IOSR Journal of Applied Chemistry (IOSR-JAC)
e-ISSN: 2278-5736.Volume 12, Issue 12 Ser. I (December. 2019), PP 43-50
www.iosrjournals.org

How to Connect a Chemical Structure to a Property: History and new developments.

Jean-Christophe Garrigues[1]
[1](*IMRCP Laboratory, CNRS, Toulouse University, France)*

Abstract: *Organic chemistry has based its knowledge on the study of structures of chemical elements and their reactivity. The classification of the various elements, atoms and molecules has evolved over time, but it is remarkable that these representations have incorporated such subjective quality concepts in connection with the perceptions that the manipulators of these substances could identify. The electronic properties of atoms connected to the nuclear behavior of these elements indeed provide each element's periodic classification. Thanks to this classification model, it was possible to envisage the existence of elements as yet not isolated. This theorization of the structure of matter by Schrödinger has opened the way to purely theoretical modeling which is in itself the basis of the in silico chemistry. But how do we link these structures to physico-chemical or biological properties? This question, which has been perplexing and uniting chemists, physicists, biologists, and mathematicians alike for more than 150 years, has given rise to molecular-level knowledge of recognition mechanisms for complementary structures and the design of new, innovative targeted drugs. With the development of quantitative structure property relationships (QSAR) models, chemists are closer than ever to understanding the paradigm of the universality of such structure models. How might one ensure that only one molecular structure is responsible for one particular property? Answering such a question could allow further development in the representation of molecular structures.*

Keywords*: QSAR, molecular descriptors, structure, property, optimization, modelling.*

Date of Submission:24-12-2019 Date of Acceptance:07-01-2020

I. Introduction

In an article read at the Royal Society of Edinburgh, in 1868, Crum-Brown and Fraser wrote, "*It is obvious that there must exist a relation between the chemical constitution and the physiological action of a substance, but as yet scarcely any attempts have been made to discover what this relation is.*" [1]. This article was the starting point of a long series of studies aimed at validating the idea that a relationship exists between a substance or chemical combination of ingredients and the biological or physicochemical property it produces. The intention to give meaning to an observation, to understand nature, to imagine a rule that would bring some order to the incredible diversity of the world that surrounds us, can be attributed to the first Greek philosophers. By seeking the "principles" of things, they worked to group substances into categories and to reduce complex phenomena down to a limited number of principles, taken in the sense of what comes "first", what is at the origin of a phenomenon.

In the 4th century Before the Common Era (BCE), Aristotle reflected on how to reduce all phenomena to 4 elements: earth, water, air and fire, which would make it possible to explain certain physicochemical properties like the fusion of metals, combustion and evaporation [2]. In this system, Aristotle attributed a "quality" to an element, which was shared with the element that followed it. This concept of element, making it possible to simplify and qualify a substance, was then deepened by Lavoisier, who used laboratory experiments to specify the quality of the compounds studied. In his "Traité élémentaire de chimie" [3], Lavoisier pointed out the fundamental importance he attached to experiments and to precise measurements of the masses of compounds that combined together, in order to define the oxidation properties of the substances. It was the study of how gaseous substances combined with one another that allowed Proust and Gay-Lussac to define simple, uniform laws explaining the relationship existing between gases. This experimental work formed the starting point of Dalton's " atomic theory" [4]. It is important to note that this atomic hypothesis was exclusively speculative. It associated quantitative simplicity drawn from the study of gases with the qualitative simplicity of an indestructible particle: the atom. This hypothesis was to give birth to the concept of chemical elements, for each of which there was an atom, characterized by its weight, and no longer by its shape as postulated by Democritus. This geometrical approach to the organization of atoms within matter allowed Dalton to explain why an element could exist in various physical forms, with the example of pure carbon, existing in the form of graphite or diamond. This atomic model was the first to provide a formal explanation for the multiplicity of the observable forms of matter by connecting them to the unique nature of a few, chemically identifiable, elements.

This theory was the first to put forward a relative scale for atomic masses, taking the mass of hydrogen as its reference. On the other hand, the atomic theory, by allowing a confusion between atoms and molecules, elements and pure compounds, did not enable all the observations to be conceptualized, in particular for water, and was debated until the beginning of the 20th century. Avogadro, resuming work of Gay-Lussac and Dalton, shed new light on the concept of the molecule. In 1811, he published "*Essai d'une manière de déterminer les masses relatives des molécules élémentaires des corps, et les proportions selon lesquelles elles entrent dans ces combinaisons* " [5], an article in which he specified the concept of the molecule as a true assembly of atoms. Confirming results obtained by Gay-Lussac and working on the volume occupied by gases, he proved that, under the same conditions of temperature and pressure, the same number of molecules, in a gaseous state, occupied the same volume. Thus, he refined the study of the relations between a property, the atomic mass, the number of moles or molar volume and the molecular structure under study.

The atomic theory was refined by Laurent [6] who wrote the first chemical equations, starting from a systematic nomenclature allowing him to represent molecules and how they combined in chemical reactions. This theory was reinforced by Clausius, who explained changes of state by the atomic theory, and Cannizzaro, who published the first list of atomic weight values for elements. This classification was then supplemented and theorized by Mendeleev, who introduced the concept of periodicity of the elements, by gathering together the chemical elements that had the same reactive property with hydrogen. This reactive property and this classification of the elements gave birth to the valence concept. The position of the elements within the periodic classification was formalized in an indisputable way by Bohr in 1913, who connected the periodicity of the elements to the atomic number and proposed a precise model for the structure of the atom. A study of the historical and philosophical development of the concept of element and the epistemological questions which were raised during the development of the theories allowing compounds not yet discovered to be predicted and known ones to be classified, makes it possible to understand the evolution of the models that are used by chemists. Paneth, recalls the conceptual birth of the components of matter from the notion of substance to the definition of the chemical element, by connecting these concepts with the evolution of how chemists perceive the "qualities" of the elements [7]. This perception has evolved from partial subjectivity (wet, dry, hot, cold,…) to the quantification of the reactivity with hydrogen (valence), or other compounds (masses and molar volume). It is a key reflection of the evolution of the epistemological status of the chemical concept of element. By integrating a particular property or responsiveness to an element, it is possible to answer the question of the persistence of the element in the substance consisting of the association of elements. A first connection may be established between the structure of a particular element and the persistence of the properties of that element at the macroscopic level. This analysis of the concept of the element can also suggest a model for each of the elementary properties in the final property of the substance.

The model defined by Bohr provided a description of spectroscopic phenomena, highlighted by Planck and Rutherford, showing that atoms absorb or emit only certain wavelengths of light. Thanks to the atomic model, Bohr envisaged the existence of elements so far unknown in the periodic table, which were later to be discovered. A few years earlier, Lewis had also worked on the representation of atomic structure, to explain the stability or the particular reactivity of certain chemical elements. He imagined the octet rule to explain this stability and, as early as 1902, conceived a theory for the covalent bond that allowed 2 atoms "to share" 2 or more electrons, to form simple or multiple connections. In 1916, he published an article "*The atom and the molecule*" [8], in which he proposed a diagrammatic representation of atoms and molecular structures that allowed chemical properties related to the combination of elements to be connected to an atomic organization within a molecule.

By integrating the postulate established by De Broglie in 1924 on waves - corpuscle dualism, Schrödinger proposed a final model in 1926, in which electrons are represented by probability clouds representing such presence. According to the quantum state of the electron, these clouds can take various forms. This model, still in use today, explains the stability of an atom or its reactivity as well as spectroscopic effects and the shapes of molecules.

By theorizing the electronic and nuclear behavior of molecules, Schrödinger opened the way to theoretical chemistry, which allows the physicochemical properties or space conformations of molecules to be described or envisaged by solving the mathematical equations of quantum mechanics [9]. This purely theoretical model is able to describe an atomic or molecular structure and to determine its physicochemical properties, *ab initio*. However, the complexity of the calculations necessary for the development of the molecular model forces chemists to confront the theoretical results using experimentation taking into account the need to classify molecular structures according to measured properties. Such an approach continues to be essential, even after the discovery of these mathematical models, and this has caused the classification tools to be improved.

II. How Connect A Property To A Structure?

The fundamental aspect that gave birth to the atomic and molecular model corresponds to a classification of structures according to properties like atomic mass or valence. It was by studying the continuous properties of elements that Gay-Lussac, Dalton or Mendeleev defined universal classification methods to order their observations. The Knowledge of the properties of each element allows chemists to understand the full range of mechanisms and reactions that underlie organic chemistry and to define how structural features can impact the energy of the intermediates formed in these processes [10]. Based on this knowledge, it is possible to segment the complexity of the chemical mechanisms studied to identify the necessity of chemical integrity through the simplest structure needed to produce a mechanism.

This question of the organization of chemical substances is a long-standing preoccupation for human beings and is found in a variety of activities. For example, in the utilization of chemical, mineral or organic compounds, the need to study the effects produced by a substance has been recognized for several millennia. Animal venom and poisons from plants were used by Man as weapons or drugs 5000 years BCE, as archaeological discoveries have shown [11]. Around 1500 BCE, the Ebers Papyrus described about 800 recipes for remedies and, towards 1500 CE, Paracelsus determined that the toxicity of a poison was related to the presence of specific chemical compounds in amounts that made them harmful. In 1813, Orfila developed modern toxicology by connecting the chemical composition of poisons with their biological properties. In 1863, Cros made a first qualitative study by noting that the toxicity of alcohols decreased with their water solubility. But it was in 1868, as mentioned in the introduction, that the first mathematical formulation of a quantitative structure-activity relationship was presented by Crum-Brown and Fraser. By studying various alkaloids, they observed that the alkylation of basic nitrogen atoms, giving quaternary ammoniums, produced very varied biological effects. They postulated that the physiological activity ϕ depended on the chemical structure C as proposed in equation 1:

$$\phi = f(C)$$

Since his chemistry and medical studies, Crum-Brown had been seeking to express chemistry through mathematics because, believing that the world had been created according to a unified plan, he found it fundamental to solve all the scientific observations with mathematical considerations.

To represent the structures of the molecules he studied, he designed a model placing the symbol of the element in the center of a circle, connected to the element that he supposed close to it by valence features, a model that enabled isomers to be represented.

In 1893, Richet showed that the toxicity of organic compounds $\Delta\phi$ was inversely proportional to their water solubility ΔC, through equation 2:

$$\Delta\phi = f(\Delta C)$$

It is important to note that, even today, only equation 2 is accessible to studies allowing a structure to be connected to a property as only differences in biological activities can be quantitatively correlated with structural modifications which it is necessary to analyze and model. This determination of a physicochemical property generated by a molecular structure and connected to a biological activity was carried out by Meyer and Overton in about 1900. They observed a linear relation between lipophilicity, expressed by the water/olive oil partition coefficient, and narcotic activity [12]. The growing idea of an additive contribution of the substituents to biological activity was highlighted by Führer in 1907, who showed that the narcotic activity of a homologous series increased following a geometric progression: 1; 3^1; 3^2; 3^3; etc. The importance of the lipophilicity parameter in property-structure studies of relationships was confirmed by many other works, which provided opportunities to refine the measurement of this physicochemical parameter and to connect it with many nonspecific biological activities. In 1939 Ferguson gave a thermodynamic interpretation explaining the threshold values of optimal lipophilicity, beyond which the activity decreased [13].

To allow quantitative structure-activity relationship (QSAR) studies to be made, it is necessary to develop indices to describe the physicochemical properties and the spatial conformation of molecules.

The foundations of QSAR studies were laid by the joint work of Hammett, Taft, Hanch, Fujita, Free and Wilson. Hammett sought to model the variations of reaction rates for aromatic compounds, substituted or not, with equation 3 which he proposed in 1937:

$$log\ k_{R\text{-}X} - log\ k_{R\text{-}H} = \rho\sigma$$

In this equation, the influence of the substituent R is described by an electronic parameter σ, defined compared to hydrogen, and a parameter ρ which describes the sensitivity of the reaction to the effects of substituents.

In 1952 Taft, taking the Hammett equation as a basis, proposed including steric, *Es*, and polar, σ^*,effects in the development of linear relations between the equilibrium constants of reactions and the influence of the substituents [14]. In 1962, Hansch published a QSAR study for plant growth regulators connected to the Hammett parameter and a new hydrophobicity parameter, which he introduced from the measurement of the octanol/water partition coefficient [15]. For this, he measured the log P coefficients of a chemical series and determined the parameter π characteristic of a substituent (*x*) relative to a lead molecule (*L*), following equation 4:

$$\pi = \log P_x - \log P_L.$$

By integrating the contributions of Hammett and Taft, Hansch and Fujita set out the starting point of the first QSAR study in 1964, "*ρ-σ-π Analysis. A method for the correlation of biological activity and chemical structure*" [16]. The first work they carried out was to determine the hydrophobic contribution of characteristic substituents of a chemical series, by comparing the partition coefficient of the various derivatives in reference to a non-substituted standard (*H*) (Equation 5):

$$\pi_x = \log P_x - \log P_H$$

By combining the hydrophobic with the electronic effect, represented by the Hammett constant, they connected the biological activity of growth factors with their molecular structure through equation 6:

$$\log \frac{1}{C} = a\pi + b\sigma + ck,$$

in which the biological activity is represented by the logarithm of the reverse of the molar concentration producing a given response.

A few hundreds of studies later, and in order to take account of observations showing the existence of a parabolic relationship with an optimal value for the lipophilic parameter, Hansch and Fujita proposed the following general relation (Equation 7):

$$\log \frac{1}{C} = a\,(\log P)^2 + b \log P + c\sigma + k$$

in which the hydrophobic parameter of a substituent is now represented by the total lipophilic contribution of the molecule, described by the partition coefficient measured in experiments between water and octanol.

In the same year, Free and Wilson presented a descriptive method allowing a structure to be connected to a property, in which no molecular descriptor for physicochemical properties (π,σ…) was used: "*A mathematical contribution to structure activity studies* " [17]. The study quantified the contribution (a_{ij}) of a substituent (X_i) to the final activity, in reference to a lead molecule with activity μ and took the additivity of the contributions into account in equation 8:

$$\log \frac{1}{C} = \Sigma\, a_{ij} + \mu.$$

These two methods still constitute two frames of reference for connecting a property to a chemical structure in the analysis of a homogeneous series. They enable us to model the intuitive concept that postulates that similar structures have similar properties. It was in 1919 that Langmuir, studying the electronic distribution of various atoms within a variety of molecules, noticed that certain elements or ions possessed similarities in their physical and chemical properties if they had the same number and the same distribution of electrons [18]. He named this similarity phenomenon "Isosterism". Later, Grimm extended this concept of isosterism to groups of atoms having the same number of valence electrons [19]. These groups are interchangeable, by simply modifying the number of hydrogen atoms present on the substituents. This approach was used again by Hansch in defining the concept of bioisosterism, postulating that some substituents are bioequivalent. This concept opened the way to the first targeted synthesis of active antibacterial compounds of the sulfamide family, by substituting para-amino-sulfonic acid for para-amino-benzoic acid. If it is possible to modulate a molecular

structure while preserving the biological properties of the lead molecule, it is therefore also possible to optimize the activity by substituting certain positions, but which ?

To answer this question, it was necessary to define and study the concept of the chemical receptor. The idea of chemical compounds interacting with biological substances having a specific affinity was born from work by Langley in about 1878 [20]. He showed that pilocarpine present in plant extracts could decrease the heart rate in animals and that the effect was reversible by injection of another alkaloid: atropine. He concluded that the 2 compounds formed a complex with substances present in the tissues, the drug reacting with the cells to activate (agonist effect) or to inhibit (antagonist effect) a function. He showed the existence of "receptive substances", present in the muscle cells, causing contraction in the presence of a chemical stimulant. The stereo-specificity of the interactions was highlighted by Fischer in 1894, who worked on the anomeric cleavage reactions of 2 enzymes: invertase and emulsine. He developed the lock-key concept, in 1894, to explain the specific action of an enzyme on a substrate [21]. It was Ehrlich who defined the concept of receptor, in 1897, before Langley clarified their the physicochemical characteristics and molecular aspect in 1905. But it was not until the late 1960's that studies were developed for the molecular characterization and identification of the membrane proteins responsible for the recognition and transduction of the message and only in the late 1990's that nucleic acid receptors were identified (aptamers) [22].

III. From structure to molecular parameters used in QSAR studies.

Since the 1990's, when receptor molecules became available, chemists and biologists have developed QSAR studies along two main lines: firstly using X-ray crystallography, which allows the three-dimensional coordinates of many complexes including a ligand and a receptor to be obtained, thus determining the geometrical characteristics of the action site of the receptor. However this "rigid" model is unable to take the allosteric effects of some ligands into account, which involve changes in the receptor conformation. It was this observation that led chemists to consider a deformable or flexible ligand/receptor model. In order to determine the interactions existing between two molecules and to quantify the intensity of the "forces" that stabilized these supramolecular structures, physical chemists developed molecular modeling tools to generate, handle and represent the three-dimensional structure of a molecule, starting from quantum theory, molecular mechanics and molecular dynamics. The common point of all these tools was that they could calculate the energy of a molecule and determine the most stable conformations under given conditions. As these structures were very difficult to handle and compare for molecules having a large number of degrees of freedom, a molecule was reduced to a set of topological descriptors and physicochemical parameters, calculated from the molecular model.
The three types of parameters that were used in the first QSAR studies were the descriptors related to the hydrophobic, electronic and steric interactions.

The lipophilicity parameter.

Historically, it was the lipophilicity parameter, defined in 1872 by Berthelot and Jungfleisch that was first used, by Meyer and Overton, to correlate molecular structures with their narcotic activity. In 1964, Hansch proposed the use of octanol as the lipid model and the logarithm of the partition coefficient between water and octanol (log P) as the index of lipophilicity of a molecule. Today, many methods exist for determining or calculating log P but it is primarily by high performance liquid chromatography (HPLC) that this parameter is experimentally measured [23]. Calculation utilizes various methods related either to fragment additivity as described by Hansch and Fujita [24], or to atomic additivity as described by Viswanadhan [25]. In 1986, a new calculating method: Molecular Lipophilic Potential (MLP) was proposed by Audry. It allows the lipophilic potential of a molecule to be determined in various parts of adjacent space [26].

Electronic parameters.

These parameters are essential for the modeling of the ligand-receptor interactions and are connected to the reactivity of the compounds studied. These parameters vary considerably in their nature and are represented by the Hammett constant [27], the energies of the electron orbitals (HOMO, LUMO), the dipole moment or the pKa of the ionizable functions.

Steric parameters.

The steric effects of a molecule are complex, in particular in the determination of the effects related to the size or the shape of a chemical compound. The first parameter related to steric effects used in QSAR is the Taft constant [28], to which can be added molar refractivity or molecular mass.

Parameters related to graph theory.

Euler developed graph theory as early as 1780 [29]. A graph allows the topology of a molecule to be represented without consideration of the exact geometry that characterizes it. Thus, starting from the atoms and

the connections, which keep them together, a matrix is created to represent the number of "pathways" present in the structure. These matrices are the starting point for many indices like the Wiener number [30], representing the sum of the connection distances, the Balaban index [31] characteristic of the average of the connection distances, or the Randic index [32], 1975), corresponding to the sum of the balanced connections of the molecule.

Tools used to connect the structure to the property.

The limiting stage of a QSAR study is the obtention of precise, reproducible data related to the property studied. This difficulty is reinforced, in studies using biological material, by the variety of the answers obtained, from the study of toxicity related to the mortality of animals, to the analysis of equilibrium constants of proteins. In order to obtain representative data, it is initially advisable to develop a biological model to quantify an effect, related, for example, to the determination of a percentage of growth inhibition in a living organism. The biological data obtained can be equilibrium constants, inhibition constants, affinity constants, percentages of inhibition (IC50%), or lethal amounts (LD50%). From these biological data or physicochemical properties measured experimentally, the QSAR study will consist in identifying a statistical method allowing the molecular descriptors to be connected to the property. The most commonly used technique is Multiple Regression Analysis (MRA) of the type:

$$Yi = b + aXi + Ei .$$

Thus, the majority of QSAR studies connecting biological activity to a parameter characteristic of the molecular structure are of the type:

$$\log 1/C = a \log P + b,$$

showing the importance of hydrophobic interactions.
In many studies using a lipophilicity parameter, linear models do not apply. The representation of the model approaches a parabola and the equation is then represented by a polynomial of degree 2 as described in equation 7.

Use of artificial neural networks in QSAR studies.

When no mathematical model is applicable to connect the molecular structure descriptors to the property, it is possible to use systems of artificial neural networks, such as those described by Mc Culloch and Pitts in 1940 [33]. Their work consisted in modeling the operation of a biological neuron. They were the first to formalize a theoretical operation for networks of simple formal neurons able to fulfill switching, arithmetic and complex symbolic functions. However, it was not until 1957 that Rosenblatt was able to develop the "Perceptron", the first "neuro-computer" based on the model described by Mc Culloch and Pitts and used in the field of pattern recognition [34]. These systems were not used in chemistry until 1988, with a QSAR study described by Hoskins [35]. Currently, there are several algorithms in existence, which modulate the information circulating within the neural network, making it possible to process very varied data, in particular related to the description of molecular structures for QSAR or QSPR studies [36].

IV. Conclusion

Following the thread of history that summarizes the evolution of how the structure of matter has been perceived and how the physicochemical properties connected to a definite elementary architecture have been analyzed raises the question of the process chemists have needed to apply in order to discover laws which appear obvious today.

The scientific methods that Man has developed to understand and represent the structure of matter have evolved since the Greek philosophers pondered the question in the 4th century BC. Significant observation of the matter constituting the world and analysis of the "qualities" connected to characteristic substances led philosophers to imagine a form of reasoning that could organize the observations and classify the elements. The significant and subjective description of matter led the first chemists to seek causes that could explain surprising phenomena. Purely speculative theories gave rise to the concept of element and atom and led chemists to imagine provisional hypotheses, which were verified or corrected on the basis of results of experiments conducted with ever more sensitive, precise and exact measurements. It was indeed experimentation that confirmed the concept of the molecule and furthered studies of the physicochemical properties related to the various chemical compounds. By studying these properties, and in particular the reactivity of elements, chemists succeeded in classifying compounds according to common properties. It then became possible to transform

matter into mathematical data, opening the way to quantitative studies of the relationships existing between a chemical structure and a property. The experimental foundations of QSAR studies, laid by the work of Hansch and Fujita in 1964, improved with the development of analytical and spectroscopic techniques like X-ray crystallography, which allowed the chemist to obtain a 3D image of the detailed structure of a receptor and a drug producing an effect. Today, the handling of these 3D structures has become routine, with software being used to represent the structures. Many active molecules have been imagined before being actually synthesized, one example being Bifenthrine, an insecticide of the Pyrethrinoid family, which was produced starting from a rational 3D QSAR method. The QSAR studies which led to the development of a consistent model representing the structure of matter helped to develop a science of property-structure relationships, with the formalization of observations and experiments (biological model), and data processing (theoretical chemistry) which makes it possible to conceive quantified relations between molecular parameters in the aim of envisaging a property before even having the molecule to study. QSAR Studies have also contributed to our knowledge of the fundamental phenomena related to the mode of action of a drugs. The concept of Hydrophobicity, born from QSAR studies, revolutionized the design of drugs while creating a new industry related to the calculation or experimental determination of this property. Certain erroneous observations helped QSAR analyses to reach maturity, by allowing scientists to refine their models by taking into consideration the possibility that equivalent phenomena and serendipity might lead to false but coherent deductions. Does the concept of structural and property equivalence validate a universal chemical contingency theory? Which would mean that theoretical models linking single structures to unique properties would be only local convergences and that there would be another possible conformation/structure point in an infinite chemical space. This question refers to the methods used by chemists to determine analogies between chemicals [37]. The chemical matter is not continuous, and the elementary structures are modified by thermodynamic parameters governing intermolecular and intra-molecular interactions (temperature, solvent effect …).

Tomorrow, QSAR tools will take the dynamic aspect of the property-structure relation into account through the development of tools to quantify molecular flexibility or allostery. These tools will make it possible to change scale, to pass from the study of molecular structures, of a few Angströms (10^{-10}m), to the supramolecular scale of the order of a nanometer (10^{-9}m). It is this comprehensive study of the phenomena that ensure the cohesion of matter within living organisms which will open the doors to the design of the chemical compounds of the 21st century, with rigorously selected properties, that will allow mankind to survive while respecting the world around us. But, for the moment, chemists are still unable to define a law representing all the properties related to a molecular structure. Even if quantum chemistry makes it possible to calculate certain essential properties, related to the spatial conformation or the reactivity of the molecule, it quickly proves unable to predict a particular biological activity or a retention time on an analytical system. It will become possible to envisage a law formalizing all causalities only from the moment when these causalities, these property-structure relations, have been characterized without any ambiguity related to the experimenter or the experiment. To help to formulate a general law for calculating the properties of matter, it is clear that a comparative analysis of the sum of the molecular descriptors available must be carried out in a systematic way, on a large scale. This analysis would identify the universalizable molecular descriptors having sense for all the molecules listed in the structural databases. From these universalizable descriptors, the chemist will be able to identify the fundamental criteria connected with the properties of matter, and take the first steps towards identifying universalizable causalities, which will direct us towards the definition of a law connecting measurable physical quantities to a purely mathematical structure.

.

References

[1]. Crum-Brown, A., Fraser, T.R., *J. Anat. Physiol.*, vol. 2, pp. 224–242, 1868.
[2]. Viano, C., La matière des choses: le livre IV des Météorologiques d'Aristote et son interprétation par Olympiodore, Librairie Philosophique J. Vrin, Paris, 2006.
[3]. De Lavoisier, A.L., Traité élémentaire de chimie, Cuchet., Paris, 1789.
[4]. Kubbinga, H., L'histoire du concept de "molécule", Springer, Paris, 2001.
[5]. Avogadro, A., *Journal de Physique, de Chimie, d'Histoire naturelle et des Arts*, Paris, pp. 58-76, 1811.
[6]. Laurent, A., Méthode de chimie, Mallet-Bachelier, Paris, 1854.
[7]. Paneth, F.A., *Br. J. Philos. Sci.*, vol. 13, pp. 1-14, 1962.
[8]. Lewis, G.N., *J. Am. Chem. Soc*, vol. 38, pp. 762-785, 1916.
[9]. Schrodinger, E., *Phys. Rev*. vol. 28, pp. 1049-1070, 1926.
[10]. Goodwin, W., *Br. J. Philos. Sci.*, vol. 60, pp. 271-301, 2009.
[11]. Cronin, M.T.D., Livingstone, D. J., *CRC Press*, London, 2004.
[12]. Hintzenstern, U.V., Schwarz, W., Goerig, M., Petermann, H., *Int. Congr. Ser.*, vol. 1242, pp. 609-612, 2002.
[13]. Ferguson, J., *Proc. R. Soc. Lond. B Biol. Sci.*, vol. 127, pp. 387-404, 1939.
[14]. Taft, R.W., *J. Am. Chem. Soc.*, vol. 74, pp. 2729-2732, 1952.
[15]. Hansch, C., Maloney, P.P., Fujita, T., Muir, R.M., *Nature*, vol. 194, pp. 178-180, 1962.
[16]. Hansch, C., Fujita, T., *J. Am. Chem. Soc.*, vol. 86, pp. 1616-1626, 1964.
[17]. Free, S.M., Wilson, J.W., *J. Med. Chem.* vol. 7, pp. 395-399, 1964.
[18]. Langmuir, I., *J. Am. Chem. Soc.*, vol. 41, pp. 868-934, 1919.

[19]. Grim, H.G., Ludwigshafen, A., *Naturwissenschaften*, vol. 17, pp. 535-540, 1929.
[20]. Langley, J., *J. Physiol.*, vol. 1, pp. 96-103, 1878.
[21]. Kubinyi, H., QSAR: Hansch Analysis and Related Approaches, Weinheim - New York, 1993.
[22]. Wang, K. Y., Swaminathan, S., Bolton, P. H., *Biochemistry*, vol. 33, pp. 7517-1527, 1994.
[23]. Lipinski, A., Lombardo, F., Dominy, B.W., Feeney, P.J., *Adv. Drug Deliv. Rev*., vol. 23, pp. 3-25, 1997.
[24]. Fujita, T., Iwasan J., Hanschn C., *J. Am. Chem. Soc.*, vol. 86, pp. 5175-5180, 1964.
[25]. Viswanadhan, V.N., Ghose, A.K., Revankar, G.R., Robins, R.K., *J. Chem. Inf. Comput. Sci.*, vol. 29, pp. 163–172, 1989.
[26]. Audry, E., Dubost, J.P., Colleter, J.C., Dallet, P, *Eur. J. Med. Chem.* vol. 21, pp. 71-72, 1986.
[27]. Hammett, L.P., *Chem. Rev.*, vol. 17, 125-136, 1935.
[28]. Taft, R.W., *J. Am. Chem. Soc.*, vol. 74, pp. 3120-3128, 1952.
[29]. Hansen, P.J., Jurs, P.C., *J. Chem. Educ.*, vol. 65, pp. 574-580, 1988.
[30]. Wiener, H., *J. Am. Chem. Soc*, vol. 69, pp. 2636-2638, 1947.
[31]. Balaban, A.T., *Chem. Phys. Lett.*, vol. 89, pp. 399-404, 1982.
[32]. Randic, M., *J. Am. Chem. Soc.*, vol. 97, pp. 6609-6615, 1975.
[33]. Mc Culloch, W.S., Pitts, W., *Bull. Math Biophys.*, vol. 5, pp. 115-133, 1943.
[34]. Rosenblatt, F., *Psychol. Rev.*, vol. 65, pp. 386-408, 1958.
[35]. Hoskins, J.C., Himmelbau, D. M., *Comput. Chem. Eng*., vol. 12, pp. 881-890, 1988.
[36]. Lafossas, C., Benoit-Marquié, F., Garrigues, J.C., *Talanta*, vol. 198, pp. 550–559, 2019.
[37]. Needham, P., *Br. J. Philos. Sci*., vol. 68, pp. 953-979, 2017.

Jean-Christophe Garrigues. "How to Connect a Chemical Structure to a Property: History and new developments." IOSR Journal of Applied Chemistry (IOSR-JAC) 12.12 (2019): 43-50.