Busch, Michael; Ahlberg, Elisabet; Laasonen, Kari

Universal Trends between Acid Dissociation Constants in Protic and Aprotic Solvents

Published in:
Chemistry - A European Journal

DOI:
10.1002/chem.202201667

Published: 21/10/2022

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY-NC

Please cite the original version:
Take Advantage and Publish Open Access

By publishing your paper open access, you’ll be making it immediately freely available to anyone everywhere in the world.

That’s maximum access and visibility worldwide with the same rigor of peer review you would expect from any high-quality journal.

Submit your paper today.

www.chemistry-europe.org
Universal Trends between Acid Dissociation Constants in Protic and Aprotic Solvents

Michael Busch,[a, b] Elisabet Ahlberg,[c] and Kari Laasonen[b]

Abstract: pKa values in non-aqueous solvents are of critical importance in many areas of chemistry. Our knowledge is, despite their relevance, still limited to the most fundamental properties and few pKa values in the most common solvents. Taking advantage of a recently introduced computationally efficient procedure we computed the pKa values of 182 compounds in 21 solvents. This data set is used to establish for the first time universal trends between all solvents. Our computations indicate, that the total charge of the molecule and the charge of the acidic group combined with the Kamlet-Taft solvatochromic parameters are sufficient to predict pKa values with at least semi-quantitative accuracy. We find, that neutral acids such as alcohols are strongly affected by the solvent properties. This is contrasted by cationic acids like ammonium ions whose pKa is often almost completely independent from the choice of solvent.

1. Introduction

Acids have been known and used within the European cultural area already for more than 2000 years. Initially, the (a)chemists of these times were restricted to only rather weak organic acids like vinegar. Reported uses comprise for example food preparation and preservation[1] but also extremely questionable applications like calming the stormy seas.[2] From these humble beginnings it took more than a millennium until the first mineral acids were synthesized by Arab and European alchemists through heating niter (KNO3), vitriol ((Fe,Cu) SO4) and/or alum (KAl(SO4)2·12H2O) to afford for example impure sulfuric acid or nitric acid.[3,4] Naturally, these new acids immediately gained significant attention but little was known regarding their true nature. However, over the following centuries the fundamental understanding of chemistry developed to a level which allowed Davy and von Liebig to suggest that hydrogen dissolved protons were later shown to display a rather complex solvation shell.[7]

Building on these insights we nowadays posses an excellent understanding of acid-base chemistry in water and determining pKa values in water experimentally[8–10] or through computations[11–18] is a routine task. For non-aqueous solvents on the other hand no comparable level of understanding has been achieved so far. Indeed, in solvents other than water our knowledge is restricted to the most fundamental properties. It is by now agreed upon, that acid base reactions in non-aqueous solvents do, in contrast to water, not necessarily proceed through proton-solvent complexes.[11,19,20] Instead, they are exchanged through a direct interaction of the acidic and basic group.[19,20] Naturally this significantly alters the kinetic profile of the acid-base reaction. The reason for this change in mechanism is the less negative energy of solvation of protons in non-aqueous solvents[11,19,31–24] which renders their presence unfavorable.

Additionally, also a less comprehensive set of pKa values in only a rather limited number of solvents has been determined so far experimentally.[9,25–31] Indeed, most of the reported data refer to the most common solvents such as dimethyl sulfoxide (DMSO) or tetrahydrofurane (THF). However, even these experimental data are often of questionable accuracy and can vary significantly between different measurements.[21] Taking for example the pKa of picric acid in 1,2-dichloroethane, pKa values of 13.7[20] and 45[19] have been suggested. Similar but less extreme variations are also observed in other solvents.[25] Furthermore, measured and computed pKa values in dimethyl sulfoxide suggest, that acids should possess comparable or lower pKa values than in water.[22,26,27] The origin of this unexpected behavior can be traced back to the predicted higher proton solvation energy of the proton in DMSO.[23,22,25,24,35–39] The proton solvation energy corresponds to the energy required to transfer a proton from vacuum into the solvent of interest. The validity
of these counter-intuitive results has recently been questioned.[14-46] These deviations are among others the result of experimental complications such as the almost unavoidable presence of impurities like water[41,42] or the possible formation of ion pairs which is corrected for in some but not all measurements.[26] However, these contributions typically cancel within a series of measurements and therefore affect the obtained pKa values to a lesser extend.[11,26] More critical are errors associated with the need to convert the measured relative pKa values to an absolute scale. This conversion relies on the pKa of a reference compound which has already been determined experimentally or computationally in an earlier study.[26] Through statistical analysis comparing computed and measured pKa values, this conversion was identified as the main source of uncertainties.[11] The situation is complicated further by the fact, that also computational modeling typically requires a reference value. In most cases an experimental pKa value, which has been determined in the solvent of interest, is used for this purpose.[44,45,46] Owing to this direct relationship with experimentally obtained pKa values also these computational results are subject to the problems associated with determining a suitable reference to offset the measured pKa values. In the context of the “electrostatic transform method” this problem was avoided by using the accurate proton solvation energy in the solvent of interest to convert the pKa values between different solvents.[47,48] However, proton solvation energies are again only known in the most common solvents and often also subject to significant uncertainties between different studies.[11,21,22,43-39] Experimentally, proton solvation energies are obtained for example through electrochemical measurements in a Kenrick cell[30,49-51] and tend to vary strongly between different experiments.[11,38,49,55] Computations of the proton solvation energy on the other hand typically rely on the experimental pKa of selected molecules in the solvent of interest[21,22,37] and thus, again suffer from the aforementioned problem of requiring a reference pKa in the experiments. Similarly, also meta-dynamics simulations, which in principle do not require an experimental pKa,[52,53] are only partly suitable for determining a reference pKa. The reason for this is that these simulations require the accurate description of the dissolved proton in the solvent of interest. These simulations are already complex and expensive in water and, owing to the rather large size and complexity of solvent molecules, unfeasible in most non-aqueous solvents and ionic liquids. Building on these rather limited data for pKa values obtained from experiments and computations some trends between different solvents have been suggested,[26] but a truly global perspective is still missing.

In what follows we will develop for the first time a general semi-quantitative understanding of the behavior of different organic acids in a wide range of solvents. In contrast to earlier work we rely on pKa values which have been computed using a recently developed procedure for pKa prediction which only requires an experimental pKa value in water combined with density functional theory (DFT) computations and a suitable solvation model as input.[11,46] Thus, any experimental or computational shortcomings of the above described methods can be avoided. Our results indicate, that acids can be divided in different groups based on the charge of the acidic functional group. The pKa of carboxylic acids like ammonium or iminium ions remains almost completely unaffected by the choice of solvent whereas the pKa of neutral acids like alcohols or carboxylic acids strongly increases when moving from highly polarizable protic solvents like water to apolar aprotic media like cyclohexane. Zwitterions finally are found to display an intermediate behavior.

Methods

Computational Details

pKa values reported in this contribution were either taken from our previous work[11] or computed using Gaussian 16 Rev C.01.[54] Independent of the origin of the data, the M06-2X functional[54] was used in combination with a triple-ζ 6–311++G* basis set with diffuse and polarization functions on all atoms and the SMD solvation model as implemented into Gaussian 16.[56] The ground state structures were considered to be converged if the lowest vibrational mode was larger than −100 cm⁻¹. Following earlier work, pKa values were computed using the method described below and in references 11 and 40. The dissociation of formic acid (pKa in water: 3.77[53]) was used as reference reaction. Note that in principle also any other reference point could be used. However, owing to a method inherent error cancellation[11] this should not affect the relative stability between protons in different solvents. Kamlet-Taft parameters for the fits were taken from experiments.[57,58] Solvation energies were computed by explicitly optimizing the structures both in vacuum and the implicit solvent. The test set of molecules was chosen such, that it spans a large range of pKa values for a diverse set of molecules and functional groups of compounds of chemical and pharmaceutical interest. Similarly, also the solvents were selected such that they comprise a set of diverse protic and aprotic polar and apolar media.

Computations of Acid Dissociation Constants

Acid dissociation constants were computed using our recently introduced procedure to compute proton solvation energies in any solvent using only a reference pKa in water.[46] The method relies on a thermodynamic cycle which expresses the acid dissociation as a series of a proton-coupled electron transfer (PCET) oxidation step (reaction 1)

$$\text{A} - \text{H} \rightarrow \text{A} \cdot + \text{H}^+ + \text{e}^-$$

(1)

and an electron transfer (ET) reduction reaction (reactio 2) in the solvent of interest (e.g. toluene).

$$\text{A} \cdot + \text{e}^- \rightarrow \text{A}^-$$

(2)

Computing the electrochemical potential of the ET step requires knowledge of the (effective) absolute potential (E_{abs,ET} (SHE)) which connects the absolute reaction energy to the standard hydrogen electrode (SHE). For the sake of improved error cancellation an effective absolute potential is computed for a given computational setup using equation 3[59]:
Here, \( G(\text{A}^-) \) and \( G(\text{A}^-) \) correspond to the total Gibbs free energy of the reference acid in water and the non-aqueous solvent need to be computed. This can easily be achieved using a standard implicit solvation model. The obtained relative stabilities using this procedure are robust with respect to different implicit solvation models and the choice of ab-initio or DFT method.[11,40]

By combining the relative stability of the proton with the effective proton solvation energy using the above described computational setup, the pKa can be computed. Owing to the shortcomings of implicit solvation models, it is nevertheless still necessary to scale the obtained raw pKa values. For protic solvents, Equation 9 was found most appropriate when using the the above computational setup and formic acid as reference.[13]

\[
pKa(Exp) = 0.49pKa(Calc) + 3.2
\]  

(9)

pKa values in aprotic media were corrected using Equation 10.[11]

\[
pKa(Exp) = 0.75pKa(Calc) + 3.2
\]  

(10)

Note that the scaling relations can vary between different functional groups and solvents.[11] Considering, however, that the experimental data required to determine the scaling relations are often of questionable accuracy in non-aqueous solvents[11] or even completely missing we opted for this general approach. Tests for solvents for which reasonably reliable experimental data were available showed that this ansatz is sufficiently accurate to at least reproduce semi-quantitative trends between solvents.[11] An increased accuracy would require the use of explicit solvation models which do not suffer from the same shortcomings of the SMD model used in the present study. Using explicit solvation would, however, require extensive molecular dynamics simulations which are not affordable, owing to the size of non-aqueous solvents and the number of solvents in the present study. Owing to the poor reliability of experimental data no exact error estimated for the presented pKa values can be provided. Nevertheless, experimental and computed pKa values typically differ by 1 to 3 pKa units but larger deviations are possible. This is, however, not necessarily a sign of the poor quality of our data.[11]

2. Results and Discussion

2.1. Acid Dissociation Constants in Non-aqueous Solvents

General trends between the acid dissociation constants were extracted from a test set comprising 182 molecules (240 pKa values) in 21 solvents. This extensive data set contains carboxylic acids, alcohols, primary/secondary/tertiary amines, imines and CH acids. Additionally, also few thiols and mineral acids were included. The pKa values of these compounds were computed in a wide range of solvents which covers the full spectrum from highly polar (e.g. DMSO or formamide) to extremely nonpolar (e.g. cyclohexane or dimethylamine) protic and aprotic solvents.

The behavior of one representative compound for some of the most common functional groups in different solvents is summarized in Figure 1. For convenience the solvent properties,
plotted at the x-axis, are described by simply using the sum of the Kamlet-Taft solvatochromic parameters $\alpha$, $\beta$, and $\pi^*$. $\alpha$ and $\beta$ correspond to the solvent's ability for hydrogen bond donation and acceptance while $\pi^*$ corresponds to the polarizability. Using this descriptor, aprotic apolar solvents such as cyclohexane or toluene are placed at the left in Figure 1 while protic polar solvents such as water or formamide are placed at the far right in Figure 1. Solvents displaying intermediate properties such as protic apolar solvents like acetic acid or polar aprotic media like acetonitrile can be found between these two extremes.

Our data clearly indicate, that the trends in the pKa vary strongly between different classes of acids when moving from protic/polar solvents such as water to apolar/aprotic solvents like cyclohexane. Carboxylic acids are represented by ibuprofen, which is typically used to treat weak to moderate pain. For ibuprofen we predict a pKa of 5.8 in water which is in reasonable agreement with the experimentally obtained value of 4.5. Upon moving towards less polar and aprotic solvents, the acid strength decreases significantly (Figure 1a). In isopropanol for example, which is still a rather polar/protic solvent, the pKa rises to 10.7. This is in good agreement with the experimentally observed increase to 11.8. Even higher pKa values of 21.1 and 22.1 are observed for acetonitrile and DMSO. It is noteworthy, that the pKa values in these two solvents are almost identical. This is contrasted by earlier studies which suggested, that pKa values in DMSO should be significantly lower than those in acetonitrile.

In line with these trends, also protons in DMSO were predicted to be equally or more stable than in water while our computations again indicated, that protons in both media should be approximately equally stable. These results were confirmed by us through...
CCSD(T) computations and by using different DFT functionals.\[^{[11,40]}\] Furthermore, they were also found to be stable with respect to the choice of implicit solvation model and its parametrization.\[^{[11,40]}\] We are therefore reasonably confident, that our method is able to capture the trends between DMSO and acetonitrile correctly. Using statistical analysis the experimental pKa values in non-aqueous solvents on the other hand were found to be of questionable quality.\[^{[11]}\] In cyclohexane, which has the least ability for hydrogen bonding or dipole interactions\[^{[57]}\] of all solvents considered in this study, the pKa increases even by 42.7 units to 48.5.

The observed behavior for carboxylic acids is not unexpected when considering, that it dissociates into two charged species, a carboxylate anion and a solvated proton. The computed relative solvation energies in different solvents for all involved species are summarized in Figure 2. Both the carboxylate anion and the solvated proton are well stabilized in water owing to its high polarizability and ability for hydrogen bonding (Figure 2a). Moving towards less polarizable solvents, which eventually also possess a reduced ability for hydrogen bonding, their relative solvation energies increase which corresponds to a decrease in their overall stability. The acid on the other hand is only to a minor degree affected by the change of solvent, i.e. the solvation energy remains almost constant independent of whether Ibuprofen is dissolved in water or cyclohexane. As a result of this, the relative pKa in less polar solvents, which is completely determined by the difference in the solvation energies (\(\Delta\Delta G_{\text{solv}}\)), increases.

Similar to carboxylic acids also the dissociation of alcohols results in two charged species, an alcoholate anion and a proton. It is therefore not surprising that an almost identical dependence of the pKa from the choice of solvent is observed (Figure 1b). This is exemplified for the second dissociation step of morphine. Morphine is used to treat strong pain and needs to be transported across the lipophilic blood-brain barrier, which renders its acid-base properties both in aqueous and non-aqueous media of relevance.\[^{[65,66]}\] According to our calculations and in agreement with experiment\[^{[66]}\] Morphine possesses a pKa of approximately 9.8 in water. When dissolved in cyclohexane, the acid dissociation constant rises, identical to the carboxylic acids, by approximately 42 pKa units to 52. A qualitatively similar behavior is also observed for C–H acids.

The behavior of the alcohol and the carboxylic acid is contrasted by protonated ammonium and iminium ions. Considering for example aniline (primary amine), Propranolol (secondary amine), Morphine (tertiary amine) and Diazepam (imine) (Figure 1c to f) we find that the relative changes to the pKa are affected by the change to extremely apolar, aprotic solvents to a much lesser degree. For primary amines the pKa remains, with a shift of only 4 pKa units upon moving from water to cyclohexane, almost constant. For secondary and tertiary amines as well as imines the shift rises to roughly 10 pKa units. Similar shifts are also observed for all other considered non-aqueous media. This strongly suggests, that the pKa is almost independent from the choice of solvent. Thus, similar minor changes in the acidity can also be expected under...
physiological conditions such as in the lipophilic blood-brain barrier which needs to be crossed by Diazepam and Morphine. The significant difference in the behavior of the pKa of ammonium and iminium ions in different solvents compared to carboxylic acids, alcohols and CH acids can again be understood from the changes in the solvation energies of the acids and their corresponding base. Considering Morphine, which was taken as an example for a tertiary amine (Figure 1f), we find that the positively charged ammonium ion behaves qualitatively similar to the carboxylate anion, e.g. it is destabilized in cyclohexane by roughly 1.5 eV compared to water (Figure 2b). This is again not surprising when considering that the apolar, aprotic solvent is to a much lesser degree able to stabilize the charged molecule. Since the changes in the proton solvation energy and the difference in the solvation energies between the acid and its corresponding base (Figure 2d) possess opposite signs and an approximately equal magnitude no significant changes in the pKa of the compounds in cyclohexane and water are observed.

Note that the scatter observed in Figure 1 is mainly the result of the incompatibility between the effective solvatochromic parameter which simply corresponds to the sum of α, β and π and the true relationship with the solvatochromic parameters. Indeed, using the exact expression to predict the difference in solvation energies between the acid and its corresponding base results in many cases in errors which are comparable to those found for carboxylic acids (Figure 2c and d). It must, however, be cautioned that the detailed relationship between the three parameters depends to some extent on the size of the dataset and in parts also on the fitting method which prohibits any more detailed interpretation of the results.

2.2. Exploring the Extreme: Trends in Cyclohexane

Equivalent trends are also found for the full data set, e.g. the pKa values of all carboxylic acids, alcohols and CH acids are strongly affected by the choice of solvent while it changes only slightly for most ammonium and iminium cations. However, while the qualitative trends are very similar within a class of functional groups, significant differences in the relative pKa shift are apparent between different molecules. Taking for example alcohols, the relative pKa values in cyclohexane compared to water vary between 24 pKa units and 100 pKa units (Figure 3). Strong variations are also observed in other solvents. In line with this, the mean absolute error (MAE) associated with the predicted pKa shift using only the Kamlet-Taft solvatochromic parameters is with 4.7 pKa units high. Owing to the presence of the very high bandwidth of relative pKa values, the standard deviation is with 7.8 pKa units even worse. A similarly poor performance of this oversimplified model is also observed for other functional groups (see SI). This strong scatter between the molecules can be rationalized from a detailed analysis of their origin.

2.2.1. Carboxylic Acids

Let us start the discussion with the carboxylic acids. Here, we find in many cases only rather minor deviations (Figure 3), e.g. 17 of the 31 considered acids display a relative shift which is within ±2 pKa units compared to what is found for Ibuprofen. More extreme deviations (>10 pKa units) are only observed for the second deprotonation step of oxalic and phthalic acid, Bendamustine, Cephradine, L–DOPA and L-Ornithine. Qualitatively similar trends are, with the exception of L–DOPA and L-Ornithine also observed in all other solvents.

The dissociation of Bendamustine (reaction 11) and Cephradine (reaction 12) results in the formation of a zwitterion with a cationic ammonium or iminium species and an anionic carboxylate.
This likely allows for a charge compensation effect which in turn reduces the impact of the formation of the carboxylate group on the solvation energy. As a result of this, reduced positive pKa shifts of only 21 (Bendamustine) and 27 pKa units (Cephradine) are observed. The slightly larger impact on Bendamustine might be a result of the fact, that the cationic group is somewhat closer to the R-COOH in this molecule compared to Cephradine but a larger data set containing more versatile zwitterions would be required to confirm this hypothesis. Even more extreme deviations are found for L-DOPA and L-Ornithine, which behave similar to ammonium and iminium ions. It is also noteworthy, that this effect is only present for toluene and cyclohexane which comprise the two most apolar, aprotic solvents. The origin of this anomalous behavior can be traced back to the structure of the two molecules. Here, the ammonium group is placed in α-position to the carboxylic acid which enables a proton transfer from the positively charged functional group to the carboxylate. This in turn quenches the negative charge and allows for the formation of a more favorable charge neutral system for L-DOPA (reaction 13).

\[
\text{O} \quad \text{NH}_2^+ \quad \text{Toluene or Cyclohexane} \quad -\text{H}^+ \quad \text{O} \quad \text{NH}_2
\]

(13)

For L-Ornithine on the other hand the full deprotonation sequence is reversed with the protons at the ammonium groups leaving first followed by the deprotonation of the carboxylic acid (reaction 14). This again minimizes the charge of the system at each deprotonation step.

\[
\text{H}_2\text{N} \quad \text{O} \quad \text{H}^+ \quad \text{Toluene or Cyclohexane} \quad -\text{H}^+ \quad \text{H}_2\text{N} \quad \text{NH}_2 \quad \text{O} \quad \text{NH}_2
\]

(14)

Naturally a similar inversion of the deprotonation sequence may also be observed in reality at other molecules which display both a carboxylic acid group and positively charged ammonium or iminium species.

Exchanging these data points with the respective formal dissociation of the ammonium ions which in reality correspond to a dissociation of a carboxylic acid, these compounds display shifts of approximately 46 pKa units which is in agreement with the other carboxylic acids. A similar proton transfer is likely also thermodynamically favorable for the other zwitterions but may be hindered kinetically in solvents which lack free protons. For solvents where this proton shift is absent, the carboxylic groups behave similar to the other zwitterions included in this data set. It is noteworthy, that the shift is even less positive than that observed for the other zwitterions which would support our hypothesis, that the proximity between the charged groups is decisive for the magnitude of the effect.

In contrast to the above cases, an increase of the relative pKa shift is observed for the second deprotonation step of oxalic and phthalic acid (reactions 15 and 16).

\[
\begin{align*}
\text{O} \quad \text{OH} \quad \text{H}^+ \quad \text{Toluene or Cyclohexane} \quad -\text{H}^+ \quad \text{O} \quad \text{OH} \\
\end{align*}
\]

(15)

\[
\begin{align*}
\text{O} \quad \text{NH} \quad \text{H}^+ \quad \text{Toluene or Cyclohexane} \quad -\text{H}^+ \quad \text{O} \quad \text{NH} \\
\end{align*}
\]

(16)

It is noteworthy, that the first deprotonation steps roughly behave similar to ibuprofen which indicates that these molecules themselves do not possess any special structural features which could explain these differences. Instead, the increase in the relative pKa shift coincides with the presence of a negatively charged functional group from the first deprotonation step. Opposite to the presence of positively charged ammonium and iminium groups, which allow for charge compensation and the formation of a net charge neutral system, this additional negative charge results in a further destabilization of the alcoholate in apolar/aprotic solvents such as cyclohexane.

2.2.2. Alcohols

More significant scatter is observed for the alcohols. Among the 53 data points only 17 display a relative pKa shift in cyclohexane which is within 5 pKa units of that observed for Ibuprofen. All other compounds are spread between a relative pKa shift of 24 pKa units and 100 pKa units. The rather large spread can partly be explained by the much stronger sensitivity to changes in the chemical environment which allows for the delocalization of the charge in the R–O– group. For the more extreme examples which display shifts below 35 pKa units, this is typically complemented by the presence of one or more electron withdrawing groups such as cyano (R–C≡N) or nitro (R–NO₂) groups.

The trends observed in the relative pKa shifts between water cyclohexane are also mirrored by the relative pKa values between the molecules in water which varies between that of very strong acids and very strong bases. Taking for example phenol (reaction 17), which has a computed pKa of 10 in water, we observe an increase by approximately 42 pKa units to 52 in cyclohexane.

\[
\begin{align*}
\text{OH} \quad -\text{H}^+ \quad \text{O}^-
\end{align*}
\]

(17)
Similar increases are also found for other compounds such as Morphine (reaction 18) whose alcoholic functional group possesses an almost identical pKa in water (computed pKa: 9.8).

Moving towards more acidic alcohols, the shift relative to water drops significantly. The most extreme example in our data set is pentacyanophenol (reaction 19) which has a computed pKa of 0 in water.

For this compound the pKa increases by only 24 pKa units. This decrease in the relative shift coincides with the presence of a delocalized \( \pi \) electron system in the phenol ring and 5 strongly electron withdrawing cyan groups as substituents which is also responsible for the more acidic pKa in water. Comparable effects are also observed for picric acid or dinoseb where the acidic OH group is directly attached to an electron-poor aromatic \( \pi \) electron system.

The presence of electron donating groups at the ring on the other hand pushes additional electron density to the negatively charged phenolate group which in turn results in a higher pKa in water. In non-aqueous solvents such as cyclohexane, this translates into a more positive shift of the relative pKa with respect to water. Examples of this are the second deprotonation steps of Propranolol (reaction 20) and Scopolamine (reaction 21).

Both compounds lack a conjugated \( \pi \) electron system close to the alcoholic group and are therefore unable to delocalize the negative charge of the alcoholate anion which renders them strong bases with pKa values of roughly 15 in water. In line with this, the relative shift of the pKa in cyclohexane versus water increases from 42 for phenol to approximately 50 for both compounds. The most extreme example is finally the deprotonation of water to hydroxide. Despite a pKa of 14 in water, which is slightly lower than those of the other examples of non-phenolic alcohols, it displays an even more positive shift of 57 pKa units in cyclohexane with respect to water. This extreme behavior is not completely unexpected when considering that the hydroxide ion displays a significantly higher negative Mulliken charge of roughly \(-1.26\) electrons at the oxygen in water which is only slightly reduced to \(-1.14\) in cyclohexane. The alkolate ions of Propranolol and Scopolamine on the other hand posses only Mulliken charges of \(-0.53\) and \(-0.58\) electrons at the oxygen adjacent to the acidic proton and are therefore somewhat less affected by highly apolar solvents. The importance of the molecules’ ability to delocalize the negative charge at the oxygen anion for the relative shift has also been reported in previous work for other solvents.\(^{[28]}\)

Identical to the carboxylic acids, compounds which display shifts of 55 or more pKa units in cyclohexane compared to water typically possess corresponding bases with two or more negatively charged functional groups. An example of such a system is 2,4-Dihydroxybenzoic acid (reaction 22) which displays an increase of 60 pKa units for the second proton resulting in the formation of an alkolate group in para position to the anionic carboxylate group.

Interestingly, the influence of the negatively charged functional group decreases with its increased distance to the acidic proton. This can for example be seen for the deprotonation of the alcoholic group in Salicylic acid, where the negatively charged carboxylate is in ortho position (reaction 23).

Accordingly, its effect on the alcohol is increased as shown by the increased relative shift in the pKa to 73 pKa units. This
increased shift again correlates with the increased computed pKa of the phenolic group in water which increases from 9.2 for the second deprotonation step in 2,4-Dihydroxybenzoic acid to 12.5 for Salicylic acid. Moving towards bases with even higher negative charges such as the third deprotonation step of 2,4-Dihydroxybenzoic acid, the relative shift of the pKa increases even further. For this final deprotonation step a relative shift of 89 pKa units compared to water is found. A similarly high shift is also observed for Dimercaprol where 2 thiol groups are deprotonated prior to the alcohol. Here, an increase of the pKa of the alcoholate group by 99 units compared to water is found. The increased shift is in line with the established trends in multi-protonic acids which indicate that each subsequent deprotonation step has a higher pKa.

2.2.3. Ammonium and Iminium Acids

The trends observed for the individual ammonium and iminium cations (Figure 1) are also found for most other compounds of this class, e.g. the shift in the pKa from water to cyclohexane is typically of the order of 5 to 10 pKa units. All acids which follow these trends display only a single positively charged functional group. If a second positively charged group is present, the shift becomes negative by approximately −10 pKa units which indicates that these systems are stronger acids in cyclohexane compared to water provided they could be dissolved in this highly apolar and aprotic solvent. Even more extreme negative shifts are observed if more than 2 positively charged groups are present in the acid. One of the most extreme negative shifts is for example observed for Chloroquine’s first deprotonation step (reaction 24) which possesses in total three positively charged ammonium or iminium groups and displays a negative shift of −32 pKa units in cyclohexane compared to water.

This increased shift is likely again a result of the interaction between the positively charged functional groups which render the acid less stable in aprotic/apolar solvents. In line with this hypothesis we find, that such an interaction is only possible if these groups are sufficiently close. The opposite trend is observed for zwitterionic ammonium and iminium compounds and neutral amines which react further to negatively charged amid anions. In case of the former we observe a shift which is comparable to that observed in the case of the formation of zwitterionic bases, e.g. the pKa has a moderate positive shift of the order of 20 pKa units. The latter systems on the other hand behave similar to carboxylic acids and alcohols as indicated by a shift of ≈40 pKa units for the first deprotonation step forming a single negatively charged group and more than 50 pKa units for the second step.

2.2.4. C-H Acids

The last class of interest are C–H acids. In contrast to the previous classes of acids we observe only a rather minor scattering between the data points, e.g. the relative pKa shifts from water to cyclohexane varies by only ≈8 pKa units. This is not surprising when considering that the included acids are rather homogeneous with respect to the chemical environment. In practice this means that our data set only includes C–H acids where the negative charge is strongly stabilized through a π electron system and electron withdrawing groups. Furthermore, no zwitterionic molecules or compounds with multiple charged groups are present. As an additional result of the good stabilization of the negative charge in the corresponding base, pKa values in water are, in comparison to the common acidity of C–H groups, with pKa values between −1 and 10 already rather low. Following the above discussion for the alcohols, this already indicates, that the relative pKa shifts will be at the lower end of the spectrum that could be expected for neutral acids. This is indeed observed for all considered compounds, e.g. the relative shift is of the order of 30 to 35 pKa units.

2.3. Global Perspective

Building on the detailed analysis in cyclohexane, a universal picture of the behavior of different acids in non-aqueous solvents can be developed. Our data clearly show that the relative shift of the pKa in different solvents mostly depends on the presence of additional anionic or cationic groups and the charge of the acidic functional group. A similar dependence has also been suggested by Rossini et al.48 Other factors like the exact chemical environment or the nature of the functional group have a much smaller impact. In the absence of any other charged functional groups neutral acids display a shift of ≈40 pKa units in cyclohexane (Figure 4). Naturally the impact is smaller for more polar/protic solvents and also decreases with stabilizing factors such as conjugated π electron systems and electron withdrawing groups. Equivalent shifts are also found for the thiol and the mineral acids which are present in our data set but were neglected during the previous analysis. In line with our above analysis we again find rather minor shifts for triflic acid and fluorsulfonic acid whose corresponding bases are both stabilized by strongly electron withdrawing groups. If one or more anionic groups are already present in the acid, the relative shift increases by ≈15 pKa units per anionic group. In the presence of a cationic group, which corresponds to the formation of a zwitterionic base, the shift decreases by the same amount (Figure 4). A shift similar to that associated with the formation of a zwitterionic base is also observed for the deprotonation of a cationic functional group in a zwitterionic acid (Figure 4). In the absence of an anionic group, the relative
shift decreases additionally by 10 to 20 pKa units to roughly 5 to 10 pKa units. A similar decrease is also approximately observed for each additional cationic group in the molecule (Figure 4).

The trends observed for cyclohexane can be generalized to other solvents through an empirical model. For the sake of simplicity we only include the absolute shift in the absence of any other charged groups and the electrostatic interactions, which we identified as the main factor for determining the relative shift of the pKa with respect to water. These contributions are scaled by the solvatochromic parameters which describe the solvents’ properties. The fitting parameters are obtained from a LAD fit using the solvatochromic parameters reported in reference [57] and our computed pKa values. We additionally divided all acids into 2 groups based on the charge of the acidic functional group. Group I acids possess a cationic acidic group while group II acids are characterized by a neutral acidic group. This division is necessary in order to properly describe the behavior of zwitterionic acids and bases. Other contributions such as the detailed chemical structure into which the acidic group is embedded in are neglected. Similarly, also the influence of the spatial distance between charged groups is not taken into account. Including these factors would require a significantly larger and more diverse data set and could potentially be achieved using machine learning or related algorithms. Using our simple empirical relationship we obtain the equations 25 (group I acids) and 26 (group II acids).

**Group I (Cationic acids)**:

\[
pKa(SolvX) = (1.41\alpha^* - 6.44\alpha + 0.63\beta + 6.95) + (Z_{acid} - 1)(8.58\pi^* + 3.54\pi + 8.98\beta - 15.76) + pKa(Water)
\]

**Group II (Neutral acids)**:

\[
pKa(SolvX) = -10.85\pi^* - 12.48\alpha - 13.46\beta + 31.23 + pKa(Water)
\]

Here, \(Z_{acid}\) corresponds to the effective charge of the acid. In most cases the effective charge corresponds to the total charge of the molecule but deviations are possible if some of the innocent charged groups are unable to interact with the acidic group or its corresponding base owing to the spatial distance between them or other inhibiting factors.

The validity of the fits is tested through comparison of predicted and explicitly computed pKa values in different solvents (Figure 5). Considering the simplicity of the model we observe in many cases a surprisingly good agreement between the predictions obtained from DFT and the model summarized in Equations 25 and 26 (Figure 5). For non-zwitterionic group I acids the predicted pKa values are within a mean absolute error (MAE) of 2.5 pKa units (\(\alpha = 3.5\) pKa units). The error decreases even to a MAE of 1.9 pKa units (\(\alpha = 2.5\) pKa units) for zwitterionic group I acids but is increased to a very high MAE of 3.9 pKa units (\(\alpha = 5.9\) pKa units) for group II acids. The rather high error bars observed for the latter class of compounds reflects the high diversity of pKa values for alcohols and neutral imines. This is especially visible in the very high standard deviations which points at the existence of significant outliers which are also apparent from visually evaluating the agreement between predictions and computations (Figure 5). These deviations are a result of the direct interactions with conjugated \(\pi\) electron systems and electron donating or withdrawing functional groups. In line with this, we do find a rather good agreement for cationic acids which typically are unable to stabilize the charge through interaction with a \(\pi\) electron system. Naturally, these effects are not captured by our simple model which only relies on the solvatochromic parameters which describe the solvent properties and the total charge of the molecule. Overall, this simple model nevertheless provides at least a qualitative description of the expected trends which often agree surprisingly well with the predictions obtained from...
DFT modeling. However, the deviations of the model increase when moving towards more extreme pKa values. Thus, it is clear that an improved machine learning model which also includes structural information would be needed to achieve sufficiently good accuracy for most applications.

Nevertheless, the qualitative trends may be used to obtain some global understanding on the factors which affect the pKa shifts the most. Evaluating the prefactors to the solvatochromic parameters in equations 25 and 26, we see that both classes of acids behave very differently. According to our fit, the relative pKa shift is, in the absence of any additional charged groups, almost exclusively affected by the hydrogen bond donation ability ($\alpha$) which results in a decrease. As a result of this, all aprotic solvents behave rather similar while a significant stabilization resulting in a reduced pKa shift is only observed for protic solvents. However, in the presence of charged groups, this picture changes. Here, the solvents ability for hydrogen bond acceptance ($\beta$) and its polarizability are most important for minimizing the influence of charged functional groups while the hydrogen bond donation is less relevant.

Neutral group II acids on the other hand are affected equally by the solvents polarizability and its ability for hydrogen bond acceptance and donation. As a result of this, the relative pKa shift continuously increases with a decrease in the hydrogen bond ability of the solvent and its polarizability. Similarly also the effect of charged groups is countered equally by the polarizability and the hydrogen bond donation and acceptance properties of the solvent. Thus, the additional shift in the presence of additional charged groups increases again upon moving towards more aprotic apolar solvents. It must, however, be cautioned that the pKa shifts also depend on structural details of the molecule such as the closeness between the charged groups or the ability for delocalizing the charges which renders the fit not necessarily fully reliable.

**3. Conclusions**

In summary we have shown global trends between the pKa values of more than 180 compounds of general and pharmacological interest in 21 solvents. Our computations indicate, that the influence of the solvent on the pKa is mostly determined by the charge of the functional group and the presence of other charged groups in the molecule. The pKa of neutral acids such as alcohols or carboxylic acids are strongly affected by the solvents' ability to stabilize the involved species through dipole interactions or hydrogen bonding whereas this effect cancels for cationic acids like ammonium or iminium ions. Structural details such as the possibility to delocalize through a conjugated $\pi$ electron system or the presence of electron donating or withdrawing groups are of relevance but typically have a much lower influence. Overall, the magnitude of the relative shift between water and non-aqueous solvents resembles the trends between pKa values of the different compounds in water.

Based on our results we developed a simple empirical model to estimate the pKa of any acid in a freely chosen solvent based on its pKa in water and the solvatochromic parameters. This allows for easy prediction of pKa values in all solvents with fair accuracy. Further improvements of the model to also include more subtle contributions such as mesomeric and inductive effects or the distance between charged groups will be needed to achieve chemical accuracy.

**Acknowledgments**

Calculations were performed at the Chalmers Centre for Computational Science and Engineering (C3SE) and the Finnish IT centre for science (CSC). M.B. and K.L. acknowledge financial support from the Jane and Aatos Erkko Foundation through the “Renewable energy storage to high value chemicals” project. M.B. also acknowledges support from the Dr. Barbara-Mez-Starck foundation. Open Access funding enabled and organized by Projekt DEAL.