Quantum mechanics/molecular mechanics modeling of covalent addition between EGFR-cysteine 797 and N -(4-Anilinoquinazolin-6-yl) acrylamide

Quantum Mechanics/Molecular Mechanics Modeling of Covalent Addition between EGFR-cysteine 797 and *N*-(4-anilinoquinazolin-6-yl) acrylamide

Luigi Capoferri,[1,¶] *Alessio Lodola,*[*,1] *Silvia Rivara,*[1] *and Marco Mor*[*,1]

[1]Dipartimento di Farmacia, Università degli Studi di Parma, Viale delle Scienze 27/A, I-43124, Parma, Italy

ABSTRACT

Irreversible EGFR inhibitors can circumvent resistance to first-generation ATP-competitive inhibitors in the treatment of non-small-cell lung cancer. They covalently bind a non-catalytic cysteine (Cys797) at the surface of EGFR active site by an acrylamide warhead. Herein, we used a hybrid quantum mechanics/molecular mechanics potential (QM/MM) in combination with umbrella sampling in the path-collective variable space to investigate the mechanism of alkylation of Cys797 by the prototypical covalent inhibitor *N*-(4-anilinoquinazolin-6-yl)acrylamide. Calculations show that Cys797 reacts with the acrylamide group of the inhibitor through a direct addition mechanism, with Asp800 acting as a general base/general acid in distinct steps of the reaction. The obtained reaction free energy is negative (ΔA = -12 kcal/mol) consistent with the spontaneous and irreversible alkylation of Cys797 by *N*-(4-anilinoquinazolin-6-yl) acrylamide. Our calculations identify desolvation of Cys797 thiolate anion as a key step of the alkylation process, indicating that changes in the intrinsic reactivity of the acrylamide would have only a minor impact on the inhibitor potency.

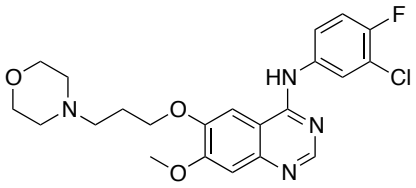
INTRODUCTION

The epidermal growth factor receptor (EGFR) is a cell surface receptor which is primarily activated by the epidermal growth factor (EGF).[1] Upon binding by EGF, EGFR undergoes a transition from a monomeric to a homodimeric form, leading to the activation of its intracellular kinase domain.[2] These events trigger intracellular signals that are needed to control cell proliferation, survival, and differentiation.[3] Overexpression of the wild-type form of EGFR or insurgence of specific mutations in the EGFR sequence conferring abnormal kinase activity (i.e. activating mutations L858R or G719S) have been related to cell division and insurgence of several type of carcinomas, including lung, breast and head and neck cancers.[4,5] Small molecule inhibitors of EGFR kinase activity are currently used for the therapeutic treatment of non small-cell lung cancer (NSCLC) and breast cancer.[6] The first generation of clinically approved EGFR inhibitors includes the 4-anilinoquinazoline derivatives gefitinib (Iressa, **1**)[7] and erlotinib (Tarceva, **2**),[8] which reversibly and competitively bind the ATP-binding site of the EGFR kinase domain. Gefitinib and erlotinib are effective inhibitors of the drug-sensitive L858R mutant associated with NSCLC. However, their clinical efficacy is limited by acquired mutations or by the activation of alternative signaling pathways.[9] Indeed, mutation at the gatekeeper residue threonine 790 with methionine (T790M mutation) in the catalytic domain of EGFR was found in nearly 50% of the NSCLC patients treated with gefitinib or erlotinib.[10] Compared to EGFR harboring the single mutation L858R, the double mutant T790M-L858R shows increased affinity for ATP and decreased affinity for gefitinib.[11] Although these mutations can affect various factors related to EGFR signaling, including equilibrium between active and inactive conformation,[12] receptor dimerization and hyperphosphorylation of its intra-cellular domain,[13] competition with ATP plays a major role in reducing the in vivo efficacy of reversible EGFR inhibitors toward tumors also harboring T790M mutation.

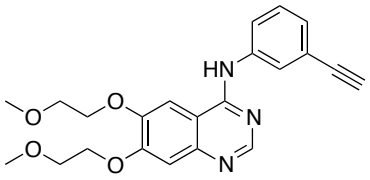
The insurgence of cancer cell resistance to gefitinib and erlotinib has prompted the development of a second generation of inhibitors having a Michael acceptor group, generally in position 6 of an anilinoquinazoline or anilino-3-cyanoquinoline scaffold. Such a warhead was introduced with the aim to

alkylate a non-catalytic cysteine (Cys797) in the ATP binding site of EGFR.[14] Formation of a stable carbon-sulfur bond between these inhibitors and EGFR[15] confers an irreversible mechanism of action to the class, along with significant advantages over reversible inhibitors.[16] In fact, irreversible inhibitors of EGFR display a remarkably high receptor residence time, circumventing competition with ATP and being active also on EGFR T790M mutant resistant to gefinitib and erlontinib. The *N*-(4-((3-bromophenyl)amino)quinazolin-6-yl)acrylamide PD168393 (**3**)[17] can be considered as the prototypical irreversible inhibitor of EGFR, and starting from its structure several second generation inhibitors were developed with dacomitinib (**4**),[18] advanced to phase III studies, and afatinib (**5**),[19] recently approved by the Food and Drug Administration for the treatment of NSCLC.[20] A third generation of EGFR inhibitors has been recently developed. This includes the mono anilino-pyrimidines AZD-9291,[21] WZ4002[22] and CO-1686[23] which are potent inhibitors of the T790M-L858R double mutant form of EGFR with selectivity over the wild type. These inhibitors might provide clinical benefit in T790M-positive patients, as they should devoid the dose-limiting toxicities due to the inhibition of EGFR wild type. Although EGFR covalent inhibitors are emerging as clinically superior to reversible drugs,[24] the detailed mechanism of their reaction with the target cysteine has not been investigated so far.

Reversible, non-covalent inhibitors of EGFR

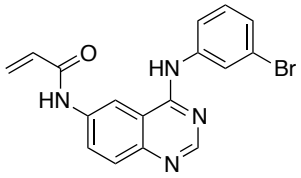


gefitinib (Iressa), **1**



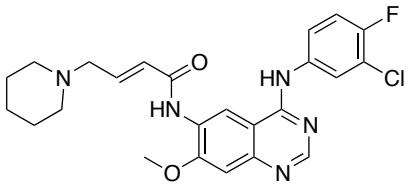
erlotinib (Tarceva), **2**

Irreversible, covalent inhibitors of EGFR



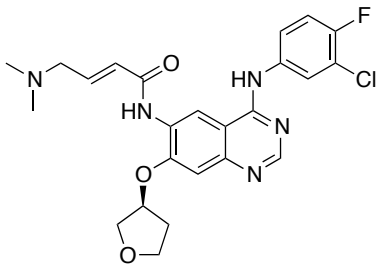
PD168393, **3**

k_{inact}/K_i = 1.5 μM^{-1} s^{-1}



dacomitinib, **4**

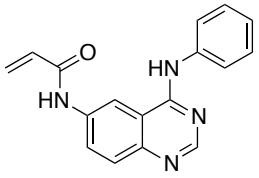
k_{inact}/K_i = 2.8 μM^{-1} s^{-1}



afatinib, **5**

k_{inact}/K_i = 15 μM^{-1} s^{-1}

Covalent inhibitor of EGFR considered in this study



13-JAB, **6**

Figure 1. Chemical structures of selected EGFR inhibitors. For reference covalent inhibitors **3**-**5**, the overall biochemical potency on L858R/T790M EGFR (expressed as k_{inact}/K_i) is also reported.[27]

While the X-ray structures of a number of covalent adducts involving Cys797 of EGFR and acrylamide-based alkylating groups have been disclosed,[15,25,26] only very recently the kinetics of EGFR inactivation by several classes of covalent inhibitors appeared in the literature.[27] These data show that the intrinsic reactivity of the acrylamide warhead linked to the 4-anilinoquinazoline scaffold is only one of the chemical determinant affecting EGFR inhibition and that other factors, such as reactant alignment, enforced local concentration, and cysteine pKa also contribute to the inactivation rate constants (k_{inact}) observed.[27]

Prompted by these findings we simulated at atomic level, by applying a quantum mechanics/molecular mechanics (QM/MM) approach,[28] the reaction that can occur between Cys797 of EGFR and the acrylamide fragment of a covalent inhibitor. To keep the interaction system as simple as possible, we started from *N*-(4-(phenylamino)quinazolin-6-yl)acrylamide (13-JAB, compound **6**, Figure 1), a congener of PD168393. The crystal structure of the covalent adduct formed by compound **6** and EGFR had been previously published.[15] Calculations of this kind had already proved to give signification information for drug discovery,[29] i.e. clarifying the mechanism of action of covalent inhibitors,[30] understanding drug resistance,[31] and predicting drug metabolites.[32]

In the present work, we propose a mechanism for the reaction involving EGFR and the prototypical irreversible inhibitor **6** and we report the corresponding free energy surface obtained combining a hybrid SCC-DFTB/AMBER potential[33] with umbrella sampling in the space of path collective variables (PCVs).[34] This computational strategy has been shown to be effective in studying reaction mechanisms[35,36] and to explicitly account for conformational and solvent reorganization events coupled with chemical transformation in solvent-exposed active sites.[37]

RESULTS AND DISCUSSION

Identification of Asp800 as key residue in the addition mechanism

The possible reaction paths for the addition of Cys797 to the acrylamide of compound **6** are depicted in Figure 2. While at physiological pH cysteine thiols are usually present in their neutral form (Cys-SH),[38] the thiolate (Cys-S^-) anion is much more reactive toward α,β-unsaturated carbonyls.[39] Thus, the presence of a basic group in the proximity of Cys797 could activate the thiol group by abstracting its proton. The resulting thiolate could then react with the acrylamide following one of two alternative mechanisms. In the *direct addition* mechanism (Figure 2, path A) the thiolate attacks Cβ of the acrylamide and a proton is transferred from the protonated base at the EGFR active site to the Cα of the acrylamide. In this way, a beta-substituted amide is obtained as the final product of the reaction. In the

1,4 addition mechanism (Figure 2, path B), after the nucleophilic attack and the formation of enolate intermediate, the proton is transferred from the protonated base to the oxygen atom, leading to the formation of the enolic form of the amide. Solvent-assisted tautomerization ("ketonization") is then required to obtain the final product of the reaction. A recent investigation by Paasche *et al* on the reactivity of unsaturated carbonyl substrates showed that, differently from α,β-unsaturated aldehydes, α,β-unsaturated amides react with thiols following preferentially *direct addition* mechanism, as the ketonization process after *1,4 addition* is an unlikely event, due to high energy content of the enol intermediate.[40] The preference for the *direct addition* is also supported by experimental findings showing that addition of thiols to substituted α,β-unsaturated carboxylic acid derivatives at room temperature is featured by high diastereoselectivity, consistent with anti addition of thiolate and proton, respectively.[41]

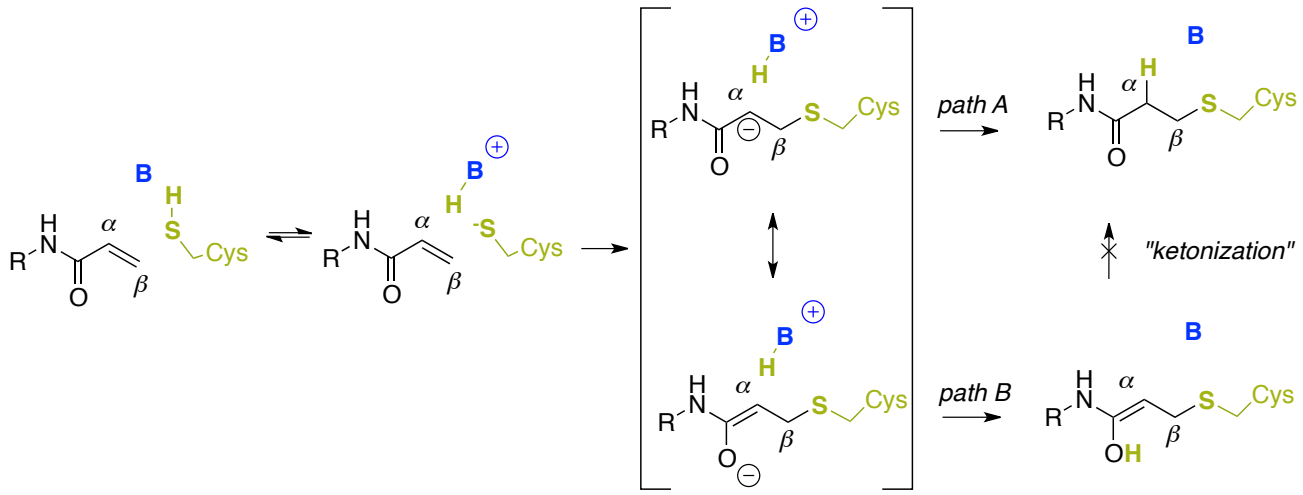


Figure 2. Possible reaction mechanisms for the addition of cysteine to a substituted acrylamide, in the presence of a residue acting as general base (B)/general acid (BH^+).

We thus hypothesized that acrylamide-based inhibitors, such as compound **6**, bind Cys797 of EGFR through a direct addiction mechanism. We visually inspected a number of X-ray structures of EGFR bound to covalent and non-covalent 4-anilinoquinazoline inhibitors deposited in the PDB to search for a residue in the proximity of Cys797 that may act as general base/general acid catalyst. Our inspection (which is summarized in Table 1) identified Asp800, a residue placed at the boundary between the

quinazoline binding site and the solvent-exposed region, as a key residue involved in a proton shuttle mechanism. We noticed that the carboxylate group of Asp800 crystallized close to the thiol group of Cys797 when EGFR was in complex with reversible inhibitors (i.e. gefitinib and iressa). Conversely, it crystallized close to the Cα of the former acrylamide group when Cys797 sulfur atom of EGFR was covalently bound to irreversible inhibitors (i.e. 13-JAB, PD168393, afatinib and daconitinib).

Table 1. Selected X-ray structures of EGFR in complex with non covalent (rows **I**-**V**) or covalent 4-anilinoquinazoline inhibitors (**VI**-**XI**). The distance between the closest carboxyl oxygen of Asp800 and the sulfur atom of Cys797 is reported for structures **I**-**V,** while the distance between the closest carboxyl oxygen of Asp800 and Cα of the acrylamide of the inhibitor is reported for structures **VI**-**XI**.

#	PDB Code	Protein-Ligand Complex or Adduct	RES. (Å)	Distance C797-D800 (Å)	Distance Cα-D800 (Å)
I	2ITY[42]	EGFR-IRESSA	3.4	3.41	-
II	3UG2[43]	EGFR (G719S) - GEFITINIB	2.5	4.06	-
III	2ITO[42]	EGFR (G719S) - IRESSA	3.2	3.81	-
IV	2ITZ[42]	EGFR (L858R) - IRESSA	2.7	3.98	-
V	1M17[44]	EGFR - ERLOTINIB	2.6	3.54	-
VI	4I24[26]	EGFR (T790M) - DACONITINIB ADDUCT	1.8	-	3.46
VII	4G5J[25]	EGFR - AFATINIB ADDUCT	2.8	-	4.05
VIII	4G5P[25]	EGFR (T790M) - AFATINIB ADDUCT	3.2	-	3.14
IX	2J5E[15]	EGFR - 13JAB ADDUCT	3.1	-	4.25
X	2J5F[15]	EGFR - PD168393 ADDUCT	3.0	-	4.52
XI	4LQM[45]	EGFR (L858R) - PD168393 ADDUCT	2.5	-	3.58

To assess if the observed arrangements of Asp800 in EGFR active site correspond to free-energy minima we performed molecular modeling studies using the empirical force field AMBER99SB.[46] Starting from the crystal structure of the covalent adduct EGFR-compound **6** (2J5E.pdb),[15] we firstly prepared a molecular model of a putative Michaelis-like complex in which the *N*-(4-anilinoquinazolin-6-yl) acrylamide **6** binds the EGFR active site only through non-covalent interactions. Compound **6** was modeled in the active of EGFR with the 6-acrylamide group in s-*cis* conformation. Previously performed docking simulations showed that this arrangement is preferred over the s-*trans*

conformations.[47] Furthermore, the s-*cis* conformer allowed to accommodate the oleofinic double bond close to Cys797, with the β-carbon atom at only 4 Å from the reactive thiol group of the cysteine.

This non-covalent complex was submitted to energy minimization followed by 100 ns of unbiased molecular dynamics (MD) simulations in NVT conditions. The quality of the MD simulations was assessed by determining the root-mean-square deviation (RMSD) of the frames with respect to the first frame of the simulation as well as by checking the stability of defined secondary structure of the protein (DSSP). Detailed plots of RMSD values and DSSP secondary structure analysis *vs* the simulation time are reported in the Supporting Information (SI). After the first 5 ns of MD the complex was stable and showed average RMSD value of 1.8 Å over the whole simulation time of 100 ns (Figure S1). Furthermore, analysis of the secondary structure indicates that the structural elements of the protein were retained over the entire simulation (Figure S2). The inhibitor remains tightly bound to the hinge region of EGFR, with the olefinic double bond stably placed in s-*cis* conformation, and no conversion in the s-*trans* form was observed. Conversely, Asp800 fluctuated between two alternative conformations. In the first one, Asp800 maintained its crystallographic position, with its carboxylate close to Cys797 ("IN" conformation, Figure 3A). In the second one, the carboxylate pointed toward the outer side of the ATP-binding site, where its negatively charge favored interactions with water molecules ("OUT" conformation, Figure 3B). The same simulation also showed that the transition between IN and OUT conformations occurred through rotation of the χ_1 dihedral angle of Asp800 (Figure 3C). Frequency analysis of χ_1 values obtained from the simulation trajectory, applying the Boltzmann law, allowed us to estimate a Helmholtz free energy difference (ΔA) of 1.6 kcal mol^{-1} between the two conformations. Although the "IN" conformation is predicted to be less favorable than the "OUT", the low energy difference indicates that they are both energetically accessible, suggesting that Asp800 can remain close to Cys797 for sufficient amount of time, and potentially act as the general base deprotonating the thiol group of EGFR Cys797.

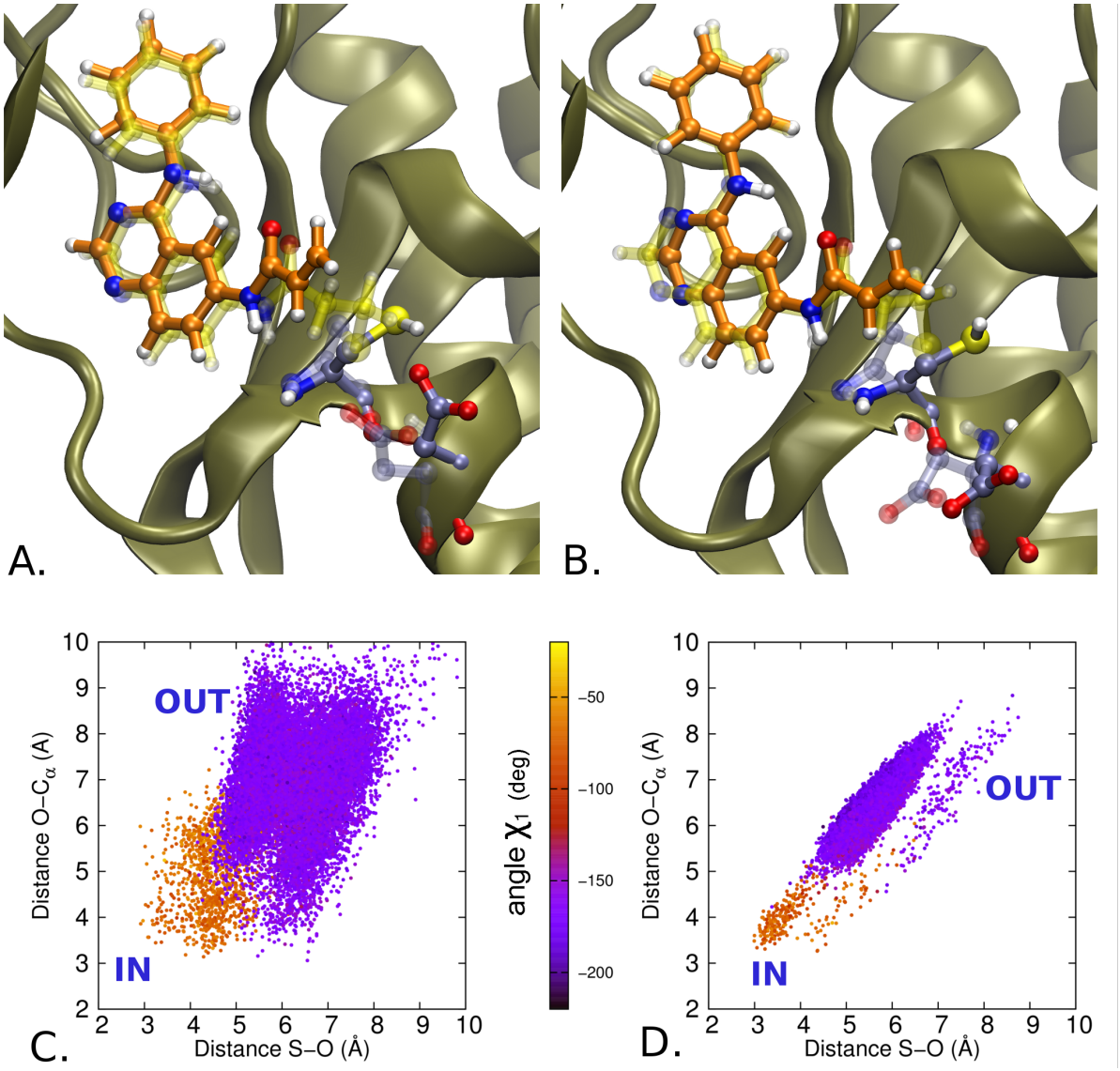


Figure 3. Representative "IN" (Panel A) and "OUT" (Panel B) conformations of Asp800 obtained from classical MD simulations of EGFR in complex with compound **6**. Conformations for the Michaelis-like complexes and covalent adducts are depicted as solid (orange carbon atoms) and transparent (yellow carbon atoms) ball-and-sticks, respectively. Plots of the distances between Asp800, Cys797 and the acrylamide fragment of compound **6** are reported in function of the angle χ_1 for the simulation of the Michaelis-like complex (Panel C) and the covalent complex (Panel D).

For comparison, we also performed 100 ns of unbiased MD simulations (NVT regime) of the EGFR-**6** covalent adduct, starting from the crystal structure. Also in this case, the structural features of the protein were retained over the whole simulation time (Figure S3) showing an average RMSD values for the alpha carbon of 2.1 Å (Figure S4). As expected, the anilinoquinazoline and the propanamide fragments were rather stable, showing very small fluctuation along the whole simulation. Also in this

simulation Asp800 assumed the two alternative conformations described above (Figure 3D). The frequency analysis of χ_1 indicates that the "IN" conformation is less stable than the "OUT" one by nearly 3 kcal/mol and this energy difference still indicates that these two states are both accessible at the simulated temperature of 300 K (Figure 3D). In the "IN" conformation Asp800 pointed toward one of the two hydrogen atoms emerging from the Cα of the former acrylamide.

Reaction path calculations

Also based on the previously described simulations, we hypothesized Asp800 as a key actor in the reaction leading to EGFR inhibition (Figure 4). We thus focused our attention on a mechanism in which Asp800 worked as a base deprotonating Cys797 in an early phase of the alkylation process (step **1**), and as an acid after the nucleophile attack at the Cβ of the acrylamide (step **2**), protonating the Cα of the acrylamide portion (step **3**).

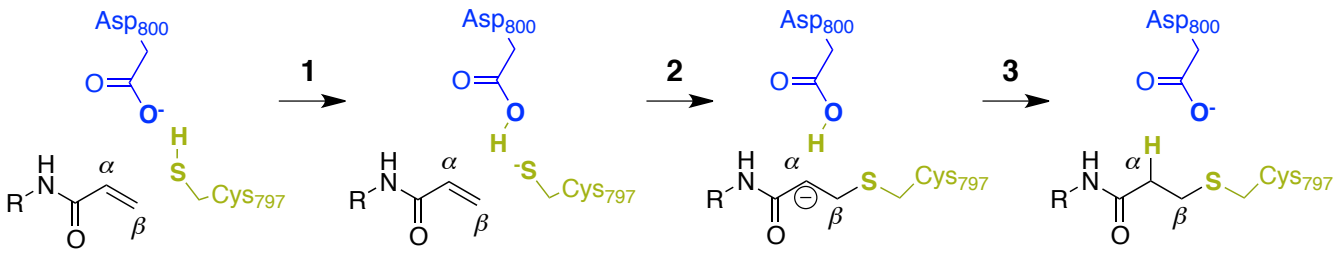


Figure 4. Proposed reaction mechanisms for the addition of Cys797 to the 6-acrylamide group of compound **6** assisted by Asp800.

Free energy simulations of multi-step reactions are often carried out through a convenient separation of the overall chemical transformation into discrete and consecutive events. While rather practical, this approach does not capture the concerted nature of chemical reactions.[48] The path method here applied overcomes this limitation using two collective variables representing the progression along (*S*) and the distance from (*Z*) a reference reaction pathway (see experimental section for details). To obtain a reasonable reference pathway, we initially simulated the *directed addition* mechanism of Figure 4 in a stepwise manner (**1**->**2**->**3**), using simple distances among atoms as reaction coordinates, i.e. $d(O_{Asp800}\text{-}H_{Cys797})$ for step **1**; $d(S_{Cys797}\text{-}C\beta_{acrylamide})$ for step **2**, and $d(H_{Cys797}\text{-}C\alpha_{acrylamide})$ for step **3**. These simulations

were performed combining a SCC-DFTB/AMBER99SB potential with the steered-MD approach. The role of these QM/MM steered-MD simulations was only to provide a collection of putative geometries lying along an ideal reaction coordinate separating the EGFR-compound **6** non-covalent complex (the reactants **R**) from the covalent adduct (the product **P**). The collection of geometries connecting **R** to **P** constituted the guess-path which was subject to an iterative geometrical optimization procedure, similar to that reported in references 35 and 37. In brief, employing the guess-path as a "reference route", the alkylation of Cys797 by compound **6** was simulated in the *S* and *Z* space using a series of independent steered-MD simulations applying the SCC-DFTB/AMBER99SB potential. While the variable *S* ensured the progress of the reaction up to the formation of **P**, the variable *Z* allowed the system to explore new geometries and thus alternative pathways in the proximity of the guess-path. The collection of geometries defining the guess-path was updated at every steered-MD run, allowing the exploration of reaction paths with lower energy. When we reached convergence the optimization process was stopped. Finally, to have a better description of the reaction energetics, umbrella-sampling (US) simulations were carried out along *S* on the converged reaction pathway to estimate the free energy associated with alkylation of Cys797. The exploration with US of the converged minimum free-energy path resulted in a region of low *Z* values, indicating that the optimized path was reasonable in term of explored geometries with the SCC-DFTB/AMBER99SB Hamiltonian (Figure S5). Furthermore, the work-curve obtained with the final steered-MD and the free-energy curve computed with US shared the same shape (Figure S6) indicating that the position of the key stationary points was not affected by the sampling methods.

Based on the free energy profile reported in Figure 5A, the *direct addition* mechanism can be divided into 5 chemically relevant regions: *i)* the Michaelis-like complex area where the 4-aniniloquinazoline moiety of the inhibitor occupies the ATP binding site in the "crystallographic" pose, with the 6-acrylamido group in proximity of Cys797 and Asp800 in the so-called "IN" conformation; (*ii*) a rather large area which accounts for the first transition state (TS), where the proton of Cys797 thiol group is being transferred to the carboxylate group of Asp800; (*iii*) a wide area where Cys797 lives as thiolate and Asp800 as aspartic acid; (*iv*) a rather broad area, which accounts for a second TS where Cys797

thiolate undergoes desolvation before attacking the Cβ of the acrylamide generating a transient carbanion, which is readily protonated at the Cα by Asp800; the last region (*v*) which accounts for the products of the overall reaction, where Cys797 is alkylated by compound **6** and Asp800 is present in its anionic form.

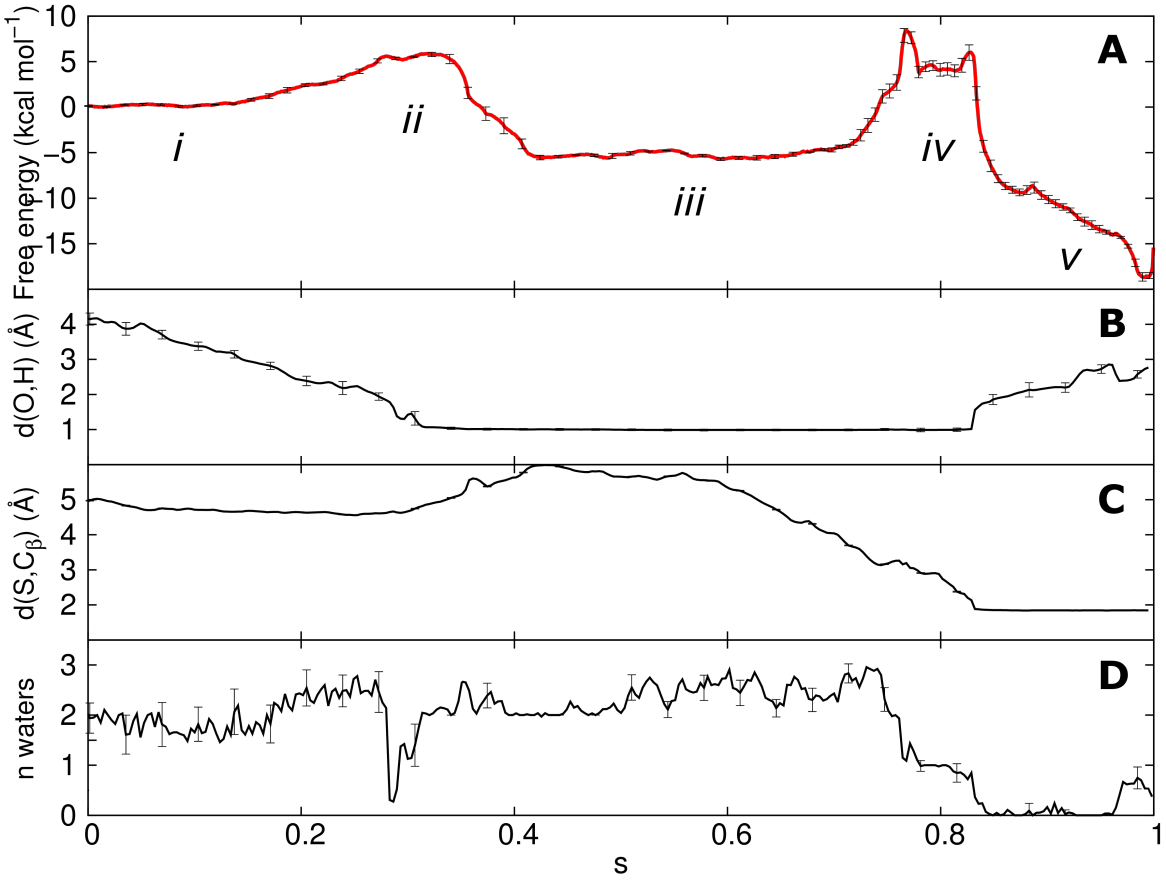


Figure 5. Alkylation of EGFR by compound **6**. Panel A, Free energy profile of the reaction with relevant steps (*i-v*) highlighted. Panel B, evolution of the distance between O_{Asp797} and the H_{Cys797} along the collective variable *S*. Panel C, evolution of distance between the S_{Cys797} and the $C\beta_{acrylamide}$ along *S*. Panel D, number of waters within 3.5 A of the thiolate S. Values are represented as average with error bars representing the standard deviations.

In Figure 5B and 5C, the evolution of relevant interatomic distances is plotted as a function of the collective variable *S* to monitor the progress along the path. At values of *S* in the range between 0 and 0.2, Cys797 and Asp800 interacted tightly, at first through the formation of water mediated H-bonds (configuration **A**, Figure 6A) and then with a proton being directly transferred (*S* between 0.2 and 0.4,

configuration **B**, Figure 6B) from Cys797-SH to Asp800-COO^- generating the Cys797-S^-/Asp800-COOH pair (configuration **C,** Figure 6C). A range of different geometries with similar energies (i.e. within a k_bT from the top barriers) can be considered possible transition-state (TS) structures along the 'edge' dividing configurations **A** and **C**. Overall, the free energy barrier is 6 kcal/mol at SCC-DFTB/AMBER99SB level of theory. Also considering the 1.6 kcal/mol estimated for moving the aspartate toward cysteine, this barrier is indicative of a fast reaction. The regions of the path corresponding to *S* values between 0.4 and 0.6 are featured by the presence of the Cys797-S^-/Asp800-COOH pair with the inhibitor maintaining its interaction with the hinge region of EGFR. The free-energy of this intermediate (configuration **C**) is 6 kcal/mol lower than that of the reactants. Visual inspection of the reaction-path trajectory relative to transformation of **A** to **C** showed that two water molecules tightly interacted with Cys797-S^- accounting for the stabilization of the thiolate anion. These results suggest that Cys797-S^-/Asp800-COOH is the most stable and abundant ionization state for this pair within EGFR, at least in this specific holo form.

The region of *S* between 0.6 and 0.7 described small but significant movements of Cys797, Asp800 and the acrylamide that are necessary to place their key atoms in a reactive configuration. In particular, the thiolate group of Cys797 approached the β carbon of the inhibitor acrylamide (S-C distance from ~4.5 Å to ~3.2 Å, Figure 5C), while Asp800 places its acid proton toward the α carbon of the acrylamide (Cα-H distance from ~5.0 Å to ~2.8 Å). The shortening of these distances did not significantly affect the *s*-cis conformation of the acrylamide double bond of **6**.

In the *S* interval 0.7-0.75, the free energy of the system increased due to a further shortening of the S-C and Cα-H distances and to a significant reduction in the number of H-bonds formed by Cys-S^- and its neighboring water molecules (Figure 5D). The free-energy cost for this step resulted in 8.6 ± 0.2 kcal/mol (14.6 kcal/mol compared to **C**, the most stable configuration of the system) and resulted in the highest free-energy point along the path (configuration **D**, Figure 6D). Desolvation of the thiolate anion is a costly process due to the ability of the negatively charged sulfur to tightly bind water molecules.

Similar results have been reported for the enzyme glutathione S-transferase, which also has a cysteine thiolate nucleophile placed in a solvent-exposed environment.[49]

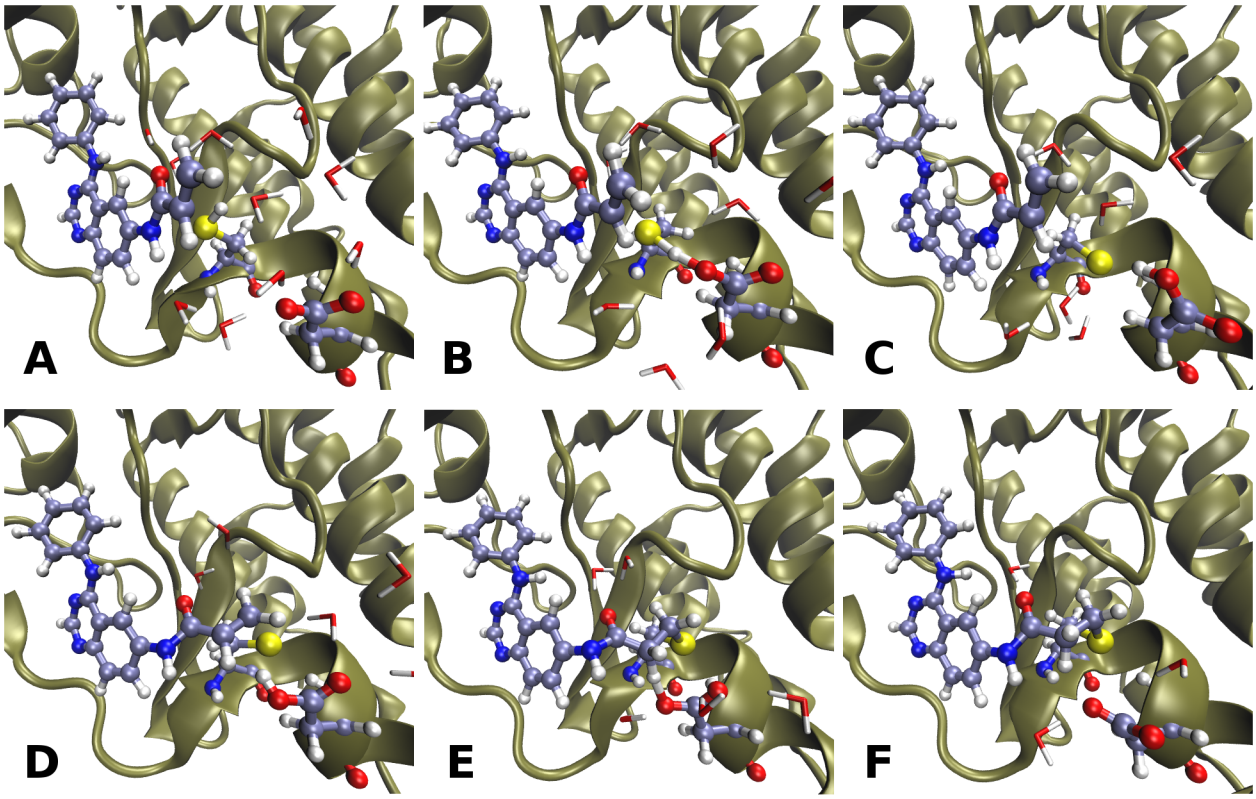


Figure 6. Structures of important configurations (**A-F**) identified along the QM/MM reaction pathway for the alkylation of Cys797 (pink carbon atoms) of EGFR (brown ribbons) by compound **6** (pink carbon atoms. Asp800 (pink carbon atoms) and key water molecules interacting with the Cys-Asp dyad are also displayed. Figur

The so-obtained activation barrier (E*act* = 14.6 kcal/mol) is in reasonable agreement with the experimental one (≈ 20 kcal/mol),[27] considering that the semi-empirical SCC-DFTB Hamiltonian has been reported to underestimate chemical barriers by about 20-25%.[50] This remains true when the energy cost for moving the key Asp800 from the out-conformation to the in-conformation (1.6 kcal/mol) is considered, as the overall barrier for the reaction becomes 16.2 kcal/mol.

At the S value of 0.76, the chemical events leading to the *direct addition* mechanism began (S-C distance = 2.80 Å; Cα-H distance = 2.30 Å), ending at *S* = 0.82, with the complete formation of an S-C

bond (S-C distance = 1.85 Å) and of Cα-H bond (Cα-H distance 1.10 Å). With the progressive reduction of the S-C distance in the 0.78-0.80 range of *S,* a significant elongation of the C=O bond of the acrylamide group was observed (i.e. from the reference value of 1.23 Å to 1.28 Å). This transient carbanion/enolate species was a high-energy configuration representing a third transition state along the path (configuration **E**, Figure 6E). The energy barrier necessary to overcome **E** (S = 0.8, S-C distance = 2.13 Å; Cα-H distance = 1.96 Å; C=O distance = 1.28 Å) was 6 kcal/mol (12.6 kcal/mol compared to configuration **C**). The *direct addiction* reaction took place in a concerted manner where nucleophile addition at Cβ and protonation at Cα were synchronous events. These results are consistent with potential energy calculations performed by Grazioso *et al* on a similar system (i.e. alkylation of falcipain-2 cysteine by an α,β-unsaturated methyl ester inhibitor) where thiolate attack and Cα protonation by a protonated hystidine were also concerted.[51]

In the last region of the reaction path (*S* values: 0.83-1.0) the EGFR-compound **6** adduct underwent to a minor conformational relaxation that led to the highly stable product **F**. The free-energy of the product is significantly lower than that of configurations **A** (-18 kcal/mol) and **C** (-12 kcal/mol), indicating that alkylation of Cys797 by the acrylamide inhibitor **6** is a spontaneous and irreversible process, which is in agreement with kinetic,[27,52] structural[15] and pharmacological data.[17]

CONCLUSIONS

Herein we investigated the mechanism of EGFR inhibition operated by the prototypical covalent inhibitor *N*-(4-anilinoquinazolin-6-yl)acrylamide by applying a free-energy based method combining a hybrid SCC-DFTB/AMBER potential[53] with umbrella sampling and the path collective variable approach. Our simulations support a *direct addition* mechanism in which Asp800 acts both as a base and as an acid in distinct steps of Cys797 alkylation. Specifically, the reaction between EGFR and acrylamide **6** starts with deprotonation of the thiol group of Cys797 by Asp800. After the key desolvation of the nucleophile Cys-S^-, the thiolate attacks Cβ of the acrylamide fragment of the inhibitor

while Asp800, in a concerted and synchronous manner, protonates the adjacent Cα leading to a stable tioether product. The proposed mechanism is energetically reasonable with a reaction free-energy (ΔA = -12 kcal/mol), consistent with the spontaneous and irreversible alkylation of Cys797. Kinetic experiments aimed at determining the k_{inact} of acrylamide derivative **6** on EGFR D800A or D800N mutants could confirm or reject this direct addition mechanism. In the meanwhile, we believe our that results shed light on two recent experimental findings relative to the functioning and inhibition of EGFR. First, the existence of a free-energy minimum state of EGFR featured by the presence of a Cys797-S^-/Asp800-COOH pair is consistent with the easy oxidation of Cys797 to sulfenic acid in presence of oxidative stimuli, which leads to EGFR activation.[54] Thus, not only could the presence of the side chain of Asp800 be important for irreversible inhibition of EGFR by cysteine-trapping agents, but it could also play a role in EGFR activation by reactive oxygen species. Secondly, the identification of desolvation of Cys797 thiolate as a key event in the alkylation process makes reason of the similar rate constants observed for EGFR inactivation by a series of acrylamide-based inhibitors including afatinib, dacomitinib and PD168393 (Figure 1), displaying k_{inact} values of 0.9 ms^{-1}, 1.5 ms^{-1} and 2.3 ms^{-1}, respectively, either having or not a terminal basic amine.[27] The *direct addition* mechanism here reported, along with these kinetic data, suggests that the terminal tertiary base attached to the acrylamide moiety of second generation EGFR inhibitors is not likely to participate to chemical step of Cys797 alkylation (i.e. by deprotonating thiol group of nucleophile)[14] but rather it improves the recognition phase and thus the non-covalent binding affinity for EGFR.[27]

EXPERIMENTAL SECTION

Model building and equilibration of the system by MD simulations

The EGFR-compound **6** non covalent complex was built starting from the crystal structure of the corresponding covalent adduct 2J5E.pdb by conveniently changing atom and bond types of the 4-anilinoquinazoline inhibitor as well as of Cys797. The resulting complex was immersed in a box of TIP3P water molecules[55] and neutralized with 2 Cl^- ions by using the xleap tool implemented in

AMBER11.[56] The total system size amounted to 52429 atoms (box size of 78.2 Å x 91.7 Å x 73.2 Å). The system was energy-minimized and gradually heated to 300 K in NVT ensemble and equilibrated at pressure of 1 atm in NPT ensemble. The system was then submitted to a MD simulation in the NVT ensemble for 100 ns applying the AMBER99SB force field[46] for the protein and the generalized Amber force field (GAFF)[57] for the ligand. The pmemd module of AMBER11 was used to perform these tasks. Full electrostatic and Van der Waals interactions were computed within a cut-off of 10 Å and long-range electrostatic interactions were treated using the particle mesh Ewald (PME)[58] with 128x128x128 grid points. The covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm and a time-step of 2 fs was applied. The same approach was employed for the system of the EGFR-compound **6** covalent adduct. The final system was composed by 52425 atoms (box size of 81.5 Å x 95.4 Å x 76.4 Å), of which 2 Cl^- ions to neutralize the total charge. MD simulations of the system were performed applying the same protocol described for the non-covalent complex.

Application of the QM/MM potential

The hybrid quantum mechanical and molecular mechanical approach (QM/MM) has been extensively applied to study enzymatic reaction mechanisms. This approach has the advantage that large system such as an enzyme can be investigated with an affordable computation cost. The fundamental idea behind the QM/MM scheme is the partition of the simulated system into two regions: *i*) a small QM region, where the chemical bond breaking and forming occur, and *ii)* and a large MM region which surrounds the QM atoms. In the present work we used the self-consistent charge- density functional tight binding (SCC-DFTB) model[59] to describe the QM region and the AMBER99SB force field[46] to describe the MM region. The SCC-DFTB approach is based on the second-order expansion of the total DFT energy with respect to the charge-density variation. It is more efficient than *ab initio* QM approaches and therefore suitable for extensive QM/MM free-energy simulations. Furthermore, hybrid

QM/MM potentials derived from the SCC-DFTB theory have been widely applied to enzyme catalysis and reported to give reasonable descriptions of reaction geometries and energetics. [60]

In the EGFR-compound **6** system, side chain atoms of Cys797, Asp800 and the whole compound **6** were treated with the SCC-DFTB method with dispersion correction. All the other atoms of the system were described with AMBER99SB force field. The resulting QM system was composed of 49 atoms including two link atoms, placed along the C-C bond connecting Cβ of Cys797 and Asp800 to their backbone Cα. The adjust_q function of AMBER was applied to conserve the total charge of the system. During QM/MM MD simulations, all the atoms of the system (including hydrogens) were free to move, and a time step of 0.2 fs was used to integrate the equation of motion. The SHAKE option was turned off for the QM region during all QM/MM simulations. A modified PME approach was used to treat the QM/MM long-range electrostatic interactions.

Definition of the reaction path

A preliminary guess-path between reactants and products was generated using steered-MD and simple reaction coordinates. The reaction was initially divided into three consecutive steps using simple distances among atoms as reaction coordinates, i.e $d(O_{Asp800}\text{-}H_{Cys797})$ for step **1**; $d(S_{Cys797}\text{-}C\beta_{acrylamide})$ for step **2**, and $d(H_{Cys797}\text{-}\ C\alpha_{acrylamide})$ for step **3** (Figure 5). In all these steps, the RCs were pulled to their target values by using a force constant of 300 kcal/mol $Å^{-2}$ for a total simulation time of 15 ps. From the resulting steered-MD trajectory, a set of frames, representing the preliminary path between reactants and products, was extracted. 12 representative atoms of the side chains of Cys797 and Asp800 and the 6-acrylamide fragment of compound **6** were used as template structure for the frame selection procedure for PCVs.[34] We obtained 70 equally spaced frames that were used to define the collective variables S and Z in a subsequent steered-MD simulation. In detail, S and Z describe the position of a point in the configurational space (R) relative to an initial path, and are defined as follows:

$$S(R) = \frac{\sum_{i=1}^{P} i e^{-\lambda(R-R(i))^2}}{\sum_{i=1}^{P} e^{-\lambda(R-R(i))^2}}$$

$$Z(R) = -\frac{1}{\lambda}\ln\left(\sum_{i=1}^{P} e^{-\lambda(R-R(i))^2}\right),$$

where *i* is a discrete index (in this study ranging from 0 to 1), and $(R - R(i))^2$ is calculated as the mean square displacement from the initial path. The system was thus pulled along *S* from 0 (reactants) to 1 (products) in 550 ps, applying a force constant of 300 kcal/mol. The variable *Z* was constrained by a quartic wall to allow a relative freedom of the system to relax, while preventing it to escape from the reactive region. The upper limit over *Z* was set to 0.05 $Å^2$ with a force constant of 300 kcal/mol $Å^{-4}$. At the end of the steered-MD/PCVs simulation, a new set of frames was extracted and employed as a novel reference path for a subsequent steered-MD/PCVs run. This procedure was repeated iteratively until the reaction path did not change passing from a steered-MD/PCVs run to a new one. Convergence of the reaction path was achieved after 8 iterative cycles (Figures S7 and S8).

Free Energy calculation

PCVs were employed as reaction coordinates in umbrella sampling (US) simulations, using the converged steered-MD configurations as the reference path. US simulations were carried out using a spring constant of 10 kcal/mol on *S* and an upper limit on Z was set at 0.05 with force constant of 300 kcal/mol $Å^{-4}$. 40 ps of simulation per umbrella were performed, of which the first 10 ps were discarded to ensure the equilibration of the system. The normal distribution of the configurations along *S* was evaluated for each umbrella and a total of 4.4 ns of productive simulation was employed to estimate the FES by the weighted histogram analysis method (WHAM).[61,62] Average error on the FES was estimated to be 0.3 kcal/mol (see SI for details). All the steered-MD and US simulations were carried out with sander module of AMBER11 patched with the PLUMED 1.3 package.[63]

ASSOCIATED CONTENT

Supporting Information

RMSD and DSSP plot for classical MD simulations. Additional plots relative to the optimization of the guess-path by steered-MD simulations at the SCC-DFTB/AMBER99SB level of theory. Work curve for the alkylation of Cys797 by compound **6** obtained from the last round of steered-MD optimization at SCC-DFTB/AMBER99SB level of theory

This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

Alessio Lodola

E-mail: alessio.lodola@unipr.it

Marco Mor

E-mail: marco.mor@unipr.it

Author Contributions

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

Notes

[¶] Current address: AIMMS Division of Molecular Toxicology, Department of Chemistry and Pharmaceutical sciences, Vrije Universiteit, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands.

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) and University of Parma. Dr. Davide Branduardi (Crucell Vaccine Institute, Leiden, The Netherlands), Prof. Adrian J. Mulholland (School of Chemistry, University of Bristol, UK) and Dr. Brion W. Murray (Oncology Research Unit, Pfizer, San Diego, California) are acknowledged for useful discussions.

ABBREVIATIONS USED

QM/MM, quantum mechanics/molecular mechanics; SMD, steeered molecular dynamics; SCC-DFTB, self-consistent charge-density functional tight binding; TS, transition state.

REFERENCES

[1] Roskoski, R.; Jr. The ErbB/HER Family of Protein-Tyrosine Kinases and Cancer. *Pharmacol. Res.* **2014**, 79, 34-74.

[2] Olayioye, M. A.; Neve, R. M.; Lane, H. A.; Hynes, N. E. The ErbB Signaling Network: Receptor Heterodimerization in Development and Cancer. *EMBO J.* **2000**, *19*, 3159–3167.

[3] Yarden, Y.; Sliwknowski, M. X. Untangling the ErbB Signaling Network. *Nat. Rev. Mol. Cell. Biol.* **2001**, *2*, 127–137.

[4] Sharma, S. V.; Bell, D. W.; Settleman, J.; Haber, D. A. Epidermal Growth Factor Receptor Mutations in Lung Cancer. *Nat. Rev. Cancer* **2007**, *7*, 169–181.

[5] Normanno, N.; De Luca, A.; Bianco, C.; Strizzi, L.; Mancino, M.; Maiello, M. R.; Carotenuto, A.; De Feo, G.; Caponigro, F.; Salomon, D. S. Epidermal Growth Factor Receptor (EGFR) Signaling in Cancer. *Gene* **2006**, *366*, 2–16.

[6] Roskoski, R.; Jr. ErbB/HER Protein-Tyrosine Kinases: Structures and Small Molecule Inhibitors. *Pharmacol. Res.* **2014**, 87 C, 42-59.

[7] Barker, A. J.; Gibson, K. H.; Grundy, W.; Godfrey, A. A.; Barlow, J. J.; Healy, M. P.; Woodburn, J. R.; Ashton, S. E.; Curry, B. J.; Sarlett, L.; Henthorn, L.; Richards, L. Studies Leading to the Identification of ZD1839 (Iressa): an Orally Active, Selective Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Targeted to the Treatment of Cancer. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1911–1914.

[8] Moyer, J. D.; Barbacci, E. G.; Iwata, K. K.; Arnold, L.; Boman, B.; Cunningham, A.; Di Orio, C.; Doty, J.; Morin, M. J.; Moyer, M. P.; Neveu, M.; Pollack, V. A.; Pustilnick, L. R.; Reynolds, M. M.; Sloan, D.; Theleman, A.; Miller, P. Induction of Apoptosis and Cell Cycle Arrest by CP-358,774, an Inhibitor of Epidermal Growth Factor Receptor Tyrosine Kinase. *Cancer Res.* **1997**, *57*, 4838–4848

[9] Kobayashi, S.; Boggon, T. J.; Dayaram, T.; Janïne, P. A.; Kocher, O.; Meyerson, M.; Johnson, B. E.; Eck, M. J.; Tenen, D. G.; Halmos, B. EGFR Mutation and Resistance of Non-Small-Cell Lung Cancer to Gefitinib. *New Engl. J. Med.* **2005**, *352*, 786–792.

[10] Engelman, J. A.; Janïne, P. A. Mechanisms of Acquired Resistance to Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. *Clin. Cancer Res.* **2008**, *14*, 2895–2899.

[11] Yun, C.-H.; Mengwasser, K. E.; Toms, A. V.; Woo, M. S.; Greulich, H.; Wong, K.-K.; Meyerson, M.; Eck, M. J. The T790M Mutation in EGFR Kinase Causes Drug Resistance by Increasing the Affinity for ATP. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2070–2075.

[12] Eck, M. J.; Yun , C. H. Structural and Mechanistic Underpinnings of the Differential Drug Sensitivity of EGFR Mutations in Non-Small Cell Lung Cancer. *Biochim. Biophys. Acta.* **2010**, 1804, 559-566.

[13] Red Brewer, M.; Yun, C. H.; Lai, D.; Lemmon, M. A.; Eck, M. J.; Pao, W. Mechanism for Activation of Mutated Epidermal Growth Factor Receptors in Lung Cancer. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E3595-E3604.

[14] Carmi, C.; Lodola, A.; Rivara, S.; Vacondio, F.; Cavazzoni, A.; Alfieri, R. R.; Ardizzoni, A.; Petronini, P. G.; Mor, M. Epidermal Growth Factor Receptor Irreversible Inhibitors: Chemical Exploration of the Cysteine-Trap Portion. *Mini-Rev. Med. Chem.* **2011**, *11*, 1019–1030.

[15] Blair, J. A.; Rauh, D.; Kung, C.; Yun, C.-H.; Fan, Q.-W.; Rode, H.; Zhang, C.; Eck, M. J.; Weiss, W. A.; Shokat, K. M. Structure-Guided Development of Affinity Probes for Tyrosine Kinases Using Chemical Genetics. *Nat. Chem. Biol.* **2007**, *3*, 229–238.

[16] Carmi, C.; Mor, M.; Petronini, P. G.; Alfieri, R. Clinical Perspectives for Irreversible Tyrosine Kinase Inhibitors in cancer. *Biochem. Pharmacol.* **2012**, *84*, 1388-1399.

[17] Fry, D. W.; Bridges, A. J.; Denny, W. A.; Doherty, A.; Greis, K. D.; Hicks, J. L.; Hook, K. E.; Keller, P. R.; Leopold, W. R.; Loo, J. A.; McNamara, D. J.; Nelson, J. M.; Sherwood, V.; Smaill, J. B.; Trumpp-Kallmeyer, S.; Dobrusin, E. M. Specific, Irreversible Inactivation of the Epidermal Growth Factor Receptor and erbB2, by a New Class of Tyrosine Kinase Inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12022–12027.

[18] Reckamp, K. L.; Giaccone, G.; Camidge, D. R.; Gadgeel, S. M.; Khuri, F. R.; Engelman, J. A.; Koczywas, M.; Rajan, A.; Campbell, A. K.; Gernhardt, D.; Ruiz-Garcia, A.; Letrent, S.; Liang, J.; Taylor, I.; O'Connell, J. P.; Jänne, P. A. A Phase 2 Trial of Dacomitinib (PF-00299804), an Oral, Irreversible pan-HER (Human Epidermal Growth Factor Receptor) Inhibitor, in Patients with Advanced

Non-Small Cell Lung Cancer after Failure of Prior Chemotherapy and Erlotinib. *Cancer* **2014**, *120*, 1145-1154.

[19] Ferrarotto, R.; Gold, K. A. Afatinib in the Treatment of Head and Neck Squamous Cell Carcinoma. *Exp. Opin. Investig. Drugs* **2014**, *23*, 135-143.

[20] Dungo, R. T.; Keating, G. M. Afatinib: First Global Approval. *Drugs*. **2013**, *73*, 1503-1515.

[21] Cross, D. A. E.; Ashton, S.; Ghiorghiu, S.; Eberlein, C.; Nebhan, C.; Spitzler, P.; Orme, J.; Finlay, M. R. V.; Ward, R. A.; Mellor, M.; Hughes, G.; Rahi, A.; Jacobs, V.; Red Brewer, M.; Ichihara, E.; Sun, J.; Jin, H.; Ballard, P.; Al-Kadhimi, K.; Rowlinson, R.; Klinowska, T.; Richmond, G.; Cantarini, M.; Kim, D.-W.; Ranson, M.; Pao, W. AZD9291, an Irreversible EGFR TKI, Overcomes T790M-Mediated Resistance to EGFR Inhibitors in Lung Cancer. *Cancer Discovery* **2014**, *4*, 1046–1061.

[22] Zhou, W.; Ercan, D.; Chen, L.; Yun, C.-H.; Li, D.; Capelletti, M.; Cortot, A. B.; Chirieac, L.; Iacob, R. E.; Padera, R.; Engen, J. R.; Wong, K.-K.; Eck, M. J.; Gray, N. S.; Janne, P. A. Novel Mutant-Selective EGFR Kinase Inhibitors Against EGFR T790M. *Nature* **2009**, *462*, 1070–1074.

[23] Walter, A. O.; Sjin, R. T. T.; Haringsma, H. J.; Ohashi, K.; Sun, J.; Lee, K.; Dubrovskiy, A.; Labenski, M.; Zhu, Z.; Wang, Z.; Sheets, M.; St Martin, T.; Karp, R.; van Kalken, D.; Chaturvedi, P.; Niu, D.; Nacht, M.; Petter, R. C.; Westlin, W.; Lin, K.; Jaw-Tsai, S.; Raponi, M.; Van Dyke, T.; Etter, J.; Weaver, Z.; Pao, W.; Singh, J.; Simmons, A. D.; Harding, T. C.; Allen, A. Discovery of a Mutant-Selective Covalent Inhibitor of EGFR that Overcomes T790M-MediatedR in NSCLC. *Cancer Discovery* **2013**, *3*, 1404–1415.

[24] Kwak, E. The Role of Irreversible HER Family Inhibition in the Treatment of Patients with Non-Small Cell Lung Cancer. *Oncologist*. **2011**, *16*, 1498-1507.

[25] Solca, F.; Dahl, G.; Zoephel, A.; Bader, G.; Sanderson, M.; Klein, C.; Kraemer, O.; Himmelsbach, F.; Haaksma, E.; Adolf, G. R. Target Binding Properties and Cellular Activity of Afatinib (BIBW 2992), an Irreversible ErbB Family Blocker. *J. Pharmacol. Exp. Ther.* **2012**, *343*, 342-350.

[26] Gajiwala, K. S.; Feng, J.; Ferre, R.; Ryan, K.; Brodsky, O.; Weinrich, S.; Kath, J. C.; Stewart A. Insights into the Aberrant Activity of Mutant EGFR Kinase Domain and Drug recognition. *Structure* **2013**, *21*, 209-219.

[27] Schwartz, P. A.; Kuzmic, P.; Solowiej, J.; Bergqvist, S.; Bolanos, B.; Almaden, C.; Nagata, A.; Ryan, K.; Feng, J.; Dalvie, D.; Kath, J. C.; Xu, M.; Wani, R.; Murray, B. W. Covalent EGFR Inhibitor Analysis Reveals Importance of Reversible Interactions to Potency and Mechanisms of Drug Resistance. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 173**–**178.

[28] (a) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249. (b) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700–733. (c) Gao, J. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. R.; Eds.; VCH Publishers: New York, **1996**; Vol. 7, pp 119–185. (d) Warshel, A. Computer Simulations of Enzyme Catalysis: Methods, Progress, and Insights. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *32*, 425–443.

[29] (a) Lodola, A.; De Vivo, M. The Increasing Role of QM/MM in Drug Discovery. *Adv. Protein Chem. Struct. Biol.* **2012**, *87*, 337–362. (b) De Vivo, M. Bridging Quantum Mechanics and Structure-Based Drug Design. *Front. Biosci.* **2011**, *1*, 1619–1633. (c) Mulholland, A. J. Modelling Enzyme Reaction Mechanisms, Specificity and Catalysis. *Drug Discov. Today* **2005**, *10*, 1393–1402.

[30] (a) Lodola, A.; Mor, M.; Rivara, S.; Christov, C.; Tarzia, G.; Piomelli, D.; Mulholland, A. J. Identification of Productive Inhibitor Binding Orientation in Fatty Acid Amide Hydrolase (FAAH) by QM/MM Mechanistic Modelling. *Chem. Commun.* **2008**, 214–216. (b) Lodola, A.; Capoferri, L.; Rivara, S.; Tarzia, G.; Piomelli, D.; Mulholland, A.; Mor, M. Quantum Mechanics/Molecular Mechanics Modeling of Fatty Acid Amide Hydrolase Reactivation Distinguishes Substrate from Irreversible Covalent Inhibitors. *J. Med. Chem.* **2013**, *56*, 2500-2512. c) Van der Kamp, M. W.; Mc Geagh, J. D.; Mulholland, A. J. "Lethal Synthesis" of Fluorocitrate by Citrate Synthase Explained through QM/MM Modeling. *Angew. Chem. Int. Ed.* **2012**, *50*, 10349–10351.

[31] (a) Hermann, J. C.; Hensen, C.; Ridder, L.; Mulholland, A. J.; Höltje, H. D. Mechanisms of Antibiotic Resistance: QM/MM Modeling of the Acylation Reaction of a Class A Beta-Lactamase with Benzylpenicillin. *J. Am. Chem. Soc.* **2005**, *127*, 4454–4465. (b) Gatti, D. L. Biapenem Inactivation by B2 Metallo beta-Lactamases: Energy Landscape of the Post-Hydrolysis Reactions. *PLoS One*, **2012**, *7*, No. e30079. (c) Fonseca, F.; Chudyk, E. I.; van der Kamp, M. W.; Correia, A.; Mulholland, A. J.; Spencer, J. The Basis for Carbapenem Hydrolysis by Class A β-Lactamases: A Combined Investigation using Crystallography and Simulations. *J. Am. Chem. Soc.* **2012**, *134*, 18275–18285. (d) Sgrignani, J.; Grazioso, G.; De Amici, M.; Colombo, G. Inactivation of TEM-1 by Avibactam (NXL-104): Insights from Quantum Mechanics/Molecular Mechanics Metadynamics simulations. *Biochemistry*, **2014,** *53*, 5174-5185.

[32] (a) Olah, J.; Mulholland, A. J.; Harvey, J. N. Understanding the Determinants of Selectivity in Drug Metabolism through Modeling of Dextromethorphan Oxidation by Cytochrome P450. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 6050–6055. (b) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. P450 Enzymes: Their Structure, Reactivity, and Selectivity-Modeled by QM/MM Calculations. *Chem. Rev.* **2010**, *110*, 949–1017.

[33] Seabra, G. M.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package. *J. Phys. Chem A*. **2007**, *111*, 5655-5664.

[34] Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in Free Energy Space. *J. Chem. Phys.* **2007**, *126*, 054103.

[35] Branduardi, D.; De Vivo, M.; Rega, N.; Barone, V.; Cavalli, A. Methyl Phosphate Dianion Hydrolysis in Solution Characterized by Path Collective Variables Coupled with DFT-Based Enhanced Sampling Simulations. *J. Chem. Theory Comput*. **2011**, *7*, 539-543.

[36] Zinovjev, K.; Martí, S.; Tuñón, I. A Collective Coordinate to Obtain Free Energy Profiles for Complex Reactions in Condensed Phases. *J. Chem. Theory Comput*. **2012**, *8*, 1795-1801.

[37] Lodola, A.; Branduardi, D.; Capoferri, L.; De Vivo, M.; Mor, M.; Piomelli, D. A Catalytic Mechanism for Cysteine N-terminal Nucleophile Hydrolases, as Revealed by Free energy simulations. *PLoS One* **2012**, *7*, e32397.

[38] Benesch, R. E.; Benesch, R. The Acid Strenght of the –SH Group in Cysteine and Related Compounds. *J. Am. Chem. Soc*. **1955**, *77*, 5877-5881.

[39] Bednar R. A. Reactivity and pH Dependence of Thiol Conjugation to *N*-ethylmaleimide: Detection of a Conformational Change in Chalcone Isomerase. *Biochemistry*, **1990**, *29*, 3684-3690.

[40] Paasche, A.; Schiller, M.; Schirmeister, T.; Engels, B. Mechanistic Study of the Reaction of Thiol-Containing Enzymes with α,β-Unsaturated Carbonyl Substrates by Computation and Chemoassays. *ChemMedChem* **2010**, *5*, 869–880.

[41] Miyata, O.; Shinada, T.; Ninomiya, I.; Naito, T.; Date, T.; Okamura, K.; Inagaki, S. Stereospecific nucleophilic addition reactions to olefins. Addition of Thiols to α,β-Unsaturated Carboxylic Acid Derivatives. *J. Org. Chem.* **1991**, *56*, 6556-6564.

[42] Yun, C. H.; Boggon, T. J.; Li, Y.; Woo, M. S.; Greulich, H.; Meyerson, M.; Eck, M. J. Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. *Cancer Cell* **2007**, *11*, 217-227.

[43] Yoshikawa, S.; Kukimoto-Niino, M.; Parker, L.; Handa, N.; Terada, T.; Fujimoto, T.; Terazawa, Y.; Wakiyama, M.; Sato, M.; Sano, S.; Kobayashi, T.; Tanaka, T.; Chen, L.; Liu, Z. J.; Wang, B. C.; Shirouzu, M.; Kawa, S.; Semba, K.; Yamamoto, T.; Yokoyama, S. Structural Basis for the Altered Drug Sensitivities of Non-small Cell Lung Cancer-Associated Mutants of Human Epidermal Growth Factor Receptor. *Oncogene* **2013**, *32*, 27-38.

[44] Stamos, J.; Sliwkowski, M. X.; Eigenbrot, C. Structure of the Epidermal Growth Factor Receptor Kinase Domain Alone and in Complex with a 4-Anilinoquinazoline Inhibitor. *J. Biol. Chem.* **2002**, *277*, 46265-46272.

[45] Yasuda, H.; Park, E.; Yun, C. H.; Sng, N. J.; Lucena-Araujo, A. R.; Yeo, W. L.; Huberman, M. S.; Cohen, D. W.; Nakayama, S.; Ishioka, K.; Yamaguchi, N.; Hanna, M.; Oxnard, G. R.; Lathan, C. S.; Moran, T.; Sequist, L. V.; Chaft, J. E.; Riely, G. J.; Arcila, M. E.; Soo, R. A.; Meyerson, M.; Eck, M. J.; Kobayashi, S. S.; Costa, D. B. Structural, Biochemical, and Clinical Characterization of Epidermal Growth Factor Receptor (EGFR) exon 20 Insertion Mutations in Lung Cancer. *Sci. Transl. Med.* **2013**, *5*, 216ra177

[46] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712-725.

[47] Carmi, C.; Cavazzoni, A.; Vezzosi, S.; Bordi, F.; Vacondio, F.; Silva, C.; Rivara, S.; Lodola, A.; Alfieri, R. R.; La Monica, S.; Galetti, M.; Ardizzoni, A.; Petronini, P. G.; Mor, M. Novel Irreversible Epidermal Growth Factor Receptor Inhibitors by Chemical Modulation of the Cysteine-Trap Portion. *J. Med. Chem.* **2010**, *53*, 2038–2050.

[48] Ensing, B.; De Vivo, M.; Liu, Z.; Moore, P.; Klein. M. L. Metadynamics as a Tool for Exploring Free Energy Landscapes of Chemical Reactions. *Acc. Chem. Res.* **2006**, *39*, 73-81.

[49] a) Huskey, S. W.; Huskey, W. P.; Lu, A. Y. H. Contributions of Thiolate "Desolvation" to Catalysis by Glutathione S-transferase Isozymes 1-1 and 2-2: Evidence from Kinetic Solvent Isotope Effects. *J. Am. Chem. Soc.* **1991**, *113*, 2283-2290. b) Ridder, L.; Rietjens, I. M. C. M.; Vervoort, J.; Mulholland, A. J. QM/MM Free Energy Simulation of the Glutathione S-Transferase (M1-1) Reaction with Phenanthrene 9,10-oxide. *J. Am. Chem. Soc.* **2002**, *124*, 9926-9936.

[50] a) Mujika, J. I.; Lopez, X.; Mulholland, A. J. Mechanism of C-terminal Cleavage in Protein Splicing from QM/MM molecular dynamics simulations. *Org. Biomol. Chem.* **2012**, *10*, 1207-1218. b) Xu, D.; Guo, H.; Cui, Q. Antibiotic Deactivation by a Dizinc Beta-Lactamase: Mechanistic Insights from QM/MM and DFT Studies. *J. Am. Chem. Soc.* **2007**, *129*, 10814–10822. c) Woodcock, H. L.; Hodoscek, M.; Brooks, B. R. Exploring SCC-DFTB paths for Mapping QM/MM Reaction mechanisms. *J. Phys. Chem. A.* **2007**, *111*, 5720-5728.

[51] Grazioso, G.; Legnani, L.; Toma, L.; Ettari, R.; Micale, N.; De Micheli C. Mechanism of Falcipain-2 Inhibition by α,β-unsaturated Benzo[1,4]diazepin-2-one Methyl Ester. *J Comput Aided Mol Des.* **2012**, *26*, 1035-1043.

[52] Carmi, C.; Galvani, E.; Vacondio, F.; Rivara, S.; Lodola, A.; Russo, S.; Aiello, S.; Bordi, F.; Costantino, G.; Cavazzoni, A.; Alfieri, R. R.; Ardizzoni, A.; Petronini, P. G.; Mor, M. Irreversible

Inhibition of Epidermal Growth Factor Receptor Activity by 3-aminopropanamides. *J. Med. Chem.* **2012**, *55*, 2251–2264.

[53] Seabra, G. M.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package. *J. Phys. Chem A.* **2007**, *111*, 5655-5664.

[54] Paulsen, C. E.; Truong, T. H.; Garcia, F. J.; Homann, A.; Gupta, V.; Leonard, S. E.; Carroll, K. S. Peroxide-Dependent Sulfenylation of the EGFR Catalytic Site Enhances Kinase Activity. *Nat. Chem. Biol.* **2012**, *8*, 57–64.

[55] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

[56] Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B. P.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. AMBER11, University of California, San Francisco.

[57] Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[58] Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - an N Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[59] Elstner, M. The SCC-DFTB Method and its Application to Biological systems. *Theor. Chem. Acc.* **2006**, *116*, 316–325.

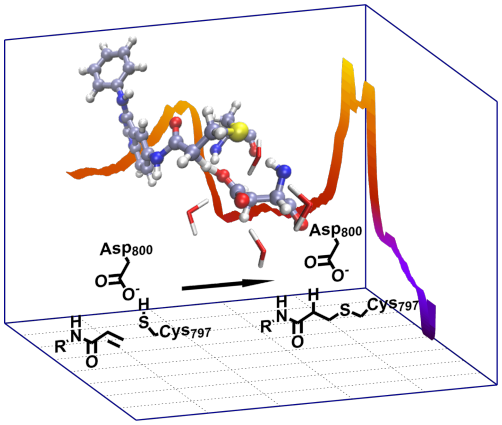
[60] a) Bayse, C. A.; Merz, M. K. Mechanistic Insights into Mg^{2+}-Independent Prenylation by CloQ from Classical Molecular Mechanics and Hybrid Quantum Mechanics/ Molecular Mechanics Molecular Dynamics Simulations. *Biochemistry*, **2014**, *53*, 5034–5041 b) Yang, Y.; Miao, Y.; Wang, B.; Cui, G.; Merz, K. M. Catalytic Mechanism of Aromatic Prenylation by NphB. *Biochemistry*, **2012**, *51*, 2606–2618. c) Capoferri, L.; Mor. M.; Sirirak, J.; Chudyk, E.; Mulholland, A. J.; Lodola A. *J. Mol. Model*. **2011**, *17*, 2375-2383. d) Xu, D. G.; Wu, S. S.; Guo, H. QM/MM Studies of Monozinc beta-Lactamase CphA Suggest That the Crystal Structure of an Enzyme-Intermediate Complex Represents a Minor Pathway. *J. Am. Chem. Soc*. **2010**, *132*, 17986–17988. e) Pierdominici-Sottile, G.; Horenstein, N. A.; Roitberg, A. E. Free Energy Study of the Catalytic Mechanism of Trypanosoma cruzi trans-Sialidase. From the Michaelis Complex to the Covalent Intermediate. *Biochemistry*, **2010**, *50*, 10150–10158. f) Xu, D. G.; Guo, H. Quantum Mechanical/Molecular Mechanical and Density Functional Theory Studies of a Prototypical Zinc Peptidase (Carboxypeptidase A) Suggest a General Acid-General Base Mechanism. *J. Am. Chem. Soc*. **2009**, *131*, 9780–9788.

[61] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem*. **1992**, *13*, 1011-1021.

[62] Grossfield, A, WHAM: the weighted histogram analysis method, version 2.0.6, http://membrane.urmc.rochester.edu/content/wham

[63] Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun*. **2009**, *180*, 1961–1972.

Table of Content Graphic


Asp800
Cys797
Asp800
Cys797