FACULTAD DE ESTUDIOS SUPERIORES ZARAGOZA
ÁREA DE QUÍMICA COMPUTACIONAL Y MODELADO MOLECULAR

Asunto: Informe de Miztli 2016

COORDINACIÓN DE SUPERCÓMPUTO,
UNAM.
P R E S E N T E

A través de la presente, hago constar el envío de la información solicitada referente al uso de los recursos en el equipo de supercómputo (Miztli), para la realización de los cálculos de la estructura electrónica que se efectuaron en diversos proyectos que conforman la línea de investigación: "***Estructura electrónica de sistemas de interés biológico***" (número: LI-FESZ-200506). En el Área de Química Computacional y Modelado Molecular de la Facultad de Estudios Superiores-Zaragoza (FES-Zaragoza), UNAM.

En el presente informe se incluyen: proyectos de investigación, artículos con factor de impacto y en JCR, presentaciones en congresos y formación de recursos humanos en el Área de Química Computacional de la FES-Zaragoza. Se adjunta a la presente, los probatorios de los productos generados durante el informe anual en el periodo de febrero-2016 a enero-2017.

Sin más por el momento, aprovecho la ocasión para hacerle llegar un cordial saludo.

A T E N T A M E N T E,

"POR MI RAZA HABLARÁ EL ESPÍRITU"
México, D.F., a 26 de enero de 2017

Dra. Catalina Soriano Correa
Profesora de Carrera Titular "C" TC

INFORMACIÓN DEL PERIODO febrero-2016-enero-2017

RESPONSABLE DE LÍNEA DE INVESTIGACIÓN

"*Estructura electrónica de sistemas de interés biológico*" registrada ante el comité de investigación con el número: LI-FESZ-200506.

PROYECTOS: **Responsable:**

2.- **Estudio de la estructura electrónica, reactividad química, toxicidad, síntesis y efecto terapéutico de moléculas antichagásicas.**
PROYECTO PAPIIT: IN114715 (2015-2017)

1.- ***Estudio teórico de las propiedades estructurales y fisicoquímicas de compuestos anti-bacterianos de tipo sulfonamidas.***
Número: FESZ-RP000/07-247.

PRODUCTIVIDAD EN INVESTIGACIÓN

ARTÍCULOS EN REVISTAS INTERNACIONALES Y CON FACTOR DE IMPACTO
Journal Citation Reports® (Thomson Reuters, 2016).

3.- **Computational study of substituent effects on the acidity, toxicity and chemical reactivity of selected bacteriostatic sulfonamides: Implications for drug design.**
C. Soriano-Correa, C. Barrientos-Salcedo, M. Francisco-Márquez
Journal of Computer -Aided Molecular Desing: Sometido, **2016.**
ISSN:0920-654X (Print) 1573-4951 (Online)

2.- ***Adsorption of Sulfonamides on Phyllosilicate Surfaces by Molecular Modeling Calculations.***
Francisco-Marquez, Misaela; Soriano-Correa, Catalina; Sainz-Diaz, Claro Ignacio.
J. Phys. Chem. C, Manuscri: DOI: 10.1021/acs.jpcc.6b12467 Publication Date (Web): 17 Jan 2017.
F.I = 4.509

1.- ***Study of the Chemical Space of selected Bacteriostatic Sulfonamides from an information-theoretical point of view***
S. López-Rosa; M. Molina-Espíritu; R.O. Esquivel; C. Soriano-Correa; J. S. Dehesa
ChemPhysChem: 17, 4003-4010, 2016. DOI: 10.1002/cphc. 201600790 F.I = 3.138

Área de Química Computacional y Modelado Molecular, Edif-A5, Lab-508B, FES-Zaragoza, UNAM. Tel: (0155) 5773 6333, email: csorico@comunidad.unam.mx.

CAPÍTULOS DE LIBROS CON ISBN

1.- **Information- theoretic Representation of the Chemical Space of many electron systems**
R. O. Esquivel., S. López-Rosa, M. Molina-Espíritu., C. Soriano-Correa, J. C. Angulo, J. S. Dehesa.
In book: Frontiers in Computational Chemistry, Chapter: Chapter 6, Vol. 3, Publisher: 2016 Bentham Science Publishers, Editors: Zaheer Ul-Haq and Jeffry D. Madura (Eds, pp.3-46. Zaheer Ul-Haq and Jeffry D. Madura (Eds.)

PARTICIPACIÓN EN CONGRESOS

2- **XXI Congreso Nacional de Parasitología.** *"Diseño, síntesis y evaluación biológica de análogos de benznidazol para el tratamiento de la Tripasomiasis americana". Campos-Fernández LV.,* Soriano-Correa, C., Padilla-Martínez II., Trujillo-Ferrara, JG.; Cuevas-Hernández, RI. Facultad de Ciencias, UNAM. Ciudad de México, septiembre de 2016.

FORMACIÓN DE RECURSOS HUMANOS

Tesis de Doctorado

2. Directora: *"**Diseño de péptidos antineoplásicos y análisis de su efecto celular y molecular en líneas celulares de cáncer de mama**",* del alumno M. C. Hugo Arana Vidal. Doctorado en Ciencias Biomédicas, Universidad Veracruzana, 2012-2016. En proceso

1.- Directora: *"**Análisis de la vía de NF-kB regulada por la acción del péptido anti-inflamatorio CNS en ratas con glioma**",* del alumno Q.C. Oskar Ibsan Valerio Hernández. Centro de Investigaciones Cerebrales (CICE), Universidad Veracruzana, 2015-2019. En proceso

Tesis de maestría

2.- Directora: ***"Estudio teórico-experimental en el diseño de nuevos fármacos derivados de imidazoles con potencial actividad antichagásica",*** de la alumna QFB. Linda Verónica Campos Fernández. Maestría en Ciencias en Farmacología. Escuela Superior de Medicina – Instituto Politécnico Nacional, 2015-2016.

Concluida: obtención del grado, enero 25 de 2017.

1.- Directora: ***"Diseño, síntesis, evaluación biológica de moléculas derivadas de imidazol con potencial antichagásico",*** del alumno QFB. Juan Andres Salazar Alvarado. Maestría en Ciencias en Farmacología. Escuela Superior de Medicina – Instituto Politécnico Nacional, 2015-2017. En proceso

Tesis de Licenciatura

1.- Directora: *"***Análisis *in silico* de proteínas de membranas de *Trypanosoma cruzi"***,* de la alumna Joscelyn Hernández Santiago. Facultad de Bioanálisis Región Veracruz, Universidad Veracruzana.
Concluida: obtención del grado enero 17 de 2017

FACULTAD DE ESTUDIOS SUPERIORES ZARAGOZA

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

OFICIO FESZ/DEPI/230/14

ASUNTO: Renovación de Registro de Línea de Investigación

DRA. CATALINA SORIANO CORREA
RESPONSABLE DE LA LÍNEA DE INVESTIGACIÓN
ESTRUCTURA ELECTRÓNICA DE SISTEMAS DE INTERÉS BIOLÓGICO
P R E S E N T E

Me es grato comunicarle que el Comité de Investigación, en su sesión ordinaria del 26 de agosto de 2014, **acordó aprobar** la renovación por **cinco años** del registro de la Línea de Investigación **"Estructura electrónica de sistemas de interés biológico con el número LI-FESZ-200506"** del cual usted es responsable, en virtud que cumple con lo estipulado en el Reglamento General de Investigación vigente en nuestra Facultad, en virtud de su alta productividad científica.

Participantes de la Línea de Investigación:

- **Dra. Carolina Barrientos Salcedo**
- **Dra. Angélica Beatriz Raya Rangel**
- **Dr. Rodolfo O. Esquivel Olea**

Sin otro particular, aprovecho la ocasión para enviarle un saludo cordial.

A t e n t a m e n t e
"POR MI RAZA HABLARÁ EL ESPÍRITU"
México, D. F., 24 de septiembre de 2014
EL JEFE DE LA DIVISIÓN

DR. EDELMIRO SANTIAGO OSORIO
C.c.p. Dr. Vicente Jesús Hernández Abad, Secretario de Consejo Técnico
C.c.p. Comités Académicos de Carrera
C.c.p. Archivo Comisiones de Evaluación del PRIDE
C.c.p. Archivo Comisiones Dictaminadoras
ESO/MISC/AAZ/ ign*

Dirección General de Asuntos del Personal Académico

Dirección de Desarrollo Académico
Oficio DGAP/ DDA/ 1078 /2016

Dra. Catalina Soriano Correa
Profesora
Facultad de Estudios Superiores "Zaragoza"
P r e s e n t e

Estimada doctora Soriano:

Por instrucciones del doctor Carlos Arámburo de la Hoz, Director General de Asuntos del Personal Académico, me permito comunicar a usted, que después de una revisión cuidadosa, el Comité Evaluador del Área de las Ciencias Físico Matemáticas y de las Ingenierías del Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), acordó aprobar la renovación de su proyecto **IN114715** *Estudio de la estructura electrónica, reactividad química, toxicidad, síntesis y efecto terapéutico de moléculas antichagásicas*.

Por lo anterior, deberá atender los siguientes puntos:

1. Consultar e imprimir el dictamen que se encuentra en el apartado de "Dictamen";
2. Imprimir, firmar y entregar el presupuesto autorizado, a más tardar el **2 de diciembre de 2016**, con un horario de 9:00 a 15:00 y de 16:00 a 19:00 horas en las oficinas de la Dirección de Desarrollo Académico, ubicadas en el 4º piso del edificio C, zona cultural.
3. Le recomendamos leer cuidadosamente las observaciones indicadas en su presupuesto.

Sin más por el momento, aprovecho para enviarle un cordial saludo.

A t e n t a m e n t e
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., 24 de noviembre de 2016

La Directora

Dra. Claudia C. Mendoza Rosales

VNIVERSIDAD NACIONAL AVTÓNOMA DE MEXICO

FACULTAD DE ESTUDIOS SUPERIORES ZARAGOZA
COORDINACIÓN DE INVESTIGACIÓN
OFICIO FESZ/CI/004/009
ASUNTO: Registro de proyecto de investigación.

DRA. CATALINA SORIANO CORREA
P r e s e n t e

Tengo el agrado de informarle que su proyecto de investigación: "ESTUDIO TÉORICO DE LAS PROPIEADES ESTRUCTURALES Y FISICOQUÍMICAS DE COMPUESTOS ANTIBACTERIANOS DE TIPO SULFONAMIDA", ha quedado registrado ante esta División con el número FESZ-RP000/07-247.

Ruego a usted considere que para mantener el registro del proyecto deberá entregar a esta División un breve informe semestral de los productos obtenidos del proyecto de investigación mencionado, en el formato anexo a este oficio. En el caso de que su proyecto esté financiado (CONACYT, DGAPA, etc.) y deba presentar un informe anual del mismo, le solicito no tome en cuenta la entrega del informe semestral, pero considere hacer llegar a esta División una fotocopia del informe anual de resultados. El primer informe semestral se recibirá en el transcurso del mes de Diciembre de 2007.

Sin otro particular, le envío un cordial saludo.

A t e n t a m e n t e.

"POR MI RAZA HABLARÁ EL ESPÍRITU"
México, D.F., a 26 de marzo de 2007.

EL COORDINADOR DE INVESTIGACIÓN.

M. en C. JOSÉ LUIS TREJO MIRANDA

JLTM/JCVLVaa

ChemPubSoc Europe

DOI: 10.1002/cphc.201600790

CHEMPHYSCHEM Articles

Study of the Chemical Space of Selected Bacteriostatic Sulfonamides from an Information Theory Point of View

Sheila López-Rosa,*[a, b] Moyocoyani Molina-Espíritu,[c] Rodolfo O. Esquivel,[b, c] Catalina Soriano-Correa,[d] and Jésus S. Dehesa[b, e]

The relative structural location of a selected group of 27 sulfonamide-like molecules in a chemical space defined by three information theory quantities (Shannon entropy, Fisher information, and disequilibrium) is discussed. This group is composed of 15 active bacteriostatic molecules, 11 theoretically designed ones, and *para*-aminobenzoic acid. This endeavor allows molecules that share common chemical properties through the molecular backbone, but with significant differences in the identity of the chemical substituents, which might result in bacteriostatic activity, to be structurally classified and characterized. This is performed by quantifying the structural changes on the electron density distribution due to different functional groups and number of electrons. The macroscopic molecular features are described by means of the entropy-like notions of spatial electronic delocalization, order, and uniformity. Hence, an information theory three-dimensional space (IT-3D) emerges that allows molecules with common properties to be gathered. This space witnesses the biological activity of the sulfonamides. Some structural aspects and information theory properties can be associated, as a result of the IT-3D chemical space, with the bacteriostatic activity of these molecules. Most interesting is that the active bacteriostatic molecules are more similar to *para*-aminobenzoic acid than to the theoretically designed analogues.

1. Introduction

A chemical space could be spanned by all possible organic molecules and chemical compounds formed from all isomeric stoichiometric combinations that result in energetically stable electronic systems, including those present in biological systems.[1] A deeper understanding of this vast set of molecules will advance our knowledge of biochemical processes. Understanding its relative location in the framework of the chemical space, which allows its systematic and rational classification, is crucial for chemical applications in materials science and pharmaceutical research. The analysis and exploration of the chemical space represent a highly demanding computational task due to the immense number of possible stable molecules[2] and the number of parameters available for describing a specific molecule, which turns this challenge into a multidimensional problem. To circumvent efforts of navigating the chemical space, many studies have been undertaken, aside from the conventional methodologies derived from the field of bio-/chemoinformatics. It seems natural that the exploration of chemical space has improved by merging these tools with visual data analysis techniques;[3–5] or the combination of quantitative structure–activity relationship (QSAR) methods with computational chemistry calculations.[6–8] Inverse quantum chemical approaches are becoming popular to find a chemical structure departing from a set of desired properties;[9] an outstanding example of these kind of methodologies is the alchemical potential approximation.[10–15]

Recently, a different approach has emerged for exploring the chemical space. This new alternative is driven by an information theory methodology, and is based on quantification of the structural shape of electron density distribution of molecules.[16,17] These studies suggest that molecules with similar informational properties (e.g. narrowness, localizability, and uniformity) are related to organic compounds that share similarities in 3D structures and chemical properties. Indeed, from simple physical systems to complex biological ensembles, the properties of atoms and molecules strongly depend on the spread of the one-electron density, $\rho(\vec{r})$, of the corresponding quantum-mechanical state.[18,19] Furthermore, the information theory of quantum systems provides an entropy-based characterization of the atomic and molecular systems, which complements the energy-based representation obtained through wave function and density functional methods. Measures of

[a] *Prof. S. López-Rosa*
Departamento de Fsica Aplicada II, Universidad de Sevilla
41012 Sevilla (Spain)
E-mail: slopezrosa@us.es

[b] *Prof. S. López-Rosa, Prof. R. O. Esquivel, Prof. J. S. Dehesa*
Instituto Carlos I de Fsica Teórica y Computacional
Universidad de Granada, 18071 Granada (Spain)

[c] *Dr. M. Molina-Espíritu, Prof. R. O. Esquivel*
Departamento de Química, Universidad Autónoma Metropolitana
09340-Mexico City (México)

[d] *Prof. C. Soriano-Correa*
Química Computacional, FES-Zaragoza
Universidad Nacional Autónoma de México
09230-Iztapalapa, Mexico City (México)

[e] *Prof. J. S. Dehesa*
Departamento de Fsica Atómica, Molecular y Nuclear
Universidad de Granada, 18071 Granada (Spain)

Supporting Information and ORCID identification number(s) for the author(s) of this article can be found under http://dx.doi.org/10.1002/cphc.201600790.

uncertainty, randomness, disorder, and localization are basic ingredients that play a relevant role in the identification and description of numerous quantum phenomena in physical systems and chemical processes. In particular, the fundamental energy quantities can be expressed by means of the Shannon entropy, disequilibrium, and Fisher information.[20–22]

These macroscopic features of delocalization, uniformity, and order can be quantified by the information theory measures of Shannon entropy, disequilibrium, and Fisher information of the corresponding electron densities of the system, respectively. Taking into account these entropy-like notions, an information theory space in three dimensions (IT-3D) can be defined. This space has been introduced in recent work,[17] in which a great variety of systems were analyzed. Therein, we showed that the entropic concepts of delocalizability, uniformity, and order, together with their associated dyadic products of complexity [measures of Fisher–Shannon (FS) and López-Ruiz–Mancini–Calbet (LMC)], allowed us to differentiate regions of physical, chemical, or biological qualities.

On the other hand, sulfonamides are synthetic antimicrobial agents with a wide spectrum that encompasses most gram-positive and many gram-negative organisms. They are used in the prevention and treatment of bacterial infections, diabetes mellitus, edema, and hypertension, among others. These drugs were the first efficient treatments to be employed chemotherapeutically for the prevention and cure of bacterial infections. Furthermore, sulfonamides possess organic sulfur compounds containing the $-SO_2NH_2$ group (the amides of sulfonic acids). One of the reported mechanisms of action for the sulfonamides dictates that their bacteriostatic activity is due to interference with the metabolic processes in bacteria that require *para*-aminobenzoic acid (PABA), see Figure 1) in the synthesis of folic acid and ultimately of purine and DNA.[23,24] By taking this mechanism into account, it is expected that our set of molecules with bacteriostatic activity will share more common informational properties with PABA than the rest of the molecules.

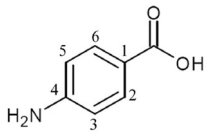


Figure 1. The structural formula of *para*-aminobenzoic acid (PABA).

Numerous attempts have been made to correlate the antibacterial activity of sulfonamides with various physicochemical and pharmacological properties, such as pK_a, acidity, basicity, protein binding, electronic charge distribution, lipophilicity, and solubility.[25–29] However, to date, despite the great deal of experimental and theoretical work performed on sulfonamides, there is scarce information about the structure–activity relationships to help understand its mechanism of action. Therefore, it is worth exploring novel alternatives, such as the chemical space framework, which provides information about the molecular
structure and its relationship with the biological behavior.

Herein, we have undertaken an investigation to apply information theory concepts to the molecular densities to explore the chemical space of a group of selected sulfonamides. This should serve as a probe for testing our information theory space. The latter is achieved by measuring the structural changes on the electron density distribution due to different functional groups and different number of electrons. Thus, we can construct an IT-3D space that gathers molecules with common properties. One of the goals of our study is to test the information theory space as a witness of biological activity through the global and local behavior of the electron density distribution. Therefore, the purpose of our study is twofold: 1) to explore a small subset of the chemical space to relate information theory quantities with chemical structures and bacteriostatic activity, and 2) to assess the effect of functional groups on the spread of the electron density distribution on the derivatives of 4-aminobenzensulfonamide. Atomic units are used throughout the paper.

2. Information Theory Measures

The ground-state one-electron density, $\rho(\vec{r})$, is a physical observable that can be obtained experimentally or calculated by using ab initio or DFT methods. It is known that the physical and chemical properties of atoms and molecules are determined by $\rho(\vec{r})$.[30] Moreover, the information theory of many-electron systems allows us to characterize the physical and chemical properties of these systems by means of functionals of their one-density electron density, which are generally referenced by information theory measures.[19,31–36] Broadly speaking, these measures have a global or local character, depending on whether they are scarcely or very sensitive to fluctuations of the density, respectively.

Global measures quantify the total extent of the probability density in various ways, according to their different analytical structures; in other words, they are described by means of logarithmic (Shannon entropy, $S[\rho]$) and power (e.g. disequilibrium, $D[\rho]$) functionals of the density. Local measures, for example, Fisher information, $I[\rho]$, quantify the gradient of the density. Thus, Fisher information is very sensitive to changes of the density over a small region of the distribution. For instance, a molecule characterized by a broader electron density distribution has several peaks of similar heights, whereas different physical and chemical properties can be expected from a different molecule characterized by a narrower electron density with several peaks of different heights. Briefly, the structural features of the electron density can be described by utilizing various information theory measures, which together qualify and quantify the main topographical features. Let us now define the most relevant information theory measures used in our study.

The Shannon entropy of the unity-normalized electronic probability distribution, $\rho(\vec{r})$, of a molecule is defined by Equation (1):[37]

$$S[\rho] = -\int \rho(\vec{r}) ln \rho(\vec{r}) d\vec{r} \qquad (1)$$

This quantity quantifies the total electronic spread in the molecular configuration space, so it constitutes a measure of the delocalization (lack of structure) of the electron density.

The disequilibrium (also called information energy in other physical contexts), which quantifies the departure of the probability density from uniformity (equiprobability), is defined by Equation (2):[38]

$$D[\rho] = \int \rho^2(\vec{r})d\vec{r} \tag{2}$$

The Fisher information is defined by the gradient density functional given by Equation (3):[18,39]

$$I[\rho] = \int \frac{\left|\vec{\nabla}\rho(\vec{r})\right|^2}{\rho(\vec{r})} d\vec{r} \tag{3}$$

This quantity measures the spatial pointwise concentration of the electronic probability cloud, and it quantifies the gradient content of the electron distribution; hence revealing irregularities of the density and providing a quantitative estimation of its fluctuations. Additionally, according to the localized/delocalized features of the distributions, the Fisher information can be interpreted as a measure of the departure of the probability density from disorder.

These three quantities measure different macroscopic aspects or facets of the electron density distribution, so that each molecule has a distinctive set of these three information theory parameters.[16,17] If we mutually compare two or more molecules, we expect that those with similar physicochemical properties will have electron density distributions with a similar spread. Therefore, an IT-3D space composed by $S \equiv S[\rho]$, $I \equiv I[\rho]$, and $D \equiv D[\rho]$ could probably be an appropriate threefold pattern to group and disentangle specific information theory regions of molecules that share similar spatial regions of delocalizabilty, uniformity, and order.

Aside from the properties of the IT-3D measures mentioned above, let us also consider two additional information theory measures that have proven to be very useful for quantifying the complexity of physical systems: the FS and LMC complexity measures.[19,40–42] These two complexity quantities, which simultaneously quantify two facets of the electron density of the system, have a number of relevant common properties: dimensionless, invariance under replication, translation and scaling transformations, and minimality (i.e. they have minimum values at the two extreme cases of perfect order and maximum disorder).

The FS measure of complexity,[40,42] C_{FS}, is defined in Equation (4) by the product of two single-facet entropic quantities of local and global character (the Fisher information measure, I, and the Shannon entropy, S, appropriately modified to preserve the common complexity properties):

$$C_{FS} = I[\rho]J[\rho], \tag{4}$$

in which $J[\rho] = \frac{1}{2\pi e} e^{\frac{2}{3}S[\rho]}$ is the Shannon power entropy. This quantity measures the combined balance of the total spreading and narrowness of the electron density (i.e. the delocalization and order features). It can be considered as a measure of electronic correlation in atomic systems.[40] Moreover, it is bounded from below as $C_{FS} \geq 3$ for any 3D probability density.

The LMC measure of complexity,[41,43] C_{LMC}, is defined by Equation (5) as the product of two single-facet entropy measures of global character (disequilibrium, D, and the Shannon exponential entropy, e^S) as:

$$C_{LMC} = D[\rho]e^{S[\rho]}, \tag{5}$$

which jointly describes the average height and extent of the density (i.e. the uniformity and delocalization features). This quantity satisfies the bound $C_{LMC} \geq 1$ for any 3D probability density.

3. Results and Discussion

A selected group of 27 sulfonamide-type molecules was chosen as follows: 15 of them possessed bacteriostatic activity and have been studied previously,[25] and 11 molecules were computationally designed as structural analogues of sulfonamide. The latter systems have not been reported before, to the best of our knowledge.[44] Additionally, PABA was also considered (Figure 1). As mentioned above, chemical analogues to sulfonamide can be constructed by substituting one hydrogen atom in the amino group by a different R^1 group. The chemical structure for these molecules is shown in Figure 2. To analyze the effect of the R group in a deeper way, we have generated these 11 novel compounds with different R groups that do not present bacteriostatic activity, but analysis is useful to describe sulfonamides from an information theory point of view. Results for the information theory analysis of the chemical space for these molecules are presented below and the values are summarized in the table in the Supporting Information. We assume that the electron density of all molecular systems are

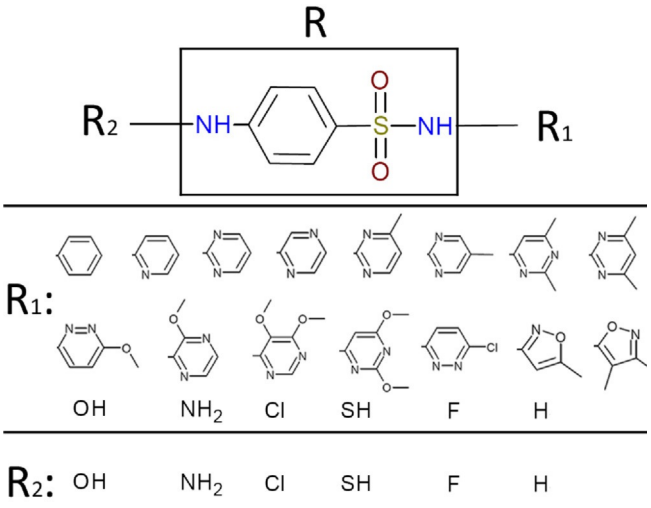


Figure 2. Construction of the molecular models. R corresponds to 4-aminobenzensulfonamide; R^1 and R^2 indicate different functional group substituents on the molecule.

normalized to the number of electrons, N, of the system under consideration.

The molecular wave functions were obtained through electronic structure calculations performed with the Gaussian 98 and Gaussian 09[45] suite of programs. The structures were optimized at the restricted Hartree–Fock (RHF) level of theory with a 6-311++G** basis set. Then, single-point calculation were performed on the optimized structures at the HF/6-31++G(d,p) and B3LYP/6-31++G(d,p) levels of theory; a frequency analysis was performed to corroborate that the obtained structure corresponded to an equilibrium geometry. Details of the electronic calculations can be found in Ref. [46]. Furthermore, molecular electron densities were obtained along with all information theory measures, and complexity dyadic products in position space, as defined in Section 2 (i.e. S, D, I, C_{LMC}, C_{FS}) by employing software developed in our laboratory, along with 3D numerical integration routines[47] and the DGRID suite of programs.[48]

The IT-3D space for all sulfonamide-type molecules and PABA is shown in Figure 3. We observe that molecules have been grouped together into five different regions of the chemical space. This IT-3D chemical gathering is discussed below; most of the essential features depicted in Figure 3 are linked with Figure 4 and Table 1. For instance, in Figure 4, we have depicted these five groups described above. One of the most remarkable features of Figure 4 is that all molecules with bacteriostatic activity belong to groups 3 and 5, apart from the sulfonamide molecule, which possesses the least bacteriostatic activity and belongs to group 2.

It is noteworthy that the grouping obtained by the IT-3D S, I and D measures (Figure 3) has been validated by use of multivariable analysis techniques. Thus, to corroborate and quantify the visual analysis depicted in Figure 3, a k-mean cluster analysis on the raw data was performed (Table 1); the search for five possible groups was specified as a model probe in the input. Interestingly, the euclidean distance reveals that the group of molecules with reported bacteriostatic activity (group 5 in Figure 4) is closer to PABA (see Table 1), that is, in terms of the global and local behavior of the spread of electron density distribution, molecules in group 5 results to be closely related to PABA in this IT-3D space.

Table 1. k-mean analysis to corroborate the grouping observed in Figure 3.

	Group 1	Group 2	Group 3	Group 4	Group 5
Group 1	0.0	14.463	19.532	25.328	9.163
Group 2	14.463	0.0	5.077	10.865	5.337
Group 3	19.532	5.077	0.0	5.815	10.380
Group 4	25.328	10.865	5.815	0.0	16.189
Group 5	9.163	5.334	10.380	16.189	0.0

Additional testing confirmed the reliability of Shannon entropy, Fisher information, and desequilibrium as intrinsic descriptors of molecular systems. A principal component analysis (PCA) was performed on the rescaled data (z score of the raw data and considering only two principal components), followed by a hierarchical cluster analysis (searching for five clusters and evaluating the euclidean distance between means). This kind of testing corroborates the existence of the five main groups mentioned above (data are available on request). The

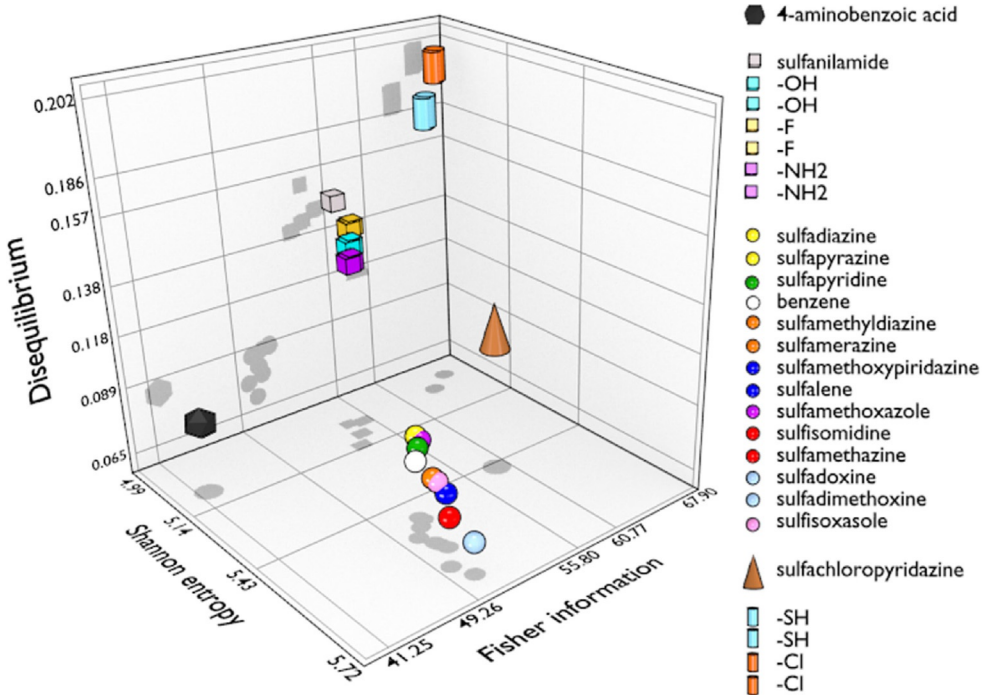


Figure 3. IT-3D space for the 27 analogues of sulfonamide and PABA. Atomic units are employed.

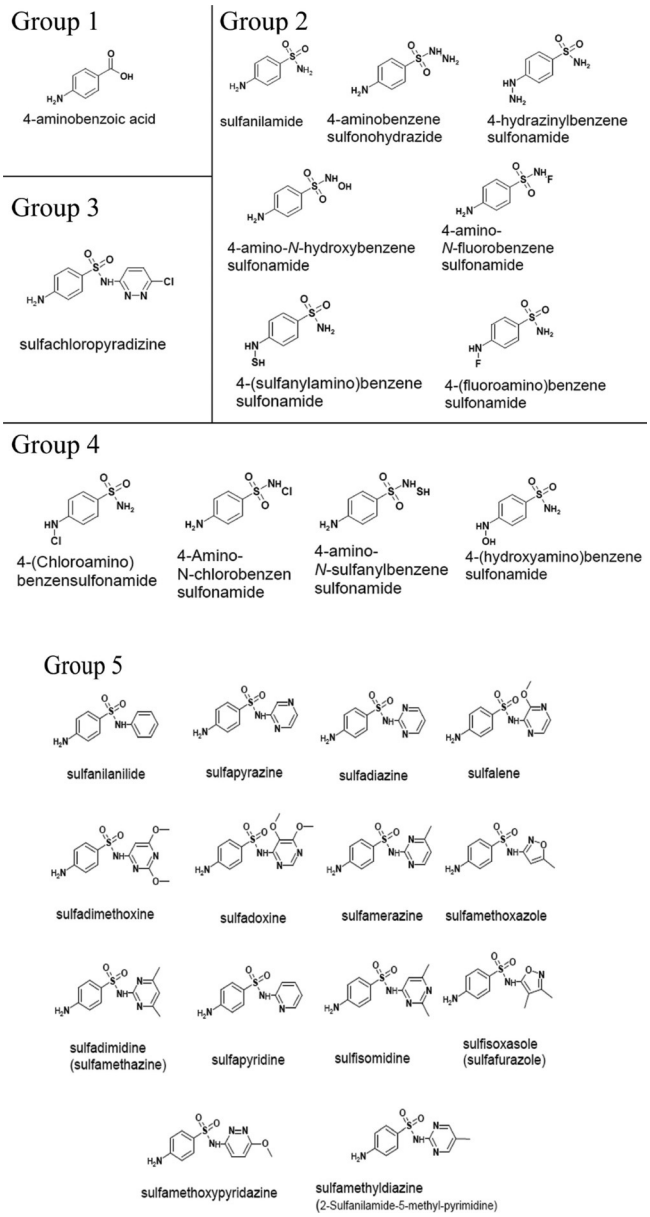


Figure 4. Groups of sulfonamide-type molecules, along with PABA, in the IT-3D chemical space (Figure 3).

multivariate analysis was performed by using the Origin Pro 2015 program.[49]

A closer look at Figure 3 reveals a definite pattern among the information theory aspects of *I* (organization) and *S* (uncertainty). In searching for this pattern, we have found it useful to analyze the set of studied molecules by plotting the contribution of each one of the information measures *I* (organization) and *J* (uncertainty) through the FS complexity. Notably, the Shannon entropy, *S*, is included in the FS complexity through the power entropy, *J*, as defined in Section 2. Hence, the FS information plane is depicted in Figure 5. The structure of the molecules and number of electrons have also been represented.

We have found that, concomitant with the IT-3D chemical pattern, there is additional characterization based upon the R^1 and R^2 groups, that is, the chemical character of the substituents, regardless of its chemical structure. The most apparent features and patterns shown in Figure 5, are described below:

1) Including R^1 as an N^1-heteroaromatic substituent (an aromatic heterocyclic compound) implies a drastic decrease in the disequilibrium of the molecule, that is, the electron density becomes closer to uniformity, which could be associated with the presence of the lone electron pair in the nitrogen atoms. In these molecules, R^1 corresponds to pyridine- (sulfapyridine) or diazine-type (sufadiazine and sulfapyrazine) compounds and R^2 to H atom.
2) Molecules with an extra methyl (sulfamerazine, sulfamethyldiazine, sulfamethazine and sulfisomidine) or methoxy group (sulfalene, sulfamethoxypyridazine, sulfadimethoxine and sulfadoxine) possess lower values of Fisher information relative to molecules with no extra subtitution of the diazine group. In other words, the substituent groups increase the rigidity of the molecules,[25] and thus, induce less ordered systems.
3) Differences between molecules with methoxy and methyl groups are due to the stronger electron-donating power of the former than the latter. This is detected by the Shannon entropy. The methyl group has a smaller value of *S*, which means that the probability density is more localized than the density of OCH_3. This behavior might be associated with the stronger electron-donating power of the methoxy group than that of the methyl group.
4) The Fisher information and disequilibrium decrease with the number of CH_3 or OCH_3 groups present in the molecules. The density probability becomes more uniform and ordered; this effect might be, in turn, related to the rigidity of the molecules. The probability density of these molecules is affected by the size of the substituents.
5) The rest of the modified molecules show higher disequilibrium and Fisher information values, which, in turn, are higher than those in group 5. The electron density of these molecules becomes less uniform and more disordered; this behavior is accentuated when the atoms in R^1 and R^2 belong to larger groups.

All compounds considered herein are characterized through their information content, featuring any of the three different IT qualities aforementioned, which make them unique among the rest of the structures. An extension of the analysis for the information content of the chemical space might be to take into account two further information theory measures that represent the complexity of the systems (namely, FS and LMC measures) to jointly grasp composite aspects of the electronic structure of the systems.

The IT-3D space and LMC or FS complexity measures are displayed in Figure 6. Notably, larger values of C_{FS} are depicted in shades of red and smaller ones in shades of blue. It is interesting to emphasize that, with some exceptions, the complexity measures show the opposite behavior. The main results are summarized below:

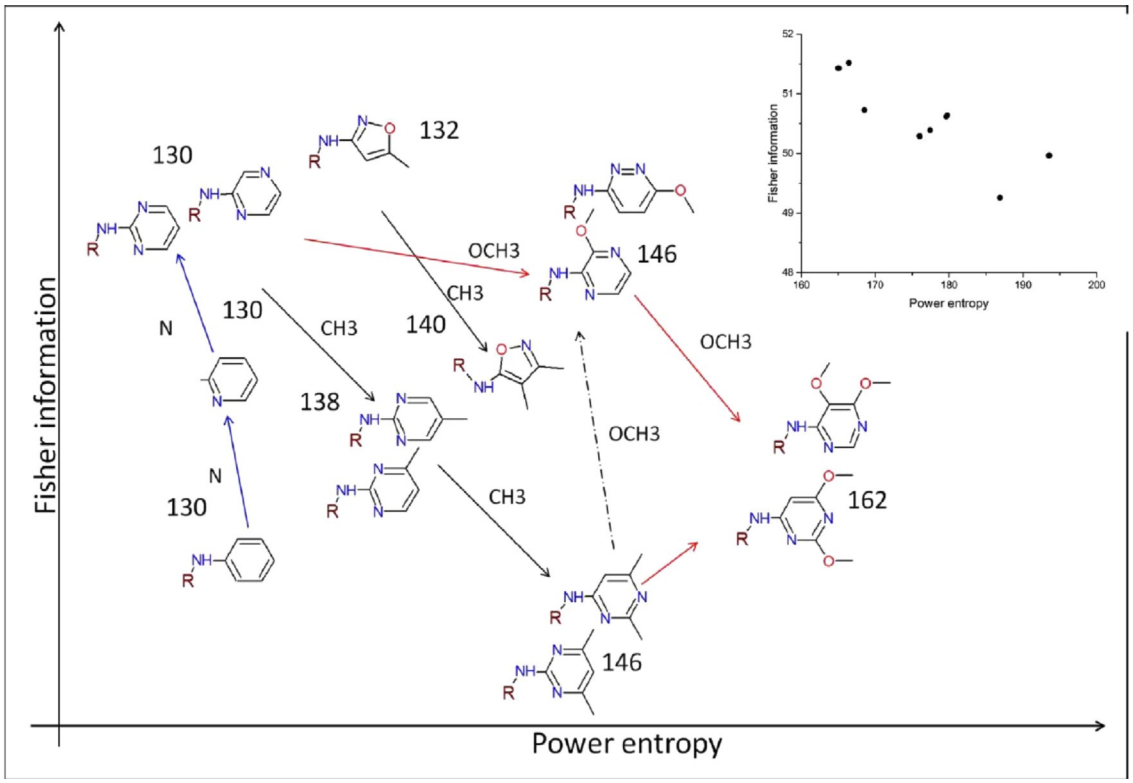


Figure 5. The FS plane (inset box) and its molecular representation of sulfonamide-type bacteriostatic compounds along with PABA. The structures of the molecules and number of electrons have also been included.

1) PABA, which is the only natural compound under consideration herein, has the lowest value of the complexity measures.
2) Including cyclic groups in the molecule increases (decreases) the C_{FS} (C_{LMC}) complexity. Sulfadiazine and sulfapyrazine possess larger (lower) values of C_{FS} (C_{LMC}) relative to that of sulfapydirine due to the presence of two nitrogen atoms instead one in the R^1 group.
3) Molecules with CH_3 (sulfamerazine, sulfamethyldiazine, sulfamethazine, and sulfisomidine) or OCH_3 groups (sulfalene, sulfamethoxypyridazine, sulfadimethoxine, and sulfadoxine) possess higher (smaller) values of C_{FS} (C_{LMC}) compared with molecules with only a benzene group.
4) The number of CH_3 or OCH_3 groups present in the molecule increases (decreases) the C_{FS} (C_{LMC}) complexity.

4. Conclusions

We applied information theory concepts to the molecular densities to explore the chemical space of a group of selected sulfonamide-type molecules. This was achieved by measuring the structural changes on the electron density distribution due to different functional groups and number of electrons. In this manner, we constructed an IT-3D space that gathered molecules with similar properties. One of the main aims of our study was to test the information theory space as an indicator of biological activity through the global and local behavior of the electron density distribution.

We found that, concomitant with the IT-3D space, there was additional characterization based upon the R^1 and R^2 groups, that is, the chemical character of the substituents, regardless of its chemical structure. Throughout the study, we were able to associate some structural aspects and information theory properties that arose from the IT-3D chemical space to the bacteriostatic activity of these molecules. An interesting result of the analysis reveals that reported bacteriostatic molecules are more similar to the information theory properties of PABA than those designed theoretically (see the Table of Contents graphical abstract for illustration).

Acknowledgements

We acknowledge partial financial support through the UGR-GENIL 2015 and Mexican grants from CONACyT (grant number 80244) and the Spanish MINECO project FIS2014-59311-P (cofinanced by FEDER), and grants FQM-7276, FQM-020, FQM-239, and FQM-207 from the Junta de Andaluca. We are grateful to the Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) of the Universidad Nacional Autónoma de México (UNAM) for an allocation of computer time on the Miztli supercomputer.

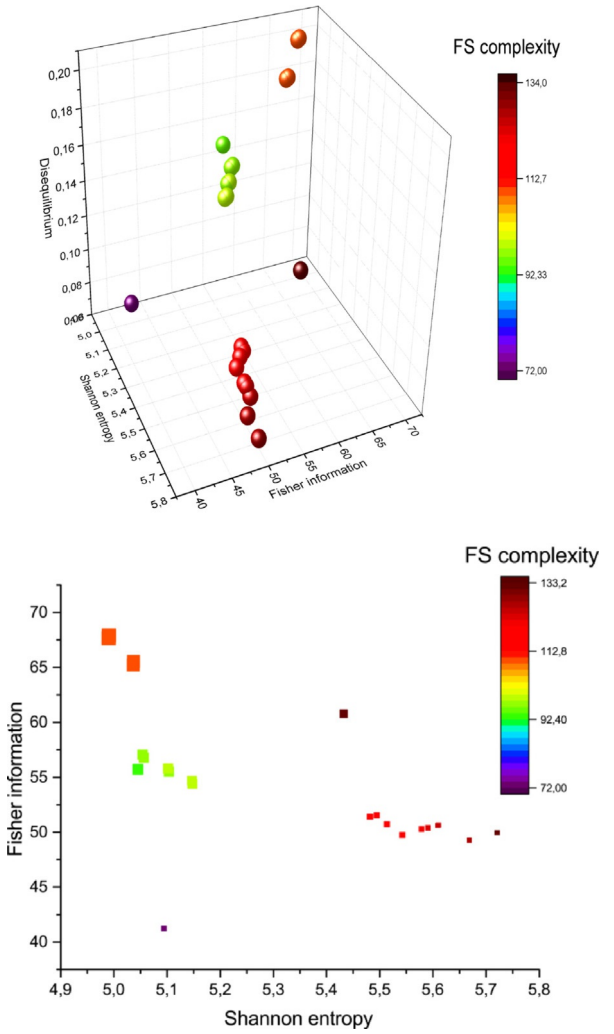


Figure 6. Top: IT-3D space for sulfonamides. Larger values of C_{FS} are depicted in reddish colors, whereas smaller ones are depicted in blueish ones. Bottom: F-S plane. Atomic units are employed.

Keywords: chemical space · electronic structure · information theory · structure–activity relationships · sulfonamides

[1] C. M. Dobson, *Nature* **2004**, *432*, 824 – 828.
[2] C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855 – 861.
[3] T. I. Oprea, J. Gottfrie, *J. Comb. Chem.* **2001**, *3*, 157 – 166.
[4] J. L. Medina-Franco, K. Martínez-Mayorga, M. A. Giulianotti, R. A. Houghten, C. Pinilla, *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 322 – 333.
[5] J. L. Reymond, *Acc. Chem. Res.* **2015**, *48*, 722 – 730.
[6] D. Xiao, W. Yang, D. N. Beratan, *J. Chem. Phys.* **2008**, *129*, 044106.
[7] V. K. Shukla, A. K. Sachan, S. K. Pathak, R. Srivastava, O. Prasad, L. Sinha, *J. Mol. Struct.* **2016**, *1106*, 265 – 276.
[8] A. R. Finkelmann, A. H. Göller, G. Schneider, *Chem. Commun.* **2016**, *52*, 681 – 684.
[9] T. Weymuth, M. Reiher, *Int. J. Quantum Chem.* **2014**, *114*, 823 – 837.
[10] O. A. Von Lilienfeld, M. E. Tuckerman, *J. Chem. Theory Comput.* **2007**, *3*, 1083 – 1090.
[11] O. A. von Lilienfeld, *J. Chem. Phys.* **2009**, *131*, 164102.
[12] D. Sheppard, G. Henkelman, O. A. von Lilienfeld, *J. Chem. Phys.* **2010**, *133*, 084104.
[13] R. Balawender, M. A. Welearegay, M. Lesiuk, F. De Proft, P. Geerlings, *J. Chem. Theory Comput.* **2013**, *9*, 5327 – 5340.
[14] O. A. von Lilienfeld, *Int. J. Quantum Chem.* **2013**, *113*, 1676 – 1689.
[15] M. to Baben, J. O. Achenbach, O. A. von Lilienfeld, *J. Chem. Phys.* **2016**, *144*, 104103.
[16] R. O. Esquivel, M. Molina-Espíritu, S. López-Rosa, C. Soriano-Correa, C. Barrientos-Salcedo, M. Kohout, J. S. Dehesa, *ChemPhysChem* **2015**, *16*, 2571 – 2581.
[17] R. O. Esquivel, S. López-Rosa, M. Molina-Espíritu, J. C. Angulo, J. S. Dehesa, *Theor. Chem. Acc.* **2016**, DOI: 10.1007/s00214-016-2002-x.
[18] B. R. Frieden, *Science from Fisher Information*, Cambridge University Press, Cambridge, **2004**.
[19] J. S. Dehesa, S. López-Rosa, D. Manzano, *Entropy and Complexity Analyses of D-Dimensional Quantum Systems*, Springer, Berlin, **2010**.
[20] L. M. Ghiringhelli, I. P. Hamilton, L. Delle Site, *J. Chem. Phys.* **2010**, *132*, 014106.
[21] A. Nágy, *Chem. Phys. Lett.* **2013**, *556*, 355 – 358.
[22] C. Rong, T. Lu, PW. Ayers, P. K. Chattaraj, S. Liu, *Phys. Chem. Chem. Phys.* **2015**, *17*, 4977 – 4988.
[23] E. Pérez-Trallero, L. Iglesias, *Enfermedades infecciosas y microbiologia clinica* **2003**, *21*, 520 – 529.
[24] R. B. Silverman, *The Organic Chemistry of Drug Design and Drug Action*, Academic Press, New York, **1992**.
[25] M. E. Wolff, *Burger's Medicinal Chemistry, Part 2*, 4th. ed., Wiley, New York, **1979**.
[26] Exploring QSAR. Fundamentals and Applications in Chemistry and Biology, C. Hansch, A. Leo, *ACS Professional Reference Book*, American Chemical Society, Washington, DC, **1995**.
[27] J. K. Seydel, *J. Pharm. Sci.* **1968**, *57*, 1455 – 1478.
[28] L. H. M. Janssen, M. J. B. Mengelers, P. E. Hougee, A. S. J. P. A. M. van Miert, *J. Vet. Pharmacol. Ther.* **1997**, *20*, 276 – 283.
[29] M. Remko, *J. Mol. Struct.* **2009**, *897*, 73 – 82.
[30] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864 – B871.
[31] C. Arndt, *Information Measures*, Springer, Berlin, **2013**.
[32] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, **1994**.
[33] S. R. Gadre in *Reviews of Modern Quantum Chemistry: A Celebration in the Contributions of Robert G. Parr, Vol. 1*, World Scientific, Singapore, **2003**.
[34] M. Molina-Espíritu, Ph.D. Thesis, Universidad Autónoma Metropolitana (Mexico), **2015**.
[35] S. López-Rosa, Ph.D. Thesis, Universidad de Granada (Spain), **2010**.
[36] R. Nalewajski, *Quantum Information Theory of Molecular States*, Nova Biomedical Books, Hauppage, NY, **2016**.
[37] C. E. Shannon, *Bell Syst. Tech. J.* **1948**, *27*, 379.
[38] O. Onicescu, *C. R. Acad. Sci. Paris A* **1966**, *263*, 25.
[39] R. A. Fisher, *Proc. Cambridge Philos. Soc.* **1925**, *22*, 700 – 725; reprinted in *Collected Papers of R. A. Fisher* (Ed.: J. H. Bennet), University of Adelaide Press, Adelaide, South Australia, **1972**, pp. 15 – 40.
[40] E. Romera, J. S. Dehesa, *J. Chem. Phys.* **2004**, *120*, 8906 – 8912.
[41] R. López-Ruiz, H. L. Mancini, X. Calbet, *Phys. Lett. A* **1995**, *209*, 321 – 326.
[42] J. C. Angulo, J. Antolín, R. O. Esquivel, Chapter 6 " *Atomic and Molecular Complexities: Their Physical and Chemical Interpretations", in Statistical Complexity: applications in electronic structure*, Springer, Berlin, **2010**, pp. 167 – 214.
[43] C. Anteneodo, A. R. Plastino, *Phys. Lett.* **1996**, *223*, 348.
[44] Geometrical data, along with the wave function of the 11 novel molecules, are available on request.
[45] Gaussian 09, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels,. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian Inc. Wallingford CT **2009**.

[46] C. Soriano-Correa, R. O. Esquivel, R. P. Sagar, *Int. J. Quantum Chem.* **2003**, *94*, 165 – 172.
[47] J. M. Pérez-Jordá, A. D. Becke, E. San-Fabián, *J. Chem. Phys.* **1994**, *100*, 6520.
[48] M. Kohut, DGRID, version 4.6, **2007**; modified version from the author.
[49] Origin 2015 (OriginLab Northampton, MA, **2015**).

Manuscript received: July 19, 2016
Final Article published: November 7, 2016

THE JOURNAL OF PHYSICAL CHEMISTRY C

Article

Adsorption of Sulfonamides on Phyllosilicate Surfaces by Molecular Modeling Calculations

Misaela Francisco-Marquez, Catalina Soriano-Correa, and Claro Ignacio Sainz-Diaz

J. Phys. Chem. C, **Just Accepted Manuscript** • DOI: 10.1021/acs.jpcc.6b12467 • Publication Date (Web): 17 Jan 2017

ACS Publications

The Journal of Physical Chemistry C is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Adsorption of Sulfonamides on Phyllosilicate Surfaces by Molecular Modeling Calculations

Misaela Francisco-Márquez[1], Catalina Soriano-Correa[2], C. Ignacio Sainz-Díaz[3*]

[1]Instituto Politécnico Nacional-UPIICSA. Té 950, Col. Granjas México, C.P. 08400 México City, México

[2] Química Computacional, FES-Zaragoza, Universidad Nacional Autónoma de México, Iztapalapa, C.P. 09230 México City, México

[3] IACT, CSIC-Universidad de Granada, Av. Fuentenueva s/n, 18002-Granada, Spain.

(ignacio.sainz@iact.ugr-csic.es)

ABSTRACT

Sulfamides belong to a great group of antibiotics used during decades for human therapeutics and farm livestock. This intensive use has generated an important environmental risk due to the presence of these compounds in wastewaters and soils close to worldwide farm industries. The molecular structure and physico-chemical properties of these molecules are interesting for evaluating the molecular interactions with surfaces of clay minerals. These interactions are important to know the mobility of these compounds through soils and to design new nanomaterials for drug delivery. The molecular structure and conformational analysis of the sulfonamides, sulfamethoxypyridazine and sulfamethoxydiazine, as models of sulfonamides, have been studied with quantum mechanical calculations and force fields based on empirical interatomic potentials calculations. The adsorption of these drugs on a phyllosilicate surface of (001) plane has been investigated finding that is an exothermic process for both sulfonamides.

INTRODUCTION

Sulfamides form a great group of antibiotics widely used since decades. They are effective antimicrobial drugs for the prevention of infections in cattle, poultry, and swine (prophylaxis), to treat veterinary diseases, and to promote growth. As a result of the extensive use of sulfonamides in the animal industry, residues of these drugs in food samples are a major concern because of their contribution to the development of antibiotic resistant pathogenic bacteria. One important sub-group of these antibiotics is the sulfonamides that have a sulfonic group joined to one amino group in the molecular structure. In general, the sulfone group is joined to a 4-aminophenyl moiety and the amino group is monosubstituted by different groups

The biggest use of sulfonamides is in veterinary for intensive livestock production. These compounds and metabolites are eliminated through animal excretions and hence they are persistent in the manure liquids. This intensive use affects to the soil contamination through the use of manure as fertilizer.[1] The presence of these antibiotics and their derivatives in soil can alter the bacterial resistance to animals and humans through the soil, water and food chains.[2,3] One critical component of soil is the clay mineral fraction. This fraction has a great absorption capacity and it is one of the main responsible in the absorption of pollutants, such as, these antibiotics, in soils. On the other hand, the application of clay barriers as a filter of pollutants in localized farm area wastewaters can be a useful natural environmental treatment. Other approach to this environmental study is to consider the origin of the problem and avoid the abusive traditional use of these antibiotics and to apply release controlled system for these antibiotics. Clay minerals can be a useful natural component for

this application. Previous works of sorption of pesticides [4] and other pollutants[5] on phyllosilicates have been reported. Therefore, for all these applications the knowledge of the interactions of sulfonamides on clay surface can be interesting for enhancing their development.

In this work, the sulfonamides, Sulfamethoxypyridazine **S1** (4-amino-N-(6-methoxypyridazin-3-yl)benzenesulfonamide) and Sulfamethoxydiazine **S2** (sulphameter or sulfamethoxypirimidine, or 4-amino-N-(5-methoxy-2-pyrimidinyl)benzenesulfonamide) are studied. These compounds are widely used in therapeutics and represent two families of sulfonamides, sulfonamido-pyrimidines and sulfonamido-pyridazine. These molecules are similar with the same substituents, an p-aminobenzenic moiety joined to the sulfoxy group and a methoxy group joined in a *para* position of the heterocyclic ring. The only difference is the relative disposition of N atoms in the heterocyclic ring (Fig. 1). The bioactivity of **S2** is twice higher than that of **S1**. [6,7] These molecules adopt different conformations in their crystal polymorphs.[8,9] There are several lacks of knowledge in understanding properties of these molecules, however no study at molecular level of these pharmaceutical drugs has been found.

One of the aims of this work is to explore the interatomic interactions between these sulfonamides and the surface of phyllosilicate minerals by means of atomistic calculations.

Models and Computational Methodology

Ab initio quantum chemical calculations of the organic molecules were performed using the density functional theory (DFT) and Hartree-Fock methods. The electronic

structure of these molecules was calculated with a triple-ζ basis set with polarization functions for all atoms, including H atoms, augmented with diffuse functions with the hybrid functional BHandHLYP/6-311+G** [10,11] and the M06-2X functional with double amount of nonlocal exchange.[12] We included also electron correlation effects by means of the second-order Moeller-Plesset approximation at the all electron MP2 level (MP2/6-311G**) as implemented in the Gaussian09 program package.[13] All geometries were fully optimized at these levels using the Berny analytical gradient method. No geometry constraint is applied on the surface and reactants atoms. Normal mode analyses were carried out at the same level to confirm the nature of the various stationary points, finding only positive eigenvalues for minima.

Since the quantum mechanical methods have difficulties to describe the dispersive interactions and weak adsorption interactions, empirical interatomic potentials have been also used for comparison and for scanning the molecules across the pyrophyllite (001) surface, in order to determine the energy landscape. These empirical potentials were used within Force Fields (FF) previously optimized by Heinz et al. [14] and also recently by us named PCFFH. [15] These FF yielded good results in interactions of organics with surfaces of phyllosilicates.[16] The atomic charges were taken from Heinz et al. [14] for the phyllosilicate surface and from charges calculated at ab initio MP2/6-311G** level and associated to the electrostatic field (ESP) with the method of Merz and Kollman [17] for each compound and conformer. Different calculation conditions were tuned with respect to the 9-6 Lennard-Jones potentials and with the van der Waals and Coulomb interactions. The molecular geometry of organic molecules does not vary with different parameters for van der Waals and coulomb interactions within the same FF. We found that the van der Waals atom based interactions with a cutoff of 15.5 Å and the Ewald summation for Coulomb interactions yielded the best results and we will use these conditions in this work. For these classical calculations, the Discover program was used within the Material Studio package. [18] Molecular dynamics simulations were performed with this FF in the NVT ensemble with 1fs steps for 5 ps at 298 K.

In all cases the adsorption energy has been calculated in the usual way: $U_{adsorption} = U_{molecule+surface} - (U_{molecule} + U_{surface})$, where U is the internal energy of the system.

A periodic model of the crystal structure of phyllosilicates was generated and used for our calculations. Experimental atomic coordinates and lattice cell parameters of pyrophyllite were taken for the crystal structure model.[19] All octahedral cations are Al and all tetrahedral cations are Si and no interlayer cation was included. The H atom coordinates were included manually and optimised previously by means of interatomic empirical potential calculations. [20] This structure is dioctahedral in the trans-vacant crystal form, and completely dry in our simulations. Although the experimental adsorption used to be performed in wet conditions, this dry model is useful for focussing this investigation on the interaction between sulphonamides and mineral surface avoiding additional interactions that will be studied in further works. In order to generate a surface along the (001) plane, an additional vacuum space of 23 Å along the *c* axis was created. However, a unit cell is too small for the adsorption calculations because the axis *a* value is too small (5.2 Å) and intermolecular interactions between adsorbates from vicinal cells cannot be neglected. Hence a 4x2x1 supercell model was generated in a balance between the minimum inter-adsorbate interaction and computational effort.

For models of sulphamide, the molecular structures were taken from crystallographic data of Sulfamethoxypyridazine (**S1**)[8,21]and Sulfamethoxydiazine (**S2**)[9].

RESULTS AND DISCUSSION

Molecular structure of organics

In order to compare several quantum-mechanical methods, the molecular structure of sulfonamides (Figure 2) was optimized at BHandHLYP/6-311+G**, MP2/6-311G** and M06-2X/6-311+G** levels with GAUSSIAN09. These results were compared with those calculated with empirical interatomic potentials using the PCFFH force field.

Conformational Analysis

To choose the initial geometry of one molecule from crystal structures of these compounds, several conformers exist in these crystals. In the crystal form of **S1** from Basak et al. [8], the molecules have the methoxy group oriented to the same side of the heterocyclic N atoms, that we named *syn* conformation. Besdies, the N-H bond can be oriented to the same side, that we named *cis* conformation, or the opposite side, that we named *trans* conformation, with respect to the methoxy group. On the other hand, in the molecules in the crystal form of **S1** from Haridas et al.,[21] the methoxy group can be also oriented to the opposite side with respect to the heterocyclic N atoms, named as *anti* conformation. Hence this molecule can have at least four conformations: *syn-cis*, *syn-trans*, *anti-cis*, and *anti-trans*. However, each one of these conformers can have also two possible conformations due to the relative disposition of the aromatic rings each other. Maintaining the heterocyclic N atoms as a reference point to define conformers in **S1** and considering the disposition of these N atoms in the left part of the heterocyclic ring viewed from methoxy group, the aromatic ring of the aniline moiety can be oriented in the up-side or in the down-side with respect to a horizontal disposition of the heterocyclic ring. Therefore, eight conformations can be considered for this molecule: *syn-cis-up*, *syn-cis-down*, *syn-trans-up*, *syn-trans-down*, *anti-cis-up*, *anti-cis-down*, *anti-trans-up,* and *anti-trans-down*. All these conformations were fully optimized as isolated molecules in gas phase (figure 2). These conformers are

characterized by the dihedral angles, H_3C-O-C-N, H-N(S)-C-N, S-N-C-N, and C-N-S-C. In the *syn* conformers the H_3C-O-C-N angle is close to 0º and in the *anti* one this angle is close to 180º. In the *cis* conformers, the H-N(S)-C-N angle is close to 0º and in the *trans* forms, this angle is close to 180º. In the *cis* conformers, the planes of aromatic rings form an angle smaller than 90º and they are slightly displaced where the aminobenzenic ring is closer to the heterocyclic N atoms probably owing to stabilizing π-π interactions. Whereas in the *trans* conformers, these rings are more separated and twisted each other, probably due to the interactions between the H atoms of the aminobenzenic ring and the heterocyclic N atom.

In all calculation levels used, the most stable conformer is the *syn-cis* one. The conformers *trans* are less stable than the *cis* ones with an energy difference of 3.05-3.97 kcal/mol in the group of *syn* conformers and 3.25-4.61 kcal/mol for the *anti* conformers group. The lower energy of the *cis* conformers can be explained by a higher stabilizing π-π interaction between the aromatic rings that are closer than in *trans*. The conformation *syn*/*anti* yields a greater energy differences being the *syn* ones 5.53-6.02 kcal/mol more stable than the *anti* ones for the *cis* conformers group. This energy difference is similar in the *trans* conformers group (Table 1). The *anti* conformers have higher energy due to the repulsion between the lone pair electrons of the O and N atoms (Fig. 2).

On the other hand, the conformers *up/down* have similar energy within the same conformers group. This happens in all conformation groups and calculation levels being the energy difference lower than 0.22 kcal/mol. However, the transition from the *up* to the *down* conformers is very complex depending at least on the S-N-C-N and C-N-S-C torsional angles. A scanning exploration of the rotation of the θ(S-N-C-N) dihedral angle in the *syn*-

cis conformers group with 30º steps with full optimization in each point showed a energy barrier close to 3.7 kcal/mol between two minima with similar energy where the atoms S-N-C-N are co-planar (Figure 3a). The conformers from aromatic ring disposition show similar intramolecular interactions in all conformers and the energy differences between the *up* and *down* forms are very low. The existence of these different conformers can be due to the intermolecular interactions in the crystal packing or a high-energy barrier in the rotation of these rings. A similar scanning calculation of the rotation of the θ(C-S-N-C) dihedral angle in steps of 30º confirmed the existence of only these conformers with an energy barrier around 3.5 kcal/mol (at BHandHLYP/6-311+G** level) (Figure 3b).

In the **S2** molecule, the heterocyclic ring is more symmetric than in **S1** and the methoxy group can adopt the *syn* and *anti* conformations with respect to the central N-H bond (Figure 4). The energy differences between these conformers are smaller than in **S1** because the effect of methoxy on the NH group is weaker in **S2**. The most stable conformer is the *syn* one being only 0.05-0.09 kcal/mol more stable than the *anti* one. Maintaining the NH group with the heterocyclic ring as a reference point to define conformers in **S2** and considering the disposition of the NH group in the left part of the heterocyclic ring viewed from methoxy group, the aromatic ring of the aminobenzenic moiety can be oriented in the up-side or in the down-side with respect to a horizontal disposition of the heterocyclic ring (Figure 4). Nevertheless, the energy difference between the *up* and *down* conformers is less than 5 cal/mol (Table 2).

An exploratory scanning of rotation of the C-C-O-CH_3 angle showed a low energy barrier, around 1.6 kcal/mol, between the coplanar *syn/anti* conformers (Figure 5a).

However, a scanning exploration of the S-N(H)-C-N rotation in steps of 30º showed an energy barrier around 8.0 kcal/mol (at BHandHLYP/6-311G** level) in the interaction of the NH group and the arylsulfoxy moiety with the heterocyclic ring and only one conformer is possible (Fig. 5b). Hence, additional rotations will be necessary to reach the *anti* conformer.

The molecules **S1** and **S2** are positional isomers, where the differences are in the relative positions of the heterocyclic N atoms. Then, the relative energy was calculated finding that **S2** is 12.68 kcal/mol more stable than **S1** at MP2/6-311G** level.

Considering the dipole moment of these conformers of **S1**, the most polar conformer is the *syn-trans-up* one (the highest μ in Table 1). The transition from *syn* to *anti* did not produce a significant change of μ, however the transition from *cis* to *trans* produced a larger μ difference. The changes from *up* to *down* produced significant variations in μ only in the *trans* conformers group, being the *syn-trans-down* conformer the less polar one. This can indicate that in a non-polar medium the population of this last conformer would be favored. In **S2** the conformer *anti* has a higher μ than the *syn* one. The calculation levels used have yielded similar relative values of μ. However, the μ value of the *anti-trans-down* conformer of **S1** calculated at MP2 level was too high with respect to the rest of calculations. This high value can be explained taking into account the geometry of the conformer where at MP2 level the aromatic rings approach each other and the O atoms of the sulfoxy group are in one side of the molecule polarizing this conformer.

Molecular Geometry

The optimization of the most stable conformers of these sulfonamide molecules yielded a molecular geometry consistent with the experimental bond distances obtained from crystal structures (Tables 3 and 4). The differences observed between experimental and theoretical values can be due to that the experimental molecule is into a crystal structure where the intermolecular interactions can change some bond lengths or bond angles, whereas the calculated models are isolated molecules without intermolecular interactions. Nevertheless, some variations are observed between conformers.

In **S1** compound, the *trans* conformers group have a slightly longer S-N bond length (1.681 Å at BHandHLYP/6-311+G** level) and shorter SN-C bond length (1.398 Å at BHandHLYP/6-311+G** level), and a larger S-N-C bond angle (123.6º at BHandHLYP/6-311+G** level) than the *cis* conformers group (1.669 Å, 1.414 Å, and 114.3º, respectively at BHandHLYP/6-311+G** level), probably due to repulsions between the electrons of the sulfoxy O atoms and the heterocyclic N atoms. The H atom of the SN(H)C moiety is slightly out of the coplanar disposition with the heterocyclic ring, where the dihedral angle (H-N-C-N) is 8-12º and 162-178º in the *cis* and *trans* conformers, respectively. In the *cis* conformers the p-aminobencenic ring is more twisted from the heterocyclic plane with a larger (S-N-C-N) torsional angle than in *trans*. In the *trans* conformers the planes of the aromatic rings form a larger angle than in *cis* conformers, and the heterocyclic N atoms are twisted and more oriented towards the aminobenzenic ring with a shorter CH…N non-bonding distance. The *anti* conformers have a shorter N-N bond (1.309 Å at BHandHLYP/6-311+G** level), a longer NC-OCH_3 bond (1.337 Å at BHandHLYP/6-311+G** level), and a smaller N-C-OCH_3 bond angle (114º at BHandHLYP/6-311+G** level) than the *syn* group (1.325 Å, 1.332 Å, and 120º, respectively at BHandHLYP/6-311+G** level). The methoxy group is

coplanar in all conformers, being close to 0º and 180º in the *syn* and *anti* conformers, respectively. No significant differences in the bond lengths and bond angles are detected between *up* and *down* conformers only in the S-N-C-N torsion angle.

The main geometrical features of these molecules calculated with quantum mechanical methods at several levels and also with the PCFFH force field were compared (Table 3 and Table S1). The bond distances calculated with PCFFH are consistent with those calculated with other methods. The tendency of the main geometric features differences between conformers is similar for all calculations levels used. On the other hand, no relationship between the relative energy and the intramolecular non-bonding interactions was found between conformers.

In **S2** sulfonamide, no significant difference in the main geometrical features was observed between conformers, explaining the low energy differences (Table 4). The structure of the most stable conformer especially the relative disposition of the aromatic rings is different that in **S1**, being consistent with previous calculations of similar sulfadiazine **S2** derivative without methoxy group.[22]

Mineral Surface and adsorption of sulphonamides

The periodic model of pyrophyllite was fully optimized with PCFFH obtaining cell parameters (a = 5.24 Å, b = 9.09 Å, c = 9.32 Å, α = 90.5º, β = 100.8º, γ = 89.9º) close to the experimental values (a = 5.16 Å, b = 8.97 Å, c = 9.35 Å, α = 91.2º, β = 100.5º, γ = 89.6º).[23] The averaged value of Si-O bond length is 1.67 Å that is slightly longer than the experimental value of 1.62 Å. The average value of Al-O and Al-OH bond lengths are 1.94

and 1.92 Å reproducing the experimental values.[23] The OH bond length is 0.979 Å with an inclination angle ρ of 32º that is consistent with experimental data. From this optimized structure, a 4x2x1 supercell was generated and a surface on the (001) plane was created with a *c* parameter value of 30 Å.

The molecule of sulfonamide optimized at BHandHLYP/6-311+G** was added at a distance of 3 Å from the basal oxygens of the pyrophyllite surface in a parallel position with respect to the surface. Preliminary molecular dynamics simulations of the most stable conformer of **S1**, *syn-cis-down*, located over the pyrophyllite (001) surface were performed with PCFFH. During the equilibrium runs of the simulations, the sulphonamide molecule was surfing along the surface maintaining an aromatic ring parallel to the surface and no specific disposition was preferential. Several adsorption complexes were sampled from this simulation trajectory and used as initial steps of optimizations. The adsorption complex, surface + adsorbate, was fully optimized at constant volume with PCFFH (lattice cell parameters: a = 10.5 Å, b = 9.10 Å, c = 30 Å, α = 90.1º, β = 100.8º, γ = 89.9º). Several optimized adsorption complexes were obtained where **S1** adopts different orientations: P1, with the aminobenzene ring parallel to the surface; P2 with the heterocyclic ring parallel to the surface; P3 in apical position, where the aromatic rings are not parallel to surface and the amino and methoxy groups are oriented to the surface; P4 with the heterocyclic ring parallel to the surface but slightly twisted, and P5 with the most part of the adsorbate molecule is open and parallel to the mineral surface (Figure 6).

Placing the most stable conformer of **S1** (*syn-cis-down*) on the mineral surface with the aminobenzene ring parallel to surface, the optimization of the adsorption complex

maintained the aminobenzene ring parallel to surface forming a P1 complex with the H atoms of amino group oriented to the basal tetrahedral O atoms of the phyllosilicate surface forming H bonds *d*(NH…OSi) = 1.71, 2.38 Å. The optimization of the adsorption complex of **S1** with the heterocyclic ring in a parallel orientation with respect to the mineral surface maintained the same orientation yielding a P2 complex with a strong interaction of the H atom of SNH group with a basal O atom of surface *d*(NH…OSi) = 2.12 Å. The methoxy group has a electrostatic interaction of one H atom with a basal O atom of surface *d*(NH…OSi) = 2.60 Å and the heterocyclic N-N moiety are placed on the centre of the tetrahedral pseudohexagonal ring of tetrahedra minimizing the repulsion with the surface O atoms.

Starting with an initial geometry with **S1** placed in a perpendicular orientation with respect to the surface with the amino group oriented to the surface, the optimization of adsorption complex maintained a perpendicular orientation of the adsorbate forming the P3 type of complex with the H atoms of amino group oriented to the surface O atoms *d*(NH…OSi) = 2.27, 3.19 Å. On the contrary, placing the **S1** in a perpendicular orientation with the surface but in the opposite sense with the sulfoxy and SNH oriented to the surface forming a strong H bond between the SNH H atom and the surface O atom, *d*(NH…OSi) = 1.30 Å, the optimization yielded a adsorption complex with the adsorbate with the heterocyclic ring quasiparallel to the surface forming a P4 type of complex with the SNH H atom forming a strong interaction with surface O atoms *d*(NH…OSi) = 2.00 Å. The higher adsorption energy of the **S1** *syn-cis-down* conformer corresponds to the P1 and P4 adsorption complex. A general tendency of higher adsorption energy with a shorter H bond between adsorbate and mineral surface is observed (Figures in Supporting Information).

However, no good linear correlation was found indicating that other interactions contribute to the adsorption energy.

For the second most stable conformer of **S1**, *syn-trans-up*, the most stable adsorption complex has the adsorbate in a P5 configuration with H atoms of amino groups oriented to the surface, *d*(NH…OSi) = 2.35 Å., and the sulfoxy O atoms away from surface. Both aromatic rings are open with an $H_2N…S…COCH_3$ angle of 122.2º and both rings are almost parallel to the surface. This configuration is obtained by optimization of a similar one and also after optimization of the initial complex with the adsorbate in a perpendicular configuration with a side of both aromatic rings oriented along surface with the NH group oriented to surface (heterocyclic N atoms away from surface). Starting with the adsorbate in a P1 disposition on surface, the optimization maintained this P1 configuration. Placing the adsorbate in a P1 configuration with a different orientation respect to the surface, the optimization of the adsorption complex yielded the P2 configuration. The optimization of an initial adsorption complex in P2 maintained the same configuration P2. The adsorption energies of all adsorption complex configurations explored for this conformer are similar within the range of 8.77-10.95 kcal/mol, being the most stable the P5 one and the rest are in the same level (Table 5).

The adsorption of the sulfamide **S2** was similar. The most stable adsorption complex has the adsorbate in a configuration P5, where both aromatic rings are almost parallel to the mineral surface with the amino and methoxy H atoms oriented to the surface *d*(NH…OSi) = 2.197 Å, and *d*(OCH…OSi) = 2.680 Å. The optimization of the adsorption complex with **S2** in a perpendicular configuration with the heterocycle and SNH group oriented to the surface

yielded also the configuration P5. Starting with a configuration P1, with the aminobenzene ring parallel to surface, but with amino H atoms oriented away from surface, the optimization of the adsorption complex maintained this configuration P1 with the amino H atoms oriented to the surface, *d*(NH…OSi) = 2.56 and 3.15 Å. With an initial adsorption complex with the adsorbate in a configuration P2, with the heterocycle parallel to the surface, the optimization maintained the configuration P2. From a configuration P1 with the amino H atoms oriented to surface and one sulphoxy O atom oriented to a Si atom of surface, the optimization obtained a adsorption complex with a configuration type P5 with both aromatic rings twisted.

The adsorption energy for all adsorption complexes of **S1** and **S2** are similar in the range of –(9.5-11.5) kcal/mol (Table 5). This value is similar that obtained previously in calculations of the adsorption of dioxins on phyllosilicates (-11.5 kcal/mol)[24] and the experimental data of the sorption of aromatics on siloxane surfaces (-10 kcal/mol).[25] In general, the most stable adsorption complexes are with the adsorbate in the configuration P5, where both aromatic rings are open and oriented to the surface.

CONCLUSIONS

The conformational analysis of the sulfamides presented in this work has allowed explaining the different conformers found in the crystal polymorphs of these drugs. The energy differences between conformers in gas phase are very low however the rotational energy barrier is significant.

Both sulfonamides, **S1** and **S2**, are likely to be adsorbed on the phyllosilicate (001) surface through an exothermic process in the same level than other organics like dioxins. This is the first work related with adsorption of these antibiotics on phyllosilicates in a simple model of pyrophyllite. This work presents the interaction of two models of sulfonamides in a neutral form on the surface of pyrophyllite as a model of phyllosilicates that are components of soils. These sulfonamides belong to the antibiotics group most used worldwide. Sulfonamides are typical amphoteric compounds and the neutral form is the most predominant one to be considered in standard soils. The proportion of neutral form will depend on the kind of sulfonamide molecule and the properties of the soil. Further investigations will be performed including new variables, like cation substitutions effect in mineral and the presence of solvents and cations in the interlayer space. Hence, phyllosilicates can be excellent traps for these drugs, especially when they are polluting soils and wastewaters of livestock industries. Besides, the results presented in this work open the possibility to explore the use of phyllosilicates as nanocarriers of sulfonamide antibiotics for a possible controlled release of these drugs.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Ignacio.sainz@iact.ugr-csic.es

ORCID:

C. Ignacio Sainz-Díaz:

0000-0001-8612-3826

Notes

The authors declare no competing financial interest.

SUPPORTING INFORMATION

A table with additional calculated geometrical features and an additional figure are published separately.

ACKNOWLEDGEMENTS

Authors are thankful to A. Hernández-Laguna for the help in computing facilities, to the Andalusian Government projects (RNM1897 and RNM363) and the MINECO and FEDER project (FIS2013-48444-C2-2-P) for the financial support, to the Supercomputational Center of the Granada University (UGRGRID), to the Centro Técnico de Informática del CSIC, to the Mexican agency CONACYT for a grant, and to the Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) of the Universidad Nacional Autónoma de México (UNAM) for allocation of computer time on the Miztli supercomputer.

[1] Wehrhan, A.; Kasteel, R.; Smunek, J.; Groeneweg, J.; Vereecken, H. Transport of sulfadiazine in soil columns-Experiments and modelling approaches. *J. Contaminant Hydrol*. **2007**, *89*, 107-135.
[2] Wang, Z. H.; Zhang, S. X.; Nesterenko, I. S.; Eremin, S. A.; Shen, J. Z. Monoclonal Antibody-Based Fluorescence Polarization Immunoassay for Sulfamethoxypyridazine and Sulfachloropyridazine. *J. Agric. Food Chem.* **2007**, *55*, 6871-6878.
[3] Sukul, P.; Spiteller, M. Sulfonamides in the Environment as Veterinary Drugs. *Rev. Environ. Contam. Toxicol.* **2006**, *187*, 67–101.
[4] Azarkan, S.; Peña, A.; Draoui, K.; Sainz-Díaz, C. I. Adsorption of two fungicides on natural clays of Morocco. *Appl. Clay Sci*. **2016**, *123*, 37-46.
[5] Muñoz-Santiburcio, D.; Ortega-Castro, J.; Sainz-Díaz, C. I.; Huertas, F. J.; Hernández-Laguna, A. Theoretical study of the adsorption of 2-nitro-1-propanol on smectite surface models. *J. Mol. Struct. (THEOCHEM)* **2009**, *912*, 95-104.
[6] Wolff, M. E. Burger's Medicinal Chemistry, Part II. 4th ed.; Wiley & Sons: New York, Vol. 2, Chapter 13, pp 1-40, 1979.
[7] Mengelers, M. J. B.; Hougee, P. E.; Janssen, L. H. M.; Van Miert A. S. J. P. A. M. Structure-activity relationships between antibacterial activities and physicochemical properties of sulfonamides *J. Vet. Pharmacol. Therap.* **1997**, *20*, 276-283.
[8] Basak, A. K.; Mazumdar, S. K.; Chaudhuri, S. Structure of N'-(6-Methoxy-3-pyridazinyl)sulfanilamide (Sulfamethoxypyridazine) *Acta Cryst. C*. **1987**, *43*, 735.
[9] Guiseppetti, G., Tadini, C., Bettinetti, G.P., Giordano, F. 2-Sulfanilamido-5-methoxypyrimidine. *Cryst.Struct.Commun.* **1977**, *6*, 263-264.
[10] Becke, A. D. A new mixing of Hartree–Fock and local density□functional theories. *J. Chem. Phys.* **1993**, *98*, 1372.
[11] Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Frisch, M. J.; Frisch, A. GAUSSIAN 98 User's Reference; Gaussian Inc.: Pittsburgh, PA, 1998.
[12] Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoret. Chem. Acc*. **2008,** *120*, 215–241.
[13] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Chesseman, J.R., Zarzewki, V.G., Montgomery, J.A., Stratmann, R.E., Burant, J.C. et al. Gaussian 03 (RevisionA.1), Gaussian, Inc., Pittsburgh, PA, 2004.
[14] Heinz, H.; Koerner, H.; Anderson, K. L.; Vaia, R. A., Farmer, B. L. Force Field for Mica-Type Silicates and Dynamics of Octadecylammonium Chains Grafted to Montmorillonite. *Chem. Mater*. **2005**, *17*, 5658-5669.
[15] Sainz-Díaz, C.I.; Francisco-Márquez, M.; Vivier-Bunge, A. Adsorption of polyaromatic heterocycles on phyllosilicate surfaces by means of different theoretical approaches. *Environ. Chem*. **2011**, *8*, 429-440.
[16] Heinz, H.; Vaia, R. A.; Farmer, B. L. Interaction energy and surface reconstruction between sheets of layered silicates. *J. Chem. Phys*. **2006**, *124*, 224713.
[17] Besler, B.H., Merz Jr., K. M., Kollman, P. A. Atomic charges derived from semiempirical methods, *J. Comp. Chem. 1990*, **11**, 431-439.
[18] Accelrys Software inc., Materials Studio Release Notes, Release 5.0, San Diego: Accelrys Software Inc., 2009.
[19] Wardle, R.; Brindley, G.W. The crystal structure of pyrophyllite, 1Tc, and of its dehydroxylate. *Am. Mineral*. **1972**, *57*, 732-750.
[20] Sainz-Díaz, C. I.; Palin, E. J.; Dove, M. T.; Hernández-Laguna, A. Monte Carlo simulations of ordering of Al, Fe, and Mg cations in the octahedral sheet of smectites and illites. *Am. Mineral.* **2003**, *88*, 1033-1045.
[21] Haridas, M., Singh, T. P. Crystal and molecular structure of sulfamethoxypyridazine. *Indian J. Chem. ,Sect.A:Inorg.,Bio-inorg., Phys.,Theor.Anal.Chem*. **1986**, *25*, 707-713.
[22] Ogruc-Ildiz, G.; Akyuz, S., Ozel, A. E. Experimental, ab initio and density functional theory studies on sulfadiazine. *J. Mol. Struc.* **2009**, *924-926*, 514-522.
[23] Lee, J. H.; Guggenheim, S. Single crystal X-ray refinement of pyrophyllite-1Tc. *Am. Mineral*. **1981**, *66*, 350-357.

[24] Austen, K. F.; White, T. O. H.; Marmier, A.; Parker, S. C.; Artacho, E.; Dove, M. T. Electrostatic versus polarization effects in the adsorption of aromatic molecules of varied polarity on an insulating hydrophobic surface. *J. Phys. Condens. Matter*. **2008**, *20*, 035215.

[25] Pelmenschikov, A.; Leszczynski, J. Adsorption of 1,3,5-Trinitrobenzene on the Siloxane Sites of Clay Minerals: Ab Initio Calculations of Molecular Models. *J. Phys. Chem. B* **1999**, *103*, 6886-6890.

Captions of Figures

Figure 1. Molecular structure of the sulfonamides **S1** (a) and **S2** (b).

Figure 2. Conformers of **S1**: *syn-cis-up* (a), *syn-cis-down* (b), *syn-trans-up* (c), *syn-trans-down* (d), *anti-cis-up* (e), *anti-cis-down* (f), *anti-trans-up* (g), and *anti-trans-down* (h), optimized at M062X/6-311+G** level.

Figure 3. Conformational analysis of the rotation of torsional angles S-N-C-N (a) and C-S-N-C (b) of **S1** *syn-cis* conformers at BHandHLYP/6-311+G** level.

Figure 4.- Conformers of **S2**: *syn-up* (a), *syn-down* (b), and *anti-down* (c).

Figure 5.- Scanning of the torsional angle H_3C-O-C-C (a) and S-N-C-N (b) of **S2** at BHandHLYP/6-311+G** level.

Figure 6. Adsorption complexes of **S1** on pyrophyllite (001) surface optimized with PCFFH with the adsorbate conformer *syn-cis-down* in configurations: P1 (a); P2 (b); P3 (c); and P4 (d); and the conformer *syn-trans-up* in a configuration P5 (e). Adsorption energies have been included in brackets below each model and they are in kcal/mol (Table 5).

Figure 7. Adsorption complexes of **S2** on pyrophyllite (001) surface optimized with PCFFH with the adsorbate in configurations P5 (a); P1 (b); and P5 twisted. Adsorption energies are in brackets below each model in kcal/mol (Table 5).

Table 1.- Relative energy (kcal/mol) and dipolar moment (Debyes) of the conformers of **S1** respect to the most stable one.

conformers	$E^{a,b}$	$E^{a,c}$	$E^{a,d}$	μ^{b}	μ^{c}	μ^{d}
syn-cis-down	0.00	0.14	-	6.21	6.13	-
syn-cis-up	0.0006	0.00	0.00	6.39	6.63	7.61
syn-trans-up	3.05	3.50	3.97	6.85	6.92	6.78
syn-trans-down	3.11	3.56	-	4.86	4.67	-
anti-cis-up	6.02	5.68	-	6.41	6.74	-
anti-cis-down	6.00	5.90	5.53	6.45	6.53	7.42
anti-trans-down	9.27	9.45	10.14	5.16	5.11	9.47
anti-trans-up	9.29	9.41	-	6.70	6.73	-

[a] Energy with zero-point correction. [b] Calculated at BHandHLYP/6-311+G(d,p) level. [c] Calculated at M06-2X/6-311+G(d,p) level. [d] Calculated at MP2/6-311G(d,p) level.

Table 2.- Relative energy (kcal/mol) and dipolar moment (Debyes) of the conformers of **S2** respect to the most stable one.

conformers	$E^{a,b}$	$E^{a,c}$	$E^{a,d}$	μ^{b}	μ^{c}	μ^{d}
syn-up	0.00	0.00	0.00	6.40	6.31	10.53
syn-down	0.003	0.005	-	6.40	6.40	-
anti-down	0.05	0.06	0.09	8.34	8.22	10.03

[a] Energy with zero-point correction. [b] Calculated at BHandHLYP/6-311+G(d,p) level. [c] Calculated at M06-2X/6-311+G(d,p) level. [d] Calculated at MP2/6-311G(d,p) level.

Table 3.- Main geometrical features (bond lengths (Å) and angles (deg)) of the molecular structure of the main conformers of **S1** calculated at different calculation levels.

	Exp.	MP2/6-311G**				BHandHLYP/6-311+G**				
	Syn-cis-up [a]	*Anti-trans-down*	*Anti-cis-down*	*Syn-cis-up*	*Syn-trans-up*	*Anti-trans-down*	*Anti-cis-down*	*Syn-cis-up*	*Syn-trans-down*	*Syn-trans-up*
S-O	1.428	1.442	1.448	1.447	1.447	1.435	1.436	1.436	1.434	1.434
S-O	1.433	1.452	1.449	1.449	1.453	1.442	1.439	1.439	1.442	1.442
S-N	1.658	1.702	1.705	1.701	1.704	1.681	1.669	1.671	1.679	1.681
N-N	1.356	1.333	1.332	1.343	1.343	1.310	1.309	1.325	1.326	1.326
N-H (NH_2)	0.997	1.011	1.011	1.011	1.011	0.999	0.999	0.999	0.999	0.999
N-C (NH_2)	1.367	1.396	1.394	1.394	1.396	1.376	1.373	1.374	1.376	1.376
SN-C	1.408	1.398	1.414	1.413	1.401	1.394	1.405	1.406	1.393	1.393
SN-H	0.936	1.013	1.020	1.019	1.015	1.003	1.007	1.007	1.002	1.002
C-S	1.742	1.764	1.764	1.765	1.760	1.754	1.760	1.759	1.754	1.754
NC-O	1.350	1.350	1.350	1.344	1.345	1.339	1.337	1.332	1.334	1.334
H_3C-O	1.454	1.420	1.421	1.432	1.432	1.405	1.406	1.418	1.418	1.418
C-S-N	107.5	104.1	102.1	102.7	102.8	106.2	107.6	107.4	106.2	106.1
S-N-C	124.1	123.6	114.3	115.6	121.8	124.2	122.9	122.8	124.9	124.8
N-C-OCH_3	119.7	112.3	112.1	119.2	119.3	113.7	113.6	119.8	120.1	120.1
N-C-O-CH_3	3.3	179.5	178.5	0.7	0.7	178.5	179.0	0.6	0.6	0.6
H-N-C-N	0.0	162.1	12.4	9.4	170.9	176.9	8.8	-8.7	178.4	178.7
S-N-C-N	139.9	-21.9	-109.6	114.0	-35.0	-38.7	-127.2	126.7	34.4	-34.9
C-N-S-C	-57.5	-47.1	48.5	-49.3	83.8	85.7	-63.3	-62.7	83.6	83.5
CH…N	4.284	2.667	4.020	4.119	2.765	2.721	4.372	4.301	2.743	2.769
NH…OS	2.467	2.416	2.577	2.572	2.422	2.380	2.559	2.556	2.381	2.380
CH…OS	2.514	2.585	2.535	2.535	2.592	2.512	2.591	2.545	2.518	2.528
CH(syn)…OS[b]	2.736	2.584	2.653	2.647	2.574	2.634	2.535	2.578	2.624	2.613
CH(hetero)…OS[c]	2.519	-	2.585	2.519	-	-	2.384	2.376	-	-
H_2N…S…$COCH_3$	86.0	73.9	63.1	67.0	93.7	98.2	86.2	86.4	97.6	97.8

[a] From experimental crystallographic data.[8] [b] Aminobencenic H atom localized in the same side of the heterocyclic N atoms. [c] H atom of the heterocyclic moiety; in the *trans* conformers the N-H bond is in the middle way between this CH bond and the OS group.

Table 4.- Main geometrical features of the molecular structure of conformers of **S2** calculated at several levels and experimental values. Bond lengths are in Å and angles in (º) degrees.

	BHandHLYP/6-311+G**			PCFFH	Exp [a]
	syn-down	*Anti-down*	*Syn-up*	*Syn-up*	*Syn-up*
S-O	1.441	1.441	1.441	1.427	1.450
S-O	1.435	1.435	1.435	1.426	1.433
S-N	1.674	1.673	1.674	1.578	1.645
SNC-N(syn)[b]	1.316	1.328	1.316	1.296	1.331
SNC-N	1.324	1.311	1.324	1.291	1.315
SN-C	1.387	1.386	1.387	1.327	1.400
SN-H	1.004	1.004	1.004	0.984	0.980
C-S	1.755	1.756	1.755	1.753	1.737
CC-O	1.347	1.347	1.347	1.325	1.359
H_3C-O	1.406	1.406	1.406	1.425	1.417
C-S-N	105.7	105.6	105.7	112.4	106.5
S-N-C	125.8	126.0	125.8	116.7	124.9
C-C-O-CH_3 [c]	0.7	1.4	0.7	2.1	1.5
H-N-C-N (syn)[b]	4.7	5.0	4.7	-8.6	-26.5
S-N-C-N(syn)[d]	-29.5	-27.7	29.5	38.1	-27.9
C-N-S-C	81.7	81.2	81.7	-79.3	-57.8
CH…N	2.919	2.926	2.917	2.945	3.424
NH…OS	2.416	2.417	2.416	2.281	2.474
CH…OS	2.555	2.558	2.556	2.819	2.806
CH(syn)…OS [e]	2.578	2.574	2.577	2.677	2.555
H_2N…S…$COCH_3$	97.4	98.0	97.4	98.3	84.8

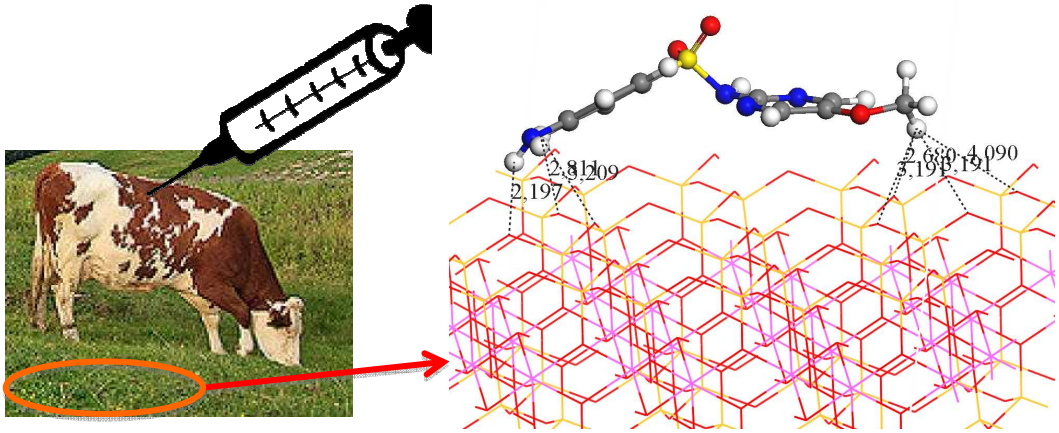
[a] From crystallographic data.[9] [b] The C-N bond at the same side of the N-H bond. [c] With the same side of methoxy group. [d] With the N atom of the same side of sulfoxy group. [e] Aminobencenic H atom localized in the same side of the heterocyclic N atoms.

Table 5.- Adsorption energy (kcal/mol) of **S1** and **S2** on the pyrophyllite surface in several configurations, calculated with PCFFH.

Adsorption complex	**S1** (*syn-cis-down*)	**S1** (*syn-trans-up*)	**S2**	
P1	-9.75	-9.00	-9.41	
P2	-7.93	-8.77	-10.55	
P3	-8.25	-	-	
P4	-9.59	-9.15	-	
P5	-	-10.66, -10.95	-11.27, -10.84[a]	

[a] The aromatic rings are parallel to the mineral surface but slightly twisted.

TOC Graphic



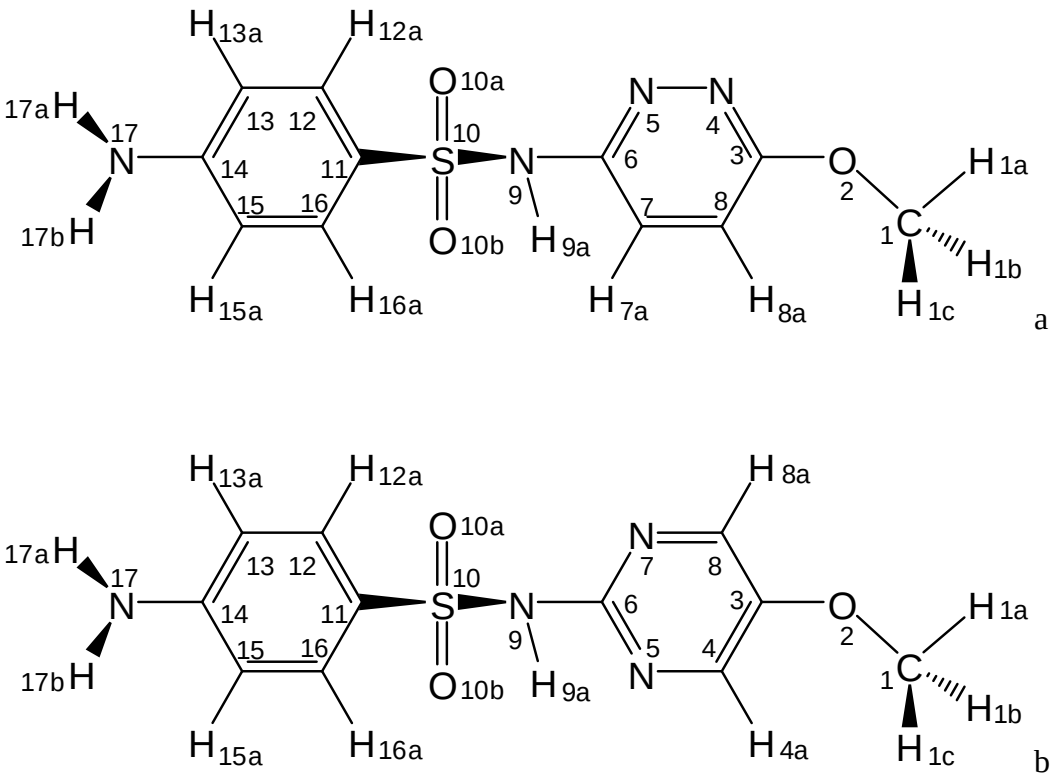


Figure 1.- Molecular structure of the sulfonamides **S1** (a) and **S2** (b).

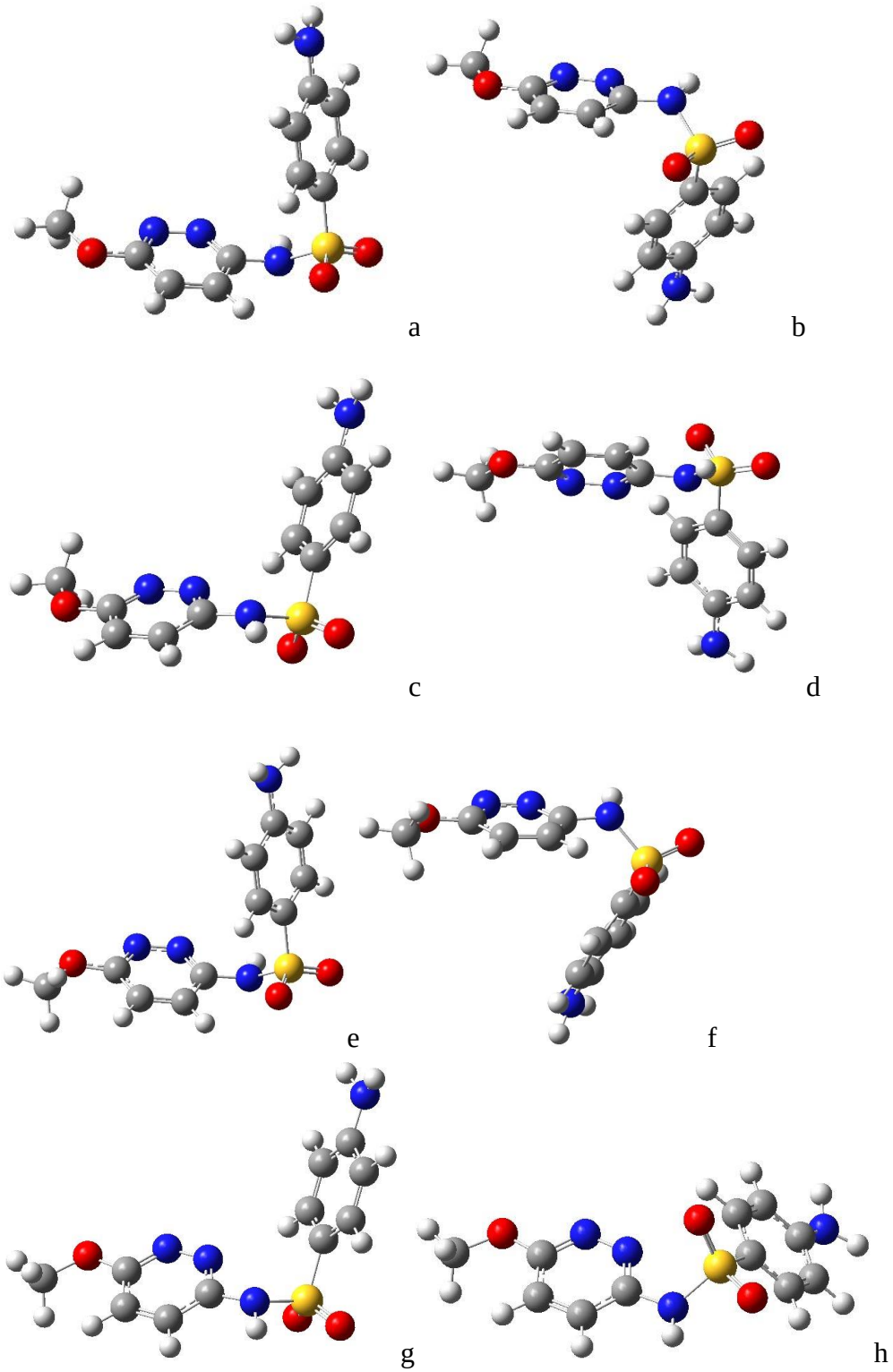
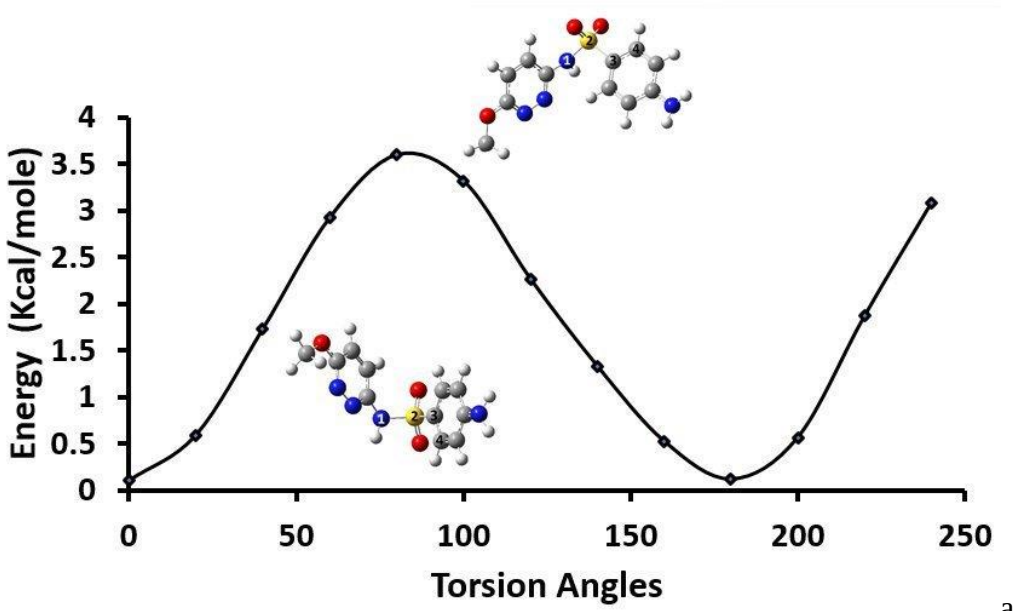
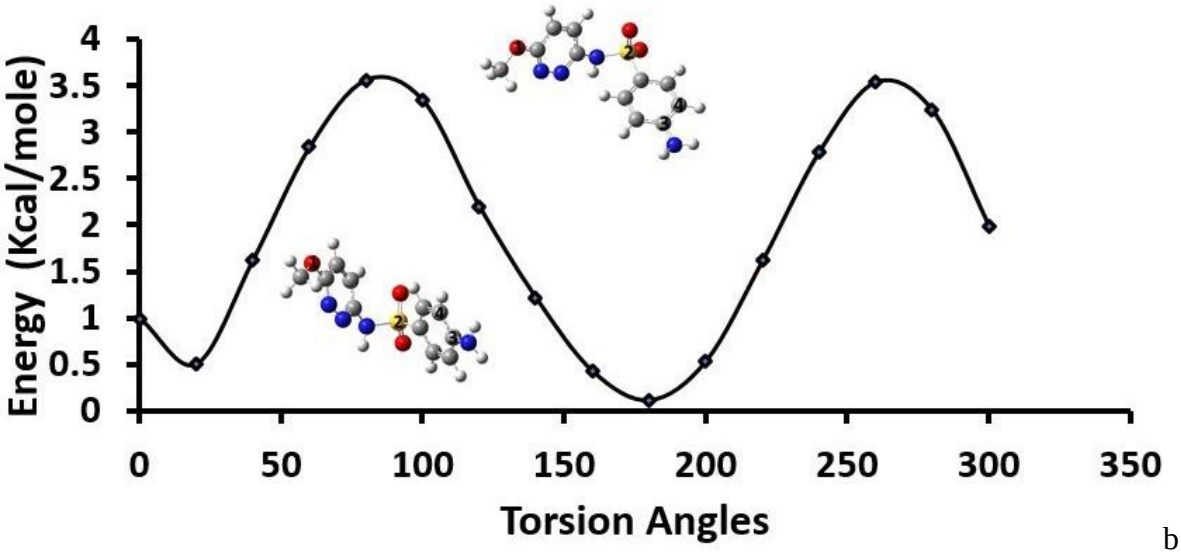


Figure 2.- Conformers of **S1**: *syn-cis-up* (a), *syn-cis-down* (b), *syn-trans-up* (c), *syn-trans-down* (d), *anti-cis-up* (e), *anti-cis-down* (f), *anti-trans-up* (g), and *anti-trans-down* (h), optimized at M062X/6-311+G** level.



a



b

Figure 3. Conformational analysis of the rotation of torsional angles S-N-C-N (a) and C-S-N-C (b) of **S1** *syn-cis* conformers at BHandHLYP/6-311+G** level.

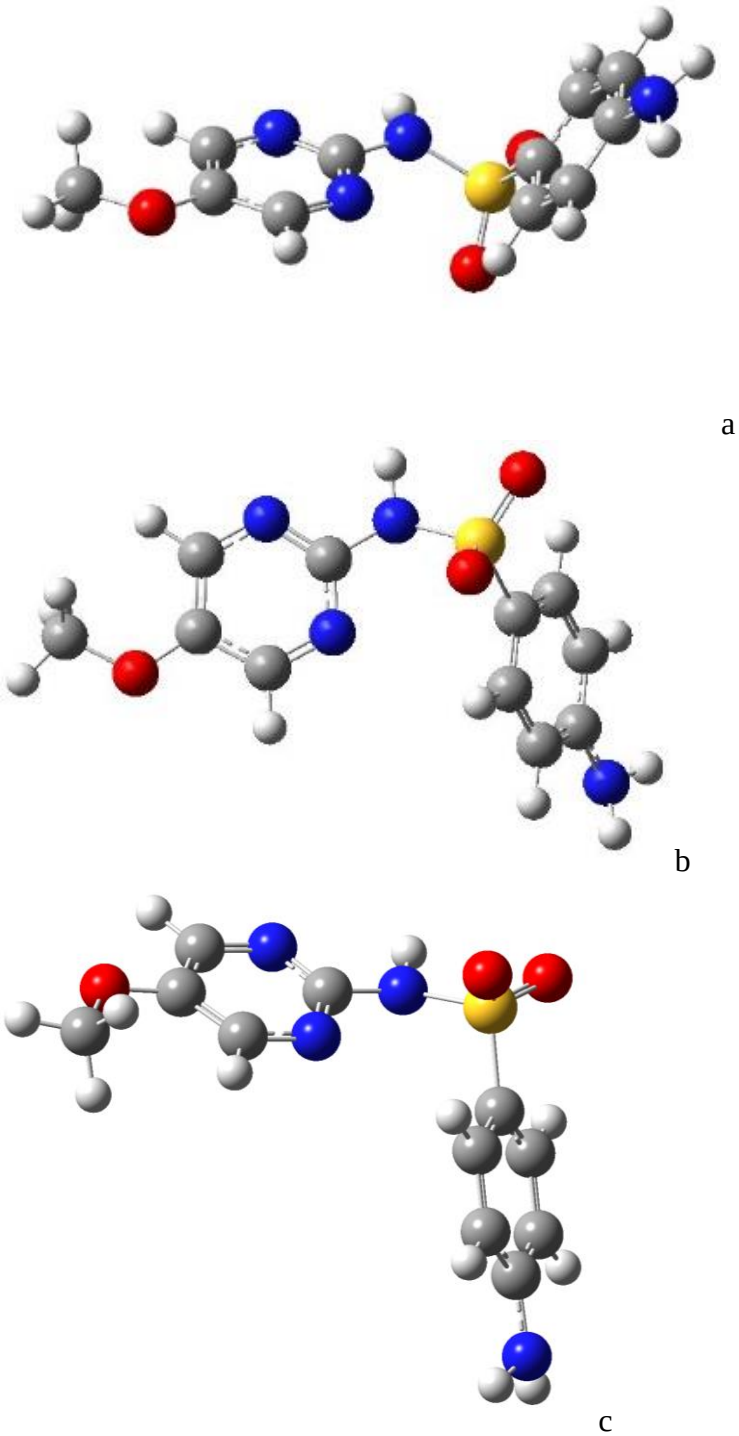
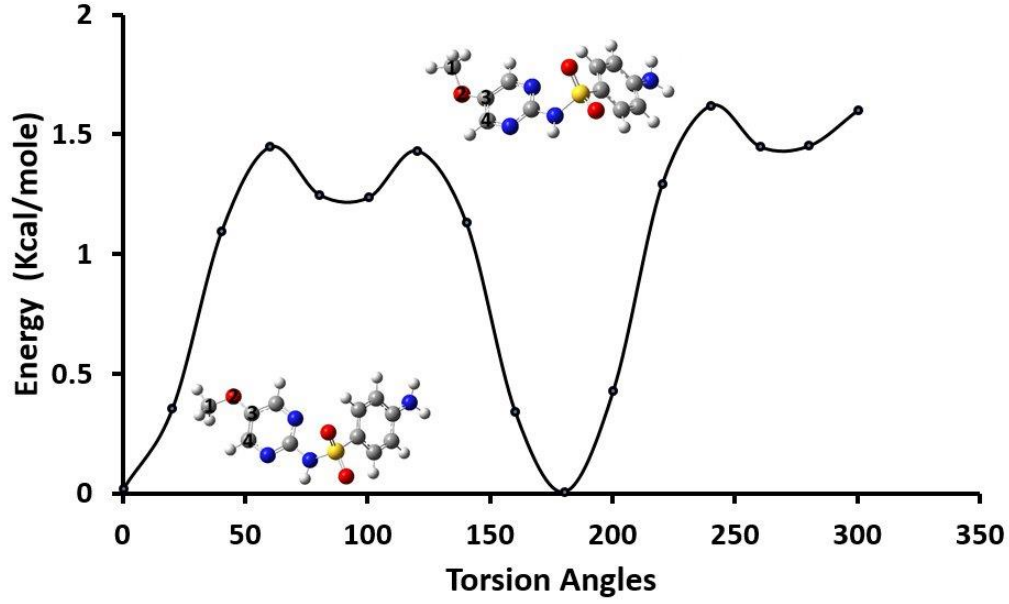
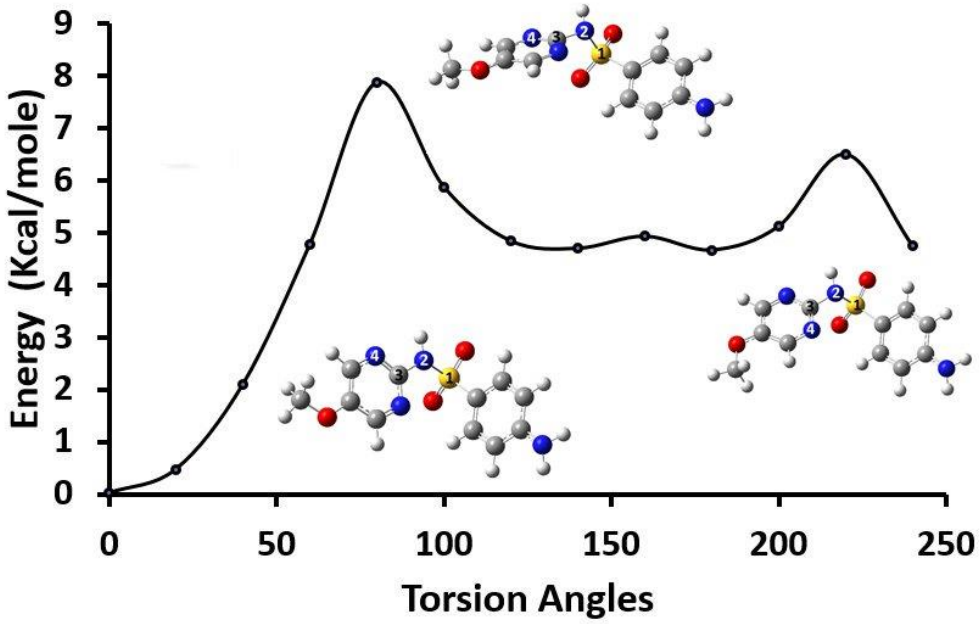


Figure 4.- Conformers of **S2**: *syn-up* (a), *syn-down* (b), and *anti-down* (c).



a



b

Figure 5.- Scanning of the torsional angle H_3C-O-C-C (a) and S-N-C-N (b) of **S2** at BHandHLYP/6-311+G** level.

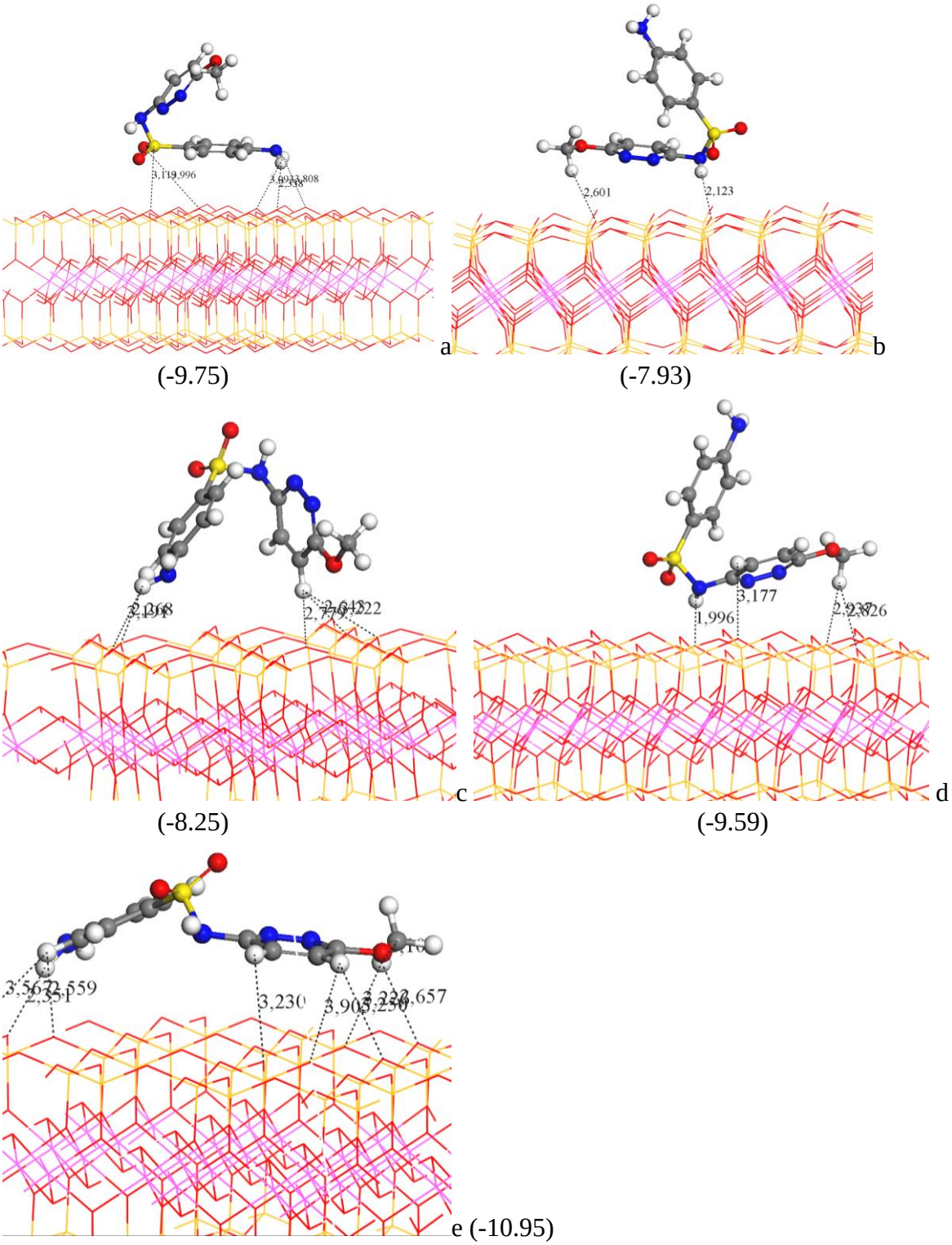
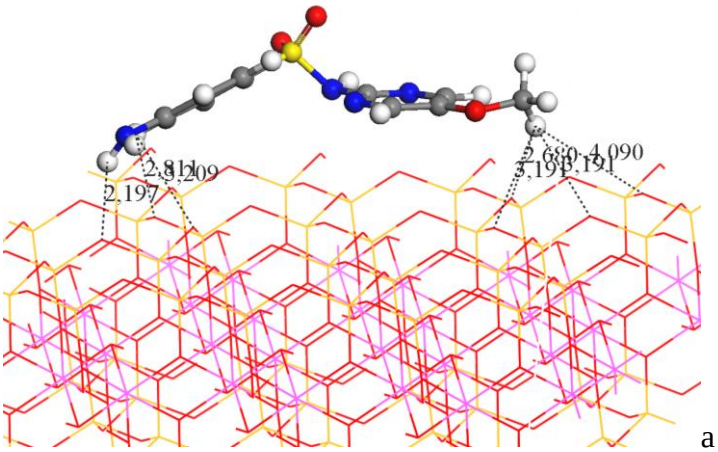
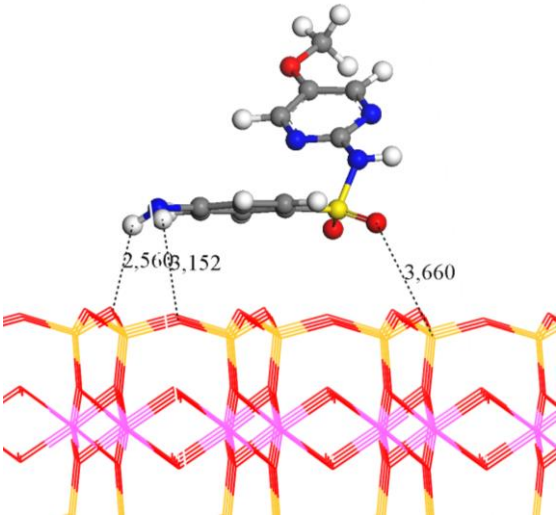


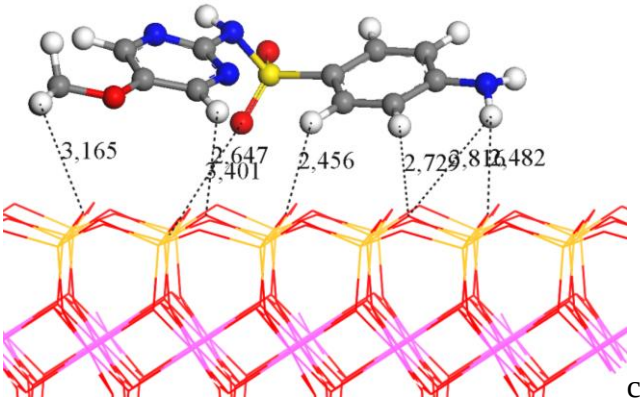
Figure 6. Adsorption complexes of **S1** on pyrophyllite (001) surface optimized with PCFFH with the adsorbate conformer *syn-cis-down* in configurations: P1 (a); P2 (b); P3 (c); and P4 (d); and the conformer *syn-trans-up* in a configuration P5 (e). Adsorption energies have been included in brackets below each model and they are in kcal/mol (Table 5).



a (-11.27)



b (-9.41)



c (-10.84)

Figure 7. Adsorption complexes of **S2** on pyrophyllite (001) surface optimized with PCFFH with the adsorbate in configurations P5 (a); P1 (b); and P5 twisted. Adsorption energies are in brackets below each model in kcal/mol (Table 5).

Journal of Computer-Aided Molecular Design

Computational study of substituent effects on the acidity, toxicity and chemical reactivity of selected bacteriostatic sulfonamides: Implications for drug design

--Manuscript Draft--

Manuscript Number:	
Full Title:	Computational study of substituent effects on the acidity, toxicity and chemical reactivity of selected bacteriostatic sulfonamides: Implications for drug design
Article Type:	Original Research Article
Keywords:	Aqueous solution acidity; toxicity; bacteriostatic sulfonamides; dihydropteroeato synthase (DHPS); DFT-M06-2X calculations.
Corresponding Author:	Catalina Soriano-Correa, Ph.D Universidad Nacional Autonoma de Mexico Mexico City, MEXICO
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universidad Nacional Autonoma de Mexico
Corresponding Author's Secondary Institution:	
First Author:	Catalina Soriano-Correa, Ph.D
First Author Secondary Information:	
Order of Authors:	Catalina Soriano-Correa, Ph.D
	Carolina Barrientos-Salcedo
	Misaela Francisco-Marquéz
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	A relationship among the physicochemical properties, the chemical structure and antibacterial activity of sulfonamides is not still completely explicated; nevertheless, from a therapeutics and prodrugs design point of view, it has been observed that the substituent group modify the electronic structure, the physicochemical features and chemical reactivity which are critical for the biological activity. In this work, we analyzed the substituent effects on the physicochemical properties, toxicity, chemical reactivity and its relation with the bacteriostatic activity of selected sulfonamides; through DFT-M06-2X calculations in aqueous solution, employing quantum chemical and docking descriptors. The results showed a correlation between the theoretical acidity and the pKa experimental values. The active sulfonamides have a larger acidity. The acidity increases with electron-withdrawing substituents. The main reactivity takes place on N4 atoms linked to aromatic ring, and in SO2NH moiety, which are influenced by substituents. Docking descriptors showed binding affinities between the sulfonamides and target (DHPS).

Computational study of substituent effects on the acidity, toxicity and chemical reactivity of selected bacteriostatic sulfonamides: Implications for drug design

Catalina Soriano-Correa[a,*], Carolina Barrientos-Salcedo[b] and Misaela Francisco-Márquez[c]

[a] Química Computacional, Facultad de Estudios Superiores Zaragoza, Universidad Nacional Autónoma de México (UNAM), Iztapalapa, C.P. 09230 Mexico City, Mexico

[b] Laboratorio de Química Médica y Quimiogenómica, Facultad de Bioanálisis-Veracruz, Universidad Veracruzana, C.P. 91700, Veracruz, Mexico

[c] Instituto Politécnico Nacional, UPIICSA, Av. Té #950, Col. Granjas México, C.P. 08400 Mexico City, Mexico

Corresponding author: *Catalina Soriano-Correa., Química Computacional, Facultad de Estudios Zaragoza (FES-Zaragoza), UNAM, Iztapalapa, 09230 Mexico City, Mexico. Phone: (+52) (55) 5773-6333; E-mail: csorico@comunidad.unam.mx

Abstract

A relationship among the physicochemical properties, the chemical structure and antibacterial activity of sulfonamides is not still completely explicated; nevertheless, from a therapeutics and prodrugs design point of view, it has been observed that the substituent group modify the electronic structure, the physicochemical features and chemical reactivity which are critical for the biological activity. In this work, we analyzed the substituent effects on the physicochemical properties, toxicity, chemical reactivity and its relation with the bacteriostatic activity of selected sulfonamides; through DFT-M06-2X calculations in aqueous solution, employing quantum chemical and docking descriptors. The results showed a correlation between the theoretical acidity and the pKa experimental values. The active sulfonamides have a larger acidity. The acidity increases with electron-withdrawing substituents. The main reactivity takes place on N4 atoms linked to aromatic ring, and in SO_2NH moiety, which are influenced by substituents. Docking descriptors showed binding affinities between the sulfonamides and target (DHPS).

Keywords. Aqueous solution acidity; toxicity; bacteriostatic sulfonamides; dihydropteroeato synthase (DHPS); DFT-M06-2X calculations.

Introduction

Since the discovery of the *in vitro* and *in vivo* antibacterial activity of sulfanilamide (*p*-aminobenzene sulfonamide), many N1-substituted sulfonamides have been tested for their bacteriostatic potency against human and veterinary pathogens [1-4]; such that, for many years the sulfonamides have been widely studied, because are highly applicable due to their pharmaceutical importance and clinical application, as antibacterial, antitrypanosomal and antiviral agents among other [1, 4-9]. A lot of studies of sulfonamides have showed a relationship among the bacteriostatic activity and the acidity, pKa, dissociation constant, charges, hydrophobic and electronic parameters, etc.; in particular, in those studies that have been shown a competition between *para*-aminobenzoic acid (PABA) and the sulfonamides by the Dihydropteroate synthase (DHPS), that is an enzyme used as target for sulfonamide antibiotics [10-17]. Besides, experimental and theoretical studies have pointed out that the physicochemical properties in these molecules have great influence on the physiological process and consequently on the biological activity; such that, a variety of chemical modifications have been made in aromatic and aliphatic sulfonamides, which have produced many other active agents; as potent antimicrobial sulfonamides, that have been obtained by the N1-substituted derivatives in the sulfonamide group (-SO_2HN1-) with, alkyl, acyl, heterocyclic and aromatic compounds which enhances the antibacterial properties and their clinical applications [11-20]. In this context, recently Papadopoulou and co-workers have examined several 3-nitrotriazoles-based sulfonamides [7, 8]; such that, these molecules have showed activity against *T. cruzi* and without toxic effects towards the host cells [8].

According to previously mentioned; in this work, we performed a study at the M06-2X/6-311++G(2d,2p)//6-311+(d,p) levels of theory in aqueous solution, for some selected

bacteriostatic N1-substitued sulfonamides, in order to analyze the electronic nature and influence of the substituent groups, as well as its relationship among acidity and the chemical structure with their bacteriostatic activity, using quantum chemical descriptors such as: acidity, the aromaticity and electrophilicity indexes, hardness, atomic charges, electrostatic potential and Fukui functions, as well as docking descriptors: free energies binding and polar interaction; with the purpose of designing drugs with specific chemical and biological characteristics, and less chemical toxicity.

Computational details

The bacteriostatic sulfonamides employed in this study, comprise the following: the Sulfanilamide (SNA), Sulfamerazine (SMZ), Sulfamethoxypyridazine (SMP), Sulfamethazine (SMT), Sulfamethoxydiazine (SMDZ), Isosulfamethyldiazine (ISMD), Sulfadiazine (SDZ), Sulfacloropyridazine, (SCP), Sulfamethoxazole, (SMX); Sulfacetamide (SCM), and the generic chemical structures are displayed in Figure 1. Electronic structure calculations were carried out using the Gaussian 09 program suite [21]. All neutral structures were optimized at the M06-2X functional, a hybrid meta exchange-correlation functional that is a good functional for aromatic groups, and describe well the no-covalent interactions [22, 23] and with the 6-311+G(d,p) basis set [24], whereas deprotonated (anionic) structures were optimized at unrestricted UM06-2X level, all these with a 6-311+G(d,p) basis set, respectively. Optimized geometries (neutral and anion structures) were characterized by harmonic vibrational frequencies, which confirmed that the structures obtained are minima on the potential energy surface. Single-point calculations were performed on optimized structures (neutral and anionic) at the M06-2X level of theory with a 6-311++G(2d,2p) basis

set. All calculations were undertaken in an aqueous solution; the solvent effect was described trough SMD model [25].

In order to determine the acidic character of the sulfonamide group, the acidity (ΔG^0) for all molecules was calculated by deprotonating the H14 hydrogen atom (-SO_2NH14R, see Figure 1), such as the following scheme [19, 26]:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \qquad (1)$$

$$\mathrm{M-H} \rightarrow M^- + \mathrm{H}^+ \qquad (2)$$

where,

$$\Delta H^0 = H^0(M^-) + H^0(H^+) - H^0(M-H) \qquad (3)$$

Such that, the enthalpy of the proton is reduced to the sum of the contribution from translational motion ($H_T = 3/2RT$) at T= 298.15 K and from the pV term.

So, the equation (3) may be written as

$$\Delta H^0 = H^0(M^-) - H^0(M-H) + \frac{5}{2}RT$$

Thus, small ΔG^0 values imply a more acidic of the -SO_2NH- moiety. This thermodynamic property has the same meaning that the deprotonation energy [18, 27-29]; that is, the ΔG^0 is

a measure of the promptness of the sulfonamide group to lose one proton in a chemical reaction or in a biological process [26].

Moreover, to depict a more reliable the chemical reactivity, some global reactivity descriptors as hardness (η) and electrophilicity index (ω) were performed. The electrophilicity index was calculated through the electronic chemical potential and the chemical hardness. For an N-electron system with external potential v(r) and total energy E, the electronic chemical potential μ, that is, the negative of the electronegativity χ, is defined as the partial derivative of the energy to the number of electrons at constant external potential [30-32].

$$\mu = -\chi = \left(\frac{\partial E}{\partial N}\right)_{v(r)} \qquad (4)$$

Parr and Pearson *et. al* [30-32] proposed, for calculations of hardness (η) the following expression:

$$\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_{v(r)} \qquad (5)$$

Using finite-difference approximation, the Eqs. (4) and (5) would be:

$$\eta \approx \frac{(E_{N+1} - 2E_N + E_{N-1})}{2} \approx I - A$$

$$-\mu \approx \frac{(E_{N-1} - E_{N+1})}{2} \approx \frac{I + A}{2}$$

where E, E_{N-1} and E_{N+1}, are the energies of N, (N-1) and (N+1) electron systems; I and A are the adiabatic ionization potential and the electron affinity, respectively.

Additionally, Parr and Yang [31], Parr *et al.* [33, 34] and Chattaraj *et al.* [35, 36] have defined another descriptor in order to quantify the global electrophilic power of the molecules, named electrophilicity index (ω), which defines a quantitative classification of the global electrophilic nature of a molecule within a relative scale. The electrophilicity index of a system in terms of its chemical potential and hardness is given by the following expression:

$$\omega = \frac{\mu^2}{2\eta}$$

Moreover, the atomic charges fitted to the electrostatic potential (ESP), were determined using standard model (CHELPG) [37], in order to analyze the N1-linked to substituent and to examine the electron donor or electron withdrawing character of the substituent group.

Furthermore, we evaluated the local density functional descriptor of reactivity that is the Fukui function, used to analyze chemical reactivity and site selectivity [38-40]. The Fukui

function values were obtained from Hirshfeld's population scheme which provides positive values of the function [41]. The Fukui function is defined as the derivative of the electron density $\rho(\vec{r})$ with respect to the total number of electrons N in the system, at constant external potential $v(\vec{r}\,)$ acting on an electron due to all the nuclei in the system [38, 39]

$$f(\vec{\boldsymbol{r}}) = \left[\frac{\delta\mu}{\delta v(\vec{\boldsymbol{r}})}\right]_N = \left[\frac{\partial\rho(\vec{\boldsymbol{r}})}{\partial N}\right]_{v(\vec{\boldsymbol{r}})}$$

Also, the Fukui function provide information about the site reactivity within a molecule. Depending on the electron transfer, three types of Fukui function are defined [39]. A common simplification of the Fukui function is to condense its value to individual atoms in the molecule [39, 40]. The condensed Fukui function were calculated using finite difference approach proposed by Yang and Mortier [42], by the following expressions.

$f_k^- = q_k(N) - q_k(N-1)$ for electrophilic attack,

$f_k^+ = q_k(N+1) - q_k(N)$ for nucleophilic attack,

$f_k^0 = [q_k(N+1) - q_k(N-1)]/2$ for radical attack,

where, q_k is the electronic population of atom k in a molecule.

Besides, in order to identify a possible relationship between carcinogenicity and aromaticity of sulfonamides, we estimated the energetic difference between molecular orbitals HOMO and HOMO-1, taking this parameter, Δ, as an indirect measure of the compounds' carcinogenicity according to the Barone rules and with previous theoretical works [29, 43-45]. Additionally, molecular docking studies were carried out using Docking Server [46-50], with the purpose to explore the affinities binding between the sulfonamides (ligands) and protein (target of the sulfonamides class of antibacterial) Dihydropteroate synthase from Staphylococcus aureus subsp. aureus Mu50, (Protein Data Bank ID: 4HB7).

Results and discussion

The experimental bacteriostatic activities *in vitro* and the pKa experimental values in Table 1, were taken from Ref. [1]. Such that, the bacteriostatic activities *in vitro* is defined as the minimum inhibitory concentration (MIC), which represents the minimum concentration of drug necessary to cause a bacteriostatic effect against *Escherichia coli* when the organisms are grown in a synthetic medium. So, smaller values imply higher bacteriostatic activity; throughout the present discussion, we will consider that Sulfanilamide (SNA) as the inactive molecule with respect to its analogues. The values of the acidity (ΔG^0), hardness (η), electrophilicity index (ω), aromatic index (Δ), as well as, the experimental values of the bacteriostatic activity and the pKa values for selected sulfonamides, free energy binding and polar interaction are shown in Table 1. We analyzed the ΔG^0 of the (sulfonamide group, -SO_2NH14R), by evaluating the rapidity or easily of the sulfonamide group to lose the H14 hydrogen atom in the molecule (see Figure 1); thus, it can be observed from Table 1 that SNA possesses the largest value for ΔG^0 and hence it presents the lowest acidity when compared with its analogs. It is noteworthy to note, that our theoretical results of ΔG^0 values

present a good correlation with the pKa experiment values (See Figure 2 and Table 1), these findings illustrate that the active sulfonamides possess acidic characteristics. These results are in agreement with previous theoretical studies [18, 28, 29]. Additionally, Figure 3 displays the isosurfaces of the electrostatic potential showing the acidic character of the H14 atoms, which increase with the electron-withdrawing effect of the substituent linked to N1.

Moreover, Table 1 shows the values for the global-reactivity descriptors calculated: hardness (η), electrophilic index (ω) and aromatic index (Δ). From Table 1, we observed that the SNA, SMX and SCM have small hardness values, indicating that these molecules are the more reactive that its analogues. The results of the electrophilic index indicated that the SNA, SMX and SCM have a larger predisposition for acquiring an additional electronic charge than the rest of their analogs, these results are in agreement with the previous descriptor of hardness. Also, from Table 1, the aromatic index showed that SNA, SCP and SDZ have larger values, whereas, SMX and SMDZ have smaller values than their analogs. So, the SMX and SMDZ might have less toxic effects than their analogues. These findings are in agreement with the Barone rules and with previous work [27, 43-45, 51].

Table 2 collects the atomic charges (fitted to ESP) values for the following atoms: N4; C7; C6; C3; N1; SO_2 and R, in order to estimate the nucleophilic/electrophilic nature sites, and to know which atoms might contribute to the chemical reactivity and with bacteriostatic effect. From Table 2, we explore the negative charge on N4, C6, C3 and N1 atoms; it was observed that SNA and SMP have the largest negative charge on N1 atoms and at the same time, these atoms have the smaller negative charge on N4 atoms than their analogues. In relation to the negative charge on the SO_2 group, the SNA has the largest value than its analogues. So, these results suggest that the atoms and groups with larger negative charge are more exposed to

electrophilic attack. Contrarily, the SDZ and SNA have the smaller positive charges on the C7 than the other sulfonamides; these findings suggest that atoms with larger positive charges are more exposed to nucleophilic attacks. Moreover, we investigated the substituent effect between the charges of R-substituent and the acidity. This result exhibited that molecules with a smaller positive charge, are those that have more acidity (Table 2 and Figure 3); so, we may note that acidity enhances with electron-withdrawing substituents and at the same time the substituent exerts an electro donor effect of the atomic charges on N4 atoms.

The Figure 4 displays the values of condensed Fukui functions obtained from Hirshfeld population scheme (Figure 4) for the following atoms: C3; C5; C7; C8; N1 and N4; from Figure 4, we observed that Fukui Function (f^-) values for C3; C5; C7; C8 and N4 are possible electrophilic attacks in all molecules; whereas, for Fukui Function (f^+) values obtained predict for N1; atoms as possible nucleophilic attacks, mainly for SNA and SMX; additionally, for Fukui Function (f^0) values suggest to C3, C5, C8 and N4 atoms as possible radical attacks to for all molecules. These results permitted us to know which are the more potential reactive sites for electrophilic, nucleophilic or radical attacks, respectively.

From Table 1 and Figure 5, we observed that the SMX with bacteriostatic activity has the most negative free energy binding than their analogs; contrarily, the SNA has the smallest value. In this sense, we can observe as the main interaction between the sulfonamide and DHPS is through the N4 atoms that have larger negative atomic charge in all active molecules. Such that, the bulky substituents increase the rigid of the molecules and therefore the interaction among the sulfonamides and DHPS. Also, the results illustrated a stronger polar interaction between the *p*-amino group of SMX and the residue of Lys251 than their analogs. These results are in agreement with previous studies of the importance of the N4

atoms of sulfonamides [11, 18], and with a recently studies of the chemical space of sulfonamides [52] and of the adsorption of the sulfonamides on phyllosilicates surfaces [53]. Our results exemplify how a substituent with specific chemical structure could modify the toxicity, acidity, chemical reactivity and consequently its biological functionality in the pharmacological molecules.

Conclusions

We have performed a theoretical study at the M06-2X in aqueous solution, to analyze the substituent effect on the acidity, toxicity and the chemical reactivity of bacteriostatic sulfonamides. According to our findings, we summarize the following points:

The results exhibited a correlation between the theoretical acidity (ΔG^0) and pKa experimental values. Also, the more acidic sulfonamides are those that have larger bacteriostatic activity. The acidity enhances with electron withdrawing substituents. Such that, the substituent group plays a critical role on the physicochemical features.

The chemical reactivity analysis indicated that, the electron withdrawing substituent groups decrease the aromatic index (less toxic). Also, sulfonamides more reactive (small hardness values) are capable of accepting additional electronic charge (larger electrophilic index values). Through the atomic charges and the condensed Fukui Function, we noticed that the main chemical reactivity takes place on aromatic ring, as well as on N4 atoms and in -SO_2NH- moiety. Complementary docking descriptors illustrated binding affinities between the sulfonamides and Dihydropteroate Synthase, that are in good agreement with the quantum chemical descriptors.

Our results of the acidity, atomic charges, Fukui functions, aromatic index, electrophilicity index, hardness and docking descriptors, led us to analyze and to characterize the reactive sites, which altogether might contribute to understand the bacteriostatic activity of the sulfonamides. Also, these findings might be useful in the designing of novel molecules with specific biological activity and less chemical toxicity.

Acknowledgments

We are grateful to the Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) of the Universidad Nacional Autónoma de México (UNAM) for allocation of computer time on the Miztli supercomputer. This research was partially supported by DGAPA-UNAM grant PAPIIT-IN114715.

References

1. Wolff ME (1979**)** Burger's Medicinal Chemistry, Part II. 4th ed.; Wiley & Sons: New York, Vol. 2, Chapter 13, pp 1-40.

2. Silverman RB, Holladay MW (2014**)** The Organic Chemistry of Drug Design and Drug Action. Academic Press.

3. Sukul P, Spiteller M (2006) Sulfonamides in the Environment as Veterinary Drugs. Rev. Environ Contam Toxicol 187:67-101.

4. Mengelers MJB, Hougee PE, Janssen LHM, Van Miert ASJPAM (1997) Structure-activity relationships between antibacterial activities and physicochemical properties of sulfonamides. J. vet Pharmacol. Therap. 20:276-283. doi:10.1046/j.1365-2885.1997. 00063.x

5. Stuper AJ, Brügger WE.; Jurs PC (1979) Computer Assisted Studies of Chemical Structure and Biological Function; Wiley-Interscience: New York

6. Zani F, Vicini P (1998) Antimicrobial Activity of Some 1, 2-Benzisothiazoles Having a Benzene Sulfonamide Moiety. Arch. Pharm. 331:219-223. doi:10.1002/(SICI)1521-4184(199806)331:6<219

7. Papadopoulou MV, Bloomer WD, Rosenzweig HS, Ashworth R, Wilkinson SR, Kaiser M, Andriani G, Rodriguez A (2013) Novel 3-nitro-1*H*-1,2,4-triazole-based compounds as potential anti-Chagasic drugs:*in vivo* studies. Future Med. Chem. 5:1763-1776. doi: 10.4155/fmc.13.108

8. Papadopoulou MV, Bloomer WD, Rosenzweig HS, Chatelain E, Kaiser M, Wilkinson SR, McKenzie C, Ioset J-R (2012) Novel 3-Nitro-1*H*-1,2,4-triazole-Based Amides and Sulfonamides as Potential Antitrypanosomal Agents. J. Med. Chem. 55, 5554-5565. doi:10.1021/jm300508n

9. Zarfl C, Matthies M, Klasmeier J (2008) A mechanistical model for the uptake of sulfonamides by bacteria. Chemosphere 70:753-760. doi: 10.1016/j.chemosphere.2007.07.045

10. Bell PH, Roblin RO (1942**)** Studies in Chemotherapy. VII. A Theory of the Relation of Structure to Activity of Sulfanilamide Type Compounds1. J. Am. Chem. Soc. 64:2905-2917.

11. Seydel JK (1968**)** Sulfonamides, structure-activity relationship, and mode of action. Structural problems of the antibacterial action of 4-aminobenzoic acid (PABA) antagonists. J. Pharm. Sci. 57:1455-1478. doi:10.1002/jps.2600570902

12. Yamazaki M, Kakeya N, Morishita T, Kamada A, Aoki M (1970) Biological activity of drugs. X. Relation of structure to the bacteriostatic activity of sulfonamides. (1). Chem. Pharm. Bull. 18:702-707. doi.org/10.1248/cpb.18.702

13. Biagi GL, Barbaro AM, Guerra MC, Forti GC, Fracasso ME (1974) Relation between. pi. and Rm values of sulfonamides. J. Med. Chem. 17:28-33. doi:10.1021/jm00247a007

14. Hansch C, Leo A (1995) Exploring QSAR. Fundamentals and Applications in Chemistry and Biology; ACS Professional Reference Book; American Chemical Society: Washington, DC.

15. Avendaño LMC (2001) Introducción a la química farmacéutica. McGraw-Hill Interamericana. Segunda edición. España.

16. Valderas MW, Andi B, Barrow WW, Cook PF (2008) Examination of intrinsic sulfonamide resistance in Bacillus anthracis: A novel assay for dihydropteroate synthase. Biochimica et Biophysica Acta 1780:848-853. doi: 10.1016/j.bbagen.2008.02.003

17. Şanli S, Altun Y, Şanli N, Alsancak G, Beltran JL (2009) Solvent Effects on pKa values of Some Substituted Sulfonamides in Acetonitrile-Water Binary Mixtures by the UV-Spectroscopy Method. J. Chem. Eng. Data. 54:3014–3021. doi:10.1021/je9000813

18. Soriano-Correa C, Esquivel RO, Sagar RP (2003) Physicochemical and structural properties of bacteriostatic sulfonamides: Theoretical study. Int. J. Quantum Chem. 94:165-172. doi:10.1002/qua.10597

19. Gomes JRB, Gomes P (2005) Gas-phase acidity of sulfonamides: implications for reactivity and prodrug design. Tetrahedron 61:2705-2712. doi: 10.1016/j.tet.2005.01.034

20. Aidas K, Lanevskij K, Kubilius R, Juška L, Petkevičius D, Japertas P (2015) Quantum chemical predictions based on density functional theory and SMD. J. Comput Chem. 36:2158-2167. doi:10.1002/jcc.23998

21. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2010) Gaussian 09, Revision C.01, Gaussian Inc: Wallingford CT.

22. Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor. Chem. Acc. 120:215-241. doi:10.1007/s00214-007-0310-x

23. Wheeler SE, Houk KN (2010) Integration grid errors for meta-GGA-predicted reaction energies: Origin of grid errors for the M06 suite of functionals, J. Chem. Theory Comput. 6:395-404. doi:10.1021/ct900639j

24. Hehre WJ, Radom L, Schleyer PVR, Pople JA (1986**)** Ab Initio Molecular Orbital Theory, John Wiley & Sons, New York. doi:10.1002/jcc.540070314

25. Marenich AV, Cramer CJ, Truhlar DG (2009) Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. J. Phys. Chem. B 113:6378-6396. doi:10.1021/jp810292n

26. Tinoco Jr., I, Sauer K, Wang JC (1995) Physical Chemistry: Principles and Applications in Biological Sciences, New Jersey: Prentice Hall.

27. Soriano-Correa C, Raya A, Esquivel RO (2008) Characterization of electronic structure and physicochemical properties of antiparasitic nifurtimox analogues: A theoretical study. Int. J. Quantum Chem. 108:1369-1379. doi:10.1002/qua.21633

28. Soriano-Correa C, Barrientos-Salcedo C, Raya A, Rubio-Póo C, Esquivel RO (2010) The influence of electron donor and electron acceptor groups on the electronic structure of the anti-inflammatory tripeptide Cys-Asn-Ser. Int. J. Quantum Chem. 110:2398-2410. doi:10.1002/qua.22673

29. Soriano-Correa C, Raya A, Barrientos-Salcedo C, Esquivel RO (2014) Influence of the physicochemical and aromatic properties on the chemical reactivity and its relation with carcinogenic and anticoagulant effect of 17 β-aminoestrogens. Chem. Phys. 438:48-59. doi.org/10.1016/j.chemphys.2014.04.012

30. Parr RG, Pearson RG (1983) Absolute hardness: companion parameter to absolute electronegativity. J. Am. Chem. Soc. 105:7512-7516. doi:10.1021/ja00364a005

31. Parr RG, Yang W (1989) Density-Functional Theory of Atoms and Molecules, Oxford University Press, New York.

32. Pearson RG (1997) Chemical Hardness: Applications from Molecules to Solids, Wily-VCH, Weinheim, Germany.

33. Parr RG, Donnelly RA, Levy M, Palke WE (1978) Electronegativity: the density functional viewpoint. J. Chem. Phys. 68:3801-3807. doi.org/10.1063/1.436185

34. Parr RG, Szentpály LV, Liu S (1999**)** Electrophilicity index. J. Am. Chem. Soc. 121:1922-1924. doi:10.1021/ja983494x

35. Chattaraj PK, Chakraborty A, Giri S (2009) Net electrophilicity. J. Phys. Chem. A. 113:10068-10074. doi:10.1021/jp904674x

36. Chattaraj PK, Giri S, Duley S (2011) Update 2 of: Electrophilicity index. Chem. Rev. 111:PR43-PR75.

37. Politzer P, Truhlar DG (Eds.) (2013) Chemical applications of atomic and molecular electrostatic potentials: reactivity, structure, scattering, and energetics of organic, inorganic, and biological systems. Springer Science & Business Media.

38. Fukui K (1982) Role of Frontier Orbitals in Chemical Reaction Science. 218:747-754.

39. Parr RG, Yang W (1984) Density functional approach to the frontier-electron theory of chemical reactivity. J. Am. Chem. Soc. 106:4049-4050. doi:10.1021/ja00326a036

40. Padmanabhan J, Parthasarathi R, Sarkar U, Subramanian V, Chattaraj PK (2004) Effect of solvation on the condensed Fukui function and the generalized philicity index. Chem. Phys. Lett. 383:122-128. doi.org/10.1016/j.cplett.2003.11.013

41. Hirshfeld FL (1977) Bonded-atom fragments for describing molecular charge densities. Theor. Chim. Acta 44:129-138. doi:10.1007/BF00549096

42. Yang W, Mortimer WJ (1986) The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines. J. Am. Chem. Soc. 108:5708-5711. doi:10.1021/ja00279a008

43. P.M.V.B. Barone, A. Camilo Jr., D.S. Galvão (1996) Theoretical approach to identify carcinogenic activity of polycyclic aromatic hydrocarbons. *Phys. Rev. Lett*. 77:1186-189. doi.org/10.1103/PhysRevLett.77.1186

44. Braga RS, Barone PMVB, Galvão DS (1999) Identifying carcinogenic activity of methylated polycyclic aromatic hydrocarbons (PAHs) J. Mol. Struct. (THEOCHEM) 464:257-266. doi.org/10.1016/S0166-1280(98)00557-0

45. Vendrame R, Braga RS, Takahata Y, Galvão DS (2001) Structure–carcinogenic activity relationship studies of polycyclic aromatic hydrocarbons (PAHs) with pattern-recognition methods. J. Mol. Struct. (THEOCHEM) 539:253-265. doi.org/10.1016/S0166-1280(00)00795-8

46. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parametrization, and performance of MMFF94. Journal of Computational Chemistry 17:490-519. doi:10.1002/(SICI)1096-987X(199604)17:5/6<490: AID-JCC1>3.0.CO;2-P

47. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart E, Belew RK, Olso AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. Journal of Computational Chemistry 19:1639-1662.

48. Bikadi Z, Hazai E (2009) Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. Journal of Cheminformatics 1:15. doi:10.1186/1758-2946-1-15.

49. Solis FJ, Wets RJB (1981) Minimization by Random Search Techniques. Mathematics of Operations Research 6:19-30. doi.org/10.1287/moor.6.1.19

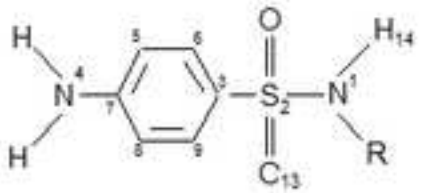
50. Humphrey W, Dalke A, Schulten K (1996) VMD - Visual Molecular Dynamics J. Mol. Graph.14:33-38. doi:10.1016/0263-7855(96)00018-5

51. Raya A, Barrientos-Salcedo C, Rubio-Póo C, Soriano-Correa C (2011) Electronic structure evaluation through quantum chemical descriptors of 17β-aminoestrogens with an anticoagulant effect. Eur. J. Med. Chem. 46, 2463-2468. doi:10.1016/j.ejmech.2011.03.032

52. López-Rosa S, Molina-Espíritu M, Esquivel RO, Soriano-Correa C, Dehesa J (2016) Study of the Chemical Space of Selected Bacteriostatic Sulfonamides from an Information-Theoretical Point of View. ChemPhysChem. 17,4003-4010. doi: 10.1002/ cphc.201600790

53. Francisco-Marquez, Misaela; Soriano-Correa, Catalina; Sainz-Diaz, Claro Ignacio (2016) Adsorption of Sulfonamides on Phyllosilicate Surfaces by Molecular Modeling Calculations., Submitted, Manuscript ID: jp-2016-124677, JP.

Figure_1 Click here to download Figure Figure_1.tif



-R

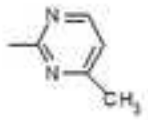
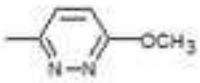
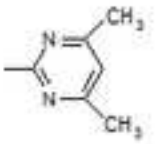
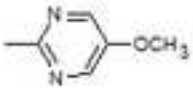
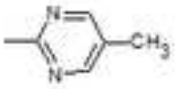
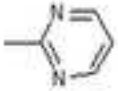
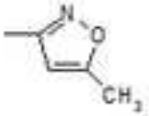
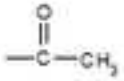
Sulfanilamide (SNA)	
Sulfamerazine (SMZ)	
Sulfamethoxypyridazine (SMP)	
Sulfamethazine (SMT)	
Sulfamethoxydiazine (SMDZ)	
Isosulfamethyldiazine (ISMD)	
Sulfadiazine (SDZ)	
Sulfacloropyridazine (SCP)	
Sulfamethoxazole (SMX)	
Sulfacetamide (SCM)	

Figure_2

Click here to download Figure Figure_2.tif

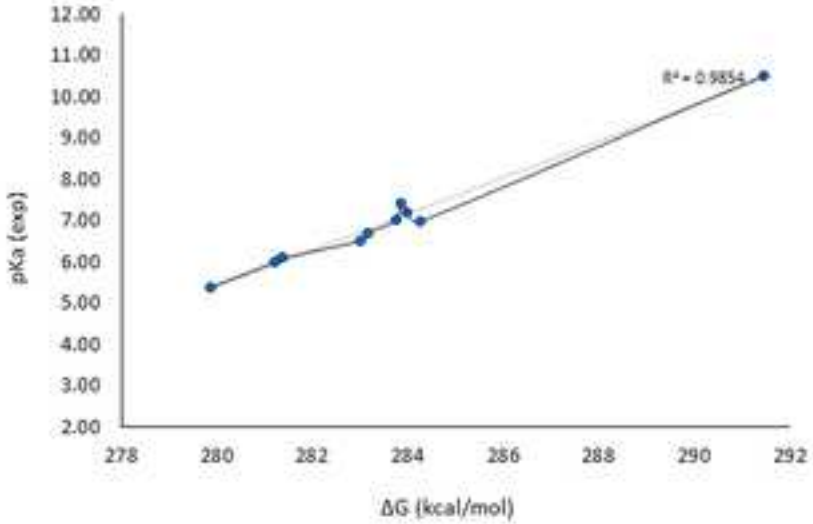


Figure_3

Click here to download Figure Figure_3.tif

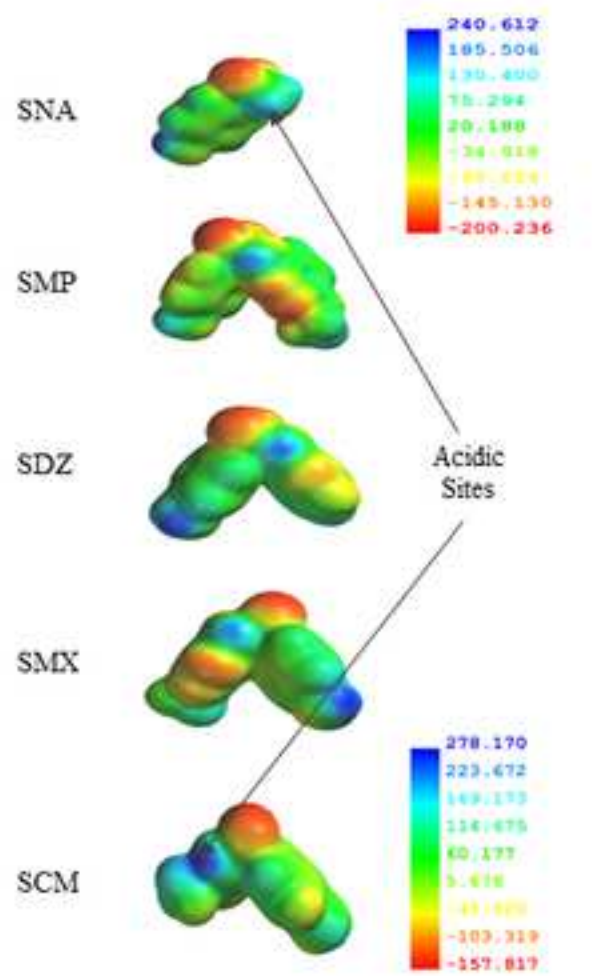


Figure_4

Click here to download Figure Figure_4.tif

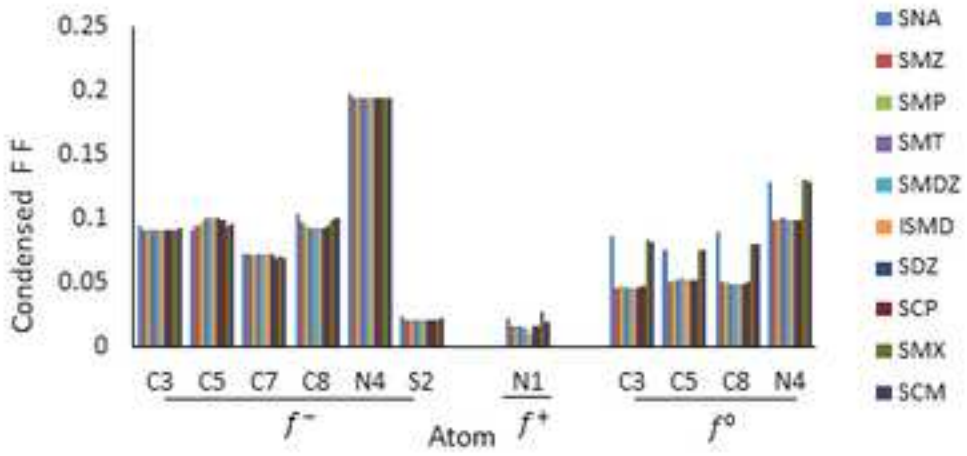


Figure_5 Click here to download Figure Figure_5.tif

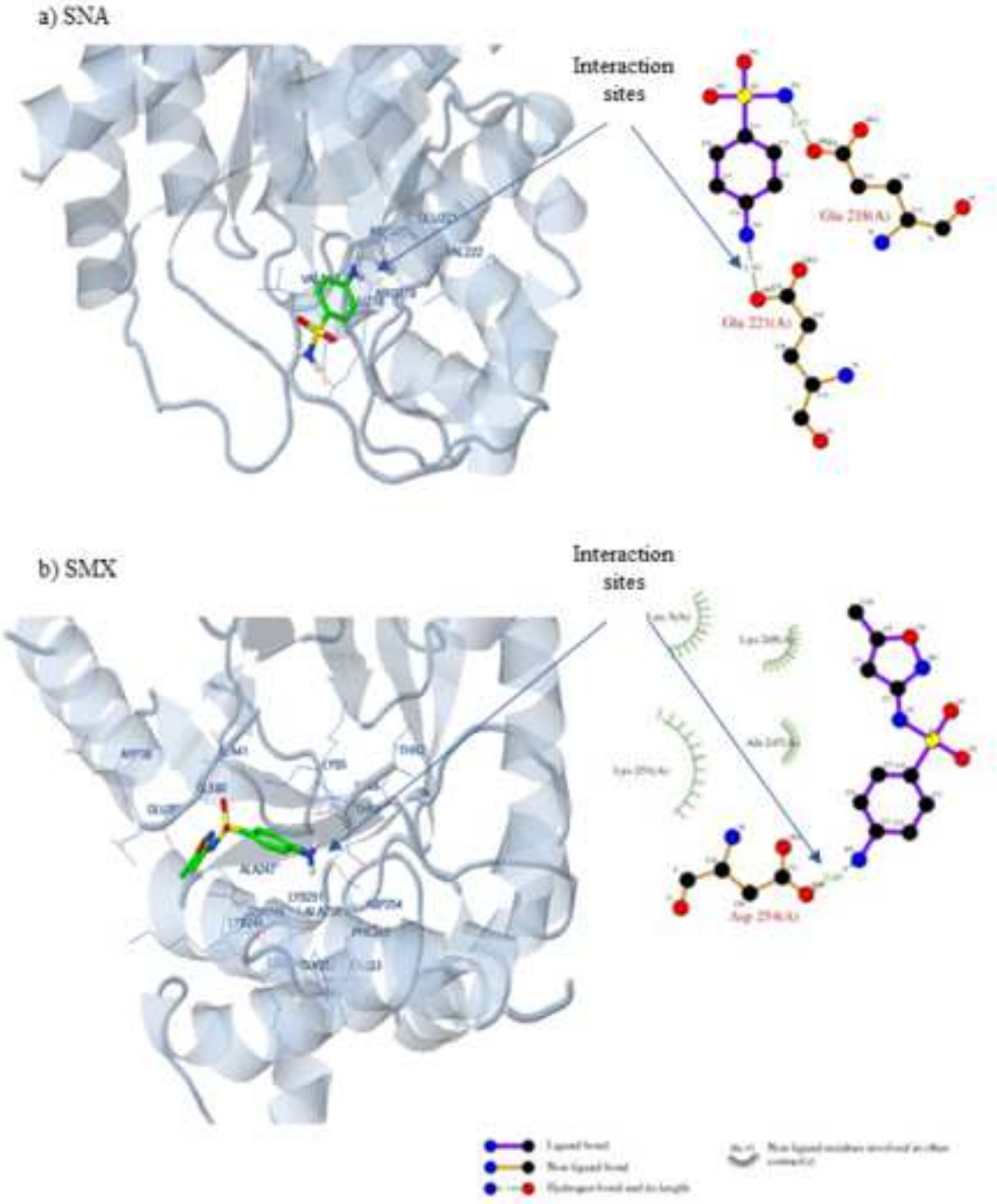


FIGURE CAPTIONS

Figure 1. Generic Structure and Selected Bacteriostatic Sulfonamides.

Figure 2. Correlation between Acidity (ΔG^0) *vs* pKa experimental values of bacteriostatic selected sulfonamides

Figure 3. Isosurfaces obtained for the electrostatic potential showing the acidic character of the H14 for selected bacteriostatic sulfonamides at the M06/6-311+G(d,p) level. Negative ESP values (colored in shades of red); positive ESP values (colored in shades of blue) go from – 157.817 to 278.170 for SCM.

Figure 4. Condensed Fukui function values for bacteriostatic sulfonamides, using Hirshfeld population scheme.

Figure 5. Hydrogens bond interactions between sulfonamide and Dihydropteroate Synthase, and hydrophobic interactions with specific residues for: a) SNA and b) SMX.

Table 1. Acidity (ΔG^0) and quantum chemical descriptors of selected bacteriostatic sulfonamides at the M06-2X/6–311++G(2d,2p) // M06-2X/6–311+G(d,p) levels.

Sulfonamide	ΔG^0 kcal/mol	η (e.V)	ω (e.V)	Δ (e.V)	[a]pKa (Exp)	[a]*In vitro* Activity *E.coli* μmol/L	[b]Free Energy Binding kcal/mol	[b]Polar Interaction kcal/mol Lys251
SNA	291.46	7.38	0.41	1.42	10.5	128.0	-1.19	-
SMZ	284.25	7.84	0.31	1.13	6.98	0.95	-1.82	-
SMP	283.97	8.21	0.25	0.99	7.2	1.0	-2.09	-1.79
SMT	283.86	7.74	0.33	1.07	7.4	1.70	-1.76	-1.79
SMDZ	283.75	8.08	0.27	0.52	7.0	2.0	-1.27	-0.88
ISMD	283. 15	7.96	0.29	0.91	6.7	1.0	-1.30	-0.83
SDZ	283.02	7.97	0.29	1.23	6.52	0.9	-1.80	-1.97
SCP	281.38	8.49	0.21	1.39	6.10	-	-2.44	-1.39
SMX	281.21	7.49	0.39	0.28	6.0	0.8	-2.39	-2.31
SCM	279.88	7.52	0.38	1.44	5.4	2.3	-1.42	-0.69

[a]The experimental bacteriostatic activities *in vitro* and the pKa experimental values, were taken from Ref. [1].
[b]Free energy binding and polar interaction were performed by Docking calculations.

Table 2. Atomic charges for selected bacteriostatic sulfonamides at the M06-2X/6–311++G(2d,2p)// M06-2X/6–311+G(d,p) levels.

Sulfonamide	N4	C7	C6	C3	N1	SO_2	R
SNA	-0.907	0.479	-0.079	-0.059	-0.841	-0.223	0.403
SMZ	-0.929	0.512	-0.060	-0.163	-0.640	-0.188	0.239
SMP	-0.908	0.500	-0.062	-0.105	-0.714	-0.186	0.293
SMT	-0.915	0.486	-0.043	-0.106	0.644	-0.194	0.231
SMDZ	-0.928	0.504	-0.036	-0.129	-0.680	-0.180	0.261
ISMD	-0.930	0.506	-0.013	-0.138	-0.678	-0.178	0.265
SDZ	-0.930	0.478	-0.053	-0.084	-0.666	-0.199	0.251
SCP	-0.929	0.510	0.055	-0.119	-0.706	-0.169	0.284
SMX	-0.932	0.508	-0.087	-0.107	-0.661	-0.177	0.219
SCM	-0.921	0.514	-0.077	-0.087	-0.481	-0.237	0.090

Values of atomic and group charges are given in a.u.

CHAPTER 6

Information-Theoretic Representation of the Chemical Space of Many Electron Systems

R.O. Esquivel[a,b,*], S. López-Rosa[b,c], M. Molina-Espíritu[a], C. Soriano-Correa[d], J.C. Angulo[b,e] and **J.S. Dehesa[b,e]**

[a] *Departamento de Química, Universidad Autónoma Metropolitana, 09340-México, CDMX, México*

[b] *Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, 18071-Granada, Spain*

[c] *Departamento de Física Aplicada II, Universidad de Sevilla, 41012-Sevilla, Spain*

[d] *Química Computacional, FES-Zaragoza, Universidad Nacional Autónoma de México, 09230-Iztapalapa, México, CDMX, México*

[e] *Departamento de Física Atómica, Molecular y Nuclear, Universidad de Granada, 18071-Granada, Spain*

Abstract: In this chapter we review the utility of an information-theoretic three-dimensional (IT-3D) space to unveil the unique physical, chemical and biological aspects of a great diversity of many electron systems, ranging from neutral and ionized atomic systems and simple molecules to much more complex ones such as amino-acids and pharmacological molecular ensembles. This space is generated from the Shannon entropy, the Fisher information and the disequilibrium measures along with their corresponding Fisher-Shannon and López-Ruíz-Mancini-Calvet (LMC) complexity measures. To achieve it we start from the theoretical ground that atoms and molecules can be described by means of the basic information-theoretical notions of delocalization, order, uniformity and complexity; thus, revealing the possible existence of an universal three-dimensional information-theoretic space for all systems in Nature. On the other hand, we discuss the abilities of the Shannon entropy, Fisher information and disequilibrium to capture the spatial spreading features of delocalizability, order and uniformity of biological molecules. Indeed, these three entropic measures are

[*] **Corresponding author R.O. Esquivel:** Departamento de Química, Universidad Autónoma Metropolitana, 09340-México, D.F., México; Tel/Fax: +525558044675; E-mail: esquivel@ xanum.uam.mx

Zaheer Ul-Haq and Jeffry D. Madura (Eds.)

found to uniquely characterize all amino acids, and some selected pharmacological systems, through a predominant information-theoretic quality scheme (PIQS) which gathers all chemical families by means of three major spreading features: delocalization, narrowness and uniformity. This scheme is shown to recognize 4 chemical groups characterized by this entropic scheme: delocalized (aliphatic and aromatic), narrowed (electro-attractive) and uniform (tiny). Chemical groups are differentiated according to their energy classifications. Also, it is shown that information planes produce interesting patterns associated to the PIQS scheme.

1.1. INTRODUCTION

Most physical theories pursue to describe the most basic aspects of the macroscopic world through simple models, predicting some parameters that are assumed or taken from experiments. In consequence, the prediction of these parameters cannot be predicted by simple theoretical models. Obviously, to gain insight of all physical features requires to analyse the features of the systems in smaller scales where the simplest processes correspond to the lowest level of knowledge. It is advantageous to go to a deeper level since it reduces the number of unspecified parameters, and hence the corresponding theory is considered to be complete and fairly adequate. A typical example of this kind of theories is Molecular Biology which is ultimately based on quantum chemistry and molecular dynamics. Notwithstanding that more comprehension of the lower level is achieved, it is practically impossible to attain a full description of the molecular processes taking place in living systems, hence the intricacy of the large set of parameters makes the endeavour a very difficult one. Considering an alternative approach to extracting the essential features of biological processes by use of Information Theory (IT) concepts has proven to be a succesful one. Moreover, the rapidly evolving field of Quantum-information biology [1 - 3], which employs information-theoretic concepts, is gaining wide attention to comprehend some of the most basic and yet unsolved questions of molecular biology.

There has been an increasing interest in characterizing and classifying different physical systems in terms of a few fundamental properties, not only in Physics but also in Chemistry and Biology. Perhaps, quantitative structure activity relation (QSAR) and quantitative structure properties relation (QSPR) constitute the most commonly approaches employed to relate molecular structures with physical

properties and biological activity. The beginning of these techniques could be remounted to mid 60's of the last century, when Hansch and Fujita proposed a connection between biological activity and chemical structure [4]. Furthermore, they argued that similar molecules share similar solubility, expecting that the relative polarity of molecules could be crucial in order to find a parameter relating structure and activity. Based on the ideas of Robert Muir and the Hammet equation [5], the structural changes might be correlated by means of parameters allusive with the partition coefficient to numerically analyse structure-activity problems [6] of biomolecules. Consequently, QSAR has evolved from simple regression methods to the analysis of very large sets of data comprising thousands of diverse molecular structures, and uses a wide variety of statistical and machine learning techniques. These advances have found broad application on QSAR methods in chemistry, material and nano-material, and life sciences to assess potential impacts on ecological systems [7]. One of the most promising application of this methodology resides in the chemical space [8]. The concept of chemical space emerges as a metaphor, and suggests the existence for a chemical universe which contains millions of organic compounds [9]. Although chemical space has not been well defined, it considers a multidimensional descriptor space in the sense of a region defined by a particular choice of descriptors to characterize as many chemical compounds as possible, and relate similar molecular structures with desired physicochemical properties and biological activity. In that respect, the relevance of any region of the chemical space must be judged by its ability to group compounds with similar bioactivity together [10].

The large number of physicochemical properties to be chosen as descriptors of the chemical space is an important disadvantage, due to the risk of employing irrelevant and redundant descriptors. Moreover, different systems could be wrongly misplaced at the same point of such a space if the descriptors selection is not well chosen [7]. A deeper understanding of this vast set of molecules will advance our knowledge of biological processes; therefore, the development of a systematic and rational classification of the chemical space is crucial for the progress of chemical applications. The analysis and exploration of this space represents a highly demanding computational task due to the immense number of possible stable molecules [11]. This challenge has led to several sophisticated

methods that combine the predictability of 2D- and 3D-QSAR methods with computational chemistry calculations [12 - 16].

In recent years, quantum many body systems have been described by Information Theory concepts. This type of analyses are driving the interest of scientists in several interconnected sciences such as Physics, Chemistry, and Biology. Indeed, there is a growing interest in applying information-theoretic techniques to systems and their associated processes at different scales; either chemical, mesoscopic or biological ones. Linked with the above, multidisciplinary efforts are being undertaken to employ IT at the classical (Shannon entropy, Fisher and complexity measures) and at the quantum (witnesses of entanglement) levels, to study a diverse type of chemical and biological purposes [17, 18]. In Chemistry, the Shannon entropy, *S*, has been used to reveal important chemical regions that are not present in the energy profile [18] by using the localized (delocalized) features of the electron distributions, allowing a phenomenological description of elementary chemical and several concurrent processes (bond forming/breaking) [19]. Other studies have used the information conservation principle to calculate several quantifiers of chemical reactivity: electrophilicity, nucleophilicity, regioselectivity and electrophilic aromatic substitution [20 - 22]. Also, IT studies have been focused to analysing molecular communication of biological organisms [23]. Complementary to the Shannon description, Fisher information *I* provides a different aspect of IT by providing a quantitative estimation of the fluctuations of the electron distributions. This quantity represents the cornerstone of numerous physical problems, for instance to derive the non-relativistic quantum-mechanical equations [24, 25] by means of a minimum principle [26, 27]. Besides, characterization of atomic and molecular avoided crossings has been reported [28]. Also, the "narrowness/disorder" features of *I* have been studied for electron densities in conjugated spaces to study the steric effect of the conformational barrier of ethane [29]. Fisher information has also been employed in Biology to test a model of carcinogenesis by applying extreme physical information analysis [30].

Several measures of complexity have been proposed and widely applied in physical sciences as general indicators of pattern, structure, and correlation, either to electron systems or processes. Several notions have been proposed for

quantifying the concepts of complexity and information, including the Kolmogorov-Chaitin or algorithmic Information Theory [31, 32], the classical IT of Shannon and Weaver [33], Fisher information [26, 27], and the logical [34] and the thermodynamical [35] depths, among others. Some of them share rigorous connections with others as well as with Bayes and information theory [36]. Different meanings for the term "complexity" have been applied in the literature: algorithmic, geometrical, computational, stochastic, effective, statistical, and structural, among others. Moreover, it has been employed in many fields: dynamical systems, disordered systems, spatial patterns, language, many electron systems, cellular automata, neuronal networks, self-organization, DNA analyses, social sciences, among others [37 - 42]. The definition of complexity is not unique, its quantitative characterization has been an important subject of research and it has received considerable attention [43]. The usefulness of each definition depends on the type of system or process under study, the level of the description, and the scale of the interactions among either elementary particles, atoms, molecules, biological systems, *etc*. Fundamental concepts such as uncertainty or randomness are frequently employed in the definitions of complexity, although some other concepts like clustering, order, localization or organization might be also important for characterizing the complexity of systems or processes. It is not clear how the aforementioned concepts might intervene in the definitions so as to quantitatively assess the complexity of the system. However, recent proposals have formulated this quantity as a product of two factors, taking into account delocalization/disequilibrium and delocalization/order [44, 45].

From an information-theoretical point of view, an alternative has recently emerged by quantifying the structural shape of the electron density distribution of molecules [46, 47]. These previous studies suggest that molecules with similar informational properties (like narrowness, localizability and uniformity) are related to organic compounds sharing similarities among their 3D-structures and chemical properties. Indeed, from simple physical systems to complex biological ensembles, properties of atoms and molecules strongly depend on the spread of the one-electron density $\rho(\mathbf{r})$ which characterizes its quantum-mechanically allowed state [48, 49]. Furthermore, the information theory of quantum systems provides an entropy-based characterization of atomic and molecular systems,

which complements the energy-based representation obtained through the wave function and density functional methods. Measures of uncertainty, randomness, disorder and localization are basic ingredients encountered to play a relevant role for the identification and description of numerous quantum phenomena in physical systems and chemical processes. These features of delocalization, uniformity and order can be quantified by the information-theoretic measures of Shannon entropy, disequilibrium and Fisher information of the corresponding electron densities of the system, respectively. Taking into account these concepts, an information-theoretic space in three dimensions (IT-3D) can be defined.

This chapter is organized as follows. In Section 2, we define and justify the different information theoretic measures to be used in this analysis. In Section 3, the IT-3D chemical space of a large variety of many electron systems of physical, chemical, biological and pharmacological interest are presented. In Section 4, we analyse this IT-3D space for a group of selected sulfonamide-type molecules. In Section 5, information-theoretic measures are employed to uniquely characterize all amino acids through a predominant information-theoretic quality scheme (PIQS). Finally, in the last section some conclusions and open problems are given.

1.2. INFORMATION-THEORETICAL MEASURES

The one-electron density, $\rho(\mathbf{r})$, is a physical observable that can be obtained experimentally or calculated using ab initio or density functional theory methods. It quantifies the probability of an electron being present at an infinitesimal element of space defined in $\Re^3$

$$\rho(\mathbf{r}) = \int |\psi|^2 d\mathbf{r_2} \ldots d\mathbf{r_N}, \tag{1.1}$$

normalizes to N (the number of electrons),

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \tag{1.2}$$

and, according to Hohenberg-Kohn theorem, the physical and chemical properties of atoms and molecules strongly depend on the one-electron density of the ground

state of the system [50]. It is important to note that all electron densities are normalized to unity which is adequate to represent probability distributions which are scaled to the size of the system according to the definition of the shape function [51] ($\sigma(\mathbf{r}) = \rho(\mathbf{r})/N$). It is convenient to denote unity-normalized densities with the symbol $\rho(\mathbf{r})$.

The properties of atoms and molecules depend heavily on the spread of the probability distribution which characterizes their quantum-mechanical states. The spread might be grasped and quantified by several information-theoretic measures, which quantify different facets of the density, aside of its standard deviation or its variance. There exist two kinds of measures: global and local ones. Global measures are suit to quantify the extent of the probability density in several manners according to their corresponding analytic definition; *i.e.*, they are represented by means of density functionals of logarithmic (Shannon) and power (disequilibrium) type. So that, they are barely sensitive to density fluctuations. In contrast, the Fisher information, I, is able to grasp the local features of the electron distribution since it is a functional of the density gradient. Therefore, Fisher information is highly sensitive to the changes of the probability distribution, locally wise. Hence, we expect that different physical and chemical properties from different molecules might be characterized by narrower electron densities. The above is adequate enough to characterize all the structural features of the electron density by use of the different information-theoretic measures to qualify and quantify several of its topological aspects. These measures are defined below.

The *Shannon entropy*, S, of a probability distribution describing a specific quantum state of the system, is given by [52].

$$S[\rho] = -\int \rho(\mathbf{r}) \ln \rho(\mathbf{r}) d\mathbf{r} \qquad \textbf{(1.3)}$$

where ρ ($\mathbf{r}$) is the unity-normalized density in position space. This definition allows to assess the delocalization or lack of structure of the electronic density. It is then interpreted as follows: S [ρ] is maximal when knowledge of $\rho(\mathbf{r})$ is minimal, characterizing its delocalized topological features.

The *disequilibrium* measure, also known as *self-similarity* [53] or *information energy* [54], D, is useful to quantify the departure from uniformity (or equiprobability) of the density. The disequilibrium is defined as [54]:

$$D[\rho] = \int \rho^2(\mathbf{r}) d\mathbf{r} \tag{1.4}$$

The *Fisher information* [26, 27], I, is defined by the following functional of the gradient of the density:

$$I[\rho] = \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r}. \tag{1.5}$$

This quantity measures the spatial locally wise concentration of the density cloud and quantifies its gradient content, thus revealing the irregularities of the density and providing a quantitative estimation of its fluctuations. Besides, according to the localized/delocalized features of the distributions, Fisher information can be interpreted as a measure of the departure of the probability density from disorder.

These three quantities measure different aspects for the electron density distribution, which means that one molecule has a distinctive set of theoretic-informational parameters [46, 47]. If we compare two, or more molecules with each other, we expect that molecules with similar structure, and physicochemical properties, share a similar spread of their electron density distribution. Therefore, a 3D informational space (composed by S, I, D) will be able to reflect a combined three-fold pattern to allowing specific information-theoretical regions for molecules sharing similar spatial portions (or not) of *delocalization*, *uniformity* and *order*.

Statistical complexity measures, namely the López-Ruiz-Mancini-Calbet (LMC) and the Fisher-Shannon ones, consist of dyadic products of information-theoretic quantities, so that each complexity measure attains the combined balance of two different aspects of the density. These measures obey boundary conditions attaining minimals in the highest ordered state of the electron distribution as well as in the disordered limit. The *LMC or shape complexity* measure [44], C_{LMC} , is

defined through the product of measures of global character: the disequilibrium D (quantifying the departure of the density from uniformity) and the Shannon entropic power, $L[\rho] = e^{S[\rho]}$, (quantifying the departure of the density from localizability). For a given density, $\rho(\mathbf{r})$, the *LMC* complexity measure is defined as given by

$$C_{LMC} = D[\rho]L[\rho] \tag{1.6}$$

In contrast, the *Fisher-Shannon complexity*, C_{FS} is defined through quantities of global (Shannon) and local character: the Fisher information measure, I, and the power entropy, with factors that preserve the general complexity properties, *i.e.*, $J = \frac{1}{2\pi e} e^{\frac{2}{3}S}$. That is,

$$C_{FS} = I[\rho]J[\rho] \tag{1.7}$$

Note that in Fisher-Shannon complexity, the disequilibrium global factor D is replaced by the Fisher local one in order to quantify the departure of the density from disorder through the gradient of the distribution. This kind of complexity has been associated as a measure of atomic correlation [55] and also defined as a statistical complexity measure [45, 56].

1.3. INFORMATION-THEORETIC CHEMICAL SPACE FOR MANY ELECTRON SYSTEMS

Features of delocalization, uniformity and order are to be quantified by information-theoretical measures defined in the previous section. Upon this theoretical approach all systems are assumed to be characterized through their information content, featuring any of the three different IT qualities aforementioned which make each of them unique among the rest of the chemical/biological structures. We have introduced recently [47] an *information-theoretic space in three dimensions* (IT-3D) with the three following axes: disequilibrium D, Shannon entropy S, and Fisher information I. We focus our attention in a large variety of many electron systems of physical, chemical, biological and pharmacological interest (388 atoms and 115 molecules). The electronic calculations in this work have been performed as follows:

- For neutral and ionized atoms: Accurate near-Hartree-Fock atomic wave functions of Koga *et al.* [57, 58] were employed in order to calculate the atomic densities and their corresponding informational measures for all atoms, singly-charged ions and isoelectronic series.
- Amino acids: The electronic structure calculations were obtained with the Gaussian 03 suite of programs [59] at the *HF/*6 − 311 + *G*(*d, p*) level of theory [60].
- Molecules with pharmacological interest: The electronic structure calculations performed in this study were carried out with the Gaussian 98, 03 and 09 suite of programs [59, 61, 62] at different levels of theory: (i) restricted Hartree-Fock (RHF)/6-31+G(d) for the sulfonamides [63] and anti-chagasic [64], respectively (iii) DFT/6-311+G(d,p) using Becke's 3−parameter hybrid functional (B3LYP) and the Lee-Yang-Parr correlation for the anti-inflammatory tripeptides [65]. Single point calculations at different levels of theory were performed on the optimized structures: (i) B3LYP/6-311+G(2d,2p) for the anti-inflammatory tripeptides and anti-chagasic, (ii) HF/6-31G(d,p) for the antibiotics [66], (iii) B3LYP/6- 31+G(d,p) for the sulfonamides.
- Organic molecules: The electronic structure calculations were obtained with the Gaussian 03 suite of programs [59] at the *CISD/*6 − 311 + +*G*(3*d f,* 2*p*) level of theory.

We have calculated all information and complexity measures defined in the previous section for all molecular systems studied here by employing software developed in our laboratory along with 3D numerical integration routines [67] and the DGRID suite of programs [68].

In Fig. (**1.1**) the IT-3D space with the positions of 103 neutral atoms, 96 (positively- and negatively-charged) ions from *H* to *Cs*, and 9 isoelectronic series from *He* through *Ne* is depicted. From Fig. (**1.1**) (top) for the atomic IT-3D space we observe two different regions: the isoelectronic series of light atoms with a number of electrons $N = 2$ to $N = 10$ (balls on yellow, green, purple, *etc.*) extend far beyond the region of the neutral (grey spheres) and ionic systems (balls on blue and red) on the D-axis (disequilibrium), thus characterizing isoelectronic atoms of increasing nuclear charge Z by departing from uniformity (equiprobability) as compared to the neutral and ionic systems. Note that

information-theoretic space of anions departs from that of neutral atoms to the right on the *S*-axis, while for the cations it goes to the left, *i.e.*, negatively-charged atoms are characterized by holding delocalized distributions as compared to the neutral and the positively-charged ionized atoms.

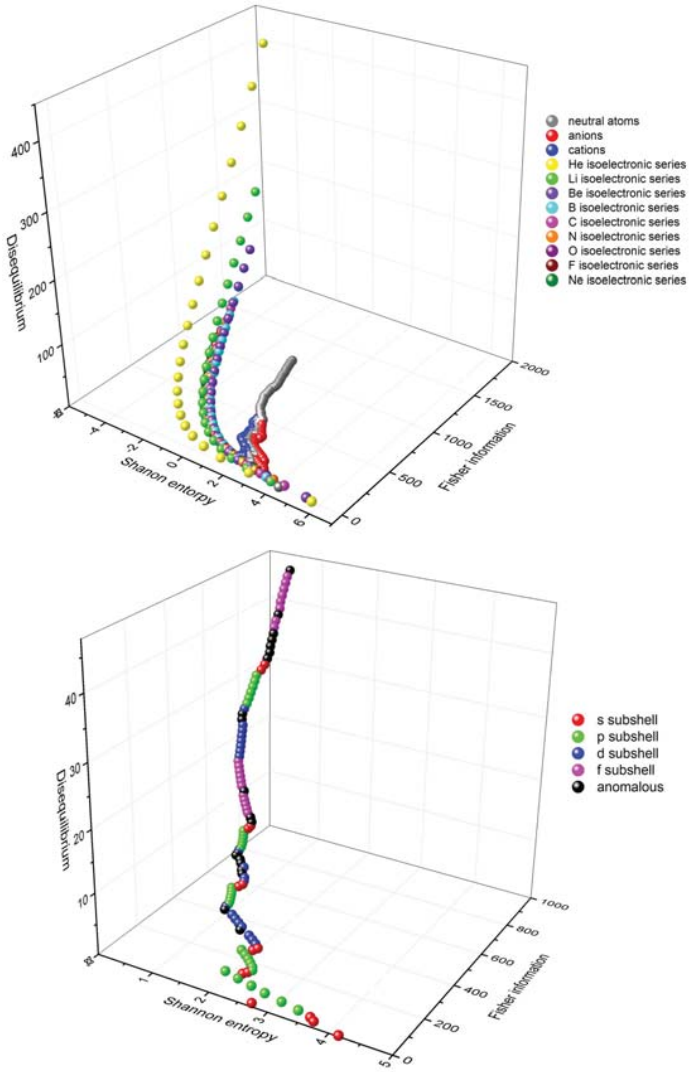


Fig. (1.1). 3*D*-information-theoretic space of 103 neutral atoms, 96 (positively- and negatively- charged) ions from *H* to *Cs*, and 9 isoelectronic series from *He* through *Ne*. Neutral atoms and positively- and negatively-charged ions are depicted by grey, blue and red balls, respectively. Iso-electronic series from $N = 2$ through $N = 10$ are depicted with balls on yellow, green, purple, *etc.* (*top*). 3*D*-information-theoretic space for neutral atoms with valence-shell filling (*bottom*). Atomic units are used.

In contrast, the atomic IT-3D space of the first 103 neutral atoms behaves as shown in Fig. (**1.1**) (bottom), which is featured by an increasing behavior on its *D*- and *I*-axes by showing segments of different slope in such a manner that reveals its periodic nature (colours indicate atoms with similar valence shell: *s, p, d, etc.* according to the code shown at the right of the Figure). Besides, the anomalous shell-filling of atoms is also observed (balls in black); see for instance atoms with $Z = 24, 29, 46$. It is remarkable to note that atoms of the first two periods (*Ne* through *Ne*) hold a somewhat constant value of *uniformity*. More specific observations might be found in Ref. [48].

In Fig. (**1.2**) the entropic IT-3D features together with the *FS* and *LMC* complexity aspects of the neutral atoms previously considered are shown. In this space it is seen that *FS*- and *LMC*-complexities (revealing features of *order-uniformity* and *delocalization-uniformity*) increase with the number of electrons. Note that in the Fig. (**1.2**), higher values of C_{FS} are depicted in reddish and smaller ones in blueish colours, similarly higher values of C_{LMC} are drawn in bigger ball sizes, and smaller ones for lower values. It is apparent that there are no relevant structural differences between complexities based on their global (C_{LMC}) or local (C_{FS}) nature. Besides, it is interesting to note that, with some exceptions, both complexities hold minima for noble gases as well as for atoms with anomalous shell-filling; see for instance atoms with $Z = 24, 29, 46$ (characterized for losing an inner *s* electron). In Fig. (**1.2**) we have depicted the 3*D*-information-theoretic space for neutral atoms with $N = 2$ to $N = 35$ to amply the topological features of lighter atoms. It is worth mentioning that a thorough analysis of the *FS*- and *LMC*-complexities for atomic systems might be found in Ref. [48].

In Fig. (**1.3**), both sets of entropic and complexity features are displayed as well for a great diversity of complex molecular systems; namely, 10 alkanes (squares), 4 ethoxides (squares), 52 organic molecules containing $C-H$, $C-O$, $C-N$, $C-S$ and $C - X$ (X stands for halogens) bonding types (hexagons), 20 amino acids (circles), and 29 molecules of different pharmacological interest: anti-cancer, bactericides, antibiotics, anti-inflammatory and anti-chagas (triangles).

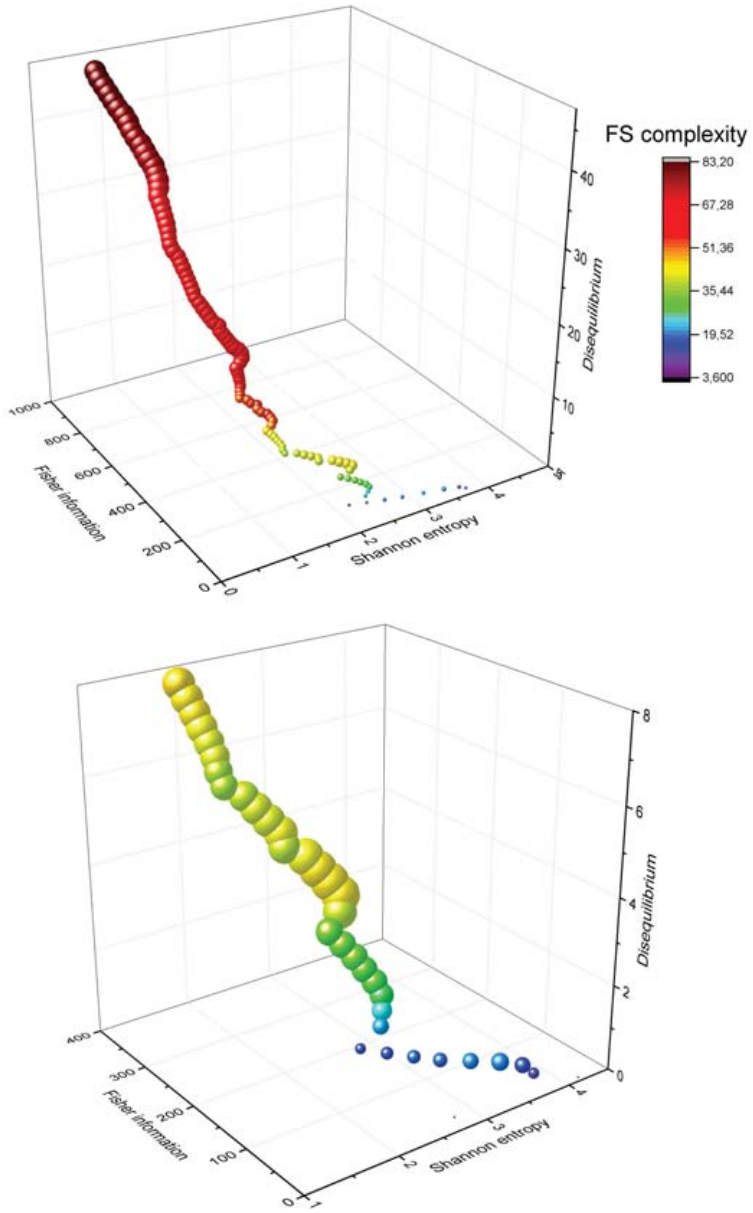


Fig. (1.2). 3*D*-information-theoretic space for neutral atoms. Higher to smaller values of C_{FS} are depicted in reddish to blueish colours. Similarly, higher to lower values of C_{LMC} are drawn with bigger to smaller ball sizes (*top*). Detail of the previous graph for the sequence of lighter atoms with $N = 2$ to $N = 35$ (*bottom*). Atomic units are used.

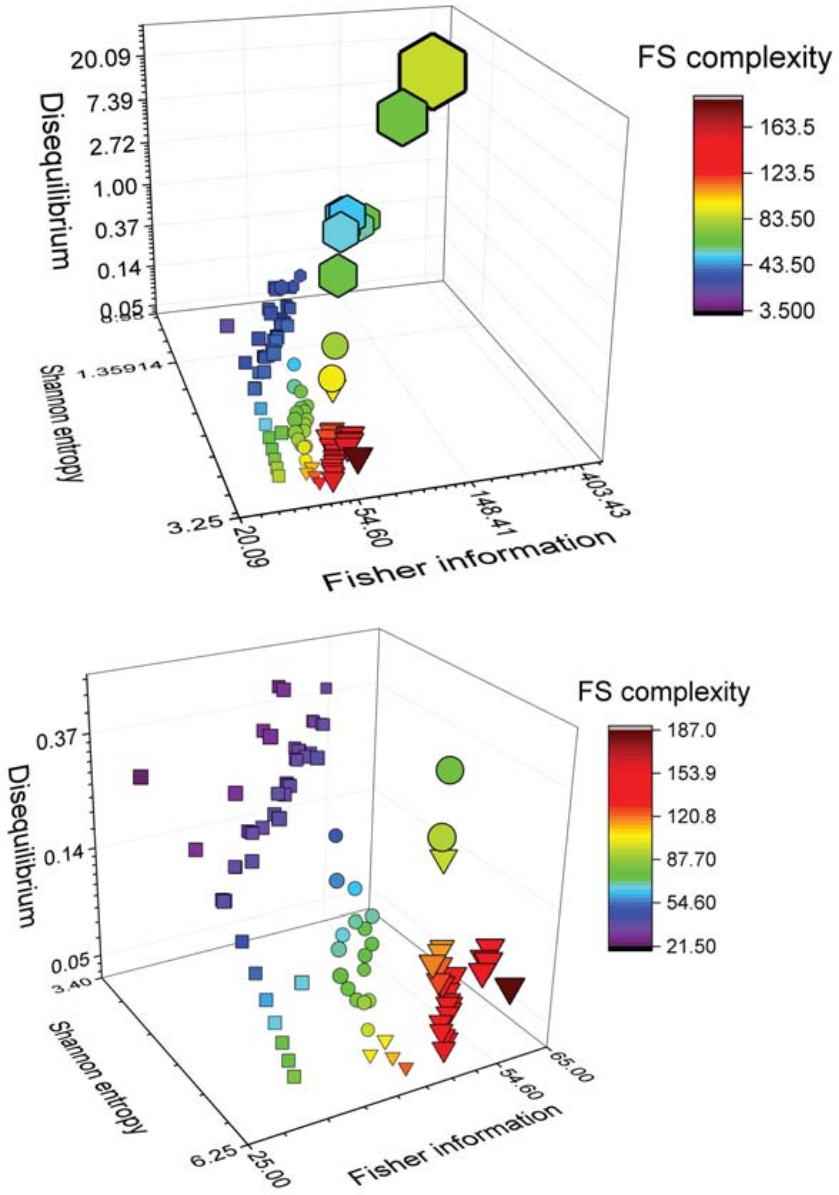


Fig. (1.3). 3*D*-information-theoretic space for organic molecules with *X* =halogens or *S* atoms (hexagons), aminoacids (circles), pharmacological molecules (triangles) and other organic molecules including alkanes and ethoxides (squares). Higher to smaller values of C_{FS} are depicted in reddish to blueish colours. Similarly, higher to lower values of C_{LMC} are drawn with bigger to smaller ball sizes (*top*). Detail of the previous graph for the same molecules except for the set of organic molecules with *X* or *S* atoms (*bottom*). Atomic units are used.

The general and most important observation from the Fig. (**1.3**) is that every set of different kind of molecules possesses a unique region of IT-space, *i.e.*, organic molecules spread from lower to higher values of *D* (departure from uniformity) as compared to the rest of the molecules, remarking that the highest values are found on organic molecules which contain heavier atoms, sulphur and halogens. Particularly interesting is to observe from a more reduced region of the *information-theoretic space* (Fig. **1.3**) that a subset of organic molecules, the alkanes, hold a linear behavior by having almost constant values of *I*, *i.e.*, order. Additionally, the alkanes, the amino acids and the pharmacological molecules show rather specific regions on the IT-space characterized by increasingly higher values of *I*, hence indicating less disordered systems. Indeed, from the alkanes to the pharmacological molecules (from left to right on the Fisher-axis) we observe that the amino acids possess intermediate values in between the alkanes (holding the lowest) to the pharmacological molecules (having the highest ones). The exception being the two amino acids which hold heavier atoms (sulphur), which possess higher values of order and disequilibrium. Moreover, the set of organic molecules seems to have IT qualities that make them behave linearly on the $D-I$ plane, grouping them together in the IT-3D space.

Next, we have extended the analysis of the molecular *information-theoretic space* by considering the information arising from the complexity measures. Here, it is remarkable that the members of each set of molecules above described possess similar complexity values for both C_{LMC} (estimated by a size scale) and C_{FS} (estimated by the colour scale) measures. For instance, all organic molecules depict blueish squares of similar size; all amino acids show greenish-yellowish circles of similar size (except for the ones containing sulphur atom which holds higher values of *LMC*-complexity). Also, the majority of the pharmacological molecules possess reddish symbols of similar size. The exceptions to this behavior are the alkanes and the organic molecules having *S* or *X* atoms, which show a wider range for *FS*-complexity, from blueish to greenish colours, although their *LMC* values seem to be constant.

Let us highlight the fact that aside of the chemical meaning of the different molecular families analysed here, one would think that most of the differences among them come from a standard chemical or biological nomenclature, even in

cases where different subsets contain the same kind of atoms and hence, they should all together be classified in the same group. It is clear from the *information-theoretic space* that these subsets in fact possess physical or chemical features that make them different and hence they occupy different regions, beyond their chemical/biological classification. See for instance the case of the organic molecules and its subset, the alkanes, the organic molecules in comparison to the amino acids subset, or the one formed by the molecules within specific families of pharmacological interest. They indeed behave differently and are grouped together.

1.4. CHEMICAL SPACE OF SELECTED BACTERIOSTATIC SULFONAMIDES

In this section we undertake an investigation of sulfonamide-type of molecules to explore their chemical space region in order to relate information-theoretic quantities with chemical structures and bacteriostatic activity [69]. Moreover, the effect of the functional groups on the spread of the electron density distribution of the derivatives of 4-aminobenzensulfonamide is analysed.

Chemically speaking, sulfonamides constitute excellent model probes because of their structural diversity, as conferred by their substituents, and yet their moderate molecular size permit a variety of theoretical studies of physicochemical and pharmacological interest. On the medical side, sulfonamides represent synthetic antimicrobial agents encompassing wide spectrum against most gram-positive and many gram-negative organisms. They have been typically employed for several clinical purposes, *e.g.*, in the prevention and treatment of bacterial infections, diabetes mellitus, edema, hypertension, among others. The chemical signature of sulfonamides resides on their organic sulfur group $-SO_2NH_2$ and at the same time on their mechanism of action. That is, the bacteriostatic activity of sulfonamides is due to interference with the metabolic processes in bacteria that requires PABA (Fig. **1.4**) in the synthesis of folic acid, and ultimately of purine and DNA [70]. As to the mechanism is concerned, we expect that molecules with bacteriostatic activity might share more similar informational-theoretic properties with PABA than the rest of the molecules.

According to the above, we have undertaken an information-theoretical analysis of sulfonamides with the following goals: (i) to explore their $IT-3D$ region on the chemical space to investigate its link with their chemical structure and bacteriostatic activity, and (ii) to assess the effect of their substituents so as to analyse how they gather in different possible chemical subsets. The selected group of sulfonamide-type of molecules that we employed for the study was chosen as follows: fifteen of them possess bacteriostatic activity and have been studied previously [71], eleven molecules were computationally designed as structural analogues of sulfonamide.

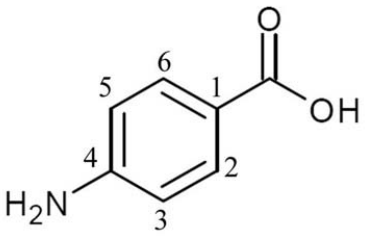


Fig. (1.4). The structural formula for the p-aminobenzoic acid (PABA).

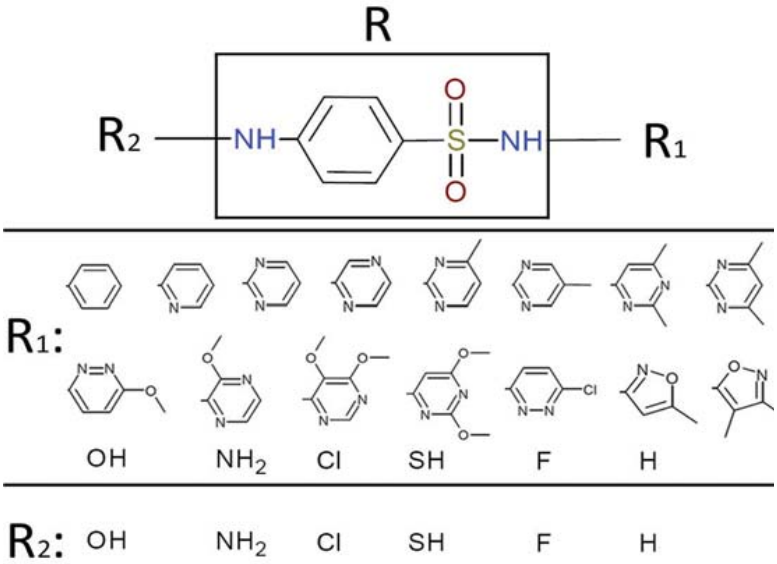


Fig. (1.5). Construction of the molecular models. **R** corresponds to 4-aminobenzensulfonamide; R_1 and R_2 are substituted by different functional groups.

The latter systems have not been reported before to the best of our knowledge [72]. Besides, the 4-aminobenzoic acid (PABA) was also considered (Fig. **1.4**). As mentioned above, chemical analogues to sulfonamide can be constructed by substituting one hydrogen atom in the amino group by a different R_1 group. The chemical structure for these molecules is shown in Fig. (**1.4**). In order to analyse the effect of the *R* group in a deeper way, we have generated these eleven novel compounds with different R groups, which do not present bacteriostatic activity, but their analysis can be useful to describe sulfonamides from an information-theoretic point of view. Results for the information-theoretic analysis of the chemical space for these type of pharmacological molecules is now presented in the rest of this section.

Standard computational programs for electronic structure calculations were employed (G98 and G09 suite of programs [62]) to obtain the molecular wave functions for the sulfonamide-type of molecules analysed here. The chemical structures were optimized at the restricted Hartree-Fock (RHF) level of theory with a 6 − 311 + *G*(*d, p*) basis set. Then, single-point calculation were performed on the optimized structures at the HF/6 − 31 + +*G*(*d, p*) and B3LYP/6 − 31 + +*G*(*d, p*) levels of theory; a frequency analysis was performed so as to corroborate that the obtained structure corresponds to an equilibrium geometry. Details of the electronic calculations can be found in Ref. [63]. Furthermore, molecular electron densities were obtained along with all information-theoretic measures, and complexity dyadic products as defined in the previous section (*i.e.*, *S*, *D*, *I*, C_{LMC} , C_{FS}) by employing software developed in our laboratory along with 3*D* numerical integration routines [67] and the DGRID suite of programs [73].

The 3*D*-information-theoretic (IT-3D) space for all sulfonamide-type molecules and the p-aminobenzoic acid (PABA) is shown in Fig. (**1.6**). Accordingly, we have observed that molecules have grouped together into five different regions of the chemical space. Besides, it is worthy to remark from this Figure that all molecules with bacteriostatic activity belong to only two of the 5 groups. Aside of the sulfonamide molecule, that holds the least bacteriostatic activity, which belongs to a different group (Fig. **1.6**). As expected, the group of molecules with reported bacteriostatic activity is much similar to PABA according to the IT-3D chemical space. It is worthy to mention that grouping observed above has been

validated by use of a different technique employing multivariate mathematical techniques.

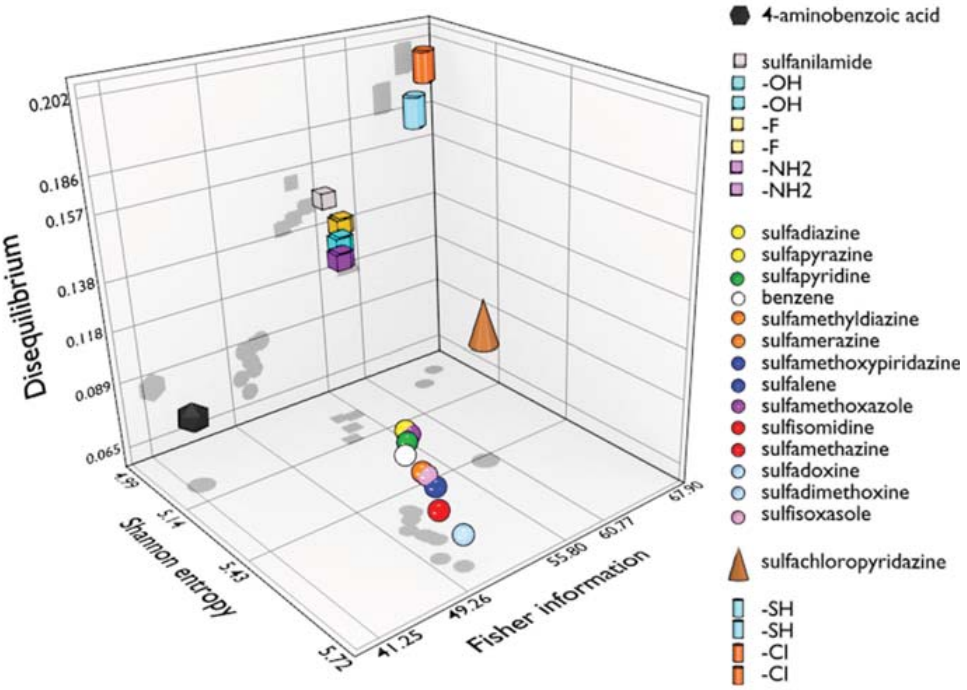


Fig. (1.6). 3*D*-information-theoretic space for the 27 analogues of sulfonamide and the p-aminobenzoic acid (PABA). Each symbol represents one of the five specific regions of the chemical space where the molecules are located. Atomic units are employed.

Furthermore, we have found that concomitant with the IT-3D chemical pattern aforementioned, there is an additional scheme of characterization based upon the R_1 and R_2 groups, *i.e.*, the chemical character of the substituents, regardless of its chemical structure. The most apparent features are reported below:

- It is observed that inclusion of R_1 as an aromatic heterocyclic compound, implies a drastic decrease of the disequilibrium of the molecule, *i.e.*, the electron density becomes closer to uniformity. In these molecules, R_1 corresponds to a pyridine (sulfapyridine) or a diazine (sufadiazine and sulfapyrazine) compounds and R_2 to H atom.
- Molecules with an extra methyl group $-CH_3$ (sulfamerazine, sulfamethyldiazine,

sulfamethazine and sulfisomidina) or a methoxy group $-OCH_3$ (sulfalene, sulfamethoxypyridazine, sulfadimethoxine and sulfadoxine), possess lower values of Fisher information compared to molecules with only no substitution over diazine group. This indicating that these systems are less disordered.

- Differences between molecules with a methyl ($-CH_3$) and a methoxy ($-OCH_3$) group are given by their Shannon entropy. The former have smaller values of *S*, which means their probability density is more localized than the densities of the latter.
- The Fisher information and the disequilibrium decrease with the number of $-CH_3$ or $-OCH_3$ groups present in the molecules. The density probability becomes much more uniform and ordered.
- The rest of the modified molecules show higher disequilibrium and Fisher information values, in turn higher to the sulfanilamide ones. The electron density of these molecules becomes less uniform and more disordered, this behavior is accentuated when the atoms in R_1 and R_2 belong to larger groups.

All compounds considered here are characterized through their information content, featuring any of the three different IT qualities aforementioned, which make them unique among the rest of the structures. An extension of the information-theoretical analysis of the chemical space, is naturally performed by use of complexity concepts through the *FS* and *LMC* dyadic definitions provided above. These measures allow to grasp composite *IT* aspects of the systems (delocalization/uniformity and disorder/uniformity of the electron molecular distributions).

The IT-3D space and the *FS* complexity measure are displayed all together in Fig. (**1.7**) (top). Note that larger values of C_{FS} are depicted in reddish and smaller ones in blueish colours. In Fig. (**1.7**) (bottom) the Fisher-Shannon projection of this IT-3D space is depicted. It is interesting to emphasize that, with some exceptions, both *LMC* and *FS* complexity measures show opposite behaviors. Most relevant aspects of the Figure are reported next:

- The p-amino benzoic acid (PABA), the only natural compound under consideration in this study, has the lowest value of both complexity measures.
- Including a cyclic radical in the molecule increases (decreases) the C_{FS} (C_{LMC})

complexity. Sulfadiazine and sulfapyrazine possess larger (lower) values of C_{FS} (C_{LMC}) compared to the sulfapydirine due to the presence of two nitrogen atoms instead one in the R_1 group.

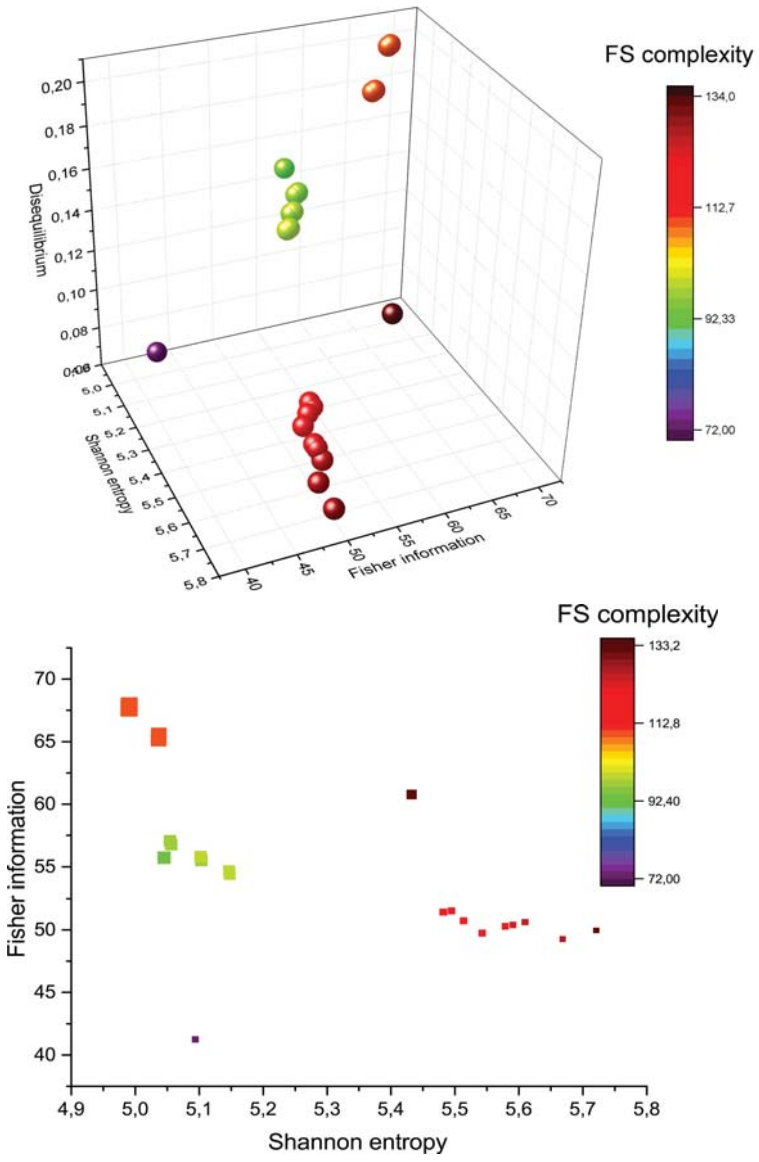


Fig. (1.7). 3*D*-information-theoretic space (*top*) for sulfonamides. Larger values of C_{FS} are depicted in redish colours whereas smaller ones are depicted in blueish ones. Fisher-Shannon plane. (*bottom*) Atomic units are employed.

- Molecules with a methyl group $-CH_3$ (sulfamerazine, sulfamethyldiazine, sulfamethazine and sulfisomidine) or a $-OCH_3$ group (sulfalene, sulfamethoxypyridazine, sulfadimethoxine and sulfadoxine), possess higher (smaller) values of C_{FS} (C_{LMC}) compared to molecules with only a benzene group.
- The number of CH_3 or OCH_3 groups within the molecule increases (decreases) the C_{FS} (C_{LMC}) complexity.

1.5. PREDOMINANT INFORMATION QUALITY SCHEME FOR THE ESSENTIAL AMINO ACIDS

Amino acids are the essential units of biological organisms which are encoded in proteins and nucleic acids (DNA and RNA). These molecules hold complex patterns controlling all biological functions. In recent years, great advances have been achieved to understand how these molecules evolved. For instance, what is appearance order of amino acids over evolution?, how many amino acids appeared at the time when life initiated?, has the number of essential amino acids changed over time?, what is the source of homochirality?. Yet, there are other interesting questions that can be addressed: Is there a fundamental classification of the essential amino acids within the vast world of other biological systems?, is it possible to find a sort of parameters that could describe and classify these systems?, at what extent quantum phenomena, such as entanglement, affect biological systems? These questions await to be properly addressed by standard theories or by use of new approaches. It is our goal to show that concepts of Information Theory constitute such an approach.

Increasing attention has been paid to investigate theoretical and experimental properties of α -amino acids [74]. For instance, studies of ground state energies and ionization potentials of the essential L-α -amino acids were performed [74 - 78]. In attempting to classify amino acids and to be able to differentiate them according to their properties, several classification schemes have been proposed by utilizing different chemical properties (acidity/basicity, hydrophobicity/hydrophilicity, charge/neutrality, polarity/non-polarity, aliphaticity/aromaticity) and also by use of biological functionality: proteic (essential, chemically modified) and non-proteic (D-amino acids, $\alpha-$amino acids, $\omega-$amino acids). In Fig. (**1.8**) a Venn diagram is depicted which shows grouping of amino acids in

accordance to their properties. This scheme was adapted from Refs. [79, 80]. It is difficult to associate relevant physicochemical properties by identifying their similarities, *i.e.*, the substitution of one amino acid by a different residue has a greater chance of being accepted if the two molecules possess similar properties. Accordingly, Taylor adopted a graphical approach to develop Venn diagrams of amino acids families. The unions and intersections within the subsets of the Venn diagram allow characterization of groups of amino acids that might be gather together from similar structural reasons [80]. More recently, Kosiol *et al.* developed a criterion to classify amino acids, starting from their substituting matrices, which utilize a Markov model of protein evolution [81]. Classification of amino acids according to the Taylor [80, 79] critera is one of the most accepted methods in Biochemistry.

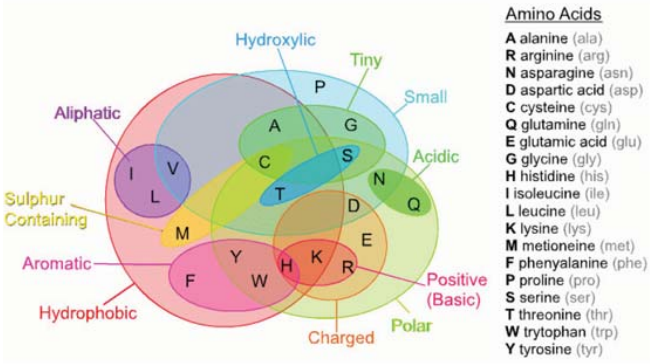


Fig. (1.8). Venn diagram of amino acids according to the representations given in Refs. [80, 79].

To the best of our knowledge, there have been no attempts to characterize the essential amino acids by use of information-theoretic measures from protein data. The material of this section is presented by following these goals: (i) to characterize amino acids according to its information content (delocalization, narrowness, and order), (ii) to design a scheme which associate amino acids based upon its biochemical properties, and (iii) to investigate possible links among information measures and reactivity parameters. Technicalities for this work might be found in Ref. [46].

To the purpose of characterizing the information-theoretic measures of amino acids under study, we have calculated several reactivity properties, *e.g.* the ionization potential (IP), the total dipole moment, the hardness (η) and the electrophilicity index (ω). To be consistent, we have performed numerical calculations to obtain molecular orbital energies at the density functional theory (DFT) level, by utilizing Janak's theorem [82], which is analogous to the Koopmans' theorem [83] at the Hartree-Fock level, hence conceptual DFT properties are straightforwardly obtained. Hardness (η) is computed in this framework [51] through

$$\eta = \frac{1}{S} \sim \frac{\varepsilon_{LUMO} - \varepsilon_{HOMO}}{2}, \tag{1.8}$$

where ε stands for the frontier molecular orbital energies and S denotes the softness of the system. It is worth mentioning that the factor 1/2 in Eq. (1.8) is set only for symmetrical ressemblance with the chemical potential [51]

$$\mu = -\left(\frac{\partial E}{\partial N}\right)_{\mu(\mathbf{r})} = -\frac{\varepsilon_{LUMO} + \varepsilon_{HOMO}}{2}, \tag{1.9}$$

and it has been already disowned [84] though. Generally speaking, chemical hardness and softness are good indicators of chemical reactivity. Hardness has been employed as an indicator of chemical reactivity since it measures the resistance of the electron density to change [84, 85]. Thus, molecules with larger values of η are less reactive molecules. The S index quantifies the polarizability of the molecule [86]; thus, soft molecules are more polarizable and then prone to acquire electronic charge [87]. These concepts form part of the reactivity theory based upon the hard and soft acids and bases principle [88].

The electrophilicity index [89] is defined in terms of its chemical potential and its hardness as

$$\omega = \frac{\mu^2}{2\eta}. \tag{1.10}$$

According to conceptual DFT, electrophilicity is a good descriptor of chemical

reactivity, quantifying the global electrophilic power of the molecules (proneness to acquire additional electronic charge) [87].

Electron structure calculations for the amino acids under study were performed by use of Gaussian 09 suite of programs [62] at the HF, DFT-M062X and CISD(Full) levels of theory on the 6 − 311 + *G*(*d, p*) basis set. Geometric data for the 18 amino acids were obtained from bacteriorhodopsin protein (Protein Data Bank ID: 1C3W [90]) at five different random locations of the protein for each amino acid. This means that 5 different energy conformations were chosen directly from the natural protein (exception being for cysteine and histidine that are not present within the protein). Of course, in order to maintain the natural conformations of the protein for all the 18 aminoacids, we keep the geometric skeleton for each conformation. Therefore, geometry optimizations were constrained to optimize the attached hydrogen atoms only. Optimization was performed at the HF level with the basis set above mentioned, afterwards, single point calculations were performed at the CISD(Full) level of theory. Finally, information-theoretic measures were obtained at the CI level by employing software developed in our laboratory along with 3D numerical integration routines [67] and the DGRID suite of programs [73]. The values of the DFT conceptual properties have been obtained at the *M*062*X*/6 − 311 + *G*(*d, p*) level of theory.

In spite of the fact that a great variety of biological molecules from living organisms can be found in the Protein Data Bank (RCSB PDB, Cambridge University) [90], protein known as bacteriorhodopsin (PDB ID: 1C3W) [91] is a very interesting molecular system in itself. Bacteriorhodopsin, a retinal protein, is the major photosynthetic protein of the archaeon *Halobacterium salinarum*. It converts "green" light energy (500-650 nm, max 568 nm) into an electrochemical proton gradient, which in turn is used for ATP production by ATP synthase. Hence the protein works a light-driven machine generating protons, for transporting out of the cell [92]. In particular, the protein possesses more than 250 amino acids along its tertiary structure. It holds 18 different essential amino acids. Our interest is focused on analyzing different conformations of the same amino acid (*i.e.*, at different locations within the protein) to characterize them within its natural environment. From a biological viewpoint, the amino acid structure strongly depends on its environment (*i.e.*, the interaction with its neighbouring

amino acids), as well as on the interactions between the molecular segments surrounding a specific amino acid.

Several studies have shown that electron delocalization along the side chains of peptides and proteins could be caused by a substituent group along the peptide backbone [93, 65]. Moreover, the geometry of the conformeric structures along the protein could be affected by hydrogen-bonding from the functional groups of the side chains of the protein [94, 95]. Thus, all physicochemical properties of amino acids also depend on these factors. Further, it is likely that a given property of an amino acid does not represent its state fully, because of its conformational diversity caused by neighbouring residues; hence the property will not be adequate to characterize that particular amino acid within the protein. In contrast, we anticipate that information-theoretic measures might have a much more stable behavior to characterize amino acids within the protein environment, in spite of its conformeric diversity. This is because information-theoretic measures grasp the essential features of the electron densities of systems.

Fig. (**1.9**) shows bar plots representing the value of a given property. This is done for all amino acids of the 1C3W protein in their chosen five conformeric structures. We plotted different physicochemical quantities (such as energy (E), chemical potential (μ) and hardness (η)) and information-theoretic measures (such as Shannon entropy (S), Fisher information (I) and disequilibrum (D)). As expected, we observe that properties such as chemical potential and hardness (in green and yellow), possess numerical values with larger dispersions, since these properties strongly depend on the conformeric geometry of amino acids, as compared with the energy and the information-theoretic measures (in red and blue), which are fairly constant for all conformers. For instance, take the chemical potential, its values for the different conformations for lysine (K) and phenilalanine (F) are so similar that we can not distinguish them in terms of this property. In contrast, information-theoretic measures produce much lesser dispersion as compared to DFT properties, which allow characterizing the amino acid families in a very accurate way. We note that energy values behave in a similar fashion. It is important to note that the rest of the DFT properties (softness, ionization potential and electrophilicity) behave in a way that larger dispersion values are observed, these are not depicted in the Figure though.

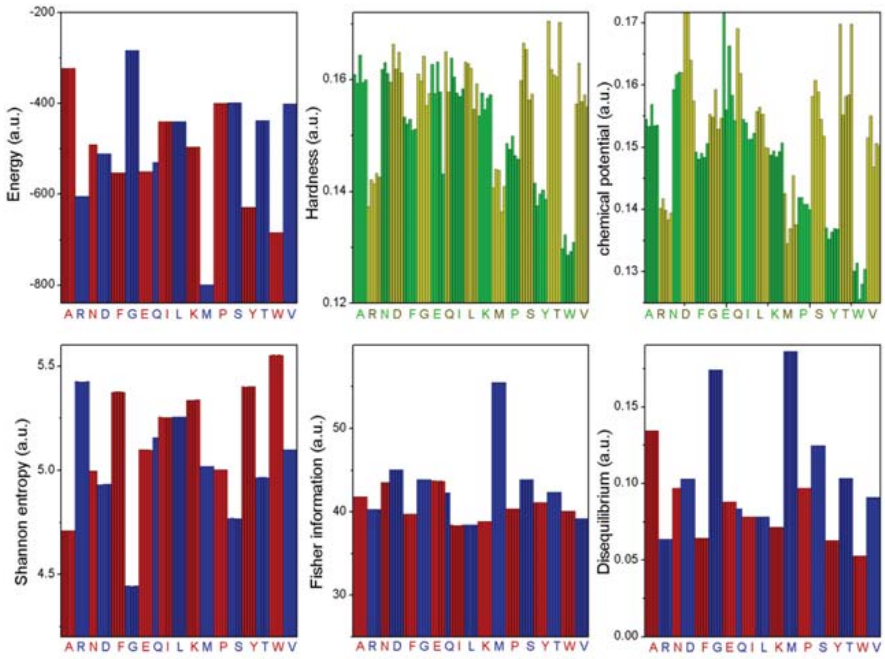


Fig. (1.9). DFT properties (on green and yellow) along with the Energy and information-theoretic measures (on red and blue) are depicted for the conformeric structures of all amino acids of the protein. Atomic units are used.

Basic notions of IT discussed above show that there are three different aspects to be emphasized from information-theoretic measures: delocalization (*S*), uniformity (*D*) and narrowness (*I*). It has been shown that all chemical structures possess different chemical or biological properties that arise from their electronic densities, and these are reflected through Shannon entropy, disequilibrium and Fisher information, respectively. Therefore, we have state that all amino acids might be characterized through their information content, through any and all of the three different qualities of "information", which characterize any amino acid to be unique among the rest of the chemical structures. Based on this observation, average values of the three information measures for each amino acid are plotted in Fig. (**1.10**); note that these average values correspond to all the conformeric structures. Let us recall as we discussed above that the dispersion of these values is very small (Fig. **1.9**) so as to justify the use of the average.

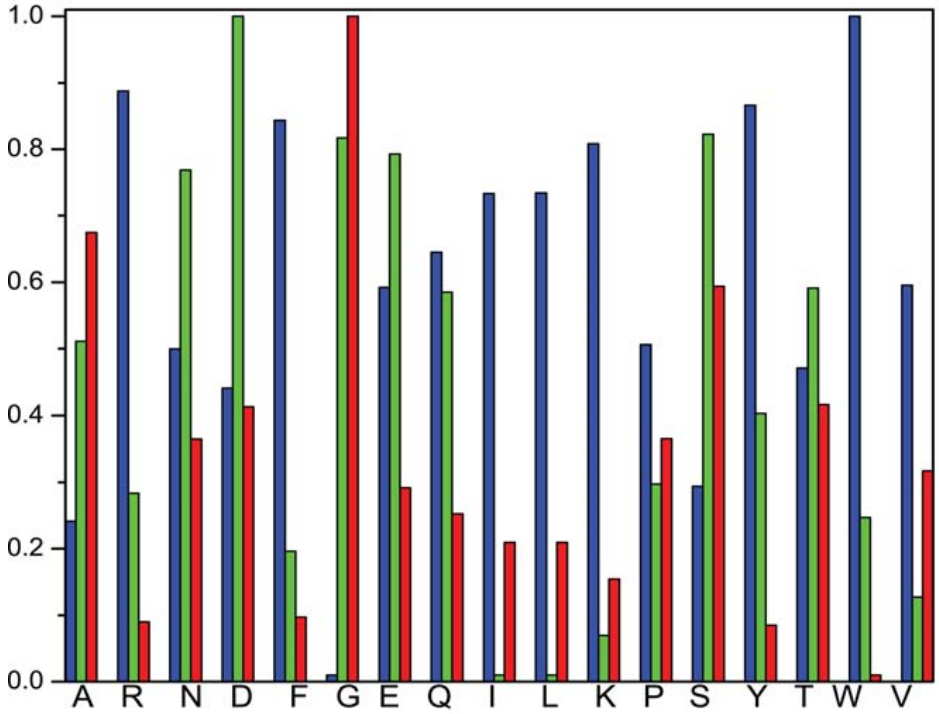


Fig. (1.10). Average values (of all conformeric structures for each amino acid) for Shannon entropy (blue), Fisher information (green) and disequilibrium (red). Note that a range-scaling-pretreatment-data method has been used according to $(x_i - x_{min}) / (x_{max} - x_{min})$ within the range $[10^{-2}, 1]$.

Accordingly, we may classify the 18 essential amino acids through the information quality scheme given in Fig. (**1.11**) where each column contains different amino acid structures characterized by three information-theoretic qualities (delocalization (*S*), uniformity (*D*) and narrowness (*I*)) ordered in a certain manner such that the "information" aspects appear in a given triplet by its relative contribution; *e.g.*, the set of five amino acids in the first column headed by the triplet *SID* reveal that Shannon entropy is the most relevant IT quality among the three, followed in turn by the disequilibrium and then by Fisher information, in decreasing order of relative contribution.

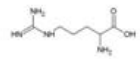
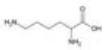
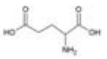
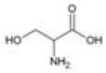
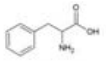
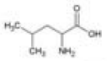
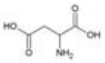
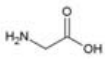
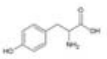
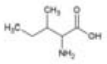
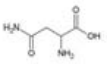
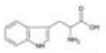
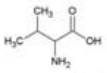
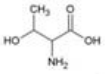
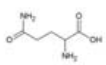
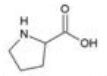
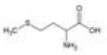
SID	SDI	ISD	IDS	DIS
 Arginine (R)	 Lysine (K)	 Glutamic acid (E)	 Serine (S)	 Alanine (A)
 Phenilalanine (F)	 Leucine (L)	 Aspartic acid (D)		 Glycine (G)
 Tyrosine (Y)	 Isoleucine (I)	 Asparagine (N)		
 Triptofane (W)	 Valine (V)	 Threonine (T)		
 Glutamine (Q)	 Proline (P)	 Methionine (M)		

Fig. (1.11). Amino acids classification according to its predominant information quality scheme (PIQS): Shannon entropy (S), Fisher information (I) and disequilibrium (D).

(i) *Delocalization (Shannon entropy) as the predominant information quality*

For the amino acids showed in the first two columns of Fig. (**1.11**), Shannon entropy appears to be the predominant quality, *i.e.*, delocalization of their electron density contributes more than the other two qualities: narrowness and uniformity. Note that delocalization (by considering this information quality only) fully characterize amino acids of the aromatic (column 1) and aliphatic (column 2) type. It is worth mentioning that we may resolve the difference between these two groups in terms of the other two information qualities, narrowness and uniformity. That is, the aromatic amino acids are featured by their narrowness (of their electronic density) which contributes relatively more than it does for the aliphatic ones, *i.e.*, the narrowness is the second more relevant information quality for the aromatic ones, and then uniformity. On the other hand, aliphatic amino acids

show uniformity as the second predominant quality, followed by narrowness as the least information quality. Some comments are pertinent:

– Note that in column 1 arginine and glutamine are also included. Despite these amino acids are not always considered as aromatic molecules in some classifications (*e.g.*, Lehninger [96], Taylor [80] and Livingstone [79]), they show tautomeric effects in the substituent group, meaning that for all the amino acids of this group, delocalization of their electronic density is yet the most important chemical effect.

– Proline pertains to the second group in column 2 and for the majority of the classifications this amino acid is not considered as aliphatic (*e.g.*, Lehninger [96], Taylor [80], Livingstone [79] and [81] classifications), its molecular structure has attached a cyclic aliphatic portion.

(ii) *Narrowness (Fisher information) as the predominant information quality*

Chemical structures shown in columns 3 and 4, larger values for Fisher, *i.e.*, narrowness (order) prevails over delocalization and uniformity. Closer look of the chemical structures sketched in Fig. (**1.11**), that the common chemical feature these molecules have in common resides on the strong electronegativity of oxygen (aspartic and glutamic acid, threonine asparagine and serine) and sulfur (methionine) provoking that the electron density of these amino acids gets ordered/narrowed. Differences between these two groups are understood from the other information qualities, delocalization and uniformity. For the amino acids showed in column 3, delocalization contributes more than for the serine amino acid in column 4, *i.e.*, the Shannon entropy is the second more relevant information quality. For the serine amino acid (column 4), delocalization is less important than its uniformity. Due to its small size, the departure from uniformity is larger than for amino acids located in column 3.

(iii) *Uniformity (disequilibrium) as the predominant information quality*

Column 5 shows results for tiny amino acids according to Livingstone [79], alanine and glycine, indicating that uniformity is the quality that contributes the most; *i.e.*, departure from uniformity prevails over delocalization and narrowness.

Thus, these two amino acids do not possess any electron attractor atoms/groups in the side chains. Note that alanine and glycine, along with serine (column 4) have been classified (*e.g.*, Livingstone classifications [79]) as tiny molecules with smaller delocalization. This explains why delocalization is the least contributing information quality for these molecules.

It is important to emphasize that the ordering found for the conformeric structures of selected amino acids is based upon a scheme of quality predominance of information. Indeed, Fig. (**1.12**) shows the five conformers of methionine through residues met-118, met-209, met-20, met-32 and met-60.

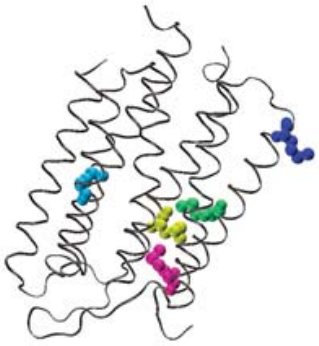


Fig. (1.12). View of the different conformeric structures of methionine at residues met-118, met-209, met-20, met-32 and met-60 for 1C3W protein.

Analysis of predominant information quality provides a useful framework to study amino acids within their natural biochemical boundaries. As shown in Fig. (**1.13**), some examples were chosen: glycine, methionine and tryptophan. We might observe from the Figure that the information-theoretic values for each conformer vary according to the location of the residue in the protein, revealing the biochemical environment of each amino acid.

We have found worth describing some of the most representative chemical classifications that are employed in the literature to characterize amino acids in terms of different criteria:

- *Lehninger* [96]: aliphatic (isoleucine, leucine, valine, glycine), aromatic (phenylalanine, tryptophan, tyrosine), polar uncharged (proline, asparagine, threonine, glutamine, serine), acids (glutamic acid, aspartic acid) and basic

(lysine, arginine).

- *Livingston* [79]: aliphatic (isoleucine, leucine, valine), aromatic (phenylalanine, tyrosine, tryptophan, histidine), sulfur containing (methionine, cysteine), hydrophobic (isoleucine, leucine, valine, methionine, cysteine, phenylalanine, tyrosine, tryptophan, histidine, threonine, alanine, lysine), hydroxylic (threonine, serine), charged (histidine, lysine, arginine, glutamic acid, aspartic acid), basic (histidine, lysine, arginine), polar (histidine, lysine, arginine, glutamic acid, aspartic acid, asparagine, glutamine, threonine, serine, cysteine, tyrosine, tryptophan), small (proline, alanine, glycine, serine, cysteine), tiny (alanine, glycine, serine, cysteine).
- *Hydrophobicity-based* [97]: hydrophobic (isoleucine, leucine, valine, alanine, glycine, proline, phenylalanine, tryptophan), hydrophilic uncharged (tyrosine, threonine, serine, asparagine), hydrophilic charged acids (glutamic acid, aspartic acids), hydrophilic charged basis (glycine, arginine).

Inspection of the list above indicates that various of the amino acids participate in several groups which represent the multicharacter of the chemical role that they play within the proteins. However it has been difficult to establish a general classification which might embrace many of these characteristics in a single property. One of the goals of our study resides in proposing a general classification which represents all the physicochemical features by means of information-theoretic measures.

Fig. (**1.14**) shows the amino acids classification proposed in this work by following the predominant information quality scheme discussed above (Fig. **1.11**). We may note that neither of the former chemical classifications on the left are grouped together, which means that there is no simple scheme that allows to classify amino acids in terms of specific chemical properties (see description above). In contrast it is readily seen that classification according to PIQS allows an easy manner of classifying amino acids within major chemical groups: delocalized (aromatic and aliphatic), narrowed (with electro-attractor atoms) and uniform (tiny); this in turn embrace all chemical species discussed above.

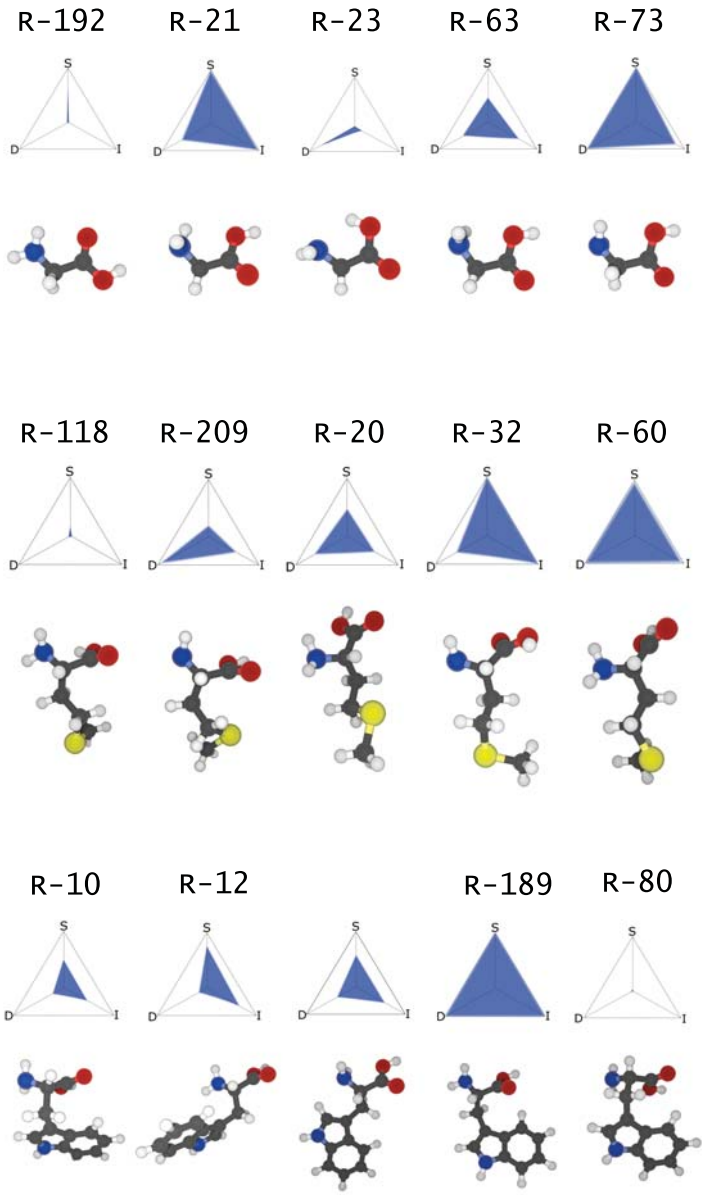


Fig. (1.13). Values of Shannon entropy (S), Fisher information (I) and Disequilibrium (D) for each conformeric structure of glycine (top), methionine (middle) and tryptophane (bottom). Note that radial variables are used by employing a range scaling pretreatment data method according to $(x_i - x_{min}) / (x_{max} - x_{min})$ within the range $[10^{-2}, 1]$. The chemical structures (in tubes) are also shown to guide the eye.

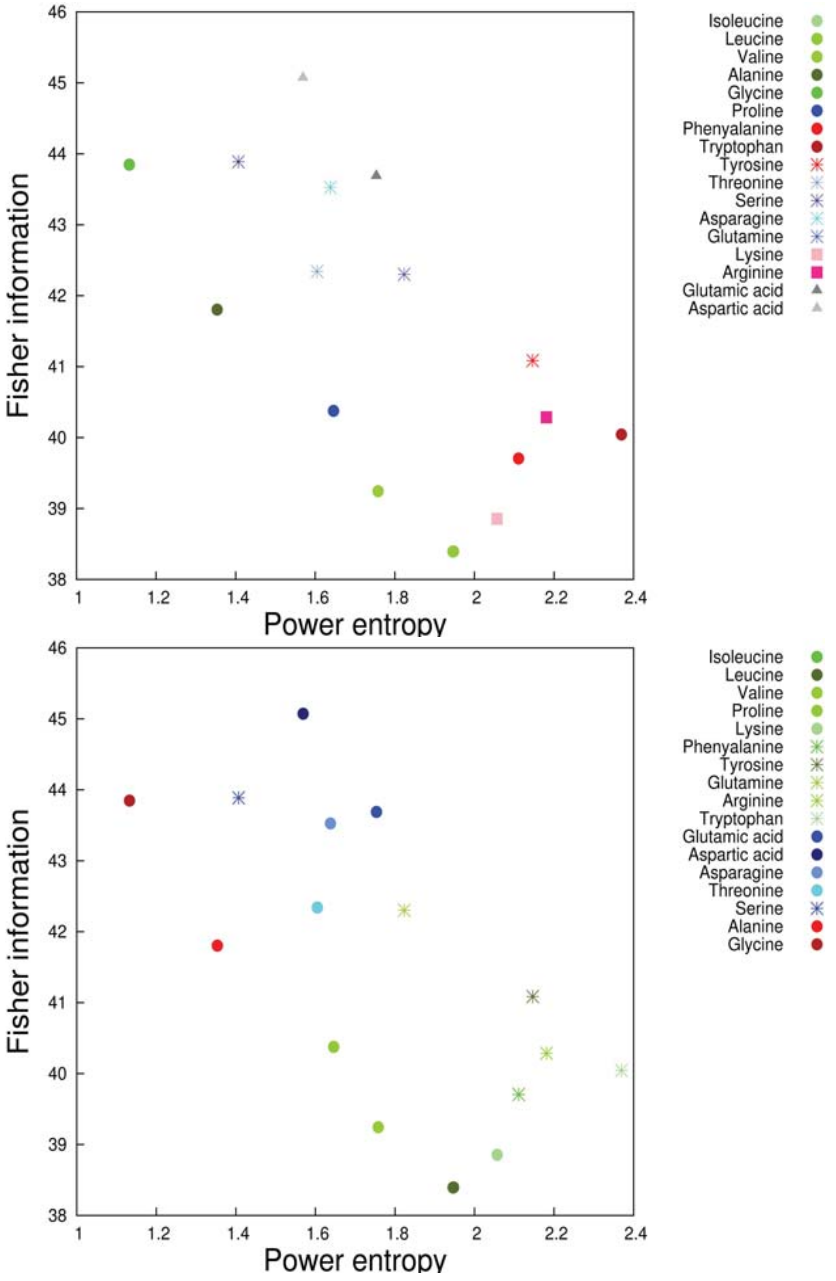


Fig. (1.14). Fisher-Shannon plane I-J for all amino acids. Amino acids in the figure (top) follow Lehninger's classification [96] (colored symbols) and also the hydrophobicity-based classification (geometric symbols). Amino acids (on the bottom) were classified according to the PIQS classification (above in Fig. (**1.11**)): SID (green circles), SDI (green stars), ISD (blue circles), IDS (blue stars) and DIS (red circles).

CONCLUSION

Throughout this Chapter, we have discussed the relevance of employing the *information-theoretical space* (IT-3D) with the following three axes: disequilibrium D, Shannon entropy S, and Fisher information I, which characterize physical features such as *uniformity*, *delocalization* and *order* respectively. Each point of this space corresponds to a natural system characterized by three macroscopic features: *uniformity*, *delocalization* and *order*, respectively. In addition, for completeness, their corresponding dyadic products of complexity (*LMC* and *FS*) are given. As revealed by the information-theoretical space, regions of physical, chemical or biological qualities are indeed differentiated, containing the numerous physical and chemical systems considered in this chapter. We have focused our attention in a large variety of many electron systems of physical, chemical, biological and pharmacological interest (388 atoms and 115 molecules), which accounts for neutral and ionized atoms, amino acids, and molecules with pharmacological interest such as the sulfonamides-type, among other organic molecules described above.

Three most interesting conclusions of this review are in order. First, the topological features mentioned above look very appealing for collecting all assortments of simple and complex systems so as to devise a general map based upon *information-theoretic concepts*. Furthermore, this *information-theoretic chemical space* might be used as a general approach for the classification of all atomic and molecular systems. Novel information-theoretic regions have been described in this study, which can be used in conjunction with standard semi-empirical techniques (*e.g.*, QSAR) to describe difference in pharmacological values of biological molecules.

Second, the sulfonamide-type of molecules group into five different regions of the information-based chemical space mentioned above. All molecules with bacteriostatic activity belong to two of these groups, aside of the sulfonamide molecule which possesses the least bacteriostatic activity and belongs to different one. We have found that concomitant with the IT-3D chemical pattern there is an additional characterization based upon the R_1 and R_2 groups, *i.e.*, the chemical character of the substituents, regardless of its conformeric structure. Throughout

the study we were able to associate some structural aspects and information-theoretical properties arisen from the IT-3D chemical space to the bacteriostatic activity of these molecules. An interesting result of the analysis reveals that bacteriostatic reported molecules are closer to the 4-aminobenzoic acid that the ones they were theoretically designed for this study.

Third, we have performed an information-theoretic analysis of 18 selected amino acids (AA) obtained from a data set of natural proteins, bacteriorhodopsin (1C3W), from the Protein Data Bank. Several conformations for each amino acid were studied as they were found in the original biochemical environments. Results show that information measures uniquely characterize all amino acids through a predominant information quality scheme (PIQS) which gathers all chemical families by means of three major information-theoretic features: delocalization, narrowness and uniformity. This scheme allows us to recognize four major chemical families: aliphatic (delocalized), aromatic (delocalized), electro-attractive (narrowed) and tiny (uniform). These in turn complement and embrace all chemical families recognized by other classifications, *i.e.*, containing sulfur, hydrophobic, hydroxylic, charged, basic, polar, small and tiny. Besides, we have also observed patterns given by dyadic products of information-theoretic measures that support the PIQS classification proposed in this work through the four major chemical features above mentioned. Finally, we have observed that, concomitant to the conformeric structures of the same amino acid pertaining to different residues of the protein, there is a pattern of stability for the information-theoretic measures and also for the energy. In contrast, it is also observed that all of the chemical properties (softness, ionization potential, chemical potential, electrophilicity) analysed in this work are highly sensitive to the chemical environment within the protein.

Finally, based on these results we can state that information-theoretic concepts provide an alternative description of the chemical space. Therefore, we propose that all natural systems could be characterized in terms of their information-theoretical quantities and grouped in different regions of this *information-theoretic chemical space* which can be associated to their physicochemical properties.

DISCLOSURE

Part of this chapter has been reproduced from authors previous publication entitled Predominant Information Quality Scheme for the Essential Amino Acids: An "Information-Theoretical Analysis" published in *ChemPhysChem*, 2015. Available at http://onlinelibrary.wiley.com/doi/10.1002/cphc.201500282/abstract.

CONFLICT OF INTEREST

The authors confirm that they have no conflict of interest to declare for this publication.

ACKNOWLEDGEMENTS

We acknowledge partial financial support through the UGR-GENIL 2015 and Mexican grants from CONACyT (grant number 80244) and the Spanish MINECO project FIS2014-59311-P (cofinanced by FEDER), and the grants FQM-7276, FQM-020, FQM-207 and FQM-239 of Junta de Andalucía (Spain).

REFERENCES

[1] Cai, J.; Plenio, M.B. Chemical compass model for avian magnetoreception as a quantum coherent device. *Phys. Rev. Lett.,* **2013**, *111*(23), 230503. [http://dx.doi.org/10.1103/PhysRevLett.111.230503] [PMID: 24476240]

[2] Bandyopadhyay, J.N.; Paterek, T.; Kaszlikowski, D. Quantum coherence and sensitivity of avian magnetoreception. *Phys. Rev. Lett.,* **2012**, *109*(11), 110502. [http://dx.doi.org/10.1103/PhysRevLett.109.110502] [PMID: 23005606]

[3] Arndt, M.; Juffmann, T.; Vedral, V. Quantum physics meets biology. *HFSP J.,* **2009**, *3*(6), 386-400. [http://dx.doi.org/10.2976/1.3244985] [PMID: 20234806]

[4] Hansch, C.; Fujita, T. p-σ-π analysis. a method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.,* **1964**, *86*, 1616-1626. [http://dx.doi.org/10.1021/ja01062a035]

[5] Hammett, L.P. Hammett. The effect of structure upon the reactions of organic compounds. Benzene derivatives. *J. Am. Chem. Soc.,* **1937**, *59*(1), 96-103. [http://dx.doi.org/10.1021/ja01280a022]

[6] Hansch, C. Quantitative approach to biochemical structure-activity relationships. *Acc. Chem. Res.,* **1969**, *2*(8), 232-239. [http://dx.doi.org/10.1021/ar50020a002]

[7] Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; Consonni, V.; Kuzmin, V.E.; Cramer, R.; Benigni, R.;

Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.,* **2014**, *57*(12), 4977-5010. [http://dx.doi.org/10.1021/jm4004285] [PMID: 24351051]

[8] Dobson, C.M. Chemical space and biology. *Nature,* **2004**, *432*(7019), 824-828. [http://dx.doi.org/10.1038/nature03192] [PMID: 15602547]

[9] Reymond, J.L.; Van Deursen, R.; Lorenz, C.; Ruddigkeit, L. Chemical space as a source for new drugs. *Med. Chem. Comm.,* **2010**, *1*(1), 30-38. [http://dx.doi.org/10.1039/c0md00020e]

[10] Reymond, J.L.; Awale, M. Exploring chemical space for drug discovery using the chemical universe database. *ACS Chem. Neurosci.,* **2012**, *3*(9), 649-657. [http://dx.doi.org/10.1021/cn3000422] [PMID: 23019491]

[11] Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature,* **2004**, *432*(7019), 855-861. [http://dx.doi.org/10.1038/nature03193] [PMID: 15602551]

[12] Xiao, D.; Yang, W.; Beratan, D.N. Inverse molecular design in a tight-binding framework. *J. Chem. Phys.,* **2008**, *129*(4), 044106. [http://dx.doi.org/10.1063/1.2955756] [PMID: 18681633]

[13] Balawender, R.; Welearegay, M.A.; Lesiuk, M.; De Proft, F.; Geerlings, P. Exploring chemical space with the alchemical derivatives. *J. Chem. Theory Comput.,* **2013**, *9*(12), 5327-5340. [http://dx.doi.org/10.1021/ct400706g] [PMID: 26592270]

[14] von Lilienfeld, O.A. First principles view on chemical compound space: Gaining rigorous atomistic control of molecular properties. *Int. J. Quantum Chem.,* **2013**, *113*, 1676-1689. [http://dx.doi.org/10.1002/qua.24375]

[15] Shukla, V.K.; Sachan, A.K.; Pathak, S.K.; Srivastava, R.; Prasad, O.; Sinha, L. Prediction of molecular properties and spectroscopic profile of Riluzole with different functionals (B3LYP, M06-2X, MPWLYP): A combined theoretical and experimental study. *J. Mol. Struct.,* **2016**, *1106*, 265-276. [http://dx.doi.org/10.1016/j.molstruc.2015.10.088]

[16] Finkelmann, A.R.; Göller, A.H.; Schneider, G. Robust molecular representations for modelling and design derived from atomic partial charges. *Chem. Commun. (Camb.),* **2016**, *52*, 681-684. [http://dx.doi.org/10.1039/C5CC07887C] [PMID: 26568131]

[17] Gatenby, R.A.; Frieden, B.R. Information theory in living systems, methods, applications, and challenges. *Bull. Math. Biol.,* **2007**, *69*(2), 635-657. [http://dx.doi.org/10.1007/s11538-006-9141-5] [PMID: 17083004]

[18] Esquivel, R.O.; Angulo, J.C.; Dehesa, J.S.; Antolín, J.; López-Rosa, S.; Flores-Gallegos, N.; Molina-Espíritu, M.; Iuga, C.; Martínez-Carrera, E. *Recent Advances Toward the Nascent Science of Quantum Information Chemistry*; Nova publisher, **2012**.

[19] Molina-Espiritu, M.; Esquivel, R.O.; Angulo, J.C.; Dehesa, J.S. Concurrent phenomena at the reaction path of the $S_N 2$ reaction $CH_3Cl + F^-$. information planes and statistical complex-ity analysis. *Entropy (Basel),* **2013**, *15*(10), 4084-4104. [http://dx.doi.org/10.3390/e15104084]

[20] Liu, S.; Rong, C.; Lu, T. Information conservation principle determines electrophilicity, nucleophilicity, and regioselectivity. *J. Phys. Chem. A,* **2014**, *118*(20), 3698-3704. [http://dx.doi.org/10.1021/jp5032702] [PMID: 24784465]

[21] Zhou, X.; Rong, C.; Lu, T.; Liu, S. Hishfeld charge as a quantitative measure of electrophilicity and nucleophilicity: nitrogen-containing systems. *Wuli Huaxue Xuebao,* **2014**, *30*, 2055-2062.

[22] Liu, S. Where does the electron go? The nature of ortho/para and meta group directing in electrophilic aromatic substitution. *J. Chem. Phys.,* **2014**, *141*(19), 194109. [http://dx.doi.org/10.1063/1.4901898] [PMID: 25416876]

[23] Schneider, T.D. A brief review of molecular information theory. *Nano Commun. Netw.,* **2010**, *1*(3), 173-180. [http://dx.doi.org/10.1016/j.nancom.2010.09.002] [PMID: 22110566]

[24] Nagy, A. Fisher information in density functional theory. *J. Chem. Phys.,* **2003**, *119*, 9401. [http://dx.doi.org/10.1063/1.1615765]

[25] Nalewajski, R.F. Information principles in the theory of electronic structure. *Chem. Phys. Lett.,* **2003**, *372*, 28. [http://dx.doi.org/10.1016/S0009-2614(03)00335-X]

[26] Fisher, R.A. Theory of statistical estimation. *Proc. Cambridge Phil. Soc.,* **1925**, *22*, pp. 700-725. Reprinted in Collected Papers of R. A. Fisher, edited by J.H. Bennet (University of Adelaide Press, South Australia), 1972, pp. 15–40.

[27] Frieden, B.R. *Science from Fisher Information*; Cambridge University Press: Cambridge, **2004**. [http://dx.doi.org/10.1017/CBO9780511616907]

[28] González-Férez, R.; Dehesa, J.S. Characterization of atomic avoided crossings by means of fishers information. *Eur. Phys. J. D,* **2005**, *32*, 39. [http://dx.doi.org/10.1140/epjd/e2004-00182-3]

[29] Esquivel, R.O.; Liu, J.C.; Angulo, S.; Dehesa, J.S.; Antolín, J.; Molina-Espıritu, M. Fisher information and steric effect: study of the internal rotation barrer of ethane. *J. Phys. Chem. A,* **2011**, *115*, 4406-4415. [http://dx.doi.org/10.1021/jp1095272] [PMID: 21473613]

[30] Gatenby, R.A.; Frieden, B.R. Application of information theory and extreme physical information to carcinogenesis. *Cancer Res.,* **2002**, *62*(13), 3675-3684. [PMID: 12097274]

[31] Kolmogorov, A.N. Three approaches to the quantitative definition of information. *Probl. Inf. Transm.,* **1965**, *1*, 3.

[32] Chaitin, O. On the length of programs for computing finite binary sequence. *J. Assoc. Comput. Mach.,* **1966**, *13*, 547. [http://dx.doi.org/10.1145/321356.321363]

[33] Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, **1949**.

[34] Bennet, C.H. *Logical Depth and Physical Complexity*; Oxford University Press, **1988**.

[35] Lloyd, S.; Pagels, H. *Ann. Phys.,* **1988**, *188*, 186. [http://dx.doi.org/10.1016/0003-4916(88)90094-2]

[36] Vitanyi, P.M.; Li, M. Minimum description length induction, bayesianism, and Kolmogorov complexity. *IEEE Trans. Inf. Theory,* **2000**, *46*, 446. [http://dx.doi.org/10.1109/18.825807]

[37] Shalizi, C.R.; Shalizi, K.L.; Haslinger, R. Quantifying self-organization with optimal predictors. *Phys. Rev. Lett.,* **2004**, *93*(11), 118701. [http://dx.doi.org/10.1103/PhysRevLett.93.118701] [PMID: 15447385]

[38] Rosso, O.A.; Martin, M.T.; Plastino, A. Brain electrical activity analysis using wavelet-based informational tools (II): Tsallis non-extensivity and complexity measures. *Physica A,* **2003**, *320*, 497. [http://dx.doi.org/10.1016/S0378-4371(02)01529-7]

[39] Chatzisavvas, K.Ch.; Moustakidis, ChC.; Panos, C.P. Information entropy, information distances, and complexity in atoms. *J. Chem. Phys.,* **2005**, *123*(17), 174111. [http://dx.doi.org/10.1063/1.2121610] [PMID: 16375521]

[40] Borgoo, A.; De Proft, F.; Geerlings, P.; Sen, K.D. Complexity of Dirac-Fock atom increases with atomic number. *Chem. Phys. Lett.,* **2007**, *444*, 186-191. [http://dx.doi.org/10.1016/j.cplett.2007.07.003]

[41] López-Rosa, S.; Esquivel, R.O.; Angulo, J.C.; Antolín, J.; Dehesa, J.S.; Flores-Gallegos, N. Analysis of complexity measures and information planes of selected molecules in position and momentum spaces. *Phys. Chem. Chem. Phys.,* **2010**, *12*, 7108-7116.

[42] Molina-Espíritu, M.; Esquivel, R.O.; Dehesa, J.S. Information-Theoretical Complexity Analysis of Selected Elementary Chemical Reactions. In: *Without bounds: A scientific canvas of nonlinearity and complex dynamics*; Springer, **2013**; pp. 525-537.

[43] Sánchez-Moreno, P.; Rudnicki, L.; Toranzo, I.V.; Dehesa, J.S. Monotone measures of statistical complexity. *Phys. Lett. A,* **2016**, *380*, 377-380.

[44] López-Ruiz, R.; Mancini, H.L.; Calbet, X. A statistical measure of complexity. *Phys. Lett. A,* **1995**, *209*, 321-326.

[45] Angulo, J.C.; Antolín, J. Atomic complexity measures in position and momentum spaces. *J. Chem. Phys.,* **2008**, *128*(16), 164109. [http://dx.doi.org/10.1063/1.2907743] [PMID: 18447423]

[46] Esquivel, R.O.; Molina-Espíritu, M.; López-Rosa, S.; Soriano-Correa, C.; Barrientos-Salcedo, C.; Kohout, M.; Dehesa, J.S. Predominant information quality scheme for the essential amino acids: an information-theoretical analysis. *ChemPhysChem,* **2015**, *16*(12), 2571-2581. [http://dx.doi.org/10.1002/cphc.201500282] [PMID: 26175003]

[47] Esquivel, R.O.; López-Rosa, S.; Molina-Espíritu, M.; Angulo, J.C.; Dehesa, J.S. Information-theoretical space from simple atomic and molecular systems to biological and phamacological molecules. *Theor. Chem. Acc.,* **2016**, *135* [http://dx.doi.org/10.1007/s00214-016-2002-x]

[48] Angulo, J.C.; Antolín, J.; Esquivel, R.O. Atomic and molecular complexities: their physical and chemical interpretations. In: *Statistical Complexities: Applications in Electronic Structures*; Sen, K.D.,

Ed.; Springer: Berlin, **2010**; pp. 167-213.

[49] Dehesa, J.S.; López-Rosa, S.; Manzano, D. Entropy and complexity analyses of *d*-dimensional quantum systems. In: *Statistical Complexities: Applications in Electronic Structures*; Sen, K.D., Ed.; Springer: Berlin, **2010**; pp. 129-166.

[50] Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.,* **1964**, *136*, B864-B871. [http://dx.doi.org/10.1103/PhysRev.136.B864]

[51] Parr, R.G.; Yang, W. Density-Functional Theory of Atoms and Molecules. In: *Horizons of Quantum Chemistry*; Fukui, K.; Pullman, B., Eds.; Springer: Berlin, **1989**; 3, pp. 5-15.

[52] Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.,* **1948**, *27*, 379. [http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x]

[53] Carbó, R.; Lleyda, L.; Arnau, M. How similar is a molecule to another? An electron density measure of similarity between two molecular structures. *Int. J. Quantum Chem.,* **1980**, *17*, 1185-1189. [http://dx.doi.org/10.1002/qua.560170612]

[54] Onicescu, O. Theorie de l'information. Energie informationelle. *C.R. Acad. Sci. Paris A,* **1966**, *263*, 25.

[55] Romera, E.; Dehesa, J.S. The Fisher-Shannon information plane, an electron correlation tool. *J. Chem. Phys.,* **2004**, *120*(19), 8906-8912. [http://dx.doi.org/10.1063/1.1697374] [PMID: 15267826]

[56] Sen, K.D.; Antolín, J.; Angulo, J.C. Fisher-Shannon analysis of ionization processes and isoelectronic series. *Phys. Rev. A,* **2007**, *76*, 032502. [http://dx.doi.org/10.1103/PhysRevA.76.032502]

[57] Koga, T.; Kanayama, K.; Watanabe, S.; Thakkar, A.J. Analytical Hartree-Fock wave functions subject to cusp and asymptotic constraints: He to Xe, Li+ to Cs+, H- to I-. *Int. J. Quantum Chem.,* **1999**, *71*, 491. [http://dx.doi.org/10.1002/(SICI)1097-461X(1999)71:6<491::AID-QUA6>3.0.CO;2-T]

[58] Koga, T.; Kanayama, K.; Watanabe, S.; Imai, T.; Thakkar, A.J. Analytical Hartree-Fock wave functions for the atoms Cs to Lr. *Theor. Chem. Acc.,* **2000**, *104*, 411. [http://dx.doi.org/10.1007/s002140000150]

[59] Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.P.; Izmaylov, A.F.; Bloino, J.; Zheng, G.; Sonnenberg, J.L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J.A., Jr; Peralta, J.E.; Ogliaro, F.; Bearpark, M.; Heyd, J.J.; Brothers, E.; Kudin, K.N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.C.; Iyengar, S.S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J.M.; Klene, M.; Knox, J.E.; Cross, J.B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J.W.; Martin, R.L.; Morokuma, K.; Zakrzewski, V.G.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Dapprich, S.; Daniels, A.D. Gaussian 03 Revision D.01. *Gaussian Inc.Wallingford CT,* **2004**.

[60] Esquivel, R.O.; Molina-Espíritu, M.; Salas, F.; Soriano, C.; Barrientos, C.; Dehesa, J.S.; Dobado, J.A.

Decoding the building blocks of life from perspective of quantum information. In: *Advances in Quantum Mechanics*; Intech, **2013**.

[61] Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R. Gaussian 98 Revision A.06. *Gaussian Inc. Pittsburgh PA,* **1999**.

[62] Frisch, M.J.; Trucks, G.W.; Schlegel, H. B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.P.; Izmaylov, A.F.; Bloino, J.; Zheng, G.; Sonnenberg, J.L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J.A., Jr; Peralta, J.E.; Ogliaro, F.; Bearpark, M.; Heyd, J.J. Gaussian 09 Revision C.01. *Gaussian Inc. Wallingford CT,* **2009**.

[63] Soriano-Correa, C.; Esquivel, R.O.; Sagar, R.P. Physicochemical and structural properties of bacteriostatic sulfonamides: Theoretical study. *Int. J. Quantum Chem.,* **2003**, *94*, 165, 172. [http://dx.doi.org/10.1002/qua.10597]

[64] Soriano-Correa, C.; Raya, A.; Esquivel, R.O. Characterization of electronic structure and physicochemical properties of antiparasitic nifurtimox analogues: A theoretical study. *Int. J. Quantum Chem.,* **2008**, *108*(8), 1369-1379. [http://dx.doi.org/10.1002/qua.21633]

[65] Soriano-Correa, C.; Barrientos-Salcedo, C.; Raya, A.; Rubio-Póo, C.; Esquivel, R.O. The influence of electron donor and electron acceptor groups on the electronic structure of the anti-inflammatory tripeptide Cys-Asn-Ser. *Int. J. Quantum Chem.,* **2010**, *110*(13), 2398-2410.

[66] Soriano-Correa, C.; Sánchez-Ruíz, J.F.; Raya, A.; Esquivel, R.o. Electronic structure and physicochemical properties of selected penicillins. *Int. J. Quantum Chem.,* **2007**, *107*(3), 628-636. [http://dx.doi.org/10.1002/qua.21165]

[67] Pérez-Jordá, J.M.; Becke, A.D.; San-Fabián, E. Automatic numerical integration techniques for polyatomic molecules. *J. Chem. Phys.,* **1994**, *100*, 6520. [http://dx.doi.org/10.1063/1.467061]

[68] Kohut, M. program DGRID. *version 4.2,* **2007**.

[69] Lopez-Rosa, S.; Molina-Espiritu, M.; Esquivel, R.O.; Soriano-Correa, C.; Dehesa, J.S. Chemical space of selected bacteriostatic sulfonamides. *An information-theoretical point of view. Preprint,* **2016**.

[70] Perez-Trallero, E.; Iglesias, L. Tetracyclines, sulfonamides and metronidazole. *Enferm. Infecc. Microbiol. Clin.,* **2003**, *21*, 520-529. [PMID: 14572387]

[71] Wolf, M.E. *Burger's Medicinal Chemistry. Part II*; Wiley and Son: New York, **1979**.

[72] *Geometrical data along with the wave function of the eleven novel molecules are available on,*

[73] Kohout, M. program DGRID, version 4.6. *modified version from the author,* **2007**.

[74] Kishor, S.; Dhayal, S.; Mathur, M.; Ramaniah, L.M. Structural and energetic properties of α-amino acids: a first principles density functional study. *Mol. Phys.,* **2008**, *106*, 2289-2300.

[75] Close, D.M. Calculated vertical ionization energies of the common *α*-amino acids in the gas phase and in solution. *J. Phys. Chem. A,* **2011**, *115*(13), 2900-2912. [http://dx.doi.org/10.1021/jp200503z] [PMID: 21410277]

[76] Ramaniah, L.M.; Chakrabarti, A.; Kshirsagar, R.J.; Kamal, C.; Banerjee, A. Density functional study of α-amino acids: structural, energetic and vibrational properties. *Mol. Phys.,* **2011**, *109*, 875-892. [http://dx.doi.org/10.1080/00268976.2011.558027]

[77] Dehareng, D.; Dive, G. Vertical ionization energies of α -l-amino acids as a function of their conformation: an ab initio study. *Int. J. Mol. Sci.,* **2004**, *5*, 301-332. [http://dx.doi.org/10.3390/i5110301]

[78] Lee, K.T.; Sung, J.; Lee, K.J.; Kim, S.K.; Park, Y.D. Conformation-dependent ionization of l-phenylalanine: structures and energetics of cationic conformers. *Chem. Phys. Lett.,* **2003**, *368*, 262-268.

[79] Livingstone, C.D.; Barton, G.J. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Comput. Appl. Biosci.,* **1993**, *9*, 745-756. [PMID: 8143162]

[80] Taylor, W.R. The classification of amino acid conservation. *J. Theor. Biol.,* **1986**, *119*, 205-218. [http://dx.doi.org/10.1016/S0022-5193(86)80075-3] [PMID: 3461222]

[81] Kosiol, C.; Goldman, N.; Buttimore, N.H. A new criterion and method for amino acid classification. *J. Theor. Biol.,* **2004**, *228*, 97-106. [http://dx.doi.org/10.1016/j.jtbi.2003.12.010] [PMID: 15064085]

[82] Janak, J.F. Proff that $\partial E/\partial n_i = \varepsilon$ in density-functional theory. *Phys. Rev. B,* **1978**, *18*, 7165. [http://dx.doi.org/10.1103/PhysRevB.18.7165]

[83] Koopmans, T.A. ber die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen eines Atoms. *Physica,* **1933**, *1*, 104.

[84] Ayers, P.W.; Parr, R.G.; Pearson, R.G. Elucidating the hard/soft acid/base principle: a perspective based on half-reactions. *J. Chem. Phys.,* **2006**, *124*, 194107. [http://dx.doi.org/10.1063/1.2196882] [PMID: 16729803]

[85] Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual density functional theory. *Chem. Rev.,* **2003**, *103*, 1793-1873. [http://dx.doi.org/10.1021/cr990029p] [PMID: 12744694]

[86] Ghanty, T.K.; Ghosh, S.K. Correlation between hardness, polarizability, and size of atoms, molecules, and clusters. *J. Chem. Phys.,* **1993**, *97*, 4951. [http://dx.doi.org/10.1021/j100121a015]

[87] Chattaraj, P.K.; Sarkar, U.; Roy, D.R. Electrophilicity index. *Chem. Rev.,* **2006**, *106*(6), 2065-2091. [http://dx.doi.org/10.1021/cr040109f] [PMID: 16771443]

[88] Pearson, R.G. Hard and soft acids and bases. *J. Am. Chem. Soc.,* **1963**, *85*, 3533. [http://dx.doi.org/10.1021/ja00905a001]

[89] Parr, R.G.; Szentpály, L.V.; liu, S. Electrophilicity index. *J. Am. Chem. Soc.,* **1999**, *122*, 1922. [http://dx.doi.org/10.1021/ja983494x]

[90] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank: A computer-based archival file for macromolecular structures. *Nucleic Acids Res.,* **2000**, *28*, 235-242.

[http://dx.doi.org/10.1093/nar/28.1.235] [PMID: 10592235]

[91] Luecke, H.; Schobert, B.; Richter, H.T.; Cartailler, J.P.; Lanyi, J.K. Structure of bacteriorhodopsin at 1.55 A resolution. *J. Mol. Biol.,* **1999**, *291*, 899-911.
[http://dx.doi.org/10.1006/jmbi.1999.3027] [PMID: 10452895]

[92] Haupts, U.; Tittor, J.; Oesterhelt, D. Closing in on bacteriorhodopsin: progress in understanding the molecule. *Annu. Rev. Biophys. Biomol. Struct.,* **1999**, *28*, 367-399.
[http://dx.doi.org/10.1146/annurev.biophys.28.1.367] [PMID: 10410806]

[93] Marino, S.M.; Gladyshev, V.N. Analysis and functional prediction of reactive cysteine residues. *J. Biol. Chem.,* **2012**, *287*, 4419-4425.
[http://dx.doi.org/10.1074/jbc.R111.275578] [PMID: 22157013]

[94] Adhikari, U.; Scheiner, S. Preferred configurations of peptide-peptide interactions. *J. Phys. Chem. A,* **2013**, *117*, 489-496.
[http://dx.doi.org/10.1021/jp310942u] [PMID: 23273150]

[95] Soriano-Correa, C.; Olivares del Valle, F.J.; Muñoz-Losa, A.; Fdez Galván, I.; Martín, M.E.; Aguilar, M.A. Theoretical study of the competition between intramolecular hydrogen bonds and solvation in the Cys-Asn-Ser tripeptide. *J. Phys. Chem. B,* **2010**, *114*, 8961-8970.
[http://dx.doi.org/10.1021/jp1035162] [PMID: 20568808]

[96] Nelson, D.L.; Lehninger, A.L.; Cox, M.M. *Lehninger principles of biochemistry*; W. H. Freeman, **2000**.

[97] Science Education Partnership. *Molymod amino acid modeling kit,* **2012**.

LA SOCIEDAD MEXICANA DE PARASITOLOGÍA A.C.

OTORGA LA PRESENTE CONSTANCIA A:

Linda Verónica Campos Fernández

Por su participación como ponente en carteles en el XXI Congreso CONAPAR que tuvo lugar del 25 al 28 de septiembre del 2016 en el Conjunto Amoxcalli de la Facultad de Ciencias de la Universidad Nacional Autónoma de México, Ciudad de México

Dr. Luis Fernando Anaya Velázquez
Presidente de la Sociedad Mexicana de Parasitología

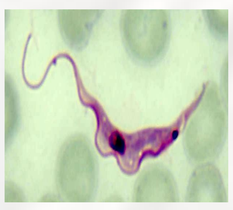
DISEÑO, SÍNTESIS Y EVALUACIÓN BIOLÓGICA DE ANÁLOGOS DE BENZNIDAZOL PARA EL TRATAMIENTO DE LA TRIPANOSOMIASIS AMERICANA

Campos-Fernández Linda Verónica.[1,2,3], Padilla-Martínez Itzia Irene.[2], Trujillo-Ferrara José Guadalupe.[1], Cuevas-Hernández Roberto I.[1], Soriano-Correa Catalina.[3]

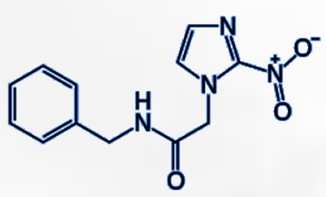
[1]Escuela Superior de Medicina-IPN.
[2]Unidad Profesional Interdisciplinaria de Biotecnología-IPN.
[3]Facultad de Estudios Superiores Zaragoza-UNAM.
linda_campos@msn.com

INTRODUCCIÓN

La Tripanosomiasis Americana es una enfermedad causada por el protozoario *Trypanosoma cruzi*. La Organización Mundial de la Salud (OMS), la cataloga como una de las 17 enfermedades tropicales olvidadas y con mayor incidencia a nivel mundial. Se calcula aproximadamente de 6 a 7 millones de personas infectadas, de las cuales, menos del 1% reciben tratamiento oportuno, causando cerca de 12000 muertes anuales [1-3].

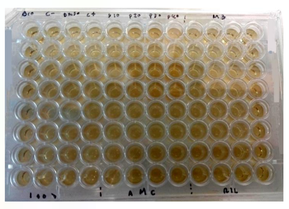


Los fármacos utilizados para el tratamiento de la tripanosomiasis son el Nifurtimox y Benznidazol; sin embargo, ambos poseen alta toxicidad, así como reacciones adversas y su efectividad disminuye hasta un 50% en la fase crónica de la enfermedad [4].



OBJETIVO

- Diseñar nuevas moléculas derivadas de benznidazol.
- Caracterizar su estructura electrónica y propiedades fisicoquímicas con la finalidad de seleccionar las moléculas con menor toxicidad teórica.
- Realizar la síntesis química de las moléculas propuestas.
- Estimar la sensibilidad de *T. cruzi* a las moléculas sintetizadas con respecto al benznidazol



METODOLOGÍA

Se realizaron cálculos de la estructura electrónica a un grupo de 50 moléculas análogas del benznidazol, utilizando Gaussian09 en fase acuosa.

→

Se seleccionaron las 8 moléculas con mejores propiedades fisicoquímicas

↓

Los descriptores químico-cuánticos locales y globales se determinaron a través de la Teoría de Funcionales de la Densidad (DFT), con el funcional híbrido M06-2X

←

Se realizó la síntesis química y caracterización fisicoquímica de las moléculas propuestas.

↓

La actividad *in vitro* de las moléculas sintetizadas fue determinada por medio lecturas de densidad óptica a 620 nm, durante 9 días

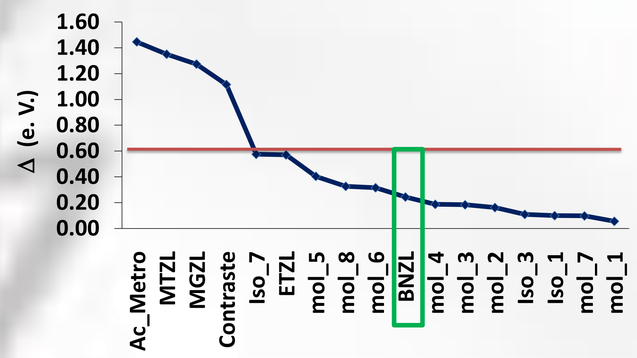
→

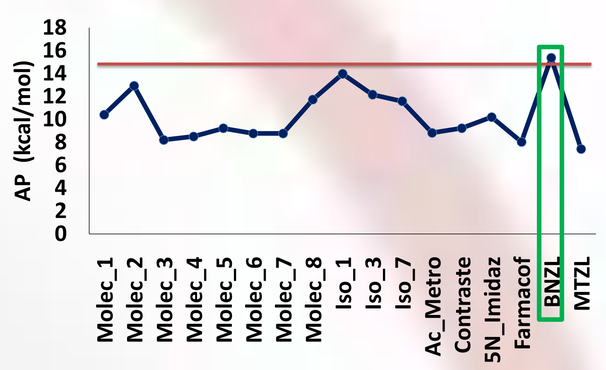
Se determinó la concentración inhibitoria 50 (CI_{50}) por medio de una curva dosis-respuesta

RESULTADOS Y DISCUSIÓN

*Teóricos

Los resultados de la estructura electrónica permitieron seleccionar a las moléculas 1, 2, 3 y 7 para ser sintetizadas, debido a que presentaron baja toxicidad teórica.

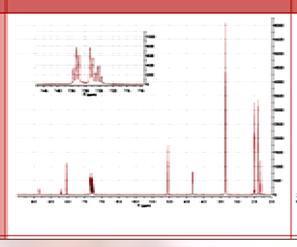
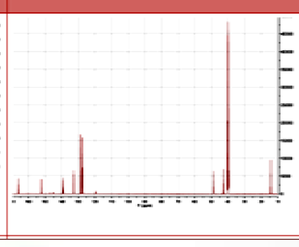
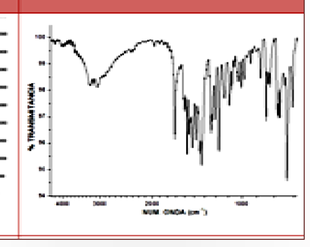




Molécula	*AP* kcal/mol	ω (e.V)	η (e.V)	Δ (e.V)	I (e.V)
BNZL	15.37	0.149	10.179	0.244	6.833
M3	8.21	0.166	10.088	0.184	6.876
Iso3	12.16	0.184	9.971	0.110	6.899
AcM	8.82	0.186	10.433	1.446	7.189

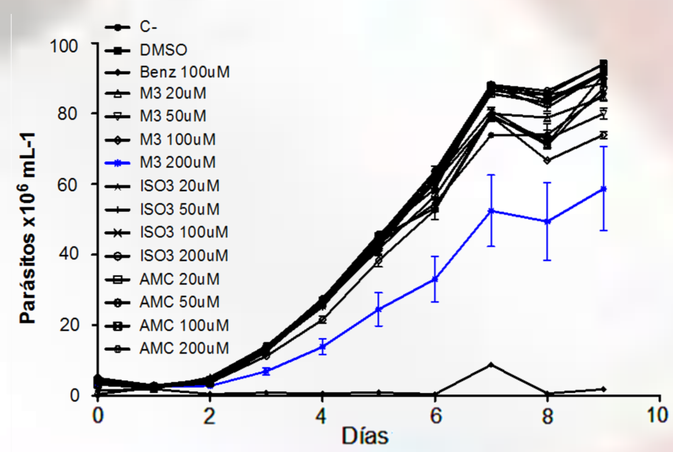
*Síntesis Química

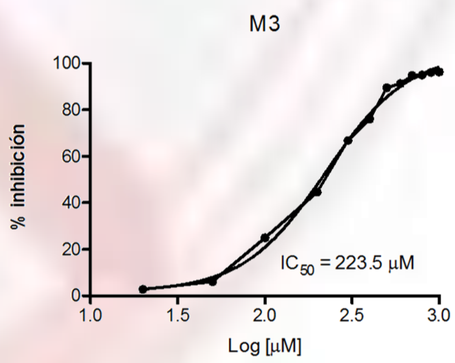
Por estrategia de síntesis, se seleccionó M3 como primer compuesto a estudiar y a ACM como molécula contraste debido a su alta toxicidad. Asimismo, se sintetizó el isómero de M3 (ISO3) con la finalidad de analizar la actividad en relación a la posición del grupo funcional.

Molécula	H^+	C^{13}	IR
M3			

*Evaluación *in vitro*

En los ensayos *in vitro* preliminares, se puede observar que la M3 presenta actividad ante el parásito. Contrariamente, las moléculas ISO3 y ACM no mostraron actividad.





PERSPECTIVAS

Estudiar la toxicidad *in vitro* e *in vivo* de M3 y realizar las modificaciones respectivas para potenciar su efecto contra *T. cruzi*

REFERENCIAS

[1] Organización Mundial de la Salud. La Enfermedad de Chagas (Tripanosomiasis Americana). Centro de Prensa de la OMS. Ginebra: OMS; 2016. [2] Mitch L. Science 2011, 333, 933. [3] Drugs for Neglected Diseases initiative. Enfermedad de Chagas. DNDi en América Latina. [4]García-Torres I, Pérez-Montfort R. REB 2011; 30(2): 68-81.

AGRADECIMIENTOS

Los autores agradecen a DGTIC-UNAM por la asignación de tiempo de cómputo en Equipo de Súper Cómputo "*Miztli*" y a DGAPA-UNAM PAPIIT-IN114715.

Universidad Veracruzana

Doctorado en CIENCIAS BIOMÉDICAS UNIVERSIDAD VERACRUZANA

Xalapa Equez., Veracruz, junio 22 de 2012.

Estimada Dra. Catalina Soriano Correa:

Por este medio nos complace informarle que el comité de admisión de nuestro posgrado, con base en los resultados de las evaluaciones especiales, dictaminó favorablemente la solicitud para ingresar al Doctorado en Ciencias Biomédicas de su estudiante **C. Hugo Arana Vidal** y formar parte de la generación 2012-2016.

Como es de su conocimiento, este programa está registrado en el PNPC del CONACYT en el nivel de reciente creación y para ascender a mejores niveles debemos de mantener una producción académica importante y cubrir satisfactoriamente los indicadores de eficiencia escolar de este consejo.

Por tal motivo, aprovechamos este medio para exhortarlo a trabajar estrechamente con su estudiante para cumplir las actividades académicas implementadas durante su trayectoria escolar. Esto beneficiará significativamente a su entidad de adscripción y al posgrado. Cabe señalar que el programa es de 8 semestres y uno de los requisitos de egreso es publicar un artículo científico derivado de su proyecto de investigación en una revista indizada en el Science Citation Index. Por lo que se les recomienda fijarse como meta, enviar el artículo en el último semestre; para no retrasar el proceso de obtención del grado académico de su estudiante y cumplir satisfactoriamente este indicador

Sin más por el momento le agradecemos su participación en el posgrado y le reiteramos nuestro apoyo esperando formar lazos sólidos de colaboración en beneficio de nuestra Universidad.

Atentamente
"Lis de Veracruz: Arte, Ciencia, Luz"

Dr. José Enrique Meza Alvarado
Coordinador del Doctorado en Ciencias Biomédicas
enmeza@uv.mx

UNIVERSIDAD VERACRUZANA
DOCTORADO
EN CIENCIAS BIOMEDICAS

Universidad Veracruzana
Doctorado en Investigaciones Cerebrales

A QUIEN CORRESPONDA:

Centro de Investigaciones Cerebrales.
Unidad de Ciencias de la Salud
Calle Médicos y Odontólogos s/n
Col. Unidad del Bosque
C.P.91010
Xalapa, Veracruz, México
www.uv.mx/dic

Teléfonos
+52(228) 841.8900
842.1700
842.2700
Ext. 16307

Correo:
cerebrales@uv.mx
luisgarcia@uv.mx

El que suscribe, Coordinador del Doctorado en Investigaciones Cerebrales de la Universidad Veracruzana

H A C E C O N S T A R

Que la **DRA. CATALINA SORIANO CORREA** -de la Facultad de Estudios Superiores Zaragoza de la Universidad Nacional Autónoma de México- funge como **Directora** del alumno Oskar Ibsan Valerio Hernández (matrícula S14025105), inscrito en el primer semestre del Doctorado, quien desarrolla el Proyecto de Investigación: *Análisis de la vía de NF-kB regulada por la acción del péptico anti inflamatorio CNS en ratas con glioma.*

A petición de la parte interesada y para los fines que a la misma convengan, se extiende la presente constancia en la ciudad de Xalapa-Enríquez, Veracruz., a los veinticuatro días del mes de febrero del año dos mil quince.

Atentamente
"Lis de Veracruz: Arte, Ciencia, Luz"

Dr. Luis Isauro García Hernández
Coordinador del Doctorado

Doctorado en Investigaciones Cerebrales

CONSTANCIA DE DIRECTOR DE TESIS

La Sección de Estudios de Posgrado e Investigación de la Escuela Superior de Medicina del Instituto Politécnico Nacional, hace constar que la **Dra. Catalina Soriano Correa,** fungió como Directora de la tesis titulada: "**Estudio teórico-experimental en el diseño de nuevos fármacos derivados de imidazoles con potencial actividad antichagásica"** de la alumna **Linda Verónica Campos Fernández (A150269),** cuyo examen fue celebrado el 25 de enero de 2017, adscrito al programa de **Maestría en Ciencias en Farmacología**.

Se extiende la presente para los fines que a la interesada convengan, en la Ciudad de México, a los 25 días del mes de enero de 2017.

A T E N T A M E N T E
"LA TÉCNICA AL SERVICIO DE LA PATRIA"

DR. GUILLERMO MANUEL CEBALLOS REYES
JEFE DE LA SECCIÓN

GMCR/MJGG/mjg*

CONSTANCIA DE DIRECTOR DE TESIS

La Sección de Estudios de Posgrado e Investigación de la Escuela Superior de Medicina del Instituto Politécnico Nacional, hace constar que la **Dra. Catalina Soriano Correa** funge como Directora de la tesis titulada: "**Diseño, síntesis, evaluación biológica y toxicológica de moléculas derivadas de imidazol con potencial antichagásico**" del alumno **Juan Andres Alvarado Salazar (B151269)**, adscrito al programa de Maestría en Ciencias en Farmacología.

Se extiende la presente para los fines legales que al interesado convengan en la Cuidad de México, a los 15 días del mes de febrero de 2016.

ATENTAMENTE
"LA TÉCNICA AL SERVICIO DE LA PATRIA"

SECRETARÍA DE EDUCACIÓN PÚBLICA
ESTADOS UNIDOS MEXICANOS
MÉXICO, D.F.

DR. GUILLERMO MANUEL CEBALLOS REYES
JEFE DE LA SECCIÓN DE ESTUDIOS
DE POSGRADO E INVESTIGACIÓN

ESCUELA SUPERIOR DE MEDICINA
I.P.N.
SECCION DE ESTUDIOS DE
POSGRADO E INVESTIGACION

GMCR/OPP/MJGG/mjg*

INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE MEDICINA
MAESTRÍA EN CIENCIAS EN FARMACOLOGÍA

ESTUDIO TEÓRICO-EXPERIMENTAL EN EL DISEÑO DE NUEVOS FÁRMACOS DERIVADOS DE IMIDAZOLES CON POTENCIAL ACTIVIDAD ANTICHAGÁSICA

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRA EN CIENCIAS EN FARMACOLOGÍA

PRESENTA:

QFB. LINDA VERÓNICA CAMPOS FERNÁNDEZ

Directoras de Tesis:

Dra. Itzia Irene Padilla Martínez

Dra. Catalina Soriano Correa

Enero, 2017.

INSTITUTO POLITÉCNICO NACIONAL

SIP-14-B S

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México siendo las 9:00 horas del dia 7 del mes de diciembre del 2016 se reunieron los miembros de la Comisión Revisora de Tesis designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de la Escuela Superior de Medicina para examinar la tesis titulada:

"Estudio teórico-experimental en el diseño de nuevos fármacos derivados de imidazoles con potencial actividad antichagásica"

Presentada por el alumno:

Campos	Fernández	Linda Verónica
Apellido paterno	Apellido materno	Nombre(s)

Con registro: A 1 5 0 2 6 9

aspirante de:

Maestria en Ciencias en Farmacologia

Después de intercambiar opiniones, los miembros de la Comisión manifestaron *APROBAR LA DEFENSA DE LA TESIS*, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Directores de tesis

Dra. Itzia Irene Padilla Martinez

Dra. Catalina Soriano Correa

Dra. Jessica Elena Mendieta Wejebe

Dra. Angélica Beatriz Raya Rangel

Dr. José Correa Basurto

PRESIDENTE DEL COLEGIO DE PROFESORES

Dr. Guillermo Manuel Ceballos Reyes

SECRETARÍA DE EDUCACIÓN PÚBLICA MÉXICO, D.F.

ESCUELA SUPERIOR DE MEDICINA
I.P.N.
SECCION DE ESTUDIOS DE POSGRADO E INVESTIGACION

Esta tesis se realizó en el Área de Química Computacional y Modelado Molecular de la Facultad de Estudios Zaragoza de la UNAM, Laboratorio de Química Orgánica, Supramolecular y Nanociencias de la Unidad Profesional Interdisciplinaria de Biotecnología del IPN y el Laboratorio de Bioquímica de la Escuela Superior de Medicina del IPN.

El proyecto fue realizado bajo la dirección de la Dra. Catalina Soriano Correa y la Dra. Itzia Irene Padilla Martínez; con las colaboraciones y asesorías del Dr. José Guadalupe Trujillo Ferrara y el Dr. Roberto Issac Cuevas Hernández.

Para el desarrollo de la maestría se contó con la beca otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACYT), con número de CVU 666127 y con la Beca de Estímulo Institucional de Formación de Investigadores (BEIFI) del IPN con ID 2697, comprendida del mes de febrero a junio del 2016.

Este trabajo de tesis contó con el financiamiento económico otorgado por el Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT-IN114715, DGAPA-UNAM) y por la Secretaría de Investigación y Posgrado (SIP/20160261) del Instituto Politécnico Nacional.

AGRADECIMIENTOS INSTITUCIONALES

A la Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) de la Universidad Nacional Autónoma de México (UNAM) por el tiempo de cómputo otorgado en la súpercomputadora "Miztli".

Al Dr. Roberto Issac Cuevas Hernández por la donación de la cepa CL-14 de *Trypanosoma cruzi* de la Escuela Superior de Medicina. Del mismo modo, por el apoyo, capacitación y disponibilidad en las pruebas de actividad tripanocida

A la Dra. Bertha J. Espinoza Gutiérrez por la donación de la cepa CL-Brener de *Trypanosoma cruzi* del Instituto de Investigaciones Biomédicas de la UNAM.

A la Dra. Araceli Contreras Rodríguez y al QBP Rodolfo Daniel Alday Reyes por su apoyo, paciencia y disponibilidad; así como, por permitirme el uso del equipo de UV Multiskan Go en el Laboratorio de Microbiología General de la Escuela Nacional de Ciencias Biológicas del IPN

ÍNDICE

UNIVERSIDAD VERACRUZANA
FACULTAD DE BIOANÁLISIS-CAMPUS VERACRUZ

Análisis *in sílico* de proteínas de superficie de *Trypanosoma cruzi*

TESIS

QUÍMICO CLÍNICO

PRESENTA

JOSCELYN HERNÁNDEZ SANTIAGO

Dra. Carolina Barrientos Salcedo
Director Interno

Dra. Catalina Soriano Correa
Director Externo

Veracruz, ver a 17 de Enero del 2017

INDICE

La tesis se llevo a cabo en el Laboratorio 6 de Química Médica y Quimiogenómica de la Facultad de Bioanálisis Región Veracruz. Se a gardece a las autoridades de la dependencia por las facilidades otorgadas.

Esta investigación fue realizada gracias al Programa UNAM-DGAPA-PAPIIT-IN114715 a traves del Proyecto cuyo Responsable es la Dra. Catalina Soriano Correa de la Facultad de Estudios Superiores, UNAM.

Resumen

Hernández Santiago Joscelyn. Lic. en Química Clínica. Universidad Veracruzana. Enero 2017. Análisis in sílico de proteínas de superficie de Trypanosoma cruzi. Directores:Interno: Dra. Carolina Barrientos Salcedo. Externo: Dra, Catalina Soriano Correa.

La enfermedad de Chagas es causada por el parásito *Trypanosoma cruzi* y se transmite al hombre a través de las heces liberadas por triatominos hematófagos. *Trypanosoma cruzi* es identificado serológicamente por fracciones proteicas; de esas los antígenos de secreción de tripomastigotes representan una alternativa para el diagnóstico de la enfermedad. De acuerdo a la OMS en el mundo hay entre 6 y 7 millones de personas infectadas por *Trypanosoma cruzi,* y solo existen dos fármacos contra la enfermedad, por lo que es importante encontrar agentes tripanocidas, tomando como blancos terapeuticos a las proteínas del parásito. En esta tesis, se eligieron algunas proteínas de la familia de las mucinas para su estudio *in silico*, ya que estas son las que se encuentran en mayor cantidad en el parasito. Las mucinas le dan protección al *Trypanosoma cruzi* contra el vector y los mecanismos del sistema inmune del hospedador, y le ayudan en la invasión al tejido y células, por lo que son un buen blanco farmacológico. Se identificaron y se bajaron de la base de datos TriTrypDB, las secuencias de todas las mucinas ahí representadas; se realizó el modelado estructural en el servidor I-Tasser de aquellas cuya estructura secundaria no se conoce experimentalmente y no se encuentra en los bancos de datos proteicos (PDB, Swiss-Prot). Y a partir del modelo se hizo un barrido conformacional y se eligió a la estructura con menor energía con el paquete computacional SwissPDBV. Para finalmente llevar a cabo el análisis de acoplamiento molecular o Docking de cada una de las 6 estructuras modelada, con los ligandos benznidazol y nifurtimox. Concluyendo que el grupo de las 6 proteínas de superficie pertenecientes a las mucinas estudiadas anteriormente, tienen características adecuadas como blancos para agentes tripanocidas.

Palabras claves: *Trypanosoma cruzi,* mucinas, bioinformática, modelado molecular, *in silico*