Torsional Energy Barriers of Biaryls could be Predicted by Electron-richness/deficiency of Aromatic Rings; Advancement of Molecular Mechanics towards Atom-Type Independence

Wanlei Wei, [1] Candide Champion,[1] Zhaomin Liu,[1,‡] Stephen J. Barigye,[1,‡] Paul Labute,[2] Nicolas Moitessier[1],*

[1]Department of Chemistry, McGill University, 801 Sherbrooke St. W., Montréal, QC, Canada H3A 0B8; [2]Chemical Computing Group Inc., 1010 Sherbrooke St. W., Montréal, QC, Canada H3A 2R7.

ABSTRACT

Biaryl molecules are one of the most ubiquitous pharmacophores found in natural products and pharmaceuticals. In spite of this, existing molecular mechanics force fields are unable to accurately reproduce their torsional energy profiles, except in a few well-parametrized cases. This effectively limits the ability of structure-based drug design methods to correctly identify hits involving biaryls with confidence (eg. during virtual screening, employing docking and/or molecular dynamics simulations). Continuing in our endeavor to quantify organic chemistry principles, we showed that the torsional energy profile of biaryl compounds could be computed on-the-fly based on the electron-richness/deficiency of the aromatic rings. This method, called H-TEQ 4.0, was developed using a set of 131 biaryls. It was subsequently validated on a separate set of 100 diverse biaryls, including multi-substituted, bicyclic and tricyclic drug-like molecules, and produced an average RMSE of 0.95 kcal·mol^{-1}. For comparison, GAFF2 produced an RMSE of 3.88 kcal·mol^{-1}, owing to problems associated with the transferability of torsion parameters. The success of H-TEQ 4.0 provided further evidence that force fields could transition to become atom type-independent, providing that the correct underlying chemical principles are used. Overall, this method solved the problem of transferability of biaryl torsion parameters, while simultaneously improving the overall accuracy of the force field.

INTRODUCTION

Biaryls. Biaryls are a class of conjugated molecules comprised of two aromatic systems, connected by a single bond. These compounds are of particular interest to the pharmaceutical field as they are found in many natural products and pharmaceuticals.[1] In fact, many essential drugs such as atorvastatin, celecoxib, nifedipine, rosuvastatin, valsartan, and others contain biaryl moieties (Figure 1). Many natural products containing biaryl fragments, extracted from plants and other organisms, have frequently been shown to possess biological activity.[2] For example,

licopyranocoumarin extracted from *Xi-bei licorice* plant, was found to inhibit the cytopathic activity of HIV.[3] These types of molecules are ubiquitous in nature and are important bioactive pharmacophores. Besides their utility in pharmaceuticals, they are also well adapted for use in building polymers, sensors, and transition metal catalysts.[1]
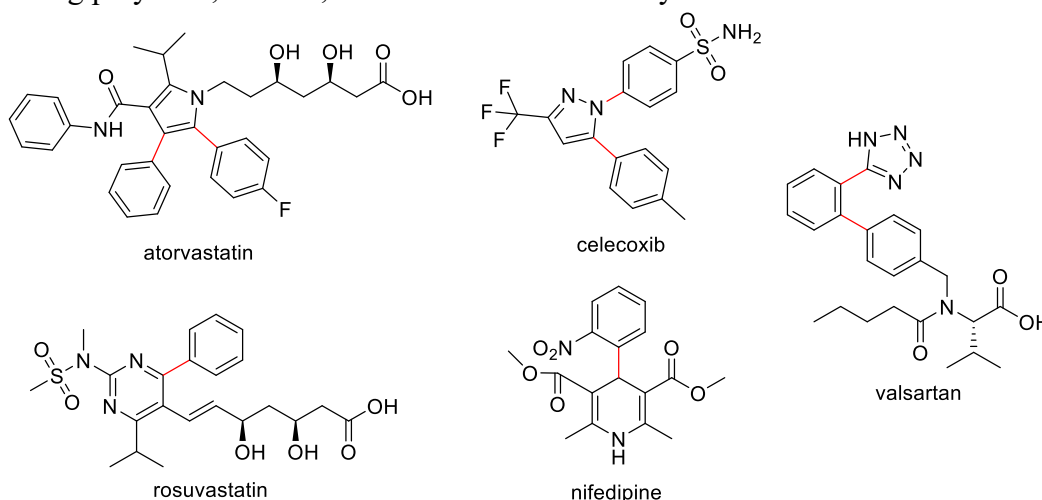
Figure 1. Commercially available drug molecules that possess biaryl fragments, with the torsion bonds of interest labelled in red.

Biaryls and molecular mechanics. Despite biaryls being promising pharmacophores and drug scaffolds, existing molecular mechanics (MM) force fields (FF) are not well-adapted to predict their conformational energy landscapes, apart from a few well-parametrized cases.[4, 5] In particular, the torsion parameters of biaryls are lacking for most compounds, due to the diversity of biaryls in drug-like molecules (Figure 1).[4] The accuracy of *in silico* docking methods directly rely on a robust torsional energy term for binding affinity calculations between the receptor and the ligand to predict the correct binding mode.[6] Consequently, this presents an interesting but difficult problem for structure-based drug design. Over the past decade, *in silico* docking methods have proven to be an indispensable tool in the field of drug discovery, allowing millions of compounds to be screened quickly and cheaply.[7] Virtual high-throughput screening (HTS) using docking has allowed an enrichment of active compounds for subsequent synthetic or biological testing efforts.[8-10]

Biaryl torsional energy profiles. Torsion energy for biaryl molecules are difficult to predict.[11] While experiments and quantum mechanical (QM) calculations have shed light on the energy profiles and preferred geometries of individual compounds on a case-by-case basis, there have been no holistic, collective approach to studying these biaryls and predicting their torsion barriers. Past work on OPLS FF,[12-14] by Jorgensen and coworkers, empirically fit the torsion parameters to match the QM profiles for thirty-three biaryl compounds most frequently found in pharmaceuticals.[5] Although this allowed the energy profiles of these aforementioned thirty-three compounds to be predicted accurately, the number of possible biaryl compounds (and diversely functionalized biaryls) are several orders of magnitudes higher and may be in the millions. At this rate, it is both impractical and impossible to parametrize each biaryl torsional barrier, individually. In addition, the torsion profile of these biaryls were found to vary drastically and cannot be transferred from one compound to another without a loss of accuracy. Consequently, separate parameters must be developed for each biaryl molecule. For example, while the torsion barrier of

biphenyl was approximately 2.1 kcal·mol^{-1}, 2-(1H-pyrrol-1-yl)pyrimidine had a much greater barrier of 11 kcal·mol^{-1}.[5] As a result, MM FF torsional parameters for most biaryl compounds are still inadequately parametrized,[4] and there still exists an unfilled need to accurately model biaryl torsion energy barriers.

In MM FFs, torsion is one of the bonded terms, along with bond and angle stretching (equation 1).[15-17] In combination with non-bonded terms (eg, electrostatics and van der Waals interactions), the linear combination of these terms give rise to the total MM energy, which determines the preferred geometry of molecules, such as *cis/trans* propensity. Each of these terms in equation 1 are modelled by different potential energy functions. In most FFs, including the General Amber force field (GAFF) 2,[18, 19] a simple elastic potential is used for bond and angle terms, while Coulomb's potential is used for electrostatic interactions, and Lennard-Jones 12-6 potential is employed for van der Waals and steric repulsion. The truncated Fourier series is most commonly employed (equation 2) for torsions. Although some FF use up to four terms within this series, only the first two (ie. $i = 1,2$) have been identified as chemically meaningful.[20, 21] The first term describes the *cis/trans* preference of a dihedral angle, with V_1 parameter denoting the energy difference between *cis* and *trans* geometry. The second term describes the height of the overall torsional barrier, which corresponds to the V_2 parameter. It should be noted that both V_1 and V_2 are divided by 2 to allow these two properties to be identified quickly. The third and fourth (ie. $i = 3,4$) are said to be correction terms.

$$E_{MM} = \underbrace{E_{bond} + E_{ang} + E_{tor}}_{bonded} + \underbrace{E_{ele} + E_{vdw}}_{non-bonded} \quad (1)$$

$$E_{tor} = \sum_{i=1}^{4} \frac{V_n}{2}(1 + \cos n\theta) \quad (2)$$

Atom-type independent FFs. In the recent past, several groups, including ours, began developing predictive MM FFs, which are atom type-independent.[20-23] Briefly, an atom type is a concept in MM, which allows parameters developed for one atom to be transferred to another by virtue of similarity of chemical environment, element identity, hybridization, and connectivity. As stated previously, the use of atom types imposes a limit on the accuracy of the FFs due to the issue of transferability. In fact, torsions suffer from one of the worst transferability within bonded interactions,[11] which prompted us to develop torsion terms without depending on atom types. Inspiration was drawn from chemical principles and knowledge accumulated by organic chemists. In the past, our first efforts were directed towards generating torsion parameters on-the-fly for molecules containing simple σ-bonds (eg. haloalkanes, alkylammonium). During that project, it was found that the strength of σ → σ* hyperconjugation played a major role in determining the height of the torsional energy barrier.[6] It was also shown that this energy barrier could be quantitatively predicted based on the electronegativity of atoms within the dihedral angle. This lead to the development of H-TEQ 1.4 (Hyperconjugation for Torsional Energy Quantification 1.4), a standalone program based on the quantification of chemical principles to derive torsion parameters on-the-fly. After this initial success, a similar strategy was used to predict the torsion barriers of molecules containing electron lone pairs in the central atom of the dihedral angle (eg. methanol, methylamine); thereby quantifying η → σ* hyperconjugation.[20] This effort lead to the development of H-TEQ 2.[20] In our last report, we began quantifying torsion interactions between benzylic and allylic bonds adjacent to conjugated systems, by looking at π → σ* and σ → π* hyperconjugation, which lead to the development of H-TEQ 3.0.[24]

UNDERSTANDING CHEMICAL ORIGINS

Qualitative predictions. Many physical and chemical properties have been employed by organic chemists to qualitatively explain observations of conformational preferences and reaction outcomes for decades. For example, hyperconjugation could explain gauche and anomeric effects,[25] while aromaticity gives a rationale for the superior stability of benzene compared to hexatriene.[26] In addition, hard and soft acids and bases (HSAB) theory has been used to account for the speeds of reaction and stability of formed products.[27] Unfortunately, despite the abundance of organic chemistry knowledge, many of these interactions remain qualitative. By understanding the underlying chemical principles, we aimed to quantify these interactions so that they could be applied to FFs.

Brief Overview of Hyperconjugation. Torsional energy profiles of molecules are primarily influenced by the effects of hyperconjugation, conjugation, sterics, and electrostatics.[4,20, 21, 28, 29] Although most existing FFs contain the two latter terms, the strengths of the two former terms have not been previously explored nor quantified.

In the context of hyperconjugation, two major factors modulate the strengths of stabilization: spatial overlap and molecular orbital energy match between the donor bonding and acceptor antibonding orbitals.[30, 31] Spatial overlap dictates that hyperconjugation is maximized when the acceptor and donor bonds are properly aligned and overlap sufficiently to allow donation of electron density from the filled to the unfilled orbital. Taking fluoroethane as an example, the donation of $\sigma_{C-H} \to \sigma^*_{C-F}$ hyperconjugation is maximum when the $\theta_{H-C-C-F}$ angle is 180°.[25] On the other hand, at 0°, this donation is nearly impossible due to having little or no overlap (Figure 2). This requirement for orbital overlap is true for all types of hyperconjugation. The orbital energy match of the donor and acceptor orbitals also plays a role in the strength of this interaction. Since the antibonding, unfilled orbitals are higher in energy than the bonding, filled orbitals, the former must be energetically accessible for donation of electron density to occur (Figure 3). Hence, a smaller energy gap increases the strength of hyperconjugation. In general, electronegative and electropositive atoms are known to lower and increase the overall energy levels of molecular orbitals, respectively.[26, 30, 32-34] Consequently, for stronger hyperconjugation interactions to occur, a more electronegative atom should comprise the acceptor orbital, while a more electropositive atom should be incorporated in the donor orbital.
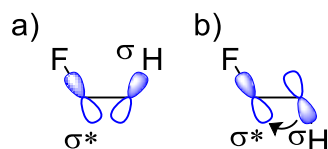

Figure 2. Two distinct conformations of fluoroethane: (a) $\theta_{H-C-C-F} = 180°$ and (b) $\theta_{H-C-C-F} = 0°$, are shown along with bonding (filled) and antibonding (unfilled) orbitals participating in $\sigma_{C-H} \to \sigma^*_{C-F}$ hyperconjugation depicted. For simplicity, not all bonds nor orbitals are shown.
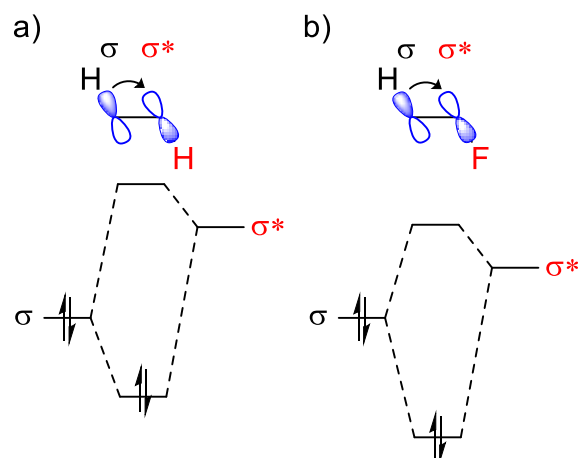
Figure 3. σ → σ* hyperconjugation in a) ethane and b) fluoroethane with selected orbitals and energies shown. In both cases, the σ-bond orbital energy level is the same, while the σ*-antibonding orbital energy for fluoroethane is lower than ethane. This causes a greater amount of hyperconjugative stabilization for fluoroethane compared to ethane.

Chemical Origins of Hyperconjugation and Conjugation in Biaryl Systems. Biaryl molecules are comprised of two aromatic cycles connected by a single bond. Consequently, various types of hyperconjugation effects exist, as exemplified by 2-(pyridin-2-yl)oxazole (Figure 4). The various orbital interactions of biaryl molecules affecting the central torsion, include σ → σ*, π → σ*, σ →π* hyperconjugation, and π → π* conjugation.[26, 35] In biaryl systems, the σ → σ* hyperconjugation (Figure 4b) and π → π* conjugation (Figure 4d) are in-phase and combined constructively. These interactions are maximum when the biaryl system is planar. In this orientation, the spatial overlap between σ and σ* orbitals and also the π-orbitals of both aromatic rings are at a maximum. Similarly, π → σ* and σ → π* hyperconjugation (Figure 4c) are also in-phase but are maximum when the two rings are orthogonal in orientation. As a result, the former, planar-favoring interactions and the latter, orthogonal-favoring interactions combine destructively, and compete with each other to modulate the overall strength of the biaryl torsional barrier.

It should also be noted that due to the geometric orientation of these orbitals, it is expected that the π and σ*-spatial overlap would be greater than the σ and π* spatial overlap (Figure 4c), leading to a greater stabilization from the latter.
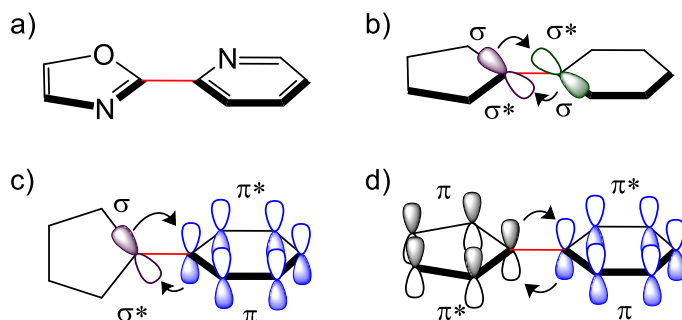
Figure 4. (a) Chemical structure of 2-(pyridin-2-yl)oxazole is shown with hyperconjugation of orbitals involving: (b) σ → σ*, (c) π → σ* and σ →π*, and (d) π → π* conjugation. The solid and

unfilled orbitals denote electron donors and acceptors, respectively. For simplicity, a maximum of one σ and σ* orbitals are shown in each subfigure. The bond of interest is colored in red.

In biaryl systems, stabilization due to π-orbitals are derived from $\pi \rightarrow \pi^*$ conjugation.[26] Unlike simple σ- and σ*-orbitals, the π- and π*-orbitals of aromatic systems are arranged in a more complex manner. In fact, Hückel molecular orbital theory predicted that the molecular orbitals of classical 5- and 6-membered aromatic cycles possess degenerate energy levels (eg. benzene and cyclopentadienyl anion).[26] Consequently, more than one possible $\pi \rightarrow \pi^*$ transition is present between the two rings, and often in either direction. For example, in a biphenyl molecule (and any other biaryls composed of two 6-membered rings), there are 18 theoretically possible conjugation interactions (Figure 5a). For 2-phenylthiophene and any other biaryls composed of a 5- and a 6-membered aromatic cycle possesses 15 theoretically possible conjugation interactions (Figure 5b). Likewise, 2,2'-bithiophene, along with other biaryls with two 5-membered aromatic rings, has 12 theoretically possible transitions (Figure 5c). In all of these three cases however, the strongest conjugation interactions come from transitions that occur between donor and acceptor orbitals with the smallest energy gaps. Consequently, $\psi_3 \rightarrow \psi_5^*$, $\psi_3 \rightarrow \psi_6^*$, $\psi_4 \rightarrow \psi_5^*$, and $\psi_4 \rightarrow \psi_6^*$ transitions are most favourable to occur (Figure 5). Despite knowing the theoretical basis from an FMO perspective, it is unclear how the substitution of one atom for a more electronegative one would affect the overall strength of conjugation. This is complicated by the number of possible transitions in conjugated systems, which will be explored in this manuscript.
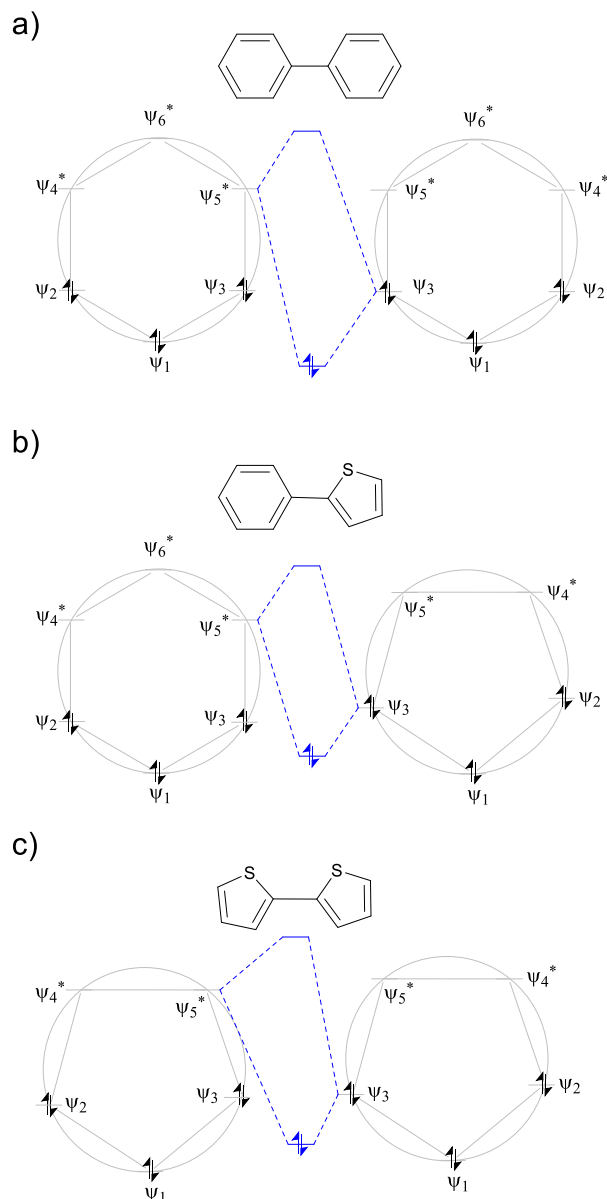
Figure 5. A qualitative molecular orbital diagram of: a) biphenyl molecule, b) 2-phenylthiophene, and c) 2,2'-bithiophene are shown, along with a major interaction, involving ψ3 → ψ5 conjugation. The right and left rings act as the donor and acceptor, respectively in this figure. For simplicity, not all possible transitions are shown.

COMPUTATIONAL METHODS

Construction of Development Set. 131 biaryl molecules of various types were included in the development set (Figure 6). The composition of this set included biaryls containing both 5- and 6-membered rings (ie., 5:5, 5:6 and 6:6 biaryls), as well as neutrally and positively charged central atoms. A suitable starting conformation was obtained for each molecule by performing a global optimization at the MP2/6-311++G** level of theory using GAMESS-US.[36] Subsequently, a constrained QM optimization was performed whereby the dihedral angle of interest was varied

7

sequentially between -180° and 180° at 10° intervals. This resulted in a total of 36 conformations for each biaryl molecule. Single point energies for each conformation was then evaluated using MP2/6-311++G** and GAFF2. AM1-BCC charges were assigned to molecules by using the default protocol implemented in Antechamber for each torsional conformation.[37] These charges were subsequently averaged for each atom throughout the conformations to preclude discontinuity in the electrostatic energy profile.

To obtain the isolated torsional energy, the van der Waals and electrostatics components of GAFF2 energies were subtracted from the total energy as computed by MP2. A Fourier regression was performed to obtain V_{1-3} parameters (equation 2) for each molecule in the development set. A variety of different chemical descriptors, including electron-richness/deficiency, central atom bond lengths, and electronegativity, were assigned to each molecule to observe their effects on the torsional barrier. The descriptor(s) which could best reproduce the torsion energy was chosen to formulate our method, which was subsequently incorporated into H-TEQ 4.0, our standalone program. To first evaluate whether these rules could reproduce the QM energy profiles of torsional rotation, H-TEQ 4.0 was first tested on the development set. The total MM energy was calculated by taking the GAFF2 energies for all terms in Equation 1, except for the torsion energy, which was computed by H-TEQ 4.0. Root-mean-square error (RMSE) was calculated for H-TEQ 4.0 and GAFF2 with the QM energy profile as a reference.
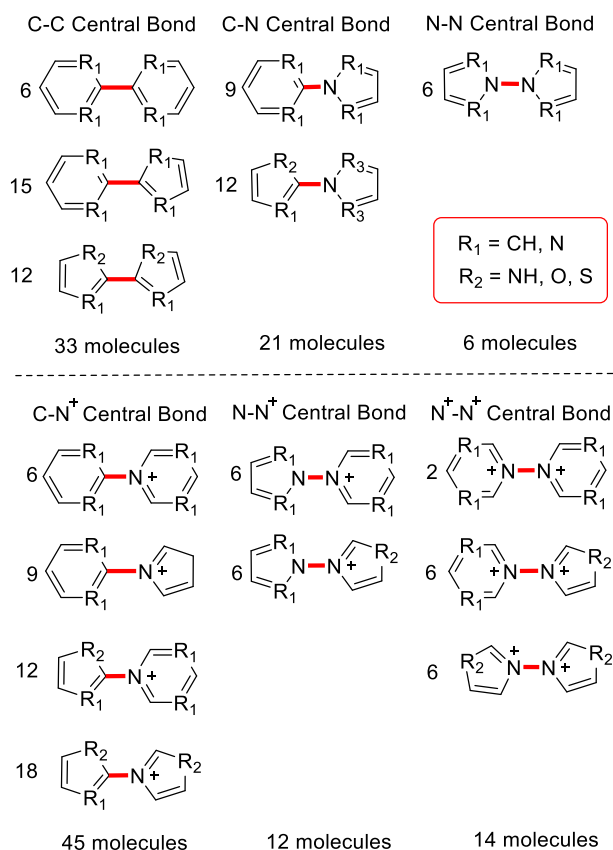
Figure 6. 131 biaryl molecules used for obtaining the torsion energy within the development set. The number on the left of each scaffold indicates the number of molecules contained in that category.

Construction of the Validation Set. 100 molecules were selected to be a part of the validation set. The composition of this set included drug-like biaryls used by Jorgenson and coworkers (20 molecules),[5] the MMFF94 set (32 molecules),[38] and a variety of molecules chosen from a previous Cytochrome P450 set (48 molecules).[39] In order to ensure the robustness of the developed method, the validation set was constructed so that it would have no overlapping molecules with the training set. From these three sources of molecules, 100 drug-like molecules were randomly selected and tested. These molecules also included those which were highly substituted, bicyclic, and tricyclic. The full validation set of molecules are shown in Figure 7.
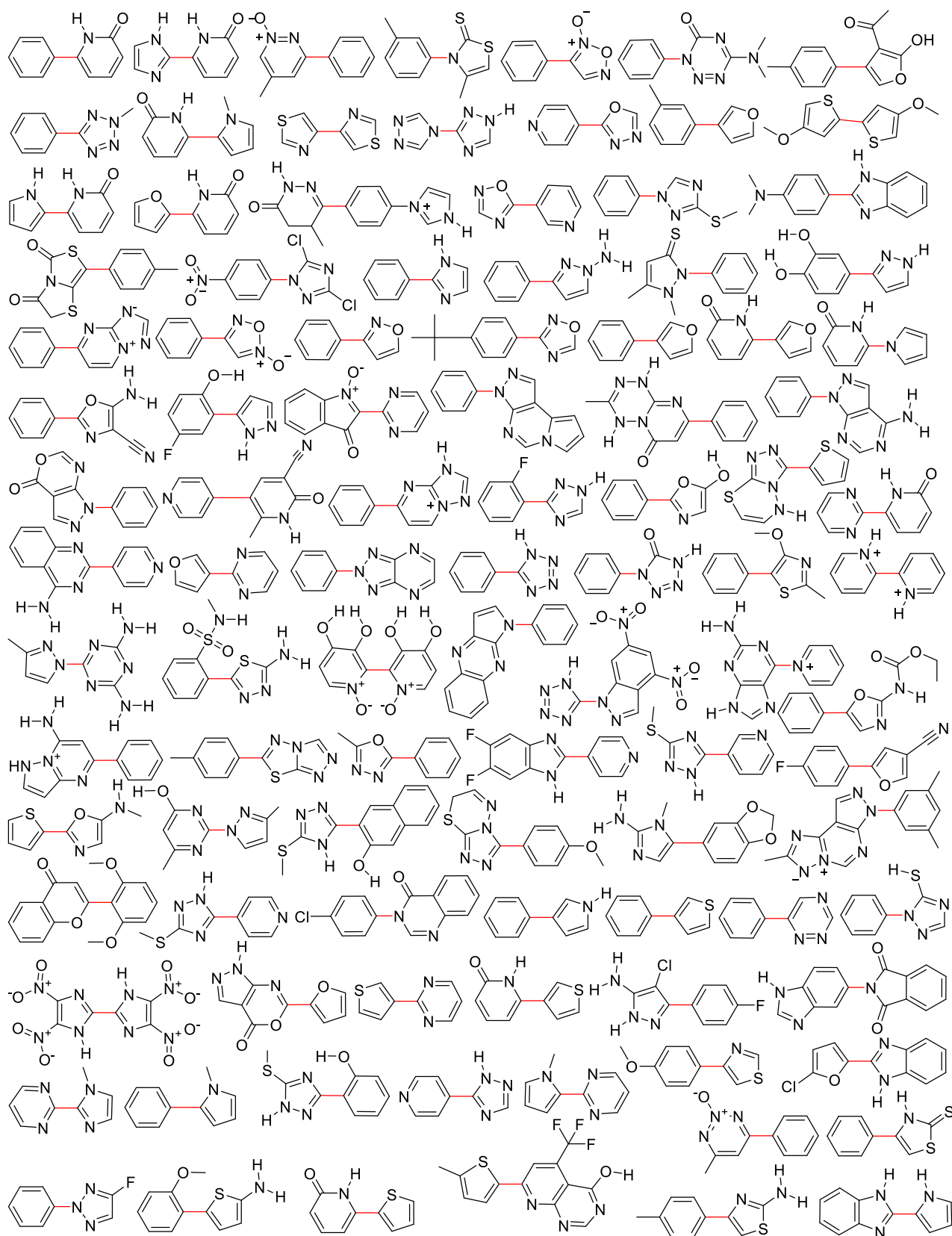
Figure 7. 100 molecules of the compiled validation set. The torsion bonds of interest are highlighted in red.

RESULTS AND DISCUSSION

Chemical Factors Modulating the Strength of Conjugation in Biaryl Molecules. It is known that the degree of π-electron delocalization in aromatic systems is influenced both by the number of available π-electrons and the electronegativity of the atoms comprising the π-system. For example, 5-member rings, such as pyrroles, are generally more electron-rich than 6-member rings, such as pyridine.[40-42] This is due to the fact that in the former, six π-electrons are distributed over five atoms as opposed to the latter where the same number of π-electrons are distributed over six atoms. As a result, it leads to a higher π-electron density in the former. In addition, it is also known that aromatic molecules with less electronegative atoms, such as pyrroles are more electron-rich than molecules with greater electronegative atoms such as furans. This is due to the propensity of more electronegative atoms to withhold π-electron density. To take both of these factors into consideration and to measure the electron-richness/deficiency of aromatic systems, a new electronegativity parameter, π-electronegativity (χ_π) was devised as shown in equation 3.

$$\chi_\pi = \frac{\sum_{i=0}^{n} \chi_i}{n(\pi)} \qquad (3)$$

In this equation, χ_i is the electronegativity of atoms comprising the conjugated system and $n(\pi)$ is the number of π electrons of this system.

Overall π-electronegativity Modulates Strength of Conjugation. Since biaryl molecules are composed of two interconnected aromatic moieties, the χ_π of both cyclic systems were computed. In all tested cases, it was found that the π-electronegativity of the entire molecule, $\chi_{\pi tot}$, correlated well with the torsional energy barrier, V_2, within each category of molecules with differing central bonds (Figure 8). More specifically, it was found that an increase in the $\chi_{\pi tot}$ correlated with an increased magnitude of V_2. Since large χ_π suggests π-electron deficiency, these results indicated that biaryls which were more π-electron deficient had a greater torsional energy barrier. Although additional calculations and experiments may be required, the increase in $\chi_{\pi tot}$ may decrease the energy gap between the ψ_2 or ψ_3 bonding and ψ_4^* or ψ_5^* antibonding molecular orbitals of biaryl systems (Figure 5). In effect, this would decrease the energy barrier required for electron donation from the bonding to the antibonding orbitals, leading to increased conjugation. This relationship between $\chi_{\pi tot}$ and V_2 was true within each category of biaryls tested, although the associated slope and y-intercept of the linear relationship and accuracy of the regression differed. For C-C, C-N, and N-N central bonds, squared correlation coefficients (R^2) of 0.48, 0.72, and 0.19 were found, respectively. For positively charged central bonds, C-N⁺, N-N⁺, and N⁺-N⁺, R^2 of 0.72, 0.65, and 0.60 were observed, respectively.

Interestingly, the slopes of the linear fit across each group varied slightly. In general, it was observed that biaryls with neutral central bonds (ie. C-C and C-N) had greater stabilization, and therefore lower V_2 than those with charged central atoms (ie. C-N⁺ and N-N⁺), despite having the same $\chi_{\pi tot}$ (Figure 8). Furthermore, both these aforementioned groups had greater conjugative stabilization when compared to those with large charge-charge repulsion within the central bond (N-N and N⁺-N⁺). Biaryls, comprising of N-N central bond (eg. 1,1'-bipyrrole), contain two adjacent lone pairs, which repel strongly when planar, as was also found in a previous study.[43] However, another explanation may be that a greater π → σ* and/or σ → π* hyperconjugation are present in these biaryls, although the exact mechanism of action is not clearly known.
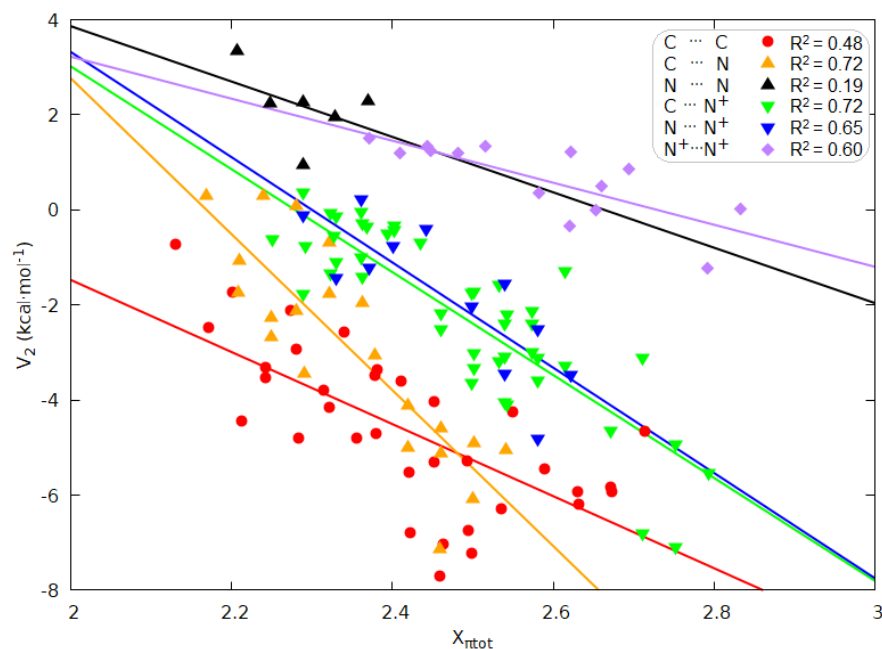
Figure 8. Height of torsional energy barrier (V_2) of 131 biaryl molecules were plotted against the total π-electron density ($\chi_{\pi tot}$), and categorized into six distinct groups, which differ by the atomic identity of the central bond. The R^2 for each linear regression is also shown in the legends.

Difference in π-electronegativity between the Two Aromatic Rings Modulates Strength of Conjugation. In addition, the difference in π-electronegativity, $\Delta\chi_\pi$, between the two aromatic groups of biaryl molecules was investigated. Intriguingly, in some cases, an increase in $\Delta\chi_\pi$ increased V_2 while in others, a decrease was observed (Figure 9). Only for N^+-N^+ central bonds, was there no observable trend, which will be mentioned and explained later in this section. For biaryls with uncharged central bonds (ie. C-C, C-N, and N-N), increases in $\Delta\chi_\pi$ increased the strength of conjugative stabilization (decrease in V_2). In fact, when V_2 was plotted linearly against $(\Delta\chi_\pi)^2$, the squared correlation coefficients of 0.28, 0.69, and 0.73 were observed for biaryls containing C-C, C-N, and N-N central bonds, respectively (Figure 9A). For biaryls containing charged central bonds, trends were also found although they seemed to be influenced by another factor—the number of atoms comprising the aromatic moiety.
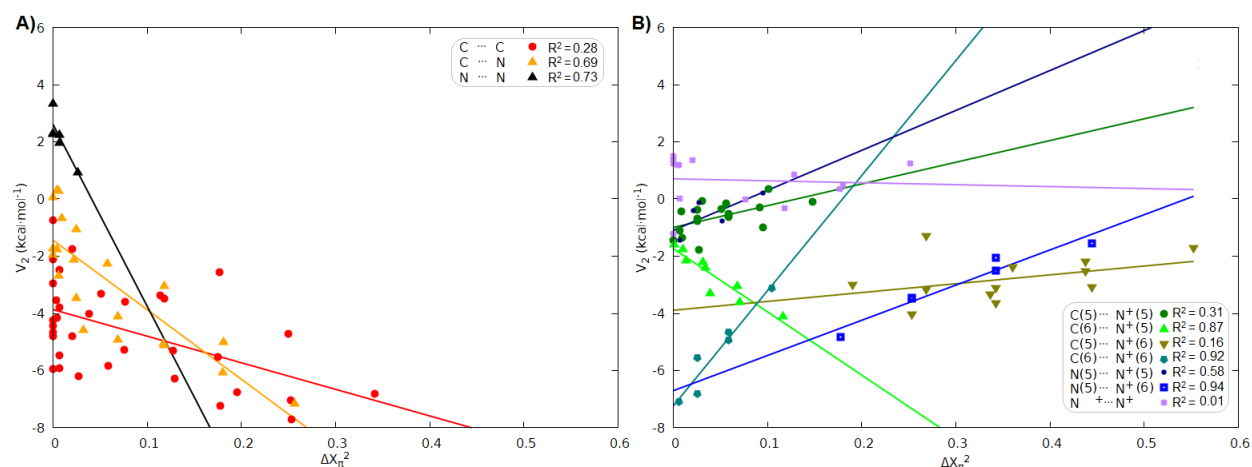
Figure 9. Height of torsional energy barrier (V2) of 131 biaryl molecules were plotted against the squared difference of π-electron density between the two connected aromatic moieties $(\Delta\chi_\pi)^2$. Biaryls were separated into: A) neutral central bonds: C-C, C-N, and N-N, and B) positively charged central bonds: C-N+, N-N+, and N+-N+. C-N+ and N-N+ was further grouped by the identity of the central atom and number of atoms on each side of the central bond. The R^2 for each linear regression is also shown in the legends.

For biaryls containing C-N$^+$ central bonds, an increase in $(\Delta\chi_\pi)^2$ decreased V_2 (Figure 9B), for those comprised of a 6-membered aromatic moiety with carbon central atom and a 5-membered aromatic moiety with a N$^+$ central atom (ie. C(5)-N$^+$(6)). All other biaryls with C-N$^+$ central bonds saw an increase in V_2 as $(\Delta\chi_\pi)^2$ increased.

Similarly, an increase in $(\Delta\chi_\pi)^2$ increased in V_2 for all N-N$^+$ central bond biaryls, regardless of the number of atoms in each ring. However, the trends for both N(5)-N$^+$(5) ($R^2 = 0.58$) differed from N(5)-N$^+$(6) ($R^2 = 0.94$) such that the slope and y-intercept differed. Although no relationship seemed to exist for N$^+$-N$^+$ central bonds, it was hypothesized that such a trend would be observed if it were sorted into N$^+$(5)-N$^+$(5), N$^+$(5)-N$^+$(6), and N$^+$(6)-N$^+$(6) central bonds. However, since these biaryls are rarely found in pharmaceuticals, the existing number of molecules was deemed sufficient for this current study, and not pursued further.

The relationship between $\Delta\chi_\pi$ and V_2 may also be caused by the lowering of the energy gap between adjacent bonding and antibonding orbitals of biaryl systems (Figure 5), leading to a more stable conjugation. In general, a greater $\Delta\chi_\pi$ in a biaryl system would suggest that the aromatic ring with the lower χ_π would act as the donor, while the aromatic ring with the higher χ_π would act as the acceptor. However, this relationship did not hold for all central bonds, possibly due to more complex molecular orbital interactions present in certain biaryls.

Substituent Effects on the Torsional Profile of Biaryls. Substituent effects of biaryl molecules on its barrier to conjugation was tested on various molecules, which were categorized into two types of substituents: inductive and conjugated. To test the effects of the inductive ligands on biaryl molecules, mono- and di-substituted biaryls were tested. More specifically, various substituents were placed at various positions on the biaryl molecules, as shown in Figure 10. In total, 15 different scaffolds and 124 substituted biaryls were tested. Although some di-substituted biaryls lead to a change in its torsion energy by as much as 1.5 kcal·mol^{-1}, most substituents had negligible effects. In fact, the average change of 8.12% was observed from the original, unsubstituted biaryl

molecule with 86% of substituted biaryls showing less than 15% change in their torsional barrier (SI Table 2.1).

To further illustrate this point, fluoryl, chloryl, methyl, and trifluoromethyl groups were placed in the meta- and para-positions of the 1,1'-biphenyl molecule, individually and in combination. These results further collaborated that inductive ligands have minor effects on the barrier to torsional rotation of biphenyl molecules. In fact, the differences in the height of substituted 1,1'-biphenyl torsion barriers to that of the unsubstituted never exceeded 0.55 kcal·mol^{-1} (Table 1). Surprisingly, the V_2 term of 2,2'-difluoro-1,1'-biphenyl and 3,3'-difluoro-1,1'-biphenyl only differed by 0.02 and 0.12 kcal·mol^{-1}, respectively, when compared to 1,1'-biphenyl, in spite of being substituted by two highly electronegative atoms. Similar negligible increases in V_2 were observed with 4-(trifluoromethyl)-1,1'-biphenyl, which had an increase of only 0.12 kcal·mol^{-1} over the unsubstituted biphenyl. This may be due to the fact that σ-orbitals of the inductive substituent and the highly delocalized π-orbitals of the biaryl system may not interact very well due to their energetic differences.

Conjugative substituents were further evaluated using a set of 1,1'-biphenyl molecules. Amine, dimethylamine, nitro, and alcohol groups were added to the para-position by themselves and in combination. The results indicate that these substituents also had little effect on the overall torsional energy barrier to rotation, despite having π-electrons which could readily interact with the π-system. In fact, the highest difference in V_2 in this class of molecules was N,N-dimethyl-4'-nitro-[1,1'-biphenyl]-4-amine, which saw an increase of 0.57 kcal·mol^{-1} over the unsubstituted biphenyl (Table 1, Figure 11). As shown at the bottom of Figure 11, this increase can be explained by the combination of a strong electron-donating group and a strong electron withdrawing group stabilizing a resonance structure leading to a strong double bond character of the central bond. However, this was somewhat surprisingly as it was expected that this combination would increase V_2 more substantially. This further confirmed that conjugative substituents had little effect on the torsional energy barrier of biphenyl.

Overall, both types of substituents: inductive and conjugated, had minor effects on the torsion profile of the biphenyl molecule which was comparable to the level of error found in MM FFs, and could potentially be caused by errors in the electrostatics or van der Waals energies. We understand that other biaryls may behave slightly differently in the context of substitution, and expect some modulation of the overall torsion energy profile. However, the presence of substituted ligands did not greatly impact the torsional energy profile of biaryls. Consequently, these effects were not further pursued nor considered to build our predictive model for estimating the torsional profiles of biaryl molecules. Only the principle aromatic rings of biaryl, directly adjacent to the σ-bond, was considered.
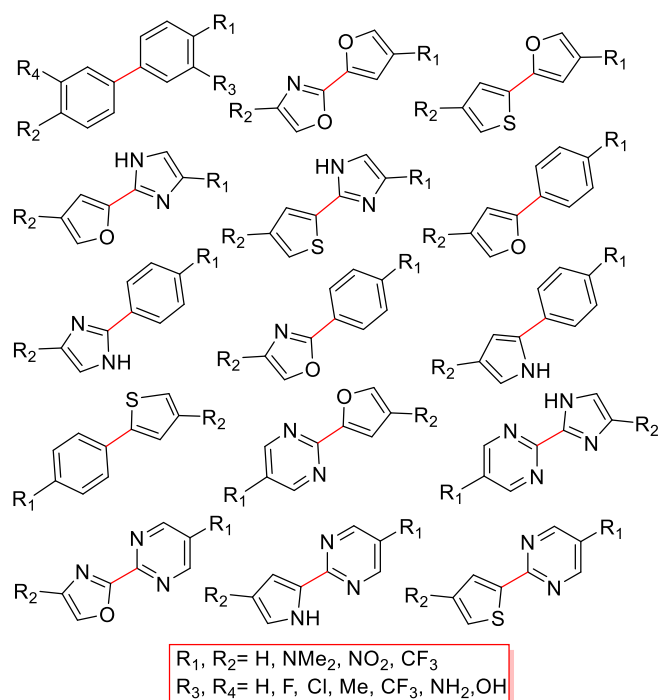
Figure 10. Biaryls used to probe the effects of substituents on its torsional energy barrier. The torsional bonds of interest are labelled in red.
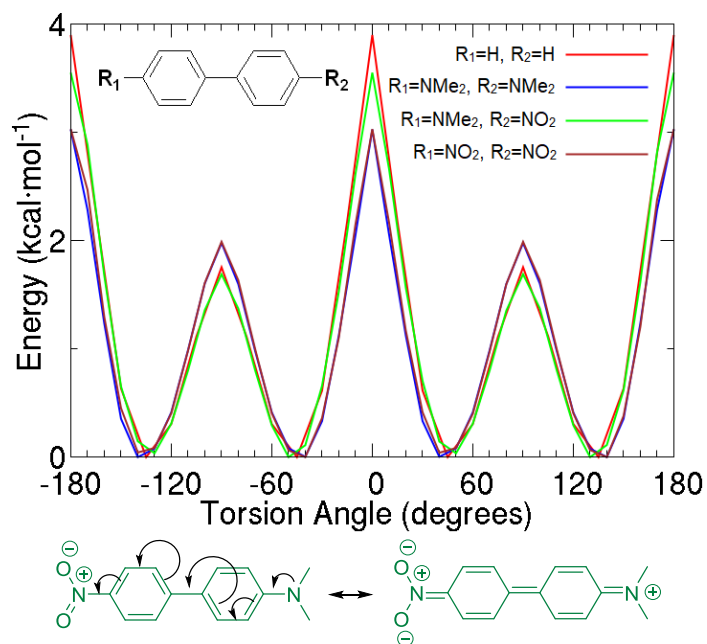
Figure 11. The torsional energy to rotation of biphenyl (red), and biphenul substituted with strong electron-withdrawing and electron donating groups as computed by MP2/6-311++G(d,p) level of theory.

Table 1. The Barrier to Torsional Rotation for Various Substituted 1,1'-Biphenyl Molecules

Biaryl Molecules	V_2 (kcal·mol^{-1})
1,1'-biphenyl	-4.25
Inductive Substituents	
3-fluoro-1,1'-biphenyl	-4.29
3-chloro-1,1'-biphenyl	-4.56
4-fluoro-1,1'-biphenyl	-4.26
4-chloro-1,1'-biphenyl	-4.64
4-methyl-1,1'-biphenyl	-4.27
4-(trifluoromethyl)-1,1'-biphenyl	-4.37
3,3'-difluoro-1,1'-biphenyl	-4.37
3,3'-dichloro-1,1'-biphenyl	-4.80
3,4'-difluoro-1,1'-biphenyl	-4.31
4,4'-difluoro-1,1'-biphenyl	-4.23
4,4'-dichloro-1,1'-biphenyl	-4.69
4,4'-dimethyl-1,1'-biphenyl	-4.23
Conjugated Substituents	
[1,1'-biphenyl]-4-amine	-4.39
[1,1'-biphenyl]-4-ol	-4.38
benzidine	-4.52
[1,1'-biphenyl]-4,4'-diol	-4.46
N4,N4,N4',N4'-tetramethyl-[1,1'-biphenyl]-4,4'-diamine	-4.28
N,N-dimethyl-4'-nitro-[1,1'-biphenyl]-4-amine	-4.82
4,4'-dinitro-1,1'-biphenyl	-4.79

***Cis/Trans* Preference of Conjugation.** The V_1 term, responsible for the cis/trans preference of a torsion term, was also investigated for the 131 biaryls in the development set. 81% of these molecules (106 out of 131 molecules) had a V_1 term between the ranges of -1 to +1 kcal·mol^{-1}, suggesting that V_1 terms do not vary greatly for the majority of molecules. For the other 19% of biaryls, no discernable trends were observed when plotted against various chemical properties.

From a chemical perspective, *cis*/*trans* preference is dictated by various chemical effects, including: electrostatics, van der Waals, and σ → σ* hyperconjugation. According to the obtained results, the strength of the V_1 energy term was found to be minor compared to electrostatics and van der Waals interactions. This is reflected in the energy profiles of the tested biaryl molecules. For example, 2,2'-bipyridine had a combined electrostatics and van der Waals energy which varied over a range of 10.65 kcal·mol^{-1} (Figure 12). In contrast, the V_1 energy term, representing the σ → σ* hyperconjugation, varied over a much smaller range of 1.81 kcal·mol^{-1}. For comparison, the V_2, term varied over a range of 5.83 kcal·mol^{-1}. These results suggest that the *cis*/*trans* preference in biaryls are predominately a result of van der Waals and electrostatic interactions, as has been shown previously.[4] In addition, although these latter terms are widely used within MM FFs, it is known that they also suffer from various sources of errors and simplifications.[11] For example, electrostatics in most common FFs cannot be polarized, while van der Waals interactions often have a steeper steric wall than in reality. As a result, the reliability of the obtained V_1 terms are questionable; they are small in value and are often below the precision of the MM method. Consequently, due to these reasons, a method to predict the V_1 variable was not included in the developed method.
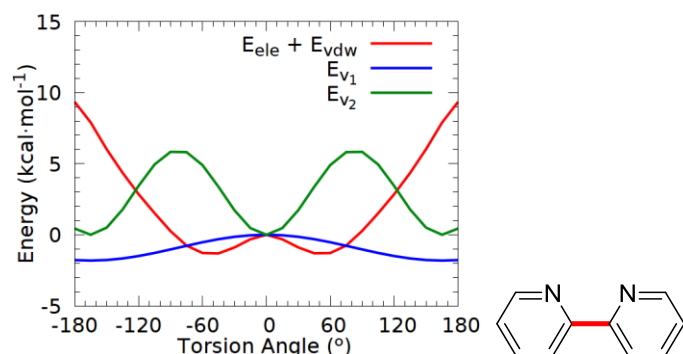
Figure 12. The energy terms extracted from a torsion scan of 2,2'-bipyridine. The red represents the sum of the GAFF2 van der Waals and electrostatics terms. The blue and green are V1 and V2, respectively, which were obtained from the torsion term as computed by $E_{QM} - (E_{vdw} + E_{ele})$.

Development of H-TEQ 4.0 Rules for Biaryl Molecules. With these observations in hand, rules were developed for all investigated biaryl molecules. In order to improve the accuracy of the method, separate rules were devised for each distinct central bond biaryls. V_2 was predicted based on a linear combination of the overall π-electronegativity and difference of π-electronegativity between the two aromatic groups, based on the equation 4, and its associated parameters (Table 2).

$$V_2 = A \cdot \chi_{\pi tot} + B \cdot (\Delta\chi_\pi)^2 + C \quad (4)$$

Table 2. Associated Parameters to Reproduce V2 for Various Categories of Biaryls

Central Atoms	A	B	C
C-C	-7.13	-7.75	13.21
C-N	-10.43	-13.37	22.38
N-N	-5.74	-63.81	15.75
$C(5)-N^+(5)$		0	24.66
$C(5)-N^+(6)$	-10.82	0	24.66
$C(6)-N^+(5)$		-20.46	25.67
$C(6)-N^+(6)$		34.36	22.60
$N(5)-N^+(5)$	-11.07	15.13	25.05
$N(5)-N^+(6)$		10.01	22.33
N^+-N^+	-5.109	0	14.22

In this equation, A and B are parameters for the weights of contribution of $\chi_{\pi tot}$ and $(\Delta\chi_\pi)^2$, respectively. These values were optimized to reproduce V_2 together; consequently, they were not the same parameters as those found in the previous section, when they were used separately. C represents the inherent strength of conjugative stabilization for a particular type of biaryl molecule. The value of these associated parameters (ie. A, B, C) were different, depending on the atomic identity of the central bond. It should also be noted that in the case of biaryls with both $C-N^+$ and $N-N^+$ central bonds, B and C parameters were also different depending on the number of atoms in each aromatic ring. Consequently, biaryls with $C-N^+$ central bonds were further separated into four categories of molecules: $C(5)-N^+(5)$, $C(6)-N^+(5)$, $C(5)-N^+(6)$, and $C(6)-N^+(6)$, and $N-N^+$ central bonds into 2 groups: $N(5)-N^+(5)$ and $N(5)-N^+(6)$.

Accuracy of H-TEQ 4.0 on Training Set- A First Validation. Prior to obtaining the accuracy of the developed method on the validation set, the accuracy of H-TEQ 4.0 was first tested on molecules found within our development set. An average RMSE of 0.71 was obtained for H-TEQ, as compared to 4.83 for GAFF2 (Figure 13) for the 131 biaryl molecules. These results indicated the robustness of the developed method. More interesting, H-TEQ 4.0 predicted a sharp peak at an RMSE of 0.5 kcal·mol^{-1}. Overall, all of the H-TEQ 4.0 predicted torsional parameters fell within an RMSE of 2.5 kcal·mol^{-1}. In contrast, GAFF2 had a broader distribution containing several maximums, with peaks found at RMSEs of 0.7, 3.7, 6.7, and 7.9. In fact, there were RMSEs of up to 19 kcal·mol^{-1}. These results suggested that while GAFF2 had been parameterized for some molecules, corresponding to the peak of RMSE at 0.7 kcal·mol^{-1}, it performed poorly for other molecules which were not parameterized. These failures may also results from a poor atom type assignment by antechamber. Within the compiled training set of 131 biaryls, GAFF2 predicted 36% of the torsional parameters below an RMSE of 2.5 kcal·mol^{-1}.
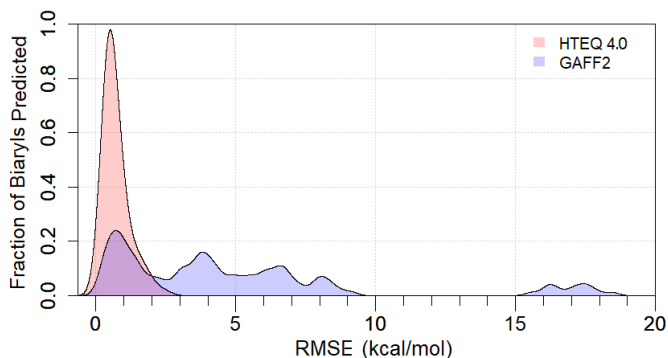
Figure 13. A smoothed histogram indicates the number of MM energy profile predictions made by H-TEQ 4.0 (pink) and GAFF2 (blue) within the training set for 131 biaryls. The RMSEs of these predictions are shown on the x-axis. The QM energy profile was used as the reference, with values closer to 0 kcal·mol^{-1} representing more accurate predictions.

Validation of H-TEQ on a Diverse Set of Drug-like Molecules. To test the accuracy of the newly developed H-TEQ 4.0, it was applied to the newly compiled set of 100 biaryl molecules (Figure 14). The results showed that the developed method had a higher overall accuracy when compared to GAFF2 (Figure 12) with mean RMSEs of 0.95 compared to 3.88, respectively, with reference to the QM torsional profile. In fact, these results indicated that the statistical mode of the RMSE distribution for H-TEQ 4.0 (0.55) was lower than GAFF2 (0.80).

In rare cases, GAFF2 performed better than H-TEQ4 for some common biaryls, such as 6-phenylpyridin-2(1H)-one (**1**), which had an RMSE of 0.17 compared to 0.97, respectively (Figure 12a). This is likely due to GAFF2 having been parametrized extensively for these molecules. When the torsion of a slightly more complex molecule, such as 6-(1H-imidazol-2-yl)pyridin-2(1H)-one (**2**), was computed however, GAFF2 drastically overpredicted the torsional barrier by approximately 25 kcal·mol^{-1}, with an RMSE of 14.15 kcal·mol^{-1} (Figure 13b). This particular biaryl had never been parametrized by GAFF2, leading to an erroneous use of a generic double bond torsion parameter, in place of a single bond parameter. In fact, 4,4'-bithiazole (**3**) and 6-(furan-2-yl)pyrazolo[3,4-d][1,3]oxazin-4(1H)-one (**5**) also had similar parameter transferability issues (Figure 13c and e), leading to an RMSE of 16.95 and 14.17 kcal·mol^{-1}, respectively when computed by GAFF2. H-TEQ 4.0, being atom type-independent, predicted a more accurate torsional barrier for these molecules, with RMSEs of 0.80, 1.10, and 0.41 for **2, 3,** and **5**, respectively.

Despite more accurately predicting the height of the torsion barrier associated with conjugation for molecule **1**, **2**, and **3**, some parts of the H-TEQ 4.0 curve also deviated slightly from that of the QM profile. This might be due to the fact that the van der Waals term in GAFF2, used as a part of the H-TEQ 4.0 energy could not capture the nuance orbital interactions at close distances.[11] When in the planar geometry, molecule **1** had close steric clash between its two pairs of hydrogen atoms, leading to an energetically high barrier at 0° and ±180° (Figure 13a). This effect was not reproduced by H-TEQ.4.0 In fact, the GAFF2 van der Waals energy, when used with H-TEQ 4.0, seemed to underpredict this barrier. It should be noted that when GAFF2 was used independently for **1**, it introduced an additional V_1 term to offset this error, thereby likely "patching" the torsional energy profile. Similarly, molecule **2** also had a pair of sterically clashing hydrogen atoms, which lead to small deviations in its H-TEQ 4.0 profile from that of the QM at dihedral angles of approximately ±180°. Its van der Waals energy was also underpredicted by 2.5 kcal·mol^{-1}. In

19

contrast, the van der Waals energy of molecule **3** was overpredicted by 2 kcal·mol^{-1} at approximately ±140° dihedral angles when its hydrogen atoms were in close proximity to each other. For biaryls without a pair of sterically clashing hydrogens, such as 6-(3-methyl-1H-pyrazol-1-yl)-1,3,5-triazine-2,4-diamine (**4**) and **5**, H-TEQ 4.0 was able to predict the QM torsional profile with still greater accuracy, with an RMSE of 0.37 and 0.41, respectively (Figure 13d and e). For reference, GAFF2 predicted an RMSE of 1.75 and 14.17, respectively. The problem of GAFF2 van der Waals seemed to extend indiscriminately to all molecules (ie. **1**, **2**, **3**, **6**, **7**, **8**, and **9**) possessing sterically clashing atoms in the ortho position with respect to the single bond connecting the two rings (Figure 13). Consequently, to further improve the accuracy of FFs, there is a need for more accurate models to describe van der Waal energies in the context of FFs.

The success of H-TEQ 4.0 on molecule **4**, **5**, and 7-(3,4-dimethylphenyl)-2-methyl-7H-pyrazolo[4,3-e][1,2,4]triazolo[1,5-c]pyrimidin-4-ium-3-ide (**9**) validated our hypothesis that both inductive and conjugated ligands had little or no effect on the torsion energy barrier of biaryls (Figure 13d, e, and i). In fact, **4** was meta-substituted by two conjugated -NH$_2$ and one inductive –CH$_3$ group. Similarly, **5** contained a bicyclic molecule and a carbonyl group. Lastly, **9** contained three methyl groups and a tricyclic aromatic moiety. In all three cases, H-TEQ 4.0 correctly computed the biaryl torsional barriers by only considering atoms within the primary aromatic rings. In fact, the RMSE of **9** was 0.29 kcal·mol^{-1} as predicted by H-TEQ 4.0, compared to 1.44 kcal·mol^{-1} for GAFF2.

The torsional energy profile of 3-(*m*-tolyl)furan (**7**) had an RMSE of 0.89 kcal·mol^{-1} as predicted by H-TEQ 4.0. The barrier to rotation at a dihedral angle of ±90° was overpredicted by approximately 0.8 kcal·mol^{-1}. This might be due to the fact that the developed method was trained on a similar molecule, 2-phenylfuran. Neglecting the presence of the methyl group on the former, the two are structural isomers. In spite of this fact, the torsional barrier was slightly different, suggesting that V_2 may be dependent on the position of the heteroatom within the aromatic ring. The reason for this is unclear, although it could be due to differences in the electron density of the biaryl system. Another reason for this might be due to the differences in the orbital energy of the σ- and σ*-orbitals adjacent to the single bond connecting the two aromatic rings. More specifically, 2-phenylfuran possesses a C-O σ*-antibonding orbital, which is a better hyperconjugation electron acceptor than the C-C σ*-orbital found in 3-phenylfuran. This phenomenon could be explored more conclusively in the future by using techniques such as energy decomposition analysis.[44, 45]

The scaffold present in the molecule, 6-isopropyl-3-phenyl-1,2,3,5-tetrazin-4(3H)-one (**8**), along with many other molecules in the validation set, was not part of the training set. In spite of this, H-TEQ 4.0 successfully computed the biaryl torsional barrier, V_2, using electronegativity of atoms comprising the ring. The RMSE for the torsion profile computed by H-TEQ 4.0 and GAFF2 was 0.67 and 2.53, respectively.

In order to also compare the accuracy of H-TEQ 4.0 to GAFF2 for those which were parametrized for the latter, a subset of biaryls below an RMSE of 5 kcal·mol^{-1} was extracted (with respect to GAFF2). Interestingly, despite only selecting for those which performed relatively well for GAFF2, the accuracy of both methods increased. In this subset, a RMSE of 0.79 was observed compared to 1.14 for H-TEQ 4.0 and GAFF2, respectively. Even for well parametrized biaryl molecules, H-TEQ 4.0 was better able to predict their torsional profiles. This further corroborates the idea that other energy terms in the FF, especially van der Waals and electrostatics, could be the source of error in certain molecules. Indeed, there has been ongoing interest in the FF community to improve the accuracy of these non-bonded interactions.[46]
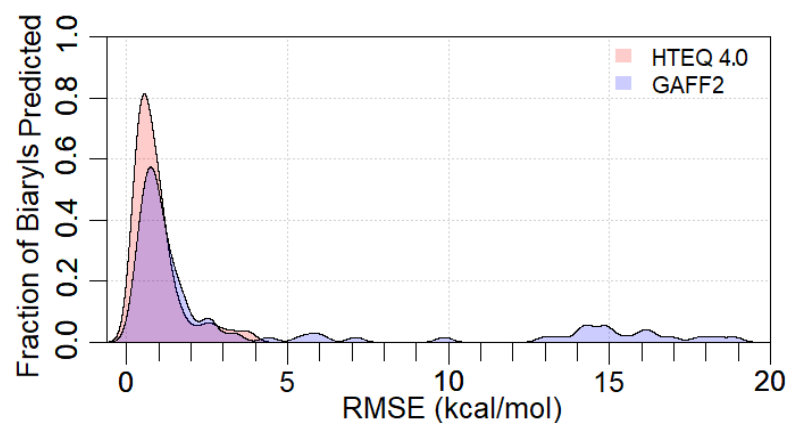
Figure 14. A smoothed histogram indicates the number of MM energy profile predictions made by H-TEQ 4.0 (pink) and GAFF2 (blue) within the validation set for 100 biaryls. The RMSEs of these predictions are shown on the x-axis. The QM energy profile was used as the reference, with values closer to 0 kcal·mol^{-1} representing more accurate predictions.
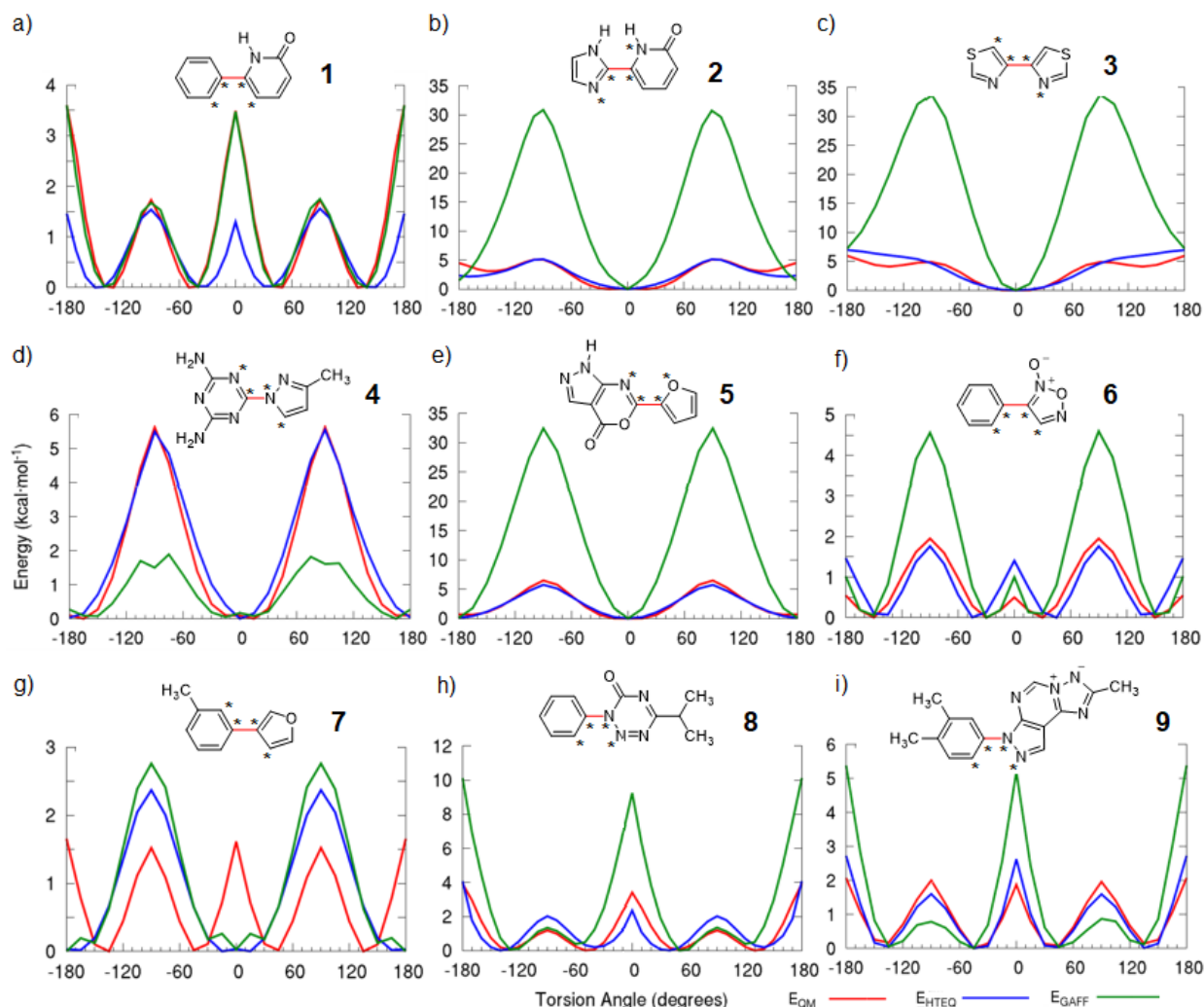
Figure 15. The torsion profiles of nine representative biaryl molecules (a-f). For each molecule, the torsion profiles as calculated by QM profile (red), HTEQ 4.0 (blue), and GAFF2 (green) are shown. The reference dihedral angle at 0° is marked by a series of four asterisks (*).

CONCLUSION

Biaryl systems are important pharmacophores, which are abundant in nature and existing drugs. Consequently, they are promising for the development of future pharmaceuticals. Unfortunately, existing FFs have difficulties in accurately reproducing their torsional profiles, due to the reliance on atom types. This limits its applicability towards virtual screening, using *in silico* docking. For example, despite GAFF2 performing well for several well-parameterized molecules, many biaryls were wrongly assigned torsion parameters, resulting in RMSEs of over 15 kcal·mol^{-1}.

Using well-established organic chemistry-based principles, such as conjugation and hyperconjugation, as a theoretical basis, we have developed an atom-type independent method to predict the torsional energies of biaryls. In fact, by understanding the underlying chemistry, we rationalized that the torsional energy of biaryl molecules were composed of electrostatic, van der Waals, and various hyperconjugation interactions. More specifically, the latter could be decomposed into: σ → σ*, π → σ*, σ →π* hyperconjugation, and π → π* conjugation. Through

a development set of 131 biaryl molecules, it was found that the strength of torsional barrier, V_2, was directly proportional to the total electron-richness of the aromatic system, $\chi_{\pi tot}$. In addition, V_2 was also found to be related to the difference in electron-richness between the two aromatic rings.

When the developed method, H-TEQ 4.0, was applied to a validation set of 100 biaryl systems, it outperformed GAFF2 in two crucial aspects. Firstly, GAFF2 suffered from transferability problems arising from atom type incompatibility issues. In fact, GAFF2 atom types were missing for several biaryl scaffolds. On the other hand, the H-TEQ 4.0, being a predictive method, more accurately computed the torsion parameters. H-TEQ 4.0 was able to solve the transferability issue associated with biaryl torsion barriers. Secondly, even for well-parametrized molecules, H-TEQ 4.0 achieved a higher accuracy than GAFF2. This proof-of-principle validation suggested that atom type-independent FFs could potentially solve the issue with transferability of atom types and improve the overall accuracy.

In the future, the chemical rationale for the strength of the torsional barrier could be studied in more detail, perhaps using high level calculations or experimental methods. In addition, research should also focus on the development of more accurate non-bonded terms in FFs, such as electrostatics and van der Waals as these were shown to be poor in the current study. Consequently, there are still many ongoing developments in FFs.[46-50] Finally, in the future, H-TEQ 4.0 should be further extended to be applicable to all dihedral angles.

Supporting Information. Additional Figures and Tables supporting and/or illustrating the conformational preferences of selected molecules are provided as supporting information, molecule sets are available as sdf files. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES
1.   Bringmann, G.; Walter, R.; Weirich, R., The Directed Synthesis of Biaryl Compounds: Modern Concepts and Strategies. *Angewandte Chemie International Edition in English* **1990**, 29, 977-991.
2.   Cayla, N. S.; Dagne, B. A.; Wu, Y.; Lu, Y.; Rodriguez, L.; Davies, D. L.; Gross, E. R.; Heifets, B. D.; Davies, M. F.; MacIver, M. B.; Bertaccini, E. J., A newly developed anesthetic based on a unique chemical core. *Proceedings of the National Academy of Sciences* **2019**, 201822076.
3.   Gill, M. Aromatic Compounds. In *The Chemistry of Natural Products*, Thomson, R. H., Ed.; Springer Netherlands: Dordrecht, 1993, pp 60-105.
4.   Sanfeliciano, S. M. G.; Schaus, J. M., Rapid assessment of conformational preferences in biaryl and aryl carbonyl fragments. *Plos One* **2018**, 13.
5.   Dahlgren, M. K.; Schyman, P.; Tirado-Rives, J.; Jorgensen, W. L., Characterization of Biaryl Torsional Energetics and its Treatment in OPLS All-Atom Force Fields. *J Chem Inf Model* **2013**, 53, 1191-1199.

6.   Roos, K.; Wu, C. J.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L. L.; Abel, R.; Friesner, R. A.; Harder, E. D., OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *JCTC* **2019**, 15, 1863-1874.

7.   Moitessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C. R., Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Brit J Pharmacol* **2008**, 153, S7-S26.

8.   Mariaule, G.; De Cesco, S.; Airaghi, F.; Kurian, J.; Schiavini, P.; Rocheleau, S.; Huskic, I.; Auclair, K.; Mittermaier, A.; Moitessier, N., 3-Oxo-hexahydro-1H-isoindole-4-carboxylic Acid as a Drug Chiral Bicyclic Scaffold: Structure-Based Design and Preparation of Conformationally Constrained Covalent and Noncovalent Prolyl Oligopeptidase Inhibitors. *J. Med. Chem.* **2016**, 59, 4221-4234.

9.   London, N.; Miller, R. M.; Krishnan, S.; Uchida, K.; Irwin, J. J.; Eidam, O.; Gibold, L.; Cimermancic, P.; Bonnet, R.; Shoichet, B. K.; Taunton, J., Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol* **2014**, 10, 1066-+.

10.   Bensinger, D.; Stubba, D.; Cremer, A.; Kohl, V.; Wassmer, T.; Stuckert, J.; Engemann, V.; Stegmaier, K.; Schmitz, K.; Schmidt, B., Virtual Screening Identifies Irreversible FMS-like Tyrosine Kinase 3 Inhibitors with Activity toward Resistance-Conferring Mutations. *J. Med. Chem.* **2019**, 62, 2428-2446.

11.   Leach, A. R. 4.05 - Ligand-Based Approaches: Core Molecular Modeling. In *Comprehensive Medicinal Chemistry II*, Taylor, J. B.; Triggle, D. J., Eds.; Elsevier: Oxford, 2007, pp 87-118.

12.   Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, 118, 11225-11236.

13.   Kaminski, G.; Jorgensen, W. L., Performance of the AMBER94, MMFF94, and OPLS-AA force fields for modeling organic liquids. *J. Phys. Chem.* **1996**, 100, 18010-18013.

14.   Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, 105, 6474-6487.

15.   Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, 30, 1545-1614.

16.   Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, 26, 1668-1688.

17.   Scott, W. R. P.; Hunenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Kruger, P.; van Gunsteren, W. F., The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **1999**, 103, 3596-3607.

18.   Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A second generation force field for the simulation of proteins, nucleic acids, and organic molecules (vol 117, pg 5179, 1995). *J. Am. Chem. Soc.* **1996**, 118, 2309-2309.

19.   Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, 25, 1157-1174.

20.   Liu, Z. M.; Barigye, S. J.; Shahamat, M.; Labute, P.; Moitessier, N., Atom Types Independent Molecular Mechanics Method for Predicting the Conformational Energy of Small Molecules. *J Chem Inf Model* **2018**, 58, 194-205.

21.   Liu, Z.; Pottel, J.; Shahamat, M.; Tomberg, A.; Labute, P.; Moitessier, N., Elucidating Hyperconjugation from Electronegativity to Predict Drug Conformational Energy in a High Throughput Manner. *J Chem Inf Model* **2016**, 56, 788-801.

22.   Mobley, D. L.; Bannan, C. C.; Rizzi, A.; Bayly, C. I.; Chodera, J. D.; Lim, V. T.; Lim, N. M.; Beauchamp, K. A.; Slochower, D. R.; Shirts, M. R.; Gilson, M. K.; Eastman, P. K., Escaping Atom Types in Force Fields Using Direct Chemical Perception. *JCTC* **2018**, 14, 6076-6092.

23.   Schmidt, J. R.; Yu, K.; McDaniel, J. G., Transferable Next-Generation Force Fields from Simple Liquids to Complex Materials. *Acc. Chem. Res.* **2015**, 48, 548-556.

24.   Champion, C. B., S; Wei, W; Liu Z; Labute P; Moitessier, N, H-TEQ3.0: Modeling the Conformational Energy of Benzylic, Allylic, and other Bonds Adjacent to Conjugated System [manuscript in preparation]. *J Chem Inf Model*.

25.   Goodman, L.; Gu, H. B.; Pophristic, V., Gauche effect in 1,2-difluoroethane. Hyperconjugation, bent bonds, steric repulsion. *J. Phys. Chem. A* **2005**, 109, 1223-1229.

26.   Fleming, I. Molecular Orbital Theory. In *Molecular Orbitals and Organic Chemical Reactions*; 2010, pp 1-67.

27.   Mayr, H.; Breugst, M.; Ofial, A. R., Farewell to the HSAB Treatment of Ambient Reactivity. *Angewandte Chemie-International Edition* **2011**, 50, 6470-6505.

28.   Pophristic, V.; Goodman, L., Hyperconjugation not steric repulsion leads to the staggered structure of ethane. *Nature* **2001**, 411, 565-568.

29.   Mo, Y. R.; Gao, J. L., Theoretical analysis of the rotational barrier of ethane. *Acc. Chem. Res.* **2007**, 40, 113-119.

30.   Houk, K. N., Frontier molecular orbital theory of cycloaddition reactions. *Acc. Chem. Res.* **1975**, 8, 361-369.

31.   Fernandez, I.; Frenking, G., Hyperconjugative stabilization in alkyl carbocations: Direct estimate of the beta-effect of group-14 elements. *J. Phys. Chem. A* **2007**, 111, 8028-8035.

32.   O'Hagan, D., Understanding organofluorine chemistry. An introduction to the C-F bond. *Chem. Soc. Rev.* **2008**, 37, 308-19.

33.   Smith, J. G.; Ho, I., Effect of electronegative substituents on the reductive dimerization of Schiff bases. Formation of vicinal dianions. *The Journal of Organic Chemistry* **1973**, 38, 2776-2779.

34.   Mo, Y. R.; Jiao, H. J.; Schleyer, P. V., Hyperconjugation effect in substituted methyl boranes: An orbital deletion procedure analysis. *J. Org. Chem.* **2004**, 69, 3493-3499.

35.   Cappel, D.; Tullmann, S.; Krapp, A.; Frenking, G., Direct estimate of the conjugative and hyperconjugative stabilization in diynes, dienes, and related compounds. *Angewandte Chemie-International Edition* **2005**, 44, 3617-3620.

36.   Gordon, M. S.; Schmidt, M. W., Advances in electronic structure theory: GAMESS a decade later. *Theory and Applications of Computational Chemistry: The First Forty Years* **2005**, 1167-1189.

37.   Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, 23, 1623-41.

38.   Halgren, T. A., Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, 17, 490-519.

39.   Campagna-Slater, V.; Pottel, J.; Therrien, E.; Cantin, L. D.; Moitessier, N., Development of a computational tool to rival experts in the prediction of sites of metabolism of xenobiotics by p450s. *J Chem Inf Model* **2012**, 52, 2471-83.

40.   Katritzky, A. R. 2.3 - Structure of Five-membered Rings with One Heteroatom**Based on Chapter 3.01 of 'Comprehensive Heterocyclic Chemistry', by C. W. Bird and G. W. H. Cheeseman, Queen Elizabeth College, University of London. In *Handbook of Heterocyclic Chemistry*, Katritzky, A. R., Ed.; Pergamon: Amsterdam, 1985, pp 53-86.

41.   Katritzky, A. R. 2.4 - Structure of Five-membered Rings with Two or More Heteroatoms**Chapter 4.01 of 'Comprehensive Heterocyclic Chemistry', by A. R. Katritzky, University of Florida, and J. M. Lagowski, The University of Texas at Austin. In *Handbook of Heterocyclic Chemistry*, Katritzky, A. R., Ed.; Pergamon: Amsterdam, 1985, pp 87-123.

42.   Katritzky, A. R. 2.2 - Structure of Six-membered Rings**Adapted from Chapter 2.01 of 'Comprehensive Heterocyclic Chemistry', by A. McKillop and A. J. Boulton, University of East Anglia. In *Handbook of Heterocyclic Chemistry*, Katritzky, A. R., Ed.; Pergamon: Amsterdam, 1985, pp 23-51.

43.   Dey, S. K.; Lightner, D. A., 1,1 '-bipyrroles: Synthesis and stereochemistry. *J. Org. Chem.* **2007**, 72, 9395-9397.

44.   Zhao, L. L.; von Hopffgarten, M.; Andrada, D. M.; Frenking, G., Energy decomposition analysis. *Wires Comput Mol Sci* **2018**, 8.

45.   Mo, Y.; Bao, P.; Gao, J., Energy decomposition analysis based on a block-localized wavefunction and multistate density functional theory. *Phys. Chem. Chem. Phys.* **2011**, 13, 6760-75.

46.   Riniker, S., Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *J Chem Inf Model* **2018**, 58, 565-578.

47.   Rusu, V. H.; Bachmann, S.; van Gunsteren, W. F., GROMOS polarisable model for acetone. *Mol. Phys.* **2016**, 114, 845-854.

48.   Qi, R.; Wang, L. P.; Wang, Q. T.; Pande, V. S.; Ren, P. Y., United polarizable multipole water model for molecular mechanics simulation. *J. Chem. Phys.* **2015**, 143.

49.   Wang, J. M.; Cieplak, P.; Li, J.; Cai, Q.; Hsieh, M. J.; Luo, R.; Duan, Y., Development of Polarizable Models for Molecular Mechanical Calculations. 4. van der Waals Parametrization. *J. Phys. Chem. B* **2012**, 116, 7088-7101.

50.   Wei, W.; Luo, J.; Waldispühl, J.; Moitessier, N., Predicting Positions of Bridging Water Molecules in Nucleic Acid–Ligand Complexes. *J Chem Inf Model* **2019**, 59, 2941-2951.

Table of Contents Graphic

$$E_{MM} = E_{bond} + E_{angle} + E_{torsion} + E_{non\text{-}bonded}$$

100 biaryls	GAFF2	H-TEQ4
$\overline{RMSE}$	3.88	0.95