HHS Public Access

Author manuscript

J Chem Inf Model. Author manuscript; available in PMC 2024 February 03.

Published in final edited form as:

J Chem Inf Model. 2023 August 14; 63(15): 4912–4923. doi:10.1021/acs.jcim.3c00720.

Quantum Descriptors for Predicting and Understanding the Structure-Activity Relationships of Michael Acceptor Warheads

Ruibin Liu[†], **Erik A. Vázquez-Montelongo**[†,‡], **Shuhua Ma**[*,¶], **Jana Shen**[*,†]

[†]Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD 21201

[‡]Joint first author

[¶]Department of Chemistry, Jess and Mildred Fisher College of Science and Mathematics, Towson University, Towson, MD 21252

Abstract

Predictive modeling and understanding chemical warhead reactivities have the potential to accelerate targeted covalent drug discovery. Recently, the carbanion formation free energies as well as other ground-state electronic properties from the density functional theory (DFT) calculations have been proposed as predictors of glutathione reactivities of Michael acceptors; however, no clear consensus exists. By profiling the thiol-Michael reactions of a diverse set of singly- and doubly-activated olefins including several model warheads related to afatinib, here we reexamined the question of whether low-cost electronic properties can be used as predictors of reaction barriers. The electronic properties related to the carbanion intermediate were found to be strong predictors, e.g., the change in the Cβ charge accompanying carbanion formation. The least expensive reactant-only properties, the electrophilicity index and the Cβ charge also show strong rank correlations, suggesting their utility as quantum descriptors. A second objective of the work is to clarify the effect of the β-dimethylaminomethyl (DMAM) substitution which is incorporated in the warheads of several FDA-approved covalent drugs. Our data suggests that the β-DMAM substitution is cationic at neutral pH in solution and it promotes acrylamide's intrinsic reactivity by enhancing the charge accumulation at Cα upon carbanion formation. In contrast, the inductive effect of the β-trimethylaminomethyl substitution is diminished due to steric hindrance. Together, these results reconcile the current views of the intrinsic reactivities of acrylamides and contribute

* sma@towson.edu; jana.shen@rx.umaryland.edu.

Author Contributions

R. L., E. A. V-M., and S. M. carried out the calculations. R. L., E. A. V-M., S. M., and J. S. performed the analysis. R. L., S. M., and J. S. wrote the manuscript.

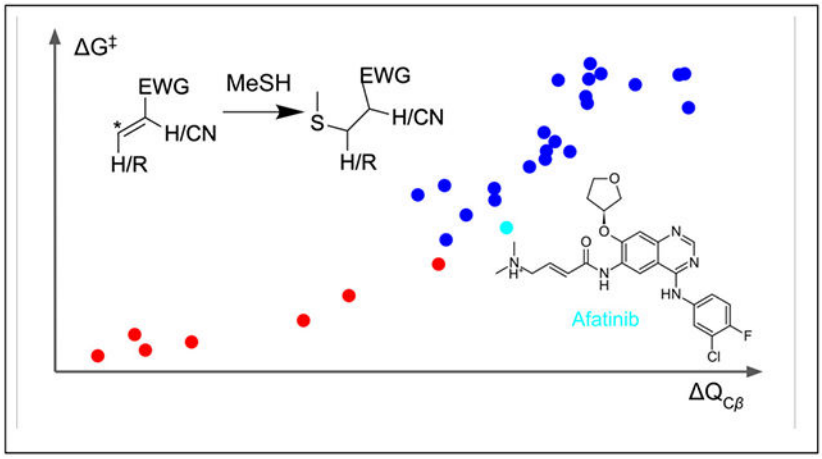
Supporting Information

Table S1 contains the comparison calculations using different GSH models. Table S2 contains the calculated reaction barriers and experimental rate constants for nine olefins. Table S3 and S4 contain the correlations between activation energies and other electronic properties. Table S5 contains the geometry details for 20, 27, 27H, and 28 in the carbanion state. A supplemental Excel sheet is provided that contains all calculated energies, natural charges, as well as other electronic properties. A supplemental Word file is provided that contains the Cartesian coordinates of the optimized structures for MeS^-, MeSH, olefins, transition states, carbanions, and thiol adducts. The single point energy, imaginary frequency, internal energy, enthalpy, and Gibbs free energy for each structure are included following the Cartesian coordinates.

The authors declare the following competing financial interest(s): J.S. is a founder and scientific advisor of ComputChem and serves as a Scientific Advisory Board member of MatchPoint Therapeutics.

to large-scale predictive modeling and understanding of the structure-activity relationships of Michael acceptors for rational TCI design.

Graphical Abstract



INTRODUCTION

In recent years, targeted covalent inhibitors (TCIs) have gained considerable interest.[1,2] The number of the U.S. Food and Drug Administration (FDA) approved TCIs against kinases for treatment of cancer, inflammatory and other diseases are steadily increasing,[3,4] and TCIs against previously undruggable targets (i.e., K-RAS^{G12C}) are now advancing to the clinical stage.[5] TCIs are typically designed by adding a bond-forming electrophilic group (also known as a warhead) to a traditional reversible inhibitor such that following the reversible binding step, the warhead is positioned to react with a specific nucleophilic residue at the target site.[2] To reduce off-target effects, an ideal warhead may have low but sufficient reactivity. In more recent years, TCIs that form reversible covalent bond have also been actively pursued.[6-9] Developing computational protocols that can make accurate, reliable, and fast prediction of reactivities and reversibilities of covalent warheads can significantly reduce medicinal chemistry efforts.

Most FDA-approved TCIs make use of warheads based on an acrylamide group, which is an olefin activated by an electron-withdrawing group (EWG, see Fig. 1a-b) at the α carbon atom. An activated olefin reacts with a cysteine via Michael addition (Scheme 1), which has been extensively studied since the 1960's.[10-14] Although alternative mechanisms have been proposed,[15] the rate-limiting step of the reaction is believed to be the nucleophilic attack of the deprotonated thiolate (either directly or via a base catalyst) on the electron-deficient β carbon, leading to a carbanion intermediate, which protonates to form the thioether adduct (Scheme 1). Thus, the intrinsic reaction rate of a thiol-Michael addition (for protein targets) is determined by the reactivity of the cysteine thiol and that of the Michael acceptor. As the thiol reactivity is related to the availability of the nucleophilic thiolate,[16,17] accurate pK_a prediction tools such as the GPU-accelerated continuous constant pH molecular dynamics simulations[18] can be used to assess cysteine reactivities in drug targets.[19,20]

In principle, the intrinsic reactivity of a Michael acceptor can be predicted by quantum mechanical (QM) calculation of the activation free energy ($\Delta G^{\ddagger}$) of the carbanion formation step in the Michael addition reaction (Scheme 1). Indeed, a number of density functional theory (DFT) studies confirmed that the calculated $\Delta G^{\ddagger}$ and $\Delta E_a^{\ddagger}$ (activation energy) of the acrylamide (or propynamide) reactions with methylthiolate are strongly correlated with the experimental half-lives of the glutathione (GSH) adduct formation.[21-26] In comparison to the available experimental data, Krenske and Houk demonstrated[27] that the substitution effects on the kinetic rate constants of of acrylamide thiol-adduct formation can be quantitatively recapitulated using the $\Delta G^{\ddagger}$ values calculated by the M06-2X DFT method.[28]

Despite the predictive value of DFT-calculated $\Delta G^{\ddagger}$ for warhead reactivities, it requires transition-state searches, which is computationally expensive and may not always be successful. Therefore, ground-state electronic properties are more desirable. Rowley and coworker proposed[30] that the rate of the Michael addition is proportional to the stability or formation free energy of the carbanion intermediate ΔG_{CF} and developed an automated protocol to establish a database of calculated ΔG_{CF} for singly-activated olefins. However, Krenske and Houk's B3LYP DFT calculations of six enones[31] showed a modest correlation between $\Delta G^{\ddagger}$ and ΔG_{CF}. Besides ΔG_{CF}, other ground-state electronic properties have also been proposed as predictors of thiol-Michael addition rates, e.g., the HOMO/LUMO energies,[32] the electrophilicity index (ω),[33,34] the overall reaction energy (ΔG_{rxn}),[24,26] and proton affinity of the carbanion intermediate (ΔG_{PA}).[35] A recent study[36] found that ΔG_{rxn} and ΔG_{PA} do not correlate with the experimental GSH half-life times of drug-like, singly activated acrylamides, although a caveat is that the B3LYP functional used in the study may not produce stable carbanion intermediates in the gas phase (see later discussion).[37] Thus, taking all these studies together, it remains unclear as to what ground-state electronic properties are reliable predictors of Michael acceptor reactivities. Furthermore, none of the aforementioned DFT studies considered both singly- and doubly-activated olefins or the correlation of a large number of electronic properties (including ΔG_{CF} and ΔG_{PA}) with $\Delta G^{\ddagger}$. In a tour-de-force study, Hermann, Weber, and coworkers trained machine learning models for predicting GSH reactivities based on over 20,000 transition-state calculations for over 4,000 singly-activated olefins.[25] Identifying ground-state surrogates of transition-state energies would allow us to drastically increase the training dataset.

When developing TCIs of EGFR and related kinases, Tsou et al. observed that substituting the Cβ of an acrylamide warhead with a dimethylaminomethyl (DMAM) group not only improves solubility but also results in a greater reactivity and biological activity.[38,39] Using competitive reactions of pairs of TCIs that differ only in the β-DMAM substitution, they found that the β-DMAM-substituted TCIs showed a higher percentage of glutathione adduct in solution and lower IC50 values in EGFR assays as compared to the corresponding unsubstituted TCIs.[38,39] Realization of this important structure-activity relationship (SAR) inspired the design of afatinib, neratinib, and dacomitinib (Fig. 1c). By comparing the glutathione (GSH) adduct formation rates of the β-DMAM and β-trimethylaminomethyl (TMAM) substituted acrylamides, Tsou and Wissner hypothesized that the activity enhancement by the former is due to the base catalyst effect, i.e., the amino group

serves as a general base that abstracts the proton from EGFR's C797 in the front pocket (which is the most popular covalent modification site among all kinases).[19] Recently, using DFT calculations Birkholz et al. suggested that the enhanced reactivity of the β-DMAM substituted acrylamides is due to an induction effect which lowers the reaction barrier.[40] However, due to the assumptions made in regards to the conformation and protonation states of the acrylamide,[40] the mechanism remains murky, and importantly the study did not explain why despite having an inductive effect, the β-TMAM substitution significantly diminishes the reactivity.[40]

To address the aforementioned controversies, the present work pursues two objectives. Using the long-range corrected meta-GGA DFT calculations of a diverse set of 30 singly- and doubly-activated olefins, we reexamined the question of whether ΔG_{CF}, ΔG_{PA}, and/or possibly other "low-cost" electronic properties can be used as predictors of $\Delta G^{\ddagger}$. Among the examined olefins are several acrylamides related to afatinib and several cyanoacrylamides in light of the recent interest in reversible warheads.[6-8,41] The second objective of this work is to explain the observed reactivity enhancement effect of the β-DMAM substitution of the acrylamide. Our data suggested that while both ΔG_{CF} and ΔG_{PA} correlate well with the reaction barriers of Michael additions, the lower-cost, olefin-only properties are strong predictors as well. Our calculations demonstrated that the β-DMAM group is protonated at neutral pH and it lowers the reaction barrier through an induction effect while a β-TMAM group that carries a permanent charge increases the reaction barrier due to steric hindrance.

RESULTS and DISCUSSION

Overview of the dataset and methods.

To study the intrinsic reactivities of activated olefins, we calculated the complete reaction profile, including the geometries and energies of the reactants, transition state, carbanion intermediate, and thiol adduct of the Michael addition reaction (Scheme 1) for 7 doubly- (**1–7**) and 23 (**8–30**) singly-activated olefins (Table 1). Olefins **1, 2, 3, 5–12, and 14** have been previously calculated by others.[6,27,31,35,42] The acrylamides **15** and **16H** are the truncated forms of afatinib, while **17** and **18H** are their respective isomers and **29** is a deactivated form with a α-methyl substitution inspired by the work of Tsou et al.[38] Starting from these olefins, additional ones taken from other literature (e.g., Ref.[22]) were added to fill in the gaps in the $\Delta G^{\ddagger}$ data range. Together, the $\Delta G^{\ddagger}$ values range from about 7 to 23 kcal/mol (Fig. S1), corresponding to a rate constant range of $7.3 \times 10^{-5} s^{-1}$ to $4.5 \times 10^{7} s^{-1}$. Following the work of others,[22,24,27,30,31,37] methanethiolate was used as the model reactive thiolate. Note, glutathione (GSH) is commonly used in the experimental studies of the intrinsic reactivities of warheads.[24,38] Test calculations showed that the gas-phase energetics of the Michael addition with methanethiol is highly correlated with larger models of GSH, such as β-mercaptoethanol or cysteamine (Table S1); however, the TS optimization with the larger GSH models are difficult. For these reasons, we proceeded with methanethiol.

All calculations employed the ωB97X-D3(BJ) / 6-311 + G(d, p) DFT method[43-45] with the Solvent Model Density (SMD) universal solvation model.[46] ωB97X-D3(BJ) treats the

dispersion effect with the Becke–Johnson damping[47] based on the ωB97X-D method,[43] which is a long-range corrected meta-GGA functional that has been demonstrated to offer very accurate isomerization, barrier height, difficult thermochemistry, and noncovalent energies in a large benchmark study.[48] Rowley and coworker found that the popular PBE and B3LYP methods cannot predict stable carbanion intermediates in the gas phase, whereas the ωB97X-D functional can reproduce the high-level CCSD(T) and MP2 geometries and energetics.[37] All species along the reaction profiles were subject to geometry optimization and frequency analysis. The latter was used to confirm the stationary points (energy minima or transition states) and to obtain zero-point energy for thermal and entropic corrections. The activation barriers and carbanion formation free energies were calculated with methylthiolate, while the reaction free energies were calculated with methanethiol. The proton affinities were calculated using a proton Gibbs free energy of −270.3 kcal/mol.[29] The calculated activation and reaction free energies are listed in Table 2 and the complete dataset of the calculated energies and geometries are given in the supplemental files.

ωB97X-D3(BJ)-calculated reaction barriers are highly predictive of the glutathione reactivities of olefins.

To further support the choice of the DFT functional ωB97X-D3(BJ), we compared the calculated activation free energies ($\Delta G^{\ddagger}$) of nine olefins with the available log-transformed pseudo first-order rate constants of glutathione adduct formation in solution. The r^2 value is 0.76 (Fig. 2), which supports the consensus that the DFT-calculated reaction barriers are highly predictive of the glutathione reactivities of olefins.[21-23,25,27,31] We note, the agreement is not perfect due to the limitations of the model thiol (methanethiol used to mimic glutathione) and the implicit-solvent model. In the latter case, the neglect of potential hydrogen bonding with explicit water may introduce an error for the reaction barrier.[49]

The reaction barrier is correlated with the carbanion formation free energy or single-point energy as well as with the proton affinity.

To examine if the free energies of carbanion formation (ΔG_{CF}) or proton affinity (ΔG_{PA}) is a good predictor of $\Delta G^{\ddagger}$, we discuss the Spearman's rank correlation coefficient ρ, as it may be most relevant in the inhibitor discovery campaigns. However, the Pearson's r and Kendall's τ values were also calculated and given in Table S2 and S3 for comparison and completeness. Note, negative values of ΔG_{PA} are used to obtain positive correlation. Interestingly, ΔG_{CF} or ΔG_{PA} show similarly strong correlation with $\Delta G^{\ddagger}$. The ρ values for ΔG_{CF} and $-\Delta G_{PA}$ of the singly-activated olefins are 0.73 and 0.77, respectively, while the corresponding ρ values including both singly- and doubly-activated olefins are 0.87 and 0.89, respectively (Fig. 3). The similarity in correlation is also evident based on the Pearson's r and Kendall's τ values (Table S3). The increased correlation with inclusion of the doubly-activated olefins is due to the nearly doubled data range. While Rowley and coworkers hypothesized that ΔG_{CF} can be used to predict $\Delta G^{\ddagger}$ of singly-activated olefins,[30] and Taunton et al. found that ΔG_{PA} correlates well with $\Delta G^{\ddagger}$ for doubly-activated olefins,[35] our data suggests that both ΔG_{CF} and ΔG_{PA} can be used as a reactivity predictor for both

singly- and doubly-activated olefins. This finding also implies that the carbanion INT is a good surrogate of the TS.

Calculation of ΔG_{CF} and $-\Delta G_{PA}$ requires frequency analysis to obtain thermal energy, which is computationally expensive, especially for a large number of compounds. Thus, we tested the correlation between $\Delta G^{\ddagger}$ and the single point energies, ΔE_{CF} and $-\Delta E_{PA}$, which require about 20% of the calculation time for free energies (Fig. 3). For the singly-activated olefins, the Spearman's ρ values are 0.70 and 0.76, respectively, while for all olefins the ρ values are 0.85 and 0.89, respectively (Fig. 3). These values are similar to those of the corresponding free energies, demonstrating that the lower-cost ΔE_{CF} and ΔE_{PA} can instead be used as reactivity predictors for activated olefins.

The reaction barrier is correlated with the electrophilicity index and its approximate form.

Although calculation of ΔE_{CF} and ΔE_{PA} requires less than 2% of the time spent on the TS calculation, electronic properties that depend only on the reactant olefin require even less computational time (at most 50% relative to ΔE_{CF} or ΔE_{PA}). Thus, we examined the correlation between $\Delta G^{\ddagger}$ and many olefin-only electronic properties (Table S3 and S4). The electrophilicity index (ω) measures the energy stabilization of a small molecule as it acquires electron density and it can be calculated from the ionization potential and electron affinity (Methods).[50,51] An activated olefin is electron deficient at C_{β} which acquires electron density from a nucleophile during the Michael addition. Therefore, the extent of the stabilization through acquiring electron density may be predictive of the olefin's reactivity towards the nucleophile. An early study found that the ω values of a number of electrophiles are well correlated with the reaction rates between the electrophiles and the zinc finger thiolates in the HIV-1 nucleocapsid protein.[50] Another study found a strong correlation between ω's of the substituted ethylenes and the activation barriers of the Diels-Alder reactions.[52] Here for the singly-activated and all olefins, the Spearman's ρ between $\Delta G^{\ddagger}$ and $-\omega$ are 0.87 and 0.71, respectively, similar to the corresponding ρ's for ΔG_{CF} (Fig. 3). Intuitively, the gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) energies is related to electrophilicity. Thus, it is not surprisingly that ω', which is an approximation of ω based on the HOMO-LUMO energy gap correlate with $\Delta G^{\ddagger}$ with similar Spearman's ρ's as ω (Fig. 3); the advantage is that ω' requires less computational time as the extra calculations of ionization potential and electron affinity are no longer needed. Palazzesi et al. found that ω' is a strong predictor of the experimental GSH adduct formation time for singly-activated terminal acrylamides.[33] Our data suggests that ω' can be used for predicting the reactivities of the doubly-activated olefins as well. Hughes et al. included HOMO and LUMO energies as descriptors in building the neural network models for predicting GSH reactivities of small molecules.[32] The Spearman's ρ between $\Delta G^{\ddagger}$ and the LUMO energy is 0.72 for all olefins and 0.57 for the singly-activated olefins. The Spearman's ρ's for the HOMO energy are 0.30 for all olefins and 0.33 for the singly-activated olefins. Thus, our data do not support the direct use of the HOMO and LUMO energies as descriptors.

Recently, Hermann et al. examined a large dataset of singly-activated acrylamides and found that the correlation between ω' and the reaction barrier can be greatly improved by using the frontier orbitals located on the warhead.[34] This insight may be used to improve the predictions for olefins with larger substituent. To test this, we recalculated ω' for the afatinib-related acrylamides **30** and **17**, which show large deviations from the prediction trend (Fig. 3f). Indeed, ρ is increased from 0.87 to 0.88; r and τ are also slightly improved (Fig. S2). Thus, our data lends further support to the use of warhead associated frontier orbitals in calculating ω' as a strong reactivity predictor.[34]

The reaction barrier is correlated with the C_β natural charge and highly correlated with its change upon carbanion formation.

Since electrophilicity can also be understood from the charge perspective, we calculated the atomic partial charges based on natural population analysis.[53] Due to the electron-withdrawing effect of the α-substituent, the negative charge is shifted from the C_β to the C_α atom in an activated olefin. This polarization effect is more pronounced in solution as compared to the gas phase (data not shown). Take the doubly-activated olefins **1** and **2** as an example. The C_β charge is nearly zero (compared to the slightly negative values in other olefins), while the reaction barriers are extremely low (7.0 and 7.3 kcal/mol, respectively, Table 2 and Fig. S1). Therefore, we examined the correlation between $\Delta G^{\ddagger}$ and the C_β charge ($Q_{C\beta}$) of the olefin. The Spearman's ρ is 0.77 for the singly-activated and 0.84 for all olefins (Fig. 3h & Table S1), similar to those of ω'. We note, the data spread for the singly-activated olefins suggests a weak linear correlation, which is confirmed by a small Pearson's r value of 0.55 (Table S3); however, Spearman's ρ which measures the rank-order correlation is strong in this case. Interestingly, $\Delta Q_{C\beta}$, the change in the C_β charge going from the olefin to the carbanion state shows an even greater correlation with $\Delta G^{\ddagger}$, with the ρ values of 0.94 and 0.87 for singly-activated and all olefins, respectively (Fig. 3). Out of all ground-state electronic properties considered here, $\Delta Q_{C\beta}$ shows the strongest correlation with $\Delta G^{\ddagger}$. We also examined the correlation of the reaction barrier with the C_α charge in the olefin ($Q_{C\alpha}$) or its change upon carbanion formation ($\Delta Q_{C\alpha}$). With $Q_{C\alpha}$, there is a modest correlation for the singly-activated olefins (ρ is 0.59) but there is no correlation if the doubly-activated olefins are included (ρ is 0.02, Table S2). With $\Delta Q_{C\alpha}$, there is nearly no correlation with the reaction barrier. Given the charge accumulation at the C_α, the lack of correlation with $Q_{C\alpha}$ or $\Delta Q_{C\alpha}$ seems puzzling and we will come back to this point in the later discussion.

The α-cyano substitution decreases both the reaction barrier and the reaction free energy.

The experiments of Taunton and coworkers showed that adding nitrile as a second EWG at the Cα of an acrylamide or acrylate accelerates the thiol adduct formation and makes it reversible at the same time.[6,35] Our data are consistent with this observation. Specifically, the $\Delta G^{\ddagger}$ values (7.0–11.7 kcal/mol) of the α-cyano-substituted doubly-activated olefins (**1–7**) are lower than those (12.9–23.1 kcal/mol) of the singly-activated olefins (**8–29**), demonstrating that adding a cyano group increases the reaction rate. The magnitude of the $\Delta G^{\ddagger}$ decrease varies depending on the other EWG (Table 2 and Fig. S1). Comparison

of **9** with **3** and **11** with **1** indicates that the reduction in $\Delta G^{\ddagger}$ for the acrylates (R_2 is a methyl ester group) is about 7–8 kcal/mol, while comparison of **14** with **4** indicates a reduction of about 10 kcal/mol for the acrylamides (R_2 is an amide group). Consistent with the experimental observation that the α-cyano substitution promotes reversibility,[6,35] the doubly-activated olefins **1–7** have a less negative ΔG_{rxn} (−5 to −9 kcal/mol) than the singly-activated olefins **8, 9, 10, and 13–30**, which have ΔG_{Rxn} of −10 to −12 kcal/mol. The two exceptions are the singly-activated olefins **11** and **12**, which have ΔG_{Rxn} of −7.6 and −8.6 kcal/mol, respectively. This can be explained by the β-phenyl substitution, which makes the thiol adduct less stable, as evident when comparing **11** and **12** with the corresponding singly-activated analogs **9** and **10**, which lack of β substitution and have ΔG_{Rxn} of −11.1 and −12.5 kcal/mol, respectively. The reduction of ΔG_{Rxn} due to the β-phenyl substitution is also evident for the doubly-activated olefins when comparing **1** with **3** and **2** with **4**. In these cases, the magnitude of ΔG_{Rxn} is reduced by 3.5 and 2.1 kcal/mol, respectively, in the presence of the β-phenyl substitution.

Adding a protonated β-dimethylaminomethyl (DMAM) substitution lowers the reaction barrier of acrylamide.

Tsou, Wissner, et al. found that a β-DMAM substitution enhances while a β-trimethylaminomethyl (TMAM) substitution diminishes the reactivity of the acrylamides. To explain this, they proposed a base-catalyzed mechanism,[38,39] in which the neutral, unprotonated DMAM group serves as a general base by extracting the proton from the sulfhydrl group, which allows a concerted nucleophilic attack of the thiolate on the β-carbon through a five-membered TS (Fig. 4). Recently, this concerted mechanism was disputed by Birkholz et al. based on the DFT calculations,[40] although the mechanism of reactivity enhancement by the DMAM but not the TMAM substitution still remains unclear.

To delineate the effect of β-DMAM substitution, we first note that the controversy[38-40] is largely due to the ambiguity about its protonation state. In the study of Tsou, Wissner et al.,[38,39] the DMAM group was considered deprotonated while in the work of Birkholz et al. both deprotonated and protonated forms were considered.[40] However, tertiary amines are expected to be protonated and charged at neutral or physiological pH, as their pK_a are highly than 8.5, e.g., the experimental pK_a is 9.76 for trimethyl amine and 10.08 for N-methyl piperidine.[54] Consistently, the DMAM group pK_a's in **16**, **18**, and **27** were estimated as 8.64, 8.62, and 8.69 using Schrödinger's Epikx program.[55] Therefore, to emphasize the protonated form, we denote these compounds as **16H**, **18H**, and **27H** in Table 1. Next, we compared the reaction barriers of **15–18**. **15** and **16H** are the truncated forms of afatinib (Fig. 1), while **17** and **18H** are the respective isoforms. Adding the protonated β-DMAM group transforms **15** to **16H** and **17** to **18H**. From our calculations, $\Delta G^{\ddagger}$ decreases by 3.5 kcal/mol from **15** to **16H** and by 1.1 kcal/mol from **17** to **18H** (Table 3), confirming that a β-substitution with a protonated DMAM group lowers the reaction barrier of acrylamide, which corroborates the calculations of Birkholz et al.[40] and provides an explanation for the experimental observation that a higher percentage of the β-DMAM substituted TCIs form glutathione adducts in solution and they are more potent in EGFR as compared to

the corresponding unsubstituted TCIs.[38,39] Consistent with the reduction in $\Delta G^{\ddagger}$, ΔG_{CF} and $-\Delta G_{PA}$ are decreased by 5.0 and 4.8 kcal/mol respectively from substituting **15** with the protonated β-DMAM. Similar decreases (4.4 and 4.9 kcal/mol) are seen from substituting **17** with the protonated β-DMAM (Table 3). **20** and **27H** are the truncated forms of **15/17** and **16H/18H**, respectively. $\Delta G^{\ddagger}$, ΔG_{CF}, and $-\Delta G_{PA}$ going from the unsubstituted **20** to the β-DMAM-substituted **27H** decrease by 5.2, 7.3, 6.2 kcal/mol, respectively (Table 3), confirming the barrier-lowering effect of the β-DMAM substitution again. It is also noteworthy that $\Delta G^{\ddagger}$ decreases from **27H** to **18H** by 2.3 kcal/mol and from **18H** to **16H** by another 2.0 kcal/mol, which demonstrates the electron withdrawing effect of the quinazoline group (as in afatinib) on the acrylamide.

Comparison between the effects of the protonated and deprotonated β-DMAM substitution.

Although the protonated (charged) form of DMAM is predominant at neutral pH, for comparison we also calculated the reaction energetics of the deprotonated (neutral) **16**, **18**, and **27**. In contrast to the charged β-DMAM substitution, adding the neutral β-DMAM substitution decreases the acrylamide's reactivity as shown by the $\Delta G^{\ddagger}$ increase of 1.9 kcal/mol going from **15** to **16** and 1.04 kcal/mol from **17** to **18** (Table 3). Consistent with the reaction barrier increase, the corresponding ΔG_{CF} and ΔG_{PA} are also increased (Table 3). Since the β-DMAM substitution is known to significantly decrease the GSH adduct formation time,[38-40,40] these data reinforce the notion that DMAM is charged at neutral pH. Furthermore, our comparison of the charged and neutral β-DMAM substitution is consistent with the data of Birkholz et al.[40] and supports their hypothesis of a charge induction mechanism for the enhanced reactivity by the charged β-DMAM substitution.

Why does the protonated β-DMAM substitution lower the reaction barrier?

To understand why the charged β-DMAM lowers the reaction barrier relative to the neutral counterpart as well as the β-trimethylaminomethyl (TMAM) substitution with a permanent charge, we closely examined the reaction profiles of the smallest model acrylamides with the deprotonated β-DMAM substitution (**27**), with the protonated β-DMAM substitution (**27H**), and with the β-TMAM substitution (**28**). We first tested the five-membered base catalyzed mechanism (Fig. 4).[38-40] In the TS structures of the three molecules, the amino nitrogen is over 4.3 Å away from the thiolate sulfur, and for **27H** the closest distance between the amino hydrogen and the sulfur is 4.74 Å, which indicates that the five-membered mechanism (Fig. 4)[38-40] is not possible. Comparing **27H** with **28**, we can see that the reaction barrier is raised by about 2 kcal/mol in the latter (Fig. 4), which is consistent with the data of Tsou et al.[38] The same trend can be seen with ΔG_{CF} and $-\Delta G_{PA}$, which are respectively increased by 5.1 and 2.8 kcal/mol going from **27H** to **28** (Fig. 4). These data are consistent with the significantly prolonged GSH lifetimes of the β-TMAM vs. β-DMAM substituted acrylamides.[38]

To explain why the barrier of **27H** is lower than **28** carrying a permanent charge, neutral **27**, or the unsubstituted **20**, we tested the hypothesis of an inductive mechanism by examining the changes of the C_β and C_α charges upon carbanion formation. Comparing **27** and **20**, ΔC_β

is less negative in **20** and it has a smaller $\Delta G^{\ddagger}$ (Fig. 4), which is in agreement with the belief that the more electron withdrawing the lower the reaction barrier and consistent with the overall correlation shown in Fig. 3. However, the decrease in the barriers of **27H** and **28** relative to **20** is inconsistent with the more negative ΔC_{β}. We attribute this to the positive charge of the β-substituent. We found that $\Delta Q_{C\alpha}$ is most negative for **27H** (-0.33 *e*) followed by **28** (-0.30 *e*) and **27** (-0.24 *e*), which suggests that the protonated β-DMAM group induces the largest charge accumulation at the C_{α}. Importantly, the decreasing order of $\Delta G^{\ddagger}$ for **27**, **28**, **27H** corresponds to the increasing order of $-\Delta Q_{Ca}$, which supports the hypothesis that the inductive effect drives the reactivity.[40]

Next, we address the question of why the inductive effect with **28** is smaller than **27H**. To explain this, we noticed some key differences in their TS geometries. In the TS of **27H**, the distance between the amino nitrogen and the Cα atom is shorter by 0.2 Å, which is consistent with the 3.4° smaller N-Cβ-Cγ angle as compared to the TS of **28** (Fig. 4). These differences indicate that the positively charged amino nitrogen is closer to the charge center C_{α} in **27H**, which increases the induction effect and stabilization of the carbanion state through electrostatic interaction. The TS geometry differences between **27H** and **28** may be explained by the extra methyl group which prevents the amino group from moving closer to the Cα atom in **28**. Comparing the TS of **27H** with **27**, we can see a similar trend with regards to the N-C_{α} distance and N-Cβ-Cγ angle. This is not surprising, as it reflects the lack of the electrostatic attraction between the neutral amino group and the Cα. It is noteworthy that the N-C_{α} distance in the TS of **27** is 0.09 Å shorter and the N-C_{β}-Cγ angle is 0.7° smaller compared to the TS of **28** (Fig. 4), which supports the aforementioned hypothesis that the bulky methyl group in **28** pushes the amino nitrogen a bit further from the charge center. We note, the carbanion-state geometries of **27H**, **27**, and **28** follow similar trends (Table S3), consistent with the notion that the carbanion state is a surrogate of the TS.

The charged β-DMAM group contributes to deprotonation and nucleophilicity increase of the thiol.

Tsou, Wissner et al. proposed that[38,39] the thiol (i.e., C797 in EGFR) is deprotonated (activated) by transferring its proton to the neutral β-DMAM group. This mechanism is incompatible not only with the pK_a of the β-DMAM group but also with the TS and carbanion geometries of **27H**, which show that the distance between the amine nitrogen and the incoming sulfur is 4.4 and 4.1 Å, respectively, much larger than that is required for proton transfer. However, these nitrogen–sulfur distances are sufficient for the charged amine to deprotonate the thiol by stabilizing its negatively charged thiolate state. In other words, the charged amine would downshift the pK_a of the thiol and increase its nucleophilicity. This may explain in part why in order to target EGFR's C797, which was found unreactive in the apo protein based on the recent constant pH molecular dynamics simulations,[19] adding a charged β-DMAM substitution is an effective strategy. Tsou, Wissner et al. showed that adding a β-DMAM substitution significantly lowers the IC50 values TCIs of EGFR,[38,39] which inspired the development of the FDA-approved drugs afatinib, neratinib and dacomitinib (Fig. 1c). We note, the warheads of more recent EGFR

TCIs such as osimertinib do not carry the β-DMAM substitution. In this case, the reaction may proceed via an alternative mechanism, e.g., C797 may be deprotonated by a different amine group (e.g., a pyrrolidine substituent) near the warhead[56] or the nearby D800 may serve as a general base independently.[57]

Concluding Discussion

reveal the structure-activity relationships (SARs) of the covalent warheads. Towards this goal, we obtained the complete reaction profiles of 30 singly- and doubly-activated olefins including several acrylamides related to the cancer drug afatinib using the ωB97X-D3(BJ) DFT calculations. The data evenly spread over a range of 7–23 kcal/mol, corresponding to the rate constants $7.3 \cdot 10^{-5}$ to $4.5 \cdot 10^{7}$ s^{-1}. We found that the reaction barrier ($\Delta G^{\ddagger}$) has similarly strong correlations with the carbanion formation free energy (ΔG_{CF}) and carbanion proton affinity (ΔG_{PA}), with Spearman's ρ of 0.87/0.89 for all and 0.73/0.77 for the singly-activated olefins. A similar level of correlation was also observed for the single-point energies of carbanion formation (ΔE_{CF}) and proton affinity (ΔE_{PA}), which require 20% of the computational time relative to the corresponding free energies. Even stronger correlation (Pearson's ρ of 0.94 for all and 0.87 for the singly-activated olefins) was found for the change in the Cβ natural charge ($\Delta Q_{C\beta}$) accompanying the carbanion formation, which requires a similar computational cost as ΔE_{CF} or ΔE_{PA}.

Interestingly, without calculation of the carbanion intermediate, three olefin-based electronic properties show similar levels of correlation (based on the Spearman's ρ) as ΔE_{CF} and ΔE_{CF}. They are the electrophilicity index (ω) based on the ionization potential and electron affinity, the approximate electrophilicity index (ω') based on the HOMO/LUMO energies, and the C_{β} charge ($Q_{C\beta}$). Calculation of these properties only requires geometry optimization and energy calculation of the olefin, which reduces the computational cost by at least 10 times relative to ΔG_{CF} or ΔG_{PA} and at least 100 times relative to $\Delta G^{\ddagger}$; thus they may serve as lower-cost reactivity descriptors in large-scale computational studies of the SARSs of TCI warheads.

Adding a β-dimethylaminomethyl (DMAM) substitution on the acrylamide warhead of EGFR TCIs significantly increases the thiol-adduct formation in solution and lowers the IC50 value in EGFR;[38,39] this important SAR inspired the development of several FDA-approved drug, e.g., afatinib, neratinib, and dacomitinib. Our data unambiguously demonstrated that the β-DMAM group is protonated (positively charged) in solution and it substantially lowers the Michael reaction barrier through induction of negative charge at the Cα atom in the TS and the carbanion INT. Although the inductive effect is also at play with the β-trimethylaminomethyl (TMAM) substitution which carries a permanent charge, the effect is diminished due to steric hindrance by the extra methyl group such that the amino group is further away from the C_{α} in the TS and INT relative to the β-DMAM substituted acrylamide. The TS and INT structures of the model acrylamides refuted the five-membered base catalyzed mechanism,[38,39] and instead suggested that the protonated DMAM downshifts the pK_a of the incoming thiol through stabilization of the deprotonated state and thereby activating it for the nucleophilic attack on the warhead. Together, the mechanism uncovered by our calculations and analysis suggests that the β-DMAM

substitution enhances the intrinsic electrophilicity of the warhead and the nucleophilicity of the cysteine. This may explain why adding a β-DMAM substitution is an important strategy for liganding EGFR C797, which was found protonated and unreactive in our recent constant pH molecular dynamics study,[19] although alternative mechanisms to activate C797 have also been proposed, e.g., an alternative amine group (e.g., a pyrrolidine substituent) near the warhead[56] or the nearby D800 independently serving as a general base.[57]

The present study has two major caveats. First, the dataset is rather small and especially the variation of the β-substitution is very limited. Our analysis of the β-DMAM substitution and related acrylamides showed that the change in the C_α charge upon carbanion formation (ΔQ_{C_α}) is an important reporter of the inductive effect of β substitution. Thus, we suggest that ΔQ_{C_α} and Q_{C_α} be included in the reactivity descriptor set for building models. A second caveat is with regards to the β-DMAM substitution effect on the Michael addition mechanism for the EGFR's C797. Our data presents strong evidence for the β-DMAM group to deprotonate the thiol in solution; however, in EGFR C797 is in the vicinity of D800 which may serve as a proton acceptor to further accelerate the reaction. A similar mechanism has been proposed for the pyrrolidine substituent to deprotonate C797 in EGFR.[56] We should note that in the absence of an amine group near the warhead, D800 may independently serve as a proton acceptor. In fact, a previous QM/MM study showed that C797-S^-/D800-COOH is the most stable and abundant ionization state for this pair within EGFR.[57] To fully understand the reaction mechanism of the β-DMAM substituted TCIs in EGFR, QM/MM simulations are underway in our group. Despite these caveats, our findings reconciled current views regarding the reaction energetics and important SARs of the thiol Michael additions. Although the present work is focused on intrinsic reactivities based on quantum calculations, the descriptors can be used in training machine learning models based on the experimental kinetics data. Our analysis offers valuable insights for predictive modeling and understanding of warhead SARs to assist covalent drug discovery.

METHODS and PROTOCOLS

Electronic structure calculations.

All electronic structure calculations were carried out using ORCA 5.0.1 software package.[58] The ωB97X-D3(BJ) density functional theory (DFT) method[43,44,47] with the 6-311+G(d,p) basis set and the Solvent Model Density (SMD) continuum solvent model[46] was used for both geometry optimization and electronic energy calculations. ωB97X-D3(BJ) is an extension of the ωB97X-D functional,[44] which is a dispersion-corrected, range-separated hybrid GGA (generalized gradient approximation) DFT method. It has been recently demonstrated as among the most accurate hybrid GGA methods for thermochemistry, kinetics, noncovalent interactions, and geometry calculations in a comprehensive benchmark study of 200 DFT functionals.[48] According to this study,[48] the M06-2X[28] (used by Krenske and Houk in their study of the thiol-Michael addition reactions[27]) and the ωB97X-D functionals[44] offer similar accuracies for barrier heights and ionization energies, but the latter has smaller errors for isomerization energies, difficult thermochemistry and noncovalent energy calculations involving strong correlation or self-interaction error, which may be attributed to the incorporation of long-range attenuation of exchange energy.

The two-dimensional structures of 30 singly- and doubly-activated olefins are given in Table 1. Some of these compounds were taken from literature,[6,27,31,35,42] and some of them are related to the FDA-approved drug afatinib or the covalent warheads for EGFR.[38] To avoid bias in the correlation study, additional model compounds were added such that the reaction barriers evenly distribute between approximately 7 and 23 kcal/mol. Following the previous work of others,[22,24,27,30,31,37] methanethiol (MeSH) and methanethiolate (MeS^-) was used as the model thiol and thiolate, respectively. The initial three-dimensional structures of MeSH, olefins, carbanions, and adducts were downloaded from Crystallography Open Database (https://www.crystallography.net/cod/) or if a X-ray structure was unavailable generated using Avogadro 1.2.0[59] and optimized with the built-in MMFF94s force field.[60] Some initial structures were modified from the afatinib crystal structure taken from the X-ray co-crystal structure of EGFR (PDB ID: 4G5J). These initial structures were subject to energy minimization and the optimized structures were checked to ensure that no imaginary frequency was found. The TS search was performed in two steps. First, we performed a potential energy scan starting from the optimized carbanion structure along the sulfur to β-carbon distance from 1.7 Å to 3 Å in 20 steps to identify a geometry that gives the maximum total energy, which then used as the initial guess for the TS geometry optimization. We conducted numerical frequency analysis in the SMD solvent model and confirmed that only one imaginary frequency was found for the TS. The thiol adduct may adopt a *syn* or *anti* conformer, which refers to the relative positions of the R1 group and sulfur atom. For **1, 2, 5–7, 29**, and **30**, both conformers for the product were optimized and the *anti* conformer was consistently found to have a lower free energy in solution. Subsequently, only the *anti* state was calculated for other molecules.

Calculation of electrophilicity index, Fukui function, and natural charges.

Electrophilicity index ω is defined as[50,51]

$$\omega = \frac{\mu^2}{2\eta}, \tag{1}$$

where μ is the chemical potential and η is the chemical hardness and they are defined using the ionization potential (IP) and electron affinity (EA) as follows

$$\mu = -\frac{IP + EA}{2};\ \eta = IP - EA. \tag{2}$$

Since μ and η can be approximated through LUMO and HOMO energies,

$$\mu \approx \frac{E_{\mathrm{LUMO}} + E_{\mathrm{HOMO}}}{2};\ \eta \approx E_{\mathrm{LUMO}} - E_{\mathrm{HOMO}}. \tag{3}$$

an approximate electrophilicity index ω' can be calculated.

Fukui (or frontier) function measures the local reactivity of a molecule.[61] The condensed Fukui function for nucleophilic attack f_i^+ is defined as[62]

$$f_i^+ = q_i(N+1) - q_i(N), \tag{4}$$

where $q_i(N)$ is the charge of atom i in the neutral state of the molecule and $q_i(N+1)$ is the charge of atom i in the anionic state of the molecule but with the neutral-state structure. We calculated f_i^+ for the C_β atom in the olefin. To calculate the atomic partial charges, we performed natural population analysis (NPA)[53] using the open source JANPA software package.[63,64]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors acknowledge National Institutes of Health (R01CA256557) for funding.

Data Availability

The ORCA quantum mechanical software package is freely available at https://orcaforum.kofo.mpg.de/index.php. The JANPA program for natural population analysis can be obtained from https://janpa.sourceforge.net. All calculated data are included in the downloadable supplemental files (see Supporting Information). All input files are available for download at https://github.com/JanaShenLab/Michael_Acceptor_Reactivity.

ABBREVIATIONS

B3LYP	Becke three parameter Lee-Yang-Parr
DFT	density functional theory
DMAM	dimethylaminomethyl
EWG	electron withdrawing group
GSH	glutathione
K-RAS	Kirsten rat sarcoma virus
EGFR	human epidermal growth factor receptor 1
MM	molecular mechanics
TCI	targeted covalent inhibitor
TMAM	trimethylaminomethyl

QM	quantum mechanical

References

(1). Zhao Z; Bourne PE Progress with Covalent Small-Molecule Kinase Inhibitors. Drug Discov. Today 2018, 23, 727–735. [PubMed: 29337202]

(2). Gehringer M; Laufer SA Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. J. Med. Chem 2019, 62, 5673–5724. [PubMed: 30565923]

(3). Ferguson FM; Gray NS Kinase Inhibitors: The Road Ahead. Nat. Rev. Drug Discov 2018, 17, 353–377. [PubMed: 29545548]

(4). Abdeldayem A; Raouf YS; Constantinescu SN; Moriggl R; Gunning PT Advances in Covalent Kinase Inhibitors. Chem. Soc. Rev 2020, 49, 2617–2687. [PubMed: 32227030]

(5). Hallin J; Engstrom LD; Hargis L; Calinisan A; Aranda R; Briere DM; Sudhakar N; Bowcut V; Baer BR; Ballard JA; Burkard MR; Fell JB; Fischer JP; Vigers GP; Xue Y; Gatto S; Fernandez-Banet J; Pavlicek A; Velastagui K; Chao RC; Barton J; Pierobon M; Baldelli E; Patricoin EF; Cassidy DP; Marx MA; Rybkin II; Johnson ML; Ou S-HI; Lito P; Papadopoulos KP; Jänne PA; Olson P; Christensen JG The KRAS ^{G12C} Inhibitor MRTX849 Provides Insight toward Therapeutic Susceptibility of KRAS-Mutant Cancers in Mouse Models and Patients. Cancer Discov. 2020, 10, 54–71. [PubMed: 31658955]

(6). Serafimova IM; Pufall MA; Krishnan S; Duda K; Cohen MS; Maglathlin RL; McFarland JM; Miller RM; Frödin M; Taunton J Reversible Targeting of Noncatalytic Cysteines with Chemically Tuned Electrophiles. Nat. Chem. Biol 2012, 8, 471–476. [PubMed: 22466421]

(7). Bradshaw JM; McFarland JM; Paavilainen VO; Bisconte A; Tam D; Phan VT; Romanov S; Finkle D; Shu J; Patel V; Ton T; Li X; Lough-head DG; Nunn PA; Karr DE; Gerritsen ME; Funk JO; Owens TD; Verner E; Brameld KA; Hill RJ; Goldstein DM; Taunton J Prolonged and Tunable Residence Time Using Reversible Covalent Kinase Inhibitors. Nat. Chem. Biol 2015, 11, 525–531. [PubMed: 26006010]

(8). Forster M; Chaikuad A; Bauer SM; Holstein J; Robers MB; Corona CR; Gehringer M; Pfaffenrot E; Ghoreschi K; Knapp S; Laufer SA Selective JAK3 Inhibitors with a Covalent Reversible Binding Mode Targeting a New Induced Fit Binding Pocket. Cell Chem. Biol 2016, 23, 1335–1340. [PubMed: 27840070]

(9). Shindo N; Fuchida H; Sato M; Watari K; Shibata T; Kuwata K; Miura C; Okamoto K; Hatsuyama Y; Tokunaga K; Sakamoto S; Morimoto S; Abe Y; Shiroishi M; Caaveiro JMM; Ueda T; Tamura T; Matsunaga N; Nakao T; Koyanagi S; Ohdo S; Yamaguchi Y; Hamachi I; Ono M; Ojida A Selective and Reversible Modification of Kinase Cysteines with Chlorofluoroacetamides. Nat. Chem. Biol 2019, 15, 250–258. [PubMed: 30643284]

(10). Allen CFH; Fournier JO; Humphlett WJ The Thermal Reversibility of The Michael Reaction: IV. Thiol Adducts. Can. J. Chem 1964, 42, 2616–2620.

(11). Dmuchovsky B; Vineyard BD; Zienty FB The Mechanism of the Base-Catalyzed Addition of Thiols to Maleic Anhydride. J. Am. Chem. Soc 1964, 86, 2874–2877.

(12). Fishbein JC; Jencks WP Elimination Reactions of β-Cyano Thioethers: Evidence for a Carbanion Intermediate and a Change in Rate-Limiting Step. J. Am. Chem. Soc 1988, 110, 5075–5086.

(13). Heo CKM; Bunting JW Rate-Determining Steps in Michael-type Additions and E1cb Reactions in Aqueous Solution. J. Org. Chem 1992, 57, 3570–3578.

(14). Nair DP; Podgórski M; Chatani S; Gong T; Xi W; Fenoli CR; Bowman CN The Thiol-Michael Addition Click Reaction: A Powerful and Widely Used Tool in Materials Chemistry. Chem. Mater 2014, 26, 724–744.

(15). Roseli RB; Keto AB; Krenske EH Mechanistic aspects of thiol additions to Michael acceptors: Insights from computations. WIREs Comput. Mol. Sci 2023, 13, e1636.

(16). Friedman M; Cavins JF Relative Nucleophilic Reactivities of Amino Groups and Mercaptide Ions in Addition Reactions with α,β-Unsaturated Compounds. J. Am. Chem. Soc 1965, 87, 3672–3682.

(17). Gennari A; Wedgwood J; Lallana E; Francini N; Tirelli N Thiol-Based Michael-Type Addition. A Systematic Evaluation of Its Controlling Factors. Tetrahedron 2020, 76, 1–6.

(18). Harris RC; Liu R; Shen J Predicting Reactive Cysteines with Implicit-Solvent-Based Continuous Constant pH Molecular Dynamics in Amber. J. Chem. Theory Comput 2020, 16, 3689–3698. [PubMed: 32330035]

(19). Liu R; Zhan S; Che Y; Shen J Reactivities of the Front Pocket N-Terminal Cap Cysteines in Human Kinases. J. Med. Chem 2022, 65, 1525–1535. [PubMed: 34647463]

(20). Liu R; Verma N; Henderson JA; Zhan S; Shen J Profiling MAP Kinase Cysteines for Targeted Covalent Inhibitor Design. RSC Med. Chem 2022, 13, 54–63. [PubMed: 35224496]

(21). Berteotti A; Vacondio F; Lodola A; Bassi M; Silva C; Mor M; Cavalli A Predicting the Reactivity of Nitrile-Carrying Compounds with Cysteine: A Combined Computational and Experimental Study. ACS Med. Chem. Lett 2014, 5, 501–505. [PubMed: 24900869]

(22). Flanagan ME; Abramite JA; Anderson DP; Aulabaugh A; Dahal UP; Gilbert AM; Li C; Montgomery J; Oppenheimer SR; Ryder T; Schuff BP; Uccello DP; Walker GS; Wu Y; Brown MF; Chen JM; Hayward MM; Noe MC; Obach RS; Philippe L; Shanmugasundaram V; Shapiro MJ; Starr J; Stroh J; Che Y Chemical and Computational Methods for the Characterization of Covalent Reactive Groups for the Prospective Design of Irreversible Inhibitors. J. Med. Chem 2014, 57, 10072–10079. [PubMed: 25375838]

(23). Cee VJ; Volak LP; Chen Y; Bartberger MD; Tegley C; Arvedson T; Mc-Carter J; Tasker AS; Fotsch C Systematic Study of the Glutathione (GSH) Reactivity of *N*-Arylacrylamides: 1. Effects of Aryl Substitution. J. Med. Chem 2015, 58, 9171–9178. [PubMed: 26580091]

(24). Lonsdale R; Burgess J; Colclough N; Davies NL; Lenz EM; Orton AL; Ward RA Expanding the Armory: Predicting and Tuning Covalent Warhead Reactivity. J. Chem. Inf. Model 2017, 57, 3124–3137. [PubMed: 29131621]

(25). Palazzesi F; Hermann MR; Grundl MA; Pautsch A; Seeliger D; Tautermann CS; Weber A BIreactive: A Machine-Learning Model to Estimate Covalent Warhead Reactivity. J. Chem. Inf. Model 2020, 60, 2915–2923. [PubMed: 32250627]

(26). Hermann MR; Tautermann CS; Sieger P; Grundl MA; Weber A BIreactive: Expanding the Scope of Reactivity Predictions to Propynamides. Pharmaceuticals 2023, 16, 116. [PubMed: 36678612]

(27). Krenske EH; Petter RC; Houk KN Kinetics and Thermodynamics of Reversible Thiol Additions to Mono- and Diactivated Michael Acceptors: Implications for the Design of Drugs That Bind Covalently to Cysteines. J. Org. Chem 2016, 81, 11726–11733. [PubMed: 27934455]

(28). Zhao Y; Truhlar DG The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals. Theor. Chem. Account 2008, 120, 215–241.

(29). Tissandier MD; Cowen KA; Feng WY; Gundlach E; Cohen MH; Earhart AD; Coe JV; Tuttle TR The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. J. Phys. Chem. A 1998, 102, 7787–7794.

(30). Smith JM; Rowley CN Automated Computational Screening of the Thiol Reactivity of Substituted Alkenes. J. Comput. Aided Mol. Des 2015, 29, 725–735. [PubMed: 26159564]

(31). Krenske EH; Petter RC; Zhu Z; Houk KN Transition States and Energetics of Nucleophilic Additions of Thiols to Substituted α,β-Unsaturated Ketones: Substituent Effects Involve Enone Stabilization, Product Branching, and Solvation. J. Org. Chem 2011, 76, 5074–5081. [PubMed: 21574592]

(32). Hughes TB; Miller GP; Swamidass SJ Site of Reactivity Models Predict Molecular Reactivity of Diverse Chemicals with Glutathione. Chem. Res. Toxicol 2015, 28, 797–809. [PubMed: 25742281]

(33). Palazzesi F; Grundl MA; Pautsch A; Weber A; Tautermann CS A Fast Ab Initio Predictor Tool for Covalent Reactivity Estimation of Acrylamides. J. Chem. Inf. Model 2019, 59, 3565–3571. [PubMed: 31246457]

(34). Hermann MR; Pautsch A; Grundl MA; Weber A; Tautermann CS Covalent Inhibitor Reactivity Prediction by the Electrophilicity Index—in and out of Scope. J. Comput. Aided Mol. Des 2021, 35, 531–539. [PubMed: 33015740]

(35). Krishnan S; Miller RM; Tian B; Mullins RD; Jacobson MP; Taunton J Design of Reversible, Cysteine-Targeted Michael Acceptors Guided by Kinetic and Computational Analysis. J. Am. Chem. Soc 2014, 136, 12624–12630. [PubMed: 25153195]

(36). Voice A; Tresadern G; van Vlijmen H; Mulholland A Limitations of Ligand-Only Approaches for Predicting the Reactivity of Covalent Inhibitors. J. Chem. Inf. Model 2019, 59, 4220–4227. [PubMed: 31498988]

(37). Smith JM; Jami Alahmadi Y; Rowley CN Range-Separated DFT Functionals Are Necessary to Model Thio-Michael Additions. J. Chem. Theory Comput 2013, 9, 4860–4865. [PubMed: 26583405]

(38). Tsou H-R; Mamuya N; Johnson BD; Reich MF; Gruber BC; Ye F; Nilakantan R; Shen R; Discafani C; DeBlanc R; Davis R; Koehn FE; Greenberger LM; Wang Y-F; Wissner A 6-Substituted-4-(3-Bromophenylamino)Quinazolines as Putative Irreversible Inhibitors of the Epidermal Growth Factor Receptor (EGFR) and Human Epidermal Growth Factor Receptor (HER-2) Tyrosine Kinases with Enhanced Antitumor Activity. J. Med. Chem 2001, 44, 2719–2734. [PubMed: 11495584]

(39). Wissner A; Overbeek E; Reich MF; Floyd MB; Johnson BD; Mamuya N; Rosfjord EC; Discafani C; Davis R; Shi X; Rabindran SK; Gruber BC; Ye F; Hallett WA; Nilakantan R; Shen R; Wang Y-F; Greenberger LM; Tsou H-R Synthesis and Structure-Activity Relationships of 6,7-Disubstituted 4-Anilinoquinoline-3-carbonitriles. The Design of an Orally Active, Irreversible Inhibitor of the Tyrosine Kinase Activity of the Epidermal Growth Factor Receptor (EGFR) and the Human Epidermal Growth Factor Receptor-2 (HER-2). J. Med. Chem 2003, 46, 49–63. [PubMed: 12502359]

(40). Birkholz A; Kopecky DJ; Volak LP; Bartberger MD; Chen Y; Tegley CM; Arvedson T; McCarter JD; Fotsch C; Cee VJ Systematic Study of the Glutathione Reactivity of *N*-Phenylacrylamides: 2. Effects of Acrylamide Substitution. J. Med. Chem 2020, 63, 11602–11614. [PubMed: 32965113]

(41). Guo W-H; Qi X; Yu X; Liu Y; Chung C-I; Bai F; Lin X; Lu D; Wang L; Chen J; Su LH; Nomie KJ; Li F; Wang MC; Shu X; Onuchic JN; Woyach JA; Wang ML; Wang J Enhancing Intracellular Accumulation and Target Engagement of PROTACs with Reversible Covalent Chemistry. Nat. Commun 2020, 11, 4268. [PubMed: 32848159]

(42). LoPachin RM; Gavin T; Geohagen BC; Das S Neurotoxic Mechanisms of Electrophilic Type-2 Alkenes: Soft Soft Interactions Described by Quantum Mechanical Parameters. Toxicol. Sci 2007, 98, 561–570. [PubMed: 17519395]

(43). Chai J-D; Head-Gordon M Systematic optimization of long-range corrected hybrid density functionals. J. Chem. Phys 2008, 128, 084106. [PubMed: 18315032]

(44). Chai J-D; Head-Gordon M Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. Phys. Chem. Chem. Phys 2008, 10, 6615–6620. [PubMed: 18989472]

(45). Grimme S; Antony J; Ehrlich S; Krieg H A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys 2010, 132, 154104. [PubMed: 20423165]

(46). Marenich AV; Cramer CJ; Truhlar DG Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. J. Phys. Chem. B 2009, 113, 6378–6396. [PubMed: 19366259]

(47). Grimme S; Ehrlich S; Goerigk L Effect of the Damping Function in Dispersion Corrected Density Functional Theory. J. Comput. Chem 2011, 32, 1456–1465. [PubMed: 21370243]

(48). Mardirossian N; Head-Gordon M Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. Mol. Phys 2017, 115, 2315–2372.

(49). Yang H; Wong MW Water-Assisted and Catalyst-Free Hetero-Michael Additions: Mechanistic Insights from DFT Investigations. Asian J. Org. Chem 2022, 11, e202100632.

(50). Maynard AT; Huang M; Rice WG; Covell DG Reactivity of the HIV-1 Nucleocapsid Protein P7 Zinc Finger Domains from the Perspective of Density-Functional Theory. Proc. Natl. Acad. Sci. USA 1998, 95, 11578–11583. [PubMed: 9751708]

(51). Parr RG; Szentpály L. v.; Liu S Electrophilicity Index. J. Am. Chem. Soc 1999, 121, 1922–1924.

(52). Domingo LR; Sáez JA Understanding the Mechanism of Polar Diels–Alder Reactions. Org. Biomol. Chem 2009, 7, 3576. [PubMed: 19675915]

(53). Reed AE; Weinstock RB; Weinhold F Natural Population Analysis. J. Chem. Phys 1985, 83, 735–746.

(54). Hall HK Correlation of the Base Strengths of Amines. J. Am. Chem. Soc 1957, 79, 5441–5444.

(55). Johnston RC; Yao K; Kaplan Z; Chelliah M; Leswing K; Seekins S; Watts S; Calkins D; Chief Elk J; Jerome SV; Repasky MP; Shelley JC Epik: pKa and Protonation State Prediction through Machine Learning. J. Chem. Theory Comput 2023, 19, 2380–2388. [PubMed: 37023332]

(56). Wood ER; Shewchuk LM; Ellis B; Brignola P; Brashear RL; Caferro TR; Dickerson SH; Dickson HD; Donaldson KH; Gaul M; Griffin RJ; Hassell AM; Keith B; Mullin R; Petrov KG; Reno MJ; Rusnak DW; Tadepalli SM; Ulrich JC; Wagner CD; Vanderwall DE; Waterson AG; Williams JD; White WL; Uehling DE 6-Ethynylthieno[3,2-d]- and 6-Ethynylthieno[2,3-d]Pyrimidin-4-Anilines as Tunable Covalent Modifiers of ErbB Kinases. Proc. Natl. Acad. Sci. USA 2008, 105, 2773–2778. [PubMed: 18287036]

(57). Capoferri L; Lodola A; Rivara S; Mor M Quantum Mechanics/Molecular Mechanics Modeling of Covalent Addition between EGFR–Cysteine 797 and *N*-(4-Anilinoquinazolin-6-Yl) Acrylamide. J. Chem. Inf. Model 2015, 55, 589–599. [PubMed: 25658136]

(58). Neese F; Wennmohs F; Becker U; Riplinger C The ORCA Quantum Chemistry Program Package. J. Chem. Phys 2020, 152, 224108. [PubMed: 32534543]

(59). Hanwell MD; Curtis DE; Lonie DC; Vandermeersch T; Zurek E; Hutchison GR Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. J. Cheminform 2012, 4, 17. [PubMed: 22889332]

(60). Halgren TA MMFF VI. MMFF94s Option for Energy Minimization Studies. J. Comput. Chem 1999, 20, 720–729. [PubMed: 34376030]

(61). Fukui K; Yonezawa T; Shingu H A Molecular Orbital Theory of Reactivity in Aromatic Hydrocarbons. J. Chem. Phys 1952, 20, 722–725.

(62). Yang Weitao.; Mortier WJ The Use of Global and Local Molecular Parameters for the Analysis of the Gas-Phase Basicity of Amines. J. Am. Chem. Soc 1986, 108, 5708–5711. [PubMed: 22175316]

(63). Nikolaienko TY; Bulavin LA; Hovorun DM JANPA: An Open Source Cross-Platform Implementation of the Natural Population Analysis on the Java Platform. Comput. Theor. Chem 2014, 1050, 15–22.

(64). Nikolaienko TY; Bulavin LA Localized Orbitals for Optimal Decomposition of Molecular Properties. Int. J. Quantum Chem 2019, 119, e25798.

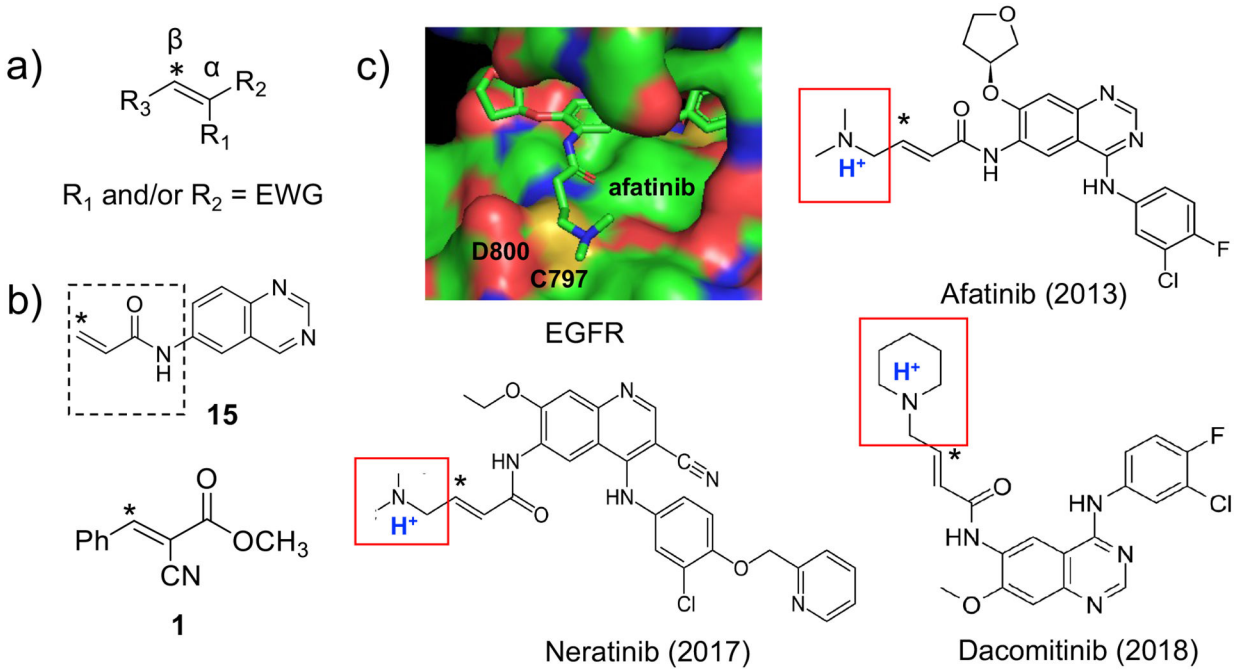


Figure 1: Example activated olefins and targeted covalent inhibitors (TCIs).

a) Structure of an activated olefin. A singly- or doubly-activated olefin refers to the α substitution with one or two electron-withdrawing groups (EWGs). The β carbon where the thiol-Michael addition takes place is indicated with an asterisk. b) Example singly-activated (**15**) and doubly-activated (**1**) olefins. The acrylamide group is highlighted by a dashed box. c) A zoomed-in view of the EGFR front pocket (in surface rendering) where afatinib is covalently attached to C797. The nearby aspartate D800 is also labeled. The structures of the three FDA-approved EGFR TCIs with the β-DMAM substitution (red box) are given below. The regulatory approval years are in parenthesis.

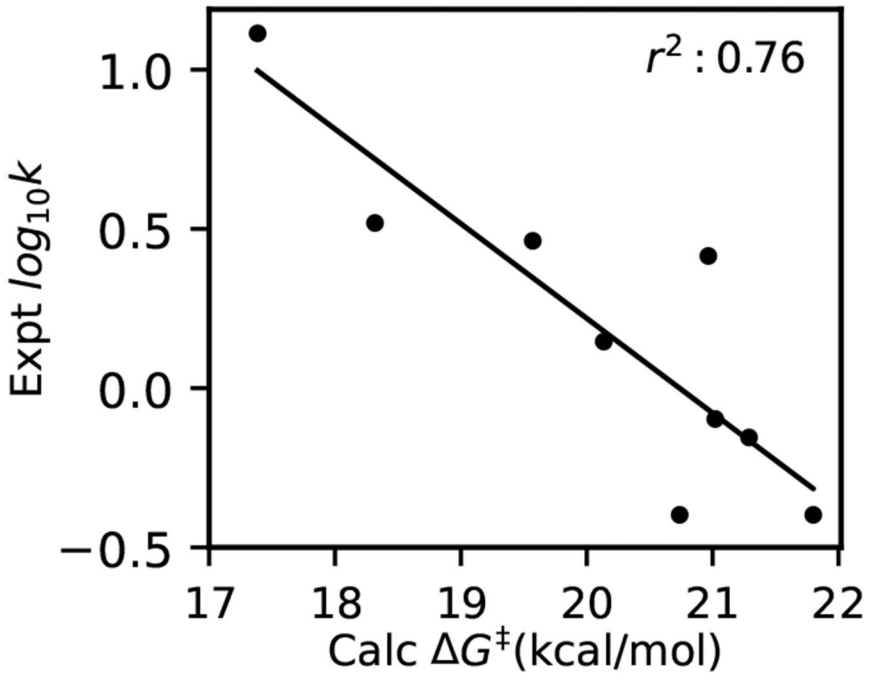


Figure 2: The calculated activation free energies are highly correlated with the experimental kinetic rate constants.

$\Delta G^{\ddagger}$ refers to the activation free energies of **13, 19, 20, 21, 22, 23, 24, 25, 26** calculated with the ωB97X-D3BJ / 6-311 + G(d, p) / SMD method in this work. k refers to the experimental pseudo first-order rate constant (in unit of 10^{-3} min^{-1}) of the solution thiol-adduct formation (data taken from Ref .[22]). The linear correlation coefficient (r^2) and the best fit line are shown. The numerical values are given in Table S2.

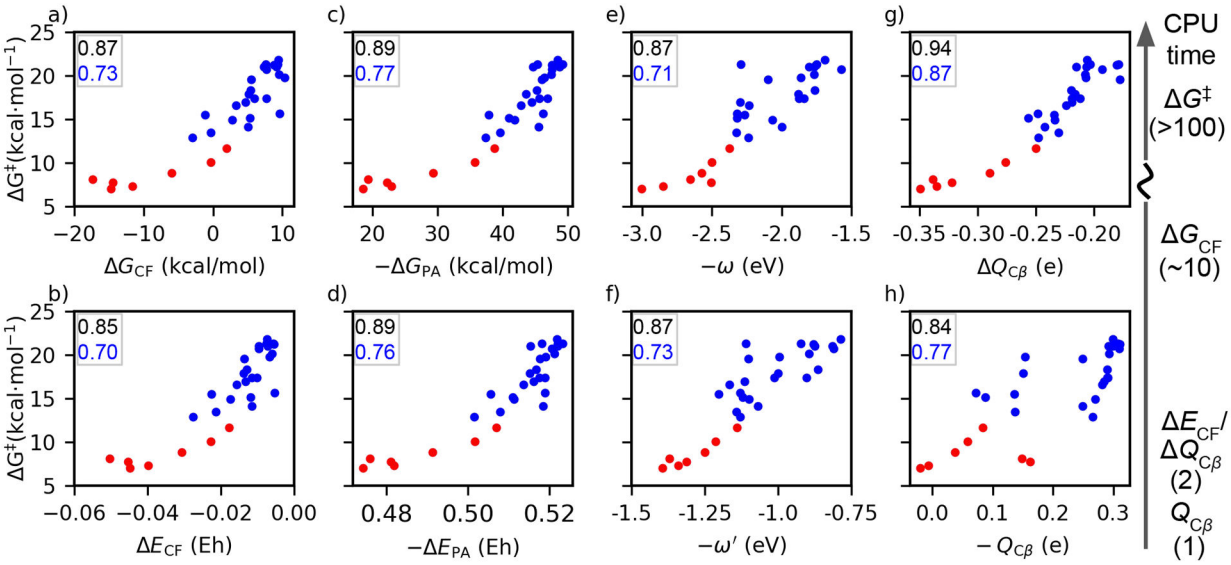


Figure 3: Correlation between the reaction barrier and ground-state electronic properties:
a) free energy of carbanion formation (ΔG_{CF}); b) single-point energy of carbanion formation (ΔE_{CF}); c) free energy of carbanion proton affinity (ΔG_{PA}); d) single-point energy of carbanion proton affinity (ΔE_{PA}); e) electrophilicity index of the olefin (ω); f) approximate electrophilicity index of the olefin (ω'); g) change in the C_β charge (from natural population analysis) going from the olefin to the carbanion state ($\Delta Q_{C\beta}$; and h) C_β atomic charge of the olefin ($Q_{C\beta}$) for the singly- (blue) and doubly-activated olefins. The Spearman's ρ values are given for all olefins in black and for singly-activated olefins in blue. Correlations with other electronic properties are given in Table S3 and S4 along with Pearson's r and Kendall's τ.

R_β =		$\Delta G^\ddagger$	$R_{S\text{-}Cb}$	$R_{N\text{-}Ca}$	$R_{N\text{-}S}$	$\Theta_{N\text{-}Cg\text{-}Cb}$	ΔG_{CF}	ΔG_{PA}	ΔQ_{Ca}	ΔQ_{Cb}
-H	**20**	21.29	2.37	NA	N/A	NA	9.36	-49.21	-0.26	-0.20
$-CH_2N(CH_3)_2$	**27**	23.09	2.37	3.16	4.39	113.3	12.45	-49.77	-0.24	-0.24
$-CH_2NH^+(CH_3)_2$	**27H**	17.85	2.38	3.05	4.38	110.6	5.11	-43.59	-0.33	-0.22
$-CH_2N^+(CH_3)_3$	**28**	19.79	2.35	3.25	4.36	114.0	10.29	-46.38	-0.30	-0.21

The units are kcal/mol (free energies), Å (bond length), ° (angles), *e* for charges.

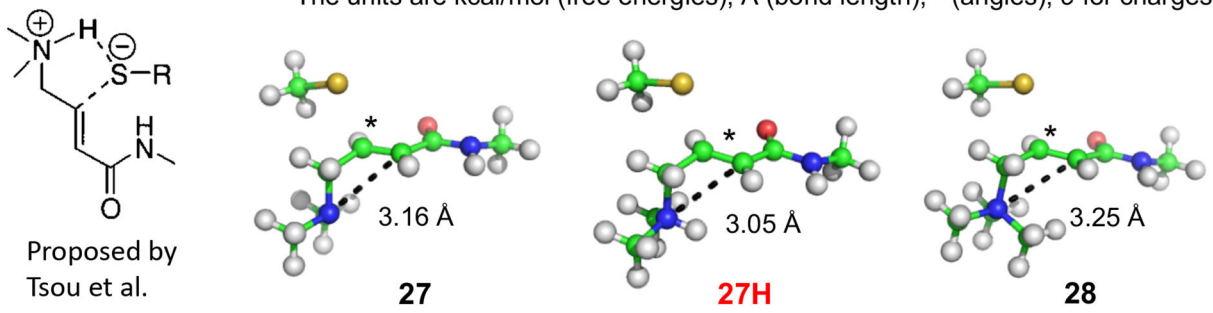
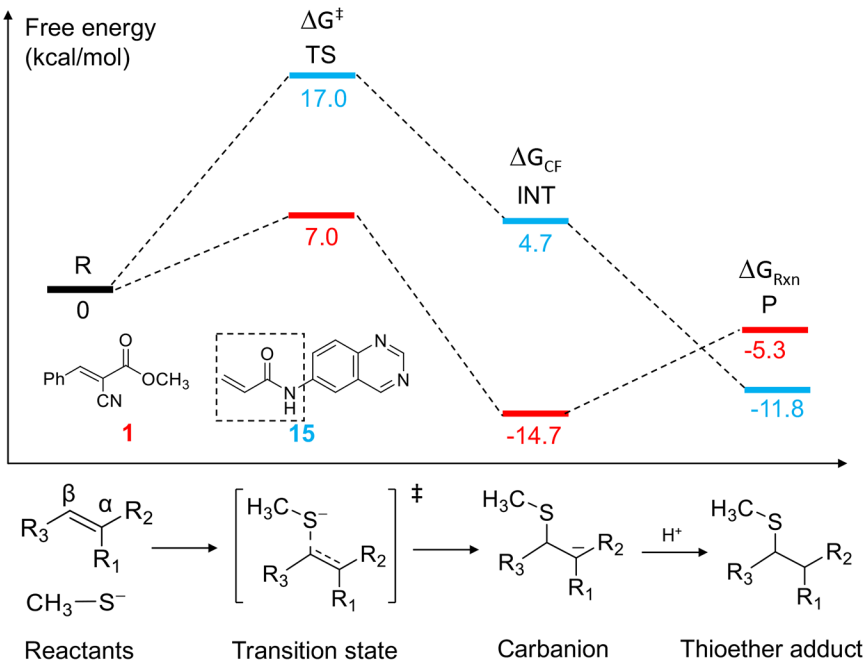


Figure 4: The protonated β-DMAM substitution lowers the reaction barrier due to an induction effect.

A comparison of the reaction energetics, TS geometries, and changes in C_α and C_β charges upon carbanion formation for the four related model compounds. The TS structures of **27, 27H**, and **28** are given below, with the distance $R_{N\cdots C\alpha}$ indicated.



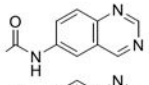
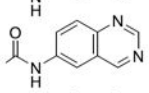
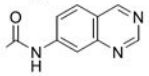
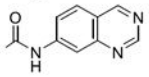
Scheme 1: Reaction profile of the thiol-Michael addition of activated olefin.

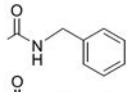
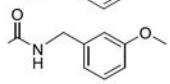
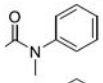
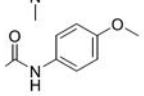
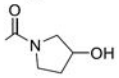
a) The reaction free energy profiles of a doubly activated olefin (**1**) and a singly-activated olefin (**15**, analog of the afatinib warhead, acrylamide group indicated with a box). b) The structures of reactants (R), transition state (TS), carbanion intermediate (INT), and product (P). $\Delta G^{\ddagger}$, ΔG_{CF}, and ΔG_{Rxn} refer to the free energy terms going from reactants to TS, INT, and P, respectively. The carbanion proton affinity ΔG_{PA} discussed in the main text refers to the free energy of protonation of INT. The proton Gibbs free energy of −270.3 kcal/mol (data from ref.[29]) was used to calculated ΔG_{PA}.

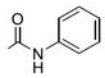
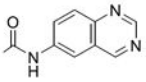
Table 1:

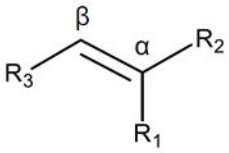
Summary of the singly- and doubly-activated olefins studied in this work[a]

	R_2 (R_1=CN)	R_3
1	-$COOCH_3$	-Ph
2	-$CONH_2$	-Ph
3	-$COOCH_3$	-H
4	-$CONH_2$	-H
5		-Ph
6		-Ph
7		-Ph

	R_2 (R_1=H)	R_3
8	-$COCH_3$	-H
9	-$COOCH_3$	-H
10	-CN	-H
11	-$COOCH_3$	-Ph
12	-CN	-Ph
13	-CONHPh	-H
14	-$CONH_2$	-H
15		-H
16H		-$CH_2NH^+(CH_3)_2$
17		-H
18H		-$CH_2NH^+(CH_3)_2$
19		-H
20	-$CONHCH_3$	-H
21	-$CONHCH_3$	-CF_3

	R_2 (R_1=H)	R_3
22		-H
23		-H
24		-H
25		-H
26		-H
27H	-$CONHCH_3$	-$CH_2NH^+(CH_3)_2$
28	-$CONHCH_3$	-$CH_2N^+(CH_3)_3$

	R_2 (R_1=CH_3)	R_3
29		-H
30		-H



Activated olefin

[a]The olefins are listed in an approximately ascending order of $\Delta G^{\ddagger}$. 16H, 18H, 27H refer to the protonated β−dimethylaminomethyl (DMAM) substitution, which is the dominant protonation state at neutral pH (see main text discussion). The deprotonated forms were additionally calculated and discussed (see Table 3).

Table 2:

Calculated solution-phase activation and reaction free energies for the singly- and doubly-activated olefins studied in this work[a]

Olefin	$\Delta G^{\ddagger}$	ΔG_{rxn}	Olefin	$\Delta G^{\ddagger}$	ΔG_{rxn}
1*	7.01	−5.28	17	16.58	−11.50
2*	7.30	−6.66	18	17.62	−9.42
3*	8.09	−8.74	18H	15.49	−11.09
4*	7.71	−8.70	19	20.74	−11.89
5*	8.80	−7.38	20	21.29	−11.88
6*	10.06	−8.07	21	19.57	−12.54
7*	11.65	−8.83	22	21.02	−11.60
8	12.89	−12.34	23	21.80	−11.09
9	14.89	−11.08	24	18.32	−11.87
10	14.14	−12.54	25	20.97	−9.37
11	15.16	−7.65	26	20.14	−9.99
12	15.63	−8.56	27	23.09	−9.34
13	17.38	−11.76	27H	17.85	−10.50
14	17.35	−11.18	28	19.79	−8.12
15	16.96	−11.79	29	21.24	−10.81
16	18.85	−10.29	30	21.28	−9.81
16H	13.50	−11.92			

[a]The unit for $\Delta G^{\ddagger}$ and ΔG_{rxn} is kcal/mol. Doubly-activated olefins are indicated by an asterisk.

Table 3:

Effect of the β-DMAM substitution on the reaction energetics[a]

	15	**16**	**16H**	**17**	**18**	**18H**	**20**	**27**	**27H**
$\Delta G^{\ddagger}$	16.96	18.85	13.50	16.58	17.62	15.49	21.29	23.09	17.85
ΔG_{CF}	4.68	6.39	−0.28	3.28	6.07	−1.16	9.36	12.45	5.11
ΔG_{PA}	−44.45	−44.64	−39.61	−42.76	−43.47	−37.91	−49.21	−49.77	−43.59

[a]In **16, 18**, and **27**, the DMAM group is deprotonated.