and a surrogate molecule representing the chemically relevant region of the reacting side-chain [29, 31]. A cysteine sidechain is usually modeled by a methyl-thiol or thiolate, while a serine sidechain by a methanol molecule, and further nucleophilic sidechains are modeled in a similar manner. The quantities of interest are the transition state (TS) barrier and the reaction energy of the model system. The former is directly related to the rate constant, while the latter is used to assess reversibility, although it may also correlate with reaction rates. Low barrier and highly negative reaction energy indicate potentially toxic compounds with off-target potency, while high barrier and positive reaction energy indicate inactivity (Figure 1).

Reaction barriers are obtained as the energy difference between the TS and the reactant state. Multiple choices are available in the computational toolbox to identify TSs and to calculate their energy. An estimation of the TS structure can be obtained by adiabatic scanning that includes constrained geometry optimizations with fixed values of the reaction coordinate. Scanning along the reaction coordinate that is typically an interatomic distance or a combination of several distances allows us to construct the potential energy curve, or surface in case of 2D scanning, and the approximate TS and its energy can be obtained [37]. An approximate TS can be used to initiate the search of a first order saddle point on the potential energy surface (PES). The Synchronous Transit guided Quasi Newton Method [38] takes into account the reactant and product structures, as well. The validation of the TS can be performed by frequency analysis and also by intrinsic reaction coordinate (IRC) calculations, confirming that the found TS connects the reactant and product states [39, 40]. The optimized geometry of the reactant state is a model for the non-covalent enzyme-ligand complex, while the product state is a model for the chemically bound ligand-enzyme complex.

Good quality reaction barriers can be obtained with DFT methods. B3LYP [41, 42] is extensively used for organic molecules, and is often extended with dispersion correction [43]. Depending on the system, precision requirement and available resources ωB97XD [44] and Minnesota type functionals [31, 45] are also used. The choice of the basis set also depends on the nature of the calculation and the applied theory. Pople

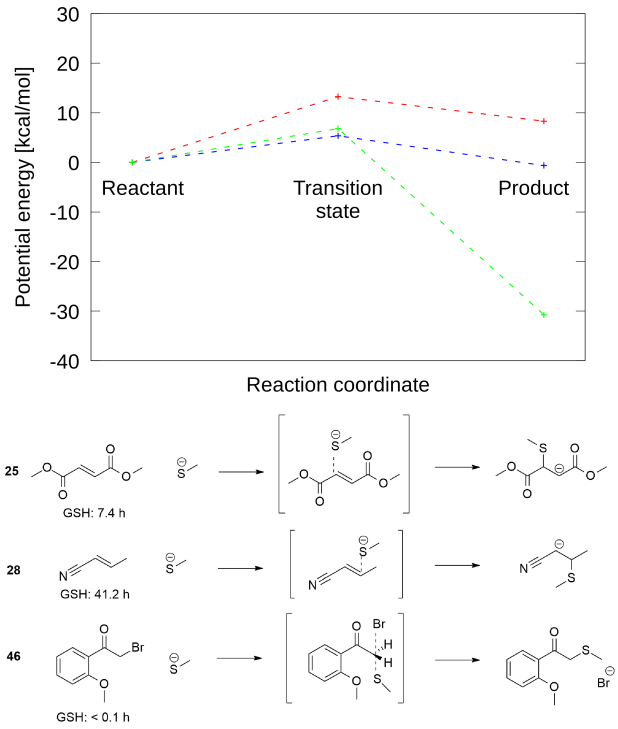


FIGURE 1 Typical energy profiles of the methyl-thiolate model reaction (top) and the mechanism of actions (bottom); inactive (red, 25), highly reactive (green, 46), potentially reversible (blue, 28) compound. Calculated energies and experimental GSH half-lifes were taken from Scarpino et al. [36]

type basis sets with diffuse and polarization functions are highly common [29], just as the use of Dunning type correlation consistent basis sets [46]. The application of implicit or explicit solvent models is also a crucial feature of model QM calculations, as they represent a better approximation to the protein environment than do vacuum calculations. Commonly used implicit water models include the polarizable continuum model (PCM) [47] and the universal solvation model (solvation model based on density [SMD]) [48]. The use of explicit water molecules may be necessary when they are directly involved in the electron reorganization or the implicit models are not sufficiently accurate owing to high charge concentration in the presence of ions [49].

Once the parameters of the model reaction have been determined, they can be compared to experimental values if the latter are available, or predictions can be made. It is notable that experimental reactivity data are often measured with a surrogate and this is typically glutathione (GSH) for cysteine nucleophiles. The half-life in adduct formation with GSH [30, 31, 50, 51] estimates the ligand reactivity against cysteine sidechains. Another option is comparing with experimental values obtained in the biological assay. In these cases, the considerations discussed in the introduction apply, and measured k_{inact} values are the most suitable to compare with the calculated rate constants. Nevertheless, IC_{50} values with constant incubation time may also be used with the assumption that the variation in inhibition is primarily affected by the barrier of the covalent reaction rather than by the equilibrium of the non-covalent binding, and the measured IC_{50} values well correlate with k_{inact}.

Although the inhibitory mechanism may include several steps, it is customary to consider the rate determining step only. The identification of the rate determining step can be based on analogy, chemical intuition, or rigorous quantum chemical calculations. The bond formation with cysteine residues typically includes negatively charged thiolate, but the deprotonation step is not necessarily included in the investigation that is justified for low-barrier, ligand independent proton transfers, for example, in cysteine proteases.

2.3 | Quantum mechanical/molecular mechanical approaches

Although QM model calculations are valuable tools in estimating ligand reactivities they are unable to take full account of the protein environment. A more sophisticated approach is to treat the solvated protein-ligand complex by quantum mechanical/molecular mechanical (QM/MM) calculations [52, 53]. QM/MM methods are hybrid technics that treat different parts of the system at different levels of complexity. They allow reducing the number of QM atoms into a computationally manageable level, while treating the rest of the system with a more affordable force-field based method. The QM/MM approaches can be categorized in various ways as it is shown in Table 1.

The total energy of the system can be evaluated either by the additive or the subtractive approach [54]. The former sums up the energy of the QM and MM region's energy and adds the interaction energy (Equation 2), while the subtractive approach estimates the whole system energy by MM, and corrects it with the difference between the QM and MM energies of the QM subsystem (Equation 3). Interestingly, energy convergence as a function of the size of the QM subsystem proved to be independent of the choice of the energy calculation scheme as long as electrostatic embedding (see later) is applied for the interactions between the two subsystems [55]. Convergence in terms of the number of sampled configurations [56] and the size of the QM region [57] was investigated by Ryde and coworkers.

$$E_{total} = E_{QM(QM)} + E_{MM(MM)} + E_{QM-MM\ interaction} \tag{2}$$

$$E_{total} = E_{MM,total} + E_{QM(QM)} - E_{MM(QM)} \tag{3}$$

The QM region usually consists of the ligand or a truncated ligand including the warhead, and the targeted sidechain, or a relevant part of it, and may contain further pieces of the protein together with water molecules. The rest of the atoms are parts of the MM subsystem. It is often inevitable to cut chemical bonds when separating the QM and MM subsystems, and this leads to fragments with dangling bonds. The main approaches to fill missing valences are the link atom [58], the frozen localized orbital [59–61] and the boundary atom [62, 63] schemes. The position of the subsystem boundary requires special attention, since cutting a polarized covalent bond or disrupting a delocalized electron system might result in serious distortions in the calculated energies and geometries. The optimal choice for separating chemically bound subsystems is to cut an apolar, aliphatic C—C bond. Concerning protein sidechains C_α—C_β bonds are suitable candidates for bond cutting yielding CH_3-SH QM fragments when cysteines are treated with the link atom approach.

The interaction between the QM and MM subsystems requires further considerations. There are two main embedding schemes applied in QM/MM calculations on enzymatic systems. The mechanical embedding (ME) performs the QM calculations in the absence of the MM region, and the interaction between the two subsystems is treated at MM level. The ONIOM method [64] developed by Morokuma and coworkers is originally an ME method, and later it was extended to use more than two layers and to treat QM-MM interactions beyond the mechanical embedding. Electrostatic embedding (EE) calculates electrostatic interactions between the QM and MM regions by the inclusion of MM charges in the QM Hamiltonian. As force fields contain an extensive set of point charges for the estimation of electrostatic interactions at the MM level, it is a typical scenario to use the atom centered point charges in the effective QM Hamiltonian.

TABLE 1 Aspects of the QM/MM calculations

Technical detail	Possibilities, examples
Energy calculation scheme	Additive Subtractive
Definition of the QM region	Atoms of the inhibitor molecule + atoms of the reactive sidechain Adding a buffer region (ONIOM)
Subsystem separation	Link atom approach Frozen orbitals Boundary atoms
QM level	DFT functionals (DFTB3, B3LYP, ωB97XD, PBE) Ab initio (MP2) Semi-empirical (DFTB3, PM3, AM1)
Force-field	AMBER (FF14SB, GLYCAM, GAFF) OPLS3 CharMM
Interaction scheme	Mechanical embedding (ME) Electrostatic embedding (EE)
Type of calculation	Geometry optimization Single point calculation Adiabatic scan Molecular dynamics
MD methods	None Steered MD Umbrella sampling Metadynamics FEP

Note: Examples are focusing on reactivity calculations.
Abbreviations: MM, molecular mechanical; QM, quantum mechanical.

A large variety of MM force fields and QM theory can be applied to QM/MM calculations of covalent enzyme inhibition. MM force fields designed to describe proteins and their interactions with small molecules include AMBER (FF14SB) [65], OPLS3 [66], CHARMM (CGenFF) [67], and GROMOS [68]. Some of these have alternative versions to better describe specific systems such as membranes, nucleic acids, lipids or carbohydrates [69]. The choice of the optimal QM theory depends on the reaction studied and the details of the computational approach. The application of DFT methods is the most common, although the use of the sophisticated long range corrected hybrid functionals with large basis sets is restricted by their high computational demand. While single point calculations and adiabatic scans can be performed with involved functionals, the large number of evaluations required in molecular dynamics (MD) simulations is compatible with computationally more efficient methods (see later).

2.3.1 | QM/MM potential energy

Reaction mechanism studies can be performed by calculating QM/MM PESs [70–74]. They explore the reaction path similarly to calculations for model systems, but the inclusion of the solvated protein-ligand complex creates a realistic environment. This ensures the structural integrity of the active site, not always straightforward to guarantee in model systems, and takes into account long range interactions, primarily electrostatics, which are known to play a substantial role in enzyme catalysis [75], and might also affect the mechanism of covalent inhibition.

QM/MM potential energy calculations require a limited number of evaluations of the QM energy and the gradient for a system with a few tens of atoms. A series of optimizations can be performed to construct a potential energy profile or minimum energy profile (MEP), analogously to the pure QM adiabatic scan. With the help of the MEP, binding events and binding mechanisms can be modeled by estimating geometries, energy barriers, reaction energies and reaction paths. These calculations are feasible with fairly involved QM methods. DFT hybrid functionals with double or triple zeta basis set, or post Hartree–Fock methods can be used that can eventually be complemented by zero-point vibrational energy and thermal corrections for stationary points. These calculations, however, miss entropic contributions that can be recovered in QM/MM MD simulations discussed in the next section.

2.3.2 | QM/MM approaches for free energy calculations

The QM/MM potential coupled with MD [76, 77] or Monte Carlo (MC) [78] methods offers the opportunity to calculate the free energy of the formation of covalent enzyme-ligand complexes. Unbiased MD is an efficient tool to investigate low energy structures and their geometries, however, the energy profiles of covalent reactions usually contain high energy regions whose structures are not well represented in unbiased MD simulations. The sampling of the region of the reaction barrier becomes feasible with the introduction of biasing potentials. There is a plethora of such methods, most notably steered molecular dynamics (SMD), umbrella sampling (US), and metadynamics, which are widely used in the simulations of biochemical systems [77] and are well suited to the investigations of the chemical bond formation of covalent inhibitors.

Steered MD [79, 80] introduces a time dependent biasing potential, which forces the reaction to proceed and enables the sampling along a reaction coordinate. SMD simulations yield work curves that can be used to generate potential of mean force (PMF) profiles according to the Jarzynski equation [81, 82]. It is more customary to use SMD simulations to generate starting structures for US calculations, the latter providing PMF curves. The key parameters for SMDs are the velocity of the pulling and the associated force constant. Fast pulling results in highly non-equilibrium processes and corresponding work curves due to the lack of relaxation during the simulation, while slow pulling requires large computational resources. High force constants obstruct the desired level of sampling and may lead to high energy structures that are subject to undesired structural changes and chemical reactions. In contrast, too low bias might cause the system unable to climb up energy barriers and stuck in a local minimum.

US [83–85] simulations use a series of differently positioned biasing potentials along the examined reaction coordinates to ensure proper sampling of high energy regions, in particular transitions states and their proximity. The weighted histogram analysis method (WHAM) [86] is perhaps the most widely used post-processing method to construct the PMF from the biased distribution of states, but other strategies, like umbrella integration [87] can also be applied. The key parameters for US are the force constant of the biasing potential and the distance between the middle of the potentials in neighboring windows. Sampling should overlap in neighboring windows to obtain reliable PMF curves and free energies.

Metadynamics [88–90] apply history dependent repulsive potentials along the discovered reaction coordinate space to aid the sampling of rare events. The valleys in the free energy surface are filled with Gaussians until a barrier free energy surface is constructed. The PMF is reconstructed from the accumulated bias potential. Essential parameters for metadynamics are the height and width of the Gaussians and the deposition time. Too frequent Gaussian addition spoils the resolution of the PMF, while too high Gaussian functions distorts the calculated barrier heights [91].

MD simulation based QM/MM methods require a large number of evaluations of the QM energy and its gradient. This represents a restriction to the computational demand and quality of the QM method. Approximate, most often semiempirical QM methods, like AM1, PM6 or DFTB3 are commonly applied [45, 92–94]. It should be noted, however, that these methods are subject to errors that are system and reaction type dependent. Smith [95] and coworkers performed benchmark calculations for model thio-Michael additions in gas phase with different DFT methods, and the results showed that a large number of functionals, including B3LYP are failing to find the correct minima of the enolate intermediate, due to the delocalization error [96, 97]. Their suggestion was to use range separated functional, such as ωB97XD which gave similar results to ab initio methods when modeling Michael-additions. A similar investigation [98] of the gas phase Michael addition reaction of selected small molecules with methyl thiolate also showed that several functionals and semiempirical Hamiltonians fail to predict stable enolate intermediates observed with sophisticated DFT functionals. However, it was also shown that the PMF curve in water as obtained with high quality DFT is significantly different from the gas phase potential energy calculated with the same method. The PMF in water was qualitatively reproduced by DFTB3 [99] suggesting that DFTB3 performs reasonably in describing Michael additions in condensed phase, a situation relevant to covalent inhibition of enzymes. Nevertheless, low level QM methods have to be validated for the system actually investigated and occasionally corrections have to be applied.

Various correction schemes have been proposed to improve the results of QM/MM MD simulations obtained for enzymatic reactions with low level QM methods. These correction schemes can be straightforwardly applied in covalent inhibition studies. The difference between the high and low level single point energies evaluated with various QM or QM/MM models has been shown to provide sensible corrections to free energy profiles [100–103].

MD based QM/MM methods are effective tools to calculate free energies profiles of covalent inhibitor binding on one hand, and to analyze the interactions and the dynamics in the system, on the other hand. These pieces of information are highly valuable to design inhibitors with improved reactivity profile. We note that the full characterization of inhibitors requires both the rate constant of the covalent step and the dissociation constant of the non-covalent complex formation (Equation 1). The dissociation constant can be obtained from the free energy change accompanying the non-covalent complex formation (Equation 4), and the rate constant is related to the free energy barrier of the chemical reaction of the covalent bond formation (Equation 5). While the latter can be calculated with the quantum chemical methods discussed above, the former can be obtained with MD based free energy perturbation (FEP) or thermodynamic integration calculations with classical force fields.

$$\Delta G = \mathrm{RTln}(K_i) \tag{4}$$

$$\Delta G^{\ddagger} = -\mathrm{RTln}\left(\frac{k_{\mathrm{inact}}}{k_{\mathrm{b}}T/h}\right) \tag{5}$$

In Equations (4) and (5), R is the universal gas constant, T is the absolute temperature, k_{b} is the Boltzmann constant and h is the Planck constant. Both K_i and k_{inact} can be experimentally measured [22, 104] and can be obtained from calculated ΔG and $\Delta G^{\ddagger}$ values, respectively. This makes it possible to compare calculated and experimental values and to validate computational methods as far as corresponding measured values are available.

The separate experimental evaluation of K_i and k_{inact} is not always feasible and then the inhibitory activity of a TCI can be characterized by k_{inact}/K_i [105, 106], or $\log(k_{\mathrm{inact}}/K_i)$ [94] values. Larger k_{inact} means a lower transition barrier, therefore higher reactivity, while lower K_i means stronger non-covalent binding to the enzyme. Compounds with high k_{inact}/K_i ratios are not only show high potency, but are expected to have better selectivity, therefore drug designing efforts are typically made to enhance the k_{inact}/K_i ratio. The simultaneous prediction of k_{inact} and K_i, or alternatively, the k_{inact}/K_i ratio by computational chemistry is an emerging topic [94, 99, 107, 108] and the efficient computation of these quantities can significantly contribute to the success of developing covalent drug candidates.

3 | CASE STUDIES

In the present section computational studies on covalent inhibition are reviewed separately for QM (Table 2) and QM/MM (Table 3) investigations.

3.1 | QM calculations on model systems

Predictions of the reactivity between covalent ligands and their protein targets have been made using model systems both in the calculations and in the experiments, the latter generate data to validate the calculations. Flanagan and coworkers [41] measured the reaction rate between a set of acrylamide derivatives and GSH, and evaluated the half-life of the selected molecules. In the next step, the barriers ($\Delta G^{\ddagger}$) of methyl-thiolate addition on the β-carbons of the examined compounds were evaluated by DFT B3LYP/6-311+G(d,p) calculations using SMD implicit solvent model [48] and scaled zero-point vibrational energy correction [143]. The comparison of the experimental $t_{1/2}$ and calculated $\Delta G^{\ddagger}$ values showed exceptionally good correlation ($R^2 = .915$). *N*-arylacrylamide reactivities were studied experimentally and computationally by Cee [29] and coworkers. The effect of aryl substitution was investigated by comparing experimental $\log(k)$ values measured against GSH with the activation energies and activation free energies against methyl-thiolate as calculated by DFT B3LYP/6-311+G(d,p) with SMD solvation model. Krenske and coworkers examined [109] the substitution effects on the reactivity of α,β-unsaturated ketones. CBS-QB3 used as reference and several commonly applied density functional methods (B3LYP, MPW1PW91, B1B95, PBE1PBE (PBE0), M06-2X, B2PLYP and B2PLYP-D), and also MP2 and SCS-MP2 methods were evaluated. M06-2X, B2PLYP-d and SCS-MP2 were found to perform best by giving results within 1 kcal/mol of the reference. Significant differences reported for the barrier free energies obtained in gas-phase and in CPCM water model [110, 111]. Lonsdale et al. [31] compared various calculated quantities, like reaction barriers, reaction energies, pK_{a} values and LUMO energies with experimental reactivities toward GSH for several cysteine targeting compound series containing acrylamide, vinyl-sulfonamide, chloroacetamide or propargylamide warheads. Fair correlations between GSH reactivities and the above quantities calculated at the M06-2X/6-31+G(d,p)-IEF-PCM level were observed for acrylamides. The correlation between reaction energies and experimental reactivities is particularly notable as this provides a computationally efficient way to predict reactivities. When investigations were extended to other warhead types then reaction energy still proved to be a useful predictor of reactivity, however, LUMO energy showed no correlation with experimental reactivity beyond aromatic compounds. Although experimental ^{13}C chemical shifts of acrylamides were found to well describe reactivities toward GSH, calculated values showed poor agreement with the experimental data. In contrast, calculated and experimental pK_{a} values exhibited good correlation for acrylamides when aromatic and nonaromatic compounds were treated separately. Encouraging correlation was also reported between the calculated pK_{a} and experimental GSH half-life. Matched molecular pair analysis (MMPA) was shown to reproduce reactivity trends upon warhead exchange, and pK_{a} calculation for acrylamides combined with MMPA was proposed to be useful in warhead selection.

The authors of this paper applied similar model calculations on the reaction of potential covalent inhibitors against methyl-thiolate, and the relation between the calculated reaction barriers and experimental GSH half-life was investigated. Reactant, transition and product state geometries were optimized at the B3LYP/6-311+G(d,p) level with PCM solvation model. Good correlation between the GSH half-life and calculated reaction barriers was found for compounds with the same warhead, but the correlation significantly dropped when compounds with different warheads were involved (Figure 2). Similar observation was made when a set of heterocyclic fragment like electrophilic compounds was investigated. Various QM descriptors were calculated and correlated with experimental reactivities toward GSH, but useful trends observed only for the

TABLE 2 QM case studies for covalent inhibition

Nucleophile	Ligands	QM method	Calculated quantity	Comments	Ref.
MeSH	Acrylamides, acrylesters, α-Cl-acetamides	B3LYP/6-311+G(d,p)	ΔG, $\Delta G^{\ddagger}$	Correlation with GSH reactivity	Present paper
MeSH and Cys	α,β-unsaturated carbonyl, carboxyl	B3LYP/6-31G**/TZVP	$\Delta G^{\ddagger}$ $\Delta E^{\ddagger}$, Hammett parameters	Keto-enol tautomerization is not rate-limiting	[23]
MeSH	Acrylamides	B3LYP/6-311+G(d,p)	$\Delta G^{\ddagger}$	Effect of aryl substitution	[29]
-	Heterocyclic electrophiles	B3LYP/6-311++(2d,2p)	QM descriptor	Correlation with GSH reactivity	[30]
MeSH	Several warhead types	B3LYP,M06-2X/6-31G(d,p)	$\Delta G^{\ddagger}$ $\Delta E^{\ddagger}$, HOMO, LUMO, NMR shift, *p*Ka, Hammett parameter	Matched molecular pair analysis	[31]
GSH	α,β-unsaturated carbonyl, carboxyl	B3LYP/6-31G**	Molecular descriptors	log k_{GSH} predictions	[32]
MeSH	Acrylamides, non-acrylamides	B3LYP/6311+G(d,p), 6311+G(3df,3pd)	$\Delta G^{\ddagger}$	Correlation with GSH reactivity	[41]
MeSH	Fragments and drug-like Michael acceptors	B3LYP-D3/6-311+G**	Proton affinity, $\Delta G^{\ddagger}$, ΔG_r	Proton affinity is not correlating with GSH reactivity of drug like molecules	[42]
MeSH	Michael acceptors	Series of DFT functionals, CCSD(T), MP2	ΔE_{net}	Large number of functionals are failing to find correct minima for enolate intermediate	[95]
MeSH	α,β-unsaturated ketones	Several commonly applied density functional, MP2, SCS-MP2	ΔG, $\Delta G^{\ddagger}$, GSH and PhSH addition rate constant	Michael acceptor substituent effect	[109]
-	Neutral small molecules	HF, MP2, B3LYP/6-31G(d), 6-311+G(d,p)	Molecular energies		[110]
-	Small molecules	HF, B3LYP/6-31G(d), 6-311+G(d,p)	Molecular free energies		[111]
-	Serine and cysteine protease inhibitors	PM6	Descriptors derived from ΔH_f	Substituent effects in a QSAR type equation	[112]
β-mercaptoethanol	Acrylnitriles	B3LYP/6-311+G(d,p)	Proton affinity, β-elimination rates		[113]
MeSH	Michael acceptors	B3LYP, M06-2X, ωB97XD/6-311+G(d,P), ωB97XD/aug-cc-pVTZ	Free energy profile	Reversibility calculations	[114]
SARS-CoV-2 M_{pro} inhibition model	α-ketoamide	ONIOM-B3LYP/6-31G(d):PM6, ωB97X-D/6-31G(d,p):PM6	Energy profiles, electron density		[115]
Cysteine protease model	α,β-unsaturated carbonyl compounds	B3LYP/TZVP//BLYP/TZVP	Potential energy surface calculations		[116]
MeSH	Michael acceptors	ωB97XD/aug-cc-pVTZ	ΔG_{int}, ΔG_{prod}, ΔG_{total}	Automated screening of thiol reactivity	[117]
MeSH	Acrylamides, chloroacetamides	ωB97XD/cc-pVDZ	$\Delta G^{\ddagger}$	Machine learning algorithm for the prediction of thiol reactivity	[118]
-	Acrylamides	B3LYP/6-31+G(d,p)	Eletrophilicity index, HOMO, LUMO		[119]
Cysteamine	Nitril based compounds	B3LYP/6-311++G(d,p)	$\Delta G^{\ddagger}$	Strong correlation with GSH, cysteine and cysteamine reactivity	[120]
MeSH	Cyanamides, aryl nitriles, aminoacetonitriles	B3LYP/6-311G(d,p)	ΔG_r, electrophilicity	Reaction energy versus cysteine adduct formation	[121]
MeSH	Aryl nitriles	B3LYP/6-311G	ΔH_r	Structure activity relationships	[122]

Abbreviation: QM, quantum mechanical.

electronic chemical potential and the natural charge on the reacting carbon atom within subsets of compounds reacting with the same mechanism [30].

Descriptors derived from semiempirical PM6 heat of formation of the first intermediate of the covalent bond formation were proposed [112] for serine and cysteine protease inhibitors. Substituent effects were described in a QSAR type equation and the different relationship observed for serine and cysteine proteases were explained.

It is worth noting that not all cysteine reactivity studies contain methyl-thiolate as a probe, and GSH is not used exclusively in reactivity assays [51]. Krishnan and coworkers [113] used β-mercaptoethanol as a model for the cysteine sidechain and calculated the proton affinity of a set of acrylonitrile-mercaptoethanol adducts at the B3LYP/6-311+G (d) level of theory with PCM water model. Experimental β-elimination rates exhibited excellent correlation with calculated proton affinities, and this was suggested to enabling the tuning of the reversibility of the reaction in a predictable manner. The effect of α-substitution on a set of acrylonitrile-based Michael acceptors was investigated [114] by calculating the free energy profile in model reactions with various DFT methods including B3LYP/6-31+G(d), M06-2X/6-31+G(d) and ωB97X-D5/aug-cc-pVTZ using CPCM continuum solvent model. Electron withdrawing groups were shown to affect reversibility by destabilizing the neutral adduct, on one hand, and stabilizing the anionic TS and the intermediate of the Michael addition, on the other hand.

Inhibition of cysteine proteases by α-ketoamide compounds was investigated [115] by reaction mechanism calculations on model systems using ONIOM-B3LYP/6-31G(d):PM6 geometry optimizations and single point ωB97X-D/6-31G(d,p):PM6 energy evaluations. It was found that the deprotonation of the cysteine thiol and the nucleophilic attack of sulfur onto the α-ketoamide warhead can proceed both with stepwise and concerted mechanisms. Paasche and coworkers [116] examined the reaction of α,β-unsaturated carbonyl compound with cysteine protease models. B3LYP/TZVP//BLYP/TZVP calculations with the COSMO continuum model were performed. Results emphasize the importance of base catalysis for thiol reactivity and computed data explained substituent effect on reactivity.

Smith [117] and coworkers developed an automated computational screening method of the thiol reactivity of substituted alkenes. The proposed workflow contains the generation, modeling, and analysis of the reactants, intermediate and products of the thiol addition using ωB97XD/aug-cc-pVTZ level of theory. They performed multiple linear regression for the calculated intermediate and product stability based on the reactant properties. Recently, Palazzesi and coworkers [118] implemented a machine learning algorithm for the prediction of the thiol reactivity. A large amount of barrier calculations (ωB97XD/cc-pVDZ) against methyl-thiolate was carried out using a warhead aligning scheme to lower the computational time. The calculated barriers were used to train the machine learning model. The ab initio and model calculated barriers showed good correlations with an R^2 of .85 for acrylamides and .69 for chloroacetamides.

The electrophilicity index based on HOMO and LUMO energies was calculated [119] at the B3LYP/6-31+G(d,p) level with CPCM water model for acrylamides and was correlated with experimental reactivity toward GSH. It was found that the reactivity of small molecules (MW < 250 Da) can be directly estimated by calculated electrophilicities, while for larger molecules a truncation algorithm is introduced to ensure orbital localization on the warhead.

Berteotti [120] calculated activation energies of the reactions between nitril based compounds and zwitterionic cysteamine. Strong correlation was found between B3LYP/6-311++G(d,p) activation energies obtained with CPCM water model and experimental half-life in reactions against cysteine, cysteamine and GSH. A reactivity index of nitril-containing compounds was calculated [121] by evaluating the energy difference between the thioimidate adduct and the precursor methanethiol and the nitrile molecule at the B3LYP/6-311G(d,p) level with PCM water model. The calculated index reproduced the trend of cysteine adduct formation. Similarly, the reaction enthalpy of aryl nitrile compounds with methanethiol was calculated [122] and enthalpies obtained were correlated with inhibition constants measured against two cysteine proteases, rhodesain and human cathepsin L, and also with cytotoxicity, and clear structure activity relationships were reported.

Voice [42] and coworkers investigated how compound size affects the correlation between GSH reactivity and either proton affinity or reaction energy of α,β-unsaturated carbonyl compounds. The latter quantities were calculated with the density functional B3LYP-D3/6-311+G**//B3LYP-D3/6-31G* method using the Poisson−Boltzmann Finite Element implicit water solvation model. It was found that the strong correlation between the calculated quantities and GSH reactivity observed for small compounds seriously deteriorates for drug-like compounds. The authors suggest that the inadequate description of available conformations and their variation during the reaction contribute to the failure of ligand-only approaches for larger molecules.

3.2 | QM/MM calculations

Quantum chemical calculations for model systems are able to provide useful estimations of various molecular properties including reactivity. However, they cannot account for several important aspects of the reactions in extended systems. A more complete description of covalent inhibition of proteins requires the inclusion of the protein environment in the computational studies. Multiscale methodologies allow the modeling of chemical reactions in the active site at reasonable computational cost. In particular, QM/MM schemes calculate the electron reorganization at QM level, while the effect of the protein environment is taken into account by molecular mechanics. There is a diverse set of QM/MM approaches applied in covalent inhibitor studies. In the following section, we first discuss applications that produce energies without generating

TABLE 3 QM/MM case studies for covalent inhibition

Biomolecule	Ligand(s)	QM method	MM method	Calculation type	Comments	Ref.
MAO-B	Serotonin	M06-2x:PM6		Structure optimizations	ONIOM (QM:QM)	[70]
Cysteine protease	Epoxides	RI-BLYP	CHARMM	PES with HDLC optimizer	Stereoselectivity study	[71]
Rhodesain	Vinyl-sulfones	BLYP, B3LYP	AMBER	Model with methylthiol, PES calculation, Docking, MD	Four steps workflow	[72]
HIV1 protease	Aziridine, epoxide	BLYP	CHARMM	Aiding covalent docking	Workflow for docking	[73]
Human cyclooxigenase-1	Aspirin	HF, B3LYP, MP2, b97D	AMBER	PES calculation.	ONIOM	[74]
Rhodesain	Nitriles	PM6	CHARMM	PMF US	Correlation between reaction free energy and experimental binding affinity	[92]
EGFR	Acrylamides	DFTB3	AMBER	US	Path collective variables were applied	[93]
Multiple targets	Multiple set of compounds	B3LYP-D3	OPLS3	FEP+, FEP	Predicting log (k_{inact}/K_i)	[94]
Multiple targets	Multiple set of compounds	DFTB3	AMBER	US, TI	Complete description of the covalent inhibition	[99]
BTK	Cyanoacrylamide	ωB97XD	CHARMM	FEP, US, Opt	Complete free energy profile, ONIOM	[107]
SARS-CoV-2 Mpro	α-ketoamide	M06-2X	AMBER	FEP, US, Opt.	Complete free energy profile, ONIOM	[108]
TAK1 kinase	Acrylesters	M06-2X	AMBER	MD Opt.	ONIOM	[45]
Cysteine protease	Epoxides and aziridines	RI-BLYP	CHARMM	PES calculation.	Regioselectivity study	[123]
FAAH	Carbamates	SCC-DFTB	CHARMM	PES calculation	QM correction with B3LYP/6-31+G(d)	[124]
Cysteine protease	Nitriles	B3LYP	AMBER	Geometry optimizations	ONIOM	[125]
Multiple targets	Multiple set of compounds	SCC-DFTB	CHARMM	Docking, QM/MM minimization	New scoring function, attracting cavities	[126]
Schistosoma mansoni cathepsin B1	Vinyl-sulfones	DFT-S/PM6-D3HX4	AMBER	Restrained optimizations	Quantum mechanics based scoring function	[127]
EGFR (L718Q mutant)	Osimertinib	SCC-DFTB	AMBER	US, RETI, Waterswap		[128]
MurA	Multiple set of compounds	DFTB3	AMBER	1D and 2D US, MD	QM correction with ωB97XD	[100]
Aspartic protease	Epoxide	SCC-DFTB	AMBER	2D US	QM correction with M06-2X	[129]
Glycoside hydrolase	Carbasugars	AM1	OPLS-AA	2D PES calculation, 1D PMF calculation, Electrostatic potential mapping	QM Correction with M06-2X	[130]
SARS-CoV-2 Mpro	N3 peptidyl Michael-acceptors	AM1	AMBER	US	QM correction with M06-2X	[131]
Cruzain	Peptidyl halomethyl-ketones	AM1d	OPLS-AA	PES calculation, US	QM correction with M06-2X	[132]
Cruzain	Vinyl sulfones, nitriles	DFTB3	AMBER	2D-FES US, adiabatic mapping	QM correction with MP2	[133]
Cruzain	Dipeptidyl nitriles	AM1/d-Phot	CHARMM	2D US	QM correction with M06-2X	[134]
β-lactamase	Carbapenem	SCC-DFTB	AMBER	2D US, MEP analysis		[135]
β-lactamase	Clavulanate	SCC-DFTB	AMBER	2D PMF US, MEP		[136]

TABLE 3 (Continued)

Biomolecule	Ligand(s)	QM method	MM method	Calculation type	Comments	Ref.
Human tissue transglutaminase (TG2)	Acrylamides	PM3, SCC-DFTB	AMBER	US, charge analysis		[137]
AChE	Soman	B3LYP	AMBER	Ab-initio Born Oppenheimer US (frozen QM region)	Reaction coordinate driving method	[138]
GH43 xylanolytic enzyme	Own substrate	PBE and Troullier-Martins ab initio pseudopotential	AMBER	Metadynamics		[139]
WT and mouse AChE	Acetophenones	B3LYP	AMBER	Scan, FEP (frozen QM), TI	Pseudobond ab initio QM/MM approach	[140]
Proteasome	Epoxomycin	B3LYP	AMBER	MEP calculation, FEP		[141]
Proteasome	Syringolin A	B3LYP	AMBER	MEP calculation, FEP		[142]

Abbreviations: MM, molecular mechanical; QM, quantum mechanical.

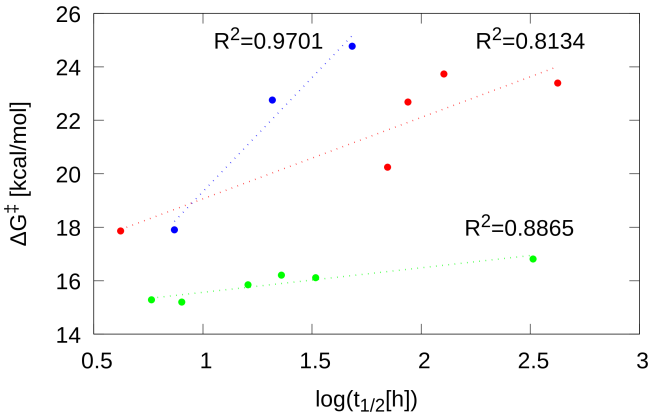


FIGURE 2 Experimental GSH half-life [36] versus calculated activation free energies against methyl-thiolate for compound sets with acrylamide (red), acrylester (blue) and α-Cl-acetamide (green) warheads. For further information about data points please refer to Table S1

conformational ensembles. Next, we focused to applications producing free energies by MD based sampling. Finally, studies are presented with the combined evaluations of the binding free energy of the non-covalent binding and the reaction barrier of the covalent complex formation.

3.2.1 | Potential energy based studies

PES based studies attempts to evaluate reaction mechanisms in a basically static protein environment, although the energy calculations are occasionally complemented with thermal corrections to generate Gibbs free energies. ONIOM [64] is often applied for QM/MM energy calculations as it allows the characterization of the stationary points, a feature not widely available in QM/MM codes. Tóth and coworkers examined [74] the mechanism of the irreversible inhibition of human cyclooxygenase-1 (COX-1) by the well-known drug aspirin (ASA). They performed a series of ONIOM-type QM/MM calculations at HF, B3LYP, MP2 and B97-D level of theories while treating the MM part with the AMBER force field. All methods applied predict the same one step reaction mechanism for the ASA-COX-1 trans-esterification reaction but with significant differences in the activation energies. The mechanism of catalysis and covalent inhibition of monoamine oxidase B (MAO-B) was investigated [70] by ONIOM (QM:QM) calculations at the M06-2X/6-31+G(d,p):PM6 and PM6 levels. A new catalytic pathway and a FAD binding mechanism of reversible covalent inhibitors was proposed for both MAO-A and MAO-B isozymes. The full free energy profile was calculated for the binding of two SARS-CoV-2 main protease (M^{pro}) inhibitors. The covalent reaction was evaluated by ONIOM at M06-2X/def2-TZVP:FF99SB-ILDNP level of theory with thermal corrections [108]. Significant differences were found between the two inhibitors both in their non-covalent and covalent binding free energies. Toviwek and co-workers investigated [45] the covalent inhibition of TAK1 kinase. Michael addition mechanism to a reactive

cysteine was studied at QM/MM ONIOM(M06-2X/6-31G(d):FF99SB) level with vibrational correction. The nucleophilic attack by the thiolate on the Michael acceptor was found to be the rate determining step that is preceded by a lower barrier cysteine deprotonation. Mladenovic and coworkers evaluated [71, 123] the stereoselectivity and regiospecificity of epoxide based cysteine protease inhibitors by applying the BLYP/TZVP//CHARMM methodology for the potential energy scan. Their results show that the different hydrogen-bonding network of the stereoisomers is responsible for the observed selectivity differences. Fatty acid amide hydrolase (FAAH) inhibition mechanism by two covalent inhibitors was studied [124] by SCC-DTFB/CHARMM27 potential with B3LYP/6-31+G(d) corrections. The inhibitors react with Ser241 and form carbamoylated intermediates. Calculated relative reaction barriers of the decarbamoylation of the inhibitor complexes are in line with the observed reversibility of one of them and the irreversibility of the other. Moreover, both barriers exceed that of the substrate bound deacylation owing to an impaired stabilization of the TS in the presence of inhibitors. Quesne et al. investigated [125] the mechanism of cysteine protease inhibition by nitril-based compounds both by DFT model calculations and by QM/MM. Concerted nucleophilic addition by thiolate, and proton transfer from histidine was found by all methods applied. DFT model calculations using B3LYP with 6-31G* and 311++G** basis sets with and without continuum solvent model led to important geometry distortions incompatible with the enzymatic environment. QM/MM studies with mechanical versus electrostatic embedding resulted in large geometric and thermodynamic differences. The authors emphasized the importance of the model in investigating this biochemical reaction mechanism. Chaskar and coworkers proposed [126] a docking protocol using their attractive cavities algorithm [144] and the SCC-DFTB/CHARMM optimization of the docking poses. The docking protocol is designed for non-covalent complexes, but it is claimed to well describe ligand-iron binding with covalent character. Fanfrlík and coworkers proposed [127] the extension of their formerly introduced QM-based scoring function with QM/SQM restrained optimizations to treat covalent inhibitor binding in docking and scoring. A multiple step design protocol of covalent inhibitors was proposed, which includes QM and QM/MM optimizations, docking, and MD simulations to evaluate reaction barriers and reaction energies of designed compounds [72, 73].

3.2.2 | QM/MM coupled with MD

Callegari [128] and coworkers investigated the deprotonation mechanism of Cys797 of the L718Q mutant EGFR tyrosine kinase, and the alkylation step by osimertinib by constructing PMF using US simulations at the SCC-DFTB/FF99SB level of theory. Their study proved that the mutation of L718Q does not affect the affinity of osimertinib toward EGFR. Capoferri [93] and coworkers simulated the inhibition of EGFR by an acrylamide derivative at the DFTB3/FF99SB level using path collective variables [145, 146]. The simulation protocol consisted of a series of iterative reaction path optimizations by SMD. After the path was converged US was performed starting from the structures extracted from the converged SMD trajectory. The results of the calculations identified the desolvation of Cys797 as a key step during the inhibition process.

Studies [100] on MurA covalent inhibitors using simulations at the DFTB3/FF14SB level suggested a deprotonation mechanism for the targeted Cys115 by His394, and produced reaction free energy profiles for nine potential covalent inhibitors. Active and inactive compounds were successfully distinguished based on the activation free energy of the covalent reaction. A QM correction at ωB97XD/cc-aug-pVTZ level was necessary to improve DFTB3 based simulation results.

Inhibition of aspartic proteases by epoxide-based irreversible covalent inhibitors was proposed to proceed with a two-step mechanism, where the oxyanion intermediate and the corresponding TS are stabilized by a group of water molecules. HIV protease and pepsin were studied with DFTB/FF14SB simulations corrected by M06-2X/631G(d,p) single point calculations [129].

The covalent labeling of glycoside hydrolase by carbasugar inhibitors was investigated [130] by experimental and theoretical methods including the calculation of the free energy profile of the reaction by AM1/OPLS-AA potential with M06-2X/6-31+G(d,p) correction. S_N2 TS was predicted, and a new direction of inhibitor design is proposed based on the location of charge development of the TS structure.

Most recently, Arafet [131] and coworkers constructed the PMF of the inhibition reaction of SARS-CoV-2 M^{pro} by a peptidyl Michael-acceptor. The AM1/FF03 level free energies were corrected by M06-2X/6-31+G(d,p) level optimizations and energy calculations. QM/MM modeling was also used to propose new compounds with potential inhibitory effect via covalent mechanism. Free energy surface by PM6/CHARMM potential was generated [92] for a set of heteroaryl nitrile reversible covalent inhibitors of rhodesain, a cysteine protease. Strong correlation between calculated reaction free energies and experimental binding affinity was reported. The covalent inhibition of cruzain, a cysteine protease by peptidyl halomethyl ketones was investigated [132] by PMF calculations with AM1d/OPLS-AA potential corrected by M06-2X functional with 6-31+G(d,p) basis set. Concerted attack of the deprotonated cysteine and the leaving of the halogen ion was observed. Both reaction barriers and reaction energies favor inhibition by peptidyl chloro-methyl ketones over fluoro analogs. Another QM/MM study of covalent cruzain inhibitors [133] using DFTB3/FF14SB potential with adiabatic mapping MP2/FF14SB corrections found that the nucleophilic attack by cysteine and the proton transfer from histidine occur in a single step for a reversible inhibitor and in consecutive steps for an irreversible inhibitor. Moreover, calculated reaction free energy was significantly more exothermic for an irreversible inhibitor than for a reversible inhibitor acting with the same covalent mechanism. Nitrile-based covalent cruzain inhibitors were also investigated [134]

with AM1/d-PhoT Hamiltonian combined with CHARMM force field to produce reaction free energy profiles. The mechanism of reversible inhibitors included a single step of cysteine attack and proton transfer from histidine in accordance with the previously cited study [133]. Calculated reaction free energy was found to be in agreement with experimental binding affinities.

DFTB/FF12SB simulations [135] of eight class A β-lactamases, although underestimated the free activation energy, were shown to well separate enzymes with carbapenem deacylation activity from enzymes lacking this activity. Based on this finding, the simulations can serve as an in silico assay for carbapenemase activity in class A β-lactamases. The covalent inhibition of KPC-2 and TEM-1 class A β-lactamases by clavulanate was investigated [136] at the DFTB/FF14SB level simulations. The decarboxylated trans-enamine of the deacylation reaction was shown to be responsible for the inhibition and the relative magnitude of calculated free energy barriers reflected the experimental inhibitor potencies.

Acrylamide based inhibitors against human tissue transglutaminase (TG2) were studied [137] by semiempirical QM/MM simulations. DFTB and PM3 was used as QM methods combined with FF99SB MM potential. US simulations suggested two step Michael addition reaction of acrylamides against TG2, and calculated activation energies exhibited good correlation with a wide range of experimental IC_{50} values.

The application of semiempirical QM methods during the MD simulations significantly reduces the computational time at the expense of reduced reliability and consistency of the results. The application of more involved methods to generate the PMF of the inhibitory reaction was also reported. Sirin and coworkers applied [138] the reaction coordinate driving method to obtain the minimum energy path of the acetylcholine esterase (AChE) inhibition by the nerve toxin soman. Once a feasible reaction path was obtained, they performed MM relaxation, while keeping the QM subsystem frozen. The relaxed structures were subject to ab-initio Born–Oppenheimer QM/MM US applying B3LYP/6-31G* QM and FF99SB MM potential. The resulted free energy profile indicates an addition-elimination inhibitory mechanism in which the His440 initiates the phosphorylation reaction by deprotonating the Ser200 residue. Recently, the full free energy profile of the covalent modification of Bruton's tyrosine kinase (BTK) by a cyanoacrylamide TCI was reported [107]. US simulation at the ωB97X-D/def2- level was applied for the nucleophilic attack of the first addition reaction between Cys481 of BTK and the cyanoacrylamide compound. Finally, ONIOM-based geometry optimization for the thioether product yielded by proton transfer were applied to recover the reaction energy of the covalent step [107]. These high level calculations are expected to yield accurate free-energy profiles, although the increased computational demand currently hinders their comparative application to several ligands.

Two distinct catalytic pathways for GH43 xylanolytic enzymes were identified [139] by Morais and coworkers using experimental data and metadynamics. They used Car-Parrinello MD with PBE functional using planewave basis and Troullier−Martins ab initio pseudopotentials coupled with FF14SB/GLYCAM06 MM potential to construct the reaction PMF. The complete hydrolysis mechanism of the GH43 enzymes was revealed and was proposed to expand the general model of catalytic mechanism of glycosidases and to facilitate the design of TS-like inhibitors.

In the work of Cheng and coworkers, the inhibition of wild type and mutant mouse AChE by trifluoroacetophenone derivatives was modeled [140] by first constructing the PES of the reaction. For each stationary point, the system was QM/MM minimized and QM subsystem fluctuations to free energy change were evaluated from the frequencies using a harmonic approximation. QM/MM interaction contributions were determined from FEP calculations along the reaction coordinate using B3LYP/6-31G*/AMBER potential. Calculations revealed barrierless nucleophilic addition of serine to the carbonyl-C of the ligand with a simultaneous serine-histidine proton transfer. Wei and coworkers also used FEP calculations to determine [141, 142] the relative free energy changes associated with the QM/MM interactions. Similar technics were applied to study the covalent inhibition mechanisms of immunoproteasome by epoxomicin and syringolin A. The details of the inhibitor binding mechanisms were explored and the rate determining steps were identified. They constructed the minimum potential and free energy curves of the inhibition reaction.

3.2.3 | Complete description of covalent inhibitor binding

As it is discussed above and shown by Equations (1), (4) and (5), the two essential parameters of the covalent inhibition are the binding constant (K_i) and the kinetic rate constant (k_{inact}). The majority of the above studies focused on the covalent reaction. However, a complete computational description of irreversible inhibition requires the characterization of both binding steps. This necessitates the use of multiple computational techniques and multiple sampling methods. A possible scheme for computing both K_i and k_{inact} includes the treatment of the non-covalent binding by thermodynamic integration using MM force field, while the construction of the PMF of the covalent reaction by QM/MM US simulations. We applied this strategy for KRAS, EGFR and Tec family tyrosine kinase (ITK, BTK, BMX) inhibitors using DFTB3/FF14SB potential, and found good correlation between experimental and calculated $\log(k_{inact}/K_i)$ values [99]. Yu and coworkers [94] developed a method in which the covalent reaction was calculated at QM (B3LYP-D3/6-311+G*) level for a model system, and the effect of the enzyme environment on the TS was evaluated by FEP calculations using modified force-field parameters. This DFT + FEP method was able to produce $\log(k_{inact}/K_i)$ giving fair correlation with the experimental values for BTK, FAAH and KRAS inhibitors [94]. The full free energy profile was calculated for the binding of two SARS-CoV-2 inhibitors [108] and for the Michael addition of a BTK inhibitor [107]. The absolute binding free energy of the non-covalent step was evaluated by FEP. The free energy profile of the chemical reaction was obtained by ONIOM at

the M06-2X/def2-TZVP:FF99SB-ILDNP level of theory for the SARS-CoV-2 inhibitors, while for the BTK inhibitors [107] US simulation at the ωB97X-D/def2-TZVP level was applied to obtain the barrier of the nucleophilic attack, and ONIOM-based geometry optimization of the product state was applied to obtain the reaction energy.

4 | CONCLUSIONS, FUTURE ASPECTS

Recently, covalent inhibition has become a hot topic in pharmaceutical research and the computational support to develop better TCIs requires the intensive application of quantum chemical tools. Calculations of reactivity descriptors and evaluations of reaction barriers in model systems are useful in the screening and in the optimization of covalent ligands. More involved approaches include multiscale calculations for the solvated ligand-protein complex to predict the energy surfaces of reactions. When combined with MD based sampling, then free energy surfaces can be obtained that give information on the mechanistic details of the inhibition reaction and on the rate determining free energy barrier. These calculations are computationally intensive owing to the sampling in extended systems and to the involved QM methods needed for reliable energies and geometries. Further performance improvements are expected with the combined development of computer hardware and QM methodology. Nevertheless, the reviewed applications of covalent inhibitor design well demonstrate the usefulness and perspectives of quantum chemical calculations in the emerging field of covalent drug discovery.

AUTHOR CONTRIBUTIONS

Levente Mihalovits: Investigation; writing-original draft; writing-review & editing. **György Ferenczy:** Investigation; supervision; writing-review & editing. **György Keserű:** Conceptualization; resources; writing-review & editing.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Levente M. Mihalovits https://orcid.org/0000-0003-1022-3294
György G. Ferenczy https://orcid.org/0000-0002-5771-4616
György M. Keserű https://orcid.org/0000-0003-1039-7809

REFERENCES

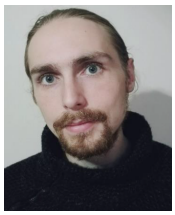
[1] S. Planken, D. C. Behenna, S. K. Nair, T. O. Johnson, A. Nagata, C. Almaden, S. Bailey, T. E. Ballard, L. Bernier, H. Cheng, S. Cho-Schultz, D. Dalvie, J. G. Deal, D. M. Dinh, M. P. Edwards, R. A. Ferre, K. S. Gajiwala, M. Hemkens, R. S. Kania, J. C. Kath, J. Matthews, B. W. Murray, S. Niessen, S. T. M. Orr, M. Pairish, N. W. Sach, H. Shen, M. Shi, J. Solowiej, K. Tran, E. Tseng, P. Vicini, Y. Wang, S. L. Weinrich, R. Zhou, M. Zientek, L. Liu, Y. Luo, S. Xin, C. Zhang, J. Lafontaine, *J. Med. Chem* **2017**, *60*, 3002.
[2] S. C. M. Lau, U. Batra, T. S. K. Mok, H. H. Loong, *Drugs* **2019**, *79*, 823.
[3] H. W. B. Johnson, E. Lowe, J. L. Anderl, A. Fan, T. Muchamuel, S. Bowers, D. C. Moebius, C. Kirk, D. L. McMinn, *J. Med. Chem.* **2018**, *61*, 11127.
[4] E. M. Huber, M. Groll, *Angew. Chem. Int. Ed.* **2012**, *51*, 8708.
[5] V. Oldfield, G. M. Keating, C. M. Perry, *Drugs* **2007**, *67*, 1725.
[6] P. Savi, P. Nurden, A. T. Nurden, S. Levy-Toledano, J.-M. Herbert, *Platelets* **1998**, *9*, 251.
[7] A. A. Adhikari, T. C. M. Seegar, S. B. Ficarro, M. D. McCurry, D. Ramachandran, L. Yao, S. N. Chaudhari, S. Ndousse-Fetter, A. S. Banks, J. A. Marto, S. C. Blacklow, A. S. Devlin, *Nat. Chem. Biol.* **2020**, *16*, 318.
[8] S. H. Olesen, D. J. Ingles, Y. Yang, E. Schönbrunn, *J. Basic Microbiol.* **2014**, *54*, 322.
[9] F. Sutanto, M. Konstantinidou, A. Dömling, *RSC Med. Chem* **2020**, *11*, 876.
[10] D. A. Shannon, E. Weerapana, *Curr. Opin. Chem. Biol.* **2015**, *24*, 18.
[11] T. A. Baillie, *Angew. Chem. Int. Ed.* **2016**, *55*, 13408.
[12] T. Zhang, J. M. Hatcher, M. Teng, N. S. Gray, M. Kostic, *Cell Chem. Biologia* **2019**, *26*, 1486.
[13] A. Miseta, P. Csutora, *Mol. Biol. Evol.* **2000**, *17*, 1232.
[14] K. Hallenbeck, D. Turner, A. Renslo, M. Arkin, *Curr. Top. Med. Chem.* **2016**, *17*, 4.
[15] Z. Zhao, P. E. Bourne, *Drug Discov. Today* **2018**, *23*, 727.
[16] H. Willems, S. De Cesco, F. Svensson, *J. Med. Chem.* **2020**, *63*, 10158.
[17] C. N. Cavasotto, M. G. Aucar, N. S. Adler, *Int. J. Quantum Chem.* **2019**, *119*, e25678.
[18] M. Garofalo, G. Grazioso, A. Cavalli, J. Sgrignani, *Molecules* **2020**, *25*, 1756.
[19] G. W. Caldwell, Z. Yan, W. Lang, J. A. Masucci, *Curr. Top. Med. Chem.* **2012**, *12*, 1282.
[20] C. Yung-Chi, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099.
[21] R. Lonsdale, R. A. Ward, *Chem. Soc. Rev.* **2018**, *47*, 3816.
[22] B.-F. Krippendorff, R. Neuhaus, P. Lienau, A. Reichel, W. Huisinga, *J. Biomol. Screen.* **2009**, *14*, 913.
[23] J. A. H. Schwöbel, J. C. Madden, M. T. D. Cronin, *SAR QSAR Environ. Res.* **2010**, *21*, 693.
[24] D. J. Ebbrell, J. C. Madden, M. T. D. Cronin, T. W. Schultz, S. J. Enoch, *Chem. Res. Toxicol.* **2016**, *29*, 1073.

[25] R. E. Amaro, A. J. Mulholland, *Nat. Rev. Chem.* **2018**, *2*, 1.
[26] M. G. Quesne, T. Borowski, S. P. de Visser, *Chem. A Eur. J.* **2016**, *22*, 2562.
[27] D. Bajusz, A. Rácz, K. Héberger, *Comprehensive Medicinal Chemistry III*, Vol. 3–8 Elsevier, Oxford **2017**, p. 329.
[28] J. Fei, Q. Mao, L. Peng, T. Ye, Y. Yang, S. Luo, *Molecules* **2018**, *23*, 2935.
[29] V. J. Cee, L. P. Volak, Y. Chen, M. D. Bartberger, C. Tegley, T. Arvedson, J. McCarter, A. S. Tasker, C. Fotsch, *J. Med. Chem.* **2015**, *58*, 9171.
[30] A. Keeley, P. Ábrányi-Balogh, G. M. Keserű, *MedChemComm* **2019**, *10*, 263.
[31] R. Lonsdale, J. Burgess, N. Colclough, N. L. Davies, E. M. Lenz, A. L. Orton, R. A. Ward, *J. Chem. Inf. Model.* **2017**, *57*, 3124.
[32] J. A. H. Schwöbel, D. Wondrousch, Y. K. Koleva, J. C. Madden, M. T. D. Cronin, G. Schüürmann, *Chem. Res. Toxicol.* **2010**, *23*, 1576.
[33] S. J. Enoch, *Challenges and Advances in Computational Chemistry and Physics*, Vol. 8 Springer, New York **2010**, p. 13.
[34] M. Larsson, B. Kumar Mishra, M. Tysklind, A. Linusson, P. L. Andersson, *SAR QSAR Environ. Res.* **2013**, *24*, 461.
[35] T. Puzyn, N. Suzuki, M. Haranczyk, J. Rak, *J. Chem. Inf. Model.* **2008**, *48*, 1174.
[36] A. Scarpino, L. Petri, D. Knez, T. Imre, P. Ábrányi-Balogh, G. G. Ferenczy, S. Gobec, G. M. Keserű, *J. Comput. Aided. Mol. Des.* **2021**, *35*, 223.
[37] P. N. Plessow, *J. Chem. Theory Comput.* **2018**, *14*, 981.
[38] T. A. Halgren, W. N. Lipscomb, *Chem. Phys. Lett.* **1977**, *49*, 225.
[39] T. Tsutsumi, Y. Ono, Z. Arai, T. Taketsugu, *J. Chem. Theory Comput.* **2018**, *14*, 4263.
[40] S. Maeda, Y. Harabuchi, Y. Ono, T. Taketsugu, K. Morokuma, *Int. J. Quantum Chem.* **2015**, *115*, 258.
[41] M. E. Flanagan, J. A. Abramite, D. P. Anderson, A. Aulabaugh, U. P. Dahal, A. M. Gilbert, C. Li, J. Montgomery, S. R. Oppenheimer, T. Ryder, B. P. Schuff, D. P. Uccello, G. S. Walker, Y. Wu, M. F. Brown, J. M. Chen, M. M. Hayward, M. C. Noe, R. S. Obach, L. Philippe, V. Shanmugasundaram, M. J. Shapiro, J. Starr, J. Stroh, Y. Che, *J. Med. Chem.* **2014**, *57*, 10072.
[42] A. Voice, G. Tresadern, H. van Vlijmen, A. Mulholland, *J. Chem. Inf. Model.* **2019**, *59*, 4220.
[43] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
[44] K. Czaja, J. Kujawski, R. Kujawski, M. K. Bernard, *Open Chem.* **2020**, *18*, 857.
[45] B. Toviwek, D. Gleeson, M. P. Gleeson, *Org. Biomol. Chem.* **2021**, *19*, 1412.
[46] J. M. L. Martin, *J. Chem. Phys.* **1994**, *100*, 8186.
[47] B. Mennucci, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 386.
[48] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378.
[49] H. Helten, T. Schirmeister, B. Engels, *J. Org. Chem.* **2005**, *70*, 233.
[50] P. Ábrányi-Balogh, L. Petri, T. Imre, P. Szijj, A. Scarpino, M. Hrast, A. Mitrović, U. P. Fonovič, K. Németh, H. Barreteau, D. I. Roper, K. Horváti, G. G. Ferenczy, J. Kos, J. Ilaš, S. Gobec, G. M. Keserű, *Eur. J. Med. Chem.* **2018**, *160*, 94.
[51] A. Keeley, L. Petri, P. Ábrányi-Balogh, G. M. Keserű, *Drug Discov. Today* **2020**, *25*, 983.
[52] M. W. Van Der Kamp, A. J. Mulholland, *Biochemistry* **2013**, *52*, 2708.
[53] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
[54] L. Cao, U. Ryde, *Front. Chem.* **2018**, *6*, 89.
[55] S. Roßbach, C. Ochsenfeld, *J. Chem. Theory Comput.* **2017**, *13*, 1102.
[56] U. Ryde, *J. Chem. Theory Comput.* **2017**, *13*, 5745.
[57] L. Hu, P. Söderhjelm, U. Ryde, *J. Chem. Theory Comput.* **2011**, *7*, 761.
[58] M. J. Field, P. A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700.
[59] V. Théry, D. Rinaldi, J. -L. Rivail, B. Maigret, G. G. Ferenczy, *J. Comput. Chem.* **1994**, *15*, 269.
[60] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227.
[61] J. Pu, J. Gao, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 632.
[62] I. Antes, W. Thiel, *J. Phys. Chem. A* **1999**, *103*, 9290.
[63] Y. Zhang, *Theor. Chem. Acc.* **2006**, *116*, 43.
[64] L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding, K. Morokuma, *Chem. Rev.* **2015**, *115*, 5678.
[65] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696.
[66] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, R. A. Friesner, *J. Chem. Theory Comput.* **2016**, *12*, 281.
[67] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell, *J. Comput. Chem.* **2010**, *31*, 671.
[68] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656.
[69] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622.
[70] K. Cakir, S. S. Erdem, V. E. Atalay, *Org. Biomol. Chem.* **2016**, *14*, 9239.
[71] M. Mladenovic, K. Ansorg, R. F. Fink, W. Thiel, T. Schirmeister, B. Engels, *J. Phys. Chem. B* **2008**, *112*, 11798.
[72] T. Schirmeister, J. Kesselring, S. Jung, T. H. Schneider, A. Weickert, J. Becker, W. Lee, D. Bamberger, P. R. Wich, U. Distler, S. Tenzer, P. Johé, U. A. Hellmich, B. Engels, *J. Am. Chem. Soc.* **2016**, *138*, 8332.
[73] T. C. Schmidt, A. Welker, M. Rieger, P. K. Sahu, C. A. Sotriffer, T. Schirmeister, B. Engels, *ChemPhysChem* **2014**, *15*, 3226.
[74] L. Tóth, L. Muszbek, I. Komáromi, *J. Mol. Graphics Modell.* **2013**, *40*, 99.
[75] A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, M. H. M. Olsson, *Chem. Rev.* **2006**, *106*, 3210.
[76] S. A. Hollingsworth, R. O. Dror, *Neuron* **2018**, *99*, 1129.
[77] M. De Vivo, M. Masetti, G. Bottegoni, A. Cavalli, *J. Med. Chem.* **2016**, *59*, 4035.
[78] M. S. Bodnarchuk, M. J. Packer, A. Haywood, *ACS Med. Chem. Lett.* **2020**, *11*, 77.
[79] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, K. Schulten, 2011, pp. 39–65.
[80] P. C. Do, E. H. Lee, L. Le, *J. Chem. Inf. Model.* **2018**, *58*, 1473.
[81] C. Jarzynski, *Phys. Rev. Lett.* **1997**, *78*, 2690.
[82] H. Xiong, A. Crespo, M. Marti, D. Estrin, A. E. Roitberg, *Theor. Chem. Acc.* **2006**, *116*, 338.

[83] G. M. Torrie, J. P. Valleau, *J. Comput. Phys.* **1977**, *23*, 187.
[84] W. You, Z. Tang, C. E. A. Chang, *J. Chem. Theory Comput.* **2019**, *15*, 2433.
[85] J. Kästner, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932.
[86] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 1011.
[87] Y. Mitsuta, S. Yamanaka, T. Kawakami, M. Okumura, *J. Comput. Chem.* **2018**, *39*, 1913.
[88] A. Laio, F. L. Gervasio, *Reports Prog. Phys.* **2008**, *71*, 126601.
[89] M. Incerti, S. Russo, D. Callegari, D. Pala, C. Giorgio, I. Zanotti, E. Barocelli, P. Vicini, F. Vacondio, S. Rivara, R. Castelli, M. Tognolini, A. Lodola, *J. Med. Chem.* **2017**, *60*, 787.
[90] G. Bussi, A. Laio, *Nat. Rev. Phys.* **2020**, *2*, 200.
[91] A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, M. Parrinello, *J. Phys. Chem. B* **2005**, *109*, 6714.
[92] C. H. S. Da Costa, V. Bonatto, A. M. Dos Santos, J. Lameira, A. Leitão, C. A. Montanari, *J. Chem. Inf. Model.* **2020**, *60*, 880.
[93] L. Capoferri, A. Lodola, S. Rivara, M. Mor, *J. Chem. Inf. Model.* **2015**, *55*, 589.
[94] H. S. Yu, C. Gao, D. Lupyan, Y. Wu, T. Kimura, C. Wu, L. Jacobson, E. Harder, R. Abel, L. Wang, *J. Chem. Inf. Model.* **2019**, *59*, 3955.
[95] J. M. Smith, Y. J. Alahmadi, C. N. Rowley, *J. Chem. Theory Comput.* **2013**, *9*, 4860.
[96] R. Baer, E. Livshits, U. Salzner, *Annu. Rev. Phys. Chem.* **2010**, *61*, 85.
[97] E. Awoonor-Williams, A. G. Walsh, C. N. Rowley, *Biochim. Biophys. Acta Proteins Proteom.* **2017**, *1865*, 1664.
[98] E. Awoonor-Williams, W. C. Isley, S. G. Dale, E. R. Johnson, H. Yu, A. D. Becke, B. Roux, C. N. Rowley, *J. Comput. Chem.* **2020**, *41*, 427.
[99] L. M. Mihalovits, G. G. Ferenczy, G. M. Keserü, *J. Chem. Inf. Model.* **2020**, *60*, 6579.
[100] L. M. Mihalovits, G. G. Ferenczy, G. M. Keserű, *J. Chem. Inf. Model.* **2019**, *59*, 5161.
[101] X. Wang, G. M. Bakanina Kissanga, E. Li, Q. Li, J. Yao, *Phys. Chem. Chem. Phys.* **2019**, *21*, 12163.
[102] A. L. Bowman, I. M. Grant, A. J. Mulholland, *Chem. Commun.* **2008**, *37*, 4425.
[103] J. J. Ruiz-Pernía, E. Silla, I. Tuñón, S. Martí, *J. Phys. Chem. B* **2006**, *110*, 17663.
[104] I. Miyahisa, T. Sameshima, M. S. Hixon, *Angew. Chem. Int. Ed.* **2015**, *54*, 14099.
[105] J. M. Strelow, *J. Biomol. Screen.* **2017**, *22*, 3.
[106] S. Ray, A. S. Murkin, *Biochemistry* **2019**, *58*, 5234.
[107] E. Awoonor-Williams, C. N. Rowley, ChemRxiv **2020**, Preprint: https://doi.org/10.26434/chemrxiv.13132814.v1
[108] E. Awoonor-Williams, A. A. A. A. Abu-Saleh, *Phys. Chem. Chem. Phys.* **2021**, *23*, 6746.
[109] E. H. Krenske, R. C. Petter, Z. Zhu, K. N. Houk, *J. Org. Chem.* **2011**, *76*, 5074.
[110] V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995.
[111] V. Barone, M. Cossi, J. Tomasi, *J. Comput. Chem.* **1998**, *19*, 404.
[112] M. Shokhen, T. Traube, S. Vijayakumar, M. Hirsch, N. Uritsky, A. Albeck, *ChemBioChem* **2011**, *12*, 1023.
[113] S. Krishnan, R. M. Miller, B. Tian, R. D. Mullins, M. P. Jacobson, J. Taunton, *J. Am. Chem. Soc.* **2014**, *136*, 12624.
[114] E. H. Krenske, R. C. Petter, K. N. Houk, *J. Org. Chem.* **2016**, *81*, 11726.
[115] S. Banerjee, *Biophys. Chem.* **2021**, *269*, 106510.
[116] A. Paasche, M. Schiller, T. Schirmeister, B. Engels, *ChemMedChem* **2010**, *5*, 869.
[117] J. M. Smith, C. N. Rowley, *J. Comput.-Aided Mol. Des.* **2015**, *29*, 725.
[118] F. Palazzesi, M. R. Hermann, M. A. Grundl, A. Pautsch, D. Seeliger, C. S. Tautermann, A. Weber, *J. Chem. Inf. Model.* **2020**, *60*, 2915.
[119] F. Palazzesi, M. A. Grundl, A. Pautsch, A. Weber, C. S. Tautermann, *J. Chem. Inf. Model.* **2019**, *59*, 3565.
[120] A. Berteotti, F. Vacondio, A. Lodola, M. Bassi, C. Silva, M. Mor, A. Cavalli, *ACS Med. Chem. Lett.* **2014**, *5*, 501.
[121] R. M. Oballa, J. F. Truchon, C. I. Bayly, N. Chauret, S. Day, S. Crane, C. Berthelette, *Bioorganic Med. Chem. Lett.* **2007**, *17*, 998.
[122] V. Ehmke, J. E. Q. Quinsaat, P. Rivera-Fuentes, C. Heindl, C. Freymond, M. Rottmann, R. Brun, T. Schirmeister, F. Diederich, *Org. Biomol. Chem.* **2012**, *10*, 5764.
[123] M. Mladenovic, K. Junold, R. F. Fink, W. Thiel, T. Schirmeister, B. Engels, *J. Phys. Chem. B* **2008**, *112*, 5458.
[124] A. Lodola, L. Capoferri, S. Rivara, G. Tarzia, D. Piomelli, A. Mulholland, M. Mor, *J. Med. Chem.* **2013**, *56*, 2500.
[125] M. G. Quesne, R. A. Ward, S. P. de Visser, *Front. Chem.* **2013**, *1*, 39.
[126] P. Chaskar, V. Zoete, U. F. Röhrig, *J. Chem. Inf. Model.* **2017**, *57*, 73.
[127] J. Fanfrlík, P. S. Brahmkshatriya, J. Řezáč, A. Jílková, M. Horn, M. Mareš, P. Hobza, M. Lepšík, *J. Phys. Chem. B* **2013**, *117*, 14973.
[128] D. Callegari, K. E. Ranaghan, C. J. Woods, R. Minari, M. Tiseo, M. Mor, A. J. Mulholland, A. Lodola, *Chem. Sci.* **2018**, *9*, 2740.
[129] M. Ahsan, S. Senapati, *J. Phys. Chem. B* **2019**, *123*, 7955.
[130] W. Ren, M. Farren-Dai, N. Sannikova, K. Świderek, Y. Wang, O. Akintola, R. Britton, V. Moliner, A. J. Bennet, *Chem. Sci.* **2020**, *11*, 10488.
[131] K. Arafet, N. Serrano-Aparicio, A. Lodola, A. J. Mulholland, F. V. González, K. Świderek, V. Moliner, *Chem. Sci.* **2021**, *12*, 1433.
[132] K. Arafet, S. Ferrer, V. Moliner, *Biochemistry* **2015**, *54*, 3381.
[133] J. R. A. Silva, L. Cianni, D. Araujo, P. H. J. Batista, D. De Vita, F. Rosini, A. Leitão, J. Lameira, C. A. Montanari, *J. Chem. Inf. Model.* **2020**, *60*, 1666.
[134] A. M. Dos Santos, L. Cianni, D. De Vita, F. Rosini, A. Leitão, C. A. Laughton, J. Lameira, C. A. Montanari, *Phys. Chem. Chem. Phys.* **2018**, *20*, 24317.
[135] E. I. Chudyk, M. A. L. Limb, C. Jones, J. Spencer, M. W. Van Der Kamp, A. J. Mulholland, *Chem. Commun.* **2014**, *50*, 14736.
[136] R. A. Fritz, J. H. Alzate-Morales, J. Spencer, A. J. Mulholland, M. W. Van Der Kamp, *Biochemistry* **2018**, *57*, 3560.
[137] M. H. Jasim, D. L. Rathbone, *J. Mol. Graphics Modell.* **2018**, *79*, 157.
[138] G. S. Sirin, Y. Zhang, *J. Phys. Chem. A* **2014**, *118*, 9132.
[139] M. A. B. Morais, J. Coines, M. N. Domingues, R. A. S. Pirolla, C. C. C. Tonoli, C. R. Santos, J. B. L. Correa, F. C. Gozzo, C. Rovira, M. T. Murakami, *Nat. Commun.* **2021**, *12*, 367.
[140] Y. H. Cheng, X. L. Cheng, Z. Radić, J. A. McCammon, *Chem.-Biol. Interact.* **2008**, *175*, 196.
[141] D. Wei, B. Lei, M. Tang, C. G. Zhan, *J. Am. Chem. Soc.* **2012**, *134*, 10436.
[142] D. Wei, M. Tang, C. G. Zhan, *Org. Biomol. Chem.* **2015**, *13*, 6857.
[143] M. P. Andersson, P. Uvdal, *J. Phys. Chem. A* **2005**, *109*, 2937.
[144] V. Zoete, T. Schuepbach, C. Bovigny, P. Chaskar, A. Daina, U. F. Röhrig, O. Michielin, *J. Comput. Chem.* **2016**, *37*, 437.

[145] D. Branduardi, F. L. Gervasio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 1.
[146] D. Branduardi, M. De Vivo, N. Rega, V. Barone, A. Cavalli, *J. Chem. Theory Comput.* **2011**, *7*, 539.

AUTHOR BIOGRAPHIES



Levente M. Mihalovits is a PhD student at the Budapest University of Technology and Economics (BME). He completed his BSc in 2016, and his MSc in 2018 at BME. He is currently working in the Medicinal Chemistry Research Group at the Research Centre for Natural Sciences as a Computational Chemist. His main focus is on the modeling of covalent inhibition reactions using QM/MM coupled with molecular dynamics simulations.

György G. Ferenczy is a senior research fellow at the Research Centre for Natural Sciences and at the Semmelweis University. He obtained a PhD in computational chemistry from the Eötvös University of Budapest and was postdoc at the University of Oxford and at the University of Nancy. Later he worked as a computational chemist at Gedeon Richter and Sanofi. His research interest includes the development and application of computational tools for extended biochemical systems relevant to drug discovery.

György M. Keserű obtained his Ph.D. at Budapest, Hungary. He worked for Sanofi and Gedeon Richter. He contributed to the discovery of the antipsychotic Vraylar® (cariprazine) that has been approved and marketed from 2016 in US and EU. From 2015 he is heading the Medicinal Chemistry Research Group at RCNS and full professor at the Budapest University of Technology and Economics. His research interests include medicinal chemistry and drug design.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: L. M. Mihalovits, G. G. Ferenczy, G. M. Keserű, *Int. J. Quantum Chem.* **2022**, *122*(8), e26768. https://doi.org/10.1002/qua.26768