

# Chemical Science

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



Journal Name

ARTICLE

## Quantitative Model for Rationalizing Solvent Effect in Noncovalent CH–aryl Interactions

Bright U. Emenike,<sup>\*a</sup> Sara N. Bey,<sup>a</sup> Brianna C. Bigelow,<sup>a</sup> and Srinivas V.S. Chakravartula<sup>b</sup>

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

The strength of CH–aryl interactions ( $\Delta G$ ) in 14 solvents was determined *via* the conformational analysis of a molecular torsion balance. The molecular balance adopted *folded* and *unfolded* conformers in which the ratio of the conformers in solution provided a quantitative measure of  $\Delta G$  as a function of solvation. While a single empirical solvent parameter based on solvent polarity failed to explain solvent effect in the molecular balance, it is shown that these  $\Delta G$  values can be correlated through a multiparameter linear solvation energy relationship (LSER) using the equation introduced by Kamlet and Taft. The resulting LSER equation [ $\Delta G = -0.24 + 0.23\alpha - 0.68\beta - 0.1\pi^* + 0.09\delta$ ]—expresses  $\Delta G$  as a function of Kamlet–Taft solvent parameters—revealed that specific solvent effects ( $\alpha$  and  $\beta$ ) are mainly responsible for “tipping” the molecular balance in favour of one conformer over the other, where  $\alpha$  represents a solvents’ hydrogen-bond acidity and  $\beta$  represents a solvents’ hydrogen-bond basicity. Furthermore, using extrapolated data ( $\alpha$  and  $\beta$ ) and the known  $\pi^*$  value for the gas phase, the LSER equation predicted  $\Delta G$  in the gas phase to be  $-0.31$  kcal/mol, which agrees with  $-0.35$  kcal/mol estimated from DFT-D calculations.

### Introduction

Noncovalent CH–aryl interactions are of fundamental importance for molecular recognition in a vast number of chemical and biological processes.<sup>1–3</sup> Although studies have revealed the physical origins of CH–aryl interactions,<sup>4, 5</sup> the effects of solvation have not been thoroughly examined, at least not from a quantitative perspective. Various theoretical calculations by density functional theory (DFT) estimate the interaction energy between a methyl group and the face of a benzene ring to range between  $-1.0$  to  $-1.5$  kcal/mol.<sup>5, 6</sup> However, the energetics of CH–aryl interactions observed in solution are comparatively much lower (less than  $-0.4$  kcal/mol in chloroform).<sup>7–9</sup> The disparity between experimental results and DFT calculations hinges on the effect from the explicit solvent interaction, which alters the net forces governing a CH–aryl interaction in solution. For example, it has been shown that a significant portion of CH–aryl interaction occurring in the gas phase originates from electrostatic and London dispersion forces.<sup>5, 6, 10, 11</sup> On the other hand, however, the net effect of London dispersion forces for interactions taking place in solution has been reported to be negligible.<sup>12, 13</sup> Although theoretical methods offer a convenient way of studying CH–aryl interactions, they seldom capture the complicated effects of solvation phenomenon, which is problematic

particularly because the explicit interactions between solute and solvent molecules are an integral part of solvation.<sup>14</sup>

The surprising lack of research interest in a quantitative model of solvent effects in CH–aryl interactions is not because solvent effects are unimportant, but, rather, due to the following reasons: (1) there is a relative lack of data on the strengths of CH–aryl interactions in various solvents, and (2) there have not been widely accepted models to quantitatively rationalize solvent effects in CH–aryl interactions. These reasons are further complicated by the energetically weak nature of CH–aryl interactions, which makes experimental measurements of CH–aryl interactions in solution a nontrivial task. Therefore, the first goal of this study is to employ a series of molecular torsion balances—capable of measuring small interaction energies—as model systems to experimentally quantify the strength of CH–aryl interactions as a function of solvation. The second objective is to explore the possibility of constructing a quantitative model—using a *linear solvation energy relationship*—to describe the energetics of CH–aryl interactions as a function of the properties of the solvating media.

### Results and discussion

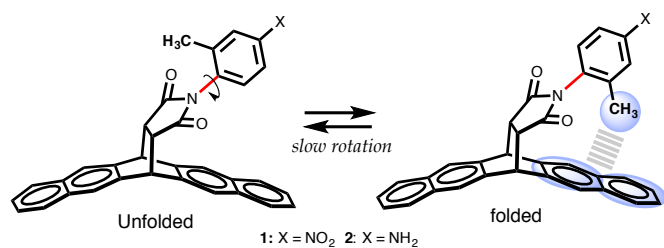
#### Model System

The model system chosen to accomplish these objectives is based on the architectural carbon framework of bicyclic *N*-arylimides, Figure 1. Previous studies by Harano *et al.*,<sup>15</sup> Shimizu

<sup>a</sup> Department of Chemistry & Physics, State University of New York, 223 Store Hill Road, Old Westbury, NY 11568, USA. E-mail: emenikeb@oldwestbury.edu

<sup>b</sup> Department of Chemistry & Biochemistry, Hunter College Graduation Center, City University of New York, 695 Park Avenue New York, NY 10065, USA

Electronic Supplementary Information (ESI) available: Correlation plots; proton and carbon NMR spectra of the molecular balances, and the xyz coordinate of the optimized geometries. [See DOI: 10.1039/x0xx00000x]



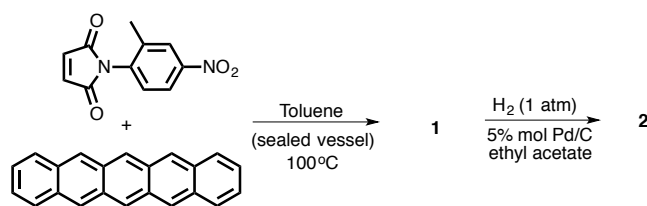
**Figure 1.** Scheme showing *folded* and *unfolded* conformational states of molecular torsional balance **1** and **2**. For structural details, see reference 16 for related single crystal structures.<sup>16</sup>

*et al.*,<sup>9, 17, 18</sup> Verma *et al.*,<sup>19</sup> Grossmann *et al.*,<sup>20</sup> and Yamada *et al.*<sup>21</sup> have demonstrated that *N*-arylimide molecular balances exhibit slow rotation about the C<sub>aryl</sub>–N<sub>imide</sub> bond at room temperature ( $\Delta G^\ddagger \sim 20$  kcal/mol).<sup>9, 22</sup> The slowly rotating C<sub>aryl</sub>–N<sub>imide</sub> bond (indicated in red in Figure 1) creates a two-state dynamic system that gives rise to two conformational states: “*folded*” and “*unfolded*”. In the *folded* state, the *ortho* methyl is positioned over the  $\pi$  electron cloud of the naphthalenyl ring in a fashion that makes it plausible for a CH–aryl interaction to take place. The population of each conformational state is determined by integrating the two distinctive peaks (using proton NMR spectroscopy) that correspond to the *ortho* methyl proton or that of the *ortho* phenyl proton in the *folded* and *unfolded* conformers. Subsequently, the free energy of the CH–aryl interaction ( $\Delta G_{\text{exp}}$  in kcal/mol) is calculated from the ratio of the population of the conformational states following the Gibbs relation, equation 1. In the absence of the CH–aryl interaction, the molecular balance is expected to have no conformational preferences, *i.e.*, the percentage of *folded* conformer ( $F_c$ ) = 50%. Consequently,  $F_c > 50\%$  is indicative of an attractive CH–aryl interaction and  $F_c < 50\%$  suggests a repulsive CH–aryl interaction.

$$\Delta G = -RT \ln K = -RT \ln [\textit{folded}] / [\textit{unfolded}] \quad (1)$$

One of the appealing features of the *N*-arylimide based molecular balances is that they can be readily synthesized,<sup>16</sup> as demonstrated in scheme 1. The convergent Diels–Alder reaction of the imide and pentacene furnished molecular balance **1** in a quantitative yield. Subsequent reductive hydrogenation of **1** using Pd/C under 1 atm of H<sub>2</sub> gas produced balance **2** in a similarly high yield.

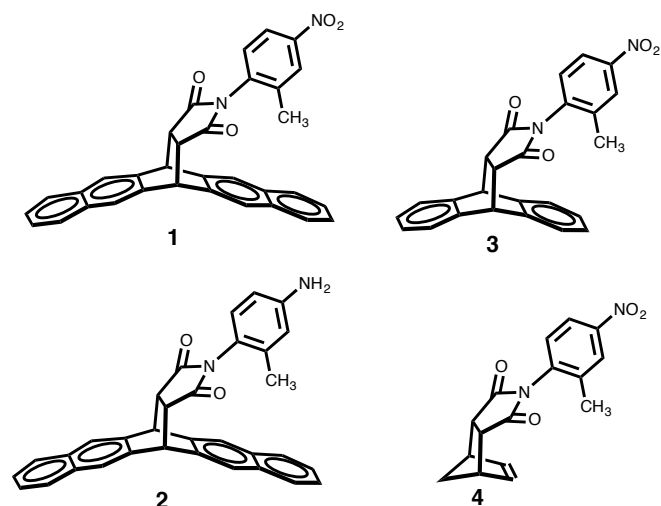
In order to test the dependence of the conformational preferences on the CH–aryl interaction, the electron density of the  $\pi$  system forming a contact with the CH donor was reduced. The balance variants in which the methyl group could form contacts with a naphthyl, phenyl and ethenyl groups were synthesized (compounds **1** to **4**, respectively).



**Scheme 1.** Synthetic route to molecular torsion balances **1** and **2**.

Because the polarity of the *ortho* methyl protons is expected to remain unperturbed by these modifications, only the electron density of the  $\pi$  face should be affected. Gratifyingly, the results from the proton NMR analysis—taken in DMSO-*d*<sub>6</sub> solution—indicated a reduction in the strength of the CH–aryl interaction as the percentage of *folded* conformer ( $F_c$ ) value dropped from 82% (for **1**) to 68% (for **3**) and to 58% (CH–ethenyl interaction for **4**). This trend can be explained by the observation that the systematic depletion of  $\pi$  electron density led to a gradual reduction in the strength of CH–aryl interaction. Concomitantly, the magnitude of the expected “shielding” effect caused by the aromatic ring current—measured as the difference in the chemical shifts ( $\delta$ ) of the methyl protons in the two conformers—is largest for balance **1** ( $\Delta\delta = 1.65$  ppm), followed by balance **3** ( $\Delta\delta = 1.08$  ppm), and smallest for balance **4** ( $\Delta\delta = 0.08$  ppm).

Following the same line of thought, the effect of CH acidity (on the conformational preferences of balance **1**) was probed by tuning the polarity of the *ortho* methyl with substituents, and, with the  $\pi$  electron density of the naphthalene ring remaining unperturbed. The nitro group placed at the rotating phenyl group (in balance **1**) is expected to polarize the CH group through a positive inductive effect, which will likely enhance the population of the *folded* conformer should the origin of the conformational preference rely on the strength of the CH–aryl interaction. Relative to balance **2**—which possess an electron-donating amino substituent—the proton NMR analysis (in DMSO-*d*<sub>6</sub> solution) indicated that the CH–aryl interaction in **1** was indeed stronger than the CH–aryl interaction in **2** because the populations of *folded* conformer ( $F_c$ ) were 82% and 77%, respectively. Although the substituent effect was only marginal, the fact that the  $F_c$  values in both balances are well above the threshold value of 50% provides indications that “strong” interaction, perhaps a CH–aryl interaction, stabilizes the *folded* state. Furthermore, DFT optimization at B3LYP/6-31G+(d) level shows that both conformers have approximately the same dipole moments (*e.g.*, 7.7 D for *folded* conformer, and 7.4 D for *unfolded* conformer), which alleviates potential concerns that solvent effects might simply be a reflection of the difference in the polarities of the conformers. Altogether, these results emphasize the importance of CH–aryl interactions on the conformational preference of the molecular balances.



**Figure 2.** Structures of molecular balances **1** - **4** used in the solution studies of CH–aryl interactions.

### CH–aryl interactions in solution

Having shown that balance **1** is a viable probe for measuring a CH–aryl interaction, we proceeded to obtain  $F_c$  values in 14 different solvents (Table 1). The results show that CH–aryl interactions are favourable in polar solvents and slightly unfavourable in nonpolar solvents. The experimental interaction energies span a range of -0.22 kcal/mol for the weakest CH–aryl interaction in cyclohexane to -0.90 kcal/mol for the strongest CH–aryl interaction in dimethylsulfoxide (DMSO). Qualitatively, these results are consistent with recent ideas that suggest solvophobic effects are responsible for interactions between nonpolar functional groups.<sup>23, 24</sup>

One of the traditional methods to attempt to rationalize solvent effects in molecular recognition studies is to simply relate the interaction energies as a function of solvent polarity.<sup>25</sup> However, this approach is not straightforward because the term “solvent polarity” is a loosely defined concept, which can be arbitrarily interpreted as the permanent dipole moment ( $\mu$ ) of the solvating molecule, the dielectric constant ( $\epsilon$ ), or the solvent polarity parameter  $E_T(30)$  of the solvating media.<sup>14, 26</sup> Nonetheless, the linear correlation of these solvent polarity parameters only led to weak correlations (Figure 3a, also see ESI). Furthermore, the possible correlation between the interaction energies and the cohesive energy density (ced) of the solvents was also explored. Cockroft *et al.* recently introduced the ced solvation model as a single universal descriptor for rationalizing solvophobic effects in nonpolar interactions.<sup>23</sup> Unfortunately, the ced model only resulted in a poor correlation ( $R^2 = 0.19$ , see ESI) when applied to balance **1**. The poor correlations between  $\Delta G_{\text{exp}}$  and the  $E_T(30)$  solvent polarity parameter and that of the ced model suggest that single and implicit solvent parameters do not produce an adequate model to quantitatively account for the observed solvent effect in the balance system. Alternatively, an approach that takes into account not only the nonspecific properties of solvation, but also specific

**Table 1.** Predicted and measured folding energies of balance **1**, and the Kamlet-Taft solvent parameters

#	Solvent <sup>b</sup>	$\alpha^e$	$\beta^e$	$\pi^{*e}$	$F_c^c$	$\Delta G_{\text{exp}}^a$	$\Delta G_{\text{pred}}^d$
1	cyclohexane	0.00	0.00	0.00	59	-0.22	-0.24
2	chloroform	0.20	0.10	0.58	61	-0.26	-0.27
3	CD <sub>2</sub> Cl <sub>2</sub>	0.13	0.10	0.82	61	-0.26	-0.31
4	CCl <sub>4</sub>	0.00	0.10	0.28	63	-0.31	-0.29
5	benzene	0.00	0.10	0.59	64	-0.34	-0.28
6	pyridine	0.00	0.64	0.87	74	-0.62	-0.67
7	DMSO	0.00	0.76	1.00	82	-0.90	-0.85
8	<i>p</i> -dioxane	0.00	0.37	0.55	71	-0.53	-0.54
9	methanol	0.98	0.66	0.60	70	-0.50	-0.52
10	acetic acid	1.12	0.45	0.64	65	-0.37	-0.35
11	acetone	0.08	0.43	0.71	73	-0.59	-0.58
12	acetonitrile	0.19	0.40	0.75	73	-0.59	-0.54
13	THF	0.00	0.55	0.58	76	-0.68	-0.67
14	CD <sub>3</sub> NO <sub>2</sub>	0.22	0.06	0.85	64	-0.34	-0.31

[a] Values in kcal/mol with associated errors of 0.03 kcal/mol.<sup>27</sup> [b] deuterated solvents were used unless stated otherwise. [c] Values are in %. [d] Values are in kcal/mol, see ESI for associated errors. [e] Values for Kamlet-Taft parameters are obtained from literature.<sup>28</sup>

aspects of solvation, appears to be a more *complete* treatment of solvation. Fortunately, the multiparameter methods—also known as *linear solvation energy relationship* (LSER)—developed by Kamlet and Taft,<sup>29, 30</sup> quantitatively partitions multiple solvent effects occurring differently or simultaneously into their respective contributors based on the solvents’ electron pair-sharing (specific) parameters and polarity (nonspecific) parameters. Equation 2 is the generalized form of the Kamlet–Taft LSER.

$$XYZ = XYZ^\circ + a\alpha + b\beta + s(\pi^* + d\delta) \quad (2)$$

In equation 2, XYZ is a solvent–dependent property of interest that usually includes rate constant, equilibrium constants or, in the context of the present study,  $\Delta G_{\text{exp}}$ .  $XYZ^\circ$  is a constant derived from the multiple linear regression of equation 2, and it equals XYZ in cyclohexane (as the reference solvent). The ability of solvent molecules to accept or donate a lone pairs of electrons towards XYZ process is represented by  $\alpha$  and  $\beta$  terms, respectively, which also denotes hydrogen–bond acidity and hydrogen–bond basicity, respectively. The effect of dipole–dipole interaction on the XYZ property is denoted by the so-called dipolarity–polarizability term  $s(\pi^* + d\delta)$ , where  $\pi^*$  is generally proportional to the molecular dipole moment of the solvent molecule with single dominant dipole moments, and  $\delta$  is a polarizability correction factor (1.0 for aromatic solvents; 0.5 for polychlorinated solvents; and 0.0 for all other solvents). The magnitudes and the signs of the coefficients ( $a$ ,  $b$ , and  $s$ ) resulting from the multiple linear regression of equation 2 provide a measure of the relative susceptibility of the XYZ physicochemical process to the indicated solvent property scales.

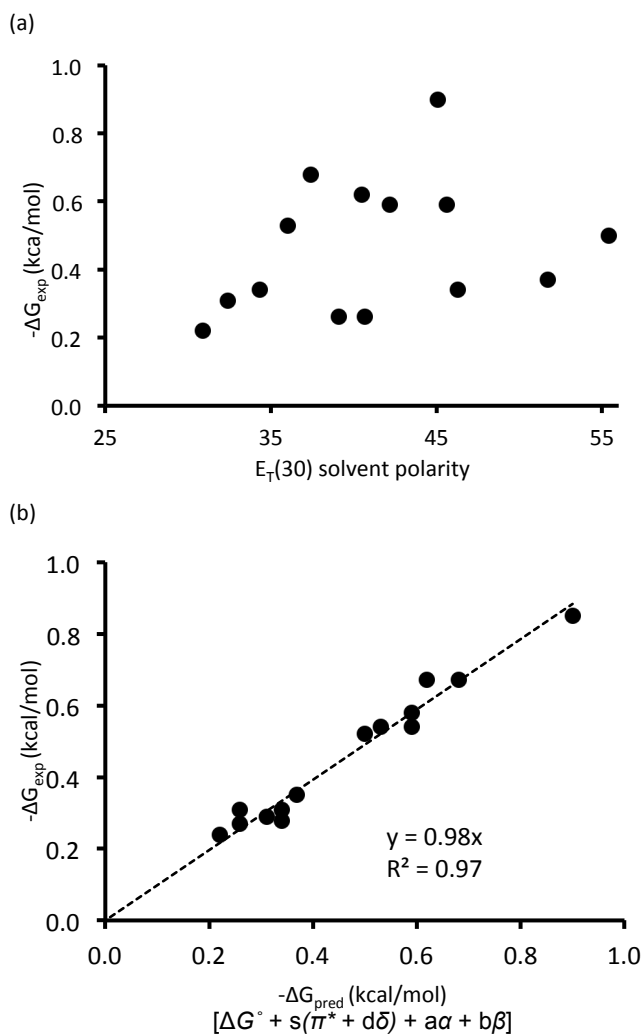
The Kamlet–Taft solvation model has been successfully used to rationalize solvent effects in a variety of processes, including but not limited to, solubility partition coefficients,<sup>31, 32</sup> reaction rates,<sup>33</sup> and conformational preferences.<sup>30, 34</sup> However, to the best of our knowledge, it has yet to be adopted as a quantitative model for rationalizing solvent effect in weak noncovalent interactions such as that of a CH–aryl interaction.

Using entries 1–10 (Table 1) as a *training* set, the multiple regression analysis of  $\Delta G_{\text{exp}}$  (dependent variable) against Kamlet–Taft parameters (independent variables), produced a linear relationship (equation 3). At 95% confidence level, and with correlation coefficient ( $R^2$ ) of 0.97 and standard deviation of 0.05 kcal/mol, the Kamlet–Taft solvation model effectively captured the solvent effect in our balance system. This correlation provided a good fit to all data on a single line without any major outliers (Figure 3b). From the point of predicting solvent effects in a CH–aryl interaction, this result represents a significant advancement over previous models. The validity of equation 3 was corroborated with four additional solvents (Table 1, Entry 11–14), in which the predicted free energy change ( $\Delta G_{\text{pred}}$ ) closely matched experimental  $\Delta G_{\text{exp}}$  values (within error). Interestingly, despite the incorporation of a nitro group in balance 1, which was expected to “polarize” the *ortho* methyl and cause the CH–aryl interaction to have a polar/ $\pi$ -like character, the observed solvent effect was nevertheless reminiscent of interactions between two nonpolar fragments.

In the context of solvophobic effect, it appears that the solvation of the “exposed” *ortho* methyl and naphthalenyl surfaces by nonpolar solvents stabilized the *unfolded* state while desolvation, facilitated by polar solvents, drove the molecule balance towards the *folded* state.

$$\Delta G_{\text{exp}} = -0.24 + 0.23\alpha - 0.68\beta - 0.1\pi^* + 0.09\delta \quad (3)$$

In spite of the weak dependence of  $\Delta G_{\text{exp}}$  on single solvent parameter (as shown in Figure 3a), the specific hydrogen bond parameters,  $\alpha$  and  $\beta$  terms, were found to be dominant contributors because the combined coefficients (a and b) were significantly larger than the values of the  $s$  coefficient. In fact,  $\Delta G_{\text{exp}}$  showed a strong linear relationship with only  $\alpha$  and  $\beta$  terms ( $R^2 = 0.92$ , see ESI Figure S2). This result is important because it demonstrates that even for interactions between nonpolar functional groups, it is apparently more accurate to treat solvation as a specific interaction between solute and solvent molecules rather than as a non-structured continuum. This result also demonstrates that solvents’ hydrogen bond parameters ( $\alpha$  and  $\beta$ ), which have electrostatic origins, are apparently more important than dispersion forces for interactions taking place in solution. The contributions of dispersion forces are estimated by the small coefficient of the  $\pi^*$  term.



**Figure 3.** (a) Correlation plot of experimental  $\Delta G_{\text{exp}}$  with solvent  $E_T(30)$  polarity scale. (b) Linear solvation energy relationship constructed with Kamlet–Taft solvent parameters.

In order to further corroborate the dependence of  $\Delta G_{\text{exp}}$  on the solvents’ hydrogen bond properties, we also tested the Hunter solvation model,<sup>35–37</sup> which is based on a set of hydrogen bond parameters ( $\alpha_s$  and  $\beta_s$ ) that are derived differently from those of the Kamlet–Taft parameters. Interestingly, the Hunter model showed a correlation of  $R^2 = 0.93$  between  $\Delta G_{\text{exp}}$  and the  $\alpha_s$  and  $\beta_s$  values (see ESI), which was almost as good as the Kamlet–Taft correlation.

Although, both Kamlet–Taft  $\alpha$  and  $\beta$  terms are dominant contributors towards the observed solvent effect, the opposite signs of the coefficients (*i.e.*, +0.23 for  $\alpha$  and -0.68 for  $\beta$ ) suggest that solvents with high  $\beta$  values stabilize the *folded* state while solvents with high  $\alpha$  values tilt the conformational preference towards the *unfolded* state. This opposing trend can be rationalized on the basis of explicit solvent–solute interactions. Solvents with high  $\beta$  values (*i.e.*, high electron-pair donors) are likely to avoid the naphthalene’s  $\pi$  cloud because of the energetically unfavourable lone pair– $\pi$  interaction, which in turn, favours the *folded* state. Or

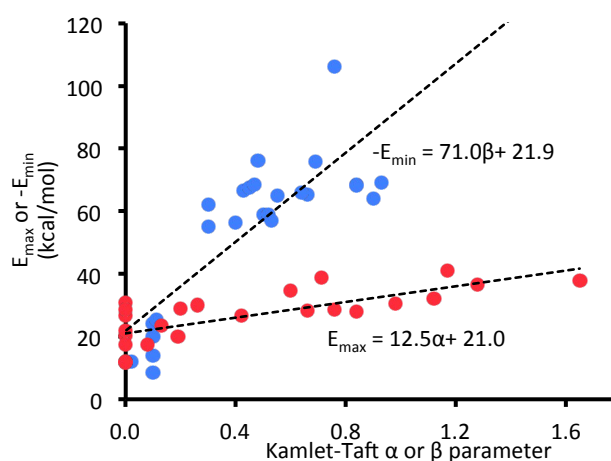
the other hand, solvents with high  $\alpha$  values (*i.e.*, high electron-pair acceptors) will engage in a favourable formal hydrogen bond with the naphthalene  $\pi$  electrons, which will bias the conformational preferences in favour of the *unfolded* state. An alternative yet complimentary rationale is based on the difference in the polarities of the *ortho* aryl CH and the *ortho* methyl group. The  $\text{NO}_2$  substituent more strongly polarizes the aromatic CH groups than the methyl group;<sup>38</sup> consequently, in solvents with high  $\beta$  constants, it is energetically preferable for the *ortho* aromatic CH group to be exposed to the solvent (*i.e.*, the folded conformation is stabilized).

### CH-aryl interactions in the gas phase

Because equation 3 describes  $\Delta G$  as a function of Kamlet-Taft parameters, it follows that if the  $\alpha$ ,  $\beta$ , and  $\pi^*$  values are known for the gas phase or conditions that mimic the gas phase, then  $\Delta G_{\text{pred}}$  should, in principle, be equal to the interaction energy expected in the gas phase. However, because Kamlet-Taft parameters were developed with cyclohexane as the reference solvent,  $\Delta G$  predicted with  $\alpha$ ,  $\beta$ , and  $\pi^*$  values all equal to zero does not represent an absence of solvent effect; rather these zero values simply represent interaction energies equivalent to those existing in a cyclohexane solution. Therefore, the questions of interest at this point are: what are the values for  $\alpha$ ,  $\beta$ , and  $\pi^*$  in the gas phase, and will such quantities accurately predict the energy of the CH-aryl interaction in the gas phase?

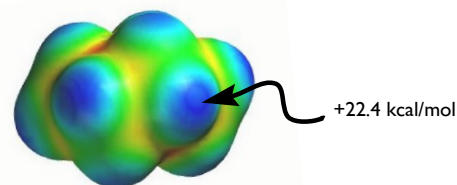
Unlike the  $\pi^*$  value of -1.1 units reported in the gas phase,<sup>39</sup> to the best of our knowledge,  $\alpha$  and  $\beta$  values in the gas phase have not yet been reported in the literature. However, few reports have shown that energy maxima ( $E_{\text{max}}$ ) and energy minima ( $E_{\text{min}}$ ) in the electrostatic potential (ESP) surface of the solvating molecule are directly proportional to solvents' hydrogen bond donor and acceptor parameters, respectively.<sup>40-42</sup> Therefore, the unknown  $\alpha$  and  $\beta$  values for a given solution can be estimated from the ESP surface of the solvating molecule. Note that the energy maxima ( $E_{\text{max}}$ ) and energy minima ( $E_{\text{min}}$ ) of most solvents are likely to be a non-zero value (*i.e.*,  $E_{\text{max}} > 0$  kcal/mol and  $E_{\text{min}} < 0$  kcal/mol), and we are assuming that an ideal solvent that closely mimics the gas phase condition or vacuum is one in which the ESP surface is "non-interactive", *i.e.*,  $E_{\text{max}}$  and  $E_{\text{min}}$  are both equal 0 kcal/mol. In other words, the intercepts on the horizontal axes in the plots of  $\alpha$  and  $\beta$  as a function of electrostatic ESP energies (shown in Figure 4) should correspond to  $\alpha$  and  $\beta$  values in the gas phase.

Even at the low level AM1 theory, the calculated  $E_{\text{max}}$  and  $E_{\text{min}}$  in the ESP surfaces correlate well with Kamlet-Taft  $\alpha$  and  $\beta$  values, respectively, which is consistent with the findings by Hunter *et al.*<sup>40</sup>



**Figure 4.** The maxima ( $E_{\text{max}}$ ) and minima ( $E_{\text{min}}$ ) in the AM1 molecular electrostatic potential surfaces of 24 solvent molecules.

Although both plots intersect the vertical axis at  $\sim 22$  kcal/mol, the  $\beta$  plot is however, much closer to the origin (offset by only 0.31 unit on the horizontal axis) than the  $\alpha$  plot, which is quite distant from zero-point origin (with an offset of 1.68 units on the horizontal axis). The reason for this can be seen by closely examining the electrostatic properties of cyclohexane—the solvent used to develop  $\alpha$  and  $\beta$  scales—as they relate to hydrogen-bond acidity and hydrogen-bond basicity. The assumption that  $\beta = 0$  is a reasonable assumption based on the facts that the surface of cyclohexane does not consist of electronegative atoms or electron pairs. However, the same argument does not hold for the  $\alpha$  scale because the hydrogen atoms on the surface of cyclohexane consisting of C-H bonds, are polarized slightly towards the carbon atom, therefore the hydrogen atoms possess partial positive charges and they are slightly acidic. As a result, the  $\alpha$  hydrogen-bond acidity—although expected to be small because the degree of C-H bond polarization in cyclohexane is also small—should not be zero (if the  $\alpha$  scale is to have a zero origin). Evidently, the ESP plotted on the van der Waals' surface of cyclohexane shows that the hydrogen atoms have  $E_{\text{max}}$  value of +22.4 kcal/mol, yet the  $\alpha$  value of cyclohexane is zero. As a result, solvents with  $E_{\text{max}}$  value less than +22.4 kcal/mol have also been assigned the value of zero even when in reality they are likely to be less acidic than cyclohexane.



**Figure 5.** Molecular electrostatic potential surface plotted on the van der Waals' surface of cyclohexane calculated using AM1 level of theory.

Using the extrapolated -1.68 and -0.31 as the approximated values for  $\alpha$  and  $\beta$  values in the gas phase, respectively, and the value of -1.1 reported elsewhere for  $\pi^*$ ,<sup>39</sup> equation 3 predicted -0.31 kcal/mol as the CH-aryl interaction energy in the gas phase using our molecular balance. Interestingly, this value compares satisfactorily with -0.3 kcal/mol reported by Datta *et al.*<sup>22</sup> for the gas-phase energy difference between *folded* and *unfolded* conformers of a structurally related molecular balance. An alternative way to verify the predicted value in the gas phase is through DFT single-point energy calculations using optimized structures of the *folded* and *unfolded* conformers. The relative energy calculation was carried out with Gaussian 09 software<sup>43</sup> at B3LYP-D3 level of theory using 6-31G+(d) as the basis set. The DFT-D results show that the *folded* conformer is more stable than the *unfolded* conformer in the amount of -0.35 kcal/mol, which is in remarkably good agreement with -0.31 kcal/mol predicted by equation 3.

Because one expects London dispersion forces to be rather significant in the gas phase, the estimated interaction energy should be in the region of -1.0 kcal/mol. However, the energy was observed to be -0.35 kcal/mol. This energy difference could be attributed to a considerable cancellation of the attractive dispersion term by an almost equally large repulsive (exchange) steric term, which is likely caused by the bent structure of the molecular balance.

The predicted  $\Delta G$  in the gas phase provides a reference point for gauging the energetic contributions of solvation in the conformational equilibrium of the molecular balance. The results (shown in Table 1) indicated that attractive CH-aryl interactions occurred in all solvents—because of the negative  $\Delta G$  values. However, relative to -0.31 (or -0.35) kcal/mol predicted in the gas phase, it is apparent that not all the solvents actually enhanced the formation of CH-aryl interactions. In fact, non-polar solvents (cyclohexane, chloroform, methylene chloride, benzene, and carbon tetrachloride) showed a destabilization effect because their  $\Delta G$  values are either equal to or less than the value in the gas phase. The small  $\alpha$  and  $\beta$  values of the non-polar solvents indicate that the CH-aryl interaction in these solvents is dominated by the  $\pi^*$  term, *i.e.*, the small diminishing effect can be attributed to competitive London dispersion forces. On the other hand, polar solvents (those with the propensity to act as hydrogen-bond donors) favoured the formation of CH-aryl interactions, which was evidenced by the resulting  $\Delta G$  values for these solvents being greater than the  $\Delta G$  value in the gas phase.

## Conclusions

In addition to providing data on the strength of CH-aryl interactions as a function of solvation, we have also demonstrated that it is possible to use the Kamlet-Taft equation to offer a molecular-level rationale for solvent effects in a *weak* CH-aryl interaction. The

magnitude and signs of the correlation coefficients in equation 3 show that specific solvent effects dominate the influence of solvation on a CH-aryl interaction. Solvents with high  $\beta$  value and low  $\alpha$  value were found to stabilize CH-aryl interactions while solvents with low  $\beta$  value and high  $\alpha$  value disfavour CH-aryl interactions. In addition, the Kamlet-Taft solvation model offers some insight into the energetics of CH-aryl interactions in the gas phase. We find that the predicted  $\Delta G$  in the gas phase lies at the boundary between  $\Delta G$  values for polar solvents and  $\Delta G$  values for non-polar solvents. Consequently, contrary to the notion that noncovalent interactions may be overestimated in the gas phase, we find that  $\Delta G$  values in the gas phase—using the molecular balance approach—are within the context of  $\Delta G$  values in solution. The quantitative partitioning of solvent effects into electron pair acceptor and electron pair donor provides a physical basis for understanding the nature of CH-aryl interactions in solution. Moving forward, we are currently investigating the universality of the Kamlet-Taft equation for modelling solvent effect in other noncovalent interactions.

## Acknowledgements

We acknowledge donors of the American Chemical Society Petroleum Research Fund for partial support of this research (ACS-PRF # 55504-UNI4). Also, we acknowledge the State University of New York Old Westbury for Faculty Development Grant. We thank Professor Robert M. Hoyte (SUNY Old Westbury) for valuable comments during the preparation of this manuscript. We also thank Professor Christopher A. Hunter for providing the Hunter  $\alpha_s$  and  $\beta$  values.

## Notes and references

1. G. B. McGaughey, M. Gagne and A. K. Rappe, *J. Biol. Chem.*, 1998, **273**, 15458-15463.
2. K. Muller-Dethlefs and P. Hobza, *Chem. Rev.*, 2000, **100**, 143-167.
3. E. A. Meyer, R. K. Castellano and F. Diederich, *Angew. Chem. Int. Ed.*, 2003, **42**, 1210-1250.
4. S. Tsuzuki and A. Fujii, *PCCP*, 2008, **10**, 2584-2594.
5. S. Tsuzuki, K. Honda, T. Uchamaru, M. Mikami and A. Fujii, *J. Phys. Chem. A*, 2006, **110**, 10163-10168.
6. A. L. Ringer, M. S. Figgs, M. O. Sinnokrot and C. D. Sherrill, *J. Phys. Chem. A*, 2006, **110**, 10822-10828.
7. B. W. Gung, B. U. Emenike, M. Lewis and K. Kirschbaum, *Chem. Eur. J.*, 2010, **16**, 12357-12362.
8. M. R. Ams, M. Fields, T. Grabnic, B. G. Janesko, M. Zeller, R. Sheridan and A. Shay, *J. Org. Chem.*, 2015, **80**, 7764-7769.
9. W. R. Carroll, C. Zhao, M. D. Smith, P. J. Pellechia and K. D. Shimizu, *Org. Lett.*, 2011, **13**, 4320-4323.
10. J. Ribas, E. Cubero, F. J. Luque and M. Orozco, *J. Org. Chem.*, 2002, **67**, 7057-7065.
11. S.-i. Morita, A. Fujii, N. Mikami and S. Tsuzuki, *J. Phys. Chem. A*, 2006, **110**, 10583-10590.

12. L. Yang, C. Adam, G. S. Nichol and S. L. Cockroft, *Nat. Chem.*, 2013, **5**, 1006-1010.
13. J. Hwang, B. E. Dial, P. Li, M. E. Kozik, M. D. Smith and K. D. Shimizu, *Chem. Sci.*, 2015, **6**, 4358-4364.
14. C. Reichardt and T. Welton, *solvents and solvent effects in organic chemistry*, Wiley-VCH, 2010.
15. Y. Yoshitake, J. Misaka, K. Setoguchi, M. Abe, T. Kawaji, M. Eto and K. Harano, *J. Chem. Soc., Perkin Trans. 2*, 2002, 1611-1619.
16. P. Li, J. Hwang, J. M. Maier, C. Zhao, D. V. Kaborda, M. D. Smith, P. J. Pellechia and K. D. Shimizu, *Cryst. Growth & Des.*, 2015, **15**, 3561-3564.
17. W. R. Carroll, P. Pellechia and K. D. Shimizu, *Org. Lett.*, 2008, **10**, 3547-3550.
18. C. Zhao, R. M. Parrish, M. D. Smith, P. J. Pellechia, C. D. Sherrill and K. D. Shimizu, *J. Am. Chem. Soc.*, 2012, **134**, 14306-14309.
19. J. Chaturvedi and S. M. Verma, *Indian Journal of Chemistry Section B-Organic Chemistry Including Medicinal Chemistry*, 1990, **29**, 9-13.
20. G. Grossmann, M. J. Potrzebowski, S. Olejniczak, N. E. Ziolkowska, G. D. Bujacz, W. Ciesielski, W. Prezdo, V. Nazarov and V. Golovko, *New J. Chem.*, 2003, **27**, 1095-1101.
21. K. Kishikawa, K. Yoshizaki, S. Kohmoto, M. Yamamoto, K. Yamaguchi and K. Yamada, *J. Chem. Soc., Perkin Trans. 1*, 1997, 1233-1239.
22. A. Nijamudheen, D. Jose, A. Shine and A. Datta, *J. Phys. Chem. Lett.*, 2012, **3**, 1493-1496.
23. C. Adam, L. Yang and S. L. Cockroft, *Angew. Chem. Int. Ed.*, 2015, **54**, 1164.
24. L. Yang, C. Adam and S. L. Cockroft, *J. Am. Chem. Soc.*, 2015, **137**, 10084-10087.
25. D. B. Smithrud and F. Diederich, *J. Am. Chem. Soc.*, 1990, **112**, 339-343.
26. A. R. Katritzky, D. C. Fara, H. F. Yang, K. Tamm, T. Tamm and M. Karelson, *Chem. Rev.*, 2004, **104**, 175-198.
27. M. Bauer, A. Bertario, G. Boccardi, X. Fontaine, R. Rao and D. Verrier, *J. Pharm. Biomed. Anal.*, 1998, **17**, 419-425.
28. Y. Marcus, *Chem. Soc. Rev.*, 1993, **22**, 409-416.
29. M. J. Kamlet, J. L. Abboud and R. W. Taft, *J. Am. Chem. Soc.*, 1977, **99**, 6027-6038.
30. M. H. Abraham, P. L. Grellier, J. L. M. Abboud, R. M. Doherty and R. W. Taft, *Can J Chem*, 1988, **66**, 2673-2686.
31. M. H. Abraham, M. J. Kamlet, R. W. Taft, R. M. Doherty and P. K. Weathersby, *J. Med. Chem.*, 1985, **28**, 865-870.
32. M. J. Kamlet, R. M. Doherty, P. W. Carr, D. Mackay, M. H. Abraham and R. W. Taft, *Environ. Sci. Technol.*, 1988, **22**, 503-509.
33. I. Correia and T. Welton, *Dalton Trans.*, 2009, 4115-4121.
34. A. T. Liu, B. U. Emenike, W. R. Carroll and J. D. Roberts, *Org. Lett.*, 2013, **15**, 760-763.
35. C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310-5324.
36. S. L. Cockroft and C. A. Hunter, *Chem. Commun.*, 2009, 3961-3963.
37. J. L. Cook, C. A. Hunter, C. M. R. Low, A. Perez-Velasco and J. G. Vinter, *Angew. Chem. Int. Ed.*, 2007, **46**, 3706-3709.
38. I. K. Mati, C. Adam and S. L. Cockroft, *Chem. Sci.*, 2013, **4**, 3965.
39. M. Essfar, G. Guiheneuf and J. L. M. Abboud, *J. Am. Chem. Soc.*, 1982, **104**, 6786-6787.
40. C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310.
41. O. Lamarche and J. A. Platts, *PCCP*, 2003, **5**, 677-684.
42. J. S. Murray, S. Ranganathan and P. Politzer, *J. Org. Chem.*, 1991, **56**, 3734-3737.
43. M. J. Frisch, G. W. Trucks and et. al., *Gaussian, Inc.*, Wallingford CT,, Editon edn., 2009.