Dimensional analysis as a tool to predict, identify and decipher the relative magnitude effects of known and unknown physicochemical properties that influence the solubilizing effectiveness of solubility modifiers and/or complexing agents.

Shireesh P. Apte*
Editorial Board, Journal of Excipients and Food Chemicals

Editorial

KEY WORDS: Excipient discovery, physico-chemical properties, solubilizing agent, association constant, dimensional analysis

"Experience may teach us what is, but never that it cannot be otherwise."

Immanuel Kant, Critique of Pure Reason

Empirical science has racked up a vast body of heuristic relationships. Theorists find scientific and logical mechanisms associated with these relationships and attempt to derive them from; or trace them to; 'first principles'. Either way, scientific reasoning has historically proceeded, and continues to proceed, a posteriori, i.e. using prior experience. Notwithstanding transcendental idealism, to aspire to Kantian synthetic a priori knowledge, knowledge without experience, requires the attributes of universality and 'mind-independent reality'. In this editorial, I ask the question: can dimensional analysis combined with brute computing power allow us to 'manufacture' such knowledge without any a priori understanding of 'first principles'?

Good mathematical models describe and make temporal and spatial predictions about natural phenomena. Dimensional analysis dictates that an equation is true only if the units on the left hand side are equal to the units on the right hand side. Therefore, would it not be possible to take any number of physicochemical properties of matter and use iterative artificial intelligence (AI) algorithms to manipulate them until the units (on one side of the equation) matched those of the quantity that is being predicted (on the other side of the equation)? It may very well turn out that the physicochemical properties being manipulated (seemingly) have no relation to the prediction quantity; however, that may very well be one of the reasons to attempt such an exercise; *viz.* to find correlations between seemingly unrelated phenomena or physicochemical properties of matter that have not yet been explored empirically. In so doing, it may be necessary to assign physical significance to any dimensions that are left-over or do not match; so that the equation is made dimensionally true.

Solubility enhancement is an important area of research in pharmaceutical development to increase the solubility of active pharmaceutical ingredients (API). The association constant between an API and a solubilizing agent may be considered to be an indication of the latter's effectiveness in solubilizing the former.

*Corresponding address: 5204 Coventry Court, Colleyville, TX 76034, Tel: 817 501 2984, E-mail: shireeshpapte@msn.com

The association constant is defined in Equation 1 as:

$$K_a = \frac{DS}{D \times S} \qquad \text{Eq. 1}$$

where, D is the molar concentration of drug, S is the molar concentration of the solubilizing agent and DS is the molar concentration of the drug-solubilizer complex in solution. The units of K_a are m^3/mol. Using dimensional analysis alone allows the construction of two Equations (2 and 3) that relate the physicochemical properties of matter to the prediction variable, K_a.

$$K_a = \frac{\text{Molar susceptibility x molar dipole moment x redox potential}}{\text{Enthalpy of vaporization x}[va\text{HLB}]} \qquad \text{Eq. 2}$$

The units of molar susceptibility are m^3/mol, those of molar dipole moment are C.m/mol, of redox potential are V or J/C and for enthalpy of vaporization are J/mol. The *va*HLB may be construed as a volume-area corrected hydrophilic lipophilic balance and is formally defined as the ratio of the volume of the hydrophobic groups of the solubilizing molecule in water in m^3 to the area of the hydrophilic groups in water in m^2. It can be seen that solving for the units of the above equation leads to K_a, the association constant between the drug and the solubilizing agent, being measured in terms of m^3/mol; which is dimensionally true. This does not necessarily verify the qualitative proportionalities between these properties and K_a, however the equation does define a possible AI iteration using dimensionally correct physicochemical properties. There may be many such possible iterations utilizing other physicochemical properties. Equation 3 (see below) represents another such dimensionally correct iteration.

$$K_a = \frac{\text{Molar attentuation constant x ebullioscopic constant x gas constant}}{\text{Lattice energy x average linear molar mass}} \qquad \text{Eq. 3}$$

In order to make Equation 2 dimensionally true, the HLB (or any other physicochemical property) needed to be expressed in the units of m^3/m^2 (see Figure 1). This deviates significantly from the dimensionless Griffin or Davies numbers. The HLB could have the units mentioned if the volume of the hydrophobic part of the solubilizing agent is incorporated along with the surface area of the hydrophilic part of the solubilizing agent in water. Since the term appears in the denominator of the equation, it is apparent that the solubilizing efficacy increases as the former decreases and the latter increases; in conformity with the current HLB paradigm. However, note that this newly defined property *va*HLB does not appear anywhere in the scientific literature. Also note that there exists no property in the equation that is solute-dependent. In other words, according to Equation 2, maximizing the solubilizing molecule's molar susceptibility, molar dipole moment and redox potential, while minimizing its *va*HLB and enthalpy of vaporization will produce a solubilizing agent with maximum solubilizing capacity for any solute.

It is apparent that the redox potential is the standard reduction potential if the drug is an electron donor so that the ability of the solubilizing agent to accept electrons is directly proportional to its association constant with the drug. Conversely, the redox potential is the standard oxidation potential if the drug is an electron acceptor so that the ability of the solubilizing agent to donate electrons is directly proportional to its association constant with the drug. Its surrogate is the H-donor and H-acceptor sites on the molecule or hydrogen-bonding.

The dipole moment of the solubilizing agent represents

Dimensional forcing of equation 2

$$\frac{m^3}{mol} = \frac{\frac{m^3}{mol} \times \frac{C.m}{mol} \times \frac{J}{C}}{\frac{J}{mol} \times \left[\frac{m^3}{m^2}\right]}$$

Left-over dimensions assigned to a new physical property: vaHLB

Dimensional forcing of equation 3

$$\frac{m^3}{mol} = \frac{\frac{m^3}{mol \,.\, m} \times \frac{K \,.\, g}{mol} \times \frac{J}{mol \,.\, K}}{\frac{J}{mol} \times \left[\frac{g}{mol \,.\, m}\right]}$$

Left-over dimensions assigned to a new physical property: average linear molar mass

Figure 1 Dimensional analysis of equations

the magnitude of the charge separation within the molecule, is hence related to the redox potential and is proportional to its drug association constant.

The enthalpy of vaporization is directly related to the cohesive energy density of the solubilizing agent and is a measure of the intermolecular attractive forces between the solubilizing agent molecules. It represents the energy needed to associate with the drug and hence is inversely proportional to the drug association constant. It surrogate is the solubility parameter.

The *va*HLB is a volume-surface area weighted measure of the hydrophilic-lipophilic balance or HLB, that is often used as a surrogate for solubilizing ability. Since the hydrophobic portion of the molecule associates with itself, a volume is a more germane representation of its shape either in water or when associated with a drug. Whereas the hydrophilic portion of the molecule associates with the hydrophilic solvent, hence a surface-area weighted parameter is a more suitable measure of its shape in water or when associated with a drug.

The association constant can be described as yet another equation (Equation 3) that combines physicochemical properties, some of which are not intuitively assignable to molecular association.

The units of the molar attenuation constant; or the molar absorptivity are L/mol.m, those of the ebullioscopic constant are K.g/mol, those for the gas constant are J/mol.K, those for the lattice energy are J/mol and those for the average linear molar mass are g/mol.m.

In order to make Equation 3 dimensionally true, a new physicochemical property called the average linear molar mass is constructed (see Figure 1). It is formally defined as the molar mass per unit length when measured along the larger of the axial or equatorial planes. It follows that the smaller this quantity; i.e., the longer the molecule in relation to its molar mass, the greater its solubilizing effectiveness and the larger the association constant. Conversely, the larger this quantity; i.e., the larger the molar mass of the molecule in relation to its length, the lesser its solubilizing effectiveness and the lesser the association constant.

The magnitude of the molar attenuation constant at the wavelength of maximum absorption effectively is a measure of electron de-localization; which, in turn, is directly proportional to its hydrogen-bonding ability. The larger the molar attenuation constant, the greater the effective molar area across which electrons are de-localized and the greater the association with the API via H-bonds. Examination of this property leads to a hitherto unexplored method to increase the solubility of a drug using a solubilizing agent whose wavelength of maximum absorption is greater than that of the API. The K_a can be increased by irradiating a stock-solution of the drug and solubilizing agent with this wavelength of EM radiation. If the association is exothermic with a significant enthalpic contribution, the complex will not dissociate post-irradiation.

The ebullioscopic constant is the only property in this equation that is dependent on the solute or drug. The greater the boiling point elevation for a given molal concentration of the solubilizing agent, the larger the deviation from Raoult's law, the greater the magnitude of intermolecular interactions and the greater the effectiveness of the solubilizing agent in increasing the solubility of the API.

Note that in Equations 2 and 3, the physico-chemical quantities are not necessarily independent of each other in terms of the values they can take, although there is some leeway during which one can be made to change without significantly changing the other(s). Quantitative structure activity relationships (QSAR) will still apply especially as these relate to macro-scale physicochemical properties. Such non-stochastic variables hence lead to an optimization problem with regard to the objective: viz. to design an effective solubilizing agent; albeit, this time, using physicochemical properties that may have been overlooked, may as yet be unknown (dimensional adjusters), or may have been deemed to be unrelated to the prediction variable.

It is interesting to note that, using only dimensional analysis as the arbiter, the author has constructed two equations without recourse to algorithmic processing,

both of which have not explicitly so appeared in the scientific literature with regard to solubility enhancement. Furthermore, both the equations reveal hitherto unknown or un-noticed properties of molecules, dimensional adjusters, that have been assigned the physicochemical properties of the *va*HLB and the average linear molar mass, that may be important to solubility enhancement. It is important to re-iterate that such 'mathematical manipulation' is neither without prior precedent, nor without considerable scientific significance, indeed, the Planck's constant; a fundamental constant of the universe, was 'discovered' as a purely mathematical abstraction or representation that built on the Rayleigh-Jeans Equation for preventing the ultraviolet catastrophe. It was assigned and ascended to its quantum physical significance either in tandem or post mathematical dimensional adjustment – not earlier.

This approach becomes even more suited to prediction when AI is used to create a 'training set' for the algorithm. For example, a mere knowledge of whether or not two or more physicochemical properties are directly or inversely related allows the algorithm to significantly improve the equation output. The algorithm then searches through a list of properties to find one that satisfies the 'left-over' dimension units or has human intelligence take over to assign physical reality to left-over dimensional mathematical abstractions.

There is no limit to the equations or predictors that can be subjected to this paradigm. It is projected to work as well in designing superior light harvesting or energy dense molecules and catalysts for fuel cells as it does with designing excipients that increase the solubility of APIs' or decrease their biological off-target effects. It is surprising that no concerted scientific effort yet exists that uses this idea. Perhaps such *a priori* knowledge will be pursued *a posteriori* to the publication of this editorial.