Design and SAR Analysis of Covalent Inhibitors Driven by Hybrid QM/MM Simulations

Design and SAR analysis of covalent inhibitors driven by hybrid QM/MM simulations

Alessio Lodola,* Donatella Callegari, Laura Scalvini, Silvia Rivara, and Marco Mor

Drug Design and Discovery Group, Department of Food and Drug, University of Parma, Parma, Italy

Running title: *QM/MM simulations in drug design*

Key Words: QM/MM, quantum mechanics/molecular mechanics, covalent inhibitor, drug design, SAR, FAAH, EGFR

Corresponding Author:

Alessio Lodola,
Associate Professor of Medicinal Chemistry,
Department of Food and Drug,
University of Parma, Parma, Italy
alessio.lodola@unipr.it

Abstract

Quantum mechanics/molecular mechanics (QM/MM) hybrid technique is emerging as a reliable computational method to investigate and characterize chemical reactions occurring in enzymes. From a drug discovery perspective, a thorough understanding of enzyme catalysis appears pivotal to assist the design of inhibitors able to covalently bind one of the residues belonging to the enzyme catalytic machinery. Thanks to the current advances in computer power, and the availability of more efficient algorithms for QM-based simulations, the use of QM/MM methodology is becoming a viable option in the field of covalent-inhibitor design. In the present review, we summarized our experience in the field of QM/MM simulations applied to drug design problems which involved the optimization of agents working on two well-known drug-targets, namely fatty acid amide hydrolase (FAAH) and epidermal growth factor receptor (EGFR). In this context, QM/MM simulations gave valuable information in term of geometry (i.e. of transition states and metastable intermediates) and reaction energetics that allowed to correctly predict inhibitor binding orientation and substituent effect on enzyme inhibition. What is more, enzyme-reaction modelling with QM/MM provided insights that were translated into the synthesis of new covalent inhibitors features by a unique combination of intrinsic reactivity, on-target activity and selectivity.

Introduction

The recognition process involving a drug and a target leads to the formation of a molecular complex capable of triggering a series of biochemical events that culminates in the insurgence of a therapeutic effect. The amplitude of this favourable pharmacological response is primarily dependent on the fraction of the molecular target occupied by the drug.[1] In condition of equilibrium, thermodynamics provides the driving force for drug binding i.e., the protein-ligand complex is formed because it possesses a lower free-energy than the protein and the ligand taken on their unbound state. However, in condition of non-equilibrium such as in cells or in living organisms, not only thermodynamics matters.[2] In tissues, where small molecules are subject to metabolic transformations and drug targets are exposed to high concentration of their physiological substrates, the kinetic stability of the protein-ligand complex becomes important for observing the pharmacological response.[3] Specifically, the residence time of the small molecule at its target should be long enough to trigger those biochemical steps which are required for the insurgence of the pharmacological effect.[4] If this is not the case, hardly a small molecule will exert a biological effect when administered *in vivo*.[5] A practical way to increase the residence time of a small molecule at its target is to install on it a reactive group able to form a stable covalent bond with the receptor. If this new covalent bond is protected by solvolysis or, more generally, if it possesses high chemical stability, the drug target will likely result occupied for its entire life, with a significant advantage in term of efficacy and duration of action.[6] Irreversible binding through covalent addition to the target is probably the most powerful strategy to obtain a sustained response *in vivo* since release from inhibition requires the re-synthesis of the engaged target.[7]

Two different strategies are currently applied to identify covalent inhibitors for a given target.[8] Both rely on the presence of a nucleophile, either member of a catalytic machinery or simply with a structural role, in the binding pocket of the protein of interest. In the first approach, briefly cited above, a scaffold of a reversible inhibitor already featured by a good

affinity of the target is identified and an electrophile warhead is inserted on its structure.[9] If the electrophile is positioned close to the nucleophile with a geometry not too far from that of the hypothetical *transition state* for the reaction, a covalent bond can be formed within a reasonable time interval. A second approach used to identify a covalent inhibitor is based on the screening of libraries of electrophilic compounds of relatively small size. Once a covalent modifier is identified (either using mass-spectrometry or biochemical assays), an optimization phase takes place. During this task, the non-reactive portion of the inhibitor (often called driver portion) is progressively modified to improve the stereo-electronic complementarity between the target and the inhibitor with the final aim of improving the potency.[10] Whether a covalent inhibitor is identified through rational design or in a screen, other runs of optimization are required to properly balance potency and selectivity. Again, this task can be done by modifying the driver portion of the inhibitor or altering the intrinsic reactivity of the warhead toward target nucleophile.[11]

In the context of discovery and optimization of covalent agents, computer simulations played and are going to play a significant role both in academia and industrial settings.[12,13] Ligand-based and structure-based computer-aided drug design techniques[14] have been successfully used to design new covalent inhibitors without explicitly accounting for their chemical mechanism of action.[15,16] Most of these applications work under the hypothesis that driver portions can be optimized independently from the inhibitor warhead.[17] This condition is expected to be valid only when modifications in the driver portion do not have an impact on the dynamic of the covalent reaction or on the stability of the reaction product. In all the other cases, a useful computational method should take into account the effect of chemical substitution both on the recognition and on the reactivity phases.[18] Only recently, a number of computational approaches aimed at driving the design and optimization of covalent inhibitors have been reported in the literature.[19,20] Among those, hybrid quantum mechanics/molecular mechanics (QM/MM) method, [21] is probably the only one that allows

to characterize the mechanism of action of a covalent inhibitor (see for instance references 22, 23, 24, 25, 26, 27) accounting for the accommodation of the driver portion and for the chemical reaction occurring between a nucleophile on the given target and the inhibitor warhead.[28] This is possible as the QM/MM approach combines the accuracy of QM methods in describing chemical transformation with the ability of MM force fields to provide a reasonable description of conformational energetics and non-bonded interactions in condensed systems.[29]

In the present paper, after summarizing the basics of the QM/MM approach and the theoretical strategy to compute-energy barriers in the context of enzyme-catalysed reaction, we will report two case studies where these methods have been applied to elucidate the mechanism of action of reference inhibitors on drug targets of current pharmacological or clinical interest. These are fatty acid amide hydrolase (FAAH),[30] one of the key enzymes involved in the metabolism of endocannabinoids and neuromodulatory fatty-acid ethanolamides[31] and epidermal growth factor receptor (EGFR),[32] a tyrosine-kinase receptor controlling cell proliferation and tissue renewal. Eventually, we show how the QM/MM approach can be used in perspective drug design.

2. The QM/MM method

2.1. Definition of the methodology

In QM/MM methodology, QM and a MM theory are coupled together to define a convenient QM/MM hybrid Hamiltonian that, applied in combination to geometry optimization, molecular dynamics or Monte Carlo sampling, is able to describe chemical reactions occurring in molecular systems composed by thousands of atoms.[33,34] Other QM-based approaches have been generated to investigate chemical reactions in condensed phase, such as the empirical-valence bond (EVB) method developed by Warshel,[35] in which

reactions are described through mixing states corresponding to classical valence-bond structures describing reactants, intermediates and products, and Car-Parrinello Molecular dynamics (CPMD),[36] where the valence electrons of molecular structures are described by plane-wave basis sets rather than atom-centred ones. For more information on these methodologies, the reader is invited to look at more specialized reviews.[37,38]

In the QM/MM approach reported here, which theory was firstly described by Warshel and Levitt in 1976[39] and then refined by Karplus and co-workers,[40] the system is divided in two regions of interests, named *reactive region* and *boundary region*. The *reactive region* of the active site is described at QM level, either based on the Hartree-Fock formalism or on the density functional theory (DFT).[41] A practical choice is to include in this QM region all the atoms participating to the breakage and forming of covalent bonds. A molecular mechanics force field is instead used to describe the *boundary region*, i.e. the rest of the system not directly involved in the chemical transformation, but which influences the *reactive region* through non-bonded interactions.[42] The interaction between *reactive region* (QM) and the *boundary region* (MM) has to be taken into account at any stage of the calculation. This is accomplished with the use of the following expression (eq.1) which includes three key terms:[43]

$$H = H_{QM} + H_{QM/MM} + H_{MM} \qquad \text{(eq. 1)}$$

where the H_{QM} is the Hamiltonian of *reactive region* treated at QM level, $H_{QM/MM}$ is the Hamiltonian that couples *reactive* and *boundary region*, and H_{MM} is the Hamiltonian of *non-reactive region*, treated at classical level. In principle, the quantum region can be described by any QM method. For instance, semi-empirical (SE) methods derived from Hartree-Fock theory, such as AM1 and PM3, can be used with the advantage of allowing very large QM systems (up to hundreds of atoms) to be treated.[44] However, these SE have known issues which may give a suboptimal description of the reaction energetics. More reliable, but highly expensive, calculations can be performed using the electron correlation methods. These

include the MP2 perturbation methods and coupled-cluster theory.[45] A computationally less demanding alternative is represented by the use of density functional theory (DFT) methods. In this context, the self-consistent charge density functional tight binding (SCC-DFTB) approach,[46] including its more recent DFTB3 variant,[47] have been proposed as a fair alternative to expensive *ab initio* methods,[48] at least in the context of free-energy simulations where an extensive sampling of the molecular system under study is required to obtain meaningful information on the reaction mechanism. Despite DFTB-based approaches are affected by systematic errors[49] (i.e. with proton affinities,[50] H-bond energetics[51] and dispersion forces[52]), they have often provided reasonable results,[53] in fair agreement with calculations performed at higher level of theory[54,] or with experimental data.[55,56]

Specialized protein force fields including CHARMM[57] and AMBER[58] are routinely used to describe atoms belonging to the *boundary* regions, and thus to account for the H_{MM} Hamiltonian. With limitations due to the empirical nature of the mathematical functions employed to describe bonding and non-bonding interactions among atoms and the local validity of parameters used to build and calibrate these functions, MM force fields can provide a reasonable description of the conformational energetics of macromolecules and of non-bonded interactions in large systems.[59] For these reasons, they are applied in nearly all available QM/MM implementations.

As anticipated above, a QM/MM hybrid scheme must combine the quantum and classical parts to obtain a meaningful Hamiltonian. In general, the $H_{QM/MM}$ term account for the QM/MM interaction energy, which involves *i.* electrostatic interactions and *ii.* van der Waals interactions between atoms of the QM and of the MM regions. Based on how the QM part and MM are coupled and therefore on how electrostatic interactions are computed, three different QM/MM approaches can be identified.[60] In the first one, known as *mechanical coupling*, the QM Hamiltonian is not influenced by the point charges of the MM system. Instead, a force field is used to account for the interactions between the QM and MM region.

Atom types are assigned also to the QM atoms, and the QM/MM interaction energy between these two regions is calculated by simply summing the electrostatic energy and the Van der Waals contribution, simply using the Coulomb law and the Lennard-Jones potential, respectively.

The second one is based on the *electrostatic coupling* approach, where the electrostatic interaction between the MM and the QM region is obtained through the inclusion of classic charges belonging to MM atoms directly in the QM Hamiltonian, i.e. by incorporating MM atomic charges directly in one-electron integrals. MM atomic partial charges also interact with the nuclei of the atoms in the QM system. The electrostatic interaction of QM electrons with MM point charges moves from the $H_{QM/MM}$ term of the *mechanical coupling* scheme (where it was described with a force field) to the quantum Hamiltonian, H_{QM} according to equation 2:[60]

$$\hat{H}^{el}_{QM(MM)} = -\sum_i^N \sum_{J\in\Phi}^L \frac{q_J}{|\boldsymbol{r}_i - \boldsymbol{R}_J|} + \sum_{\alpha\in I+L}^M \sum_{J\in\Phi}^L \frac{q_J Q_\alpha}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_J|} \quad \text{(eq. 2)}$$

where, the q_J are the MM point charges separated by the R_J distance; Q_α are the nuclear charges of the QM atoms at R_α distance; and r_i describes electron positions. The indices *i*, *J*, and α run over the *N* electrons, *L* point charges, and *M* QM nuclei, respectively.

This electrostatic coupling is probably the most popular embedding scheme used today, at least in the context of biomolecular simulations,[61] as it allows to account for the electrostatic influence of MM atoms on the QM region at a reasonable computational cost. The inclusion of this effect is likely to be important for many enzymes, given the polar nature of several enzyme active sites. The third approach is based on the *polarisation coupling* scheme, in which classical charges of the MM atoms are not fixed but change on the base of the electrostatic potential generated by QM core. The QM and MM regions experience a mutual polarisation process, which improves the description of the electrostatic forces[62] but at a greater computational cost.

In all the three coupling schemes described above, QM/MM van der Waals interactions (representing dispersion and exchange repulsion interactions between QM and MM atoms) are calculated by MM. This requires that MM van der Waals parameters have to be assigned to each QM atom. This procedure is critical as van der Waals interactions are important at short range where they affect both QM/MM interaction energies and reactant geometries. [63] A known limitation of most QM/MM implementations is that the same set of van der Waals parameters is used for the QM atoms during a simulation. The nature of the groups involved in a chemical reaction (treated by QM) is subject to change as the reaction proceeds and the van der Waals parameters assigned to the QM atoms should be modified accordingly.[64] The effect of using different van der Waals parameters in QM/ MM simulations has been investigated. It was concluded that thermodynamic quantities such as the potential of mean force are minorly affected by the van der Waals parameters. The improvement of the consistency of QM/MM methods for predicting energetic properties should thus focus on other factors.[21]

2.2 Exploration of reaction mechanism

In the case of small systems where the use of a hybrid QM/MM approach is not required, QM alone can be used to characterise the potential energy surface (PES) of a chemical process by means of geometry optimisation, coupled with calculations of second derivatives. This approach allows to identify stationary points along a minimum-energy path (MEP) and to roughly estimate their free-energy after correcting potential energy for zero-point vibration as well as for thermal and entropy effects.[65]

In this condition, the application of the *transition state theory* can allow to compare the calculated activation barrier to that one derived from the experimentally measured reaction rate.[66] If the barrier for the proposed mechanism is considerably larger than the experimental one, such a mechanism should not be trusted. On the same line, a mechanism with a calculated barrier comparable to the apparent experimental barrier is more plausible.

This workflow requires direct calculation, storage and manipulation of the Hessian matrix, which contains the topological information of the PES of interest. It becomes extremely difficult to handle with molecular systems with many degrees of freedom.[67] In context of QM/MM simulation, the exploration of the PES for a given mechanism can be performed using the adiabatic mapping approach, which does not require calculation of second derivatives. According to this method, the energy of the system is calculated after optimizing the structure of the reactants at a series of harmonically restrained values of a reaction coordinate, describing the process of interest such as the distance between two atoms. This approach is only valid if one conformation of the protein can represent the state of the system at a particular value of the reaction coordinate.[68] This is expected to be true when negligible structural changes occur at the active site with the progression of the reaction.[69] Only in this very specific condition, the minimisation of the potential energy along the specified reaction coordinate provides a reasonable approximation of the “enthalpic component” of the free energy profile for the considered reaction.[70]

In other situations, the conditions described above are not satisfied, and structural fluctuations of the protein active site including those of the substrate may represent essential movements of the catalytic process.[71] The use of multiple structures as a starting point for reaction modelling offers a partial solution to the limited structural flexibility allowed by the adiabatic mapping approach. For a given reaction path, multiple PESs can be explored and several barrier heights can be computed and compared.[72] Depending on the type of statistical distribution, a convenient averaging method can be used to obtain a representative barrier value.[73] In this condition, it is also possible to estimate the uncertainty of the computed barrier which is critical when different reaction mechanisms are compared.[74] Regardless of the statistical robustness of the computed potential energy barrier, QM/MM calculations based on energy minimization remains potentially erratic. For instance, adiabatic mapping overestimates energy barriers if atom movements connected

with the reaction are not included in the definition of the reaction coordinate, essentially because minimization alone cannot resolve steric strains caused by the progression of the system along the given reaction coordinate.[75] Despite these limitations, adiabatic mapping is useful for preliminary exploration of the PES, including the generation of models of transition states and intermediates. Several pioneering QM/MM works on enzyme reaction modelling were conducted with this approach and still preserve their validity.[21,60,68]

2.3 Calculation of Potential of Mean Force (PMF)

Computational techniques that sample multiple configurations of the molecular system along a given path have the capability to contribute a better description of a chemical transformation.[76] These approaches account for the conformational space explored by reactants, TSs and products, as well as for solvent reorganization, during the chemical transformation of interest,[77] allowing to compute the variation of the free-energy of the system as a function of the progression of the reaction. The resulting free energy profile, which is computed along a well-defined reaction coordinate, represents the potential of mean force (PMF) of the simulated process. Its knowledge is critical for enzyme reaction modelling as it allows to directly retrieve the activation free energy for the chemical transformation of interest. Thanks to the application of the *transition state theory,* this computed quantity can be meaningfully compared to the experimental barrier for the enzyme-catalysed reaction under study.[78,79]

In principle QM/MM simulations coupled with molecular dynamics (MD) might allow to calculate activation free-energy. However, classical MD does not explore phase space regions with high energy, as for transition states or transient intermediates. Monte Carlo (MC) methods overcome this limitation,[80] but the number of QM/MM MC iterations necessary to obtain converged properties is extremely high, making the overall process computationally demanding.

Conformational sampling of processes describing chemical reactions requires enhanced sampling (ES) methods in which an external bias is applied to sample high energy region. Umbrella sampling (US) is one of the method capable to achieve this task.[81] In this technique, a biasing potential (the “umbrella” potential) is applied to force the system to remain at a specific value of a given reaction coordinate.[82] Furthermore, this bias makes sure that all regions along the reaction coordinate are sampled with similar probability.[83]

QM/MM US simulations are performed along a chemical path connecting reactants and products (or stable intermediate). Very often US simulation begins with QM/MM MD of the Michaelis complex with a bias potential applied to restrain the reaction coordinate to a value corresponding to the reactant state. In subsequent US simulations, the reference value of the restraint is increased by a small quantity to ensure the progression along the reaction coordinate, while allowing the sampling of the available conformational space for a relatively short time-interval, routinely 50-100 ps with the current computational capability. As for adiabatic mapping methods (see above), the reaction coordinate is often defined in terms of distances involving bond lengths. In US simulations a typical difference between two consecutive points (or windows) is in the range of 0.1-0.2 Å. The reaction coordinate values assumed at each window of the US simulation are recorded. At the end of the whole process, the effects of the restraining potentials are removed and combined by the weighted histogram analysis method (WHAM) which gives the unbiased PMF along the reaction coordinate.[84] To obtain a continuous PMF for the modelled reaction, it is important that neighbouring US windows produce overlapping distributions in term of dispersion of the explored value of the reaction coordinate. This condition can be met by conveniently tuning the magnitude of the spring constant and/or the step-size of the reaction coordinate. As with other sampling methods, it is important to test the convergence of the PMF with respect to length of the simulation. It should be stressed that while US allows to sample multiple configurations of the system along a given path, it will fail to correctly estimate the free-

energy difference between the two states if a major rearrangement of the conformation of the proteins occurs during the reaction especially, if this conformational transition is not described by the reaction coordinate.

Sampling reaction paths with US is computationally demanding since the underlying simulation requires a very large number of QM/MM computations of the potential energy including gradients. Hence US is used in combination with QM/MM approached derived from semi-empirical Hamiltonian such as AM1, PDDG-PM3 and SCC-DFTB. These semi-empirical methods are featured by systematic errors that often lead to rather inaccurate energy barriers. The accuracy of QM/MM US barriers can be improved by applying a correction term derived from adiabatic mapping study performed at a higher level of theory. This practice is well documented in the literature and it has been shown to give reaction barriers in close agreement to the experimental ones. [21,60]

Other ES methods have been used in the context of QM/MM simulations. Non-equilibrium methods are viable alternatives to US sampling and can provide rather rapid (but coarse) estimation of the activation free-energy of a chemical transformation. One example is steered molecular dynamics (SMD),[85] a method in which an external force is applied to part of the system, e.g., the distance separating two atoms, to drive it along a predefined direction.

While specifically suited for investigating ligand unbinding from proteins with MM force field,[86] SMD has been proposed as an efficient sampling method in the context of QM/MM simulations. [87] The dissipative work performed on the system (W) during a QM/MM steered-MD simulation is calculated by numerical integration of the exercised force (F_{ex}) at each value of the reaction coordinate (eq3), where dx = v dt.

$$W_{[x(t)]} = \int_0^{x(t)} F_{ex}(t)\, dx(t) \qquad \text{(eq.3)}$$

QM/MM SMD simulations are usually carried out by sequentially pulling the system along the reaction coordinate using a harmonic potential whose centre moves with constant

velocity and whose Hooke's constant is few hundreds of kcal/mol Å^{-2}. At each QM/MM MD integration step, the equilibrium position of the harmonic potential is increased of a fixed amount. A pulling velocity in the range of 0.01-0.1 Å/ps should allow to reduce the effect of dissipative frictional work on the overall work (W) computed on the system.

SMD is useful for comparing alternative reaction mechanisms. While single SMD run can only provide a qualitative description of the change in the mechanical work performed along the reaction path, independent works calculated through multiple SMD simulations can be combined together applying the Jarzynski's equality (eq. 4), providing a means of recovering an equilibrium PMF from non-equilibrium events. [88]

$$\langle e^{-\Delta F/k_B T} \rangle = \langle e^{-W/k_B T} \rangle \qquad \text{(eq. 4)}$$

It should be pointed out that obtaining statistically converged and accurate PMF profiles is far from easy and may result in increased computational cost if the steering velocity and number of replicas are not properly selected.[89]

Example of QM/MM SMD simulations applied to elucidate reaction mechanism are represented by the seminal work of Adrian Roitberg on Chorismate Mutase,[87] and by others both in the field of organic reaction in solution[90] and in that of enzyme-catalysed processes.[91] Also metadynamics,[92] a non-equilibrium ES approach primarily used to investigate binding modes for drugs,[93,94] conformation transitions of proteins[95] and unbinding events of small-molecules[96] from enzyme[97] and receptors,[98] has been largely used in the context of chemical reaction modelling. Several examples of its use in enzyme catalysis can be found in reference 99 as well as in more recent articles.[100]

2.4 Path-collective variables in QM/MM simulations

QM/MM mechanistic modelling of complex enzymatic reactions is often carried out by separating the catalytic process into a sequence of consecutive steps. This approach is viable only when stable intermediates can be identified along a given path in a way that the

end-point of a chemical step can be used as starting-point for the subsequent one. Furthermore, the arbitrary separation of a complex chemical transformation in independent steps may prevent to capture the concerted nature of the enzyme-catalyzed reactions.

The path-collective variable (PCV) approach overcomes these issues allowing to model a multi-step process in a single simulation run. When PCVs are used, a multidimensional problem is reduced to only two dimensions. The first one, usually called S, describes the *progress* of the chemical transformation, the second one, Z, defines the *tolerance* versus the exploration of paths alternative to that tracked by S (*vide infra*).[101]

The variable S (eq. 5) is defined as a multiparametric function dependent on the position of atoms involved in the chemical transformation with respect to a reference path connecting the reactants to the products. This path is a discrete set of arbitrary geometries (called nodes) representative of the chemical transformation of interest and can be generated by extracting a set of frames from a PES of the reaction of interest, or from a series of consecutive SMD simulations in which the reactants are gradually transformed into the products using classical reaction coordinates.[102]

$$S(R) = \frac{\sum_{i=1}^{P} i e^{-\lambda(R-R(i))^2}}{\sum_{i=1}^{P} e^{-\lambda(R-R(i))^2}} \qquad \text{(eq. 5)}$$

In eq. 5, R is a vector representing the cartesian coordinates of the atoms involved in the chemical transformation. $(R-R_{(i)})^2$ corresponds to the squared difference between the coordinates assumed by the system at a given time (R) and the coordinates of the *i-th* node ($R_{(i)}$) along the reference path. λ is a tunable parameter that helps the transition from one node (i) to the following one (i+1) and its value depends on the average distances among nodes of the reference path.

A second multiparametric variable (Z) is employed in combination with S. According to its definition (eq. 6), Z is the distance from the input path and can be used to define a

confinement boundary which prevents the exploration of configurations of the system which are too far from the nodes of the reference path.

$$Z(R) = -\frac{1}{\lambda} ln \left(\sum_{i=1}^{P} e^{-\lambda (R - R(i))^2} \right) \quad \text{(eq.6)}$$

However, as the most probable reaction path does not necessarily coincide with Z = 0, it is important that the bias acting on Z is sufficiently soft to allow some deviations from the reference path. QM/MM MD simulations performed on the S/Z space using US or SMD allow to calculate the PMF for the chemical transformation of interest and to verify if a lower free-energy pathway alternative to the reference one can be identified. If this is the case, a novel reference path needs to be defined. This is done selecting a set of discrete frames from the MD trajectory along the minimum free-energy path applying the fitting procedure developed by Maragliano and Vanden-Eijnden.[103] Thus a new QM/MM simulation in the S/Z space is performed along the nodes defined by the new path. This procedure is repeated iteratively until the minimum-free energy path does not change passing from a PCV simulation to a new one. The PCV approach here briefly summarized has been initially developed to describe conformational transitions in macromolecules.[101] More recently it has been applied to model chemical reactions in solution[104] and within enzyme. [105,106]

3. QM/MM studies on relevant drug targets

3.1. Fatty acid amide hydrolase (FAAH)

3.1.1. Catalytic mechanism for acylation and substrate selectivity

Fatty acid amide hydrolase (FAAH) is a member of the amidase signature family responsible for the inactivation of biologically active fatty acid ethanolamides.[30] FAAH terminates the signal brought by the endocannabinoid anandamide, catalysing its hydrolysis to arachidonic acid and ethanolamine. FAAH catalyses the hydrolysis of other amides such as oleoylethanolamide and palmitoylethanolamide which are agonists of the peroxisome

proliferator-activated receptor (PPAR) subtype alpha. Inhibition of FAAH by small molecules has emerged as promising therapeutic strategy to treat central nervous diseases, including anxiety, inflammation and depression.[107]

FAAH possesses a unique catalytic triad composed of two serine residues (Ser217 and Ser241), and one lysine (Lys142), rather than the serine-histidine-aspartate triad found in classical serine hydrolases. This Ser-Ser-Lys triad is responsible for the remarkable ability of this enzyme to hydrolyse amides and esters at a similar rate, by a mechanism in which acylation is the rate-limiting event.[108] As the maximum hydrolytic activity of FAAH is observed in mild-basic conditions (pH = 9), it is believed that Lys142 participates to catalysis in its neutral form, i.e. with Lys142 triggering a series of deprotonation events which culminate in the formation of Ser241 alcoholate anion, the key nucleophilic species of the first step of acylation.[109] Computational investigations performed at B3LYP/6-31+G(d)//PM3/CHARMM22 level by adiabatic mapping[110] and PDDG-PM3/OPLS level by FEP in combination with MC sampling[111] support this role for Lys142.

Recently, an SCC-DFTB/CHARMM27 QM/MM potential has been used to model the process of FAAH acylation both in the presence of oleamide (OA) and oleolymethyl ester (OME), two model substrates for which experimental kcat have been measured in the same experimental conditions, allowing to test reliability of the employed QM/MM approach.[55]

PES surfaces for FAAH acylation in the presence of OA and OME have been explored by adiabatic mapping, according to the mechanism proposed by McKinney and Cravatt.[112] In the first step of acylation, Lys142 activates Ser241 nucleophile which in turn attacks the substrate carbonyl carbon, with formation of an anionic tetrahedral intermediate (TI). In the second step, the positively charged Lys142 readily protonates the leaving group, leading to its expulsion. The role of Lys142 on Ser241 activation and leaving group protonation is mediated by Ser217 which acts as key proton shuttle of the overall process (Figure 1).

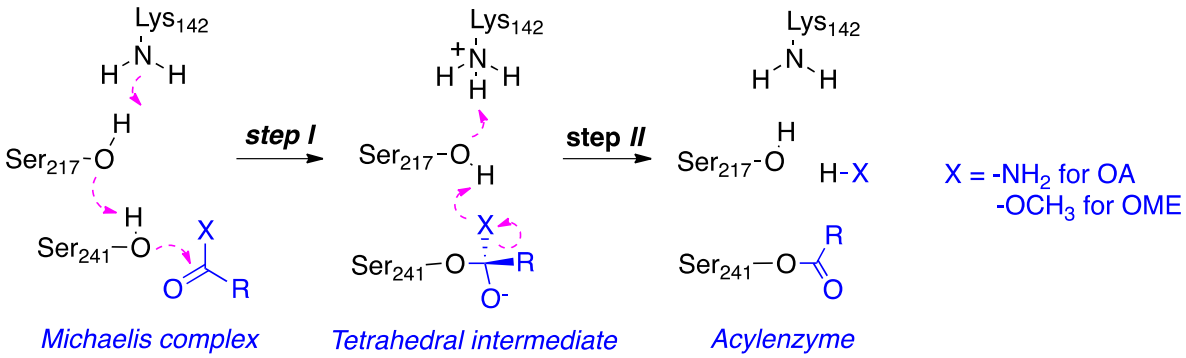

Figure 1. Acylation mechanism of FAAH in the presence of oleamide (OA) and oleoylmethyl ester (OME).

The exploration of the PESs for FAAH acylation at SCC-DFTB/CHARMM27 level indicates that the collapse of the TI is the rate-limiting step of the reaction for both OA and OME substrates, with calculated barriers of 19 and 21 kcal/mol at SCC-DFTB/CHARMM27 level. Similar findings have been obtained by Palermo et al., who modelled the hydrolysis of anandamide catalysed by FAAH using a CPMD approach.[113] CPMD simulations identified protonation of the leaving group as the rate limiting step of acylation with a free-energy barrier of 19 kcal/mol.

When SCC-DFTB/CHARMM27 energy values were averaged over multiple reaction paths (n = 6) the energy barriers became 21 ± 1 kcal/mol for OA and 25 ± 1 kcal/mol for OME,[55] still in acceptable agreement with the experimentally deduced barriers of 16 kcal/mol for OA and 17 kcal/mol for OME, respectively. In qualitative term, the relevant experimental preference showed by FAAH in hydrolysing OA at higher rate that OME is satisfactorily reproduced by SCC-DFTB/CHARMM27 calculations. Analysis of minimum potential energy path on the PES show that protonation of the leaving group promoted by Lys142 emerged as the difficult step of the acylation for both OA and OME. The higher basicity of the nitrogen atom compared to that of oxygen in the TI configuration account for the lower barrier of leaving group protonation calculated for OA with respect to OME. This finding provides a theoretical explanation for the remarkable ability of FAAH to hydrolyse amides faster than esters.

3.1.2. Identification of productive binding orientation for covalent inhibitors

A QM/MM approach has also been used to clarify the reaction of mechanism of a class of carbamic acid aryl esters designed to inhibit FAAH. The reference compound of the class, URB597 (the 3'-$CONH_2$ derivative of URB524),[114] was demonstrated to inhibit FAAH by carbamoylating Ser241.[115]

Molecular modelling investigation performed by our research group suggested URB524 and its derivatives, including URB597, can be docked within the FAAH catalytic site according to two possible orientations, both placing the carbamoylating group of this inhibitor in proximity to Ser241.[116] In the first (*binding orientation I,* Figure 2A), the *m*-biphenyl moiety of URB524 occupied the so-called acyl-chain channel of FAAH i.e., where the arachidonoyl portion of AEA is accommodated,[117] while in the second one (*binding orientation II,* Figure 2A), the *m*-biphenyl group lied in the cytosolic pocket of FAAH,[118] i.e., where the ethanolamine head of AEA lays.

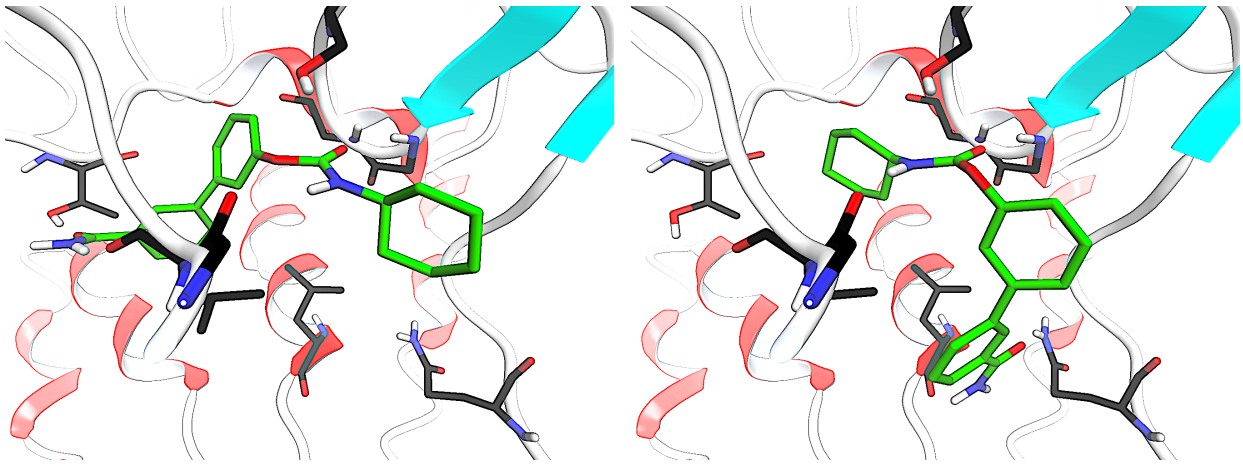

Figure 2. Binding orientations for URB597 (green carbon atoms), docked within the active site of rat FAAH (black carton atom). Left panel, *binding orientation I*; right panel, binding orientation *II.*

A QM/MM approach was used to model the inhibitor binding process, applying a PM3/CHARMM22 potential corrected at B3LYP/6-31G+(d) level. PES surfaces of Ser241 carbamoylation were built for both binding orientations according to the reaction scheme depicted in Figure 3.

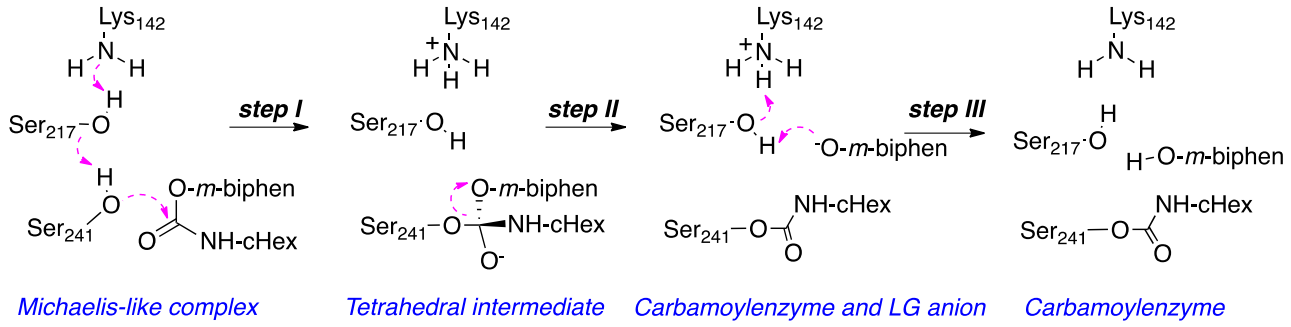

Figure 3. Modelled mechanism of carbamoylation of FAAH Ser241 by QM/MM.

Calculations showed that the carbamoylation in orientation *II* was energetically preferred (by nearly 15 kcal/mol at B3LYP/6-31+G(d)//PM3/CHARMM22 level) over that one occurring in orientation *I*, proposing the second arrangement of URB524 at FAAH active site as the productive binding mode.[119] This prediction was later confirmed by the crystallographic resolution of the FAAH-URB597 structure.[120] Furthermore, in the mechanism occurring from binding orientation *II*, activation of Ser241 emerged as difficult step of the reaction (step I, Figure 3). Nucleophilic attack performed by Ser241, which led to the formation of the TI, took place with a low barrier. Similarly, the expulsion of the byphenolate leaving group (step II, Figure 3) occurred with a negligible barrier, leading to the formation of stable carbamoylated adduct. This mechanism of inhibition suggests a possible explanation on why the modulation of the electronic properties of the *O*-aryl moiety of carbamic acid aryl esters, through the insertion of electron-donor substituents in para position, has a negligible role on the inhibitory potency on FAAH.[121]

3.1.2. Reversible deacylation vs irreversible decarbamoylation

URB597 have been shown by crystallography and by mass-spectrometry to react with Ser241 to form a stable carbamoylated intermediate.[120] The carbamoylated form of FAAH is significantly more resistant to hydrolysis than the corresponding acylated form (produced by the reaction of the enzyme with oleamide or another substrate) and thus is expected to be responsible for time-dependent and persistent inhibition of FAAH in vivo.

A QM/MM approach based on adiabatic mapping calculations performed at SCC-DTFB/CHARMM27 level with B3LYP correction was used to elucidate and compare the energetics of FAAH deacylation starting from Ser241 acylated by OA *versus* that of FAAH decarbamoylation starting from Ser241 carbamoylated by URB597.[22] These two processes were modelled according to the reaction scheme reported in Figure 4, where deacylation and decarbamoylation where promoted by the action of a "deacylating" water molecule identified in the X-ray structure of FAAH.[120]

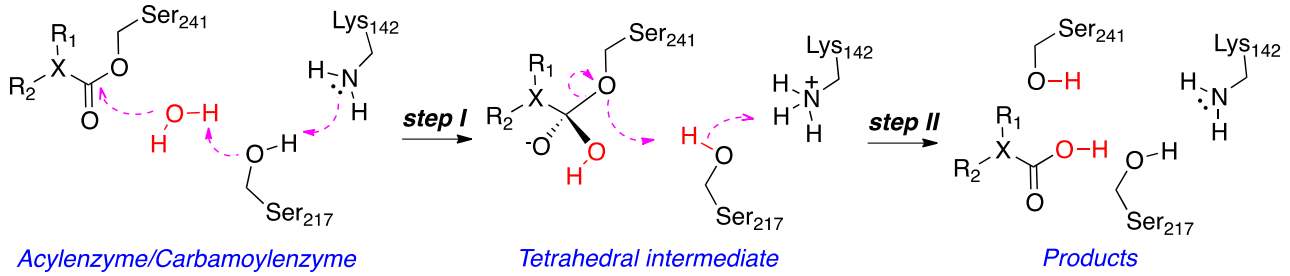

Figure 4. Modelled mechanism of FAAH Ser241 deacylation/decarbamoylation by QM/MM. In the first step of the reaction, the key water molecule was activated by deprotonation

In the first step of the reaction, the key water molecule was activated by deprotonation, promoted by a combined action of Lys142 and Ser217. This event led to the formation of a hydroxide anion which in turn attacked the carbonyl carbon of the acylenzyme/carbamoylenzime leading to an anionic TI. In the second step of the reaction, protonation of Ser241 oxygen operated by Lys142 and Ser217 triggered the collapse of TI structure leading to expulsion of Ser241 side chain. At least in the case of the substrate, this second step was expected to restore the initial configuration of the catalytic triad. QM/MM calculations showed that the step I of FAAH deacylation occurs with a concerted mechanism, in which deprotonation of the water molecule and the nucleophilic attack at the carbonyl carbon were tightly coupled. The transition state (TS1) separating acylenzyme and TI configuration had an energy of nearly 17 kcal/mol above the reactant, while the TI itself was less stable than the reactant by 12.5 kcal/mol. Step II of the deacylation was

characterized by the expulsion of Ser241 side chain from the TI, assisted by protonation reactions performed by a combined action of Ser217 and Lys142. The transition state (TS2) separating TI and reaction products (with formation of free Ser241 and oleic acid) had an energy of 14 kcal/mol, nearly 3 kcal/mol lower than that calculated for the formation of the TI. This suggested that TI formation was the rate-liming event for FAAH deacylation. Overall, reaction barriers at SCC-DFTB/CHARMM27 level are consistent with a fast deacylation and thus with an efficient FAAH catalysis in the presence of oleamide.[112]

Table 1. SCC-DFTB/CHARMM27 potential energy values, expressed in kcal mol^{-1}, for key configurations identified along the deacylation path starting from acylated-(OA) or carbamoylated- FAAH.

	OA	***URB597***
Acyl-Enz/Carb-Enz	0.0	0.0
TS1	16.6	28.3
TI	12.5	27.8
TS2	14.3	28.9
Products	8.2	9.6

When it comes to decarbamoylation of Ser241, significantly higher energy barriers were found either for TS1 or for TS2 configurations. Formation and collapse of TI required 28.3 and 28.9 kcal/mol, respectively. As expected, decarbamoylation of FAAH resulted an unlikely event to happen. In the case of URB597, visual inspection of the TS geometry for TI formation (TI1) indicated that N-cyclohexylcarbamoyl portion connected to Ser241 underwent to a major rearrangement of its geometry during the reaction with the deacylating water molecule (Figure 5, left panel), while for the oleamide, the acyl portion bound to Ser241 experienced a minimal variation of its geometry to form a covalent bond with the solvent (Figure 5, right panel). Active-site decomposition analysis[122] and gas-phase calculations indicates that significant distortion of the TS1 geometry observed in the case of URB594 reduces the electrostatic stabilization provided by FAAH active site during decarbamoylation accounting for the remarkable increase in the barrier for this process.

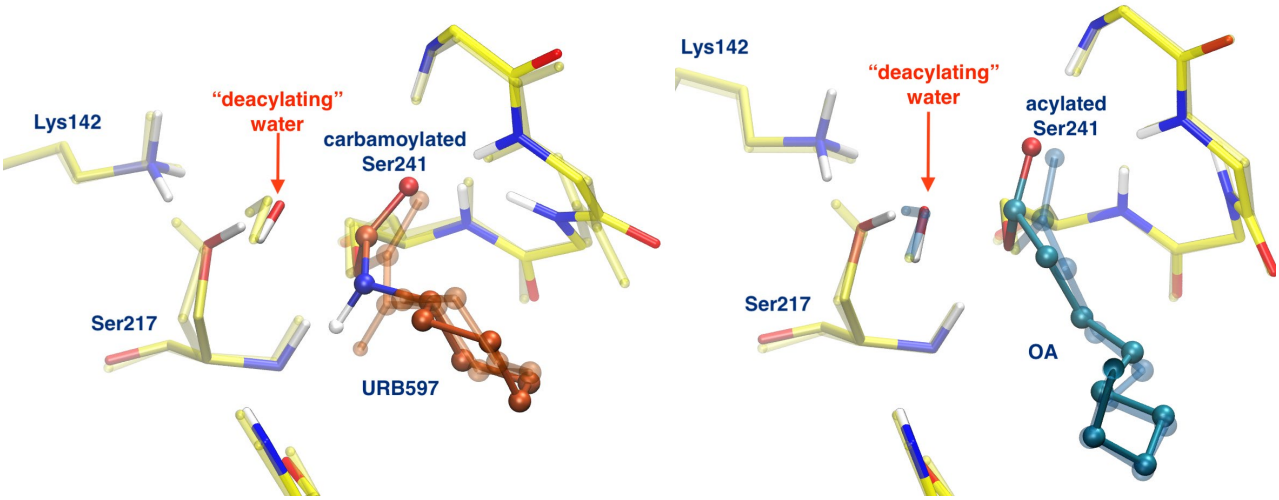

Figure 5. Geometry of TS1 for FAAH decarbamoylation (left panel) and FAAH deacylation (right panel). FAAH carbon atoms are depicted in the yellow, URB597 in orange and oleamide in cyan. The position of the N-cyclohexylcarbamoyl (left) and oleoyl (right) portion in the reactant state is displayed with a shaded ball and stick representing.

To the test validity of the computational approach SCC-DFTB/CHARMM27 calculations were repeated using alternative starting conformations of FAAH covalent adducts. A total of eight FAAH structures extracted from a QM/MM-MD trajectory were used to model deacylation or decarbamoylation at SCC-DFTB/CHARMM27 level. Although some differences were observed in the calculated energies, averaged values (Table 2) confirmed that Ser241 deacylation was significantly favoured over Ser241 decarbamoylation. In light of this finding, this DFTB-based approach could be used to detect new irreversible inhibitors of FAAH acting on Ser241.

Table 2. Average potential energy values (n=8), expressed in kcal mol^{-1}, for key configurations identified along the deacylation path starting from acylated-(OA) or carbamoylated- FAAH.

	OA	***URB597***
Acyl-Enz/Carb-Enz	0.0	0.0
TS1	19.1 ± 0.5	31.7 ± 0.9
TI	16.6 ± 0.9	31.1 ± 0.8
TS2	17.6 ± 0.7	31.0 ± 0.6
Products	8.9 ± 0.5	10.6 ± 0.6

3.2 Epidermal Growth Factor Receptor (EGFR)

3.2.1 Mechanism of Cys797 alkylation by acrylamide inhibitors

Epidermal growth factor receptor (EGFR) is a trans-membranal protein endowed with an extracellular EGF binding domain and a cytoplasmic domain which possesses a tyrosine kinase activity. EGFR activation promoted by EGF leads to receptor dimerization and autophosphorylation which in turn activates a signal transduction cascade promoting cell proliferation.[123] In non-small cell lung cancer (NSCLC), mutations in the kinase domain of EGFR have been observed in nearly the 50% of the cases. Of the known mutations, more than the 90% occurs in exon 19 or in exon 21, the first leading to deletion of the segment E746-A750, the second resulting in arginine replacing leucine at position 858 (L858R).[124] These mutations promote EGFR activation favoring the insurgence of NSCLC. The first-generation of EGFR inhibitors included the 4-anilinoquinazoline gefitinib (Figure 6) which elicited good responses in NSCLS patients.[125] However, most of them acquired drug resistance within 1-year treatment, which in several cases was due to the selection of T790M mutation at the gatekeeper position.[126] The significant variation in the hydrophobic property of the “gatekeeper” residue is believed to be responsible for the reduced potency displayed by gefitinib and other first-generation inhibitors.[127] Second-generation of EGFR inhibitors such as afatinib (Figure 6) demonstrated a good activity against T790M EGFR variants.[128] Thanks to the presence of an acrylamide warhead capable of alkylating Cys797, afatinib and other acrylamide-based analogues were able to circumvent ATP competition and thus to overcome the detrimental effect caused by the presence of a bulky and hydrophobic methionine.

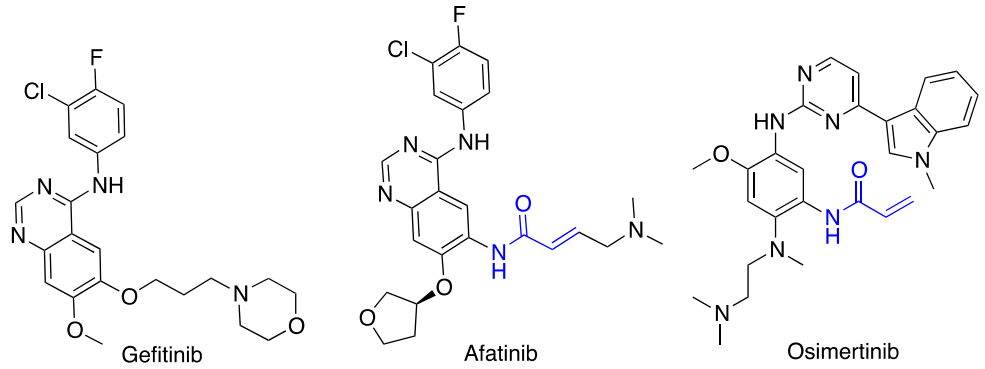

Figure 6. Chemical structure of reference EGFR inhibitors. The acrylamide fragment in afatinib and osimertinib is highlighted in blue.

The mechanism of EGFR wild-type inhibition operated by prototypical *N*-(4-anilinoquinazolin-6-yl)acrylamide has been recently elucidated on the basis of a computational investigation at SCC-DFTB/AMBER level using a PCV approach.[129] QM/MM simulations supported a mechanism (Figure 7) in which an aspartate residue has a critical role in Cys797 alkylation. In details, Asp800 could firstly act as a base deprotonating Cys797-SH and then as an acid protonating the α position of the carbanion resulting from the nucleophilic attack of Cys797-S^- at α position of the acrylamide.

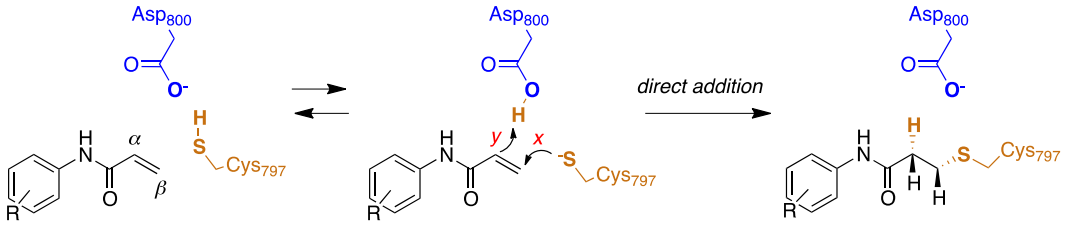

Figure 7. Mechanism of Cys797 alkylation for acrylamide inhibitors of EGFR.

QM/MM simulations indicated that the deprotonation of the thiol group of Cys797 by Asp800 carboxylate took place overcoming a free-energy barrier of 6 kcal/mol suggesting that this reaction was a fast event. Furthermore Cys797-S^-/Asp800-COOH pair was found more stable than Cys797-SH/Asp800-COO^- pair by approximately 4 kcal/mol. This suggested that the predominant form of Cys797 thiol side-chain is anionic. While *ab initio* calculations (i.e., CCSD(T)) in gas-phase indicate that SCC-DFTB underestimates the energetic cost for a cysteine residue to transfer a proton to an aspartate,[130] our SCC-

DFTB/AMBER results were consistent with two key experimental findings reported for EGFR such as, *i.* the rapid oxidation to sulfenic acid for Cys797 SH (which is favored by the presence of a thiolate species)[131]; *ii*. titration experiments indicating that Cys797 thiol has a pKa value approaching that one of a carboxylic acid.[132] We thus proposed that Cys797 participated to the reaction with its thiolate chain acting as a nucleophile. Cys797-S^- approached the electrophile β carbon of the acrylamide fragment to form a new C-S bond. The acrylamide group of the inhibitor maintained a s-*cis* configuration during the entire alkylation process. The nucleophilic attacked occurred in a concerted manner with the protonation of acrylamide α carbon performed by Asp800-COOH. These two events were required to overcome a free-energy barrier of nearly 15 kcal/mol to take place, in agreement with the fast alkylation of EGFR by *N*-(4-anilinoquinazolin-6-yl)acrylamide observed in a fluorescence-based assay.[133]

The product of the reaction was a 3-(alkylsulfanyl)propenamide derivative which resulted significantly more stable than the reactants (~12 kcal/mol) in line with the spontaneous and irreversible alkylation of Cys797.[134,135]

A similar mechanism of Cys797 alkylation has been recently reported for osimertinib,[136] a third-generation EGFR inhibitor featured by a 2-aminopirimidine scaffold (Figure 6). Osimertinib inhibits activated form of EGFR regardless of the presence of T790M mutation, but compared to afatinib, it spares wild-type EGFR thus avoiding severe side-effects in patients observed with second-generation inhibitors.[137]

In the QM/MM study involving Osimertinib, 2D-US sampling simulation at a SCC-DFTB AMBER potential was performed along two intuitive reaction coordinates, describing nucleophilic attack by Cys797-S^- and Cα protonation by Asp800-COOH, respectively. These simulations confirmed the presence of a concerted mechanism in which Cys-S^- attack at the β carbon was tightly coupled with protonation at the α carbon of osimertinib acrylamide by Asp800-COOH leading to the formation of a highly stable alkylation product (Figure 8).

Recent calculations[130] indicated that, in the framework of a Michael-type reaction and in opposition to high level *ab initio* calculations, SCC-DFTB and other popular DFT functionals failed to identify a stable enolate intermediate as a thiolate sulfur approaches the β carbon of a prototypical acrylamide. To account for this drawback, we corrected SCC-DFTB PES at higher level. The resulting MP2/cc-pVTZ//SCC-DFTB/AMBER99SB PES for the Michael addition in EGFR indicated that the *ionic pathway,* corresponding to a stepwise reaction with formation of an enolate intermediate, and the *neutral pathway,* corresponding to the concerted reaction described above, have similar barriers (manuscript in preparation) suggesting that both pathways are accessible within the enzyme.

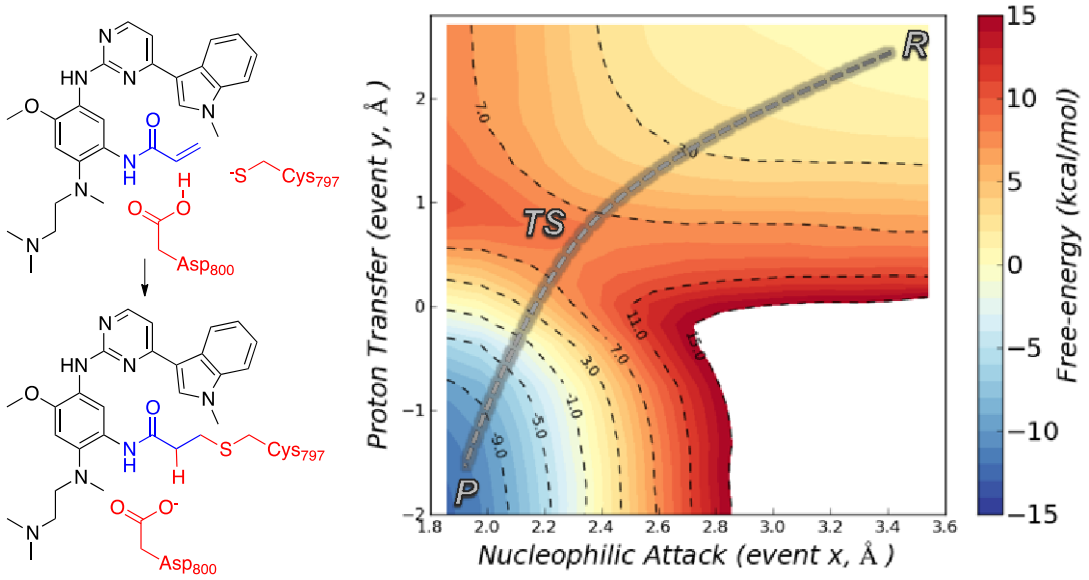

Figure 8. Mechanism of Cys797 alkylation in EGFR T790M by osimertinib.

The presence of a low free-energy path for the reaction of alkylation by osimertinib (activation free-energy 8 kcal/mol, reaction energy -12 kcal/mol) is likely due to the presence of a favorable environment in EGFR able to stabilize Cys797 in anionic form.[132] This suggests that the acrylamide itself could be effectively replaced by less electrophilic warheads thus helping the design of more selective and possibly safer inhibitors of EGFR.

3.2.2. Design of "soft" warheads targeting Cys797 in EGFR

Warheads able to form a covalent bond with a specific cysteine while sparing other nucleophilic residues in cells may help the design of inhibitors featured by high selectivity and reduced side-effects. Acetamides, if conveniently substituted, can meet this requirement possibly alkylating Cys797 by means of a S_N2 reaction (Figure 9) while showing negligible reactivity with free cysteine in solution.[138]

Figure 9. Postulated reaction mechanism for N-(4-(anilino)quinazolin-6-yl)acetamides.

The reactivity of acetamide versus thiols can be modulated working on the leaving group (LG). Reactivity studies with cysteine allowed us to identify (hetero)-aromatic thiols as poor LG substituents, thus suitable for our scope. Once inserted on an acetamide substituent, the resulting compounds were able to give the expected thioether product in presence of cysteine, but at a very low rate i.e., with a half-life time ($t_{1/2}$) > 24 h. In the same experimental conditions, a reference chloroacetamide derivative displayed a $t_{1/2}$ of 1 h.

Thanks to the environment of EGFR, able to stabilize the thiolate form of Cys797, we speculated that (hetero)-aromatic thioacetamides, once installed on a high-affinity driver such as the 4-anilinoquinazoline scaffold, might be able to efficiently alkylate this specific cysteine. To support this hypothesis, we evaluated whether acetamides could undergo attack by Cys797 via a S_N2 reaction pathway using a QM/MM approach. The reaction mechanism was simulated for two 2-(imidazol-2-ylthio)acetamides (UPR1364 and UPR1381 Figure 10) as well as for the highly reactive chloroacetamide UPR1303, employed as control.

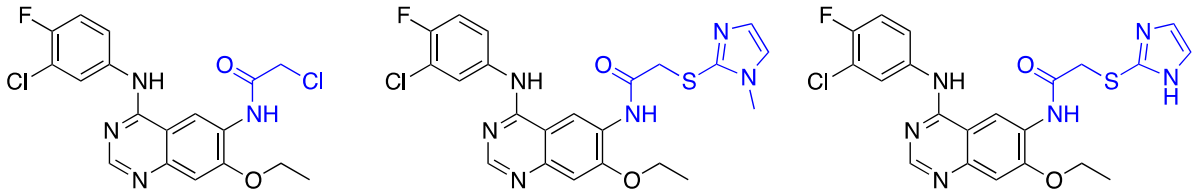

Figure 10. Chemical structure of activated acetamides described here.

The S_N2 reaction was modelled using PDDG/PM3/AMBER potential. The PDDG/PM3 Hamiltonian[139] was selected as it has been reported to give energy barriers for SN_2 reactions in qualitative agreement with *ab initio* calculations.[140] The Cys797/Asp800 pair was modelled in Cys-S-/Asp-COOH form. To simulate the nucleophilic substitution with UPR1364 and UPR1381, we applied a SMD approach applying a moving force along a reaction coordinate defined as the difference between the distance separating the acetamide Cα and the sulfur atom of the imidazol-2-ylthio group and the distance between the acetamide Cα and the sulfur atom of Cys797. In the case of UPR1303, the reaction coordinate was defined as the difference between the distance separating the acetamide Cα and the chlorine atom and the distance between the acetamide Cα and the sulfur atom of Cys797

Preliminary calculations showed that while for UPR1303 an alkylation product with low free-energy could be identified, this was not the case for UPR1364 and UPR1381. We hypothesized that protonation of the imidazole could stabilize the imidazolylthio leaving group. In fact, the carboxylic acid of Asp800 was close enough to the distal nitrogen at the imidazole ring to protonate it. For compounds UPR1364 and UPR1381, Cys797 alkylation was modelled starting from the cationic species resulting from protonation of the nitrogen atom at the imidazole ring, with deprotonated carboxylate group of Asp800 (Figure 11).

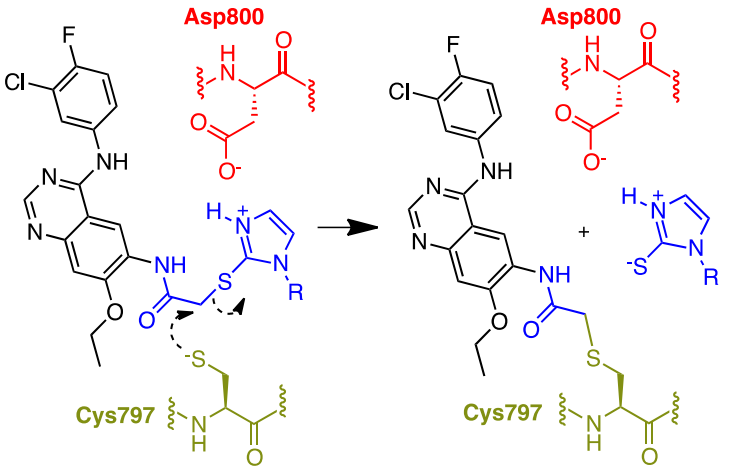

Figure 11. Reactions modelled by QM/MM simulation for UPR1364 and UPR1381 with expulsion of a protonated imidazolylthiolate group. Asp800 was modelled as a carboxylate anion. R stands for -CH_3 (UPR1364) or -H (UPR1381).

Free-energy profiles for Cys797 alkylation were obtained applying the Jarzynski equality using five independent work curves for each compound. The activation energy for Cys797 alkylation, was 27.0 ± 0.7 kcal/mol for UPR1303, 40.4 ± 1.2 kcal/mol for UPR1364 and 34.2 ± 0.7 kcal/mol for compound UPR1381. The reaction energy was negative for UPR1303 and UPR1381 (-12.6 ± 0.7 and -10.2 ± 1.2 kcal/mol, respectively) indicating that Cys797 alkylation by these inhibitors was spontaneous. The reaction energy calculated for UPR1364 was positive (0.9 ± 1.0 kcal/mol) thus indicating that for this inhibitor the reaction was not favored.

Analysis of the geometries along the SMD trajectory indicated that the TS structure for alkylation of Cys797 by UPR1303 and UPR1381 was characterized by a trigonal bipyramidal geometry with the thiolate sulfur, the acetamide Cα and the LG heteroatom forming an angle of nearly 180° (Figure 12). This was not the case for UPR1364 where the close proximity between the N-methyl group emerging from the imidazole ring to the p-loop of EGFR hindered the geometry of the TS.

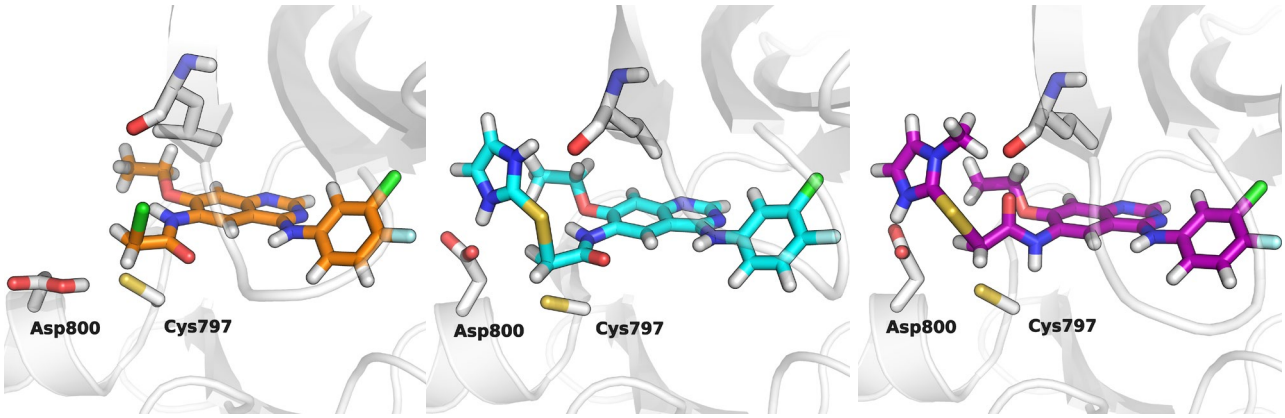

Figure 12. Representative TS structures for Cys797 alkylation by UPR1303 (left panel, orange carbon atoms), UPR1381 (central panel B, cyan carbon atoms) and UPR1364 (right panel, magenta carbon atoms).

QM/MM SMD simulations suggest that UPR1303 and UPR1381 can alkylate Cys797 leading to stable adducts. It is worth mentioning that these simulations were performed before the synthesis of the titled compounds. Computational results were essential for the decision to synthesize and test in biological assays UPR1381.

The ability of UPR1381 to inhibit EGFR L858R/T790M double mutant was assessed using a TR-FRET binding assay with and without pre-incubation. When a covalent modification occurs at an enzyme binding site, the detected IC_{50} value turns out to be dependent on the time of pre-incubation.[141] The IC_{50} value of UPR1381 displayed a significant reduction with 5-h preincubation, supporting a covalent mechanism of inhibition for this compound. In accordance with this finding, UPR1381 displayed long-lasting EGFR inhibition in A549 cells and inhibited the growth on H1975 cell line, which harbors L858R/T790M mutation, with a potency (IC_{50} = 1.4 μM) comparable to that of a reference acrylamide-based inhibitor tested in the same condition (IC_{50} = 0.70 μM).

Overall, our findings suggest that heteroaryl thioacetamide warheads can be used as alternative to acrylamides provided that their lower reactivity is counterbalanced by the presence of an appropriate scaffold able to place them in close proximity to the nucleophilic cysteine.

Conclusions

Recent advances in software and hardware development is making possible the use of the QM/MM approach to solve problems relevant for medicinal chemistry and chemical biology. Different research groups worldwide are currently trying to exploit results gathered from QM/MM simulations to design new covalent inhibitors featured by higher potency, and improved selectivity. Efforts are currently being taken to use calculations to predict the stability of the reaction adduct and thus control inhibitor k_{off} also when the formation of covalent bond is part of the inhibitory process.[25] A mechanistic understanding of the chemical reaction involving a drug target and a covalent modifier can suggest modifications that hardly other computational techniques would provide.

Here, we have reported few representative examples of QM/MM studies performed on biological targets relevant for our drug discovery programs. In the case of FAAH, we showed the ability of QM/MM simulations to *i.* predict the binding orientation of a reference inhibitor, *ii*. rationalize substituent effect on activity data, and *iii.* characterize the mechanism of enzyme reactivation discriminating a substrate from a covalent irreversible inhibitor. In the case of EGFR, QM/MM simulations allowed to *i*. characterize at atomic level the mechanism of inhibition of acrylamide-based inhibitors, *ii.* highlight the uncommon nucleophilicity possessed by the non-catalytic Cys797, and *iii.* assist the prospective design of novel acetamides warheads featured by low chemical reactivity. In our experience, QM/MM simulations can provide essential mechanistic answers that can help covalent drug design. We also believe that QM/MM methodology has the potentiality for becoming a standard and reference tool for computational chemists operating in field of medicinal chemistry and chemical biology. In these fields, one of the current topics is represented by the discovery of inhibitor targeting residues distinct from cysteine or serine[142] Again, we believe that this kind of research might benefit from a wise use of QM/MM simulations.

Acknowledgements

Adrian Mulholland, (University of Bristol, Bristol, UK), Christo Christov (Michigan Tech University, Houghton, Michigan, US) and Davide Branduardi (Schrodinger Inc, London, UK) are kindly acknowledged for insightful discussions.

Reference

[1] Copeland RA (2016) The Drug-Target Residence Time Model: a 10-year Retrospective. *Nat Rev Drug Discov*. 15:87-95.

[2] Tonge P (2018) Drug-Target Kinetics in Drug Discovery. *ACS Chem. Neurosci* 9: 29−39

[3] Nunez S, Venhorst J, Kruse CG (2012) Target-Drug Interactions: First principles and their Application to Drug Discovery. *Drug Discovery Today* 17: 10−22.

[4] Copeland RA, Pompliano DL, Meek TD (2006) Drug-target residence time and its implications for lead optimization. *Nat. Rev. Drug Discovery* 5: 730−739.

[5] Vauquelin G, Charlton, S J (2010) Long-lasting Target Binding and Rebinding as Mechanisms To Prolong in Vivo Drug Action. *Br. J. Pharmacol* 161: 488−508.

[6] Bauer RA (2015) Covalent Inhibitors in Drug Discovery: From Accidental Discoveries to Avoided Liabilities and Designed Therapies. *Drug Discov Today* 20:1061-1073.

[7] Miyahisa I, Sameshima T, Hixon MS (2015) Rapid Determination of the Specificity Constant of Irreversible Inhibitors (kinact/KI) by Means of an Endpoint Competition Assay. *Angew Chem Int Ed Engl.* 54:14099-14102.

[8] Strelow JM (2017) A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discov.* 22:3-20.

[9] Baillie TA (2016) Targeted Covalent Inhibitors for Drug Design. *Angew Chem Int Ed Engl*. 55:13408-13421

[10] Carmi C, Lodola A, Rivara S, Vacondio F, Cavazzoni A, Alfieri RR, Ardizzoni A, Petronini PG, Mor M (2011) Epidermal Growth Factor Receptor Irreversible Inhibitors: Chemical Exploration of the Cysteine-Trap Portion. *Mini Rev Med Chem*. 11:1019-1030.

[11] Singh J, Petter RC, Baillie TA, Whitty A (2011) The resurgence of covalent drugs. *Nat Rev Drug Discov* 10:307-317

[12]Jorgensen WL (2004) The Many Roles of Computation in Drug Discovery. *Science* 303:1813-1818.

[13]Jorgensen WL (2009) Efficient Drug Dead Discovery and Optimization. *Acc Chem Res* 42:724-33.

[14] London N, Miller RM, Krishnan S, Uchida K, Irwin JJ, Eidam O, Gibold L, Cimermančič P, Bonnet R, Shoichet BK, Taunton J (2014) Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol* 10:1066-72.

[15] Lonsdale R, Ward RA (2018) Structure-based design of targeted covalent inhibitors. *Chem Soc Rev.* 47:3816-3830.

[16] De Cesco S, Kurian J, Dufresne C, Mittermaier AK, Moitessier N (2017) Covalent inhibitors design and discovery. *Eur J Med Chem* 138:96-114.

[17] Schmidt TC, Welker A, Rieger M, Sahu PK, Sotriffer CA, Schirmeister T, Engels B (2014) Protocol for Rational Design of Covalently Interacting Inhibitors. *ChemPhysChem* 15:3226-3235.

[18] Zhang H, Jiang W, Chatterjee P, Luo Y (2019) Ranking Reversible Covalent Drugs: From Free Energy Perturbation to Fragment Docking. *J Chem Inf Model.* 59:2093-2102.

[19] Scarpino A, Ferenczy GG, Keserű GM (2018) Comparative Evaluation of Covalent Docking Tools. *J Chem Inf Model* 58:1441-1458.

[20] Awoonor-Williams E, Walsh AG, Rowley CN (2017) Modeling Covalent-Modifier Drugs *Biochim. Biophys. Acta* 1865:1664–1675.

[21] Van der Kamp MW, Mulholland AJ (2013) Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry*. 52:2708-2728.

[22] Lodola A, Capoferri L, Rivara S, Tarzia G, Piomelli D, Mulholland A, Mor M (2013) Quantum Mechanics/Molecular Mechanics Modeling of Fatty Acid Amide Hydrolase Reactivation Distinguishes Substrate from Irreversible Covalent Inhibitors. *J Med Chem.* 56:2500-2512.

[23] Chudyk EI, Limb MA, Jones C, Spencer J, van der Kamp MW, Mulholland AJ (2014) QM/MM simulations as an assay for carbapenemase activity in class A β-lactamases. *Chem Commun (Camb)* 50:14736-14739.
[24] Sgrignani J, Grazioso G, De Amici M, Colombo G (2014). Inactivation of TEM-1 by Avibactam (NXL-104): Insights from Quantum Mechanics/Molecular Mechanics Metadynamics Simulations. *Biochemistry* 53:5174-5185.
[25] Schirmeister T, Kesselring J, Jung S, Schneider TH, Weickert A, Becker J, Lee W, Bamberger D, Wich PR, Distler U, Tenzer S, Johé P, Hellmich UA, Engels B (2016). Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J Am Chem Soc* 138:8332-8335
[26] Arafet K, Ferrer S, González FV, Moliner V (2017) Quantum Mechanics/Molecular Mechanics Studies of the Mechanism of Cysteine Protease Inhibition by Peptidyl-2,3-epoxyketones. *Phys Chem Chem Phys* 19:12740-12748.
[27] Nutho B, Mulholland AJ, Rungrotmongkol T (2019) The Reaction Mechanism of Zika Virus NS2B/NS3 Serine Protease Inhibition by Dipeptidyl Aldehyde: a QM/MM study. *Phys Chem Chem Phys*, in press
[28] Lodola A, De Vivo M (2012) The Increasing Role of QM/MM in Drug Discovery. *Adv Protein Chem Struct Biol.* 87:337-362
[29] De Vivo M (2011) Bridging Quantum Mechanics and Structure-Based Drug Design. *Front Biosci* 16:1619-1633.
[30] Di Marzo V, Fontana A, Cadas H, Schinelli S, Cimino G, Schwartz JC, Piomelli D (1994) Formation and Inactivation of Endogenous Cannabinoid Anandamide in Central Neurons. *Nature* 372:686-691.
[31] Fegley D, Gaetani S, Duranti A, Tontini A, Mor M, Tarzia G, Piomelli D (2005) Characterization of the Fatty Acid Amide Hydrolase Inhibitor Cyclohexyl Carbamic Acid 3'-Carbamoyl-Biphenyl-3-yl Ester (URB597): Effects on Anandamide and Oleoylethanolamide deactivation. *J Pharmacol Exp Ther* 313:352-358.
[32] Roskoski R Jr (2014) The ErbB/HER family of Protein-Tyrosine Kinases and Cancer. *Pharmacol Res.* 79:34-74
[33] Świderek K, Tuñón I, Moliner V, Bertran J (2015) Computational Strategies for the Design of New Enzymatic Functions. *Arch Biochem Biophys*. 582:68-79.
[34] Amaro RE, Mulholland AJ (2018) Multiscale Methods in Drug Design Bridge Chemical and Biological Complexity in the Search for Cures. *Nat Rev Chem* 2: 0148
[35] Warshel A, Weiss RM (1980) An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J Am Chem Soc* 102:6218–6226.
[36] Car R, Parrinello M (1985) Unified approach for molecular dynamics and density-functional theory *Phys Rev Lett* 55:2471–2474.
[37] Kamerlin SCL, Warshel A (2011) The Empirical Valence Bond Model: Theory and Applications. *Wiley Interdiscip Rev: Comput Mol Sci* 1:30–45.
[38] Dal Peraro M, Ruggerone P, Raugei S, Gervasio FL, Carloni P (2007) Investigating Biological Systems Using First Principles Car-Parrinello Molecular Dynamics Simulations. *Curr Opin Struct Biol* 17:149-156.
[39] Warshel A, Levitt M (1976) Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J Mol Biol* 103:227-249.
[40] Field MJ, Bash PA, Karplus M (1990) A Combined Quantum-Mechanical and Molecular Mechanical Potential for Molecular-Dynamics Simulations*. J Comput Chem* 11:700-733.
[41] Mulholland AJ (2005) Modelling Enzyme Reaction Mechanisms, Specificity and Catalysis. *Drug Discov Today* 10:1393-1402.
[42] Ryde U. (2016) QM/MM Calculations on Proteins. *Methods Enzymol* 577:119-158.

[43] Warshel A (2003) Computer Simulations of Enzyme Catalysis: Methods, Progress, and Insights. *Annu Rev Biophys Biomol Struct*. 32:425-443.
[44] Ainsley J, Lodola A, Mulholland AJ, Christov CZ, Karabencheva-Christova TG. (2018) Combined Quantum Mechanics and Molecular Mechanics Studies of Enzymatic Reaction Mechanisms. *Adv Protein Chem Struct Biol.* 113:1-32.
[45] Claeyssens F, Harvey JN, Manby FR, Mata RA, Mulholland AJ, Ranaghan KE, Schutz M, Thiel S, Thiel W, Werner HJ (2006) High-accuracy Computation of Reaction Barriers in Enzymes. *Angew Chem Int Ed Engl* 45:6856-6859.
[46] Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *J Phys Rev B* 58:7260-7268.
[47] Gaus M, Cui Q, Elstner M (2012) DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J Chem Theory Comput* 7:931-948.
[48] Kromann JC, Christensen AS, Cui Q, Jensen JH (2016). Towards a Barrier Height Benchmark Set for Biologically Relevant Systems. *Peer J*. 4:e1994.
[49] Christensen AS, Kubař T, Cui Q, Elstner M (2016) Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem Rev* 116:5301-5337.
[50] Gaus M, Goez A, Elstner M (2013) Parametrization and Benchmark of DFTB3 for Organic Molecules. *J Chem Theory Comput* 9:338-354
[51] Domínguez A, Niehaus TA, Frauenheim T (2015) Accurate Hydrogen Bond Energies within the Density Functional Tight Binding Method. *J Phys Chem A.* 119:3535-3544.
[52] Miriyala VM, Řezáč J (2017) Description of Non-Covalent Interactions in SCC-DFTB Methods. *J Comput Chem* 38:688-697.
[53] Gruden M, Andjeklović L, Jissy AK, Stepanović S, Zlatar M, Cui Q, Elstner M (2017) Benchmarking Density Functional Tight Binding Models for Barrier Heights and Reaction Energetics of Organic Molecules. *J Comput Chem* 38:2171-2185.
[54] Elstner M (2006) The SCC-DFTB Method and its Application to Biological Systems. *Theor Chem Acc* 116:316-325.
[55] Capoferri L, Mor M, Sirirak J, Chudyk E, Mulholland AJ, Lodola A (2011) Application of a SCC-DFTB QM/MM Approach to the Investigation of the Catalytic Mechanism of Fatty Acid Amide Hydrolase. *J Mol Model* 17:2375-2383.
[56] Jitonnom J, Limb MA, Mulholland AJ (2014) QM/MM Free-Energy Simulations of Reaction in Serratia Marcescens Chitinase B Reveal the Protonation state of Asp142 and the Critical Role of Tyr214. J Phys Chem B. 118:4771-4483.
[57] Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the Biomolecular Simulation Program. *J Comp Chem* 30:1545-1614.
[58] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* 65:712-725.
[59] De Vivo M, Masetti M, Bottegoni G, Cavalli A (2016) Role of Molecular Dynamics and Related Methods in Drug Discovery. *J Med Chem.* 59:4035-4061.
[60] Senn HM, Thiel W (2009) QM/MM Methods for Biomolecular Systems. *Angew Chem Int Ed Engl* 48:1198-1229.
[61] Boulanger E, Harvey JN (2018) QM/MM Methods for Free Energies and Photochemistry. *Curr Opin Struct Biol.* 49:72-76.

[62] Ganguly A, Boulanger E, Thiel W (2017) Importance of MM Polarization in QM/MM studies of Enzymatic Reactions: Assessment of the QM/MM Drude Oscillator Model. *J Chem Theory Comput* 13:2954-2961.
[63] Ridder L, Rietjens IMCM, Vervoort J, Mulholland AJ (2002) Quantum Mechanical/ Molecular Mechanical Free Energy Simulations of the Glutathione S-transferase (M1-1) Reaction with Phenanthrene 9,10-oxide. *J Am Chem Soc* 124:9926–9936.
[64] Riccardi D, Li GH, Cui Q (2004) Importance of Van der Waals Interactions in QM/MM Simulations. *J Phys Chem B* 108:6467–6478.
[65] Himo F (2006) Quantum Chemical Modeling of Enzyme Active Sites and Reaction Mechanisms. *Theor Chem Acc* 116:232–240.
[66] Marti S, Roca M, Andres J, Moliner V, Silla E, Tunon I, Bertran J (2004) Theoretical Insights in Enzyme Catalysis. *Chem Soc Rev* 33:98–107.
[67] Lonsdale R., Harvey, JN, Mulholland, AJ (2012). A Practical Guide to Modelling Enzyme-Catalysed Reactions. *Chem Soc Rev* 41: 3025–3038.
[68] Sousa SF, Ribeiro AJM, Neves RPP, Brás NF, Cerqueira NMFSA, Fernandes PA, Ramos MJ (2017) Application of Quantum Mechanics/Molecular Mechanics Methods in the Study of Enzymatic Reaction Mechanisms. *Wiley Interdiscip Rev Comput Mol Sci 7: e1281.*
[69] Cerqueira NMFSA, Fernandes PA, Ramos MJ (2018) Protocol for Computational Enzymatic Reactivity Based on Geometry Optimisation. *ChemPhysChem.* 19:669-689.
[70] Lodola A, Mulholland AJ (2013) Computational Enzymology. *Methods Mol Biol* 924:67-89.
[71] Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On Possible Pitfalls in Ab Initio Quantum Mechanics/Molecular Mechanics Minimization Approaches for Studies of Enzymatic Reactions*. J Phys Chem B* 109:15645-15650.
[72] Christov CZ, Lodola A, Karabencheva-Christova TG, Wan S, Coveney PV, Mulholland AJ. (2013) Conformational Effects on the Pro-S Hydrogen Abstraction Reaction in Cyclooxygenase-1: an Integrated QM/MM and MD Study. *Biophys J* 104:L5-L7
[73] Ryde U. (2017) How Many Conformations Need To Be Sampled To Obtain Converged QM/MM Energies? The Curse of Exponential Averaging. *J Chem Theory Comput*. 13:5745-5752.
[74] Lodola A, Sirirak J, Fey N, Rivara S, Mor M, Mulholland AJ (2010) Structural Fluctuations in Enzyme-Catalysed Reactions: Determinants of Reactivity in Fatty Acid Amide Hydrolase from Multivariate Statistical Analysis of Quantum Mechanics/Molecular Mechanics Paths. *J Chem Theory Comput* 6:2948-2960.
[75] Mulholland AJ (2001) The QM/MM Approach to Enzymatic Reactions. *Theor Comp Chem* 9:597-653.
[76] Acevedo O, Jorgensen WL (2010) Advances in Quantum and Molecular Mechanical (QM/MM) Simulations for Organic and Enzymatic reactions. *Acc Chem Res* 43:142-151.
[77] Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MH (2006) Electrostatic basis for enzyme catalysis. *Chem Rev* 106:3210-3235.
[78] Glowacki DR, Harvey JN, Mulholland AJ (2012) Taking Ockham's Razor to Enzyme Dynamics and Catalysis. *Nat Chem* 4:169-176.
[79] Swiderek K, Tunon I, Moliner V (2014) Predicting Enzymatic Reactivity: from Theory to Design. *Wiley Interdiscip Rev Comput Mol Sci* 4:407–421.
[80] Acevedo O, Jorgensen WL (2014) Quantum and Molecular Mechanical (QM/MM) Monte Carlo Techniques for Modeling Condensed-Phase Reactions. *Wiley Interdiscip Rev Comput Mol Sci* 4:422-435.
[81] Roux B (1995) The Calculation of the Potential of Mean Force Using Computer Simulations. *Comp Phys Comm* 91: 275-282.
[82] Gao J, Ma S, Major DT, Nam K, Pu J, Truhlar DG (2006) Mechanisms and Free Energies of Enzymatic Reactions. *Chem Rev* 106:3188-3209.

[83] Kaestner J (2011) Umbrella Sampling. *Wiley Interdiscip Rev Comput Mol Sci* 1:932–942.
[84] Kumar S, Rosenberg JM, Bouzida D, Swendsen R. H, Kollman PA (1992) The Weighted Histogram Analysis Method for Free- Energy Calculations on Biomolecules. I. The Method. *J Comput Chem* 13:1011−1021.
[85] Izrailev, S., S. Stepaniants, M. Balsera, Y. Oono, K. Schulten (1997) Molecular Dynamics Study of Unbinding of the Avidin-Biotin complex. *Biophys. J.* 72:1568–1581.
[86] Murcia M, Jirouskova M, Li J, Coller BS, Filizola M. (2008) Functional and Computational Studies of the Ligand-Associated Metal Binding Site of Beta3 Integrins. *Proteins* 71:1779-1791.
[87] Crespo A, Martí MA, Estrin DA, Roitberg AE (2005) Multiple-Steering QM-MM Calculation of the Free Energy Profile in Chorismate Mutase. *J Am Chem Soc* 127, 6940-6941.
[88] Jarzynski C (1997) Non-equilibrium Equality for Free Energy Differences. *Phys Rev Lett* 78:2690–2693.
[89] Ramírez CL, Martí MA, Roitberg AE (2016) Steered Molecular Dynamics Methods Applied to Enzyme Mechanism and Energetics. *Methods Enzymol* 578:123-143.
[90] Soto-Delgado J, Tapia RA, Torras J (2016) Multiscale Treatment for the Molecular Mechanism of a Diels-Alder Reaction in Solution: A QM/MM-MD Study. *J Chem Theory Comput* 12:4735-4742
[91] Karabencheva-Christova TG, Torras J, Mulholland AJ, Lodola A, Christov CZ. (2017) Mechanistic Insights into the Reaction of Chlorination of Tryptophan Catalyzed by Tryptophan 7-Halogenase. *Sci Rep* 7:17395.
[92] Laio A, Parrinello M. Escaping Free-energy Minima (2002) *Proc Natl Acad Sci USA* 99:12562-125666.
[93] Gervasio FL, Laio A, Parrinello M (2005) Flexible Docking in Solution Using Metadynamics. *J Am Chem Soc.* 127:2600-2607.
[94] Russo S, Callegari D, Incerti M, Pala D, Giorgio C, Brunetti J, Bracci L, Vicini P, Barocelli E, Capoferri L, Rivara S, Tognolini M, Mor M, Lodola A (2016) Exploiting Free-Energy Minima to Design Novel EphA2 Protein-Protein Antagonists: From Simulation to Experiment and Return. *Chemistry* 22:8048-8052.
[95] Scalvini L, Vacondio F, Bassi M, Pala D, Lodola A, Rivara S, Jung KM, Piomelli D, Mor M (2016) Free-energy Studies Reveal a Possible Mechanism for Oxidation-Dependent Inhibition of MGL. *Sci Rep.* 6:31046.
[96] Cavalli A, Spitaleri A, Saladino G, Gervasio FL (2015) Investigating Drug-Target Association and Dissociation Mechanisms Using Metadynamics-based algorithms. *Acc Chem Res* 48:277-285.
[97] Callegari D, Lodola A, Pala D, Rivara S, Mor M, Rizzi A, Capelli AM (2017) Metadynamics Simulations Distinguish Short- and Long-Residence-Time Inhibitors of Cyclin-Dependent Kinase 8 *J Chem Inf Model.* 57:159-169.
[98] Ibrahim P, Clark T (2019) Metadynamics Simulations of Ligand Binding to GPCRs. *Curr Opin Struct Biol.* 55:129-137.
[99] Ensing B, De Vivo M, Liu Z, Moore P, Klein ML (2006) Metadynamics as a Tool for Exploring Free Energy Landscapes of Chemical Reactions. *Acc Chem Res.* 39:73-81
[100] Raich L, Nin-Hill A, Ardèvol A, Rovira C (2016) Enzymatic Cleavage of Glycosidic Bonds: Strategies on How to Set Up and Control a QM/MM Metadynamics Simulation *Methods Enzymol*. 577:159-83.
[101] Branduardi D, Gervasio FL, Parrinello M (2007). From A to B in Free Energy Space. *J Chem Phys* 126:054103.
[102] Zinovjev K, Tuñón I (2014) Exploring Chemical Reactivity of Complex Systems with Path-Based Coordinates: Role of the Distance Metric. *J Comput Chem* 35:1672-81.
[103] Maragliano L, Vanden-Eijnden, EJ (2008) Single-Sweep Methods for Free Energy Calculations. *J Chem Phys* 128:184110.

[104] Branduardi D, De Vivo M, Rega N, Barone V, Cavalli A (2011) Methylphosphate Dianion Hydrolysis in Solution Characterized by Path Collective Variables Coupled with DFT-Based Enhanced Sampling Simulations. *J Chem Theory Comput* 7:539-543.
[105] Lodola A, Branduardi D, De Vivo M, Capoferri L, Mor M, Piomelli D, Cavalli, A. (2012) A Catalytic Mechanism for Cysteine N-terminal Nucleophile Hydrolases, as Revealed by Free Energy Simulations. *PLoS One,* 7, e32397.
[106] Murillo-López J, Zinovjev K, Pereira H, Caniuguir A, Garratt R, Babul J, Recabarren R, Alzate-Morales J, Caballero J, Tuñón I, Cabrera R. (2019) Studying the Phosphoryl Transfer Mechanism of the E. coli Phosphofructokinase-2: from X-ray Structure to Quantum Mechanics/molecular Mechanics Simulations. *Chem Sci* 10:2882-2892.
[107] Lodola A, Castelli R, Mor M, Rivara S. (2015) Fatty Acid Amide Hydrolase Inhibitors: a Patent Review (2009-2014). *Expert Opin Ther Pat* 25:1247-1266.
[108] McKinney MK, Cravatt BF (2003) Evidence for Distinct Roles in Catalysis for Residues of the Serine-Serine-Lysine Catalytic Triad of Fatty Acid Amide Hydrolase. *J Biol Chem* 278:37393-37399.
[109] Lodola A, Mor M, Sirirak J, Mulholland AJ (2009) Insights Into the Mechanism and Inhibition of Fatty Acid Amide Hydrolase from Quantum Mechanics/Molecular Mechanics (QM/MM) Modelling. *Biochem Soc Trans 37*(Pt 2):363-367.
[110] Lodola A, Mor M, Hermann JC, Tarzia G, Piomelli D, Mulholland AJ (2005) QM/MM Modelling of Oleamide Hydrolysis in Fatty Acid Amide Hydrolase (FAAH) reveals a new Mechanism of Nucleophile Activation. *Chem Commun (Camb)* 4399-4401.
[111] Tubert-Brohman I, Acevedo O, Jorgensen WL (2006) Elucidation of Hydrolysis Mechanisms for Fatty Acid Amide Hydrolase and its Lys142Ala Variant via QM/MM Simulations. *J Am Chem Soc* 128:16904-16913.
[112] McKinney MK, Cravatt BF (2005) Structure and Function of Fatty Acid Amide Hydrolase. *Annu Rev Biochem* 74:411-432.
[113] Palermo G, Campomanes P, Cavalli A, Rothlisberger U, De Vivo M (2015) Anandamide Hydrolysis in FAAH Reveals a Dual Strategy for Efficient Enzyme-Assisted Amide Bond Cleavage via Nitrogen Inversion. *J Phys Chem B* 119:789-801
[114] Mor M, Rivara S, Lodola A, Plazzi PV, Tarzia G, Duranti A, Tontini A, Piersanti G, Kathuria S, Piomelli D (2004) Cyclohexylcarbamic Acid 3'- or 4'-substituted Biphenyl-3-yl Esters as Fatty Acid Amide Hydrolase Inhibitors: Synthesis, Quantitative Structure-Activity Relationships, and Molecular Modeling Studies. *J Med Chem* 47:4998-5008.
[115] Alexander JP, Cravatt BF (2005) Mechanism of Carbamate Inactivation of FAAH: Implications for the Design of Covalent Inhibitors and In Vivo Functional Probes for Enzymes. *Chem Biol* 12:1179-1187.
[116] Lodola A, Rivara S, Mor M (2011) Application of Computational Methods to the Design of Fatty Acid Amide Hydrolase (FAAH) Inhibitors Based on a Carbamic Template Structure. *Adv Protein Chem Struct Biol* 85:1-26.
[117] Bracey MH, Hanson MA, Masuda KR, Stevens RC, Cravatt BF (2002) Structural Adaptations in a Membrane Enzyme that Terminates Endocannabinoid Signaling. *Science* 298:1793-1796.
[118] Basso E, Duranti A, Mor M, Piomelli D, Tontini A, Tarzia G, Traldi P (2004) Tandem Mass Spectrometric Data-FAAH Inhibitory Activity Relationships of Some Carbamic Acid O-Aryl Esters. *J Mass Spectrom* 39:1450-1455.
[119] Lodola A, Mor M, Rivara S, Christov C, Tarzia G, Piomelli D, Mulholland AJ (2008) Identification of Productive Inhibitor Binding Orientation in Fatty Acid Amide Hydrolase (FAAH) by QM/MM Mechanistic Modelling. *Chem Commun (Camb)* 214-216.
[120] Mileni M, Kamtekar S, Wood DC, Benson TE, Cravatt BF, Stevens RC (2010) Crystal Structure of Fatty Acid Amide Hydrolase Bound to the Carbamate Inhibitor URB597:

Discovery of a Deacylating Water Molecule and Insight into Enzyme Inactivation. *J Mol Biol* 400:743-754.
[121] Lodola A, Capoferri L, Rivara S, Chudyk E, Sirirak J, Dyguda-Kazimierowicz E, Andrzej Sokalski W, Mileni M, Tarzia G, Piomelli D, Mor M, Mulholland AJ (2011) Understanding the Role of Carbamate Reactivity in Fatty Acid Amide Hydrolase Inhibition by QM/MM Mechanistic Modelling. *Chem Commun (Camb)* 47:2517-2519.
[122] Mulholland AJ, Richards G (1997) Acetyl-CoA Enolization in Citrate Synthase: a Quantum Mechanical/Molecular Mechanical (QM/MM) Study. *Proteins*, 27, 9−25.
[123] Roskoski R Jr (2014) The ErbB/HER Family of Protein-Tyrosine Kinases and Cancer. *Pharmacol Res* 79:34-74.
[124] Sharma SV, Bell DW, Settleman J, Haber DA (2007) Epidermal Growth Factor Receptor Mutations in Lung Cancer. *Nat Rev Cancer* 7:169−181.
[125] Pao W, Chmielecki J (2010) Rational, Biologically Based Treatment of EGFR-Mutant Non-Small-Cell Lung Cancer. *Nat Rev Cancer* 10:760−774.
[126] Engelman JA, Janne PA (2008) Mechanisms of Acquired Resistance to Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. *Clin Cancer Res* 14:2895-2899.
[127] Michalczyk A, Klüter S, Rode HB, Simard JR, Grütter C, Rabiller M, Rauh D (2008) Structural Insights into How Irreversible Inhibitors Can Overcome Drug Resistance in EGFR. *Bioorg Med Chem* 16:3482-3488.
[128] Hirsh V (2011) Afatinib (BIBW 2992) Development in Non-Small-Cell Lung Cancer. *Future Oncol.* 7:817-825.
[129] Capoferri L, Lodola A, Rivara S, Mor M (2015) Quantum Mechanics/Molecular Mechanics Modeling of Covalent Addition between EGFR-cysteine 797 and N-(4-anilinoquinazolin-6-yl) acrylamide. *J Chem Inf Model* 55:589-599.
[130] Awoonor-Williams E, Isley W, Dale S, Johnson E, Yu H, Becke A, Roux B, Rowley C (2019) Quantum Chemical Methods for Modeling Covalent Modification of Biological Thiols. *ChemRxiv*, doi: 10.26434/chemrxiv.8061845.v1
[131] Truong TH, Carroll KS (2012) Redox Regulation of Epidermal Growth Factor Receptor Signaling Through Cysteine Oxidation. *Biochemistry* 51:9954-9965.
[132] Truong TH, Ung PM, Palde PB, Paulsen CE, Schlessinger A, Carroll KS (2016) Molecular Basis for Redox Activation of Epidermal Growth Factor Receptor Kinase. *Cell Chem Biol* 23:837-848.
[133] Klüter S, Simard JR, Rode HB, Grütter C, Pawar V, Raaijmakers HC, Barf TA, Rabiller M, van Otterlo WA, Rauh D (2010) Characterization of Irreversible Kinase Inhibitors by Directly Detecting Covalent Bond Formation: a Tool for Dissecting Kinase Drug Resistance. *Chembiochem* 11:2557-2566.
[134] Schwartz PA, Kuzmic P, Solowiej J, Bergqvist S, Bolanos B, Almaden C, Nagata A, Ryan K, Feng J, Dalvie D, Kath JC, Xu M, Wani R, Murray BW (2014) Covalent EGFR Inhibitor Analysis Reveals Importance of Reversible Interactions to Potency and Mechanisms of Drug Resistance. *Proc Natl Acad Sci USA* 111:173-178.
[135] Carmi C, Galvani E, Vacondio F, Rivara S, Lodola A, Russo S, Aiello S, Bordi F, Costantino G, Cavazzoni A, Alfieri RR, Ardizzoni A, Petronini PG, Mor M (2012) Irreversible Inhibition of Epidermal Growth Factor Receptor Activity by 3-aminopropanamides. *J Med Chem* 55:2251-2264.
[136] Callegari D, Ranaghan KE, Woods CJ, Minari R, Tiseo M, Mor M, Mulholland AJ, Lodola A (2018) L718Q Mutant EGFR Escapes Covalent Inhibition by Stabilizing a Non-reactive Conformation of the Lung Cancer Drug Osimertinib. *Chem Sci* 9:2740-2749.
[137] Cross DA, Ashton SE, Ghiorghiu S, Eberlein C, Nebhan CA, Spitzler PJ, Orme JP, Finlay MR, Ward RA, Mellor MJ, Hughes G, Rahi A, Jacobs VN, Red Brewer M, Ichihara E, Sun J, Jin H, Ballard P, Al-Kadhimi K, Rowlinson R, Klinowska T, Richmond GH, Cantarini M, Kim

DW, Ranson MR, Pao W (2014) AZD9291, an Irreversible EGFR TKI, Overcomes T790M-mediated Resistance to EGFR Inhibitors in Lung Cancer. *Cancer Discov*. 4:1046-61.

[138] Castelli R, Bozza N, Cavazzoni A, Bonelli M, Vacondio F, Ferlenghi F, Callegari D, Silva C, Rivara S, Lodola A, Digiacomo G, Fumarola C, Alfieri R, Petronini PG, Mor M (2019) Balancing Reactivity and Antitumor Activity: Heteroarylthioacetamide Derivatives as Potent and Time-Dependent Inhibitors of EGFR. *Eur J Med Chem* 162:507-524

[139] Repasky MP, Chandrasekhar J, Jorgensen WL (2002) PDDG/PM3 and PDDG/MNDO: Improved Semiempirical Methods. *J Comput Chem* 23:1601-1622.

[140] Vayner G, Houk KN, Jorgensen WL, Brauman JI (2004) Steric Retardation of SN2 Reactions in the Gas Phase and Solution. *J Am Chem Soc* 126:9054-9058.

[141] Barf T, Kaptein A (2012) Irreversible Protein Kinase Inhibitors: Balancing the Benefits and Risks. *J Med Chem* 55, 6243− 6262.

[142] Gehringer M, Laufer SA. (2019) Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J Med Chem* In press.