MOLECULAR SIMILARITY IN THE FRAMEWORK OF A BIOISOSTERISM STUDY

ИЗСЛЕДВАНЕ НА МОЛЕКУЛЯРНА СЪВМЕСТИМОСТ В СРЕДА НА БИОИЗОСТЕРИЗЪМ

Desislava Ivanova

IICT – BAS

divanovaacomin@gmail.com

Abstract: *Bioisosterism is a strategy of Medicinal Chemistry for the rational design of new drugs (bioisosterism is also known as quantum similarity). The success of this strategy in developing new substances which are therapeutically attractive has observed a significant growth in distinct therapeutic classes, being amply used by the pharmaceutical industry to discover new analogs of therapeutic innovations commercially attractive, and also as a tool useful in the molecular modification. In this work,the concept of quantum similarity in the signed particle formulation of quantum mechanics is introduced. This concept was suggested very recently and was studied for the particular case n=1. Therefore, we extend the mathematical expression of similarity to the case n>1, with particular attention to the differences brought in this new context.*

Keywords: QUANTUM SIMILARITY, MOLECULAR DESCRIPTOR, BIOISOSTERISM.

1. Introduction

The notion of molecular quantum similarity was proposed in 1980 and, since then, a growing interest has been shown towards this concept [1]. As a matter of fact, nowadays it is utilized for practical purposes in many different fields such as organic, quantum and physical chemistry. It is interesting to note that different definitions of a quantum similarity can be provided depending on the context. For example, organic chemists may define two similar molecules when they have similar reactivity in certain reactions, while quantum chemists may focus on the molecular features, etc. Therefore, it is not surprising that it is difficult, if not impossible, to provide a universal definition of quantum similarity. In spite of that, it is important to have a practical definition even though it may be related to a particular context. For instance, in a pharmaceutical context molecular quantum similarity may provide a systematic way to investigate the relation between the structure of a molecule and its chemical reactivity, thus drastically easing the process of drug design and minimizing the occurrence of side effects.

While the concept of quantum similarity is becoming widespread in many communities and different definitions are being provided, to the best of our knowledge it seems that all directions taken so far are based on the Schrödinger (or standard) formulation of quantum mechanics [2]. However, different formulations exist which may offer a different perspective on the concept of quantum similarity [3–7]. In practice, in order to concretely (mathematically) define the concept of molecular quantum similarity, one has to define a molecular descriptor first. While an infinite set of options is available, for many practical and theoretical reasons, the electron density has become the most common choice [1].

Recently, a new definition of quantum similarity has been proposed based on the concept of quasi-distribution functions [8]. This novel approach exploits the Wigner formulation of quantum mechanics which belongs to the class of phase-space formalisms and has been implemented in the framework of the Wigner Monte Carlo method [8]. On the other hand, a new formulation of quantum mechanics has recently been suggested which is based on the intuitive concept of (signed) classical particles [10]. These particles are contemporaneously provided with a position and a momentum, and therefore this formalism belongs to the class of phase-space formulations of quantum mechanics. In particular, it is able to reconstruct the corresponding Wigner quasi-distribution function of a given quantum system which, in turn, represents the molecular indicator chosen in [8]. Again, one should note that the common molecular indicator utilized in quantum chemistry is represented in the vast majority of studies by the probability density in the real space (also referred to as electronic structure). While our definition of quantum similarity is a drastic departure from the common definition, the reader should also note that it actually corresponds to its generalization.

2. Bioisosterism

Bioisosterism is a strategy of Medicinal Chemistry for the rational design of new drugs (see reference [1] below, bioisosterism is also known as quantum similarity). The success of this strategy in developing new substances which are therapeutically attractive has observed a significant growth in distinct therapeutic classes, being amply used by the pharmaceutical industry to discover new analogs of therapeutic innovations commercially attractive, and also as a tool useful in the molecular modification. A way to achieve reliable bioisosteric analysis is by means of quantum simulation of given molecules. As a matter of fact, without simulation capabilities, this analysis would be limited to experimental observations which cannot always be provided, not to mention the high cost of such experiments. Furthermore, quantum simulations allow researchers to access to an amount of details which is practically impossible nowadays. Indeed, energy levels, electronic structure, molecular orbitals, exchange-correlation energies (just to mention a few of them) cannot be easily measured with the actual laboratory technology.

While quantum simulations for the study of bioisosterism is already available the visualization of the results obtained from these simulations is still a tricky and complex task. Indeed, it is clear that the use of a three-dimensional screen would strongly improve the understanding of bioisosteric molecules. What remains to be clarified is the way visualizations should be done to obtain an intuitive and efficient analysis. Furthermore, one of the main outcome of this investigation would be the development of a know-how that could be utilized by physicists and chemists for this particular kind of chemical/medical investigation. Finally, at least one publication about the developed technique would be published in a peer reviewed international scientific journal. An important aspect of this publication may be represented by the potential of attracting pharmaceutical and medical companies. Several simulation will be ran and a couple of several numerical experiments will be performed in order to validate the new definition. A set of molecules and crystal structures which could be used to benchmark this new framework. One of these could be the Graphene, Silicene and Titanium Dioxide which have an incredible amount of possible applications (Graphene is raising a lot of interest as one of the European H2020 Flagship Projects). The task is to create a three-dimensional visualization of meaningful molecular

orbitals which vibrate in a way which is not visible on a normal screen. This opens the door towards real-time three-dimensional visualization and manipulation of quantum objects.

A way to achieve reliable bioisosteric analysis is by means of quantum simulation of given molecules. As a matter of fact, without simulation capabilities, this analysis would be limited to experimental observations which cannot always be provided, not to mention the high cost of such experiments. Furthermore, quantum simulations allow researchers to access to an amount of details which is practically impossible nowadays. Indeed, energy levels, electronic structure, molecular orbitals, exchange-correlation energies (just to mention a few of them) cannot be easily measured with the actual laboratory technology. While quantum simulations for the study of bioisosterism is already available the visualization of the results obtained from these simulations is still a tricky and complex task. Indeed, it is clear that the use of a three-dimensional screen would strongly improve the understanding of bioisosteric molecules. What remains to be clarified is the way visualizations should be done to obtain an intuitive and efficient analysis. Furthermore, one of the main outcome of this investigation would be the development of a know-how that could be utilized by physicists and chemists for this particular kind of chemical/medical investigation. Finally, at least one publication about the developed technique would be published in a peer reviewed international scientific journal. An important aspect of this publication may be represented by the potential of attracting pharmaceutical and medical companies.

3. Molecular descriptor

In more details, in order to mathematically define the concept of molecular quantum similarity in a particular context, one must first select a significant molecular descriptor. In practical applications, this descriptor will depend on the particular simulated system and on the specific task one wants to achieve. Thus, a universal choice good for any situation is simply not possible.

In practice, one may arrange the most common descriptors in use nowadays in several main classes [10]:

• Feature counts. This is the simplest class of descriptors which consists of counting a certain feature in a molecule. An example is represented by the number of certain atomic species in a molecule.

• Physicochemical parameters. In this class, one finds the descriptors defined in terms of specific physical and/or chemical features. A good example is the partition coefficient (logP).

• Fragment descriptors. These descriptors are based on molecular two-dimensional connection tables and threedimensional descriptions of the structure for substructure searches.

• Topological indices. This represents a wide range of atom based descriptors expressed in terms of indices. Examples are represented by the Wiener index (sometimes referred to as Wiener number).

• Field-based descriptors. Descriptors in this class are based on fields such as the electron density, the electrostatic potential, etc. This represents, at the present time, the frontline in the field of molecular quantum similarity. These descriptors can be subdivided into two sub-categories: measurable and non-measurable.

• Others. Further classes can be given [11].

Although the list of descriptors seems to be enormous, to the best of our knowledge, it seems that they all have one thing in common: they are all based on the Schrödinger formulation of quantum mechanics. No attempt to define a molecular descriptor and, thus, the notion of a molecular quantum similarity in the context of the Wigner formalism seems to have been done. Moreover, the electron density ρ(x) appears, so far, to play a fundamental role in the definition of molecular quantum similarity (note that the density is usually time-independent in this context). The focus is aimed at on defining a molecular descriptor based on the Wigner quasi-distribution function $f_W = f_W(x; p; t)$, and we investigate the possible advantages in doing so (note that the function f_W is time-dependent). In this respect, we define a new field-based molecular descriptor.

In a previous research of the author was found an intuitive definition of quantum similarity which is expressed in terms of distribution functions (in the Wigner formalism) instead of a simple probability density function (as it is done usually in other formulations of quantum mechanics). In particular, a great deal of attention has been put in the visualization of the multi-dimensional distribution functions to understand what quantities should be properly exploited. The new definition defines the similarity of two quantum systems by means of the similarity they achieve in the multi-dimensional phase-space. In practice, our new definition of quantum similarity based on the utilization of quasi-distribution functions reads:

$$d_{AB}^{n}(t) = \int dx \int dp \left| \left(f_W^A(x;p;t) \right)^n - \left(f_W^B(x;p;t) \right)^n \right|^{\frac{1}{n}}$$

with n an integer (taken equal to 1 in [8]). It is trivial to see that, by simulating quantum systems by means of the signed particle formulation, one can reconstruct the quasi-distribution function of two given systems and, therefore, compute their similarity.

As a matter of fact, after running several simulations, it became quickly clear that a definition in the phase-space brings several non-negligible advantages. In fact, the existence of systems which would be considered similar when the old definition is applied, but very different when the new definition is applied (because they are not similar energetically speaking) was discovered. By means of visualization tools, in particular VisIt (LLNL) and paraview (Kitware), it was possible to explain why the new definition is actually superior.

Finally, several numerical experiments have been performed in order to further validate the new definition, focusing on simple controllable chemical systems such as a free electron in vacuum and the H2 molecule, and, in order to keep the concepts simple and clear, the validation is performed between the initial conditions and the actual conditions of a quantum system. The results clearly show that the new definition is, at least in several cases, superior.

In this project, we want to focus on the above formula and modify the integer n=1 to n=2, 3, 4. We observe that different values of this integer brings different ways of measuring the quantum similarity between two (or more) quantum systems. Such investigation will not only help us to understand what is the best possible mathematical expression of quantum similarity in the context of phase-space formulations of quantum mechanics, it will also offer a practical tool exploitable by chemistry base companies such as, for instance, pharmaceutical companies, electronics, materials, etc.

In practice, we mathematically define the new molecular quantum similarity in the following way. Given two molecules A and B described, respectively, by the quasi-distribution functions f_W^A and f_W^A , their quantum similarity is defined as:

$$d_{AB}(t) = \int dx \int dp \left| f_W^A(x;p;t) - f_W^B(x;p;t) \right|$$

where the integration is performed over the phase-space simulation domain. Obviously, this is not the only possible choice and one may consider the definition above as a special case of

$$d_{AB}^{n}(t) = \int dx \int dp \left| \left(f_W^A(x;p;t) \right)^n - \left(f_W^B(x;p;t) \right)^n \right|^{\frac{1}{n}}$$

with n = 1. It is clear from this definition that $d_{AB} \geq 0$ and two molecules are said to be similar only when their distance d_{AB} is close or equal to zero. Conversely, the bigger the number d_{AB}(t) the less similar the molecules are.

The quantum system taken into account for our numerical experiments consists of an hydrogen atom in the Bohr-Oppenheimer approximation. This system is considered going from the ground state and vice versa and the quantum similarity between this two states is computed in function of time. Fig. 1 shows what the orbitals of an electron in such system look like (these results are obtained by means of the signed particle formulation [2]). In particular, the left-hand side shows the electron in the 1s state while the right-hand side shows the same electron in the 2s state.

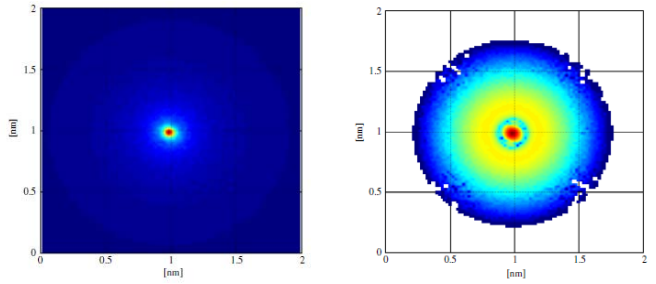

Fig. 1 *Electronic orbitals in a hydrogen atom corresponding to 1s and 2s*

4. Results

A wide range of outcomes is expected. As a preliminary application of our suggested approach, we show in Fig.2 the similarity in time coming from a system of Gaussian wave-packets interacting with an external potential [1].

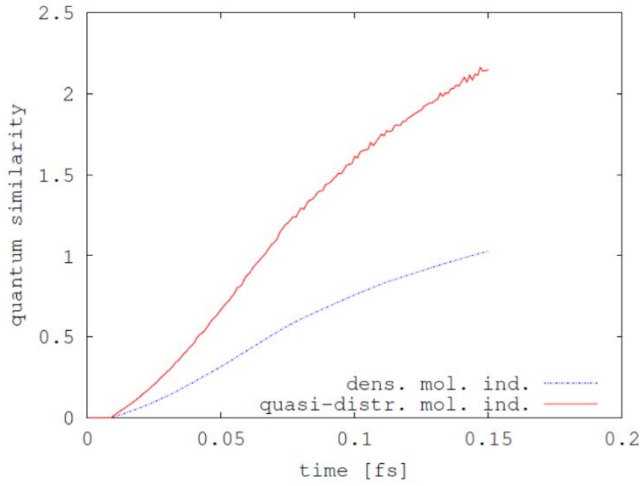

Fig.2 *Comparison of quantum similarities with two different molecular descriptors for a H2 molecule. The dashed (blue) curve corresponds to the density probability as the molecular descriptor, while the continuous (red) curve corresponds to the quasi-distribution function as the descriptor [1].*

Acknowledgments

The work is supported by the "Research program for young scientists at the Bulgarian Academy of Science 2016 funded by the Bulgarian Ministry of Education and Science - Bulgaria National Science Fund under grant DFNP - 101/04.05.2016

References

[1] R. Carbo, J. Arnau, L. Leyda, How similar is a molecule to another? An electron density measure of similarity between two molecular structures, Int. J. Quantum Chem. 17 (1980) 1185.

[2] E. Schrödinger, Quantisierung als Eigenwertproblem, Ann. Phys. 385 (1926) 437–490.

[3] E. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40 (1932) 749.

[4] R.P. Feynman, Space–time approach to non-relativistic quantum mechanics, Rev. Modern Phys. 20 (1948) 367.

[5] L.V. Keldysh, Zh. Eksp. Teor. Fiz., Sov. Phys. JETP 20 (1965).

[6] K. Husimi, Some formal properties of the density matrix, Nippon Sugaku-Buturigakkwai Kizi Dai 3 Ki 22 (4) (1940) 264–314.

[7] D. Bohm, A suggested interpretation of the quantum theory in terms of hidden variables. I, Phys. Rev. 85 (1952) 166.

[8] J.M. Sellier, D.Y. Ivanova, I. Dimov, Molecular Descriptors and Quasi-distribution Functions, Computers and Mathematics with Applications, Elsevier, Volume 70, Issue 11, Pages 2726–2731, (2015).

[9] Sellier , J.M., Ivanova, D. Y., Dimov, I. (2016). "A study on quantum similarity in the phase space". Physica A: Statistical Mechanics and its Applications 460: 116 - 120. doi:10.1016/j.physa.2016.05.003

[10] J.M. Sellier, A Signed Particle Formulation of Non-relativistic Quantum Mechanics, Journal of Computational Physics, Elsevier, Volume 297, Pages 254–265, (2015).

[11] G.M. Downs, in: P. Bultinck, H. De Winter, W. Langenaeker, J.P. Tollenaere (Eds.), Computational Medicinal Chemistry for Drug Discovery, Dekker, Inc., New York, 2003, pp. 364–386. Molecular Descriptors.