Thermodynamically effective molecular surfaces for more efficient study of condensed-phase thermodynamics

Amin Alibakhshi (alibakhshi@ptc.uni-kiel.de)
Theoretical Chemistry, Institute for Physical Chemistry, Christian-Albrechts-University

bernd hartke
Theoretical Chemistry, Institute for Physical Chemistry, Christian-Albrechts-University

Physical Sciences - Article

Keywords: condensed-phase thermodynamics, molecular surfaces, “thermodynamically effective”

DOI: https://doi.org/10.21203/rs.3.rs-816803/v1

License: Creative Commons Attribution 4.0 International License.
Read Full License
Thermodynamically effective molecular surfaces for more efficient study of condensed-phase thermodynamics

Amin Alibakhshi\textsuperscript{1,}\textsuperscript{*}, Bernd Hartke\textsuperscript{1}

Theoretical Chemistry, Institute for Physical Chemistry, Christian-Albrechts-University, Olshausenstr. 40, 24118 Kiel, Germany

Corresponding author: alibakhshi@pctc.uni-kiel.de

Abstract

Evaluation of molecular surfaces plays the key role in a wide range of cutting-edge scientific fields and technologies, due to the well-characterized dependency between molecular surfaces and condensed phase thermodynamics. Numerous methods to evaluate molecular surfaces such as van-der-Waals and solvent accessible surface areas and various parameterizations for each one, have been proposed in the literature and typically yield quite diverse estimations of molecular surfaces. Despite this diversity, numerous successful applications have been reported for each one, which has become possible via ad-hoc modifications and parametrizations employed to accommodate inappropriately defined molecular surfaces.

The main aim of the present study is to propose “thermodynamically effective” molecular surface which unlike the conventionally accepted molecular surfaces, can be defined only uniquely, can be measured experimentally for each molecule directly and straightforwardly, is defined based on a well-characterized theoretically described dependency between molecular surfaces and solution thermodynamics, and is highly accurate in evaluating various thermodynamics quantities in solution for a wide temperature range and different types of molecules, without requiring any ad-hoc modification.

1- Introduction

Molecular surface area of molecules is one of the key molecular parameters in a number of extensively used theoretical approaches for studying physical chemistry of the condensed phase. Evaluation of molecular surfaces plays a key role in a large number of cutting-edge scientific fields and technologies such as drug discovery\textsuperscript{1,2}, catalysis\textsuperscript{3}, molecular biology\textsuperscript{4-6}, molecular genetics\textsuperscript{7} and nanotechnology\textsuperscript{8,9}. One of
the main applications of molecular surface estimation is theoretical evaluation of solution thermodynamics via continuum solvation models \textsuperscript{10} which is an extensively applied method in very diverse scientific fields, ranging from catalysis \textsuperscript{11,12}, advanced nanomaterials \textsuperscript{13}, surface science\textsuperscript{14}, or mechanisms of chemical reactions in the condensed phase \textsuperscript{15,16} to unraveling the activity mechanism of coronavirus \textsuperscript{17}.

Employing molecular surfaces for unravelling scientific challenges in the condensed phase has been an active scientific area with a history of more than a century. One of the earliest attempts in this regard dates back to 1886 and was proposed by Eötvös. He suggested a proportionality between the free energy per unit of interfacial surface, i.e. the surface tension, and the surface area of liquid-phase molecules \textsuperscript{18}. This work was indeed one of the earliest examples of experimental evaluation of molecular surfaces, which was achieved by assuming solution-phase molecules as perfect spheres, allowing to evaluate molecular surface area via liquid molar volume. Although the assumption of perfect spheres is the simplest approach to get a rough estimation of molecular surfaces, it satisfactorily holds for mono-atomic molecules. Accordingly, the earliest successful applications of molecular surfaces to study solution thermodynamics both exploit surfaces determined via perfect sphere assumption and are mainly limited to noble gases \textsuperscript{19,20,21,22}.

In 1964, the van-der-Waals (vdW) surface area concept was proposed in the pioneering work of Bondi \textsuperscript{23}, which became the cornerstone of more advanced molecular surfaces such as Solvent Excluded Surfaces (SES) and Solvent Accessible Surface Area (SASA). Since then, a large number of methods and algorithms have been proposed to evaluate molecular surfaces in solution, in particular multiple variants of solvent excluded or solvent accessible surface areas. This wide variety of methods typically yields quite diverse estimations of molecular surfaces, as can be seen in a comparison of vdw and SAS molecular surfaces of ethylene, as depicted in figure 1. For a broader comparison, we provide computed molecular surfaces for 215 molecules via different parameterizations of the vdw method as supplementary material. These data imply that selecting the most appropriate method is not a trivial task.

![Figure 1- comparison of vdw (left) and SAS (right) surfaces in ethylene](image)
Surprisingly, despite the diversity of methods and molecular surface approximations they yield, there are numerous examples of reporting successful applications for each one of these methods in studying thermodynamics in solution. This is mainly because the majority of these research works are based on empirically defined relationships between solution thermodynamics and molecular surfaces \(^{24,25}\). Inaccuracies due to deviations of the employed molecular surfaces from the actual values are then corrected via ad-hoc modifications and parametrizations, mainly applied to atomic radii. For example, while the Gaussian 03 software package used SES surfaces and UA0 atomic radii as default for computing solvent effects based on the polarizable continuum solvation models, for the latest release of the same software, vDW surfaces and UFF atomic radii are considered as default \(^{26}\). Similarly, the most widely applied continuum solvation models exploit their specifically defined molecular surfaces and parameterizations of atomic radii \(^{10}\).

One main reason behind this diversity in defining and parameterizing molecular surfaces and requiring such ad-hoc modification is that although the main application of molecular surfaces are commonly for studying solution phase thermodynamics, they are typically parameterized for reproducing other target quantities. For example, the Bondi parameterization of atomic radii has been done using physical quantities like X-ray diffraction data, gas kinetic collision cross section and liquid density as target quantities\(^{23}\) while the UFF or UA0 atomic radii are parametrized against bond distances \(^{27}\).

It indeed stems from unavailability of a rigorous theoretical method which allows precise and analytical characterization of the relationship between molecular surfaces and solution thermodynamics without ad-hoc modifications or parameterization.

Obviously, such a theoretical method offers a number of advantages. First and foremost, it allows defining molecular surfaces with physical significance, without requiring ad-hoc modifications and thus uniquely definable. Furthermore, it provides the possibility of evaluating the performance of various methods and parametrizations in reproducing those reference molecular surfaces. Last but not least, it allows for better understanding and treating some of the main challenges in theoretical studies of solvation, such as appropriate treatment of solvent effects in continuum solvation models.

To achieve this goal, in the present study we exploit a theoretically derived relationship describing temperature dependence of vaporization enthalpy to the molecular surfaces in solution which is an extension of a recent study\(^{28}\) leading to a remarkable improvement of the formerly reported results.

Among all potential thermodynamics quantities of solution which can be analytically related to molecular surfaces for this purpose, the vaporization enthalpy, which is employed in the present study, possesses a number of obvious advantages. The main one is that vaporization enthalpy can be directly determined
experimentally, while free energy or entropy can only be determined indirectly and via measuring the temperature dependence of vaporization enthalpy or equilibrium vapor pressures at multiple temperatures, which implies accumulation of errors inherent in both experimental measurements and the subsequent computations. The more convenient experimental procedure for enthalpy measurement has also made accurate benchmark datasets more readily available, which is another advantage of using vaporization enthalpy. Finally, evaluation of molecular surfaces via vaporization enthalpy is not only both more accurate and less challenging but also once it is found, it can be conveniently used to obtain relationships between molecular surfaces and other thermodynamic quantities, via the fundamental thermodynamics relationships, as shown in section 4-3.

2- Theory

By considering vaporization as a dynamic process at which evaporation and condensation have the same rates and equating the rates of evaporation and condensation described by transition state theory and some manipulations, the ratio of partition functions of the gas and liquid phases is obtained as 28:

\[
\frac{Q^g}{Q^s} = \frac{N_a P_{sat}}{p} \frac{(k_B T)^{1/2}}{(2\pi m)^{1/2}} \frac{h}{k_B T |n_s|} \exp \left( \frac{\Delta\varepsilon_{sg}}{k_B T} \right),
\]

(1)

where \(N_a\) is Avogadro’s constant, \(P_{sat}\) is the saturation vapor pressure of the liquid, \(\Delta\varepsilon_{sg}\) is the energy for moving one molecule from the liquid surface to the gas phase, and \(k_B\) and \(h\) are Boltzmann and Planck constants, respectively. Using the statistical thermodynamics relationship between the energy and partition function stated as:

\[
\langle \varepsilon \rangle = k_B T^2 \frac{\partial \ln (Q)}{\partial T},
\]

(2)

and with some algebraic manipulations, it can be shown that the temperature dependence of the vaporization enthalpy follows 28:

\[
\Delta h_{vap} = \Delta\varepsilon_{bs} - \frac{k_B}{2} T \ln (T) - T \int \frac{\Delta\varepsilon_{bs}}{T^2} dT + CT,
\]

(3)

where \(C\) is a constant and \(\Delta\varepsilon_{bs}\) is the energy required for moving one molecule from the bulk of the liquid to the surface. Evaluation of \(\Delta\varepsilon_{bs}\) via experimentally measurable quantities can be achieved using the fundamental thermodynamics relationships between energy (\(\varepsilon\)), Helmholtz free energy (\(f\)) and entropy (\(s\)), which implies 28.
Δ𝜀_{bs} = ∆f_{bs} - T \frac{d(∆f_{bs})}{dT}, \quad (4)

where ∆f_{bs} is the free energy change for moving one molecule from the bulk of liquid to the surface. Another straightforward way to obtain Eq. (4) is using the Gibbs-Helmholtz equation:

\frac{d(\frac{f}{T})}{dT} = -\frac{\varepsilon}{T^2}, \quad (5)

which implies

\frac{1}{T} \frac{df}{dT} - \frac{f}{T^2} = -\frac{\varepsilon}{T^2}, \quad (6)

This then clearly yields Eq. (4) by subtracting the resulting equations for the bulk and surface states.

Exploiting the thermodynamics relationship among ∆f_{bs}, surface tension (γ) and the molecular surface area \(a_s\) which is defined as\(^28\):

\[ \Delta f_{bs} = \frac{a_s}{2} \gamma, \quad (7) \]

Eq. (4) can be rewritten as:

\[ \Delta \varepsilon_{bs} = \frac{a_s}{2} (\gamma - T \frac{dy}{dT}). \quad (8) \]

Halving the molecular surfaces \(a_s\) in Eq. (7) is considered here because in fact only one half of the molecular surfaces contribute in forming the gas-liquid interface and the other half remains in the bulk of the liquid\(^28\).

By substituting Eq. (8) in Eq. (3) and using \(\frac{d(\frac{f}{T})}{dT} = T \frac{dy}{dT} - \gamma\) and multiplying both sides by Avogadro’s number, the correlation between the surface tension and molar vaporization enthalpy (\(ΔH_{vap}\)) is obtained as:

\[ ΔH_{vap} = \frac{a_s}{2} N_A \left( 2 \gamma - T \frac{dy}{dT} \right) - \frac{R}{2} T \ln(T) + \beta T, \quad (9) \]

in which \(\beta\) is a constant. Knowing that at the critical temperature both vaporization enthalpy and surface tension approach zero, and due to continuity of the surface tension, the \(\frac{dy}{dT}\) term also approaches zero, the constant \(\beta\) is found as:

\[ \beta = \frac{R}{2} \ln(T_c). \quad (10) \]
which by substitution into Eq. (9) finally results in:

$$
\Delta H_{\text{vap}} = \frac{a_s}{2} \left( 2 \gamma - T \frac{d\gamma}{dT} \right) - \frac{R}{2} T \ln \left( \frac{T}{T_c} \right).
$$  \hspace{1cm} (11)

As discussed earlier, the main advantage of the theoretically derived relationship among vaporization enthalpy, surface tension and molecular surfaces described by Eq. (11) is the possibility of experimental determination of molecular surfaces through this relationship as well as evaluation of the currently defined molecular surfaces. For the latter goal, we have compared accuracy of evaluated vaporization enthalpies obtained via Eq. (11) using 252 differently computed molecular surfaces based on various parameterizations of vdW and SAS methods.

3- Experimental

3-1- Dataset

The theoretically derived methods were benchmarked using thermophysical data of the DIPPR801 database. Screening the initial dataset and selecting only the compounds with maximum uncertainty of 5% for vaporization enthalpy and surface tension resulted in 215 compounds from diverse chemical families, provided as supplementary materials.

For each compound, the experimentally determined data of vaporization enthalpies for 25 temperatures linearly distributed between the melting point and the critical temperature were evaluated using the provided relationships in the DIPPR database. Due to the scarcity of accurate surface tension data at all of the required data points at which vaporization enthalpy data were available, we employed the Guggenheim–Katayama relationship stated as:

$$
\gamma = \gamma^\circ \left(1 - \frac{T}{T_c}\right)^{11/9},
$$  \hspace{1cm} (12)

as a rigorous model for evaluating temperature dependence of surface tension and its temperature derivatives required by Eq. (11). One main advantage of employing the Guggenheim–Katayama relationship is that it perfectly satisfies the boundary conditions, i.e. yielding exactly zero for surface tension and its higher order derivatives with respect to temperature at the critical point, which is commonly violated by other relationships like those proposed in the DIPPR dataset. This is indeed a key requirement, as it was one of the premises of obtaining the constant $\beta$ in Eq. (11). Additionally, compared to various surface tension predictive models such as those used by the DIPPR database and the Eötvös relationship, we...
found that the Guggenheim–Katayama relationship provides the most accurate evaluation of not only temperature dependence of surface tension but also its derivate for the whole temperature range and hence, the most accurate prediction of vaporization enthalpy via Eq. (11).

To obtain surface tension data at the required temperatures via the Guggenheim–Katayama relationship, for each compound the pre-factor $\gamma^\circ$ was calculated by optimization using the available experimental data points of surface tension. The calculated values of $\gamma^\circ$ for each compound is reported in the supplementary materials.

The accuracy of the predicted vaporization enthalpy is reported as Average Absolute Deviation (AAD), defined as:

$$ AAD = \frac{1}{N} \sum \{ |y_i^{exp} - y_i^{pred}| \}. \tag{13} $$

Considering that at the critical point, vaporization enthalpy approaches zero and slight deviations in predicted enthalpies results in very large relative errors, AAD provides a more reliable evaluation of the model performances, compared to relative errors.

3-2- Computational details

To calculate well-established molecular surfaces, the geometry of each molecule was first optimized at the B3LYP/6-311+G (2d,p) level of theory. Using the optimized structures, the vdW molecular surfaces and Surface Accessible Surface Areas (SASA) for each molecule were calculated using the GEPOL algorithm implemented in the Gaussian 16 software.

For both vdW and SASA, we have calculated molecular surfaces using UA0, Pauling, Bondi and UFF parameterizations of atomic radii, as available in the Gaussian 16 software. For each one of the mentioned parameterizations, we have also studied the atomic radii scaled by factors 0.85, 0.9, 0.95, 1.0, 1.05, 1.1, 1.15, 1.2 and 1.25.

For SASA, we computed solvent radii required for calculation of solvent accessible surfaces using radii of perfect spheres with volumes equal to the vdW volume and also molar volumes of the solvents at their melting points. For each one of the employed solvent radii, in addition to the original values, radii scaled by factors of 0.7 and 0.85 were also studied. The reason of such scaling is that we noticed for most of the studied molecules, the radii computed via vdW or melting point volumes deviate from the respective solvent radii used in Gaussian 16 in calculation of continuum solvation models by a factor varying between these two values.
With all this diversity in details, in total 252 differently computed molecular surfaces, in addition to molecular surfaces approximated based on the perfect sphere assumption and molar volumes at melting point and normal boiling point, were studied for each molecule, and their performance in reproducing experimentally determined vaporization enthalpy via Eq. (11) was tested.

4- Results and discussion

4-1- Verifying the validity of the theoretically derived relationship

By studying various relationships proposed in the past century to analytically relate solution thermodynamics, surface tension and molecular surfaces, we noticed an obvious inconsistency not only between previously proposed models themselves but also with the theoretically derived relationship proposed in the present study as well. Accordingly, while a large number of studies support:

\[
\Delta G_{C…†QAD…7} = A_2 \gamma + B,
\]

(14)

where \(\Delta G_{solvation}\) is the free energy of solvation, \(A\) is the solvent excluded surface of molecules and \(B\) is a constant, many other works employ the very similar relationship to relate vaporization enthalpy, surface tension and molecular surfaces. A well-known example of the latter category is Kabo’s method which proposes:

\[
\Delta H_{QAR} = A_2 + V_{Y} \gamma^2 + B,
\]

(15)

and has been extensively applied specially in studying the phase equilibrium in ionic liquids. Noteworthy, the right hand side of Kabo’s relationship closely matches the right hand side of Eq. (14) with the only difference that in Kabo’s relationship, the pre-factor of the surface tension is in fact molecular surfaces evaluated based on the perfect sphere assumption discussed earlier, while in Eq. (14) this factor is solvent excluded surface of molecules.

Alongside the mentioned paradoxical deviations between the two models, both of them also show obvious inconsistencies compared to our theoretically derived relationship proposed in Eq. (11) which implies the necessity of a careful and rigorous verification of our model.

To that end, we first evaluated the overall accuracy of vaporization enthalpies predicted via the newly developed relationship. Accordingly, for each compound we optimized the value of the \(a_s\) parameter
required by Eq. (11) which yielded lowest error in predicting vaporization enthalpies over the whole temperature range. Via the optimized \( a_s \) parameters, which are in fact our proposed “thermodynamically effective” molecular surfaces, an AAD of 0.188 kcal/mol was obtained for the predicted vaporization enthalpies of the whole dataset. This resulting AAD is within both the chemical accuracy and the reported accuracy of the employed, experimentally determined data.

Interestingly, we noticed that the molecular surfaces evaluated via molar volumes at the melting point based on the perfect sphere assumption divided by the factor of two to include only the part of the molecular surfaces which contribute to the interface as discussed earlier, closely match the thermodynamically effective molecular surfaces with an average ratio of 1.0372, standard deviation of 0.117 and correlation coefficient of 0.95. These results clearly verify the robustness of the theoretically derived relationship and its underlying presumptions like defining the \( a_s \) parameters as molecular surfaces at interface and considering the factor of \( \frac{1}{2} \) for it.

Noteworthy, using the molecular surfaces evaluated from molar volumes at the melting point and without any adjustable parameters, Eq. (11) yielded an AAD of 0.561 kcal/mol for the whole dataset and the whole temperature range. On the other hand, using molecular surfaces obtained via molar volumes at normal boiling points resulted in AAD of 1.248 kcal/mol. These results clearly show that the perfect sphere assumption is not generally a rigorous approach for evaluating molecular surfaces. In addition to lower accuracy in describing temperature dependence of vaporization enthalpy compared to the thermodynamically effective molecular surfaces, it totally overlooks temperature dependence of molar volume and its impact on the evaluated molecular surfaces.

To evaluate the significance of the \( T \frac{dy}{dT} \) and \(-\frac{R}{2} T \ln \left( \frac{T}{T_c} \right)\) terms, which are the most obvious differences between our proposed relationship and the conventionally accepted models, we re-optimized \( a_s \) parameters for the two following relationships:

\[
\Delta H_{vap} = a_s \gamma - \frac{R}{2} T \ln \left( \frac{T}{T_c} \right),
\]

(16)

and

\[
\Delta H_{vap} = \frac{a_s}{2} \left( 2 \gamma - T \frac{dy}{dT} \right),
\]

(17)

which are two variants of Eq. (11) obtained by removing the term being studied. Using the re-optimized \( a_s \) parameters, the two abovementioned variants yielded AADs of 2.071 and 0.197 kcal/mol, respectively. These results clearly show that both terms \( \frac{dy}{dT} \) and \(-\frac{R}{2} T \ln \left( \frac{T}{T_c} \right)\) play a significant role in describing the
relationship between solution thermodynamics and molecular surfaces. Among them, the $T \frac{\partial r}{\partial r}$ term has the more significant impact and its overlooking reduces the accuracy of predicted vaporization enthalpies by one order of magnitude.

**4-2- Evaluation of molecular surfaces estimated via computer algorithms**

After verifying the validity and robustness of the theoretically derived relationship in the previous section, this section focuses on studying predictability of vaporization enthalpies obtained via various parameterizations of the vdW and solvent accessible surfaces. To that end, we have studied a total number of 252 differently computed molecular surfaces discussed in section 3-2.

According to the results, while via the solvent accessible surfaces we could not achieve any AAD better than 7.889 kcal/mol, for the vdW surfaces the best results with AAD of 0.568 kcal/mol was obtained for UA0 atomic radii scaled by 0.9. For the vdW surfaces, the results obtained for original parameterizations of atomic radii without any scaling yielded AADs of 0.613, 1.078, 1.256 and 0.867 kcal/mol for Bondi, UFF, UA0 and Pauling parameterizations, respectively.

These results show that which method is selected for approximating molecular surfaces clearly has a remarkable impact on the thermodynamic quantities evaluated with those surfaces. For a better illustration, the AADs of predicted vaporization enthalpies obtained via Eq. (11) for different scalings and parameterizations of atomic radii are compared in figure 2. As can be seen in figure 2, even slight differences in approximated molecular surfaces can have a significant impact on the evaluated vaporization enthalpies in solution.

Noteworthy, the vdW surfaces with UA0 atomic radii scaled by 0.9, which yielded better results compared to other parameterizations for vaporization enthalpy prediction, after scaling by the $\frac{1}{2}$ factor as discussed earlier, show an excellent agreement with the calculated thermodynamically effective surfaces, with correlation coefficient of 0.96, average ratio of 1.019 and standard deviation of 0.172.

This immediately suggests employing thermodynamically effective surfaces for a more rigorous parameterization of atomic radii in the currently defined molecular surfaces.
In figure 3, the temperature dependence of vaporization enthalpy of some of the most widely used solvents predicted via Eq. (11) using various molecular surfaces are compared. These results clearly show the significant importance of evaluated molecular surfaces on the accuracy of obtained results on the one hand, and the robustness and reliability of the newly derived relationship and the thermodynamically effective molecular surfaces on the other hand.
Vap. enthalpy (kcal/mol) vs. T (K) for:

- (water)
- (phenol)
- (methyl acetate)
- (benzene)

Comparison of experimental data (Exp.) and predictions from Eq. (11) for different interactions:

- TE
- melt
- vdw
Figure 3- Comparison of the impact of various molecular surfaces on the accuracy of predicted vaporization enthalpies. The results show a remarkably enhanced description of vaporization enthalpy at various temperatures using our newly proposed
As discussed in section 4-1, there are paradoxical inconsistencies between our proposed relationship for describing dependency of solution thermodynamics to molecular surfaces with those proposed in the literature. To compare our proposed relationship with conventionally accepted models, we evaluated the accuracy of vaporization enthalpies predicted via Kabo’s relationship, for which the two constants were determined by optimization using the whole vaporization enthalpy data for each compound. According to the results, while the newly derived relationship with optimum molecular surfaces yielded AAD of 0.188, Kabo’s method with two empirical parameters optimized for each compound could not yield an AAD better than 0.402 kcal/mol which is comparable to the results obtained via molecular surfaces at melting point without needing any adjustable parameters, as reported earlier. The same comparison of our proposed relationship and Eq. (14) as the other conventionally accepted model for relating solution thermodynamics to molecular surfaces is provided in section 4-3.

4-3- Evaluation of other thermodynamics quantities via thermodynamically effective surfaces

As it was discussed earlier, one of the main advantages of characterizing dependency of solution thermodynamics on molecular surfaces through temperature dependence of vaporization enthalpy is its straightforward transferability to other thermodynamics quantities.

As one of the most important thermodynamics quantities, evaluation of the solvation free energy, which is the primary goal in continuum solvation models\cite{10}, can be achieved via the Gibbs-Helmholtz relationship as follows:

$$
\Delta G_{solvation} = -T \int \frac{\Delta H_{\text{vap}}}{T^2} \ dT. \tag{18}
$$

Via the analytical relationship describing temperature dependence of vaporization enthalpy, one can solve Eq. (18) analytically or numerically. The constant of integration then needs to be determined using a reference data point. To that end, the most convenient choice is to use the free energy of solvation at the normal boiling point, where the saturation vapor pressure of the liquid becomes equal to the atmospheric pressure. It then allows to determine the solvation Gibbs free energy at the normal boiling point through
atmospheric pressure \( P_{atm} \), liquid molar volume at the normal boiling point \( V_{m,nbp} \) and normal boiling point temperature \( T_{nbp} \) via \(^{36}\):
\[
P_{atm} = \frac{R T_{nbp}}{V_{m,nbp}} \exp \left( \frac{\Delta G_{sol,nbp}}{k T_{nbp}} \right).
\] (19)

The abovementioned dependency between vaporization enthalpy and solvation free energy also allows a quantitative comparison of our theoretically derived relationship with the other conventionally accepted model, which in contrast to our model suggests linear dependency between solvation thermodynamics and the product of molecular surfaces and surface tension via Eq. (14). To that end, we studied the predictability of the solvation free energies obtained by numerically integrating the Gibbs-Helmhotz equation for the experimentally determined vaporization enthalpies as a function of temperature and used the obtained free energies to find the parameter \( B \) in Eq. (14) by optimization. For the solvation free energies predicted this way, the results showed an AAD of 1.166 kcal/mol which is again much higher than the AAD of computed vaporization enthalpies. Noteworthy, employing the newly proposed relationship for numerically integrating the Gibbs-Helmhoötz equation could reproduce the reference solvation free energies with AAD of only 0.057 kcal/mol. The much higher accuracy obtained via thermodynamically effective surfaces for predicted solvation free energies is because the solvation free energies are proportional to the area under the curve of vaporization enthalpy versus temperature and therefore are less affected by outliers. These results also suggest that considering linear proportionality between vaporization enthalpy and surface tension, which has already been empirically suggested in numerous studies like Kabo’s relationship and is supported in the present study, is more valid than the other conventionally employed relationships which consider the same proportionality between solvation free energy and surface tension.

We also studied predictability of experimentally determined standard state solvation free energies reported in the Minnesota Solvation database \(^{37}\) via thermodynamically effective surfaces for the solutions common between the DIPPR and Minnesota Solvation databases. The results, which are reported in table 2, show an excellent agreement between predicted solvation free energies obtained via the thermodynamically effective surfaces and the experimentally determined values reported in the Minnesota Solvation database. The obtained accuracy of these results shows an AAD of only 0.1215 kcal/mol which is by almost a factor of 2 more accurate than best results obtained via advanced continuum solvation models \(^{10}\).

| Table 2- Comparison of standard state solvation free energies theoretically predicted using thermodynamically effective surfaces and experimentally determined data |
|---|
| 15 |
| Compound                        | Exp. (kcal/mol) | Predicted (kcal/mol) |
|--------------------------------|-----------------|----------------------|
| Acetonitrile                   | -4.85           | -4.9191              |
| 2-methylpyridine               | -5.71           | -5.8055              |
| Acetophenone                   | -7.59           | -7.7809              |
| Aniline                        | -7.61           | -7.5768              |
| Anisole                        | -6.33           | -6.432               |
| Benzene                        | -4.55           | -4.5595              |
| Benzonitrile                   | -7.28           | -7.4975              |
| Bromoethane                    | -3.67           | -3.7014              |
| Bromoform                      | -6.21           | -6.3047              |
| 1-butanol                      | -6.03           | -6.1576              |
| Chlorobenzene                  | -5.66           | -5.742               |
| Chloroform                     | -4.13           | -4.1932              |
| Cyclohexane                    | -4.43           | -4.4176              |
| Cyclohexanone                  | -6.25           | -6.3928              |
| 1-decanol                      | -9.58           | -9.3872              |
| N,N-dimethylacetamide          | -6.77           | -6.871               |
| 2,6-dimethylpyridine           | -6.04           | -6.0984              |
| Ethanol                        | -5.04           | -5.0844              |
| Ethylbenzene                   | -5.67           | -5.7657              |
| Fluorobenzene                  | -4.6            | -4.6416              |
| 1-heptanol                     | -7.84           | -7.7992              |
| 1-hexanol                      | -7.05           | -7.2207              |
| 2-Methyl-1-Propanol            | -5.79           | -5.8164              |
| 2-Propanol                     | -4.82           | -5.0727              |
| Isopropylbenzene               | -6.04           | -6.1356              |
| m-cresol                       | -8.4            | -8.3327              |
| Mesitylene                     | -6.4            | -6.5034              |
| dichloromethane                | -3.8            | -3.8348              |
| Nitrobenzene                   | -7.94           | -8.0294              |
| Nitroethane                    | -5.53           | -5.649               |
| Nitromethane                   | -5.38           | -5.4908              |
In addition to the solvation free energy, the saturation vapor pressure \( P_{\text{sat}} \) as another extensively required thermodynamics quantity in many industrial and scientific applications can also be accurately computed via the thermodynamically effective surfaces via the Clausius–Clapeyron relation, which for ideal gases can be written as:

\[
\ln \left( \frac{P_{\text{sat}}}{P_{\text{atm}}} \right) = \int_{T_{\text{NBP}}}^{T} \frac{\Delta H_{\text{vap}}}{R T^2} \, dT.
\]

The saturation vapor pressure of non-ideal gases can also be determined the same way via evaluating the temperature dependence of gas phase molar volumes using appropriate equation of states. Nevertheless, in the present study and only for a proof of concept, we study predictability of saturation vapor pressures of ideal gases (pressures up to 2.5 atm) via Eq. (20). A comparison of predicted and experimentally determined saturation vapor pressures for a number of most widely used solvents is depicted in figure 4. The excellent agreement between the theoretically evaluated and experimental data depicted in figure 4 implies the robustness of our proposed thermodynamically effective surfaces.
Figure 4- Comparison of theoretically predicted saturation vapor pressures obtained via Eq. (11) and thermodynamically effective surfaces (denoted by Eq(11)-TE) and experimentally determined data
To sum up, in the context of the present study, we could theoretically derive a relationship which describes dependency between vaporization enthalpy, molecular surfaces and surface tension. We could demonstrate that the proposed dependency between solution thermodynamics and molecular surfaces are remarkably more reliable and rigorous compared to the other models empirically proposed for the same purpose within the last century.

Through our newly derived theoretical approach, we proposed thermodynamically effective surfaces. We demonstrated that the “thermodynamically effective” surfaces are only slightly different than empirically proposed vdW surfaces. However, this slight deviation yields a substantial improvement in predictability of multiple thermodynamics quantities. As a result, we propose the thermodynamically effective surfaces as a reliable alternative for the currently defined molecular surfaces, especially for studying the condensed phase thermodynamics.

5- References:

1 Zimmerman, M. I. et al. SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome. *Nature Chemistry*, 1-9 (2021).

2 Portelli, S., Phelan, J. E., Ascher, D. B., Clark, T. G. & Furnham, N. Understanding molecular consequences of putative drug resistant mutations in Mycobacterium tuberculosis. *Scientific reports* 8, 1-12 (2018).

3 Christoffel, F. et al. Design and evolution of chimeric streptavidin for protein-enabled dual gold catalysis. *Nature Catalysis*, doi:10.1038/s41929-021-00651-9 (2021).

4 Lee, J., Chang, I. & Yu, W. Atomic insights into the effects of pathological mutants through the disruption of hydrophobic core in the prion protein. *Scientific reports* 9, 1-14 (2019).

5 Mishra, A., Ranganathan, S., Jayaram, B. & Sattar, A. Role of solvent accessibility for aggregation-prone patches in protein folding. *Scientific reports* 8, 1-13 (2018).

6 Gristick, H. B. et al. Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nature structural & molecular biology* 23, 906-915 (2016).
Waskiewicz, E., Vasiliou, M., Corcoles-Saez, I. & Cha, R. S. Cancer genome datamining and functional genetic analysis implicate mechanisms of ATM/ATR dysfunction underpinning carcinogenesis. *Communications biology* **4**, 1-11 (2021).

Van der Verren, S. E. *et al.* A dual-constriction biological nanopore resolves homonucleotide sequences with high fidelity. *Nature biotechnology* **38**, 1415-1420 (2020).

Gainza, P. *et al.* Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods* **17**, 184-192 (2020).

Amin, A. & Bernd, H. *Improved Prediction of Solvation Free Energies by Machine-Learning Polarizable Continuum Solvation Model.* (2021).

Kim, D. H. *et al.* Selective electrochemical reduction of nitric oxide to hydroxylamine by atomically dispersed iron catalyst. *Nature communications* **12**, 1-11 (2021).

Guo, L. *et al.* Nickel-catalyzed Suzuki–Miyaura cross-couplings of aldehydes. *Nature communications* **10**, 1-6 (2019).

Liu, Y.-M. *et al.* Nanographenes as electron-deficient cores of donor-acceptor systems. *Nature communications* **9**, 1-7 (2018).

Rousseau, R., Glezakou, V.-A. & Selloni, A. Theoretical insights into the surface physics and chemistry of redox-active oxides. *Nature Reviews Materials* **5**, 460-475 (2020).

Biasin, E. *et al.* Direct observation of coherent femtosecond solvent reorganization coupled to intramolecular electron transfer. *Nature Chemistry* **13**, 343-349 (2021).

Han, Y.-X. *et al.* Biomimetically inspired asymmetric total synthesis of (+)-19-dehydroxyl arisandilactone A. *Nature communications* **8**, 1-13 (2017).

Kim, T. Y. *et al.* Platycodin D, a natural component of Platycodon grandiflorum, prevents both lysosome-and TMPRSS2-driven SARS-CoV-2 infection by hindering membrane fusion. *Experimental & molecular medicine* **53**, 956-972 (2021).

Eötvös, R. Ueber den Zusammenhang der Oberflächenspannung der Flüssigkeiten mit ihrem Molecularvolumen. *Annalen der Physik* **263**, 448-459 (1886).

Sisskind, B. & Kasarnowsky, I. Untersuchungen über die Löslichkeit der Gase. 2. Mitteilung. Löslichkeit des Argons. *Zeitschrift für anorganische und allgemeine Chemie* **214**, 385-395 (1933).

Uhlig, H. The solubilities of gases and surface tension. *Journal of Physical Chemistry* **41**, 1215-1226 (1937).

Eley, D. On the solubility of gases. Part I.—The inert gases in water. *Transactions of the Faraday Society* **35**, 1281-1293 (1939).
Clever, H. L., Battino, R., Saylor, J. & Gross, P. The solubility of helium, neon, argon and krypton in some hydrocarbon solvents. *The Journal of Physical Chemistry* **61**, 1078-1082 (1957).

Bondi, A. v. van der Waals volumes and radii. *The Journal of physical chemistry* **68**, 441-451 (1964).

Gallicchio, E., Kubo, M. & Levy, R. M. Enthalpy–entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation. *The Journal of Physical Chemistry B* **104**, 6271-6285 (2000).

Levy, R. M., Zhang, L. Y., Gallicchio, E. & Felts, A. K. On the nonpolar hydration free energy of proteins: surface area and continuum solvent models for the solute–solvent interaction energy. *Journal of the American Chemical Society* **125**, 9523-9530 (2003).

Frisch, M. *et al.* Gaussian 16. *Revision A* **3** (2016).

Rappé, A. K., Casewit, C. J., Colwell, K., Goddard III, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American chemical society* **114**, 10024-10035 (1992).

Alibakhshi, A. Enthalpy of vaporization, its temperature dependence and correlation with surface tension: a theoretical approach. *Fluid Phase Equilibria* **432**, 62-69 (2017).

Wilding, W. V., Rowley, R. L. & Oscarson, J. L. DIPPR® Project 801 evaluated process design data. *Fluid phase equilibria* **150**, 413-420 (1998).

Adam, N. (London, 1941).

Frisch, M. *et al.* (Gaussian, Inc. Wallingford, CT, 2016).

Reynolds, J. A., Gilbert, D. B. & Tanford, C. Empirical correlation between hydrophobic free energy and aqueous cavity surface area. *Proceedings of the National Academy of Sciences* **71**, 2925-2927 (1974).

Hermann, R. B. Use of solvent cavity area and number of packed solvent molecules around a solute in regard to hydrocarbon solubilities and hydrophobic interactions. *Proceedings of the National Academy of Sciences* **74**, 4144-4145 (1977).

Zacharias, M. Continuum solvent modeling of nonpolar solvation: Improvement by separating surface area dependent cavity and dispersion contributions. *The Journal of Physical Chemistry A* **107**, 3000-3004 (2003).

Zaitsau, D. H. *et al.* Experimental vapor pressures of 1-alkyl-3-methylimidazolium bis(trifluoromethylsulfonyl) imides and a correlation scheme for estimation of vaporization enthalpies of ionic liquids. *The Journal of Physical Chemistry A* **110**, 7303-7306 (2006).
36. Akkermans, R. L. Solvation Free Energy of Regular and Azeotropic Molecular Mixtures. *The Journal of Physical Chemistry B* **121**, 1675-1683 (2017).

37. Marenich, A. V. *et al.* Minnesota solvation database. *Minnesota Solvation Database version 20* (2012).
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarydata.xlsx