Quantum Chemical Methods for Modeling Covalent Modification of Biological Thiols

Ernest Awoonor-Williams[*], William C. Isley III[†],
Stephen G. Dale[‡], Erin R. Johnson[§], Haibo Yu[¶],
Axel D. Becke[‖], Benoît Roux[**], Christopher N. Rowley[††]

July 24, 2019

Abstract

Targeted covalent inhibitor drugs require computational methods that go beyond simple molecular-mechanical force fields in order to model the chemical reactions that occur when they bind to their targets. Here, several semi-empirical and density-functional theory (DFT) methods are assessed for their ability to describe the potential energy surface and reaction energies of the covalent modification of a thiol by an electrophile. Functionals such as PBE and B3LYP fail to predict a stable enolate intermediate. This is largely due to delocalization error, which spuriously stabilizes the pre-reaction complex, in which excess electron density is transferred from the thiolate to the electrophile. Functionals with a high-exact exchange component, range-separated DFT functionals, and variationally-optimized exact exchange (i.e., the LC-B05minV functional) correct this issue to various degrees. The large gradient behaviour of the exchange enhancement factor is also found to significantly affect the results, leading to the improved performance of PBE0. While ωB97X-D and M06-2X were reasonably accurate, no method provided quantitative accuracy for all three electrophiles, making this a very strenuous test of functional performance. Additionally, one drawback of M06-2X was that MD simulations using this functional were only stable if a fine integration grid was used. The low-cost semi-empirical methods, PM3, AM1, and PM7, provide a qualitatively correct description of the reaction mechanism, although the energetics are not quantitatively reliable. As a proof of concept, the potential of mean force for the addition of methylthiolate to MVK was calculated using QM/MM MD in an explicit polarizable aqueous solvent.

[*]Department of Chemistry, Memorial University of Newfoundland, St. John's, Canada

[†]Department of Biochemistry and Molecular Biology, University of Chicago

[‡]Research School of Chemistry, Australian National University, Canberra, Australia

[§]Department of Chemistry, Dalhousie University, Halifax, Canada

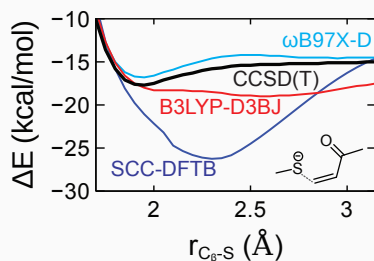[¶]School of Chemistry and Molecular Bioscience, University of Wollongong, Wollongong, Australia

[‖]Department of Chemistry, Dalhousie University, Halifax, Canada

[**]Department of Biochemistry and Molecular Biology, University of Chicago

[††]Department of Chemistry, Memorial University of Newfoundland, St. John's, Canada

Stop using B3LYP, PBE, and DFTB for reactions with cysteine thiolates

ΔE (kcal/mol)

−10
−15
−20
−25
−30

ωB97X-D
CCSD(T)
B3LYP-D3BJ
SCC-DFTB

2   2.5   3

$r_{C_\beta\text{-}S}$ (Å)

Reactions involving cysteine thiols often involve an anionic thiolate intermediate. Delocalization error in many popular DFT methods results in incorrect predictions of thio-Michael addition mechanisms of covalent-modifier drugs, while large exact-exchange components or range-separated functionals address these issues to some degree. DFTB methods also fail to predict a stable enolate, while some semi-empirical methods provide qualitative agreement with high-level theory. The thio-Michael additions are found to be an extremely stringent test of electronic-structure methods, highlighting the need for accurate treatment of both electron delocalization and non-bonded repulsion.

INTRODUCTION

Covalent modifiers are an emerging class of drugs that form a covalent bond with their desired targets, which allows for greater potency,[1] improved selectivity,[2] and extended residence times.[3] A notable class of covalent modifiers features an electrophilic functional group, often termed a warhead, that undergoes a Michael addition with the thiol sidechain of a noncatalytic cysteine residue. A vast majority of these electrophilic groups feature α-β unsaturated carbonyl compounds,[4] particularly acrylamides, which are now the standard electrophile in recent covalent drug-design efforts. One notable example is ibrutinib,[5] a drug that features an electrophilic acrylamide warhead. Ibrutinib is used in the treatment of B cell malignancies such as chronic lymphocytic leukemia, targeting Bruton's tyrosine kinase (BTK) by covalently binding to Cys481 (Figure 1).

Figure 1: Reaction mechanism for the addition of the acrylamide warhead of ibrutinib to Cys481 of Bruton's tyrosine kinase (BTK).
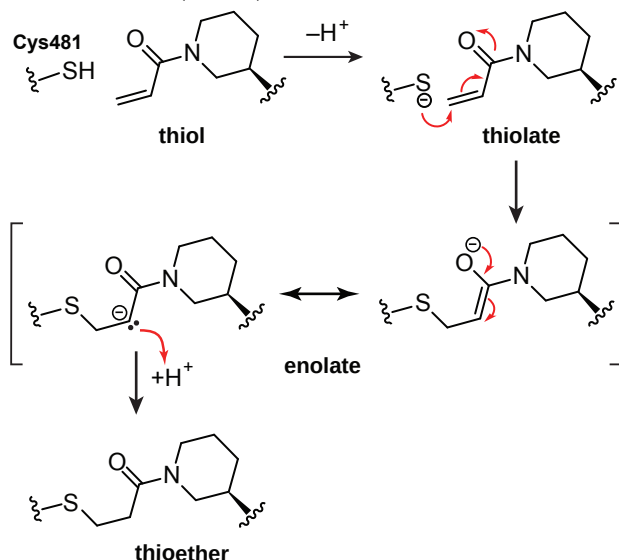
Computational modeling of covalent modification can provide insights into the binding affinity, selectivity, and kinetics of these drugs.[6] In two recent papers, Luo and coworkers have shown that the relative efficacy of covalent inhibitors can be related to the relative stabilities of the covalently bound adduct.[7,8] The relative binding affinities of the covalent-modifier drugs can be estimated using free-energy perturbation so, in these cases, simulating

the mechanism of covalent bond formation was not necessary. Nevertheless, rigorous models of the absolute binding energies and rates will require methods that can describe the covalent-bond formation steps accurately.

Selection of an optimal warhead would benefit from accurate predictions of how the chemical substituents of the electrophile affect the rates and thermodynamics of the addition reaction. This can be particularly useful in the development of reversible covalent drugs, which have an advantage over their irreversible-binding counterparts in that they can dissociate if they bind to off-target proteins and eventually reach their intended targets. The effects of chemical substituents on warhead reactivity are discussed in detail in Refs. [9–14]. Recent work by Taunton and coworkers on covalent-reversible inhibitors[3,15] has demonstrated that the reversibility of the addition of acrylamide-based electrophiles to target cysteines can be tuned by substitutions on the acrylamide. For example, introducing an electron-withdrawing nitrile group onto the α-carbon of an acrylamide and altering the steric environment around the β-carbon resulted in covalent drugs with prolonged residence times for BTK.[3] A 2015 paper by Smith *et al.*[12] used density-functional theory (DFT) calculations to show that substitution of the electrophilic group can affect the rates of reaction and the reaction energy with a model thiol, suggesting that new covalent modifier drugs could be designed using computational methods that describe the mechanism of covalent modification accurately.

Kinetic experiments indicate that the conjugate addition of a thiol to an electrophile in aqueous solutions proceeds through an ionic mechanism, consistent with the classical Michael-addition mechanism.[16] This is at odds with some computational models of these reactions, which predicted mechanisms that do not include an enolate intermediate.[17–20] In 2013, Smith *et al.*[21] showed that several popular density functionals (e.g., PBE, B3LYP) predict that there is no potential-energy minimum associated with the enolate intermediate, although high-level *ab initio* calculations (i.e., CCSD(T)) predict that the enolate intermediate is stable. The contradictory results were attributed to delocalization error in the exchange-correlation functional of these DFT methods, which spuriously stabilizes the inter-molecular charge-transfer complex over the enolate. This issue was attenuated with range-separated functionals, where a greater component of exact exchange is used to calculate the

exchange-correlation energy for larger inter-electron separations.

In this paper, we revisit the mechanism of covalent modification to further explore the reasons for the failure of several popular density functionals for the Michael-addition step. New and established density functionals are explored, along with low-cost computational methods, to identify which are suitable for modeling covalent modification of thiols by electrophiles.

METHODOLOGY

Potential Energy Surfaces and Reaction Energies

In the first section of this paper, we examine the potential energy surface along an adiabatic (energy-minimized) reaction pathway for different levels of theory. Those calculations are carried out "in vacuo" (no solvation environment). The semi-empirical, DFT, and MP2 potential energy surfaces were calculated using Gaussian 16 A03.[22] Except where noted, all calculations were performed using the aug-cc-pVTZ basis set.[23,24] The popular PBE,[25] PBE0,[26] and B3LYP[27,28] density functionals with Grimme's D3BJ dispersion correction,[29] as well as ωB97X-D,[30] and M06-2X[31], and the semi-empirical quantum-mechanical (SQM) methods AM1,[32] PM3,[33] and PM7 (MOPAC implementation),[34] were considered. Grimme's D3BJ dispersion correction was used because it is implemented in ORCA 4.0.1.2, which we use in our QM/MM MD simulations; however, alternative dispersion corrections like XDM[35] could be used instead. The reference CCSD(T) energies were calculated at the MP2-optimized geometries within the resolution of identity approximation using TURBOMOLE 7.0.[36] DFTB calculations with SCC-DFTB[37,38] and DFTB3[39,40] were also carried out at the MP2-optimized geometries using DFTB+ (version 18.1).[41]

In construction of the adiabatic, relaxed potential energy surfaces for the thiolate–Michael acceptor reactions, energies are reported relative to the separate, energy-minimized methylthiolate and Michael-acceptor fragments. The net reaction energies are calculated relative to the neutral thiol and Michael-acceptor reactants. For the surfaces involving acrylamides, the dihedral angles of the N-methyl groups were constrained to prevent the acrylamide from

undergoing a conformation change as the $C_\beta-S$ distance was varied. As a result, the minima of the curves in Figure 8 are slightly higher in energy than those reported in Table 1, which are not constrained.

The LC-B05minV calculations were performed at the MP2-optimized geometries for the case of thiolate addition to methyl vinyl ketone (MVK) and followed the same procedure described by Becke et al.[42] With this method, a series of single-point energy calculations were performed with the aug-cc-pVTZ basis set and the LC-BLYP functional,[43] as implemented in Gaussian 09.[44] The value of the range-separation parameter, ω, was systematically varied in increments of 0.01. The B05 energies[45,46] were then evaluated for each ω from the LC-BLYP wavefunction using the POSTG program.[47–49] Following our previous work,[42] 12 occupied orbitals were treated as "core" and the remainder treated as "valence." The B05 energy was minimized with respect to the value of ω for the core and valence electrons separately. This is done as the core electrons remain localized and have an optimum range-separation parameter of $\omega = 0.19$, as is typically used for computations involving neutral, main-group compounds at equilibrium geometries.[46] However, the valence electrons are involved in the non-local charge transfer between the methylthiolate and MVK to yield the enolate, where an extremely high range separation of $\omega = 1.33$ is found to minimize the B05 energy.

Potential of Mean Force

The potential energy surfaces described above correspond to "in vacuo" calculations along an adiabatic (energy-minimized) reaction pathway and do not account for the influence of solvation or thermal fluctuations (entropy). In the final section of this manuscript, we perform a proof-of-concept calculation of the potential of mean force (PMF) for the reaction of MVK and methylthiolate in aqueous solution, with explicit water molecules, using a Drude polarizable model. The parameters for MVK were generated using the GAAMP server (see Supplementary Information),[50] while the parameters of Lin et al. were used for the polarizable MM model for methylthiolate.[51] These parameters have been previously found to provide a solvation structure of methylthiolate that is in excellent agreement with QM/MM MD simulations.[52] The PMF was calculated for the interval of $r_{C_\beta-S} = 4.0 - 8.0$ Å by employing a harmonic restraint on the distance between the C_β of MVK and the S atom

Figure 2: Reaction mechanism for the addition of methylthiolate to MVK via an enolate intermediate.
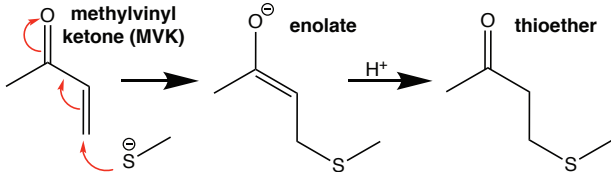
of methylthiolate. A spring constant of 2.5 kcal/mol Å^{-2} was used and the windows were separated by 1 Å, spanning $r_{C_\beta-S,0} = 3$ Å to $r_{C_\beta-S,0} = 8$ Å.

QM/pol-MM MD was used to calculate the PMF in $r = 1.7 - 4.0$ Å interval where the $C_\beta-S$ bond is formed. The molecular dynamics simulations were performed using NAMD 2.13[53] interfaced with ORCA 4.0.1.2.[54–56] The quantum-mechanical (QM) region was represented using the ωB97X-D functional and the def2-TZVP basis set. The resolution of the identify approximation was used to calculate the Coulomb integrals and chain-of-sphere integration was used for the exchange-integrals (i.e., the ORCA RICOSX keyword), which has been shown to be an efficient and reliable approximation in QM/MM MD simulations.[57–59] The standard auxiliary basis set for the def2-TZVP basis set was used. The spring constant was set to 100 kcal/mol Å^{-2} and the windows were separated by 0.1 Å, spanning $r_{C_\beta-S,0} = 1.7$ Å to $r_{C_\beta-S,0} = 4.5$ Å. A 5 ps equilibration simulation was performed for each window, prior to a 15 ps simulation that was used to construct the PMF via the Weighted Histogram Analysis Method (WHAM).[60,61]

As the restraint applied to the coordinates was in terms of an interatomic distance, it was transformed to be in terms of the radial coordinate in a spherical coordinate system using the relation[62]

$$w_r(r) = w(r) + 2RT \ln(r) + \text{constant}. \tag{1}$$

Figure 3: The $C_\beta-S$ distance used as the reaction coordinate in the calculated potential energy surfaces (left). Some QM methods predict a stable enolate intermediate with the $C_\beta-S$ bond length within a normal interval (center), while others predict a minimum energy structure with a longer $C_\beta-S$ distance that corresponds to a non-bonded change-transfer complex.
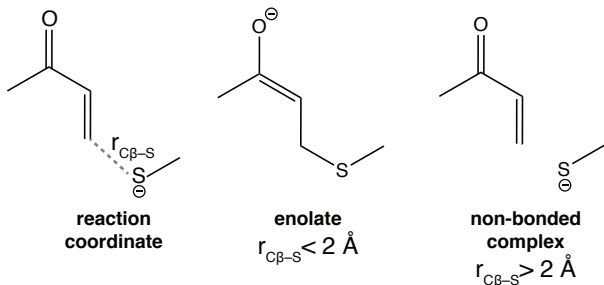
RESULTS AND DISCUSSION

Potential Energy Surfaces for Methyl Vinyl Ketone

The addition of methylthiol to methyl vinyl ketone (MVK) serves as a convenient test system to evaluate electronic structure methods for modeling thio-Michael additions. In this mechanism, the thiol is deprotonated to form a thiolate, which then attacks the C_β of the Michael acceptor to form an enolate intermediate (Figure 2). The adiabatic relaxed potential energy surfaces along the $r_{C_\beta-S}$ coordinate (Figure 3), corresponding to the formation of the enolate, are presented in Figure 4 for various DFT methods. The optimum $C_\beta-S$ bond lengths and relative energies of the enolate intermediates are collected in Table 1.

There is no potential energy minimum corresponding to an enolate intermediate (i.e., $r_{C_\beta-S} < 2.0$ Å on the potential energy surfaces calculated using the B3LYP-D3BJ and PBE-D3BJ functionals. Instead, the minimum-energy structures calculated using these methods correspond to charge-transfer complexes, where no $C_\beta-S$ bond is formed. This is contrary to the CCSD(T)//MP2 PES, which predicts a potential-energy minimum for the enolate at $r_{C_\beta-S} \approx 1.95$ Å. PBE0-D3BJ gives a potential-energy minimum at a separation comparable to the reference, but the shape of the curve is distorted, with a broader well. It should be noted that addition of Grimme's D3BJ dispersion correction[29] to B3LYP, PBE, or PBE0 serves to shift the potential energy surfaces to lower energies, but does not significantly alter

the curve shapes or positions of the minima relative to the dispersion-free base functionals. The potential energy surfaces calculated using ωB97X-D and M06-2X are in better agreement with CCSD(T). Each has a minimum near $r_{C_\beta-S} \approx 1.95$ Å and similar curve shapes for the PES.

Figure 4 also displays the Hirshfeld charges[63] on the thiolate group along the reaction coordinate. At small $C_\beta-S$ distances, the thiolate group has only a modest negative charge (ca. -0.2), which becomes more negative monotonically as the bond is stretched. At large $C_\beta-S$ distances, the thiolate bears close to a full negative charge, but there is considerable spread depending on the level of theory. Notably, ωB97X-D and M06-2X, which have the highest degrees of exact-exchange mixing, give charges close to the CCSD reference. The B3LYP and PBE0-D3BJ hybrid functionals, which have reduced exact-exchange mixing, show less negative charges on the thiolate by ca. 0.05 e^-. This is followed by PBE, which gives a charge that is ca. 0.1 e^- less negative than the CCSD reference. This tendency of functionals with reduced exact-exchange mixing to give more fractional charges for the intermolecular complex is characteristic of density-driven delocalization error.[64–66] The effect of delocalization error on the computed potential energy surface will be discussed in detail in the next section.

Resolving Delocalization Error and Non-Bonded Repulsion Issues

The failures of PBE and B3LYP to predict a stable enolate intermediate was previously attributed to delocalization error in the exchange-correlation functional.[21] Delocalization error is a result of the assumption of a localized exchange hole in the construction of the exchange functional.[67] It can also be interpreted as resulting from a spurious contribution to the exchange energy from self-interaction of electrons.[68] Delocalization error results in overstabilization of delocalized or fractional charge distributions[65,66,69] and can be reduced through inclusion of exact exchange, as in hybrid and range-separated hybrid functionals.[70–72] The classic example is a stretched H_2^+ molecule, where the dissociation limit is predicted to be far too low in energy with generalized-gradient-approximation (GGA) functionals (such as BLYP and PBE) as the delocalized charge is massively over-stabilized.[65,67,69] This is an "energy-driven" error, but delocalization error can also affect electron density distributions,

Figure 4: Potential energy surfaces for the addition of methylthiolate to MVK with selected density functionals (top). Also shown are computed Hirshfeld charges on the methylthiolate group as a function of $C_\beta-S$ interatomic separation (bottom). All calculations used the aug-cc-pVTZ basis set, except for the CCSD charges, which used the aug-cc-pVDZ basis set.
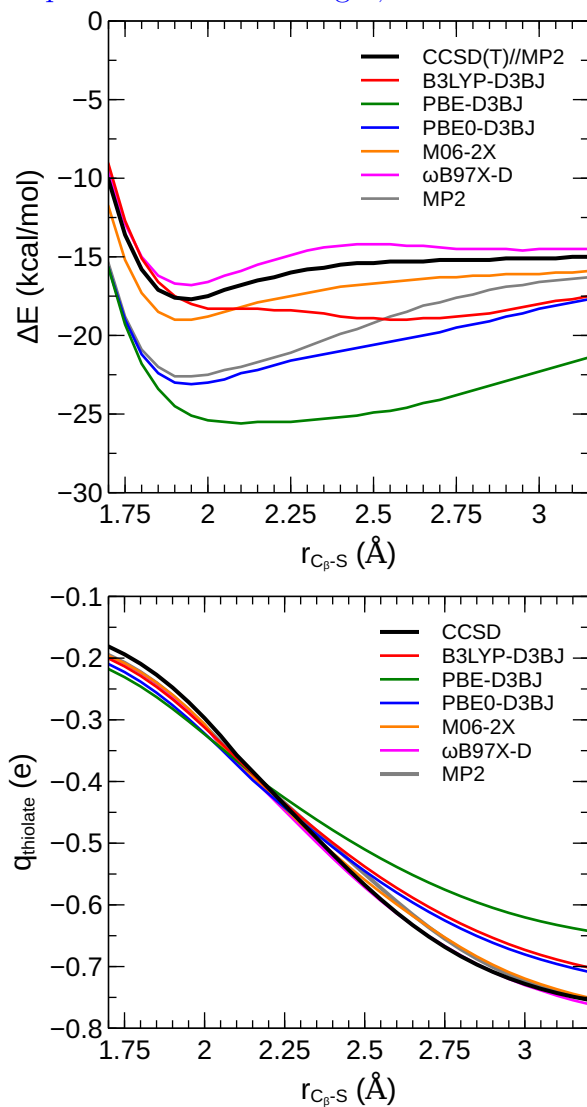
leading to an excessive spreading of charge over two or more molecular fragments in a system;[65,73,74] the most notable examples of this are for charge-transfer complexes[71] solvated electrons,[66] diffuse anions,[69,75,76] and organic salts.[77] In some cases, these "density-driven" delocalization errors can also result in erroneous geometries of intermolecular complexes and solids, as has been demonstrated for organic acid-base co-crystals[78] and for the present enolate intermediate resulting from Michael addition of thiolate to MVK.[21]

In the case of thio-enolates, delocalization error favors spurious transfer of electron density from the thiolate to the antibonding pi-orbitals of the Michael acceptors.[21] This occurs when the reactants have not yet formed a bond (i.e., $r_{C_\beta-S} > 2$ Å), so the intermolecular charge-transfer complexes are the most stable structures along the potential energy surface when GGA functionals are used. Conversely, functionals such as PBE0, M06-2X, and ωB97X-D, which reduce delocalization error via inclusion of exact-exchange mixing, predict a stable enolate intermediate. When a large component of exact exchange is included, electron density distributions where there is intermolecular charge transfer over large distances are predicted to be less stable.

However, delocalization error is not the only factor at play in controlling the shape of the potential energy surface for thiolate addition to MVK. This is clear from comparison of the B3LYP and PBE0-D3BJ results in Figure 4; both functionals include similar exact-exchange mixing (20 vs. 25%), but yield qualitatively different curves. As we are concerned with an addition reaction, the large-gradient behaviour of the exchange functional is also key in determining the shape of the potential.

It has long been established that the large-gradient behaviour of the exchange functional controls the description of non-bonded repulsion.[79,80] This occurs because atomic densities are piece-wise exponential; thus, the reduced density gradient is large far from isolated atoms or molecules. When these fragments are brought together to form a complex, the gradient is low in the intermolecular region (and identically zero at any intermolecular bond critical points). The energy difference to form the complex is therefore determined by the difference between the energy contributions from the large-gradient and low-gradient regions. As all GGAs approach the same (uniform-electron-gas) low-gradient limit, the interaction energies are consequently controlled by the large-gradient behaviour of the functional.

The exchange energy density for a general GGA functional is given by

$$E_X^{\mathrm{GGA}} = -c_X \int \rho^{4/3} F(\chi) d\mathbf{r}, \tag{2}$$

where ρ is the electron density, $\chi = |\nabla\rho|/\rho^{4/3}$ is the reduced density gradient, and c_X is a constant. It has been observed[81–83] that, for non-bonded interactions, the optimum large-gradient behaviour for the GGA enhancement factor $F(\chi)$ is

$$\lim_{\chi \to \infty} F(\chi) \sim \chi^{2/5}, \tag{3}$$

which is obtained by the PW86[84] and B86b[85] GGA exchange functionals. B88 exchange,[27] used in BLYP, has a higher large-gradient limit, which results in too much non-bonded repulsion. Conversely, PBE exchange[25] underestimates the large-gradient limit of the enhancement factor, which results in a spurious non-bonded attraction. It is notable that the Minnesota functionals, including M06-2X, underestimate this large-gradient limit, such that there is a non-bonded attraction that mimics dispersion binding near the minimum-energy separations of van-der-Waals complexes.[86] However, this attraction dies off quickly with intermolecular separation, due to the piece-wise exponential nature of the atomic densities, rather than following the known $1/R^6$ behaviour of the dispersion energy.

To demonstrate the competing effects of delocalization error and treatment of non-bonded repulsion, we consider three GGA functionals with varying large-gradient limits of the enhancement factor: PBE, PW86PBE, and BLYP. As can be seen from the results in Figure 5, all three GGAs give identical charges on the thiolate moiety for the entire range of the PES. This is consistent with delocalization error having the same effect on the degree of charge transfer from the thiolate to MVK. Nevertheless, there are significant differences in the resulting potential energy surfaces from the three GGAs, particularly at short C_β–S distances, demonstrating the impact of the exchange enhancement factor. Indeed, the effect of enhancement factor is exaggerated here, relative to neutral intermolecular complexes. This occurs because the electron density in anions is much more diffuse than in neutral molecules, so there will be a larger change in reduced gradient between the isolated monomers and the complex. With BLYP, both the delocalization and enhancement-factor errors accentuate the intermolecular repulsion, destabilizing the enolate relative to the charge-transfer

complex. However, with PBE, the delocalization error will again overstabilize the complex, but the incorrect limit of the enhancement factor will cause spurious attraction, stabilizing the enolate. These errors offset, leading to the improved performance of PBE-D3BJ, and particularly of the PBE0-D3BJ hybrid, where delocalization error is already reduced. Thus, the performance of any given density functional is determined by the interplay between these competing errors.

To decompose the effects of delocalization error and enhancement factor, additional calculations were performed with PBE0 using self-consistent Hartree–Fock (HF) densities (PBE0@HF). This approach is popularly known as density-corrected DFT[76] and the use of a HF density eliminates density-driven delocalization error.[64,74] As expected, the potential-energy well for the thiolate is deepened with PBE0@HF; higher PES values are obtained at greater $C_\beta-S$ distances in Figure 6 because of the enhancement-factor error of the PBE exchange functional.

To examine more systematically the dependence of the potential-energy curves and charge transfer on the extent of exact-exchange mixing, we consider a family of hybrid functionals based on the PW86PBE GGA.[25,84] This GGA was chosen because it has the large-gradient limit of Eqn. 3, optimum for non-bonded interactions. From Figure 5, increasing exact-exchange mixing again serves to give more localized charge on the thiolate moiety at large $r(C_\beta - S)$ distances, in improved agreement with the CCSD reference charges. This systematically reduces the excess stabilization of the charge-transfer complex. Thus, the potential-energy well corresponding to the enolate intermediate is found to deepen with increased exact-exchange mixing. The 25% exact-exchange hybrid gives a minimum that is too shallow, while the 50% hyrbid gives a minimum that is too deep, but an intermediate mixing fraction could be used to give improved agreement with the CCSD(T) reference, as shown in Figure 6. This highlights the issue with using hybrid functionals to reduce delocalization error: the optimum exact-exchange mixing fraction, or range-separation parameter, is system dependent and its choice is thus highly empirical.

A strategy to address delocalization error in a more rigorous fashion is provided by the new class of exact-exchange-based density functionals, such as B05.[45,46] As B05 uses full exact exchange, it is free of delocalization error by definition and has been shown to be highly suc-

14

Figure 5: Potential energy surfaces for the addition of methylthiolate to MVK with selected density functionals (top). Also shown are computed Hirshfeld charges on the methylthiolate group as a function of $C_\beta-S$ interatomic separation (bottom). All calculations used the aug-cc-pVTZ basis set and the MP2-optimized geometries.
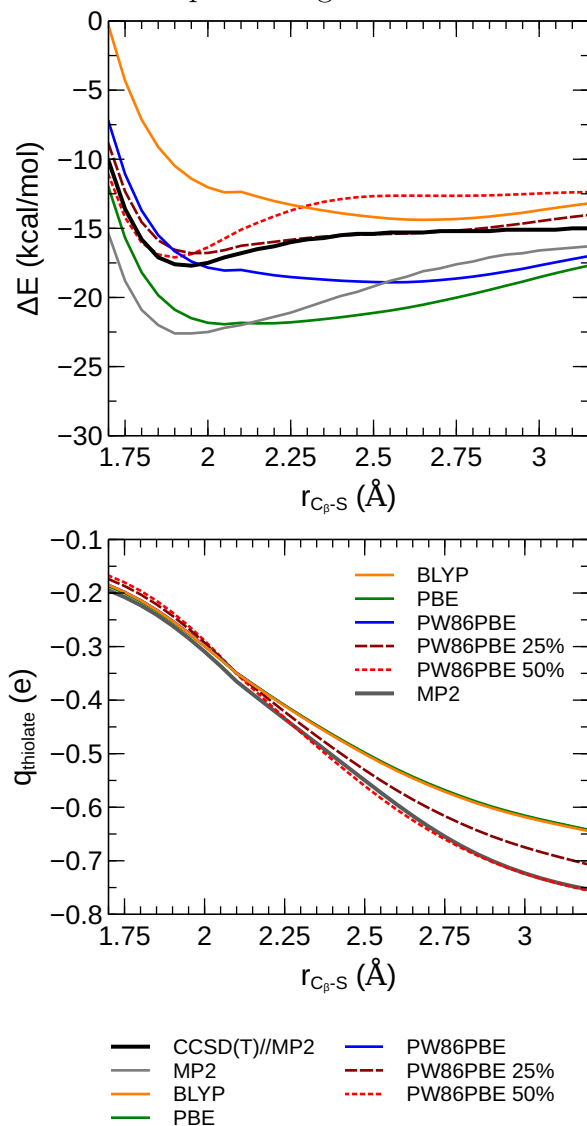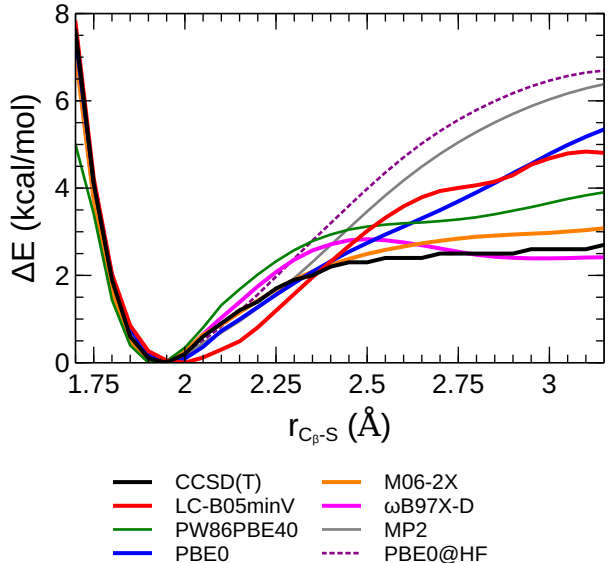
Figure 6: Potential energy surfaces of the addition of methylthiolate to MVK for methods that predict a stable enolate intermediate. The energies are referenced to the potential energy minimum.


cessful for barrier heights of radical reactions[87] and for charge transfer.[42] A self-consistent implementation of this functional is extremely complex,[88–91] so a variational approach, termed LC-B05minV,[42] has been employed as a proof of concept to illustrate that B05 can reduce or eliminate delocalization error for challenging systems. Herein, LC-B05minV was used to calculate the potential energy surface at the MP2-optimized geometries. As shown in Figure 6, LC-B05minV gives improved performance for the PES, predicting the existence of the enolate minimum at $r(C_\beta - S) \approx 1.95$ Å. LC-B05minV also gives charges that are in better agreement with the CCSD results, with a thiolate charge of ca. $-0.8e^-$ at the highest $C_\beta-S$ distances considered. Other features, such as the dissociation limit and curvature around the potential energy minimum, are less satisfactory. An additional drawback of this approach is that the computational complexity of the variational determination of the energy is considerably higher than for conventional DFT methods, where the energy can be calculated through a self-consistent. Continued development of this type of DFT functional, which is free of delocalization error, may eventually provide a functional that is generally applicable for this type of system.

## Potential Energy Surfaces for Other Electrophiles

Although MVK serves as a small, convenient system to evaluate methods for describing Michael-acceptor reactions, the Michael acceptors used in covalent-modifier drugs typically have different functional groups.[92] Specifically, acrylamide is a widely-used warhead.[4] Taunton and coworkers have also developed cyanoacrylamide warheads, which were shown to provide a prolonged residence time for an inhibitor of BTK.[3] To determine if the methods evaluated here perform the same way for reactions of these electrophiles as they do for MVK, we have calculated the reaction energies and potential energy surfaces for the addition of methylthiol to these two electrophiles. The chemical structures and model reactions for these compounds are presented in Figure 7.

Figure 7: Molecular structures of electrophilic functional groups used in pharmaceutical covalent modifiers and the structures involved in their addition to a model thiol (methylthiol).
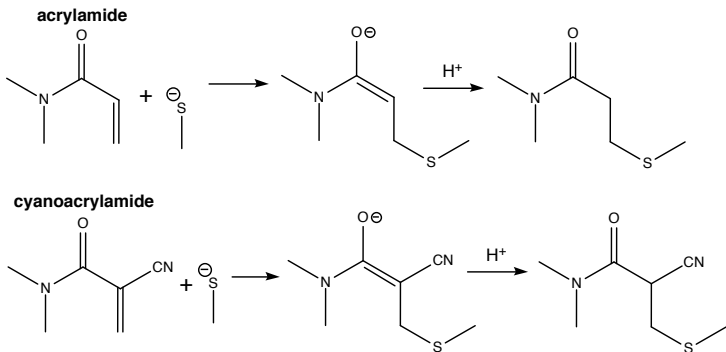
The potential energy surfaces for the addition of methylthiolate to acrylamide and cyanoacrylamide are presented in Figure 8 and the data for the minima summarized in Table 1. The trends for acrylamide are generally similar to those for MVK, with PBE-D3BJ and B3LYP failing to predict a stable enolate intermediate, while M06-2X, PBE0-D3BJ, ωB97X-D, and CCSD(T) do predict a stable enolate intermediate for the addition reaction. ωB97X-D is less accurate for this electrophile because it underestimates the stability of the enolate intermediate by roughly 2 kcal/mol. PBE0-D3BJ and M06-2X are in very good agreement with the CCSD(T) potential energy surface for this reaction. The acrylamide enolate intermediate is considerably less stable than that of the MVK enolate intermediate, with a CCSD(T) potential energy minimum of roughly −10 kcal/mol, versus −17 kcal/mol for MVK.

17

Figure 8: Potential energy surfaces for the addition of methylthiolate to acrylamide (top) and cyanoacrylamide (bottom).
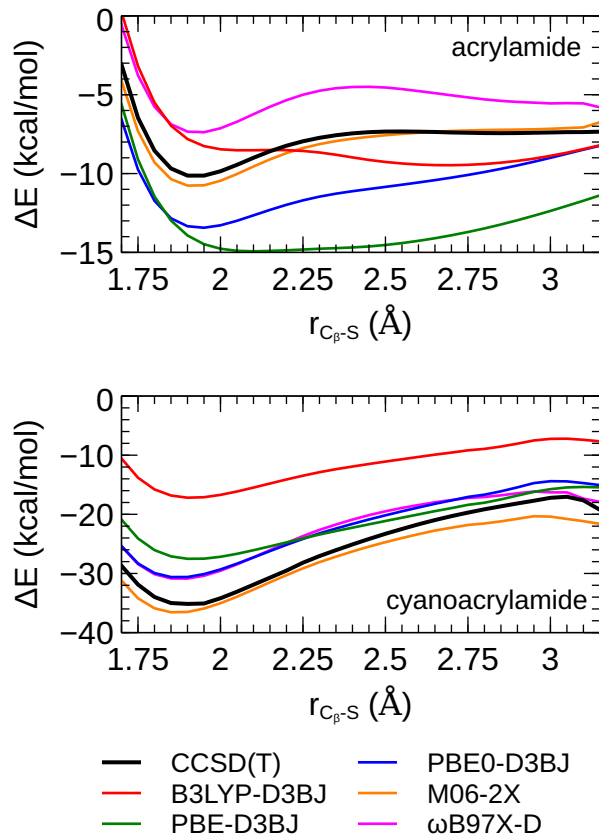
The addition of methylthiolate to cyanoacrylamide has an exceptionally stable enolate intermediate (CCSD(T): $\approx$ -35 kcal/mol). This results from delocalization of the C_α-centered anion to two electron-withdrawing functional groups (i.e., acrylamide and cyano). This intermediate is so stable that even B3LYP and PBE predict enolate minima for this reaction. Here, the PBE0, M06-2X, and ωB97X-D potential energy surfaces are all in reasonable agreement with the CCSD(T) data.

Table 1: Minimum-energy $C_\beta-S$ bond lengths (Å) and energies for formation of the enolate from the reaction of methylthiolate with MVK, acrylamide, and cyanoacylamide (kcal/mol). CCSD(T) energies were calculated at the MP2-optimized structures.

Method	MVK		acrylamide		cyanoacrylamide	
	$r_{C_\beta-S}$	$\Delta E_{\text{enolate}}$	$r_{C_\beta-S}$	$\Delta E_{\text{enolate}}$	$r_{C_\beta-S}$	$\Delta E_{\text{enolate}}$
B3LYP-D3BJ	2.57	-19.0	1.99	-10.2	1.90	-32.3
PBE-D3BJ	2.09	-25.6	2.03	-16.1	1.91	-42.6
PBE0-D3BJ	1.95	-23.1	1.92	-14.9	1.87	-38.0
ωB97X-D	1.94	-16.8	1.87	-9.4	1.87	-32.6
M06-2X	1.93	-19.0	1.87	-9.3	1.87	-38.3
CCSD(T)	1.94	-17.7	1.87	-12.2	1.86	-35.2

Low-Cost Computational Methods

QM/MM MD simulations of reactions involving biomacromolecules often employ approximate QM models with a lower computational cost than conventional DFT or *ab initio* methods. We compare the potential energy surfaces from five of these methods (AM1, PM3, PM7, SCC-DFTB, and DFTB3) to the CCSD(T) potential energy surface for the addition of methylthiolate to MVK, acrylamide, and cyanoacrylamide in Figure 9.

The semi-empirical quantum-mechanical (SQM) methods, AM1, PM3, and PM7, predict qualitatively correct potential energy surfaces for all three reactions. In each case, there is a stable enolate intermediate, with the minimum-energy C_β–S bond length in an appropriate range (1.7–1.9 Å). The quantitative agreement is poorer; none of these methods are in

Figure 9: Potential energy surfaces for the addition of methylthiolate to MVK (top), acrylamide (middle), and cyanoacrylamide (bottom), using computationally-efficient semi-empirical and DFTB methods.
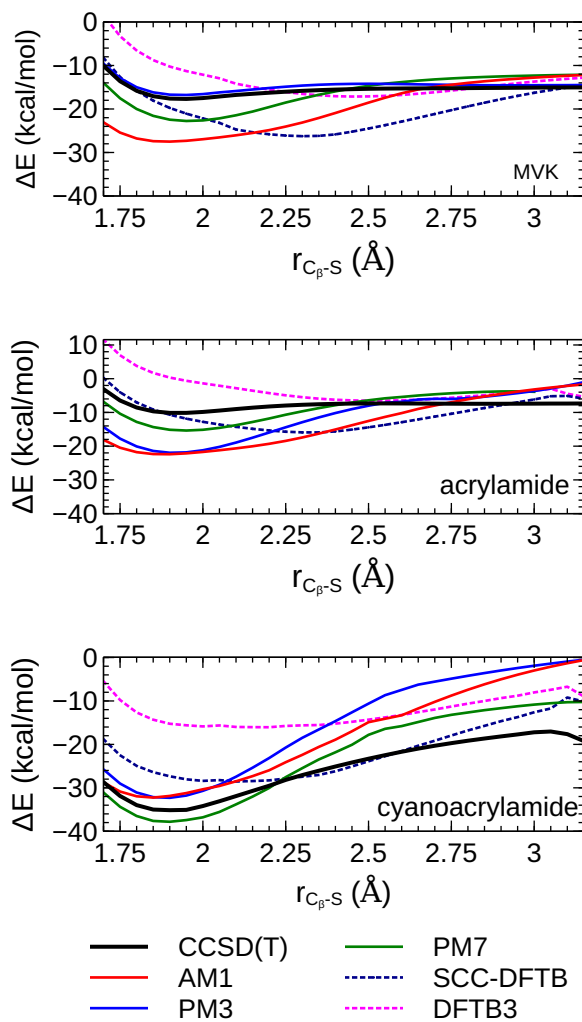
quantitative agreement with the CCSD(T) data for all three electrophiles.

The tight-binding DFT methods (SCC-DFTB and DFTB3) perform comparably poorly. Both of these methods predict broad minima corresponding to a non-covalently interacting complex, similar to those predicted by the PBE-D3BJ and B3LYP-D3BJ methods. The SCC-DFTB surfaces are systematically more negative than those of the DFTB3 surfaces, although neither method consistently agrees with the CCSD(T) surfaces. These methods appear to be generally incapable of describing the potential energy surfaces for thio-Michael addition through an ionic mechanism.

For the purposes of modeling covalent modification reactions with QM/MM simulations, the SQM methods may provide potential energy surfaces with generally correct features (i.e., a stable enolate intermediate) at a much lower computational cost. Simulations using these methods could provide an a qualitatively correct reaction profile that could then be refined using quantitatively-accurate DFT calculations. The inconsistency in predicting the stability of the enolate when the functional groups are varied suggests that these methods should only be used to compare the reaction kinetics of two different warheads if benchmark calculations show that the SQM method describes potential energy surfaces for both warheads correctly.

Reaction Energies

Depending on the electrophile, thio-Michael addition reactions can be reversible.[12,93,94] This has been utilized in the development of covalent modifiers that bind reversibly, despite the formation of a covalent linkage. Accurate computational prediction of the reversibility of a thio-Michael addition would aid continued development of these inhibitors. The net reaction energy for the addition of the thiol to the electrophile, corresponding to the energy difference between the products and reactants shown in Figure 1, provides a metric to assess the ability of these computational methods to predict this activity. The reaction energies for the addition of methylthiol to the three electrophiles with both the low-cost and high-cost computational methods are collected in Table 2.

The CCSD(T) data indicates that there are modest variations in the reaction energies, with the cyanoacrylamide addition being the most reversible and the acrylamide addition being the least. The reaction energies calculated using B3LYP-D3BJ are systematically too

21

Table 2: Reaction energies for the reaction of methylthiol with MVK, acrylamide, and cyanoacylamide. All energies are in kcal/mol.

Method	MVK	acrylamide	cyanoacrylamide
AM1	-25.9	-26.2	-16.8
PM3	-21.8	-22.5	-13.4
PM7	-20.6	-18.8	-16.7
SCC-DFTB	-34.1	-35.4	-29.1
DFTB3	-22.2	-23.8	-32.4
B3LYP-D3BJ	-20.2	-22.2	-13.7
PBE-D3BJ	-23.3	-24.8	-21.5
PBE0-D3BJ	-26.1	-27.7	-19.0
ωB97X-D	-23.1	-24.9	-16.5
M06-2X	-23.7	-25.4	-20.4
CCSD(T)	-23.1	-25.0	-20.0

low, by 3–6 kcal/mol, as expected from the long-known tendency of B3LYP to underestimate the strengths of chemical bonds not involving hydrogen,[95–97] with can be ameliorated by changing the correlation functional. Conversely, the reaction energies obtained from PBE0-D3BJ are too negative by ca. 3 kcal/mol for the reactions with MVK and acrylamide. The PBE-D3BJ and ωB97X-D functionals perform well, although the latter significantly underestimates the stability of the thioether product for the cyanoacrylamide addition. The differing quality of results for cyanoacrylamide is likely due to the more extended conjugation in this system, compared to the other ketones. The M06-2X results are in good agreement with CCSD(T) for all three reactions. In comparison to the enolate stabilities, these reaction energies are less substantially affected by delocalization error because they involve only neutral species rather than anions.

The low-cost computational methods perform inconsistently for calculation of the net reaction energy. The relative stability of the products is predicted inconsistently by the SQM methods, which all have a tendency to understate the stability of the cyanoacrylamide thioether product. SCC-DFTB significantly overestimates the net reaction energy in each

case, while DFTB3 performs well for MVK and acrylamide, but again significantly over-estimates the stability of thioether product for cyanoacrylamide. Generally, none of these methods are reliable for predicting the thermochemistry of these thiol-addition reactions.

Numerical Issues for DFT Molecular Dynamics

QM/MM molecular dynamics simulations rely on equations of motion to generate configurations of the targeted thermodynamic ensemble. As a result, the gradients calculated using DFT methods must be accurate in order to generate the next state in the ensemble correctly. The standard implementation of DFT using a Gaussian basis set requires numerically integrating the exchange-correlation term on a grid and its required fineness depends on the particular form of the exchange-correlation functional. For example, meta-GGA functionals have been shown to require finer grids than for GGA functionals, because certain terms that depend on the kinetic-energy density are ill-behaved in intermolecular regions.[86,98–100]

To identify if this could be an issue for the DFT functionals explored in this study, we performed 1 ps of velocity Verlet molecular dynamics with QM models for the acrylamide warhead using NAMD-ORCA. A simulation with accurate gradients will conserve the total energy (i.e., sample the microcanononical ensemble), but the total energy of a simulation where the gradients have significant numerical error will change over time. The total energy over the course of simulations using the PBE0, ωB97X-D3BJ, and M06-2X functionals are plotted in Figure 10 (top). The simulations using the PBE0-D3BJ and ωB97X-D functionals show no significant deviation in total energy over the course of the simulation, indicating that the gradients are computed accurately using the standard grid. The simulation using the M06-2X functional with the default grid shows large deviations in the total energy. This suggests a significant inconsistency in the calculated gradients, which could result in a non-representative ensemble of states if this method were used in QM/MM MD. This issue is partially attenuated when a finer DFT grid is used (i.e., ORCA keyword FinalGrid6), although the computational cost of calculations using this finer grid is approximately 20% higher (Figure 10, bottom). It is possible that these issues with the M06-2X functional can be rectified by using the newer range-separated hybrid-meta GGA functional, M11.[101] However, this functional was not available in the ORCA code used in our QM/MM molecular

Figure 10: (Top) Plots of the total energy from a velocity Verlet molecular dynamics simulation of acrylamide. The total energy of the simulations using PBE0-D3BJ and ωB97X-D3BJ overlap on the x-axis, indicating that the total energy is tightly conserved. The simulation using the M06-2X functional with the standard grid (FinalGrid4, blue) shows very poor energy conservation. The energy conservation is improved when a finer grid is used (FinalGrid6, green); however, there a still a small but systematic drift in the total energy over the course of the 1 ps simulation that is not present for simulations using the PBE0-D3BJ or ωB97X-D3BJ functionals. The def2-SVP basis set was used in these simulations. (Bottom) The number of MD steps calculated per hour using 4 processors on a 3.20 GHz Intel(R) Xeon(R) CPU E5-2667 v4 using NAMD interfaced to ORCA 4.0.1.2.
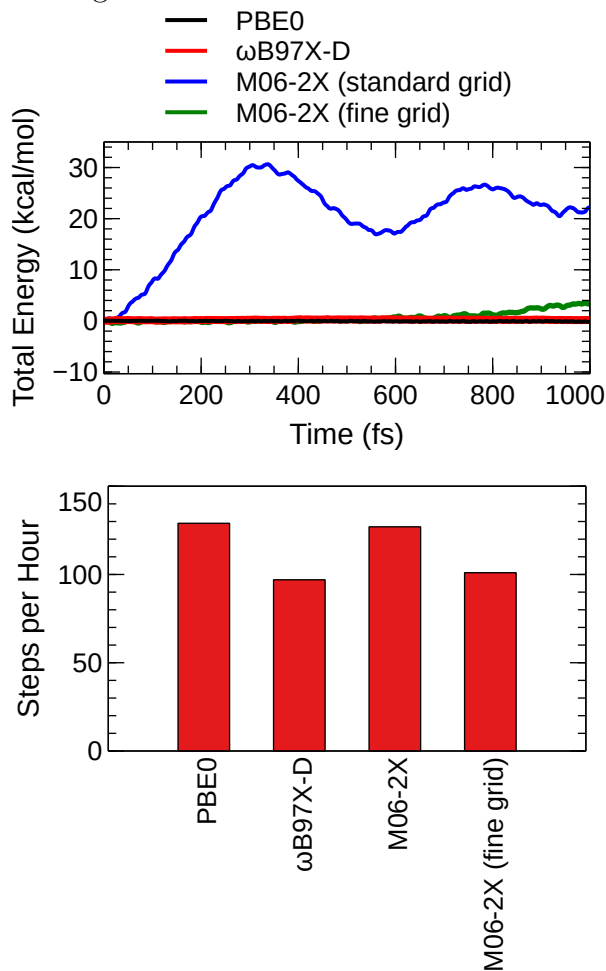
Calculation of Thiol Gas Phase Acidity

Thio-Michael additions begin with the deprotonation of the thiol to form a nucleophilic, anionic thiolate, so the propensity of a cysteine thiol to be deprotonated is an important quantity in modeling its modification.[16,102,103] The ability of computational methods to predict the absolute deprotonation energy can be tested by comparing to the experimental estimate of the gas phase acidity of methylthiol. The enthalpies of acid dissociation ($\Delta_{\text{acid}}H$) calculated using the various high- and low-cost methods are collected in Table 3. The DFT and CCSD(T) results are in good agreement with the experimental value of 357.5 kcal/mol. However, the semi-empirical methods are less reliable, with errors greater than 20 kcal/mol for all three variants.

The proton affinity of methylthiol calculated using SCC-DFTB is 6 kcal/mol lower than the experimental value. This is opposite to the typical trend for SCC-DFTB predictions of organic acids; for example, the proton affinity of acetic acid is overestimated by 11 kcal/mol.[104] This suggests that this method will underestimate the energetic cost for a cysteine residue to transfer a proton to glutamate or aspartate by roughly 17 kcal/mol. This is a significant issue in modeling covalent-modifier chemistry because glutamate and asparate residues have been previously proposed to abstract a proton from the cysteine undergoing modification.[105,106] This is counter to trend predicted by experimental pK$_a$'s, which predict that the transfer of a proton from an alkylthiol to a carboxylate will be significantly non-spontaneous. DFTB3 provides accurate gas-phase proton affinities for both methylthiol and acetic acid, so QM/MM MD simulations using this method should provide more accurate models for reactions involving cysteine proton transfers than those using SCC-DFTB.

Potential of Mean Force for the Reaction

The potential energy surfaces presented in the previous sections of this paper are adiabatic surfaces that do not account for the effects of solvation and thermal fluctuations on the reaction. Modeling the kinetics of covalent modification using transition-state theory will

Table 3: Computed gas-phase proton affinities of methylthiol, in kcal/mol. Calculations were performed using the aug-cc-pVTZ basis set. An enthalpy correction using MP2 harmonic frequency analysis was included. The rightmost column shows the difference between the calculated and experimental[107] values.

Method	$\Delta_{\mathrm{acid}}H$	$\Delta H_{\mathrm{calc}} - \Delta H_{\mathrm{exptl}}$
AM1	337.3	-20.2
PM3	379.6	22.1
PM7	285.4	-72.1
SCC-DFTB	351.0	-6.5
DFTB3	357.9	0.4
PBE-D3BJ	355.1	-2.4
B3LYP-D3BJ	356.5	-1.0
PBE0-D3BJ	357.4	-0.1
ωB97X-D	358.6	1.1
M06-2X	355.0	-2.5
CCSD(T)//MP2	357.4	-0.1
Exptl.[107]	357.5	

require the calculation of the PMF in condensed phases (e.g., solution or the protein binding site). As a proof of concept, we have used QM/pol-MM MD umbrella sampling simulations to calculate the PMF for the reaction of methylthiolate with MVK in an explicit aqueous solution. The interval where the C_β–S bond is forming was treated using a QM/pol-MM model ($r = 1.7 - 4$ Å) while the interval where the two solutes are associating ($r = 4 - 8$ Å) was treated using a polarizable MM model. The PMF calculated from these combined surfaces is presented in Figure 11. The ωB97X-D functional was used to describe the QM region, which is in generally good agreement with the CCSD(T) potential energy surface for this step of the reaction.

In the gas phase, there is a strong attractive force between the polar MVK and the anionic methylthiolate, so there is a minimum of -15.5 kcal/mol at a C_β–S distance of 3.5 Å. MVK and methylthiolate form a contact pair at this distance. The minimum corresponding

Figure 11: The PMF for the reaction of MVK and methylthiolate in an aqueous solution (blue curve) and in gas phase (red curve). The PMF in the $r = [4.0, 10.0]$ Å interval was calculated from a molecular dynamics simulation using the Drude polarizable force field. The PMF in the $r = [1.7, 4.0]$ Å interval was calculated using QM/MM MD (QM: ωB97X-D3BJ//def2-TZVP, MM: Drude).
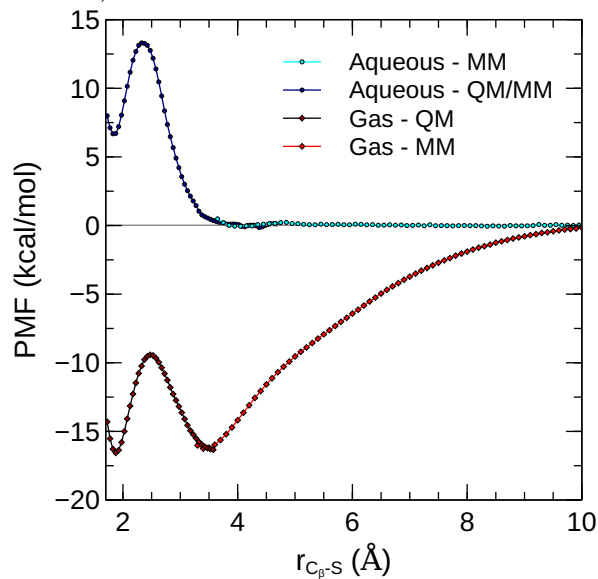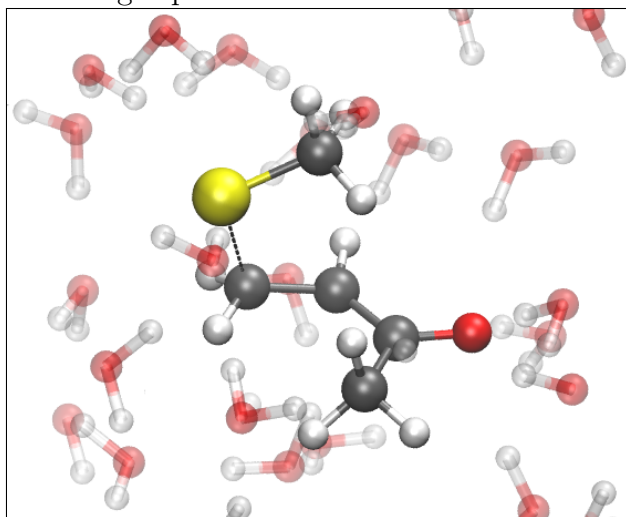
Figure 12: A representative configuration from QM/MM simulation of the transition state for the addition of methylthiolate to methyl vinyl ketone. The thiolate has a lower hydration number at the transition state than as a free ion in solution, which results in a higher barrier on the solution PMF than the gas phase PMF.

to the enolate is approximately as stable as this contact pair, although there is an activation energy of roughly 7 kcal/mol separating these minima. This barrier is not present on the potential energy surface (Figure 4), indicating that it originates from a loss of entropy.

There are no significant barriers to the association of the two reactants in solution (i.e., $r > 4$ Å). This consistent with thiolates having disordered hydration spheres,[52] so that MVK can enter into a contact pair with the methylthiolate without incurring the enthalpic cost that occurs for ions with tightly-associated water molecules.[108,109] The enolate is roughly 7 kcal/mol less stable than this contact pair and there is a significant barrier in the $r = 2.0-3.5$ Å interval, with a maximum height of 12.5 kcal/mol near $r = 2.5$ Å. A representative configuration at this point is presented in Figure 12. The hydration number of the S atom drops from ~ 5 water molecules to ~ 4 water molecules for $r < -3.5$ Å. This loss of stabilizing anion–water interactions, compared to the separated reactants, results in a higher barrier on the PMF and a less-stable enolate intermediate (1.7 Å $< r < 2$ Å), relative to the gas-phase PMF.

Figure 13: Transition state structure for the 1,2-σ addition of methylthiol to MVK. Key bond lengths are indicated adjacent to the bonds (Å)
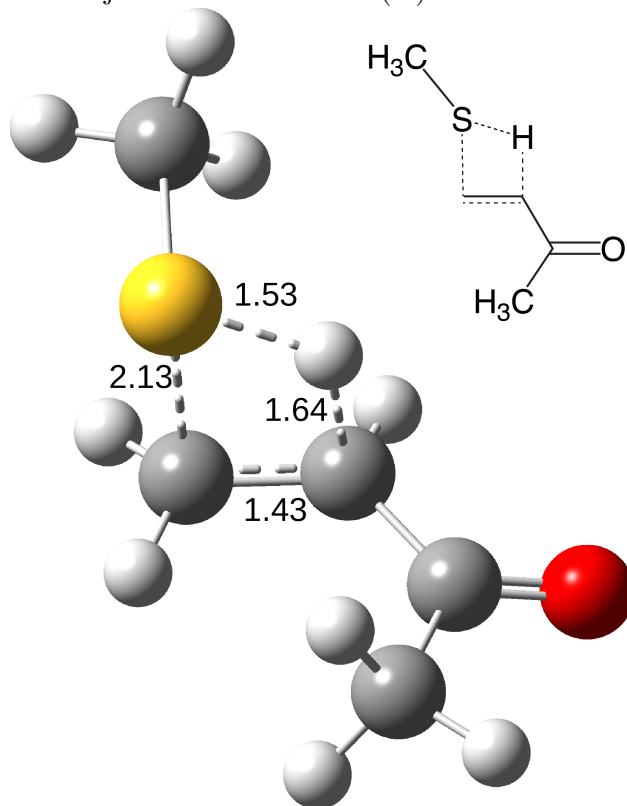
Alternative 1,2-σ Addition Mechanism

In response to questions as to whether a direct σ-addition of the alkyl thiol to the Michael acceptor is also possible, we have calculated the transition state structure for the addition of methyl thiol to MVK by this mechanism (Figure 13). The Gibbs energy of activation for this reaction, calculated using CCSD(T)/aug-cc-pVTZ//ωB97X-D/aug-cc-pVTZ, is 65.2 kcal/mol relative to methyl thiol and MVK. This high barrier is consistent with this reaction being thermally forbidden by orbital-symmetry rules. Because this barrier is extremely high, we chose to exclude this mechanism from further consideration. This mechanism is also inconsistent with the pH dependence on the thiol-Michael reaction rates that has been observed experimentally.[110]

CONCLUSIONS

Electronic-structure methods were evaluated for their ability to describe the potential energy surfaces and thermodynamics of the reaction between a model thiolate and the Michael acceptor MVK. In contrast to CCSD(T) calculations, several functionals, including the popular B3LYP and PBE fuctionals, failed to predict a potential-energy minimum for the enolate intermediate. This is caused by delocalization error, which results in the spurious stabilization of a charge-transfer complex instead of the enolate intermediate.

Various strategies to address these issues within DFT were explored. The M06-2X functional performs reasonably well due to the large component of exact exchange applied globally. The PBE0-D3BJ functional gives improved performance compared to B3LYP-D3BJ, despite a similar exact-exchange mixing. This results from a fortuitous cancellation of error, as the PBE exchange functional underestimates non-bonded repulsion due to an incorrect large-gradient limit of the exchange enhancement factor. The ωB97X-D functional also performs well because its exchange-correlation functional is range-separated, so points with larger interelectronic distances are treated with a large component of exact exchange; such functionals were designed to give an improved treatment of charge transfer compared to global hybrids. The variational LC-B05minV functional was also shown to resolve the issues stemming from delocalization error through optimization of the electron density with respect to long-range exact-exchange mixing. As the interplay of these effects has a dramatic effect on the predictions for the thiolate/MVK PES, this system serves as a challenging test of DFT methods. Notably, it highlights the necessity of an accurate treatment of both electron delocalization and non-bonded repulsion in functional construction.

The benchmark was extended to include the acrylamide and cyanoacrylamide warheads, which are more relevant to pharmaceutical chemistry. As in the case of MVK, the PBE-D3BJ and B3LYP-D3BJ functionals were generally unreliable for describing the potential energy surface of the enolate, while PBE0-D3BJ, M06-2X, and ωB97X-D performed relatively well. The ability of these methods to predict the net reaction energy for addition of the thiol to the Michael acceptor to form a thioether was also evaluated. The reaction energies predicted by the B3LYP-D3BJ functional were systematically low, while those for the PBE0-D3BJ

functional were incrementally too high. The other DFT methods generally performed well.

Low-cost computational methods were evaluated for use in QM/MM MD simulations. The AM1, PM3, and PM7 semi-empirical methods provided qualitatively correct potential energy surfaces. This suggests they may have utility for modeling the general features of covalent-modification reactions at a much reduced computational cost, in comparison to more rigorous DFT methods. The most significant drawback of these low-cost methods is that the errors in the potential energy surfaces were not systemic between MVK, acrylamide, and cyanoacylamide. This suggests that these methods are not suitable for modeling the relative substituent effects of electrophilic warheads. SCC-DFTB and DFTB3 did not predict a stable enolate intermediate and SCC-DFTB was significantly in error for the proton affinity of methylthiol. These methods are not recommended for quantitative models of thio-Michael additions.

For practical simulations, the PBE0, ωB97X-D, and M06-2X functionals all perform reasonably well. These functionals also perform well for the calculation of electric moments and polarizabilities, so they are also well-suited for modeling electrostatic effects in thiolate reactions.[111] The M06-2X functional has the drawback that the calculation of gradients is inaccurate, unless a very fine integration grid is used to evaluate the exchange-correlation terms. The PBE0-D3BJ and ωB97X-D functionals are most appropriate for modeling these reactions, although innovations in DFT functionals that address delocalization error may provide improved solutions. The ability to predict a stable enolate for a thiolate–Michael-acceptor system is a valuable test for issues related to delocalization error in new DFT functionals and is an essential test for any method used to model the covalent modification of cysteine residues.

ACKNOWLEDGMENTS

Additional Supporting Information may be found in the online version of this article.

References

1. M. H. Potashman and M. E. Duggan, J. Med. Chem. **52**, 1231 (2009).

2. M. S. Cohen, C. Zhang, K. M. Shokat, and J. Taunton, Science **308**, 1318 (2005).

3. J. M. Bradshaw, J. M. McFarland, V. O. Paavilainen, A. Bisconte, D. Tam, V. T. Phan, S. Romanov, D. Finkle, J. Shu, V. Patel, et al., Nat. Chem. Biol. **11**, 525 (2015).

4. P. A. Jackson, J. C. Widen, D. A. Harki, and K. M. Brummond, J. Med. Chem. **60**, 839 (2017).

5. L. A. Honigberg, A. M. Smith, M. Sirisawad, E. Verner, D. Loury, B. Chang, S. Li, Z. Pan, D. H. Thamm, R. A. Miller, et al., Proc. Natl. Acad. Sci. U.S.A. **107**, 13075 (2010).

6. E. Awoonor-Williams, A. G. Walsh, and C. N. Rowley, Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics **1865**, 1664 (2017).

7. P. Chatterjee, W. M. Botello-Smith, H. Zhang, L. Qian, A. Alsamarah, D. Kent, J. J. Lacroix, M. Baudry, and Y. Luo, J. Am. Chem. Soc. **139**, 17945 (2017).

8. H. Zhang, W. Jiang, P. Chatterjee, and Y. Luo, J. Chem. Inf. Model **0**, null (2019).

9. E. H. Krenske, R. C. Petter, Z. Zhu, and K. N. Houk, J. Org. Chem. **76**, 5074 (2011).

10. M. E. Flanagan, J. A. Abramite, D. P. Anderson, A. Aulabaugh, U. P. Dahal, A. M. Gilbert, C. Li, J. Montgomery, S. R. Oppenheimer, T. Ryder, et al., J. Med. Chem. **57**, 10072 (2014).

11. V. J. Cee, L. P. Volak, Y. Chen, M. D. Bartberger, C. Tegley, T. Arvedson, J. McCarter, A. S. Tasker, and C. Fotsch, J. Med. Chem. **58**, 9171 (2015).

12. J. M. Smith and C. N. Rowley, J. Comput.-Aided Mol. Des. **29**, 725 (2015).

13. D. S. Allgauer, H. Jangra, H. Asahara, Z. Li, Q. Chen, H. Zipse, A. R. Ofial, and H. Mayr, J. Am. Chem. Soc. **139**, 13318 (2017).

14. R. Lonsdale, J. Burgess, N. Colclough, N. L. Davies, E. M. Lenz, A. L. Orton, and R. A. Ward, J. Chem. Inf. Model. **57**, 3124 (2017).

15. I. M. Serafimova, M. A. Pufall, S. Krishnan, K. Duda, M. S. Cohen, R. L. Maglathlin, J. M. McFarland, R. M. Miller, M. Frödin, and J. Taunton, Nat. Chem. Biol. **8**, 471 (2012).

16. M. A. R. Raycroft, K. E. Racine, C. N. Rowley, and J. W. Keillor, J. Org. Chem. **83**, 11674 (2018).

17. D. Mulliner, D. Wondrousch, and G. Schuurmann, Org. Biomol. Chem. **9**, 8400 (2011).

18. C. Wang and C. Qi, Tetrahedron **69**, 5348 (2013).

19. S. R. Pereira, V. M. Vasconcelos, and A. Antunes, FEBS J. **280**, 674 (2013).

20. D. Callegari, K. E. Ranaghan, C. J. Woods, R. Minari, M. Tiseo, M. Mor, A. J. Mulholland, and A. Lodola, Chem. Sci. **9**, 2740 (2018).

21. J. M. Smith, Y. Jami Alahmadi, and C. N. Rowley, J. Chem. Theory Comput. **9**, 4860 (2013).

22. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., *Gaussian 16 Revision A.03* (2016), Gaussian Inc. Wallingford CT.

23. R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

24. D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).

25. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

26. C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

27. A. D. Becke, J. Chem. Phys. **88**, 2547 (1988).

28. C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

29. S. Grimme, S. Ehrlich, and L. Goerigk, J. Comp. Chem. **32**, 1456 (2011).

30. J.-D. Chai and M. Head-Gordon, Phys. Chem. Chem. Phys. **10**, 6615 (2008).

31. Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).

32. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).

33. J. J. P. Stewart, J. Comput. Chem. **10**, 209 (1989).

34. J. J. P. Stewart, J. Mol. Model. **19**, 1 (2013).

35. E. R. Johnson, in *Non-covalent Interactions in Quantum Chemistry and Physics*, edited by A. Otero-de-la-Roza and G. A. DiLabio (Elsevier, 2017), chap. 5, pp. 169–194.

36. *TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from*
`http://www.turbomole.com`.

37. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Phys. Rev. B **58**, 7260 (1998).

38. T. Niehaus, M. Elstner, T. Frauenheim, and S. Suhai, J. Mol. Struct. THEOCHEM **541**, 185 (2001).

39. M. Gaus, Q. Cui, and M. Elstner, J. Chem. Theory Comput. **7**, 931 (2011).

40. M. Gaus, X. Lu, M. Elstner, and Q. Cui, J. Chem. Theory Comput. **10**, 1518 (2014).

41. B. Aradi, B. Hourahine, and T. Frauenheim, J. Phys. Chem. A **111**, 5678 (2007).

42. A. D. Becke, S. G. Dale, and E. R. Johnson, J. Chem. Phys. **148**, 211101 (2018).

43. H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, J. Chem. Phys **115**, 3540 (2001).

44. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., *Gaussian 09 Revision D.01*, gaussian Inc. Wallingford CT 2009.

45. A. D. Becke, J. Chem. Phys. **122**, 064101 (2005).

46. S. G. Dale, E. R. Johnson, and A. D. Becke, J. Chem. Phys. **147**, 154103 (2017).

47. F. O. Kannemann and A. D. Becke, J. Chem. Theory Comput. **6**, 1081 (2010).

48. A. Otero-de-la Roza and E. R. Johnson, J. Chem. Phys. **138**, 204109 (2013).

49. The postg program is available from http://schooner.chem.dal.ca.

50. L. Huang and B. Roux, J. Chem. Theory Comput. **9**, 3543 (2013).

51. F.-Y. Lin, P. E. M. Lopes, E. Harder, B. Roux, and A. D. MacKerell, J. Chem. Inf. Model. **58**, 993 (2018).

52. E. Awoonor-Williams and C. N. Rowley, J. Chem. Phys. **149**, 045103 (2018).

53. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kal, and K. Schulten, J. Comput. Chem. **26**, 1781 (2005).

54. F. Neese, Wiley Interdiscip. Rev. Comput. Mol. Sci **2**, 73 (2012).

55. F. Neese, Wiley Interdiscip. Rev. Comput. Mol. Sci **8**, e1327 (2018).

56. M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone, et al., Nat. Methods **15**, 351 (2018).

57. F. Neese, F. Wennmohs, A. Hansen, and U. Becker, Chem. Phys. **356**, 98 (2009).

58. C. N. Rowley and B. Roux, J. Chem. Theory Comput. **8**, 3526 (2012).

59. S. Riahi and C. N. Rowley, J. Comput. Chem. **35**, 2076 (2014).

60. B. Roux, Comput. Phys. Commun. **91**, 275 (1995), ISSN 0010-4655.

61. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, J. Comput. Chem. **13**, 1011 (1992), ISSN 1096-987X.

62. D. Trzesniak, A.-P. E. Kunz, and W. F. van Gunsteren, ChemPhysChem **8**, 162 (2007).

63. F. L. Hirshfeld, Theor. Chim. Acta **44**, 129 (1977).

64. M.-C. Kim, H. Park, S. Son, E. Sim, and K. Burke, J. Phys. Chem. Lett. **6**, 3802 (2015).

65. M.-C. Kim, E. Sim, and K. Burke, Phys. Rev. Lett. **111**, 073003 (2013).

66. E. R. Johnson, A. Otero-de-la Roza, and S. G. Dale, J. Chem. Phys **139**, 184116 (2013).

67. A. D. Becke, J. Chem. Phys. **140**, 18A301 (2014).

68. A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, J. Chem. Phys. **125**, 194112 (2006).

69. A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).

70. O. A. Vydrov, J. Heyd, A. V. Krukau, and G. E. Scuseria, J. Chem. Phys. **125**, 074106 (2006).

71. S. N. Steinmann, C. Piemontesi, A. Delacht, and C. Corminboeuf, J. Chem. Theory Comput. **8**, 1629 (2012).

72. G. Sini, J. S. Sears, and J. L. Bredas, J. Chem. Theory Comput. **7**, 602 (2011).

73. B. G. Janeskoa and G. E. Scuseria, J. Chem. Phys. **128**, 244112 (2008).

74. A. Wasserman, J. Nafziger, K. Jiang, M.-C. Kim, E. Sim, and K. Burke, Ann. Rev. Phys. Chem. **68**, 555 (2017).

75. C. Zhang, T. A. Pham, F. Gygi, and G. Galli, J. Chem. Phys. **138**, 181102 (2013).

76. M.-C. Kim, E. Sim, and K. Burke, J. Chem. Phys. **140**, 18A528 (2014).

77. S. R. Whittleton, A. Otero-de-la-Roza, and E. R. Johnson, J. Chem. Theory Comput. **13**, 441 (2017).

78. L. M. LeBlanc, S. G. Dale, C. R. Taylor, A. D. Becke, G. M. Day, and E. R. Johnson, Angew. Chem. Int. Ed. **57**, 14906 (2018).

79. J. M. Perez-Jorda and A. D. Becke, Chem. Phys. Lett. **233**, 134 (1995).

80. Y. Zhang, W. Pan, and W. Yang, J. Chem. Phys. **107**, 7921 (1997).

81. D. J. Lacks and R. G. Gordon, Phys. Rev. A **47**, 4681 (1993).

82. F. O. Kannemann and A. D. Becke, J. Chem. Theory Comput. **5**, 719 (2009).

83. Éamonn D. Murray, K. Lee, and D. C. Langreth, J. Chem. Theory Comput. **5**, 2754 (2009).

84. J. Perdew and W. Yue, Phys. Rev. B **33**, 8800 (1986).

85. A. Becke, J. Chem. Phys. **85**, 7184 (1986).

86. E. R. Johnson, A. D. Becke, C. D. Sherrill, and G. A. DiLabio, J. Chem. Phys, **131**, 034111 (2009).

87. R. M. Dickson and A. D. Becke, J. Chem. Phys. **123**, 111101 (2005).

88. E. Proynov, F. Liu, Y. Shao, and J. Kong, J. Chem. Phys. **136**, 034102 (2012).

89. A. V. Arbuznikov and M. Kaupp, J. Chem. Phys. **131**, 084103 (2009).

90. E. Proynov, F. Liu, and J. Kong, Chem. Phys. Lett. **525-526**, 150 (2012).

91. E. Proynov, F. Liu, Y. Shao, and J. Kong, J. Chem. Phys. **136**, 034102 (2012).

92. M. Gehringer and S. A. Laufer, J. Med. Chem. (2018).

93. C. F. H. Allen, J. O. Fournier, and W. J. Humphlett, Can. J. Chem. **42**, 2616 (1964).

94. C. F. H. Allen and W. J. Humphlett, Can. J. Chem. **44**, 2315 (1966).

95. J. S. Wright, C. N. Rowley, and L. L. Chepelev, Mol. Phys. **103**, 815 (2005).

96. L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys **106**, 1063 (1997).

97. G. A. DiLabio and D. A. Pratt, J. Phys. Chem. A **104**, 1938 (2000).

98. J. Gräfenstein, D. Izotov, and D. Cremer, J. Chem. Phys. **127**, 214103 (2007).

99. S. E. Wheeler and K. N. Houk, J. Chem. Theory Comput. **6**, 395 (2010).

100. L. Simón and J. M. Goodman, Org. Biomol. Chem. **9**, 689 (2011).

101. R. Peverati and D. G. Truhlar, The Journal of Physical Chemistry Letters **2**, 2810 (2011).

102. E. Awoonor-Williams and C. N. Rowley, J. Chem. Theory Comput. **12**, 4662 (2016).

103. E. Awoonor-Williams and C. N. Rowley, J. Chem. Inf. Model **58**, 1935 (2018).

104. Yang, H. Yu, D. York, Q. Cui, and M. Elstner, J. Phys. Chem. A **111**, 10861 (2007).

105. E. R. Wood, L. M. Shewchuk, B. Ellis, P. Brignola, R. L. Brashear, T. R. Caferro, S. H. Dickerson, H. D. Dickson, K. H. Donaldson, M. Gaul, et al., Proc. Natl. Acad. Sci. U.S.A. **105**, 2773 (2008).

106. L. Capoferri, A. Lodola, S. Rivara, and M. Mor, J. Chem. Inf. Model. **55**, 589 (2015).

107. N. Eyet, S. M. Villano, and V. M. Bierbaum, Int. J. Mass Spectrom. **283**, 26 (2009).

108. C. J. Fennell, A. Bizjak, V. Vlachy, and K. A. Dill, J. Phys. Chem. B **113**, 6782 (2009).

109. Y. Luo, W. Jiang, H. Yu, A. D. MacKerell, and B. Roux, Faraday Discuss. **160**, 135 (2013).

110. M. P. Lutolf and J. A. Hubbell, Biomacromolecules **4**, 713 (2003).

111. A. L. Hickey and C. N. Rowley, J. Phys. Chem. A **118**, 3678 (2014).