Prediction of partition coefficients for systems of micelles using DFT

Leila Saranjam

UNIVERSITAT DE BARCELONA

A thesis submitted by Leila Saranjam for the degree of Doctor in Philosophy in theoretical chemistry and computational modeling at the University of Barcelona

Prediction of partition coefficients for systems of micelles using DFT

Tutor: Dr.Sergio Madurga Díez

(Universitat de Barcelona)

Author: Leila Saranjam

(Universitat de Barcelona)

September 2023
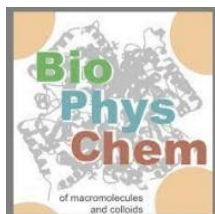
Abstract

A compound's solvent−water partition coefficient (log P) measures the equilibrium ratio of the compound's concentrations in a two-phase system: as two solvents in contact or a system of micelles in an aqueous solution.

In this thesis, the partition coefficient of three groups of small compounds (alcohol, ether, and hydrocarbons) in 10 different solvents (benzene, cyclohexane, hexane, n-Octane, toluene, carbon tetrachloride, heptane, trichloroethane, and octanol) was computed used DFT and B3LYP method with 6.31G(d), 6.311+G** and 6.311++G** basis sets. It is obtained that the partition coefficient of alcohol solutes in various solvents using the 6.31G(d) basis set indicates a satisfactory correlation with experimental values. The correlation between the experimental value and the partition coefficient of ether solutes in different solvents using the 6.311++G** basis set shows high agreement. The experimental data displayed a high correlation with the partition coefficient computed for hydrocarbon compounds in various solvents using all three basis sets: 6.31G(d), 6.311+G**, and 6.311++G**.

In addition, we have studied the correlation of the experimental partition coefficients in Sodium Dodecyl Sulfate (SDS), Hexadecyltrimethylammonium bromide (HTAB), Sodium cholate (SC), and Lithium perfluoro octane sulfonate (LPFOS) micelles with *ab initio* calculated partition coefficients in 15 different organic solvents. Specifically, the partition coefficients of a series of 63 molecules in an aqueous system of SDS, SC, HTAB, and LPFOS micelles are correlated with the partition coefficient in heptane/water, cyclohexane/water, n-dodecane/water, pyridine/water, acetic acid/water, octanol/water, acetone/water, 1-propanol/water, 2-propanol/water, methanol/water, formic acid/water, diethyl sulfide/water, decan-1-ol/water, 1-2 ethane diol/water and dimethyl sulfoxide/water systems. All calculations were performed using the Gaussian 16 Quantum Chemistry package. Molecular structures were generated in the more extended conformation using Avogadro, and geometries of all molecules were optimized using Density Functional Theory (DFT) B3LYP and MO6-2X with 6-31++G** basis set by the continuum solvation model based on density (SMD). The obtained results show that calculated partition coefficients in the alcohol/water mixture give the best correlation to predict the experimental partition coefficients in SDS, SC, and LPFOS micelles. With respect to HTAB micelle systems, a new selection of molecules is created, excluding those containing N atoms and Urea atom groups.

Interestingly, the partition coefficient of these chosen molecules exhibits a strong correlation with the experimental partition coefficient.

Finally, the partition coefficient of flexible molecules was studied by the same protocol for two solvent combinations, octanol/water and cyclohexane/water. The calculated values were compared with the experimental partition coefficients. The average partition coefficient in octanol solvent exhibited a high correlation with the experimental data. However, for the 16 compounds in the cyclohexane solvent, their partition coefficients do not exhibit significant agreement with the experimental partition coefficients.

Resumen

El coeficiente de partición disolvente-agua (logP) de un compuesto mide la relación de equilibrio de las concentraciones del compuesto en un sistema de dos fases: como dos disolventes en contacto o como un sistema de micelas en una solución acuosa.

En esta tesis, se calculó el coeficiente de partición de tres grupos de pequeños compuestos (alcohol, éter e hidrocarburos) en 10 disolventes diferentes (benceno, ciclohexano, hexano, n-octano, tolueno, tetracloruro de carbono, heptano, tricloroetano y octanol). Se utilizó el método DFT y B3LYP con conjuntos de bases 6.31G(d), 6.311+G** y 6.311++G**. Se obtiene que el coeficiente de partición de solutos alcohólicos en varios disolventes utilizando el conjunto de bases 6,31G(d) indica una correlación satisfactoria con los valores experimentales. La correlación entre el valor experimental y el coeficiente de partición de solutos de éter en diferentes disolventes utilizando el conjunto de bases 6.311++G** muestra una alta concordancia. Los datos experimentales mostraron una alta correlación con el coeficiente de partición calculado para compuestos de hidrocarburos en varios disolventes utilizando los tres conjuntos de bases: 6,31G(d), 6,311+G** y 6,311++G**.

Además, hemos estudiado la correlación de los coeficientes de partición experimentales en micelas de dodecilsulfato de sodio (SDS), bromuro de hexadeciltrimetilamonio (HTAB), colato de sodio (SC) y perfluorooctanosulfonato de litio (LPFOS) con coeficientes de partición calculados ab initio en 15 diferentes disolventes orgánicos. Específicamente, los coeficientes de partición de una serie de 63 moléculas en un sistema acuoso de micelas de SDS, SC, HTAB y LPFOS se correlacionan con el coeficiente de partición en heptano/agua, ciclohexano/agua, n-dodecano/agua, piridina/agua, ácido acético/agua, octanol/agua, acetona/agua, 1-propanol/agua, 2-propanol/agua, metanol/agua, ácido fórmico/agua, sulfuro de dietilo/agua, decan-1-ol/agua, 1-2 sistemas de etanodiol/agua y dimetilsulfóxido/agua. Todos los cálculos se realizaron utilizando el paquete Gaussian 16 Quantum Chemistry. Las estructuras moleculares se generaron en la conformación más extendida usando Avogadro, y las geometrías de todas las moléculas se optimizaron usando la Teoría Funcional de Densidad (DFT) B3LYP y MO6-2X con la base 6-31++G** establecida por el modelo de solvatación continua basado en la densidad. (SMD). Los resultados obtenidos muestran que los coeficientes de partición calculados en la mezcla alcohol/agua dan la mejor correlación para predecir los coeficientes de partición experimentales en micelas SDS, SC y

LPFOS. Con respecto a los sistemas micelares HTAB, se crea una nueva selección de moléculas, excluyendo aquellas que contienen átomos de N y grupos de átomos de urea. Curiosamente, el coeficiente de partición de estas moléculas elegidas muestra una fuerte correlación con el coeficiente de partición experimental.

Finalmente, se estudió el coeficiente de partición de moléculas flexibles mediante el mismo protocolo para dos combinaciones de disolventes, octanol/agua y ciclohexano/agua. Los valores calculados se compararon con los coeficientes de partición experimentales. El coeficiente de partición promedio en disolvente octanol mostró una alta correlación con los datos experimentales. Sin embargo, para los 16 compuestos en el disolvente ciclohexano, sus coeficientes de partición no muestran una concordancia significativa con los coeficientes de partición experimentales.

Resum

El coeficient de partició dissolvent-aigua (logP) d'un compost mesura la relació d'equilibri de les concentracions del compost en un sistema de dues fases: com dos dissolvents en contacte o un sistema de micel·les en una solució aquosa.

En aquesta tesi, es va calcular el coeficient de partició de tres grups de compostos petits (alcohol, èter i hidrocarburs) en 10 dissolvents diferents (benzè, ciclohexà, hexà, n-octà, toluè, tetraclorur de carboni, heptà, tricloroetano i octanol). va utilitzar el mètode DFT i B3LYP amb conjunts bàsics 6.31G(d), 6.311+G** i 6.311++G**. S'obté que el coeficient de partició dels soluts d'alcohol en diversos dissolvents utilitzant el conjunt de bases 6.31G(d) indica una correlació satisfactòria amb els valors experimentals. La correlació entre el valor experimental i el coeficient de partició dels soluts d'èter en diferents dissolvents utilitzant el conjunt de bases 6.311++G** mostra un alt acord. Les dades experimentals van mostrar una alta correlació amb el coeficient de partició calculat per als compostos d'hidrocarburs en diversos dissolvents utilitzant els tres conjunts de bases: 6.31G (d), 6.311 + G** i 6.311 ++ G**.

A més, hem estudiat la correlació dels coeficients de partició experimentals en micel·les de dodecil sulfat de sodi (SDS), bromur d'hexadeciltrimetilamoni (HTAB), colat de sodi (SC) i perfluorooctà sulfonat de liti (LPFOS) amb coeficients de partició calculats ab initio en 15 diferents dissolvents orgànics. Concretament, els coeficients de partició d'una sèrie de 63 molècules en un sistema aquós de micel·les SDS, SC, HTAB i LPFOS estan correlacionats amb el coeficient de partició en heptà/aigua, ciclohexà/aigua, n-dodecà/aigua, piridina/aigua, àcid acètic/aigua, octanol/aigua, acetona/aigua, 1-propanol/aigua, 2-propanol/aigua, metanol/aigua, àcid fòrmic/aigua, sulfur de dietil/aigua, decan-1-ol/aigua, 1-2 sistemes d'etan diol/aigua i dimetilsulfòxid/aigua. Tots els càlculs es van realitzar mitjançant el paquet Gaussian 16 Quantum Chemistry. Les estructures moleculars es van generar en la conformació més estesa mitjançant Avogadro, i les geometries de totes les molècules es van optimitzar mitjançant la Teoria Funcional de la Densitat (DFT) B3LYP i MO6-2X amb una base 6-31++G** establerta pel model de solvació contínua basada en la densitat. (SMD). Els resultats obtinguts mostren que els coeficients de partició calculats a la barreja alcohol/aigua donen la millor correlació per predir els coeficients de partició experimentals en micel·les SDS, SC i LPFOS. Pel que fa als sistemes de micelles HTAB, es crea una nova selecció de molècules, excloent les que contenen àtoms N i grups àtoms d'urea. Curiosament, el coeficient

de partició d'aquestes molècules escollides presenta una forta correlació amb el coeficient de partició experimental.

Finalment, el coeficient de partició de molècules flexibles es va estudiar pel mateix protocol per a dues combinacions de dissolvents, octanol/aigua i ciclohexà/aigua. Els valors calculats es van comparar amb els coeficients de partició experimentals. El coeficient de partició mitjà en dissolvent d'octanol va mostrar una alta correlació amb les dades experimentals. Tanmateix, per als 16 compostos del dissolvent de ciclohexà, els seus coeficients de partició no mostren un acord significatiu amb els coeficients de partició experimentals.

Acknowledgments

The first person I would like to thank sincerely is Dr. Sergio for all his help and guidance in chemistry. He has been my best teacher and guide; he has always shown me a clear way to learn with his kindness and support. And thanks to him, I started this academic course.

His patience and kindness know no bounds. He consistently responded to my inquiries with truthfulness and courtesy, aiding me in uncovering scientific truths. As I study in a foreign land away from home, his compassion demonstrated that success can be achieved collectively through shared humanity and determination, rather than being influenced by nationality or ethnicity. He guided me through the challenges, providing unwavering support to help me navigate this arduous journey confidently. His words consistently uplifted my spirits, and his valuable teachings continually benefited me. This illustrates his commitment to improving our world and promoting peace through his dedicated efforts and actions.

Dr. Francesc Mas has always been a caring father by my side. His reassuring gaze and advice made me feel like, I had a welcoming and supportive family here. I benefited from his guidance and attention.

I deeply appreciate Pablo for his uplifting vibes and radiant smiles. His warm demeanor consistently fills me with happiness, and I consider myself extremely fortunate to have crossed paths with such a compassionate and considerate individual as him right from the start of our educational journey.

I am filled with gratitude for Haruna, who is not only a colleague but also a friend that I've had the privilege of knowing and working with over the years. His compassionate nature and kindness have consistently brightened my days, and I am truly thankful for his positive impact on my life.

I always remember that my education period was the best and most beautiful period of my life and all of this was derived from the compassion, support, and teachings of these loved ones. I thank them for giving me the honor to benefit from their knowledge.

I want to express my heartfelt thanks to my fiancé, Milad, who constantly fills me with positivity, confidence, and unwavering support. Having him in my life is a wonderful blessing, and I truly appreciate that he has always been there for me, providing comfort during tough times and even

when we are apart. Milad reminds me that I'm not alone and that he will always be there to support and be with me. This reassurance is one of the most precious gifts in my life, and I am truly grateful for his endless love.

I wish to express my heartfelt gratitude to my parents for their steadfast backing, limitless affection, and significant selflessness. My father has been a constant and unwavering presence in my journey. His unwavering support and constant encouragement have spurred me to chase my ambitions and advance. I am deeply thankful for his unyielding commitment and steadfastness; I value every single thing he does for me.

My mom has shown me what it means to give and be kind. I deeply respect her for her dedication and for prioritizing others ahead of her care and love have been my constant source of encouragement and positivity. I'm genuinely thankful for all that she has accomplished for me. Also, I am grateful to my dear sister and grandmother who always supported me and helped me with their kindness. Thanks to her kindness and companionship, I always felt that I had a supportive family behind me, which greatly motivated my progress.

The love and support of my family have always been the best gift of my life and they have taught me great lessons about humanity and love. Without them, I could not have reached here and enjoyed these days full of hope and light.

As a final note, I want to express my sincere appreciation to Mahyar, Farah, Mohammad, Roya, Zahra, Mohsen, Fatemeh, Masoomeh, Tahereh, and Davood for their unwavering and consistent support, along with their boundless love and affection.

I dedicate this thesis to my beloved grandmother 'Iran Dokht' and my dear professor 'Sergio Madurga Díez' with deep devotion and gratitude. Your constant support, guidance, and love have truly inspired me. This humble work is my way of showing appreciation for all your love and efforts.

The world, all of it is nothing, and the people of the world, all of them are nothing, oh nothing, do not entangle yourself with nothingness.

Do you know what remains after death? Love remains, affection remains, and everything else is nothing.

El mundo, todo es nada, y la gente del mundo, todos son nada,

Oh nada, no te enredes con la nada.

¿Sabes qué queda después de la muerte?

El amor queda, y la afecto queda, y todo lo demás es nada.

Table of Contents

XIII

Abbreviations

p	Partition coefficient
σ	Sigma bond
π	Pi bond
2D	Two dimensions
3D	Three dimensions
ps	Picoseconds
K	Kelvin
DFT	Density functional theory
SMD	Solvation model based on density
atm	Atmosphere
R²	R-squared
MAE	Mean absolute error
RMSE	Root mean squared error
ASE	Average signed error
Log P$_{Experimental}$	Experimental partition coefficient
Log P$_{Calculation}$	Calculated partition coefficient
O/W	Octanol/Water
C/W	Cyclohexane/Water
CMC	Critical micelle concentration
ADME	Absorption, Distribution, Metabolism, and Excretion
Solv/Wat	Solvent/water
T	Temperature
MEKC	Micellar electrokinetic chromatography
LFERs	Linear free energy relationships
V	Volume
CT	Total concentration
EFF	Empirical force field
HK	Hohenberg-Kohn's
KS	Kohen's theorem
LSDA	Local-Spin Density approximation
GGA	Generalized gradient Approximation
STO	Slater-type orbital
SCRF	Self-consistent reaction field
COSMO	Conductor-like screening model
PVC	Polyvinyl chloride
PET	Polyethylene terephthalate
MDR	Multidrug resistance
HPLC	High-performance liquid chromatography
TLC	Thin-layer chromatography
MD	Molecular dynamics

CHAPTER

I

INTRODUCTION

Chapter
Introduction

1.1 Micelles
1.1.1 Properties of micelle systems

The word "Micelle" appeared in the scientific literature in the early 19th century, meaning a small particle. "Monomers" are individual surfactant molecules present in the solution but do not form a micelle. Micelles are a type of molecular assembly in which each component is in thermodynamic equilibrium with all other monomers of the same species present in the environment. In aqueous media, molecules with both hydrophilic and hydrophobic structures can associate to create dynamic aggregates known as micelles. Since Hartley's initial research was published, interest in these aggregates has increased over time. Hartley described the micelle geometric model, which is still in use today [1,2]. Micelles have different uses according to their charge (anionic, cationic, and non-ionic). The results can lead to improvements that weren't initially anticipated and can be both unexpected and illuminating. Amphiphilic molecular aggregates known as micelles are dynamic.

An amphiphilic molecule has both hydrophobic and hydrophilic areas. "Micelle" refers to a surfactant system that has been dissolved in an aqueous media. The polar structure of the amphiphilic is known as the (hydrophilic) head group, whereas the nonpolar component of the molecule is often referred to as the (hydrophobic) tail known (as normal micelles). As seen in Figure 1, these amphiphilic aggregations assemble with the polar head groups forming a border zone between the nonpolar core of the micelle and the isotropic (polar) aqueous solution beyond. The tails of the molecules are packed together in the interior, or core, of the micelle. The Stern layer of ionic micelles refers to the charged interfacial region. Nonionic micelles present polar structures like polyoxyethylene groups to the bulk solution instead of having charged head groups. Sheath refers to the polyoxyethylene head groups of the nonionic micelles.
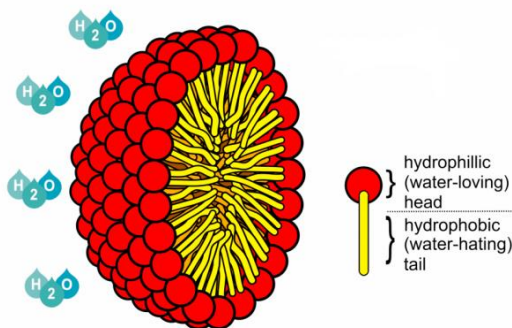
Figure 1. The aggregation of N monomers forms a normal, aqueous micelle. The open circles represent polar head pups and may be anionic, cationic, or nonionic.

One line of thinking believes that the bulk aqueous system's micelle is a hydrophobic phase droplet [3]. These findings are usually the result of NMR studies of line widths and relaxation durations. This concept has been called the "Reef model" of micelle structure because it prevents deep water penetration into the micelle interior [4]. The other widely accepted theory suggests that the micelle provides fairly deep routes into the core of the aggregation. This theory is supported by molecular models and fluorescence measurements, this model is known as the Fiord model [3]. The type of amphiphilic and the solution's composition both affect the micelle's properties.

According to the nature of the head group, micelles are divided into three categories: anionic, cationic, and non-ionic.

Samples of micelle having an anionic head group are the carboxylic acid, sulphuric acid, and phosphoric acid salts found in alkali and alkaline earth metals.

Cationic micelles have quaternary nitrogen head groups which are stable. Nonionic micelles often include polyoxyethylene or polyoxypropylene chains as their polar head groups [5]. The molecular geometry of a micelle's surfactant molecules, as well as factors like surfactant concentration, temperature, pH, and ionic strength, influence the micelle's form, size, and properties [6,7,3].

Aqueous-micellar solutions may therefore solubilize both polar and nonpolar compounds because of the existing polarity gradient [8]. The form of micelles is about spherical. Other phases are also possible, including ellipsoids, cylinders, and bilayers. According to their polymorphism, many lipids display the phase behavior known as micellization, which is the process of producing micelles [9]. Ionic surfactant-based micelles are electrostatically attracted to the countering that is

present in the solution around them. Even though the nearest counterions partially surround a charged micelle, the micelle charge has an influence on the solvent's structure up to a large distance from the micelle. One of the characteristics of ionic micelles is electrical conductivity. Salts can decrease electrostatic interactions in colloid-containing micelles and encourage the formation of larger ionic micelles [10]. Micelles can only form when the system temperature is above the critical micelle temperature, also known as the Krafft temperature, and the surfactant concentration exceeds the critical micelle concentration (CMC) [11]. Thermodynamics may be used to explain the creation of micelles: Micelles can form spontaneously when entropy and enthalpy are balanced. Even though assembling surfactant molecules is unfavorable for the system's enthalpy and entropy, the hydrophobic effect is what causes micelle production in water. Only monomers are found in the solution when the surfactant concentration is very minimal [12]. As the surfactant concentrations increase, there arrives a point where the negative entropy contribution from clustering the molecules' hydrophobic tails is outweighed by an increase in entropy from the release of the solvation shells around the surfactant tails. At this step, a portion of the surfactants' lipid tails must be separated from the water. Micelles begin to develop as a result. In general, above the CMC, the entropy gain from releasing the water molecules that were "trapped" in the solvation shells of the surfactant monomers is greater than the entropy increase associated with the assembly of the surfactant molecules [13-15]. Enthalpy factors, such as the electrostatic interactions between the charged components of surfactants, are also significant. Surfactants can act as emulsifiers, causing a normally insoluble molecule (in the solvent being employed) to dissolve, when they are present above the critical micelle concentration (CMC). This is possible due to the incorporation of the insoluble species into the micelle core, which is then able to solubilize in the bulk solvent because of the favorable interactions of the head groups with solvent species.

Detergents are the most common example of this phenomenon, which helps to dissolve substances that are not soluble in water. Also, the detergent makes it easier to remove water from the surfaces by reducing the water level. The emulsifying property of surfactants is also the basis of emulsion polymerization.

Micelles could have significant roles in chemical reactions. In some cases, multi-step chemical synthesis may be more feasible because of micellar chemistry, which employs the inside of micelles to harbor chemical reactions [16,17]. By doing this, one may increase reaction yield, make

environments that are better suited for particular reaction products (such as hydrophobic compounds), and use fewer solvents, side products, and conditions overall (e.g., extreme pH). Micellar chemistry is therefore categorized as a type of Green chemistry because of these advantages [18]. However, the production of micelles can also prevent chemical reactions from occurring, as when reacting molecules generate micelles that protect a molecular component that is susceptible to oxidation.

Micelle production is necessary for the body to absorb complex fats and fat-soluble vitamins. Bile acids and lecithin assist in the solubilization of cholesterol by generating mixed micelles. When the critical concentration of micellar form as normal micelles nation (CMC) is reached, bile salts act as detergents and produce micelles. The hydrophobicity and kind of bile salts have an impact on CMC. The concentration and makeup of bile acids determine the proportion of bile-mixed micelles and vesicles. The vesicles are a stable structure that is important for the solubilization and transportation of cholesterol into the bile together with micelles [19]. This allows the small intestine to absorb complex lipids such as lecithin and lipid-soluble vitamins like A, D, E, and K from the micelle. Micelles can also be used to deliver drugs to specific regions [20].

1.1.2 Partition coefficients

The partition coefficient relates to the molecule's physicochemical characteristics.

The partition coefficient (P) is the ratio of the concentrations of a compound in a mixture of two immiscible solvents at equilibrium [21]. A comparison of the solute solubility in these two liquids is consequently represented by this ratio. Partition coefficients refer to the concentration ratios of all species of a compound (ionized and non-ionized). However, Distribution coefficients usually refer to the concentration ratios of a compound's non-ionized species (ionized) [22].

Thus, the partition coefficient evaluates a chemical substance's hydrophilic (love of water) or hydrophobia (fear of water) [23,24].

The partition coefficient plays a significant role in drug delivery [25]. Hydrophobic molecules with a high partition coefficient exhibit the transfer of the molecule to the non-aqueous region, and hydrophilic molecules with a low partition coefficient indicate the transfer of the molecule to the aqueous parts.

The partition coefficient is usually related to the quantitative determination of the relationship between structure and activity (biological) as well as other physical properties such as electronic and steric factors.

Partition coefficients have many uses in various fields, such as:

- **Pharmacology**

The ability of a drug to reach its intended target in the body is significantly influenced by its partition coefficient. Hydrophobic drugs with high partition coefficients are transported in preferentially hydrophobic regions such as cellular lipid bilayers. In contrast, watery environments like blood serum are where you'll typically find hydrophilic drugs (low partition coefficients) [26,27].

- **Pharmacokinetics**

In this case, the ADME characteristics are strongly influenced by the partition coefficient. Therefore, the hydrophobicity of the molecule is an important factor. Additionally, drugs must cross the lipid bilayer of the intestinal epithelium to be orally absorbed. For effective transport, it must be suitably hydrophobic to partition into a lipid bilayer. However, it shouldn't be too much hydrophobic. When drugs are absorbed into the body, hydrophobicity plays an important role in determining how they are dispersed throughout the body.

- **Pharmacodynamics**

Drugs that have a hydrophobic effect attach to their target receptors, but long-term storage makes hydrophobic drugs extremely harmful [28,29].

- **Agrochemical analysis**

Hydrophobic herbicides and insecticides are usually more effective. Hydrophobic pesticides often have long half-lives, making them more likely to have negative environmental impacts [30].

- **Environmental science**

The hydrophobicity of a compound can indicate how easily a compound can absorb groundwater. The movement of dissolved hydrophobic organic compounds in soil and groundwater is predicted and modelled in the science of hydrogeology using the partition coefficient.

- **Product generation of consumers**

Properties of consumer goods including hair colors, cosmetics, topical ointments, color, makeup formulations, and many more products are affected by their partition coefficient [31].

There are several experimental methods for calculating partition coefficients such as Electrochemical [32-34] , HPLC [35], Chromatography [36-38], and Shake flask [39,40] or calculated using a variety of techniques (fragment-based [41], atom-based [42,43], etc.).

In fact, the experimental log P of compounds estimation is expensive, time-consuming, and difficult.

The computational estimation of the partition coefficient (log P) by computer simulation using free energy was promoted for this purpose because of the importance of the partition coefficient in several research fields, including drug delivery. By using computer simulations in which one compound is changed into the other while the simulation is running, these calculations enable the partition coefficient's (log P) difference to be calculated from the first principles. The molecular dynamics simulations, in both the aqueous and organic phases, are based on a classical mechanical force field and explicitly treat each atom.

A compound solute dispersed between two immiscible solvents seems to have a partition coefficient that is defined as [44,45]:
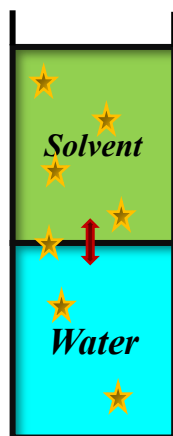
Figure 2. Schematic diagram of the analyte partition between the water and solvent layers upon equilibration.

Where solutes refer to neutral solutes in both solvents. Usually, the ratio of solute concentrations is shown using the logarithm P (log P).
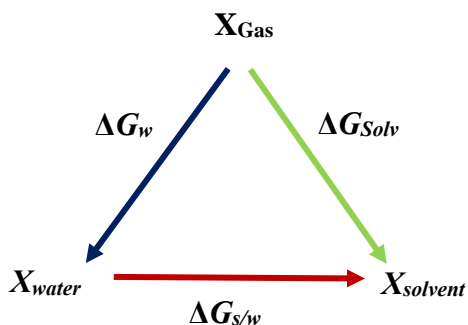
Figure 3. Thermodynamic cycle used to determine the transfer free energy of a compound (X) between two immiscible solvents.

The Gibbs free energy of transfer ($\Delta G_{solv/wat}$) between water and a certain solvent is related to the partition coefficient of a molecule. This free energy difference is calculated using absolute solvation-free energies in both media by

$$\Delta G^{\circ}_{solv/wat} = \Delta G^{\circ}_{solvent} - \Delta G^{\circ}_{water} \tag{1.1}$$

$$logP = \frac{-\Delta G^{\circ}_{solv/wat}}{RTln10} \tag{1.2}$$

where R is the molar gas constant and T is the temperature (298.15 K). It should be noted that a similar solvation energy of right-hand terms of Equation (1.1) will correspond to a null Gibbs free energy of transfer, meaning a similar distribution of the molecule between both solvents [46].

1.1.3 Experimental method for micelles

A common separation technique based on electrophoretic and electroosmotic principles is micellar electrokinetic chromatography (MEKC) whose distinctive characteristic is the addition of a surfactant over its critical micelle concentration to the separation buffer. Thus, solutes can be separated from complicated mixtures of neutral and ionized solutes not only by migration but also by distribution between the aqueous phase (bulk electrolyte) and the pseudo-stationary phase (charged micelles). The distribution constants between the aqueous and micellar phases are used to separate the uncharged solutes, while a distribution between phases and electrophoretic mobility are employed to separate the charged solutes [1,2,47]. The main benefit of MEKC is the ability to change the migratory behavior and separation in a very simple and flexible way by simply changing the characteristics of the pseudo-stationary phase. The solution can be separated by using the suitable surfactant type, as well as by adding complexing agents or organic solvents [47-52]. Therefore, it is widely acknowledged that the choice of surfactant is the key factor for varying the medium's chemical components and optimizing selectivity [52,53]. Since it enables a better understanding of the types and relative strengths of chemical interactions, the solubility parameter model is advised for characterizing MEKC systems [54]. In addition to describing how distinct the defined systems are, Abraham's model [55] also measures the strength of various interactions between the phases (in the case of MEKC, the aqueous phase, and the micellar phase) and neutral solutes. Based on solute descriptors, linear free energy relationships (LFERs) have been defined.

The solvation parameter model based on Abraham's solute descriptors of excess molar refraction (E), solute dipolarity/polarizability (S), effective hydrogen-bond acidity and hydrogen-bond basicity (A and B, respectively), and the McGowan's characteristic volume (V) is one method for characterizing the micellar systems and predicting retention [54,56,36].

$$logK = c + eE + sS + aA + bB + vV \tag{1.3}$$

where the system constants c, e, s, a, b, and v are related to the phase ratio (c), the capacity of the buffer and micelles to interact with solute (e), the difference in dipolarity/polarizability between

the two phases (*s*), the difference in hydrogen-bond basicity and acidity between the buffer and micelles (*a* and *b*, respectively), and the relative ease of forming a cavity (*v*). By using multiple linear regression to analyze the experimental *log K* values obtained for a variety of different solutes with known descriptor values, the coefficients of the correlation equation for each MEKC system may be determined [53,54,37]. The literature offers several suggestions since it is critical to choose a significant set of solutes to properly calculate the coefficients of equation (1.3). For the chosen group of compounds, well-characterized descriptor values with the least amount of uncertainty are also necessary to get a decent fit of the solvation parameter model. Many different MEKC surfactants have been evaluated using the solvation parameter model because of the wide range of surfactants that are commercially accessible [47,48,52,57]. A similar model has also been used to describe micellar systems. The reason why mixed micelles have received so much interest is that the properties of the pseudo stationary phase, and consequently the coefficients of equation (1.3) for the MEKC system, may be continually changed by adjusting the number of surfactants.

This is particularly intriguing because, as has been found in certain research to categorize the chemical selectivity of electrokinetic chromatography systems, it can make selectivity fine-tuned more easily. In this work, we characterize micellar systems of anionic micelles sodium dodecyl sulfate (SDS), sodium cholate (SC), lithium perfluoro octane sulfonate (LPFOS), and cationic micelles hexadecyltrimethylammonium bromide (HTAB).

In MEKC, neutral molecules are separated according to their distribution between the aqueous phase and the micellar phase. The retention factor, *k*, of a compound, can be calculated according to the following equation:

$$k = \frac{t_R - t_0}{\left(1 - \left(\frac{t_R}{t_m}\right)\right)t_0} \tag{1.4}$$

where t_R is the retention time of the compound of interest, and t_0 and t_m the retention times of the electro-osmotic flow and micellar markers (methanol and phenyl-undecylketone, respectively). The partition coefficients were obtained from:

$$k = P\frac{v(C_T - CMC)}{1 - v(C_T - CMC)} \tag{1.5}$$

Where P is the partition coefficient of a compound distributed between the micellar and the aqueous phase, *v* is the partial molar volume of the surfactant, C_T is the total concentration of

surfactant, CMC is the critical micellar concentration and k is the MEKC retention factor of the test compound. CMC is obtained from the conductimetric analysis. The partition coefficients P are determined on 40mM micelles solution, in 20mM phosphate buffer at pH 7 at $25^{\circ}C$.

1.2 References

[1] Terabe, S., & Otsuka, K. (1994). Review: Separation techniques of enantiomers by capillary electrophoretic. Analytical Chemistry, 666, 295–319. doi:10.1016/0021-9673(94)80392-7.

[2] Buszewski, B., K, E., Dahm, H., Ró, H., Szeliga, J., & Jackowski, M. (2007). Rapid identification of Helicobacter pylori by capillary electrophoresis: An overview. Journal of Chromatography B, 122(1), 116–122. doi:10.1002/bmc.

[3] Taylor, P., Mclntire, G. L., Dorsey, J. G., & Mclntire, G. L. (2006). Critical Reviews in Analytical Chemistry: Micelles in Analytical Chemistry Micelles in Analytical Chemistry, 37–41. doi:10.1016/S0223-5234(02)01384-3.

[4] Menger, F. M., Jerkunica, J. M., & Johnston, J. C. (1978). The water content of a micelle interior. The Fjord vs. Reef models. Journal of the American Chemical Society, 100(15), 4676-4678. doi:10.1021/ja00483a008.

[5] Helenius, A. R. I., & Simons, K. A. I. (1975). Solubilization of membrane by detergents. October, 29-79. doi: 10.1016/0304-4157(75)90016-7.

[6] Armstrong, D. W. (2006). Separation & Purification Reviews: Micelles in Separations: Practical and Theoretical Review. November, 37–41. doi:10.1080/03602548508068421.

[7] Sepi, L., & August, R. (1984). Counterion Association in Mixed Micelles of Cationic and Nonionic Detergents. Journal of Colloid and Interface Science, 99(2), 536–542. doi:10.1016/0021-9797(84)90140-1.

[8] Schwarze, M., Volovych, I., Wille, S., Mokrushina, L., Arlt, W., & Schomäcker, R. (2012). Partition coefficients of itaconates in aqueous-micellar solutions: Measurements and predictions with COSMO-RS. Industrial & Engineering Chemistry Research, 51(4), 1846–1852. doi:10.1021/ie2006565.

[9] Ruths, M. (2009). A Review of: 'Introduction to Soft Matter: Synthetic and Biological Self-Assembling Materials, by I. W. Hamley'. Molecular Crystals and Liquid Crystals, 506(1), 171–173. doi:10.1080/15421400902841346.

[10] Turro, N. J., & Yekta, A. (1978). Luminescent Probes for Detergent Solutions. A Simple Procedure for Determination of the Mean Aggregation Number of Micelles. Journal of the American Chemical Society, 100(18), 5951–5952. doi:10.1021/ja00486a062.

[11] Dave, N., & Joshi, T. (2017). A Concise Review on Surfactants and Its Significance. International Journal of Applied Chemistry, 13(3), 663–672. doi:10.37622/ijac/13.3.2017.663-672.

[12] Demissie, H., & Duraisamy, R. (2016). Effects of electrolytes on the surface and micellar characteristics of Sodium dodecyl sulphate surfactant solution. Journal of Scientific and Innovative Research, 5(6), 208–214. doi:10.31254/jsir.2016.5603.

[13] Ball, P. (2008). Water as an active constituent in cell biology. Chemical Reviews, 108(1), 74–108. doi:10.1021/cr068037a.

[14] Gelbart, W. M., & Ben-Shaul, A. (1996). The 'new' science of 'Complex fluids.' Journal of Physical Chemistry, 100(31), 13169–13189. doi:10.1021/jp9606570.

[15] Khan, N., & Brettmann, B. (2019). Intermolecular interactions in polyelectrolyte and surfactant complexes in solution. Polymers (Basel), 11(1). doi:10.3390/polym11010051.

[16] Paprocki, D., Madej, A., Koszelewski, D., Brodzka, A., & Ostaszewski, R. (2018). Multicomponent reactions accelerated by aqueous micelles. Frontiers in Chemistry, 6(OCT), 1–21. doi:10.3389/fchem.2018.00502.

[17] Navarro, O., Marion, N., Mei, J., & Nolan, S. P. (2006). Rapid room temperature Buchwald-Hartwig and Suzuki-Miyaura couplings of heteroaromatic compounds employing low catalyst loadings. Chemistry - A European Journal, 12(19), 5142–5148. doi:10.1002/chem.200600283.

[18] Macquarrie, D. J. (2009). l. Topics in Catalysis, 52(12), 1640–1650. doi:10.1007/s11244-009-9301-6.

[19] Reshetnyak, V. I. (2013). Physiological and molecular biochemical mechanisms of bile formation. World Journal of Gastroenterology, 19(42), 7341–7360. doi:10.3748/wjg.v19.i42.7341.

[20] Chen, X., et al. (2008). Core-shell-corona Au-micelle composites with a tunable smart hybrid shell. Langmuir, 24(15), 8198–8204. doi:10.1021/la800244g.

[21] Bannan, C. C., Calabró, G., Kyu, D. Y., & Mobley, D. L. (2016). Calculating partition coefficients of small molecules in octanol/water and cyclohexane/water. Journal of Chemical Theory and Computation, 12(8), 4015–4024. doi: 10.1021/acs.jctc.6b00449.

[22] Ingram, T., Mehling, T., & Smirnova, I. (2013). Partition coefficients of ionizable solutes in aqueous micellar two-phase systems. Chemical Engineering Journal, 218, 204–213. doi: 10.1016/j.cej.2012.12.047.

[23] Rivett, M., Drewes, J., Barrett, M., Chilton, J., Appleyard, S., Dieter, H. H., Wauchope, D., & Fastner, J. (2006). Chemicals: Health relevance, transport and attenuation. In J. Chilton, M. A. B. A. H. Bartram, & G. Howard (Eds.), Protecting groundwater for health: managing the quality of drinking-water sources. IWA Publishing.

[24] Koehler, M. G., Grigoras, S., & Dunn, W. J. (1988). The Relationship Between Chemical Structure and the Logarithm of the Partition Coefficient. Quantitative Structure-Activity Relationships, 7(3), 150–159. doi:10.1002/qsar.19880070306.

[25] Landry, M. L., & Crawford, J. J. (2020). Log D Contributions of Substituents Commonly Used in Medicinal Chemistry. ACS Medicinal Chemistry Letters, 11(1), 72–76. doi:10.1021/acsmedchemlett.9b00489.

[26] Edwards, M. P., & Price, D. A. (n.d.). (2010) Chapter 23 - Role of Physicochemical Properties and Ligand Lipophilicity Efficiency in Addressing Drug Safety Risks. In Annual Reports in Medicinal Chemistry - Volume 45 (Vol. 45). Elsevier Inc. doi:10.1016/S0065-7743(10)45023-X.

[27] Refaay, D. A., Abdel-hamid, M. I., Alyamani, A. A., Mougib, M. A., Ahmed, D. M., Negm, A., Mowafy, A. M., Ibrahim, A. A., & Mahmoud, R. M. (2022). Growth Optimization and Secondary Metabolites Evaluation of Anabaena variabilis for Acetylcholinesterase Inhibition Activity.11(6), 1–16. doi:10.3390/plants11060735.

[28] Pusey, P. N., & van Megen, W. (1986). Phase behaviour of concentrated suspensions of nearly colloidal spheres. Nature, 320, 340. doi: 1038/320340a0.

[29] Miyamoto, S. (1993). What determines the strength of noncovalent association of ligands to proteins in aqueous solution? Proceedings of the National Academy of Sciences, 90(September), 8402–8406. doi:10.1073/pnas.90.18.8402.

[30] Noble, A. (1993). Review Partition coefficients (n-octanol-water). Journal of Chromatography A, 642, 3–14. doi:10.1016/0021-9673(93)80072-G.

[31] Hansch, C. (1995). 7. properties of chemicals and estimation methodologies. Environmental Sciences Division - Publication No. 4081, 239–292. doi:10.1007/978-94-015-8520-0_7.

[32] Ulmeanu, S. M., Jensen, H., Carrupt, P., & Girault, H. H. (2003). Water–Oil Partition Profiling of Ionized Drug Molecules Using Cyclic Voltammetry and a 96-Well Microfilter Plate System. Analytical and Bioanalytical Chemistry, 375(8), 1077–1082. doi:10.1023/A:1025025804196.

[33] Lovric, M. (2000). A new access to Gibbs energies of transfer of ions across liquid N liquid interfaces and a new method to study electrochemical processes at well-defined three-phase junctions. Electrochemistry Communications, 2, 112–118. doi:10.1016/S1388-2481(99)00156-3.

[34] Bond, A. M., & Marken, F. (1994). Mechanistic aspects of the electron and ion transport processes across the electrode 1 solid 1 solvent (electrolyte) interface of microcrystalline decamethylferrocene attached mechanically to a graphite electrode. Journal of Electroanalytical Chemistry, 372, 125–135. doi:10.1016/00220728(93)03257-P.

[35] Valkó, K. (2004). Application of high-performance liquid chromatography-based measurements of lipophilicity to model biological distribution. Journal of Chromatography A, 1037, 299–310. doi: 10.1016/j.chroma.2003.10.084.

[36] Fuguet, E., Ràfols, C., Bosch, E., Abraham, M. H., & Rosés, M. (2006). Selectivity of single, mixed, and modified pseudostationary phases in electrokinetic chromatography. Electrophoresis, 27(10), 1900–1914. doi: 10.1002/elps.200500464.

[37] Fuguet, E., Rafols, C., Bosch, E., Abraham, M. H., & Roses, M. (2002). Solute-solvent interactions in micellar electrokinetic chromatography III. Characterization of the selectivity of micellar electrokinetic q chromatography systems. Chromatography A, 942(1), 237-248. doi: 10.1016/S0021-9673(01)01383-8.

[38] Fuguet, E., Ràfols, C., Bosch, E., & Rosés, M. (2003). Characterization of the solvation properties of sodium n-dodecyl sulfate micelles in buffered and unbuffered aqueous phases by solvatochromic indicators. Langmuir, 19(1), 55–62. doi:10.1021/la026307o.

[39] Dearden, J. C., & Bresnen, G. M. (1988). Review The Measurement of Partition Coefficients. Quantitative Structure-Activity Relationships, 3, 133–144. doi:10.1002/qsar.19880070304.

[40] Andrés, A., et al. (2015). European Journal of Pharmaceutical Sciences Setup and validation of shake-flask procedures for the determination of partition coefficients (log D) from low drug amounts, 76, 181–191. doi: 10.1016/j.ejps.2015.05.008.

[41] Leo, A., Hansch, C., Elkins, D., Law, A. H., & Behavior, B. N. (1971). Partition Coefficients and their Uses. Chemical Reviews, 71(6), 525–616. doi:10.1021/cr60274a001.

[42] Ghose, A. K., & Crippen, G. M. (1986). Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships I. Partition Coefficients as a Measure of Hydrophobicity. Journal of Computational Chemistry, 7(4), 565–577. doi:10.1002/jcc.540070419.

[43] Cheng, T., et al. (2007). Computation of octanol-water partition coefficients by guiding an additive model with knowledge. Journal of Chemical Information and Modeling, 47(6), 2140–2148. doi:10.1021/ci700257y.

[44] Best, S. A., Merz, K. M., & Reynolds, C. H. (1999). Free energy perturbation study of octanol/water partition coefficients: Comparison with continuum GB/SA calculations. The Journal of Physical Chemistry B, 103(4), 714–726. doi: 10.1021/jp984215v.

[45] Pagliara, A., Carrupt, P., Caron, G., Gaillard, P., Testa, B., & Ampholytes, O. (1997). Lipophilicity Profiles of Ampholytes. Chimica Acta, 2665(96). doi:10.1021/cr9601019.

[46] Nedyalkova, M. A., Madurga, S., Tobiszewski, M., & Simeonov, V. (2019). Calculating the Partition Coefficients of Organic Solvents in Octanol/Water and Octanol/Air. Journal of Chemical Information and Modeling, 59(5), 2257–2263. doi: 10.1021/acs.jcim.9b00212.

[47] Poole, C. F., & Poole, S. K. (1997). Interphase model for retention and selectivity in micellar electrokinetic chromatography. Journal of Chromatography A, 792, 89–104. doi:10.1016/S0021-9673(97)00699-7.

[48] Yang, S., & Khaledi, M. (1995). Chemical Selectivity in Micellar Electrokinetic Chromatography: Characterization of Solute-Micelle Interactions for Classification of Surfactants. Analytical Chemistry, 67(3), 499–510. doi:10.1021/ac00099a004.

[49] Poole, S. K., & Poole, C. F. (1997). Influence of Composition on the Selectivity of a Mixed-micellar Buffer in Micellar Electrokinetic Chromatography. Journal of the Chemical Society, Faraday Transactions, 34(February), 57–62. doi:10.1039/A607645I.

[50] Liu, Z., Zou, H., Ye, M., Ni, J., & Zhang, Y. (1999). Effects of organic modifiers on retention mechanism and selectivity in micellar electrokinetic capillary chromatography studied by linear solvation energy relationships. Journal of Chromatography A, 863, 69–79. doi:10.1016/S0021-9673(99)00949-8.

[51] Fu, C., & Khaledi, M. G. (2009). Selectivity patterns in micellar electrokinetic chromatography: Characterization of fluorinated and aliphatic alcohol modifiers by micellar selectivity triangle. Journal of Chromatography A, 1216, 1901–1907. doi: 10.1016/j.chroma.2009.01.041.

[52] Poole, S. K., & Poole, C. F. (1997). Characterization of surfactant selectivity in micellar electrokinetic chromatography. Journal of the Chemical Society, Faraday Transactions, 122(March), 267–274. doi:10.1039/A605799C.

[53] Poole, C. F., Poole, S. K., & Abraham, M. H. (1998). Recommendations for the determination of selectivity in micellar electrokinetic chromatography. Journal of Chromatography A, 798, 207–222. doi:10.1016/S0021-9673(97)01164-3.

[54] Vitha, M., & Carr, P. W. (2006). The chemical interpretation and practice of linear solvation energy relationships in chromatography. Journal of Chromatography A, 1126, 143–194. doi: 10.1016/j.chroma.2006.06.074.

[55] Abraham,M.H. (1992). Scales of Solute Hydrogen-bonding: Their Construction and Application to P hysicoc hem ica I and B ioc hem ica I Processes. CHEMICAL SOCIETY REVIEW, 096(5), 73-83, doi: 10.1039/CS9932200073.

[56] Poole, C. F., Poole, S. K., Abraham, M. H., & Es, I. R. O. S. (1997). Hydrogen Bonding. 42. Characterization of Reversed-Phase High-Performance Liquid Chromatographic C 18 Stationary

Phase. Journal of Planar Chromatography - Modern TLC, 10, 358–368. doi:10.1002/(SICI)1099-1395(199705)10:5<358: AID-POC907>3.0.CO;2-N.

[57] Hidalgo-Rodríguez, M., Fuguet, E., Ràfols, C., & Rosés, M. (2010). Solute–solvent interactions in micellar electrokinetic chromatography: VII. Characterization of sodium cholate–sodium deoxycholate mixed-micellar systems. Journal of Chromatography A, 1217, 1701–1708. doi: 10.1016/j.chroma.2010.01.001.

CHAPTER

II

THEORETICAL BACKGROUND

Chapter II
Theoretical background

2.1 Quantum mechanics

Quantum chemistry is the application of quantum mechanics to understand the properties and stability of atoms and molecules. Its main purpose is to study the behavior of materials based on their constituent particles: electrons and nuclei of atoms and molecules. The most important feature is energy quantization, and it is an important tool to explain chemical phenomena. One of the goals of quantum chemistry is to investigate systems and give their total energy (lowest energy) according to the stable molecular structure.

In the 20th century, new physics developed in science and covered atomic phenomena. This physic was able to resolve the issues with classical physics by discovering an explanation for the actions of inanimate objects. We have observed significant changes in physics since the turn of the twentieth century, including the creation of quantum mechanics. This science expresses motion at the atomic and molecular levels. On December 14, 1900, at a meeting of the German Physical Society, Max Planck presented a paper entitled The Theory of the Standard Energy Production Law, which describes radiant energy. Electromagnetism in a cavity requires energy to be quantized. This conference marked the beginning of a major revolution in physics. Although it took nearly a quarter of a century of quantum mechanics newly developed by Schrödinger and others, The date when this article was presented is regarded as the 'birth date' of quantum physics. This discovery was quickly followed by the application of quantization in atomic and molecular phenomena [1,2,3].

This great discovery made progress in various sciences including chemistry, physics, and biology. During the years 1925 to 1926, two different descriptions of mathematics by Heisenberg and Schrödinger were proposed for quantum mechanics. In the Heisenberg model, which is known as the matrix mechanics model, the quantum mechanical description of systems is based only on observable quantities. On the other hand, the Schrödinger model is based on a differential equation

[4]. Later, Dirac showed that the theories are the same and their difference is in the mathematical description which leads to the same results.

2.1.1 Calculations of the Energy of molecule structures

We can distinguish four main methods in computational chemistry: Molecular mechanics method, Semi-empirical methods, *ab intino*, and density functional method.

The empirical force field (EFF) simulations, also known as the molecular mechanic's method, is a computational technique for predicting the structures of very complex organic and organometallic molecules (up to several thousand atoms). This method does not require the Hamiltonian operator and molecular wave function, and the molecule is considered a set of atoms connected by bonds. The energy of the molecules is described based on the bond bending and stretching force constants and other parameters. Molecular mechanics methods are usually formulated for the molecular ground state, so it is not able to predict structures that include the process of bond formation and breaking and molecular properties related to the interaction of molecular orbitals [5]. In the semi-empirical quantum method, Hamiltonian is used which is simpler than actual molecular Hamiltonian, and the parameters used are adjusted to fit the experimental data or the results of the original calculation. In this method, several integrals are completely ignored or it is possible to replace them with experimental data. An example of these methods is Hückel molecular orbital method, which is used to investigate conjugated hydrocarbons with a one-electron Hamiltonian [5].

Ab initio means from the base or the beginning, and that indicates calculations based on fundamental principles, and in this method, the actual Hamiltonian molecules are used. No experimental parameters are used except the mass and charge of the electron and the nucleus, Planck's constant, and light speed. In a calculation of the self-consistent Hartree-Fock field, the real Hamiltonian is used, and to solve the Schördinger equation, the wave function is considered as the antisymmetric product of single-electron spin-orbitals, and a finite set is used [5,6].

2.1.2 Schrodinger and Born-Oppenheim approximation

The aim of the Born-Oppenheimer approximation is the quantum study of the multi-particle system, which consists of n electrons and N nuclei [4]. The Schördinger equation of this system is:

$$H\Psi(r_1,...,r_n, R_1,..., R_N) = E\Psi(r_1,...,r_n, R_1,..., R_N) \tag{2.1}$$

Where the Hamiltonian of the multi-particle system in the atomic unit appliance is defined as follows:

$$H = \sum_i(-\frac{1}{2M_I}\nabla_i^2) + \sum_I(-\frac{1}{2M_I}\nabla_I^2) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|r_i-r_j|} + \frac{1}{2}\sum_{I\neq J}\frac{Q_IQ_J}{|R_I-R_J|} - \sum_{iI}\frac{Q_I}{|r_i-R_J|} \tag{2.2}$$

In this Hamiltonian, r_i and R_I are the space operators of the *i*-th electron and the *I*-th nucleus, respectively. The charge and mass of electrons and *h* in atomic units are equal to 1. In this relationship, *m* and *Q* are the mass and charge of the nucleus, respectively. Born-Oppenheimer's approximation is based on the fact that the mass of the nuclei is bigger than the mass of electrons as a result, the specific speed of electrons compared to nuclei (ions) is much more. Therefore, nuclei are considered at rest in terms of time evolution compared to electrons and can be separated. So, the total wave function is the product of two parts electronic and nuclear are written:

$$\Psi(r_1,....r_n,R_1,...,R_N) = \Psi_{R_1...R_N}^{el}(r_1,...r_n)\Psi^i(R_1,...,R_N) \tag{2.3}$$

In which the electronic wave function is parametrically dependent on the location of the ions. Another used approximation is to consider nuclei like classical particles so that in the end all nucleus location operators are converted to location variables. As a result, quantum effects are limited to electronic wave functions only:

$$H^{el}\Psi_{R_1,...,R_N}^{el}(r_1,...,r_n) = E_{R_1,...,R_N}^{el}\Psi_{R_1,...,R_N}^{el}(r_1,...,r_n) \tag{2.4}$$

with the electron Hamiltonian

$$H^{el} = \sum_i(-\frac{1}{2}\nabla_i^2) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|r_i-r_j|} - \sum_{iI}\frac{Q_I}{|r_i-R_J|} \tag{2.5}$$

2.1.3 Density functional theory

Density functional theory (DFT) is a popular computational quantum mechanical modeling method that is used in different sciences, including physics, chemistry, and materials, to investigate the electronic structure of solids and their surface, and so for metal surfaces, usually in the ground state, in different systems (atoms, molecules, and dense phases). DFT is a quantum mechanics technique that is based on the electron density function. Also, it is an alternative to traditional *ab initio* techniques which are based on the wave function. In the density functional theory that we are discussing, electron density replaces the wave function as the main factor. The key benefit of DFT is more economic. The exact solution of quantum particle systems is not possible even in the Born-Oppenheimer approximation. As a result, the scientists developed different concepts to overcome these complexities and provide approximate solutions that are acceptable and relatively simple from a physical viewpoint. To solve this problem, usually, multiple approximations in the form of the theory are formed, and it is used.

In DFT calculations the system becomes simpler and its properties can be calculated more easily. The electron wave function of an n-electron molecule has on space variables and n spin variables. If the Hamiltonian operator only includes terms of one and two electrons, the molecular energy can be written as an integral equation.

$$\widehat{H_{el}} = \frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_a \frac{Z_a}{r_{ia}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} \qquad (2.6)$$

The first term in equation ($\frac{1}{2}\sum_i \nabla_i^2$) represents the sum of kinetic energy operators for Electrons, the second term is the total potential energy of attraction between electrons and nuclei, and the third term represents the repulsion energy between electrons. Condition *j>i* prevents repeated counting of electrons. On the other hand, the multi-electron wave function contains more information on physical properties. For this reason, the search for functions with a smaller number of variables compared to the wave function, which can be used to calculate energy and other properties, was given attention [7].

2.1.4 Hohenberg-kohn theorem
2.1.4.1 first Hohenberg-kohn theory

According to Hohenberg-Kohn's (HK) first theorem, two electronic systems with external potentials that differ by more than a constant cannot have ground states with the same electron density. In other words, a system's electrical density is specific to a certain external potential and the opposite.

$$\rho(r) \rightarrow V_{ext}(r) \tag{2.7}$$

$$V_{ext}(r) \rightarrow \rho(r) \tag{2.8}$$

Pirhohenberg and Water Cohen were able to determine the ground state molecular energy, wave function, and other electronic properties of molecules by using ground state electron density ρ(x, y, z) for molecules with the heterogeneous ground state [8,9].

$$\rho(x,y,z) = 2|\varphi(x,y,z,x_2,y_2,z_2)|^2 dx_2 dy_2 dz_2| \tag{2.9}$$

$$E_0 = E_v[\rho_0] \tag{2.10}$$

In equation (2.9), ρ is a function with three variables x, y, and z, and in equation (2.10), the sign 0 indicates the ground state function. E_0 is the ground state energy of an electron, which is a function of ρ_0. The density functional theory tries to calculate the energy value of the ground state (E_0) and other properties of the molecule by using the electron density of the ground state (ρ_0).

A function like $E[f]$ is a relation that will have different values for different $f(x)$ functions. For example, in the function:

$$E(f) = \int_{-\infty}^{\infty} f^*(x) f(x) dx \tag{2.11}$$

By integrating the expression $|f|^2$ for the quadratic functions of the integral $f(x)$ over the whole space, different numbers are obtained.

$$\omega[\varphi] = <\varphi|\hat{H}|\varphi>|<\varphi|\varphi> \tag{2.12}$$

Integral (2.12) as a function of the variable function φ can be obtained according to the harmonicas of φ.

Ψ_0 the electronic wave function of the ground state of an n-electron molecule is the following electronic Hamiltonian Eigen function:

$$\hat{H} = \frac{1}{2}\sum_{i=1}^{n} \nabla_i^2 + \sum_{i=1}^{n} v(r_i) + \sum_j \sum_{i>j} \frac{1}{r_{ij}} \tag{2.13}$$

$$v(r_i) = -\sum_a \frac{z_a}{r_{ia}} \tag{2.14}$$

The interaction potential energy between electron *i* and the nucleus, which is indicated by $v(r_i)$, depends on the coordinates of electron *i* in the form of x_i, y_i, and z_i and the coordinates of the nucleus. As long as the equation Schrödinger assumes the nucleus to be stationary, nuclear coordinates in the Schrödinger equation are considered constant. In DFT, $v(r_i)$ is the external potential in use to electron *i*. This is similar to the effect of external electric charges on the electronic system. When the external potential and the number of electrons n are known, the electron wave functions and the allowed energies of the molecule can be obtained by solving the Schördinger equation. Hohenberg and Cohen proved that for systems whose ground state is inequivalent, the wave function of the state ground and energy of the ground state is determined by the electronic density of the ground state. To observe this, it is sufficient to integrate equation (2.15) over the whole space.

$$\rho(r) = n\sum_{aums} \int \dots \int |\Psi(r_1, r_2, \dots r_n, m_{s_1}, \dots, m_{s_n})|^2 \, dr_2 \dots dr_n \tag{2.15}$$

Using Ψ, get the value of n from equation (2.16):

$$\int \rho_0(r)dr = n \tag{2.16}$$

The electronic energy of the ground state E_0 is a function of the function $\rho_0(r)$, which is written as. v indicates E_0 dependence on the external potential $v(r)$, which varies depending on the molecules. Sums of the kinetic energy of electrons and electron-nucleus interaction and repulsions Electron - Electron is an electronic Hamiltonian. The energy value is obtained by averaging this Hamiltonian.

$$\bar{E} = \bar{T} + \overline{V_{ne}} + \overline{V_{ee}} \tag{2.17}$$

The average values in this equation are determined from the electronic wave function of the ground state, which is also determined from $\rho_0(r)$.

$$E_0 = E_v[\rho_0] = \bar{T}[\rho_0] + \overline{V_{ne}}[\rho_0] + \overline{V_{ee}}[\rho_0] \tag{2.18}$$

$$\widehat{V_{ne}} = \sum_{i-1}^{n} v(r_i) \tag{2.19}$$

26

$$v(r_i) = -\sum_a \frac{Z_a}{r_{ia}} \tag{2.20}$$

$$\bar{V}_{ne} = \langle\Psi_0|\sum_{i=1}^n v(r_i)|\Psi_0\rangle = \int \rho_0(r)v(r)dr \tag{2.21}$$

the equation (2.21) is obtained according to the following relation:

$$\int \Psi^* \sum_{i=1}^n B(r_i)\Psi d\tau = \int \rho(r)B(r)dr \tag{2.22}$$

The function $v(r)$ is determined by the energy of the electron interaction potential at point r. As a result, only $\bar{V}_{ne}[\rho_0]$ is known, while $\bar{T}[\rho_0]$ and $\bar{V}_{ee}[\rho_0]$ are unknown functions. Therefore, the relationship is as follows:

$$E_0 = E_v[\rho_0] = \int \rho_0(r)v(r)dr + \bar{T}[\rho_0] + \bar{V}_{ee}[\rho_0] = \int \rho_0(r)v(r)dr + F[\rho_0] \tag{2.23}$$

Here $F[\rho_0]$ is determined from $F[\rho_0] = \bar{T}[\rho_0] + \bar{V}_{ee}[\rho_0]$ relation which is independent of an external potential. Because the function $F[\rho_0]$ is unknown, Equation (2.23) cannot be used to determine E_0 from ρ_0.

2.1.4.2 Second Hohenberg-kohn method

The second HK theorem states that there cannot be two different systems with the same electron in the ground state, so the energy is a universal function of the electron density, so

$$E = E[\rho(r)] \geq E_0 = E[\rho_0(r)] \tag{2.24}$$

To convert equation (2.16) into a practical relation, we need the second proven Hohenberg-Cohen theorem and the Cohen-Sham extension [10]. Any density function $\rho_{tr}(r)$ that is obtained from the integral equation $\int \rho_{tr}(r)dr = n$, the condition $\rho_{tr}(r) \geq 0$, and the inequality $E_0 \leq E_v[\rho_{tr}]$, all r follow it, E_v is the energy function in equation (2.20).

If the wave function of the ground state is normal, the energy function of $E_v[\rho_{tr}]$ will have its minimum value when $E_0 = E_v[\rho_0]$, ρ_0 is the electron density of the ground state. If we use the function Ψ_{tr} (according to ρ_{tr}) as a function for the molecule with Hamiltonian $\hat{H}$, the equation becomes (2.25).

$$\langle\Psi_{tr}|\bar{H}|\Psi_{tr}\rangle = \langle\Psi_{tr}|\hat{T} + \hat{V}_{ee} + \sum_{i=1}^n v_{ri}|\Psi_{tr}\rangle \geq E_0 = E_v[\rho_0] \tag{2.25}$$

Because the average kinetic energy and potential energies are a function of the electron density and Ψ_r has been replaced by Ψ_0 in equation (2.21), equation (2.25) can be written as follows:

$$\bar{T}(\rho_{tr}) + \bar{V}_{ee}[\rho_{tr}] + \int \rho_{tr} v(r) dr \geq E_v[\rho_0] \tag{2.26}$$

The functions $\bar{T}$ and $\bar{V}_{ee}$ are the same as functions (2.23) and (2.26), although ρ_0 is different Ψ from ρ_{tr}.

2.1.5 Kohn-Sham method

According to Hohenberg and Kohen's theorem (KS), if we know the electron density of the molecular ρ_0 ground state basically, we will be able to determine the molecular property of the target molecule in its base state even if we do not have the wave function of the molecule. Typically, in quantum mechanics, the wave function and then employ it to integrate (1.20) to calculate the electron density. The theory of Hohenberg and Cohen is unable to compute E_0 from ρ_0. (When F is unknown). Additionally, it doesn't provide a solution for determine ρ_0 without Ψ_0. in 1965 Cohen and Shem decided to develop a suitable method for determining ρ_0 in addition to calculating E_0 using ρ_0.

In this method, accurate results are obtained, but due to the existence of unknown function equations, it is an approximate method and the results are approximate. Cohen and Sham have theorized the ideal electron system, which consists of n electrons with a single external potential and no internal interactions. Equation (2.27) describes the reference system's Hamiltonian.

$$\hat{H}_s = \sum_{i=1}^{n} \left[-\frac{1}{2} \nabla_i^2 + v_s(r_i) \right] = \sum_{i=1}^{n} \hat{h}_i^{ks} \tag{2.27}$$

$$\hat{h}_{ks}^i = -\frac{1}{2} \nabla_i^2 + v_s(r_i) \tag{2.28}$$

$\hat{h}_{ks}^i$ is the one-electron Hamiltonian. According to the Pauli principle, the wave function of the ground state ($\Psi_{s,v}$) must be antisymmetric, that is, the reference Cohen-Schum orbitals U_i^{Ks} form a Slater determinant so that $\theta_i^{Ks}(r_i)$ is the special spatial part of the function of the one-electron operator $\hat{h}_i^{Ks}$.

$$\Psi_{s,0} = |U_1 U_{2...} U_n| \tag{2.29}$$

$$U_i = \theta_i^{Ks}(r_i)\sigma_i \tag{2.30}$$

$$\hat{h}_i^{Ks} \theta_i^{Ks} = \varepsilon_i^{Ks} \theta_i^{Ks} \tag{2.31}$$

σ_i is the spin function, which can be α, β, and ε_i^{Ks} is the Kuhn-Schum orbital energies. Equation (2.23) was modified by Cohen and Sham as follows:

$$\Delta \bar{T}[\rho] = \bar{T}[\rho] - \bar{T}_s[\rho] \tag{2.32}$$

$\Delta \bar{T}$ is the difference in the average kinetic energy of the ground state between the molecule and the reference system where there is no interaction between electrons. Also, the equation of (2.33):

$$\Delta \bar{V}_{ee}[\rho] = \bar{V}_{ee}[\rho] - \frac{1}{4} \iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 \, dr_2 \tag{2.33}$$

that r is the distance between two points with x_1, y_1, z_1 and x_2, y_2, z_2 coordinates.

The value of the right-hand integral in equation (2.33) represents the energy of electron-electron repulsion. If we consider the electrons as successively charged with density p.

Equation (2.23), changes like this:

$$E_v[\rho] = \int \rho(r) v(r) dr + \bar{T}_s[\rho] + \frac{1}{2} \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + \Delta \bar{T}[\rho] + \Delta \bar{V}_{ee}[\rho] \tag{2.34}$$

The functions of $\Delta \bar{T}$ and $\Delta \bar{V}_{ee}[\rho]$ are unknown. And as a result, we define Exchange-correlation energy as below:

$$E_{xc}[\rho] = \Delta \bar{T}[\rho] + \Delta \bar{V}_{ee}[\rho] \tag{2.35}$$

So, we have

$$E_0 = E_v[\rho] = \int \rho(r) V(r) dr + \bar{T}_s[\rho] + \frac{1}{2} \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + E_{xc}[\rho] \tag{2.36}$$

The first three terms on the right side of equation (2.36) can be easily calculated using ρ of the base state, but the fourth term, E_{xc}, is not easily obtained, but it is smaller compared to other terms. KSDFT plays an important role in the accuracy of molecular properties calculations in obtaining a suitable approximate value for E_{xc}.

To be able to calculate the terms of equation (2.36), we must obtain the electron density of the ground state. It should be noted that, we defined the reference system as systems with electrons without interaction with the same electron density that the molecule has in the ground state.

$$\rho_0 = \rho_s = \sum_{i=1}^{n} |\theta_i^{Ks}|^2 \tag{2.37}$$

by using equation (2.20), (2.36), and (2.37) is obtained equation (2.38):

$$E_0 = -\sum_a Z_a \int \frac{\rho(r_1)}{r_{1a}} dr_1 - \frac{1}{2}\sum_{i=1}^{n}\langle\theta_i^{Ks}|\nabla_1^2|\theta_i^{Ks}\rangle + \frac{1}{2}\frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + E_{xc}[\rho] \tag{2.38}$$

According to equation (2.38), we obtain the value of E_0 by ρ if we get $(\theta_i^{Ks})KS$ orbitals and the function of E_{xc}. Nuclear repulsion is another part of electronic energy, and it may be obtained by adding the repulsive energy V_{NN} into equation (2.32). Here, we discovered Cohen-Sham orbitals. According to the theory and method of Hohenberg and Cohen to find ρ_0, it is necessary to change the value of ρ so that the energy of the ground state of $E_v[\rho]$ reaches the minimum value. In equation (2.37), provided it is orthonormal, $(\theta_i^{Ks})KS$ orbitals can be changed instead of ρ. This is the Hartree-Fack method, which obtains the molecular energy of equation (2.39).

$$\hat{F}(1)\emptyset_i(1) = \varepsilon_i\emptyset_i(1) \tag{2.39}$$

$$\hat{F}(1) = \hat{H}^{core}(1) + \sum_{i=1}^{n/2}[2\hat{J}_i(1) - \hat{K}_i(1)] \tag{2.40}$$

$$\hat{H}^{core}(1) = -\frac{1}{2}\nabla_1^2 - \sum\frac{Z_a}{r_{1a}} \tag{2.41}$$

We may determine the ground state energy by minimizing equation (2.38).

$$\left[-\frac{1}{2}\nabla_1^2 - \sum_a\frac{Z_a}{r_{1a}} + \int\frac{\rho(r_2)}{r_{12}}dr_2 + v_{xc}(1)\right]\theta_i^{Ks}(1) = \varepsilon_i^{Ks}\theta_i^{Ks}(1) \tag{2.42}$$

The above relationship can be written in other ways:

$$\left[-\frac{1}{2}\nabla_1^2 - v_s(1)\right]\theta_i^{Ks}(1) = \varepsilon_i^{Ks}\theta_i^{Ks}(1) \tag{2.43}$$

$$\hat{h}^{Ks}(1)\theta_i^{Ks}(1) = \varepsilon_i^{Ks}\theta_i^{Ks}(1) \tag{2.44}$$

The value of exchange-correlation potential v in relation (2.42) is obtained from relation (2.45):

$$v_{xc}(r) = \frac{\delta E_x[\rho(r)]}{\delta\rho(r)} \tag{2.45}$$

$E_{xc}[\rho]$ is a function of ρ, while ρ is a function of r. Therefore, v_{xc} is a function of r and as a result a function of x, y, z. The one-electron Cohen-Sham operator in relation (2.44) ($\hat{h}^{Ks}(1)$) is the same as the Fock minimum in the Hartree-Fock equation (2.40). Exchange operator's $(-\sum_i^n = 1\hat{K}_j)$ have been replaced by v_{xc}in Fack's operator, but it should be noted that there are two exchange effects (which come from the antisymmetric of the wave function) and electron correlation.

According to the summary of the content, there is just one issue remaining to solve, which is finding the function $E_{XC[\rho]}$ to calculate E_0 and ρ_0. It may be said that the main issue in DFT is to find a suitable approximation for $E_{xc}[\rho]$. There are various approximate methods such as LDA, LSDA, GGA, etc., the most important of which are mentioned below.

2.1.6 Exchange and correlation functional

In the KS method, all the different contributions to the system's energy are known, except for the exchange and correlation energy, whose function plays an important role in the correct application of DFT. In Born-Oppenheimer's approximation, this theory is accurate but the exact form is unknown so for description it can be divided into two terms electron exchange and electron correlation.

$$E_{XC}[\rho] = E_X[\rho] + E_C[\rho] \tag{2.46}$$

The many-body wave function must be antisymmetric to allow for the exchange of both electrons with the same spin, which leads to electron exchange. The antisymmetric wave function is a general expression of the Pauli exclusion principle. But the Coulomb energy of the electronic system decreases with increasing spatial separation between electrons of the same spin. The motion of each electron is correlated with the motion of the other electrons, reducing the Coulomb energy between electrons with different spins. Electron correlation also assists in keeping the spatial separation of the electrons with odd spins. In the following, some well-known approximations for exchange and correlation functions are stated.

2.1.6.1 Local Density approximation

In the local density approximation, it is assumed that the electron density (ρ) changes very slowly with the position, and as a result, the correlation exchange energy E_{xc} is obtained from equation (2.47).

$$E_{xc} = \int \rho(r)\varepsilon_{XC}[\rho(r)]dr \tag{2.47}$$

According to this equation, Exchange correlation energy $\varepsilon_{XC}[\rho(r)]$ is constant for each electron in the homogeneous gas-electron model with density. Although equation (1.49) is approximate, it has high accuracy in predicting structural properties.

The accuracy of this method decreases with the change of the electron density in the system, and for many molecules, the use of the LDA approximation (in which the Exchange-correlation potential depends only on ρ and not on the derivative of ρ) causes the bond energy to be larger than the actual energy [68]. Local-Spin Density approximation (LSDA) also considers the spin of molecules and atoms in local density calculations. LSDA can accurately determine molecular geometry, dipole moments, and vibrational frequencies, but its results are not suitable for dissociation energies [11].

2.1.6.2 Generalized gradient Approximation (GGA)

LDA and LA approximations are based on the homogeneous gas-electron model, and as mentioned, this model is suitable for a system where the electron density changes slowly with spaces. If for heterogeneous systems (mostly molecular systems), the Exchange-correlation energy depends on the density in the adjacent volumes in addition to the local density. Therefore, Local Density Approximation is not a correct approximation for these systems. A simple way the correction of the correlation function is not to make it depends only on the local value of the density, but it also depends on the speed of density changes (density gradient) [10]. Typically, the sub-integral function includes electron density gradients for this purpose. So:

$$E_{XC}^{GGA}\left[\rho^a,\rho^\beta\right] - \int f\left[\rho^a,\rho^\beta,\nabla\rho^a,\nabla\rho^\beta\right]dr \qquad (2.48)$$

Where f is a function of spin densities and their gradients. Usually, E_{xc}^{GGA} is split into two parts, exchange, and correlation (modeled separately).

$$E_{xc}^{GGA} = E_x^{GGA} + E_c^{GGA} \qquad (2.49)$$

Based on the approximations used to describe the Exchange-correlation part, there are different pure DFT methods. The PBE method includes Perdew-Burke-Ernzerhof correction functions and it is one of the most typical approximations methods of the generalized gradient.

2.1.6.3 Hybrid methods

These methods are a combination of functions used in other DFT methods with a part of Hartree-Fock computations. One of the most widely used hybrid methods use in DFT calculations in recent years is the B3LYP calculation technique.

B represents the term E_x^{GGA} (exchange contribution suggested by Becke) and LYP defines E_c^{GGA} (correlation contribution proposed by Lee, Yang, and Parr), and number 3 symbolizes three experimental parameters.

2.1.7 DFT calculation with Dispersion corrected (DFT-D)

With computational and theoretical methods in science, density functional theory is known as a simple and reliable method of calculating the electronic and structural properties of atoms, molecules, and solids. Since in this theory approximations for the exchange-correlation functions are used for calculations and the pure density function is not employed, these weak interactions cannot be considered successful with these calculations. Non-covalent forces like hydrogen bonds and van der Waals interactions (often representing double interactions) push molecules away from each other, even though they play a significant role in the creation, stability, and application of materials. Many exchange-correlation functions are unable to compute these interactions.There are different ways to include the scattering correction and take into account van der Waals forces in density functional theory computations. The method of Ortmann, Bechstedtant, Schmidt (OBS) [12] and the approach of Jurecka et al [13], and the technic of Tkatchenko and Scheffler (TS) [14] and the Grimme ́s Scheme method [15] are among the main methods in DFT-D calculations.

The semi-empirical combination of DFT with the correction of pairs of atoms will create a significant improvement in the results of pure DFT or pure force field method in describing structures and classifying polymorphs (compounds with different crystal structures) according to their energy. In this semi-empirical method of Dispersion corrected, the missing contribution of Dispersion in interatomic interactions can be estimated with a simple uniform potential. At a large distance, this potential is shown by the term $C_{6,ij}R_{ij}^{-6}$, where C_{ij}^6 is the dispersion coefficient for i and j atoms, which are located at a distance r_{ij} from each other. The Dispersion coefficient in the OBS method depends only on the chemical species, and in the TS method, it is calculated for each

structure taking into account the dependence of the polarizability on its volume. Dispersion in a short distance is adapted by multiplying the balancing function $f(R_{ij}^0, R_{ij}$ in the above term with the DFT potential. The balancing function reduces the contribution of excess Dispersion to zero under the effective cut-off range defined by the appropriate combination R calculated from the van der Waals radii of pairs of atoms. Then, the Dispersion-corrected exchange-correlation function is formed by adding the correction potential to the normal DFT exchange-correlation functional. The total energy, carrying into account the scattering correction, is written as [16,17]:

$$E_{tot} = E_{DFT} + S_i \sum_{i=1}^N \sum_{j>i}^N f(S_R R_{ij}^0 R_{ij}) C_{6,ij} R_{ij}^{-6} \tag{2.50}$$

E_{DFT} is the standard total energy DFT and the summation is performed over all N atoms of the system. In this method, $C_{6,ij}$ and R_{ij}^0 in heterogeneous cores are obtained by approximating semi-empirically determined parameters of core types. The difference between DFT exchange-correlation functionals in Dispersion or short to medium-range interactions is taken into account with a suitable change in the correction potential through the S_6 or S_R parameter. S_6 is a parameter that has a specific value for each Exchange-correlation potential [15,18].

Conducting research on the application of DFT-D methods on the complex of mesotetraphenyl propylene and fullerene C_{60} has shown that GGA subordination using PBE corrected by Gram method, PW91A corrected by OBS method in terms of intermolecular distances gives good results compared to experimental data [19].

2.1.8 Basis set

Basis sets are mathematical functions used to represent and describe molecular orbital and the electronic space around the desired species is used [20]. Most quantum mechanical calculations begin with selecting a basis set. Ruthaan suggested the set of base functions in 1951. He considered one-electron functions as a linear combination of a complete series of known functions under the name of the ϕ_M-basis function.

$$\Psi = \sum_{M=1}^N c_{Mi} \phi_M \tag{2.51}$$

The basic functions coefficients in the molecular orbital are proposed to minimize energy. The three types of orbitals that are often applied to obtain basis functions are as follows:

- **Hydrogen Like Functions**

One possibility for a basis set is to use hydrogen atom orbitals. Naturally, in the wave functions of the hydrogen atom, there are no envelope effects and other effects related to the repulsion between electrons [59]. The form of these functions is expressed by equation (2.52):

$$\Psi_{n,l,m} = (r,\theta,\phi) = N_{n,l}\rho^{l}L_{n+1}^{2l+1}e^{\frac{\rho}{2}}Y_{l}^{m}(\theta,\phi) \qquad (2.52)$$

In this equation, n, l, and m are quantum numbers, $N_{n,l}$ is the Normalizing Constant, ρ^{l} is a power function of r, Y_{l}^{m} is the Spherical Harmonic, and L is the Laguerre Polynomial.

- **Slater Type Orbitals (STO)**

Slater-type orbitals are in the form of equation (2.53).

$$N_{nlm,\xi}(r,\theta,\varphi) = Nr^{n-1}e^{-\xi(r-R_A)}Y_{l}^{m}(\theta,\varphi) \qquad (2.53)$$

ξ is the Slater orbital symbol, n, l, and m are the quantum numbers, $Y_{l}^{m}(\theta,\varphi)$ is the Spherical Harmonic, N is the Normalizing Constant, R_A is the functional center and r is the distance of the electron from the nucleus. These functions are not orthogonal, but they produce a complete set. In these functions, a peak on the core makes these functions an acceptable description of the distribution of electrons on Provide a core. Generally, mathematical work with these functions is difficult.

- **Gaussian Type Orbitals**

These functions are shown in the general form of equation (2.54):

$$X_{nlm,a}(r,\theta,\varphi) = Nr^{2n-2-1}e^{-a(r-R_A)^2}Y_{l}^{m}(\theta,\varphi) \qquad (2.54)$$

a is the symbol of the Gaussian orbital. R_A is the center of the Gaussian function, and since this function corresponds to the atom's nucleus, R_A is the coordinate of the atom's nucleus.

The orbital symbol is a positive and non-zero number that determines the spread of the function around r in space. Calculations with GTO are simpler than STO, but they have problems such as the low speed of convergence of these functions, predicting the probability of finding the electron

at distances far from the nucleus, and not having the desired sharp peak in the nucleus and the areas close to it. Figure 1 shows the difference between GTO and STO.

Distance to the nucleus
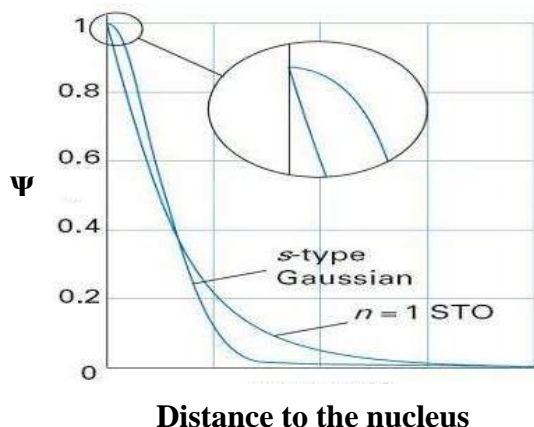
Figure 1. A representation of the form of GTO and STO and their differences [21]

The most important factor in the division of basic functions is the number of functions used in them. Based on this, the most common types of basic functions are:

2.1.8.1 Minimal basic set

This collection contains the minimum number of basic functions of atomic orbitals that are necessary to describe a system. Therefore, for H and He, a minimal basis set contains 1s orbital. For Li to Ne atoms, it also contains 1s, 2s, $2p_x$, $2p_y$, and $2p_z$ orbitals. The STO-nG basis series presented by Papel and her colleagues are minimal basis series in which n Gaussian functions are used to generate each Slater function. The most common basic sets of this type are:

STO-5G, STO-4G and STO-3G

2.1.8.2 Basis set of Double Zeta (DZ)

In a double zeta basis set, each member of a basis set is replaced by two functions. As a result, in comparison to the minimal basis set, the number of functions doubles.

Of course, in some cases, the number of collections of the double zeta basis is slightly less than double. By doubling the number of functions, better solutions will be obtained compared to the minimal basis sets.

2.1.8.3 The basis set of Split – Valence (SV)

In the Split - Valence basis set, more basis functions are indicated for each of the atomic orbitals of the valence layer. This way that for each atomic orbital of the capacity layer, two dense functions (or more) are used, and each atomic orbital of the inner layer is characterized by a Gaussian dense function. The most popular mentioned basis collections are:

6-31G, 6-21G, 4-31G and 3.21G

For example, 6-21G shows that the internal orbitals contain a principal function obtained from the linear combination of six elementary Gaussian functions, and the valence layer consists of two principal functions, the first consisting of two elementary Gaussian functions and the second containing only one primary Gaussian function.

2.1.8.4 Polarized basis set

Although double zeta sets allow orbitals to change size, there is a possibility to change the shape of the orbitals in such basis sets does not have. In polarized sets, it is possible to add orbitals with larger angular momentum than the ground-state orbitals of the atom. Polarized basis functions are functioning whose size the orbital angular motion (l) that they describe is one unit more than what appears in the space of occupied atomic orbitals, in other words, p functions for H, He atoms, and d functions for the second and third rows of the periodic table are polarized functions.

2.1.8.5 Diffuse basis function

In situations where the probability of the presence of electrons at a relative distance further away from the core increases, the functions that have been said so far do not have satisfactory results in calculations. The use of Diffuse functions, especially in excited state calculations, in anions where the electron density is more concentrated at distances far from the nucleus, in neutral systems which have free electron pairs, in systems with weak bonds such as hydrogen bonding, and also in the calculations of properties such as dipole moments and Polarizability is necessary.

In such cases, the basis sets should be supported by adding Diffuse basis orbitals. The normal and polarized capacity functions that were described do not provide enough radial flexibility to describe these cases. In such cases, the addition of Gaussian orbitals with small orbital exponents, which extend over long distances from the nucleus, is utilized, and these ground functions are called Diffuse functions.

2.1.8.6 Numerical basis sets

When the utilization of numerical three-dimensional integrals became feasible for computations, the adaptability of integration techniques led to a more efficient adjustment in employing the basis set. Numerical basis sets featuring initial convergence characteristics found application in quantum mechanics calculations.

In the numerical basis set, numerical orbitals are used for the basic functions in which each function corresponds to an atomic orbital. In the numerical calculation of basic functions, the angular part of each function is proportional to the spherical coordinate $(Y_{lm}(\theta,\phi))$,and its radial part is obtained by numerically solving atomic DFT equations. The set of the atomic basis set by determining the cut-off value, according to the quality in the calculation is selected, it will be limited. And this point will make calculations faster, especially in solid-state systems.

The possibility of complete separation of the molecule to its constituent atoms and as a result minimizing the superposition effect of these orbitals makes it possible to show excellent results even for weak bonding. The usual method for creating a numerical basis set includes uncharged atom calculations, ion with two positive charges and calculations of hydrogen orbitals. Minimum bases only with the Calculation of uncharged atoms are obtained. Orbitals that are obtained from the calculation of doubly positively charged ions are added to the minimal basis set to the double

basis set. generate a number that is similar to the concept of double zeta sets in Gaussian. Utilizing polarized functions and additional orbitals expands the basis set beyond simple numerical multiplication, enabling calculations of hydrogen orbitals and incorporating nuclear charge. For the hydrogen atom, 1s hydrogen orbitals with $\frac{1}{3}$ core charge are calculated for the minimum basis set, and 2p hydrogen orbitals with the same core charge for polarized functions [22].

Some of the numerical basis sets are:

- **MIN**

The minimal basis set, in which for each occupied atomic orbital, a function equivalent to one atomic orbital (AO) is used, which, despite increasing the calculation speed, is not accurate enough.

- **DN**

The double numerical basis set is equal to the double Gaussian Zeta (DZ) set, in which two functions are employed for each capacity orbital, and compared to the minimal basis set provides better results.

- **DND**

The double numerical basis set, which is added to consider the polarization of orbitals for non-hydrogen atoms, d-type polarization functions.

- **DNP**

This basis set includes all DND functions. In addition, it includes P-type functions for hydrogen atoms. This basis set is more than ten times faster than the corresponding basis set in Gaussian [23]. Research has displayed that the results obtained with the DNP basis set are consistent with the results obtained with the 6-31G** basis set in Gaussian, and there is no requirement to use a larger basis set to improve structural and geometric properties. The numerical basis set is more accurate in comparison with the set of the same size in the Gaussian.

- **TNP**

The numerical triple basis set with polarization, in other words, DNP, includes additional polarized functions for all atoms. This basis set is available for hydrogen to chlorine atoms (except He, Ne). The best accuracy is achieved at a high computational cost. Calculations have shown that the examination of the Basis set superposition error of the BSSE basis set with the Counterpoise correction method shows that the numerical basis set produces a much smaller BSSE than the Gaussian basis set, which is caused by the numerical production of spherical atomic orbitals [24,25]. By analysing the binding energy of molecules with a numerical basis set, it can be seen that results with high accuracy and close to the results of Gaussian functions are obtained, and the use of the TNDP basis set will reduce the error in the binding energy. Also, the results of combined TNDP research have shown that numerical basis sets are superior to Gaussian basis sets in terms of the ratio of the accuracy of results to computing facilities (the most required memory and processing time) [25].

2.2 Continuum solvent models

Previous calculations make it possible to study molecular properties in the gas phase. While in the solution phase, especially in the polar solvent, these properties change significantly. Although it is possible to model solvent effects with solvent molecules (explicit solvent coating model) in quantum calculations, it will be very difficult and expensive in terms of calculation and execution. In addition to these cluster calculations, due to the importance of long-range electrostatic interactions, it cannot accurately determine the solvation energy. The primary effect of a solvent, especially a polar solvent, is to change the dipole moment of the molecule. On the other hand, the solvent is also polarized under the influence of the dipole moment of the molecule and creates an electric field, which is called the reaction field. Considering such an effect of the solvent, it can be modelled as a continuum with a certain dielectric constant. There have been a variety of models put out to include the solvent effect in quantum calculations, each with a different level of accuracy and complexity. These models generally have a self-consistent reaction field (SCRF) algorithm. Continuum solvent models create a reaction field dependent on the electron density of the solvent. Therefore, during the convergence of the wave function, iteratively, continuously, and self-consistently, the effects of the solvent should also be considered. The interaction between a solute

molecule and the dielectric constant around it is added as a potential term to the electronic Hamiltonian of the molecule [26].

- **polarizable continuum model**

The most popular method in computational chemistry for investigating solvent effects is the polarizable continuum model. This is a low-cost calculation method.

Two types of this model have been developed:

Dielectric PCM type (D-PCM) in which the continuum is like a polarizable dielectric and Semiconductor PCM type (C-PCM) in which the continuum is considered a semiconductor. In this method, the soluble molecule is placed in a cavity that is obtained from the overlap of spheres with a radius 1/2 times the van der Waals radius and centered on the nucleus of all atoms.

Potential operator solvent and solute interaction are determined with the help of a numerical method. The surface charge of the cavity (Apparent surface charge, ASC), which is considered to interact with the solute, is approximated to point charges during surface divisions. The converged charges are used to find an estimate of the interaction potential operator as follows:

$$\hat{V} \sum_i \varphi_{in}(r_i) + \sum_a Z_a \varphi_{in}(r_a) \tag{2.55}$$

where $\varphi_{in}(r)$ is the electric potential related to the dielectric polarization and is obtained from the following equation:

$$\varphi_{in}(r) = \sum_k \frac{Q_k}{|r - r_k|} \tag{2.56}$$

In this relation, Q_k is the point charge at point k with radius r_k.

Generalized forms of PCM are also presented. Among them, there is the uniform density PCM model (IPCM), in which the size of the cavity changes based on the change in the size of the resolved electronic wave function in each iteration.

- **Conductor-like screening model (COSMO)**

This model is a generalized form of the PCM model in which the continuum is considered an electrical conductor instead of a dielectric [27]. Therefore, a special version of it is called conductive PCM, which is suitable for geometric optimization solutions. In this model, the amount of charge on the cavity surface (which is known as apparent surface charge) is approximated by an extremely simple method. If the electric charge distribution is known, the charge available on the selected parts on the inner surface of the cavity (q^*) can be determined. Then it can be reduced to $f(\varepsilon)$ according to a function of the dielectric constant of the solvent, q:

$$q = f(\varepsilon)q^*$$
(2.57)

The coefficient f is approximated as follows:

$$f(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon + 0.5}$$
(2.58)

where ε is the transmittance coefficient. The COSMO method is a relatively accurate method for solvents with a high transmittance because a solvent with a high transmittance behaves closer to an ideal conductor. By using this model for aqueous solutions, with a high transmission coefficient ($\varepsilon \approx 80$), very acceptable results are obtained. Apart from the numerical advantage, another advantage of this method in comparison with other dielectric constant methods is the significant reduction of the external charge error, which is caused by a small part of the electron density that is resolved outside the cavity [28].

2.3 Mulliken and Hirschfeld charges

- **Charge Mullikenk**

Mulliken population to describe the electronic charge distribution in the molecule and determine the nature of the bond, antibonding and nonbonding molecular orbitals are used for specific pairs of atoms. To understand this population, consider the real normal molecular orbital, which is obtained from the combination of two atomic orbitals (as in the following equation):

$$\Psi_i = c_{ij}\varphi_j + c_{ik}\varphi_k$$
(2.59)

j and k are constant coefficients that indicate the weight of each atomic orbital in the molecular orbital. The charge distribution is the probability density resulting from the second power of this wave function according to the following equation:

$$\Psi_i^2 = c_{ij}^2 \varphi_j^2 + c_{ik}^2 \varphi_k^2 + 2c_{ik} \varphi_i \varphi_j \tag{2.60}$$

By integrating all the electronic coordinates and considering that the molecular orbital and atomic orbitals are normal. we reach to the following conclusion:

$$1 = c_{ij}^2 + c_{ik}^2 + 2c_{ij} c_{ik} S_{jk} \tag{2.61}$$

S_{jk} is the overlap integral of two atomic orbitals. C_{ij}^2 and C_{ik}^2 is called atomic orbital population and $2C_{ij} C_{ik} S_{jk}$ is called overlapping population. The overlap population of molecular orbitals is positive, while it is zero for non-bonding orbitals and negative for antibonding orbitals. For a better display of these populations, these expressions are in the form of a matrix for each molecular orbital are classified. This matrix is called the Mulliken population matrix. If two electrons are in the molecular orbital, this population doubles. Every column and every row in the population matrix corresponds to an atomic orbital and can represent the diagonal elements of the atomic orbital population and the non-diagonal elements of the overlapping population. This way for the introduced system, the population matrix is written as follows:

$$Pi = \begin{pmatrix} C_{ij}^2 & 2C_{ij} C_{ik} S_{jk} \\ 2C_{ij} C_{ik} S_{jk} & C_{ik}^2 \end{pmatrix} \tag{2.62}$$

Since there is a population matrix for each molecular orbital, a large volume of Information will be created and the formation of the net population matrix will reduce this amount of information.

The total of all population matrices for occupied orbitals is the net population matrix:

$$NP = \sum_{i=occupied} Pi \tag{2.63}$$

In this matrix, the diagonal elements show the total charge of each atomic orbital, and the other element's diagonals are indicative of overlapping populations that identify each of them in the bond between two atoms.

The gross population matrix summarizes the data in another way. This matrix combines the overlap populations with the atomic orbital population for each molecular orbital. The columns of the gross population matrix correspond to molecular orbitals and the rows correspond to atomic orbitals. The elements of the matrix are obtained as follows:

$$GP_{ij} = Pi_{jj} + \frac{1}{2} \sum_{k \neq j} Pi_{jk} \tag{2.64}$$

where Pi is the population matrix for the i molecular orbital and Pi_{jj} is the atomic orbital population and Pi_{jk} is the overlap population for j and k atomic orbitals in the i molecular orbital.

To further reduce the amount of data, it is possible to use overlapping and atomic populations of atoms instead of atomic orbitals. The resulting matrix is called the reduced population matrix. Reduced population from pure population matrix plus atomic orbital population and population overlap of all atomic orbitals of the same atom is obtained. The rows and columns of the reduced matrix correspond to atoms.

Atomic orbital charges by adding the row elements of the impurity population matrix for the occupied molecular orbitals are obtained. Finally, the net charge on each atom is obtained by subtracting the atomic charge from the nuclear charge adjusted for complete shielding with ls electrons [29]. Mulliken's method is one of the most common calculation methods for the charge, spin, and bond degree analysis; But it is very sensitive to the selection of the base set [30].

- **Hirshfeld charge**

Hirschfeld population analysis defines the atomic charges by a Deformation density distribution between the atoms in the molecule. Compared to Mulliken population analysis, Hirschfeld's population analysis produces non-negative Fukui function indices (which will be discussed in the next sections) in all chemical systems and has less dependence on the basis set chosen for calculation [31-33]. Deformed density is defined as follows:

$$\rho_d(r) = \rho(r) - \sum_a \rho_a(r - R_a) \tag{2.65}$$

Where $\rho(r)$ is the molecular density according to equation (1.62) and ρ_a *(r-Ra)* is the density of the free atom placed in R_a coordinates.

$$\rho(r) = 2 \sum_i |\Psi_i^2(r)| \tag{2.66}$$

44

Using the deformation density, the effective atomic charges *(q(a))*, dipole *($\mu_x(a)$)* and quadrupole *($\mu_{xy}(a)$)* are defined according to the following relations:

$$q(a) = \int \rho_d(r)W_a(r)d^3r \tag{2.67}$$

where *q(a)* is the atomic charge.

$$\mu_x(a) = \int \rho_d(r)W_a(r)(x\text{-}x_a)d^3r \tag{2.68}$$

$$\mu_{xy}(a) = \rho_d(r)W_a(r)(x - x_a)(y\text{-}y_a)d^3r \tag{2.69}$$

that $W_a(r)$ is a fraction of the atomic density of atom *a* in the coordinate *r* and is shown by the following relationship:

$$W_a(r) = \rho_a(\text{r-}R_a)\left[\sum_a \rho_a(r - R_a)\right]^{-1} \tag{2.70}$$

2.4 References

[1] "Book Reviews." Journal of World Trade, 12(6), 564–566. 1978. doi:10.54648/trad1978050.

[2] von Bertalanffy, L. (1968). The Meaning of General System Theory. General Systems Theory, 30-53.

[3] Taylor, P., Mclntire, G. L., Dorsey, J. G., & Mclntire, G. L. (2006). Critical Reviews in Analytical Chemistry: Micelles in Analytical Chemistry Micelles in Analytical Chemistry, 37–41. doi:10.1016/S0223-5234(02)01384-3.

[4] Awad, M. K., Metwally, M. S., Soliman, S. A., El-Zomrawy, A. A., & Bedair, M. A. (2014). Experimental and quantum chemical studies of the effect of polyethylene glycol as corrosion inhibitors of aluminum surface. J. Ind. Eng. Chem., 20(3), 796–808. doi: 10.1016/j.jiec.2013.06.009.

[5] Perdew, J. P., Staroverov, V. N., Tao, J., & Scuseria, G. E. (2008). Density functional with full exact exchange, balanced nonlocality of correlation, and constraint satisfaction. Phys. Rev. A - At. Mol. Opt. Phys., 78(5), 052513-1- 052513-13. doi:10.1103/PhysRevA.78.052513.

[6] Schaefer, H. F. (1988). A history of ab initio computational quantum chemistry: 1950-1960. Tetrahedron Comput. Methodol., 1(2), 97–102. doi:10.1016/0898-5529(88)90014-0.

[7] Bader, R. F. W. (1998). A bond path: A universal indicator of bonded interactions. J. Phys. Chem. A, 102(37), 7314–7323. doi:10.1021/jp981794v.

[8] Ernzerhof, M., & Scuseria, G. E. (2000). Perspective on "inhomogeneous electron gas." Theor. Chem. Acc., 103(3-4), 259–262. doi:10.1007/s002149900030.

[9] Dong, S. S., Govoni, M., & Galli, G. (2021). Machine learning dielectric screening for the simulation of excited state properties of molecules and materials. Chem. Sci., 12(13), 4970–4980. doi:10.1039/d1sc00503k.

[10] Jung, J. Y., et al. (2006). Involvement of Bcl-2 family and caspases cascade in sodium fluoride-induced apoptosis of human gingival fibroblasts. Korean J. Physiol. Pharmacol., 10(5), 289–295. doi: 10.4196/kjpp.2006.10.5.289.

[11] Kaupp, M. (2001). Buchbesprechung: A Chemist's Guide to Density Functional Theory. Von Wolfram Koch und Max C. Holthausen. Angew. Chem., 113(5). doi: 10.1002/1521-3757(20010302)113:5<989: AID-ANGE989>3.0.CO;2-5.

[12] Ortmann, F., Bechstedt, F., & Schmidt, W. G. (2006). Semiempirical van der Waals correction to the density functional description of solids and molecular structures. Phys. Rev. B - Condens. Matter Mater. Phys., 73(20), 1–10. doi: 10.1103/PhysRevB.73.205101.

[13] Allouche, A. (2012). Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. J. Comput. Chem., 32, 174–182. doi:10.1002/jcc.

[14] Tkatchenko, A., & Scheffler, M. (2009). Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. Phys. Rev. Lett., 102(7), 6–9. [15] Allouche, A. (2012). Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. J. Comput. Chem., 32, 174–182. doi:10.1002/jcc.

[16] Delley, B. (1996). Fast calculation of electrostatics in crystals and large molecules. J. Phys. Chem., 100(15), 6107–6110. doi:10.1021/jp952713n.

[17] Zhang, X., et al. (2019). Temperature and size dependent surface energy of metallic nano-materials. J. Appl. Phys., 125(18), 1-7. doi:10.1103/PhysRevLett.102.073005.

[18] Grimme, S., Antony, J., Ehrlich, S., & Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys., 132(15), 3414-3420. doi:10.1063/1.3382344.

[19] Basiuk, V. A., & Henao-Holguin, L. V. (2014). Dispersion-corrected density functional theory calculations of meso-tetraphenylporphine-C60 complex by using DMol3 module. J. Comput. Theor. Nanosci., 11(7), 1609–1615. doi:10.1166/jctn.2014.3539.

[20] Bâldea, I. (2014). A quantum chemical study from a molecular transport perspective: Ionization and electron attachment energies for species often used to fabricate single-molecule junctions. Faraday Discuss., 174, 37–56. doi:10.1039/c4fd00101j.

[21] Ulusoy, I. S., & Wilson, A. K. (2019). Slater and Gaussian basis functions and computation of molecular integrals. In Mathematical physics in theoretical chemistry (pp. 31-61). Elsevier. doi:10.1016/B978-0-12-813651-5.00002-4.

[22] Delley, B. (1990). An all-electron numerical method for solving the local density functional for polyatomic molecules. J. Chem. Phys., 92(1), 508–517. doi: 10.1063/1.458452.

[23] Benedek, N. A., Snook, I. K., Latham, K., & Yarovsky, I. (2005). Application of numerical basis sets to hydrogen bonded systems: A density functional theory study. J. Chem. Phys., 122(14), 1-8. doi: 10.1063/1.1876152.

[24] Delley, B., & Delley, B. (2000). From molecules to solids with the DMol3 approach. 113(18), 7756-7764. doi: 10.1063/1.1316015.

[25] Allouche, A. (2012). Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. Journal of Computational Chemistry, 32, 174–182. doi:10.1002/jcc.2029.

[26] Sebens, C. T. (2021). Electron Charge Density: A Clue from Quantum Chemistry for Quantum Foundations. Foundations of Physics, 1–44. doi:10.1007/s10701-021-00480-7.

[27] Klamt, A. (1995). Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. The Journal of Chemical Physics, 102(5), 2224–2235. doi:10.1021/j100007a062.

[28] Baldridge, K., & Klamt, A. (1997). First principles implementation of solvent effects without outlying charge error. The Journal of Chemical Physics, 107(15), 6622. doi:10.1063/1.473662.

[29] Szczepanik, D. W., & Mrozek, J. (2013). Nucleophilicity Index Based on Atomic Natural Orbitals. Journal of Chemistry, 2013(1), 1-6. doi:10.1155/2013/684134.

[30] Segall, M. D., Shah, R., Pickard, C. J., & Payne, M. C. (1996). Population analysis of plane-wave electronic structure calculations of bulk materials. Physical Review B, 54(23), 16317. doi:10.1103/PhysRevB.54.16317.

[31] Guerra, L. I. A. F., Snijders, J. G., te Velde, G., & Baerends, E. J. (2003). Voronoi Deformation Density (VDD) Charges: Assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD. Journal of Computational Chemistry, 24(1), 131-146. doi:10.1002/jcc.10351.

[32] Hirshfeld, F. L. (1977). Bonded-Atom Fragments for Describing Molecular Charge Densities. Theoretica Chimica Acta, 138, 129–138. doi:10.1007/BF00549096.

[33] Saha, S., Roy, R. A. M. K., & Ayers, P. W. (2008). Population Analysis Schemes Consistent With Chemical Intuition? Journal of Computational Chemistry, 109(2009), 1790–1806. doi:10.1002/qua.

CHAPTER

III

PREDICTION OF THE PARTITION COEFFICIENT OF SMALL MOLECULES

Chapter III

Prediction of the partition coefficient of small molecules

3.1 Partition coefficients of small molecules

The partition coefficient is a measure of the distribution of a solute between two immiscible phases, typically a hydrophilic phase (such as water) and a hydrophobic phase (such as a lipid). In the case of small molecules, the partition coefficient is often used as a tool to predict how a molecule will behave in different biological systems.

As stated in the first chapter, the partition coefficient (P) is defined as the ratio of the concentration of a solute in the hydrophobic phase (C_H) to its concentration in the hydrophilic phase (C_L):

$$P = \frac{C_H}{C_L} \tag{3.1}$$

The value of the partition coefficient depends on several factors, including the nature of the solute, the nature of the two phases, and the temperature. In general, small molecules with a high partition coefficient are more likely to be hydrophobic and less likely to dissolve in water.

Since it may be used to predict the ADME (absorption, distribution, metabolism, and excretion) properties of potential drug candidates, the partition coefficient has a wide range of applications in the drug development process.

In general, the partition coefficient of small molecules is an important parameter that can provide valuable information about the behavior and properties of these molecules in different biological and environmental systems.

The literature was searched to obtain experimental data for log P $_{benzene}$, log P $_{cyclohexane}$, log P $_{hexane}$, log P $_{n\text{-}octane}$, log P $_{toluene}$, log P $_{carbon\ tetra\ chloride}$, log P $_{heptane}$, log P $_{1,2\text{-}dichloroethane}$, log P $_{octanol}$, and log P $_{1,1,1\text{-}trichloroethane}$ for 29 different compounds [1,2].

This research work presents Tables 1 and 2, which include the structures and chemical formulas of the 29 compounds and 10 solvents studied, respectively. Overall, the information presented in Tables 1 and 2 is important for understanding the study and interpreting the results. It provides a comprehensive overview of the molecular structures and chemical properties of the compounds and solvents studied, which can be used to analyze the behavior of the molecules in different systems.
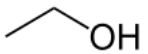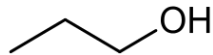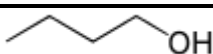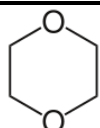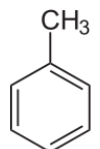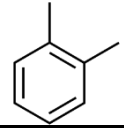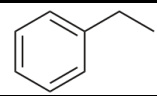
This work used density functional theory (DFT) and B3LYP functional with 6.31G(d), 6.31+G**, and 6.311++G** basis sets to calculate the partition coefficients of the 29 molecules in 10 different solvents and compare them with experimental values shown in Table 3. DFT is a widely used quantum chemistry method that can accurately predict the electronic structure and properties of molecules using the solvation model (SMD). Calculations were performed with the electronic structure program Gaussian 16 which provides a wide-ranging suite for the prediction of molecular properties.

In addition, harmonic vibrational frequencies were computed at 1 atm and 298K for each compound. These frequencies describe how the atoms in the molecule move to each other, and they are an important factor in calculating thermodynamic properties such as free energies.

Equations (1.1) and (1.2) in Chapter 1 were used to predict the partition coefficient of different compounds that are classified into three groups: alcohols, ethers, and hydrocarbons. Table 3, shows the solutes and organic solvents and the partition coefficient estimated employing the 6.31G(d), 6.31+G**, and 6.311++g** bases sets.
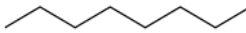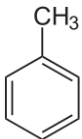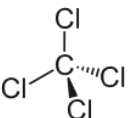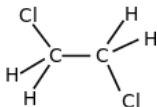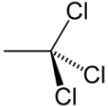
Table 1. Molecules and their structures.

Compound	Formula	Structure
Ethanol	C_2H_6O	
n-Propanol	C_3H_8O	
I-Propanol	$(CH_3)_2CHOH$	
N-Butanol	$C_4H_{10}O$	
Tetrahydrofuran	C_4H_8O	
Dioxane	$C_4H_8O_2$	
Toluene	C_7H_8	
Benzene	C_6H_6	
Xylenes	$(CH_3)_2C_6H_4$	
Ethylbenzene	C_8H_{10}	

Table 2. Structure, chemical formula, and dielectric constant of solvents. Values of dielectric constants are taken from the indicated references.

Solvents	Formula	Structure	Dielectric constant	Reference
Benzene	C_6H_6		2.25	[4]
Cyclohexane	C_6H_{12}		2.00	[4]
Hexane	C_6H_{14}		1.88	[4]
N-octane	$CH_3(CH_2)_6CH_3$		1.94	[5]
Toluene	C_7H_8		2.33	[4]
Carbon tetra chloride	CCl_4		1.74	[6]
Heptane	C_7H_{16}		1.92	[4]
1,2-Dichloroethane	$C_2H_4Cl_2$		10.36	[7]
N-octanol	$C_8H_{18}O$		10.30	[7]
1,1,1-Trichloroethane	CH_3CCl_3		7.33	[6]

Table 3. List of compounds' estimated partition coefficients using the DFT method and B3LYP functional with SMD solvent, calculated using the 6.31G(d), 6.31+G**, and 6.31++G** basis sets. Molecules' experimentally partition coefficients in organic solvents obtain from scientific publications.

Compound	Solvent	Log P Calculation 6.31G(d)	Log P Calculation 6.31+G**	Log P Calculation 6.311++G**	Log P Experimental
Ethanol	Benzene	1.25	1.59	-	0.90
Ethanol	Cyclohexane	1.71	2.08	2.07	1.36
Ethanol	Hexane	1.58	1.96	1.95	1.58
Ethanol	n-Octane	1.65	2.02	2.01	0.80
Ethanol	Toluene	1.29	1.61	1.60	1.06
N-propanol	Benzene	0.59	0.97	0.96	0.11
N-propanol	Cyclohexane	1.09	1.51	1.50	0.81
N-propanol	Hexane	0.92	1.35	1.34	-0.13
N-propanol	n-Octane	1.01	1.44	1.43	0.07
N-propanol	Toluene	0.64	1.01	1.00	0.02
I-propanol	Benzene	0.64	1.03	1.01	-0.06
I-propanol	Cyclohexane	1.15	1.57	1.56	0.90
I-propanol	Hexane	0.99	1.43	1.41	0
I-propanol	n-Octane	1.07	1.50	1.49	0.40
I-propanol	Toluene	0.69	1.07	1.05	0.06
N-Butanol	Benzene	-0.016	0.17	-	-0.72
N-Butanol	Cyclohexane	0.45	0.81	0.81	0.25
N-Butanol	Hexane	0.24	0.62	0.61	-1.39
N-Butanol	n-Octane	0.35	0.72	0.72	0.008
N-Butanol	Toluene	-0.08	0.24	0.23	-0.76
THF	Benzene	-0.38	-0.80	-0.88	-1.22
THF	Carbon tetrachloride	-0.40	-0.80	-0.87	-1.22
THF	Heptane	-0.03	-0.35	-0.43	-1.09
THF	Hexane	-0.08	-0.40	-0.48	-0.79
THF	Trichloro ethane	-0.97	-1.53	-1.58	-1.30
Toluene	Heptane	-3.40	-3.18	-3.19	-2.52
Toluene	Octanol	-2.72	-2.67	-2.67	-2.5
Xylenes	Heptane	-3.39	-3.05	-3.10	<-3.0
Xylenes	Octanol	-2.54	-2.39	-2.43	<-3.0
Benzene	Heptane	-2.90	-2.72	-2.73	<-3.0
Benzene	Octanol	-2.25	-2.21	-2.21	-2.13
Ethylbenzene	Heptane	-3.69	-3.43	-3.44	<-3.0
Ethylbenzene	Octanol	-3.08	-3.01	-3.01	-3.15

3.1.1 Alcohols

Alcohols are a class of organic compounds that contain a hydroxyl (-OH) functional group attached to a carbon atom. They are characterized by the presence of the hydroxyl group, which imparts unique chemical and physical properties to the molecule. Alcohols can be classified based on the number of carbon atoms attached to the carbon atom with the hydroxyl group. Primary alcohols have one carbon atom attached to the hydroxyl carbon, secondary alcohols have two, and tertiary alcohols have three. One of the most well-known and widely used alcohol molecules is ethanol, which is commonly found in alcoholic beverages. Ethanol has a wide range of applications in organic chemistry, alcohols are important reagents and intermediates in many chemical reactions. They can be used as reducing agents, in nucleophilic substitution reactions, and oxidation reactions. For example, primary alcohols can be oxidized to aldehydes and then to carboxylic acids, while secondary alcohols can be oxidized to ketones. In industry and research, including as a fuel additive, solvent, and disinfectant. Alcohols can also be converted into other functional groups, such as halides, esters, and ethers, through various chemical reactions. This makes them useful building blocks for the synthesis of complex organic molecules.

3.1.1.1 Ethanol

Ethanol, also known as ethyl alcohol or grain alcohol, is a clear, colorless liquid with a slightly sweet odor. It is a primary alcohol with the chemical formula C_2H_5OH, and it is one of the most widely used alcohols in the world. It is miscible in water, acetone, and other organic solvents, and it has a wide range of applications in industry, healthcare, and household settings.

The physical properties of ethanol are characterized by a boiling point of 78.5 °C, a melting point of -114.1 °C, a density of 0.789 g/mL, a molecular weight of 46.1 g/mol, and a flashpoint of 13 °C.

Ethanol has many applications in industry, including:

- As a solvent for coatings, inks, and adhesives.
- In the production of cosmetics, fragrances, and personal care products.
- In the manufacture of pharmaceuticals and medical devices.

- As a fuel additive and alternative fuel source.
- In the production of alcoholic beverages, such as beer, wine, and spirits.

3.1.1.2 N-propanol

1-Propanol, also known as n-propanol or propyl alcohol, is a colorless, flammable liquid with a mild odor. It is a primary alcohol with the chemical formula C_3H_8O, and it is one of the simplest alcohols. It is soluble in water, ethanol, and ether, and it is commonly used as a solvent, in the production of other chemicals, and as a fuel.

1-propanol is characterized by certain properties, such as a boiling point of 97.2 °C, a melting point of -126.2 °C, a density of 0.804 g/mL, a molecular weight of 60.1 g/mol, and a flashpoint of 22.2 °C.

1-Propanol has a variety of uses in industry, including:

- As a solvent for resins, gums, and other organic compounds.
- As a component of printing inks, coatings, and adhesives.
- In the production of glycerol and other chemicals.
- As a fuel additive.

It is also used in the pharmaceutical industry as a solvent for drugs and in cosmetics as a fragrance ingredient and solvent.

In healthcare, ethanol is commonly used as an antiseptic and disinfectant, as it can kill bacteria, viruses, and fungi. It is also used as a preservative for biological specimens and as a solvent for medications.

3.1.1.3 I-propanol

2-Propanol, also known as isopropyl alcohol or IPA, is a colorless, flammable liquid with a strong odor. It is a secondary alcohol with the chemical formula C_3H_8O, and it is one of the most commonly used solvents worldwide. It is miscible in water, ethanol, and other organic solvents, and it has a wide range of applications in industry, healthcare, and household settings.

The unique features of 2-propanol include a boiling point of 82.6 °C, a melting point of -89.5 °C, a density of 0.786 g/mL, a molecular weight of 60.1 g/mol, and a flashpoint of 11.7 °C. These characteristics play a role in its various applications.

2-Propanol has many applications in industry, including:

- As a solvent for coatings, inks, and adhesives.
- In the production of cosmetics, fragrances, and personal care products.
- In the manufacture of pharmaceuticals and medical devices.
- As a cleaning agent in electronic and semiconductor manufacturing.
- As a fuel additive and extraction solvent in biofuel production.

3.1.1.4 N-butanol

N-Butanol, also known as n-butyl alcohol or 1-butanol, is a colorless, flammable liquid with a mild odor. It is a primary alcohol with the chemical formula $C_4H_{10}O$, and it is one of the four isomers of butanol. It is miscible in water, ethanol, and other organic solvents, and it has a wide range of applications in industry, healthcare, and household settings.

Some properties that define n-butanol include boiling at 117.7 °C, melting at -89.8 °C, having a density of 0.81 g/mL, a molecular weight of 74.1 g/mol, and a flash point of 35 °C.

N-Butanol has many applications in industry, including:

- As a solvent for coatings, inks, and adhesives.
- In the production of plastics, textiles, and synthetic resins.
- As a chemical intermediate for the manufacture of other chemicals.
- As a fuel additive and extraction solvent in biofuel production.

In healthcare, n-butanol is used as a solvent and a preservative for biological specimens. It is also used in the pharmaceutical industry as a reagent in chemical synthesis.

3.1.1.5   Prediction partition coefficient of log P alcohol

Partition coefficients of alcohol compounds (ethanol, n-Propanol, i-Propanol, and n-Butanol) were estimated in three different basis sets, as shown in Table 3.

The DFT method with the B3LYP functional and various basis sets, such as 6.31G(d), 6.311+G*, and 6.311++G**, were used to determine the partition coefficient of ***ethanol*** in five solvents including Benzene, Cyclohexane, Hexane, Octane, and Toluene. Based on the findings presented in Table 3, the partition coefficients derived from the low basis set (6.31G(d)) show a satisfactory correlation with the experimental partition coefficient. However, there is a great difference between the log P $_{Calculation}$ in Octane solvent and the log P $_{Experimental}$ values.

With respect to ***n-propanol*** solute, the calculated data were compared with experimental results, and it was found that 6.31G(d) shows a reasonable correlation with log P $_{Experimental}$. The partition coefficient (log P) of the n-propanol compound in Hexane does not display a good correlation when compared with experimental data. The utilization of a low basis set (6.31G(d)) to predict the partition coefficient of ***i-Propanol*** in the solvents identified in Table 3 results in a stronger correlation with the experimental data compared to other basis sets.

The partition coefficient of ***n-Butanol*** in various organic solvents was calculated using different basis sets (6.31G(d), 6.311+G**, and 6.311++G**). Table 3 displays the results, revealing that the use of a low basis set to compute log P makes a better correlation with the experimental values when compared to the partition coefficient computed using other basis sets.

In general, the partition coefficient of ***alcohol*** solutes in various solvents using the 6.31G(d) basis set indicates a satisfactory correlation with experimental values.

In Table 4, one can find the equations and their corresponding coefficient of determination (R^2) and mean absolute error (MAE) for the alcohol compounds that were investigated. The data indicates that the 6.31G(d) basis set is more closely correlated with the experimental value than other sets.

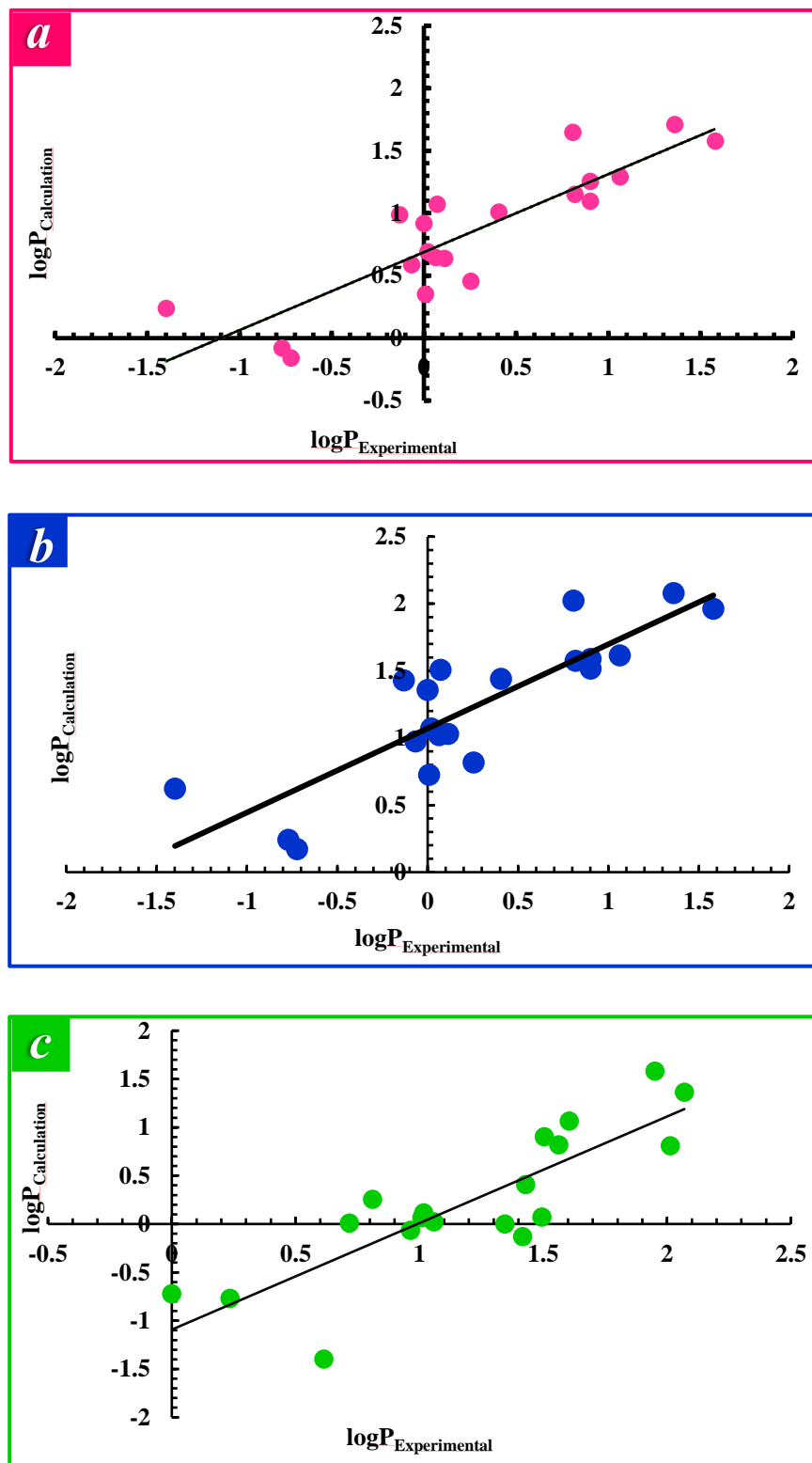**Figure 1.** Comparison of the experimental log P $_{solvent/water}$ with respect to the calculated log P $_{solvent/water}$ for Alcohol compounds for different solvent/water combinations using DFT method and B3LYP functional with a) 6.31G(d) b) 6.311+G** c) 6.311++G** basis sets.

Table 4. Linear regression, R^2 and MAE for Alcohol compounds with DFT method and B3LYP functional and 6.31G(d), 6.311+G** and 6.311++G** basis sets.

Basis set	6.31G(d)	6.311+G**	6.311++G**
Equation	Y=0.62x+0.68	Y=.62x+1.07	Y=1.10x+1.09
R^2	0.75	0.72	0.71
MAE	0.69	1.04	1.04

3.1.2 Ethers

The history of ether molecules is closely tied to the history of organic chemistry. The first ether compound to be synthesized was diethyl ether, which was synthesized by the French chemist Valerius Cordus in 1540 by distilling ethanol with sulfuric acid.

However, it was not until the 19th century that the importance of ether molecules was fully recognized. In 1846, the American dentist William Morton used diethyl ether as a general anesthetic for the first time during a surgical procedure. This marked the beginning of the use of ether as a surgical anesthetic, which revolutionized medicine and made many surgical procedures possible. The study of organic chemistry was fundamentally altered by Friedrich August Kekulé's novel theory of chemical structure around the turn of the 20th century. Kekulé postulated that chains of carbon atoms with attached functional groups, such as ethers, make up organic molecules and that carbon atoms may create numerous bonds with one another.

At that time, research on ether molecules remained a crucial component of organic chemistry. The discovery of novel techniques for the synthesis and cleavage of ether molecules has resulted in the production of new materials and medications. Ethers are used as solvents in several industrial processes.

Ethers are a type of compound in organic chemistry that have an ether group, which is an oxygen atom bonded to two alkyl or aryl groups. They are described by the generic formula R-O-R′, where R and R′ stand for the alkyl or aryl groups.

Ethers may also be classified into two groups: simple or symmetrical ethers and mixed or unsymmetrical ethers. Simple ethers are defined as having the same alkyl or aryl group on both sides of the oxygen atom. The alpha hydrogens of ethers are more acidic than those of simple

hydrocarbons because oxygen is more electronegative than carbon. The alpha hydrogens of carbonyl groups (seen in ketones or aldehydes, for example) are far more acidic than these, though.

One of the unique properties of ether molecules is their ability to act as solvents. Due to the polar nature of the oxygen atom, ethers can dissolve a wide variety of polar and nonpolar substances. For this reason, ethers are commonly used as solvents in laboratory procedures and industrial processes. Ethers also have low boiling points, making them useful as refrigerants and as a starting material for the production of other chemicals. Additionally, ether molecules are relatively stable and unreactive, making them useful as protecting groups in organic synthesis.

Ethers can also undergo a process called ether cleavage, in which the oxygen atom is cleaved from the molecule to form two alkyl or aryl radicals. This reaction is important in the synthesis of complex organic molecules.

Overall, ether molecules are a versatile class of compounds with a range of applications in both laboratory and industrial settings. Their unique properties make them valuable in various fields, including chemistry, pharmaceuticals, and materials science.

Ethers may be produced in various methods. Aryl ethers often require metal catalysts, while alkyl ethers typically form more easily. Since the 13th century, it has been known that ethanol and sulfuric acid may be used to synthesize diethyl ether.

3.1.2.1 THF

This work uses Tetrahydrofuran (THF), which is a cyclic ether molecule with the chemical formula C_4H_8O. It is a clear, colorless, and highly flammable liquid with a characteristic odor. THF is an important industrial solvent that is widely used in organic chemistry reactions. One of the most important properties of THF is its ability to dissolve many organic compounds, including polar and nonpolar compounds. This makes it a versatile solvent that can be used in a wide range of chemical reactions.

THF is commonly used in Grignard reactions, which are important in the synthesis of organic compounds. It is also used as a solvent in polymer chemistry, particularly in the synthesis of polyvinyl chloride (PVC). THF can also act as a ligand in coordination chemistry. In this role,

THF can form complexes with metal ions, which can be useful in the synthesis of metal-organic compounds and catalysts.

However, THF can also pose some health and safety risks. It is highly flammable and can react with air to form explosive peroxides. THF can also be toxic if ingested or inhaled in large quantities.

3.1.2.2 Prediction of log P THF

In this work, estimate the partition coefficient of **THF** in Carbon tetrachloride, Heptane, Hexane, and Trichloroethane solvents using three different basis sets (6.31G(d), 6.311+G**, and 6.311++G**). Results presented in Table 5 show R^2=0.70 and mean absolute error =0.35, indicating that the calculation of log P $_{Solv/Wat}$ using a higher basis set demonstrates a suitable correlation with log P $_{Experimental}$.
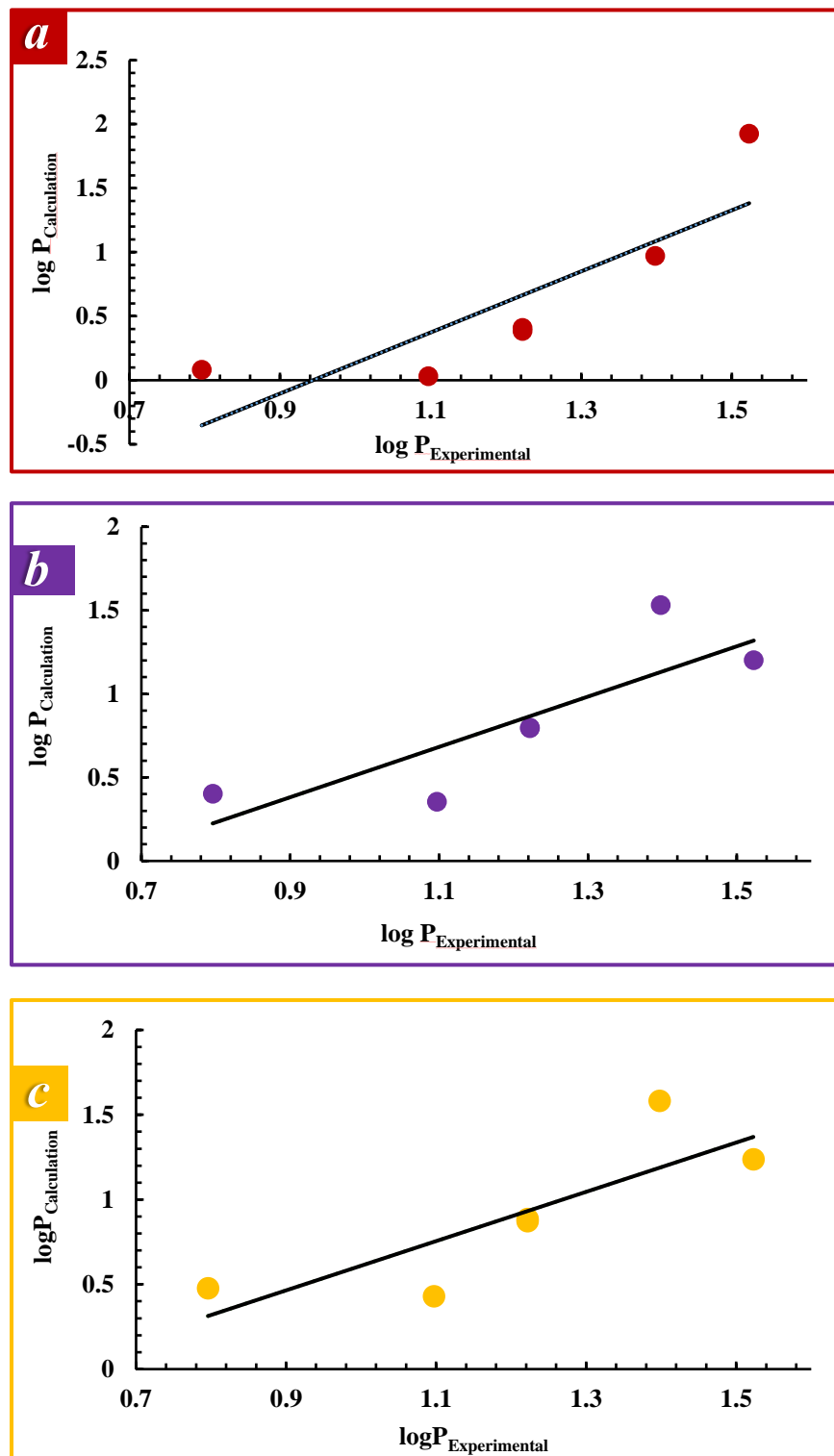
Figure 2. Comparison of the experimental log P $_{solvent/water}$ with respect to the calculated log P $_{solvent/water}$ for the THF ether compound for different solvent/water combinations using DFT method and B3LYP functional with a) 6.31G(d) b) 6.311+G** c) 6.311++G** basis sets.

Basis set	6.31G(d)	6.311+G**	6.311++G**
Equation	Y=2.39x-2.25	Y=1.50-0.97	Y=1.45x-0.84
R^2	0.70	0.69	0.69
MAE	0.71	0.40	0.35

Table 5. Linear regression, R^2 and MAE for Ether compound with DFT method and B3LYP functional and 6.31G(d), 6.311+G** and 6.311++G** basis sets.

3.1.3 Hydrocarbons

Hydrocarbons are organic compounds that contain only carbon and hydrogen atoms. They are the primary components of fossil fuels such as coal, oil, and natural gas, which have been used as sources of energy for centuries. Hydrocarbons can also be found in many other sources, including plants and animals, and they play a vital role Hydrocarbons can be categorized into different types, such as alkanes, alkenes, and alkynes, based on their chemical structure in the functioning of many natural systems.

Hydrocarbons can also play a significant role in:

•   Drug delivery using hydrocarbons: Drugs can be delivered to specific bodily parts using hydrocarbons as vehicles. Using liposomes, which are small vesicles formed of phospholipid molecules (which include hydrocarbon chains) that may carry medicines and target them to target organs, is one example of this.

•   Hydrocarbons in pharmaceutical synthesis: Many pharmaceuticals are synthesized using hydrocarbons, including alkanes, alkenes, and alkynes. For example, many antibiotics are made using aromatic hydrocarbons such as benzene.

•   Hydrocarbon-based drug discovery: Some natural hydrocarbons, such as terpenes, have been found to have medicinal properties and are being investigated as potential sources of new drugs. Additionally, synthetic hydrocarbons can be designed and tested for specific pharmacological activities.

Hydrocarbon-based biomaterials: Hydrocarbons can also be used to create biomaterials that are used in medical devices and implants. For example, polyethylene is a hydrocarbon-based material commonly used in orthopedic implants.

3.1.3.1 Toluene

Toluene, also known as methylbenzene or phenyl methane, is a clear, colorless liquid with a sweet, pungent odor. It is an aromatic hydrocarbon with the chemical formula C_7H_8, and it is commonly used as a solvent in industry, healthcare, and household settings. It is insoluble in water but is miscible in many organic solvents.

Toluene possesses certain characteristics, such as a boiling point of 110.6 °C, a melting point of -95 °C, a density of 0.87 g/mL, a molecular weight of 92.1 g/mol, and a flash point of 4.4 °C.

Toluene has many applications in industry, including:

- As a solvent for coatings, adhesives, and polymers.
- In the production of chemicals, such as benzene, phenol, and TNT.
- As a fuel additive and a solvent for oil refining.
- In the printing and leather industries.

In healthcare, toluene is used as a solvent for medications and as a preservative for biological specimens. It is also used as a component in medical imaging agents.

3.1.3.2 Xylenes

Xylenes are a group of isomeric hydrocarbons that have the molecular formula C_8H_{10}. They are colorless, flammable liquids with a sweet odor.

Some of the properties of toluene include:

•   The boiling point of xylenes varies depending on the isomer. Here are the boiling points of the three isomers of xylenes at atmospheric pressure:

•   Ortho-xylene: 144.4 °C (291.9 °F).

•   Meta-xylene: 139.1 °C (282.4 °F).

•   Para-xylene: 138.4 °C (281.1 °F).

Xylenes are liquids at room temperature and do not have a well-defined melting point. Instead, they solidify into a glassy or crystalline state at low temperatures. The temperature at which this occurs depends on the isomer and the purity of the sample.

For example, the glass transition temperature of meta-xylene is reported to be around -47 °C (-53 °F), while the melting point of a high-purity para-xylene crystal is around 13.2 °C (55.8 °F). However, it is important to note that the solidification behavior of xylenes can be complex and depends on factors such as the cooling rate, impurities, and the presence of other compounds.

The density of xylenes varies depending on the isomer and temperature. Here are the approximate densities of the three isomers of xylenes at 20 °C (68 °F):

- Ortho-xylene: 0.88 g/cm^3.
- Meta-xylene: 0.86 g/cm^3.
- Para-xylene: 0.86 g/cm^3.

As you can see, the density of ortho-xylene is slightly higher than that of the other two isomers. The density of xylenes is an important property that affects their behavior in various industrial applications, including their solubility and vapor pressure.

The molecular weight of xylenes varies depending on the isomer. Here are the molecular weights of the three isomers of xylenes:

- Ortho-xylene: 106.17 g/mol.
- Meta-xylene: 106.17 g/mol.
- Para-xylene: 106.17 g/mol.

The flash point of xylenes varies depending on the isomer and the method used to measure it. Here are the approximate flash points of the three isomers of xylenes according to different sources:

- Ortho-xylene: 25-28 °C (77-82 °F).
- Meta-xylene: 24-26 °C (75-79 °F).
- Para-xylene: 17-18 °C (63-64 °F).

The flash point is the lowest temperature at which a liquid gives off enough vapor to ignite when exposed to an ignition source such as a flame or spark. The flash point is an important safety parameter for handling and storing xylenes, as they are highly flammable and can ignite easily at room temperature if exposed to a source of ignition.

xylenes are used in a wide range of industrial applications, including:

- Production of plastics: Xylenes are used as a feedstock in the production of various types of plastics, including polyethylene terephthalate (PET) and polyvinyl chloride (PVC). These plastics are used in a wide range of applications, including packaging, construction materials, and textiles.
- Manufacturing of synthetic fibers: Xylenes are used in the production of synthetic fibers such as polyester and nylon. These fibers are used in the manufacture of clothing, carpets, and other textiles.
- Manufacturing of synthetic fibers: Xylenes are used in the production of synthetic fibers such as polyester and nylon. These fibers are used in the manufacture of clothing, carpets, and other textiles.
- Manufacturing of synthetic fibers: Xylenes are used in the production of synthetic fibers such as polyester and nylon. These fibers are used in the manufacture of clothing, carpets, and other textiles.

3.1.3.3 Benzene

Benzene is a highly important organic chemical compound with the chemical formula C_6H_6. It is a colorless and highly flammable liquid that has a sweet and aromatic smell. Benzene is a cyclic hydrocarbon and is classified as an aromatic compound because of its unique molecular structure.

Toluene is characterized by a molecular weight of 78.11 g/mol, a density of 0.88 g/cm³ at 20°C, a melting point of 5.5°C, a boiling point of 80.1°C, and a flashpoint of -11°C.

Benzenes are used in a wide range of industrial applications, including:

- Production of plastics: Benzene is used as a feedstock to produce various plastics, such as polystyrene, polyurethane, and nylon.

- Production of synthetic fibers: Benzene is used in the production of synthetic fibers, including nylon and polyester.

- Production of rubber: Benzene is used to produce synthetic rubber, which is used in the manufacture of tires, hoses, and other rubber products.

- Production of dyes and pigments: Benzene is used as a solvent in the production of dyes and pigments.

- Production of pharmaceuticals: Benzene is used as a starting material in the synthesis of various pharmaceuticals, including antibiotics, antihistamines, and pain relievers.

- Fuel production: Benzene is blended with gasoline to increase its octane rating, which improves engine performance.

- Extraction of oils: Benzene is used as a solvent to extract oils from seeds, such as soybeans and corn.

3.1.3.4 *Ethylbenzene*

Ethylbenzene is an organic compound with the chemical formula C_8H_{10}. It is a colorless liquid that is commonly used as a feedstock in the production of styrene, which is used to make various plastics and synthetic materials.

Ethylbenzene displays several features, such as a molecular weight of 106.17 g/mol, a melting point of -95.3°C, a boiling point of 136°C, a density of 0.867 g/mL at 20°C, and a flash point of 26°C.

Ethylbenzene is used in a wide range of industrial applications, including:

- Production of styrene: Ethylbenzene is used as a feedstock in the production of styrene, which is used to make various plastics and synthetic materials such as polystyrene, ABS, and styrene-butadiene rubber.

- Solvent: Ethylbenzene is used as a solvent for various industrial applications, such as in the production of coatings, resins, adhesives, and inks.

▪ Fuel additive: Ethylbenzene is added to gasoline as an octane booster to increase its combustion efficiency and reduce engine knock.

Production of phenyl ethylene: Ethylbenzene can be used to produce phenyl ethylene, which is used as a fragrance and flavoring agent.

▪ Production of chemicals: Ethylbenzene can be used to produce various chemicals such as ethylbenzene diisocyanate, which is used in the production of polyurethane foam.

3.1.3.5 Prediction of log P hydrocarbons

The partition coefficient of ***Toluene*** was evaluated in Heptane and n-Octanol using low, medium, and high basis sets. Table 3 shows that the calculated partition coefficient for toluene indicates a stronger correlation with experimental data when using a medium or high base set, compared to a low set.

In addition, Ortho ***Xylene*** was computed for its partition coefficient and compared to the experimental value. It was found that the log P calculation in Octanol using the 6.31G(d) basis set resulted in better agreement with experimental data than other basis sets. However, the use of three different basis sets in log P $_{Calculation}$ for Xylenes in Heptane demonstrates a strong correlation in comparing experimental partition coefficients.

Furthermore*, **Benzene*** solute's partition coefficients were calculated using low, medium, and high basis sets in Heptane and Octanol solvents, resulting in a satisfactory agreement with the experimental log P values.

Also, Table 3 displays the computed partition coefficient of ***Ethylbenzene*** in two solvents, Octanol and Heptane, using various basis sets. The results indicate that all basis sets used in the calculation are appropriate for estimating the partition coefficient and comparing it with the experimental data.

Unfortunately, it was not possible to analyze a plot for hydrocarbon compounds as no precise experimental data was available for these partition coefficients. Despite this limitation, our findings based on log P calculations using all three basis sets suggest a reasonable correlation with the experimental values.

3.2 References

[1] Smallwood, I. M., & Smallwood, I. M. (1966). Handbook of Organic Solvent Properties. doi:10.1016/C2009-0-23646-4.

[2] Niederlehner, B. R., Cairns, J., & Smith, E. P. (1998). Modeling Acute and Chronic Toxicity of Nonpolar Narcotic Chemicals and Mixtures to Ceriodaphnia dubia. Journal of Ecotoxicology and Environmental Safety, 39(146), 136–146. doi:10.1006/eesa.1997.1621.

[3] Nath, J., & Tripathi, A. D. (1984). Binary systems of 1,1,2,2-tetrachloroethane with benzene, toluene, p-xylene, acetone, and cyclohexane. 1. Excess volumes, ultrasonic velocities, and adiabatic compressibility's at 298.15 and 308.15 K. Journal of Chemical Engineering, 28, 1517–1524. doi.org/10.1021/je00032a038.

[4] Horne, S., & Murthy, K. (2005). Full Papers Control of Product Quality in Batch Crystallization of Pharmaceuticals and Fine Chemicals. Part 1: Design of the Crystallization Process and the Effect of solvent. Journal of Organic Process Research and Development, 9, 858−872. doi.org/10.1021/op050049v.

[5] Laurence, C., Nicolet, P., Dalati, M. T., Abboud, J. M., & Notario, R. (1994). The Empirical Treatment of Solvent-Solute Interactions: 15 Years of pi. Journal of Physical Chemistry, 98, 5807–5816. doi.org/10.1021/j100074a003.

[6] Kocsis, L. S., Kagalwala, H. N., Mutto, S., Godugu, B., Bernhard, S., Tantillo, D. J., & Brummond, K. M. (2015). Mechanistic Insight into the Dehydro-Diels-Alder Reaction of Styrene-ynes. Journal of Organic Chemistry, 80, 11686-11698. doi: 10.1021/acs.joc.5b00200.

[7] Zhuravlev, V. I., & Usacheva, T. M. (2020). Dielectric Properties of the Liquid Phase of Higher Alkanols on the Line of Vapor – Liquid Equilibrium, 94, 687–692. doi: 10.1134/S0036024420040299.

CHAPTER
IV

PREDICTION OF PARTITION COEFFICIENT IN MICELLE SYSTEMS

Chapter IV

Prediction of partition coefficient in micelle systems

4.1   Properties and applications of micelles

The concept of micelles has gained significant attention in the field of chemistry and pharmaceutical sciences. A micelle is a little grouping or cluster of surface-active agent molecules, also called surfactants, present in a liquid solution. Surfactants are compounds that have both hydrophilic (water-loving) and hydrophobic (water-repellent) properties. When surfactant molecules are added to a liquid, such as water, they tend to be made into micelles to minimize their exposure to the surrounding solvent [1].

The formation of micelles occurs due to the amphiphilic nature of surfactant molecules. In a liquid solution, surfactant molecules orient themselves in a way that the hydrophobic "tails" of the molecules group together, shielded from the surrounding water, while the hydrophilic "heads" remain exposed to the water [2,3]. This arrangement allows the micelle to have a hydrophilic outer shell and a hydrophobic core. Micelles are typically spherical; this structure is energetically favorable because it reduces the exposure of the hydrophobic portions of the surfactant molecules to water [4].

Typically, micelles display a spherical shape, which is advantageous from an energy perspective as it minimizes the interaction between the hydrophobic parts of the surfactant molecules and the surrounding environment [5].

4.1.1   Important Factors in micelle formation
Several factors play a crucial role in micelles' formation, stability, and properties. Some of the important factors to consider in the context of micelles are:

- **Amphiphilic Molecules***: Micelles are formed by amphiphilic molecules, which have both hydrophilic (water-loving) and hydrophobic (water-repellent) regions. The balance between these two regions is essential for the formation and stability of micelles. The structure, size, and shape of the amphiphilic molecules can greatly influence the properties of the micelles formed [6,7].

- **Critical Micelle Concentration (CMC):** The CMC is the minimum concentration of amphiphilic molecules in a solution at which micelle formation occurs. It is an important parameter that determines the thermodynamics of micellization. Below the CMC, individual amphiphilic molecules are dispersed in the solvent, while above the CMC, micelles start to form [8,9].

- **Solvent Properties:** The nature and properties of the solvent in which micelles form can significantly impact micelle formation and stability. Solvent polarity, temperature, and ionic strength can influence the interactions between amphiphilic molecules and affect micelle size, shape, and aggregation behavior [10].

- **Temperature and pH:** Temperature and pH play a role in the creation and durability of micelles. Changes in temperature can affect how hydrophobic molecules within the micelle core interact, resulting in alterations to the micelles' size and structure. Fluctuations in pH can impact the charge of the micelles' hydrophilic sections, consequently influencing their stability and properties [11].

- **Solubilization and Encapsulation:** Solubilization and encapsulation are significant functions of micelles, as they can effectively dissolve hydrophobic substances such as drugs within their hydrophobic core. This capability to encapsulate and transport hydrophobic molecules is crucial for utilizing micelles in drug delivery systems. When considering micelles for this purpose, important factors to consider are the efficiency of encapsulation and the kinetics of substance release from the encapsulated state [12,13].

- **Size and Shape:** The stability, capacity to load drugs, and interaction with biological systems of micelles can be influenced by their size and shape. Micelles can exist in various forms, such as small spherical structures, elongated shapes, or branched configurations, depending on the specific amphiphilic molecule and the surrounding environment. Controlling the size of micelles is essential for drug delivery purposes, as it ensures optimal distribution within the body and efficient uptake by cells [14,15].

4.1.2   Applications in drug delivery systems

Micelles have emerged as promising vehicles for drug delivery due to their unique properties and versatility. Some key applications of micelles in drug delivery systems include:

- **Solubilization of Hydrophobic Drugs:** Micelles can solubilize hydrophobic drugs within their hydrophobic core, improving their solubility and bioavailability. By encapsulating the drug molecules, micelles enhance their stability and protect them from degradation, facilitating their delivery to the target site [16].

- **Enhanced Permeability and Retention (EPR) Effect***:* Micelles can take advantage of the EPR effect, which refers to the passive accumulation of drug-loaded micelles in tumor tissues. The leaky vasculature and impaired lymphatic drainage in tumors allow for the selective accumulation of micelles, leading to improved drug delivery to cancer cells while minimizing systemic toxicity [17,18].

- **Controlled Drug Release:** Micelles can be engineered to achieve controlled drug release profiles. By manipulating the composition and structure of the micelles, drug release can be modulated based on factors such as pH, temperature, or the presence of specific enzymes or stimuli. This enables sustained and targeted delivery of drugs, improving therapeutic efficacy and reducing side effects [19,14].

- **Theranostic Applications:** Micelles can be integrated with imaging agents, such as fluorescent dyes or contrast agents, to enable both drug delivery and real-time monitoring of treatment efficacy. These theranostic micelles offer the potential for personalized medicine, enabling simultaneous diagnosis and therapy [21].

- **Overcoming Multidrug Resistance:** Multidrug resistance (MDR) is a significant challenge in chemotherapy. Micelles can be designed to address MDR by incorporating strategies such as the co-delivery of drug and MDR modulators, promoting drug efflux inhibition, or utilizing stimuli-responsive systems to overcome drug resistance mechanisms [22].

- **Localized Drug Delivery:** Micelles can be utilized for localized drug delivery in specific anatomical sites or tissues. For example, micelles can be designed to target ocular, intranasal, or topical drug delivery, improving drug bioavailability and minimizing systemic side effects.

Micellar drug delivery systems offer tremendous potential for improving the efficacy and safety of medicinal interventions. Ongoing research and advancements in micelle design and formulation are expected to further grow their applications and effect in the field of drug delivery [23].

4.2   The relationship between the partition coefficient and micelles

The partition coefficient, also known as the distribution coefficient, is a measure of the relative solubility of a compound in two immiscible phases, typically a hydrophobic solvent and water. It is defined as the ratio of the concentration of a compound in one phase to its concentration in the other phase at equilibrium and it can calculate with equations (1.1) and (1.2).

When it comes to micelles, the partition coefficient can play a role in the solubilization and transportation of hydrophobic compounds. Micelles can encapsulate hydrophobic molecules within their hydrophobic core, allowing them to become more soluble in water and be transported through an aqueous environment.

The presence of micelles can alter the partition coefficient of a compound between water and a hydrophobic solvent. When a compound is solubilized in a micelle, its effective concentration in the water phase is increased, which affects its partitioning behavior. The compound is essentially "trapped" within the micelle, and its distribution between the micelle and the aqueous phase determines its overall partition coefficient.

The partition coefficient of a compound in the presence of micelles depends on factors such as the size and structure of the micelles, the nature of the hydrophobic compound, and the concentration of surfactant molecules. Generally, the partition coefficient of a compound in the presence of micelles can be higher compared to its partition coefficient between water and the hydrophobic solvent alone.

Overall, the relationship between the partition coefficient and micelles is complex and depends on several factors, but micelles can influence the solubility and distribution of hydrophobic compounds, providing a valuable tool for various applications.

One determining factor in this kind of application is the partition coefficient of the potential drug. Thus, the Log P parameter indicates the capacity of a potential drug to aggregate with the micelles

as it is related to the thermodynamic equilibrium of the distribution of the compound between the water and the system of micelles. In aqueous solutions, micelles consist of a hydrophobic core and a hydrophilic external layer, which is created by the surfactant head groups, directly in contact with the surrounding aqueous phase. Thus, because of the existing polarity gradient, aqueous–micellar solutions can solubilize polar and nonpolar materials [24,25].

4.2.1 Experimental methods to calculate log P

Experimental methods can be employed to obtain the partition coefficient of molecules in a micellar system. The choice of method depends on the specific system and the properties of the molecules under investigation. Here are a few commonly used techniques:

- **Shake Flask Method:** This method is a conventional approach that entails achieving equilibrium between a compound of interest and a combination of an aqueous phase and an organic solvent (typically an organic solvent that is immiscible). By measuring the compound's concentrations in both phases, it becomes possible to calculate the partition coefficient. When micelles are present, the organic solvent phase can contain both freely dispersed compounds and compounds that are encapsulated within the micelles [26].

- **Dialysis Method:** In this method, a dialysis membrane is used to separate the micellar solution from an aqueous phase. The compound of interest is added to the micellar solution, and over time, it diffuses across the membrane into the aqueous phase. The concentrations of the compound in both phases are determined, allowing the calculation of the partition coefficient [27].

- **Chromatographic Methods:** Techniques such as high-performance liquid chromatography (HPLC) or thin-layer chromatography (TLC) or Micellar Electrokinetic Chromatography (MEKC) can be used to determine the partition coefficient. The compound is introduced into a chromatographic system containing micellar mobile and stationary phases. The compound will partition between the micelles in the mobile phase and the stationary phase, allowing for separation and measurement of the partition coefficient [28,24,29].

This work uses Micellar Electrokinetic Chromatography (MEKC). MEKC is a chromatographic technique that combines electrophoresis and micellar solutions as the separation medium. It has

gained significant attention in the field of analytical chemistry due to its unique capabilities for the separation and analysis of diverse analytes.

In a MEKC system, a capillary filled with an electrolyte solution containing surfactants is used. The surfactant molecules self-assemble to form micelles, which consist of a hydrophobic core and a hydrophilic shell. The micelles act as a pseudo-stationary phase, providing additional retention and separation mechanisms beyond the conventional electrophoretic mobility. The separation in MEKC is achieved through the combined effects of electrophoretic mobility, electroosmotic flow, and micellar solubilization. When an electric field is applied across the capillary, analytes migrate based on their charge-to-size ratio. The hydrophobic analytes can interact with the hydrophobic core of the micelles through hydrophobic interactions, leading to differential partitioning behavior. This interaction with the micelles allows for the separation of analytes based on their hydrophobicity in addition to their charge [28].

MEKC offers several advantages for analytical applications. It provides high separation efficiency, allowing for the resolution of complex mixtures with multiple components. MEKC can analyze a wide range of analytes, including small molecules, pharmaceuticals, natural products, proteins, and peptides. It also offers the ability to simultaneously analyze hydrophobic and hydrophilic compounds within a single run. In terms of methodology, MEKC requires the optimization of several parameters, including the type and concentration of surfactants, buffer pH, ionic strength, and capillary temperature. These parameters can be adjusted to achieve the desired separation and optimize the resolution and sensitivity of the method [29].

It includes different uses, such as pharmaceutical analysis, environmental monitoring, food analysis, forensic sciences, and bioanalysis.

From a computational point of view, the prediction of log P in systems of micelles could be performed by using molecular structures of micelles obtained by dynamic simulations (MDs) [30]. MD simulations starting from pre-assembled micelles (SDS, SC, LPFOS, and HTAB) CTAB (cetyltrimethylammonium bromide), C12E10, Brij35 (C12E23), Triton X-114 and Triton X-100 micelles were used to obtain micelles structures to be used as input for COSMOmic to predict micelle water partition coefficients [31]. Highly precise results were achieved when forecasting the log P values of both neutral and charged solutes within micellar systems by employing a combination of molecular dynamics (MD) simulations and umbrella sampling techniques [32]. Recently, the partition coefficients were calculated for the combined system of sodium laureth

sulfate (SLES) and fatty acids [33]. In this work, good agreement with the experimental data for the neutral solutes of capric acid and palmitic acid was obtained using both MD and COSMOmic approaches. However, for simulating charged solutes in anionic surfactant micelles, the use of an accurate polarizable force field was crucial to predict log P for the anion forms.

In this study, we introduce a computational approach that utilizes density functional theory (DFT) calculations to forecast the log P values for micellar systems. Our methodology focuses on predicting partition coefficients in pure solvents, offering a more time-efficient and universally applicable approach [34,35]. It must be noted that there are several approaches to determining the partition coefficient between pure phase [36,37]. Various methods employ quantum calculations or fragment-based approaches to estimate partition coefficients. In most cases, these methodologies rely on computing the transfer of free energy between two phases to determine the partition coefficients of molecules. In the fragment-based approach, the logarithm of a solute's partition coefficient (log P) is calculated by summing the fragment constants associated with each constituent molecular fragment. This method involves assigning specific values to each fragment and combining them to obtain the overall log P value. However, fragment-based methodologies are typically parameterized for a specified set of solvents, especially octanol/water systems, making them less applicable for predicting properties in various solvent environments. In the proposed methodology, the partitioning of molecules between the aqueous phase and the micelles can be correlated with the partition coefficient between the aqueous phase and a hypothetical solvent that possesses comparable physicochemical characteristics to the micellar system.

The present study employed density functional theory (DFT) calculations to predict the partition coefficients between organic solvents and water in a collection of 15 solvents. The methodology is utilized to calculate the partition coefficient of a compound within a solution containing micelles of Sodium Dodecyl Sulphate (SDS), hexadecyltrimethylammonium bromide (HTAB), sodium cholate (SC) and lithium perfluorooctanesulfonate (LPFOS). The Solvation Model Based on Density (SMD) is a suitable model for estimating solvation-free energy [38,39]. This method is implemented using the Becke three-parameter Lee-Yang-Parr (B3LYP) [34,40] and M06-2X functionals [41].

Finally, a comparison is conducted with experimental octanol/water partition coefficients to assess the accuracy of the methodology.

4.3   Experimental materials and methods

4.3.1   Apparatus and experimental conditions

The experimental determinations have been performed by Elisabet Fuguet of the Department of Chemical Engineering and Analytical Chemistry of the Universitat de Barcelona. MEKC determinations were conducted using a Beckman P/ACE System 5500 capillary electrophoresis instrument, which was equipped with a UV diode array detector. A fused silica capillary with a total length of 47 cm (40 cm effective length) and an internal diameter of 50 µm was employed. The measurements were carried out at a temperature of 25°C and an applied voltage of +15 kV. The detection wavelength was set at 214 nm. The test compounds were introduced into the capillary via pressure injection, with an application of 0.5 p.s.i. for 1 second.

The capillary was prepared by following this conditioning procedure: It was rinsed with water for 5 minutes, treated with a 1 M sodium hydroxide solution for 20 minutes, rinsed again with water for 10 minutes, treated with a 0.1 M sodium hydroxide solution for 10 minutes, and finally equilibrated with the separation buffer for 20 minutes. Before each injection, the capillary was flushed with the separation buffer for 5 minutes.

Four different micelle solutions were prepared at pH 7 With 40 mM SDS, 80 mM of SC, 40 mM of LPFOS, and 20 mM of HTAB, all three in 20 mM phosphate buffer. Test compounds were solved in a methanol solution (used as an electroosmotic flow marker), which already contained 2 mg ml^{-1} of phenyl-undecyl ketone (used as a micellar marker). The concentration of the test compounds was 2 mg mL^{-1}. All solutions were filtered through 0.45 µm nylon syringe filters (Albet). All measurements were performed by triplicate.

4.3.2   Experimental determination of partition coefficients in micelles (SDS, SC, LPFOS, HTAB)

In MEKC, the separation of neutral molecules relies on their partitioning between the aqueous phase and the micellar phase. The retention factor, k, of a compound, can be determined using the following equation:

$$k = \frac{t_R - t_0}{t_0\left(1 - \frac{t_R}{t_m}\right)} \tag{4.1}$$
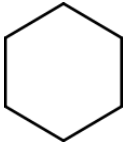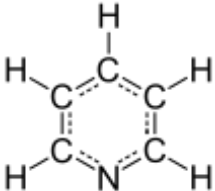
In the equation, the retention time (t_R) of the compound being analyzed is compared to the retention times of the electroosmotic flow (t_0) and the micellar markers (methanol and phenyl-undecyl ketone, denoted as t_m).

In this study, previously determined retention times were utilized to calculate the partition coefficients between water and SDS micelles using the following relationship [28,34]:
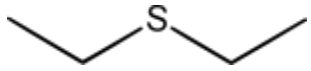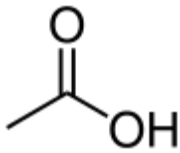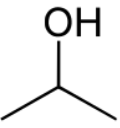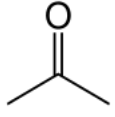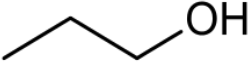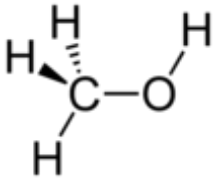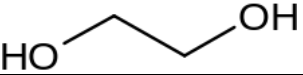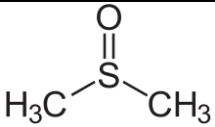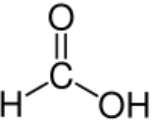
$$k = P \frac{v(C_T - CMC)}{1 - v(C_T - CMC)} \tag{4.2}$$

In the equation, P represents the partition coefficient of a compound between the micellar phase and the aqueous phase. v denotes the partial molar volume of the surfactant [42], C_T represents the total concentration of the surfactant, CMC stands for the critical micellar concentration, and k is the MEKC retention factor of the compound under investigation. The CMC value is obtained through conductimetric analysis [43]. Therefore, all partition coefficients listed in Table 2 were determined under consistent conditions, specifically in a 40 mM SDS micelle solution within a 20 mM phosphate buffer at pH=7 and a temperature of 25 °C.

Table 1. Structure of the solvents employed.

Solvent	Structure
Heptane	
Cyclohexane	
N-dodecane	
Pyridine	

Diethyl sulfide	
Acetic acid	
Decan-1-ol	
Octanol	
Propan-2-ol	
Acetone	
Propan-1-ol	
Methanol	
1,2-ethane diol	
Dimethyl sulfoxide	
Formic acid	

Table 2. List of experimental partition coefficients of compounds in SDS micelles (Log PSDS, SC, HTAB, LPFOS) deter-mined from retention factors obtained from MECK experiments with 40 mM SDS micelles solution, in 20 mM phosphate buffer at pH 7 at 25 ∘C using Equation (3.2).

Compound	log P$_{SDS}$	log P$_{SC}$	log P$_{HTAB}$	log P$_{LPFOS}$
Ethylbenzene	2.71	2.50	3.00	2.06
Propylbenzene	3.20	2.94	3.42	2.39
Butylbenzene	3.70	3.26	3.71	2.71
1-phenylethanone	2.08	1.33	2.03	2.19
1-phenylpropan-1-one	2.41	1.65	2.42	2.44
1-phenylbutan-1-one	2.77	2.01	2.80	2.72
1-phenylpentan-1-one	3.18	2.41	3.24	3.01
1-phenylheptan-1-one	4.17	3.15	-	3.68
Furan	1.26	0.77	1.48	1.19
2-nitroaniline	2.16	1.59	2.67	1.80
2,3-benzofuran	2.44	2.12	2.82	1.82
Diphenylmethanone	3.25	2.48	3.28	3.01
Benzamide	1.60	1.06	1.72	1.50
4-chloroaniline	2.18	1.69	2.69	1.44
2,3-dimethylphenol	2.31	1.90	3.15	1.66
Naphthalen-2-ol	2.73	2.31	-	1.73
4-aminobenzamide	1.27	0.98	1.11	1.76
3-methylphenol	1.97	1.53	2.78	1.43
2,4-dimethylphenol	2.38	1.93	3.17	1.02

Naphthalene	3.04	2.67	3.47	2.09
Pyrimidine	0.78	0.56	-	1.27
Benzaldehyde	1.90	1.20	1.91	1.91
3-chloroaniline	2.15	1.63	2.72	1.41
Pyrrole	1.04	0.68	1.65	0.72
3-nitroaniline	1.91	1.38	2.42	1.53
4-chlorophenol	2.25	2.00	3.24	1.30
Phenol	1.58	1.21	2.35	1.08
Methylbenzoate	2.40	1.71	2.39	2.36
Bromobenzene	2.60	2.37	2.95	1.80
1,4-xylene	2.77	2.51	3.04	2.10
Benzene-1,3-diol	1.27	1.21	2.48	0.75
2-methylaniline	1.90	1.17	2.15	1.59
1-methoxy-2-nitrobenzene	2.17	1.55	2.37	2.26
N-4chlorophenylacetamide	2.43	2.03	2.80	1.84
Aniline	1.59	0.92	1.83	1.34
Nitrobenzene	1.99	1.47	2.21	1.94
Chlorobenzene	2.44	2.21	2.77	1.77
N-phenylacetamide	1.78	1.25	1.98	1.58
4-nitroaniline	1.94	1.52	2.50	1.45
Anisole	2.10	1.66	2.31	1.83
Benzonitrile	1.91	1.21	1.96	1.95

1-ethyl-4-nitrobenzene	2.83	2.19	3.02	2.68
1-methoxy-4-nitrobenzene	2.36	1.69	2.58	2.20
N,N-diethyl-4-nitroaniline	3.42	2.44	3.56	3.36
Benzyl benzoate	3.96	2.99	-	3.18
Caffeine	1.67	1.11	1.32	1.85
Corticosterone	3.99	1.94	3.69	3.64
Cortisone	3.39	1.72	3.16	3.37
β-Estradiol	4.08	2.77	-	2.84
Estriol	3.12	2.32	3.52	2.01
Cortisol	3.44	1.83	3.39	2.89
Hydroquinone	1.03	1.09	1.94	0.19
Quinoline	2.59	1.65	2.36	2.68
Atrazine	2.84	1.86	1.90	2.71
Diuron	2.92	2.46	2.19	2.34
Fluometuron	2.56	2.01	1.92	2.57
Isoproturon	2.83	2.19	1.95	2.61
Linuron	3.08	2.59	2.24	2.50
Metobromuron	2.66	2.16	2.03	2.22
Monuron	2.34	1.81	1.73	2.03
Metoxuron	2.44	1.69	1.46	2.34
Phenyl urea	1.78	1.20	1.20	1.38
Propazine	3.13	2.02	2.08	3.03

4.3.3 Calculation of partition coefficients

In this study, quantum chemistry serves as an ideal approach for accurately determining the partition coefficients of a specific group of compounds. The main objective is to investigate the partitioning behavior of 63 molecules. To ensure efficiency, each compound is represented by a single conformation, specifically selecting the most elongated conformation for flexible molecules. This approach allows for a focused analysis of the partitioning characteristics within the given set of compounds.

The geometries of all compounds were optimized using Density Functional Theory (DFT) employing B3LYP and M06-2X [41] methods, with a 6-31++G** basis set. Additionally, the continuum solvation model based on density (SMD) was employed [38,44]. As a universal solvation model, SMD can be utilized for any type of solvent, accommodating both charged and uncharged solutes.

The solvation-free energy in this model is divided into two primary parts: the bulk electrostatic portion and the cavity dispersion portion.

4.3.3.1 Bulk electrostatic portion

The bulk electrostatic portion in solvation-free energy refers to the contribution of electrostatic interactions between the solute and the surrounding solvent molecules in the bulk phase. It represents the energetic cost or gain associated with the redistribution of charges when the solute is introduced into the solvent. This component takes into account the Coulombic interactions between charged species, including ions or polar molecules, and the surrounding solvent molecules, accounting for the solvent's dielectric constant. In essence, the bulk electrostatic portion quantifies the influence of electrostatic forces on the overall solvation process [45].

4.3.3.2 Cavity dispersion portion

The cavity dispersion portion refers to the contribution of dispersion forces within the solvation process. Dispersion forces, also known as London dispersion forces or van der Waals forces, arise from temporary fluctuations in electron density and induce attractive interactions between molecules [46].

In the context of solvation, the cavity dispersion portion quantifies the energetic cost or gain associated with the formation of cavities or voids in the solvent to accommodate the solute molecule. It takes into account the interactions between the solute and the surrounding solvent molecules, considering both the attractive dispersion forces and any repulsive interactions that may arise due to steric hindrance. This component accounts for the nonpolar interactions between the solute and solvent molecules, which are crucial for understanding the solvation behavior of nonpolar or hydrophobic solutes [47].

4.3.4 *Application of DFT calculation*

The primary advantage of employing Density Functional Theory (DFT) in this methodology is its universal applicability, as it has the potential to be used for any type of molecule. Calculations were performed with the electronic structure program Gaussian 16 [50] which provides a wide-ranging suite for prediction of molecular properties.
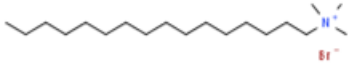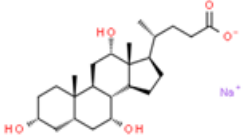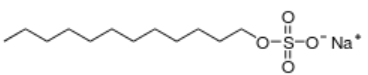
The initial molecular structures were created using the freely available, cross-platform molecule editor Avogadro [48]. Only solvation energies derived from minimizations that produced all positive frequencies were taken into account. Harmonic vibrational frequencies were calculated for all compounds. The thermochemical values were determined at a temperature of 298.15 K and a pressure of 1.0 atm.

The partition coefficient of a molecule is related to the Gibbs free energy of transfer, $\Delta G^{\circ}_{solv/wat}$, between water and a particular solvent. This free energy is, therefore, obtained via the calculation of absolute solvation-free energies in the respective media by equations (1.1) and (1.2).

4.3.5 *Statistical analysis*

Linear regression analysis was conducted using the regression data analysis tool in Microsoft Excel. This analysis yielded coefficients, confidence intervals, standard errors, F statistics, significant F and p-values, as well as the Pearson correlation coefficient (R^2). As part of the evaluation process for predicting experimental octanol/water log P values, various statistical measures were computed, including the mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE).

Table 3. Structure of micelles.

Micelle	Symbol	Structure
Hexadecyltrimethylammonium bromide	HTAB	
Lithium perfluorooctanesulfonate	LPFOS	
Sodium cholate	SC	
Sodium dodecyl sulfate	SDS	

4.3.5.1   *Identification of best solvents for prediction of log P in micelles*

This research involved predicting the water partition coefficients of 63 compounds (Listed in Table 2) across 15 different solvents as shown in Table 1. The experimental partition coefficients in SDS, SC, LPFOS, and HTAB micelle of those molecules were correlated with respect to the predicted partition coefficients in all 15 solvents with respect to water (Listed in Table 2).

Table 4 displays the correlation coefficients obtained from B3LYP and M06-2X calculations for SDS, SC, LPFOS, and HTAB. The results of both B3LYP and M06-2X methods exhibit a comparable partition coefficient, indicating negligible disparities between the two approaches. An

analysis was conducted to compare the calculated and experimental partition coefficients for HTAB, SC, SDS, and LPFOS micelles.

As seen in Table 4, in SDS micelles, Propan-2-ol, propan-1-ol, and methanol exhibit the highest correlations with experimental data ($R^2 \geq 0.7$).

The best correlation with experimental values ($R^2 \geq 0.67$) is obtained for Propan-1-ol for SC micelles.

Propan-1-ol and Propan-2-ol show the highest correlation to compare experimental data for LPFOS micelles $R2 \geq 0.52$ and $R2 \geq 0.53$, respectively.

The partition coefficient calculated for propane-1-ol compared to the experimental partition coefficient of SC and SDS micelles provided the best correlation and both micelles show similar result in alcohol solvents.

Observations show that SDS, SC, and LPFOS micelles behave similarly to alcoholic solvents with dielectric constants ranging from 20 to 33. It should be noted that since these solvents are miscible with water, the partition coefficients of these solvents cannot be determined using the traditional shake flask technique. However, these coefficients can still be calculated by employing suitable thermodynamic cycles and utilizing immiscible solvents. Additionally, according to the data in Table 4, the behavior of SDS is not effectively replicated by a solvent like N-dodecane (with an R^2 value less than 0.1), even though N-dodecane shares a chemical structure resembling the hydrophobic tail of SDS. Instead, it appears that the hydrophilic characteristics of the sulfate group in SDS and the hydrophobic properties of the dodecyl tail are better represented by alcohols mixture solvent with moderate dielectric constants, such as Propan-1-ol, Propan-2-ol, and methanol.

However, log P in HTAB is not correlated with any combination of solvents.

Figure 5 and Table 5 shows the pairwise correlation between experimental and calculated log P values but with the exclusion of compounds containing nitrogen in an aromatic ring or the urea group (they referred N set). It is observed that all experimental log P values, including log P HTAB, show a high correlation with 2-propanol and 1-propanol. Additionally, for log P SC and log P HTAB, a high correlation with methanol is also observed. This suggests that the excluded

compounds may have a different mechanism for describing the partition coefficient of the HTAB micelle.

Also, Table 4 highlights the bold values representing the highest correlation between the experimental and calculated data.

Table 4. Linear regression parameters were obtained for the correlation of the calculated Log P in 15 different solvents with respect to the experimental partition coefficients in SDS, SC, LPFOS, and HTAB micelles. Solvents are sorted as a function of their dielectric constant.

Solvent	Dielectric Constant	R^2 SDS B3LYP	R^2 SDS M06-2X	R^2 SC B3LYP	R^2 SC M06-2X	R^2 LPFOS B3LYP	R^2 LPFOS M06-2X	R^2 HTAB B3LYP	R^2 HTAB M06-2X
Heptane	1.92	0.04	0.07	0.002	0.26	0.28	0.26	0.11	0.15
Cyclohexane	2.02	0.03	0.06	0.3	0.14	0.14	0.02	0.14	0.006
N-dodecane	2.03	0.08	0.04	0.2	0.15	0.15	0.02	0.09	0.01
Pyridine	2.35	0.20	0.30	0.3	0.29	0.18	0.28	0.0034	0.001
Diethyl sulfide	6.14	0.12	0.22	0.26	0.26	0.09	0.18	0.0019	0.008
Acetic acid	6.20	0.03	0.09	0.16	0.17	0.01	0.00	0.01	0.03
Decan-1-ol	7.53	0.26	0.34	0.46	0.47	0.11	0.15	0.06	0.03
Octanol	10.30	0.40	0.48	0.46	0.59	0.11	0.28	0.09	0.12
Propan-2-ol	19.26	**0.68**	0.68	**0.64**	0.61	**0.53**	0.47	0.1	0.15
Acetone	20.16	0.19	0.35	0.32	0.35	0.17	0.35	0.01	0.003
Propan-1-ol	21.03	**0.73**	0.69	**0.67**	0.58	**0.52**	0.47	0.13	0.06
Methanol	32.61	**0.71**	0.62	**0.58**	0.55	**0.43**	0.41	0.13	0.04
1,2-ethan diol	40.24	0.12	0.26	0.17	0.33	0.0003	0.04	0.06	0.007
Dimethyl sulfoxide	46	0.13	0.23	0.22	0.28	0.13	0.52	0.0008	0.001
Formic acid	58	0.06	0.10	0.17	0.10	0.0004	0.00.5201	0.015	0.006

The results of the linear regression analysis for the partition coefficients of 1-propanol-water, 2-propanol, and Methanol for SDS, SC, LPFOS, HTAB, and HTAB without aromatic N and NCON group (Urea) atoms are presented in Table 5.

Table 5: Best linear regressions were obtained to predict the log P in SC, SDS, LPFOS, and HTAB micelles using DFT calculations. Results from B3LYP and M062-X functionals with 6-31++G** basis set using SC, SDS, LPFOS, HTAB and HTAB without N set (molecules with Nitrogen in an aromatic ring or with the urea or carbamide group) for propan-1-ol, propan-2-ol, and methanol are indicated. x refers to the predicted log P alcohol/water, and y refers to the predicted log P in micelles.

Micelle	Solvent	B3LYP
LPFOS	Propan-1-ol	$y=0.46x + 0.77$ $R^2=0.52$ MAE=0.87
	Propan-2-ol	$y=0.49x + 0.61$ $R^2=0.53$ MAE=0.92
	Methanol	$y=0.41x + 0.90$ $R^2=0.43$ MAE=0.86
SC	Propan-1-ol	$y=0.47x + 0.55$ $R^2=0.67$ MAE=0.92
	Propan-2-ol	$y=0.46x + 0.51$ $R^2=0.64$ MAE=1.08
	Methanol	$y=0.41x + 0.68$ $R^2=0.58$ MAE=0.89
HTAB	Propan-1-ol	$y=0.23x + 1.80$ $R^2=0.13$ MAE=0.74
	Propan-2-ol	$y=0.22x + 1.83$ $R^2=0.1$ MAE=0.72
	Methanol	$y=0.24x + 1.78$ $R^2=0.13$ MAE=0.72
HTAB without N set	Propan-1-ol	$y=0.56x+1.24$ $R^2=0.66$ MAE=0.45
	Propan-2-ol	$y=0.54x+1.24$ $R^2=0.62$ MAE=0.43
	Methanol	$y=0.56x+1.26$ $R^2=0.63$ MAE=0.46

The linear regression results for the alcohol-water partition coefficients, regardless of the functional employed in the DFT calculations (B3LYP and M06-2X), are presented in Table 5. The obtained slopes have similar values similarity, suggesting comparable predictions for both calculation sets.

In addition, a general trend is observed, wherein more hydrophobic compounds tend to display higher log P values, while those with lower values are more hydrophilic. Consequently, these log P values for the four types of micelles appear to serve as a measure of the lipophilicity or hydrophobicity of the respective compounds.

Among the three solvent systems analyzed, the log $P_{propan-1-ol/water}$ exhibited the highest degree of linear correlation with the experimental log P_{SDS} in SDS micelles, as shown in Figure 1. The log $P_{propan-1-ol/water}$ values obtained using the B3LYP function and 6-311++G** basis set demonstrate the strongest correlation among the two functionals assessed, with an estimated R^2 value of 0.73. In conclusion, the results of this study display the establishment of a linear correlation between the partition coefficient of the molecules under analysis in SDS micelles and the log P partition coefficient in propan-1-ol/water, propan-2-ol/water, and methanol/water (Figure 1).
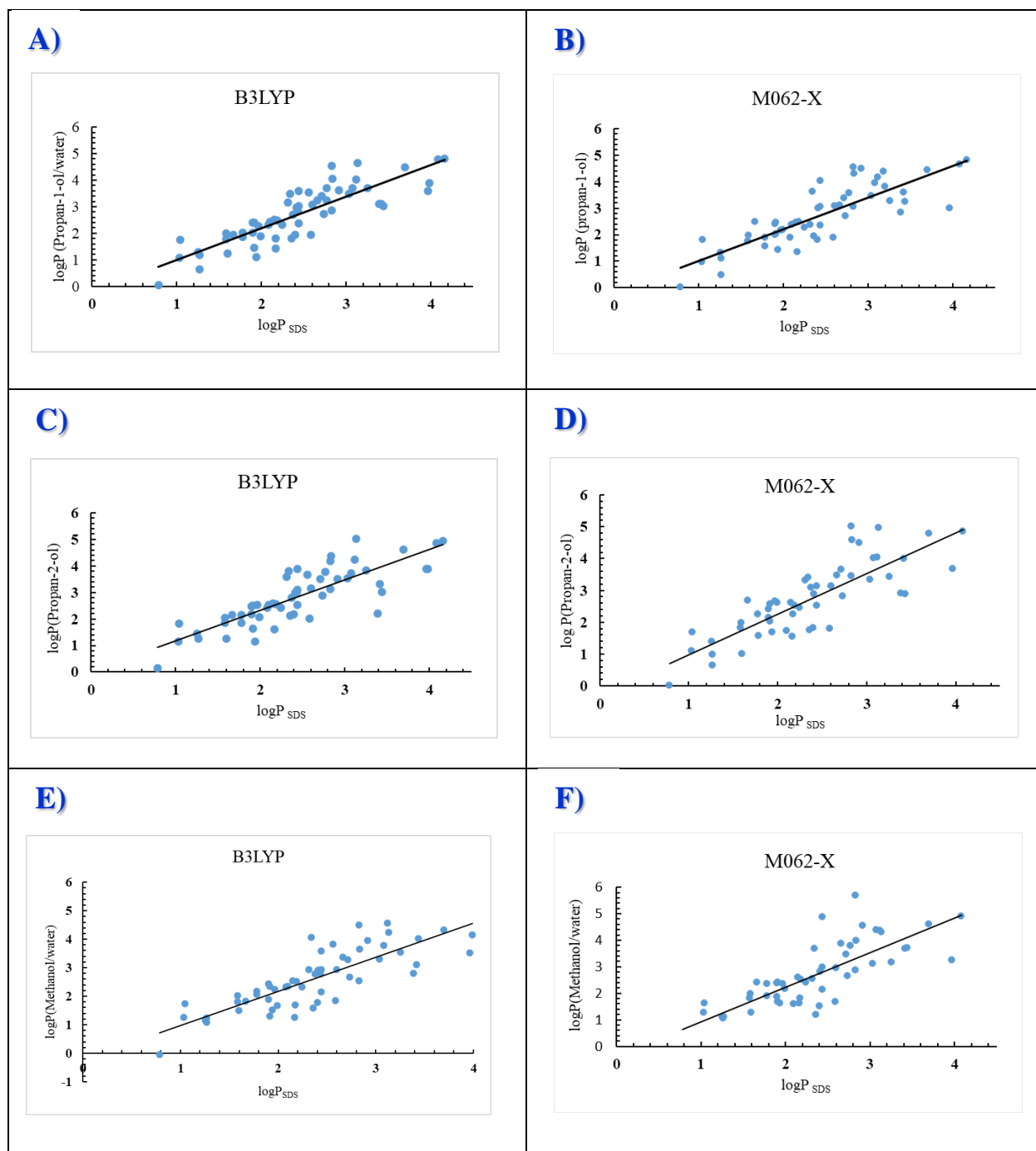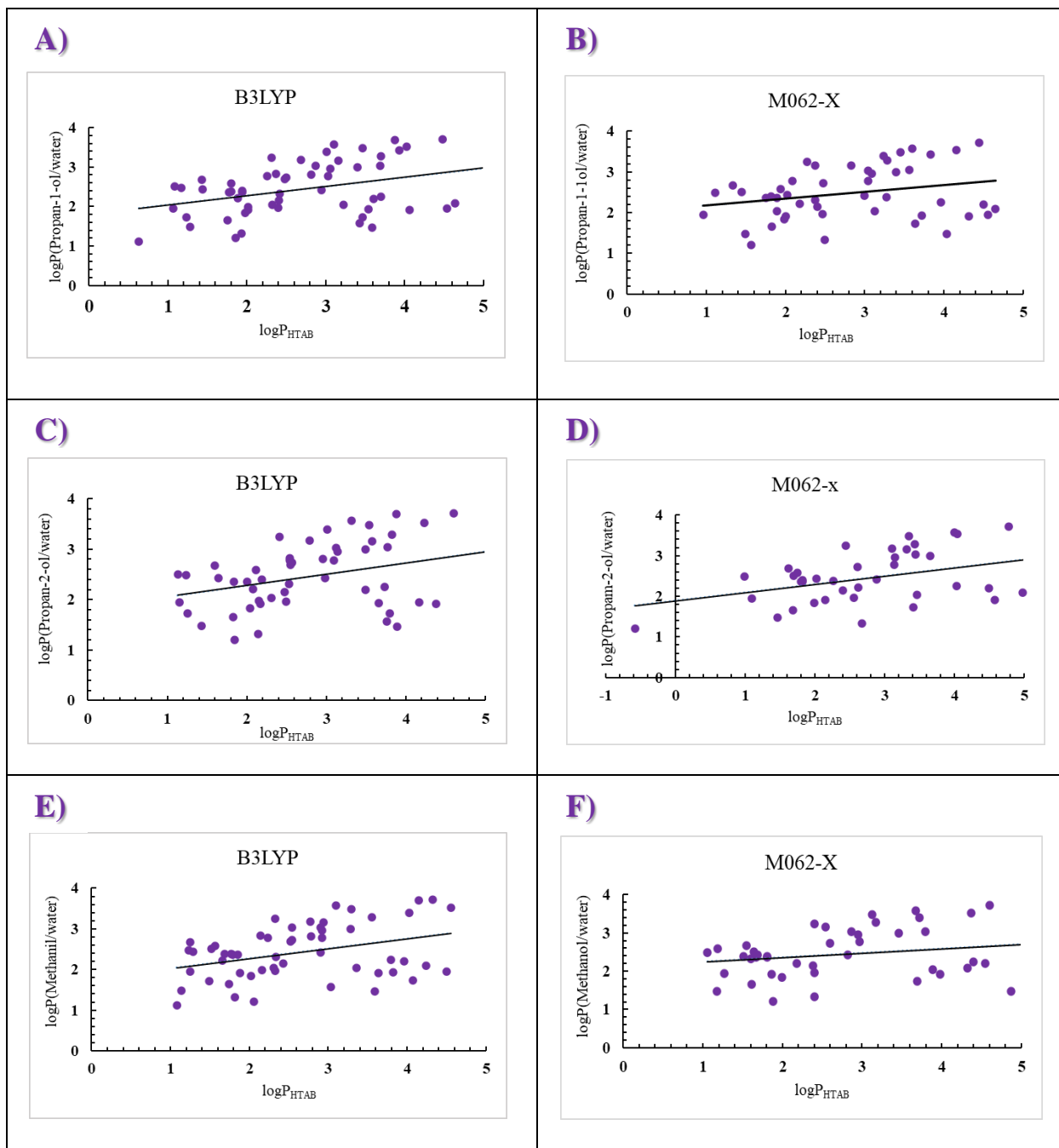
Figure 1. Comparison of the Experimental log P (SDS) with **A)** the calculated log P (Propan-1-ol/water) for B3LYP and **B)** M06-2X **C)** the calculated log P (Propan-2-ol/water) for B3LYP **D)** M062-X **E)** the calculated log P (Methanol/water) for B3LYP **F)** M062-X DFT calculations.

The best correlation was observed between the experimental log Psc and logP$_{\text{1-propanol/water}}$, among the investigated micelles and alcohol systems (1-propanol, 2-propanol, and methanol), as shown in Figure 2.
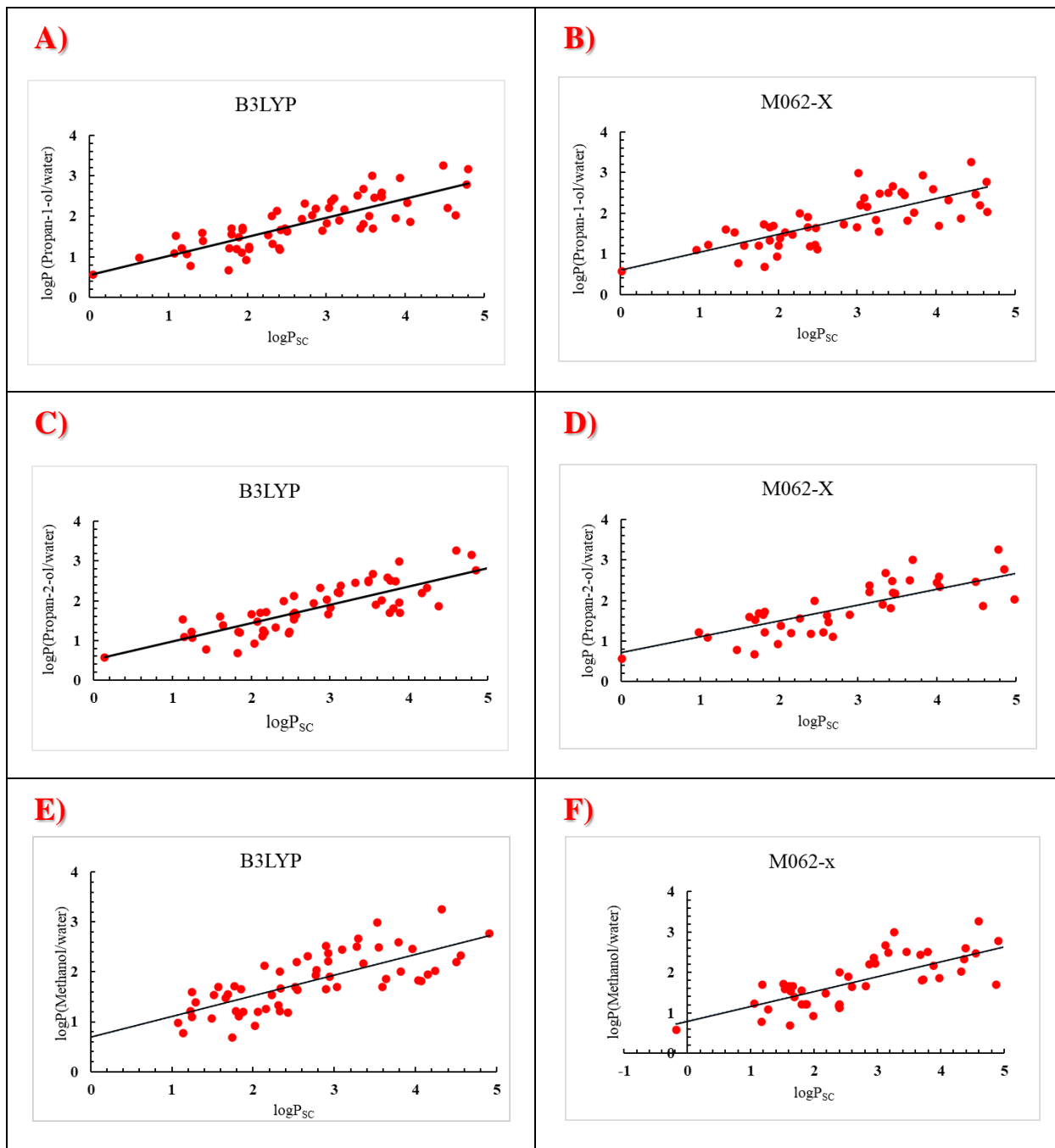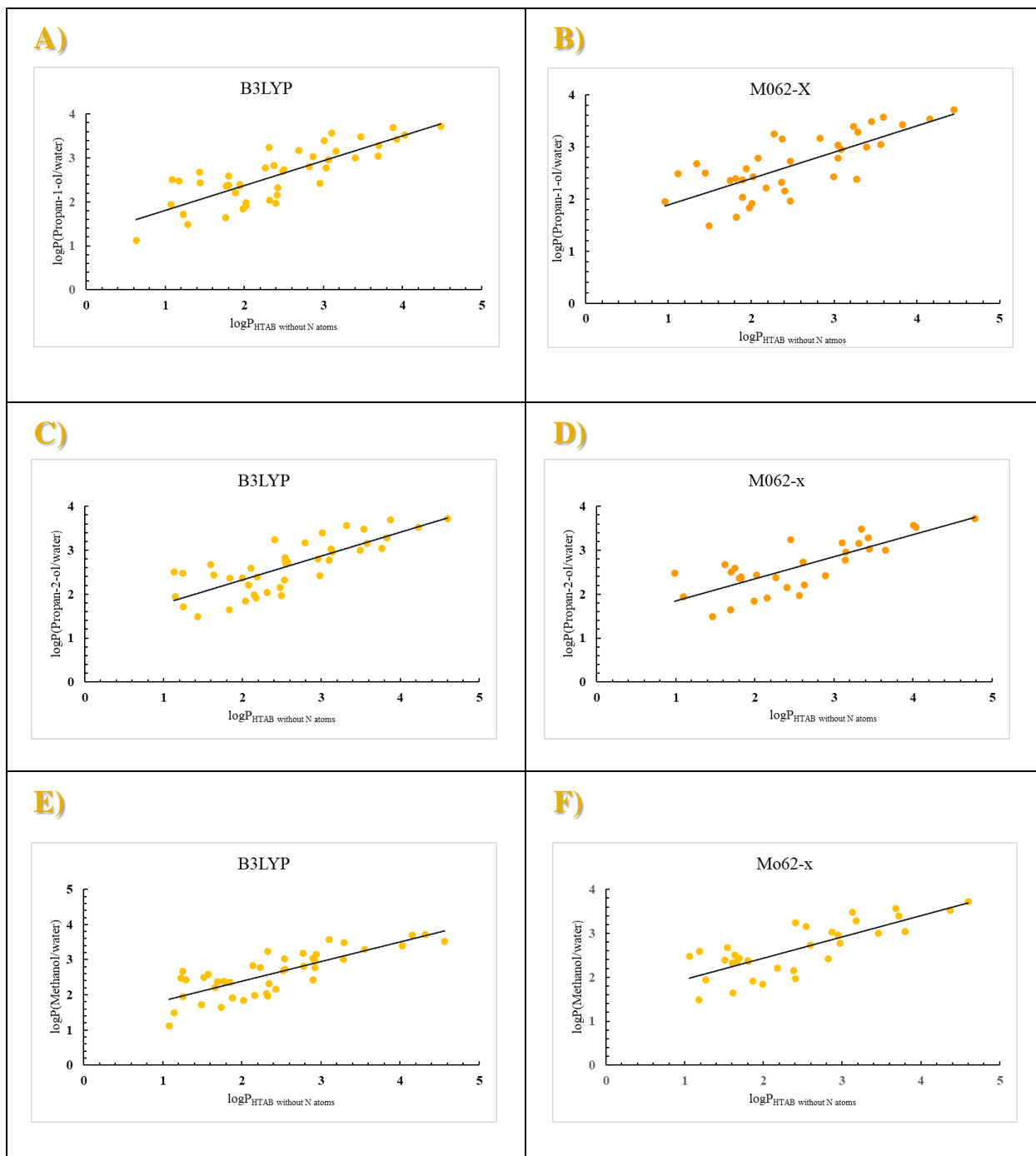

Figure 2. Comparison of the Experimental logP (SC) with **A)** the calculated log P (Propan-1-ol/water) for B3LYP and **B)** M06-2X **C)** the calculated logP (Propan-2-ol/water) for B3LYP **D)** M062-X **E)** the calculated log P (Methanol/water) for B3LYP **F)** M062-X DFT calculations.

Regardless of the DFT calculation method used, both the B3LYP and M06-2X methods yielded similar slopes, indicating a comparable distribution coefficient prediction. The best correlation was observed between the experimental log P_{LPFOS} and $logP_{1-propanol/water}$, among the investigated micelles and alcohol systems (1-propanol, 2-propanol, and methanol), as shown in Figure 3.
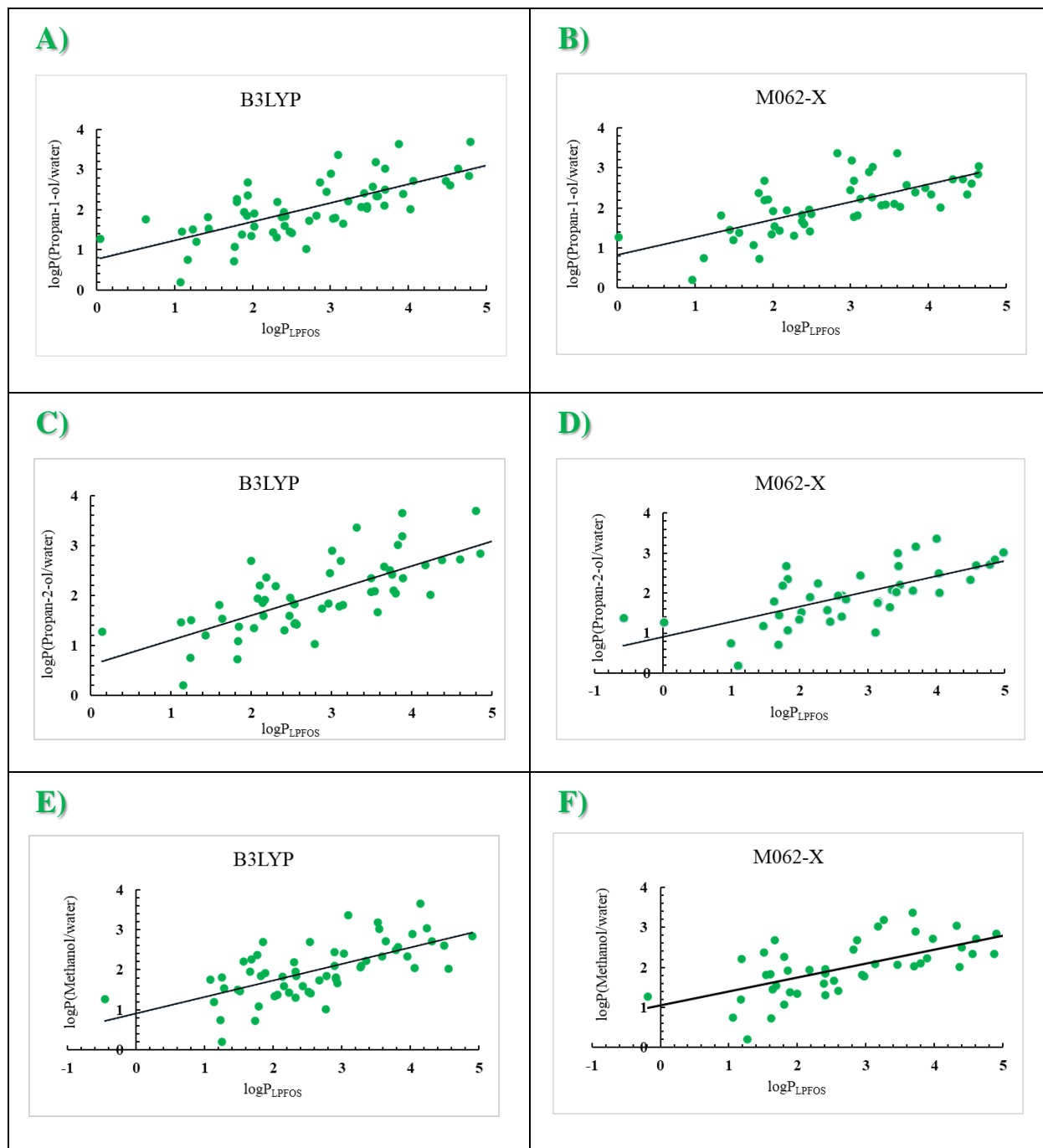
Figure 3. Comparison of the Experimental log P (LPFOS) with **A)** the calculated log P (Propan-1-ol/water) for B3LYP and **B)** M06-2X **C)** the calculated log P (Propan-2-ol/water) for B3LYP **D)** M06-2X **E)** the calculated log P (Methanol/water) for B3LYP **F)** M062-X DFT calculations.

When estimating all compound sets, the findings exhibited a strong association between experimental log P values in SDS, SC, and LPFOS for 1-propanol/water or 2-propanol/water solvent mixtures. However, the experimental log P in HTAB exhibited no correlation with any of the computed log P values. Figure 4 shows the poor correlation between experimental partition coefficient and calculated partition coefficients in alcohol solvents.

Figure 4. Comparison of the Experimental log P (HTAB) with **A)** the calculated log P (Propan-1-ol/water) for B3LYP and **B)** M06-2X **C)** the calculated log P (Propan-2-ol/water) for B3LYP **D)** M06-2X **E)** the calculated log P (Methanol/water) for B3LYP **F)** M062-X DFT calculations.

From the data presented in Table 4 and Figure 5, it is apparent that HTAB micelles displayed no correlation in all solvents. Consequently, different analysis was performed to select groups of

molecules that could have a significant correlation. In that way, a new set of molecules were defined excluding molecules having aromatic N atoms or those molecules having the urea group.

Figure 5. Comparison of the Experimental logP (HTAB without N set) and with **A)** the calculated logP (Propan-1-ol/water) for B3LYP and **B)** M06-2X **C)** the calculated logP (Propan-2-ol/water) for B3LYP **D)** M06-2X **E)** the calculated logP (Methanol/water) for B3LYP **F)** M062-X DFT calculations.

Summarizing, it has been determined the experimental partition coefficients (Log P values) of a diverse set of compounds in four different types of micelles: SDS, SC, LPFOS, and HTAB. The obtained Log P values were used to parametrize computational methodologies for each type of micelle. Correlations between experimental and calculated log P values were investigated using simple DFT calculations.

When considering the entire set of compounds, the results showed a high correlation between experimental log P values in SDS, SC, and LPFOS for propan-1-ol/water or propan-2-ol/water solvent mixtures. However, log P values in HTAB were only well correlated with a specific selection of compound.

4.3.6   Comparison with experimental Octanol/Water partition coefficient

To assess the accuracy of the calculated partition coefficients in various solvents, a comparison was made between the predicted and experimental partition coefficients in the octanol/water system (Table 5, Figure 6).

Table 6. Linear Regression Parameters Obtained for the Calculated Log $P_{O/W}$ (octanol/water) with Respect to the Experimental partition coefficients in octanol/water. Statistical Error Assessment of the Linear Regression Terms Based on Applied Computational Models are indicated. Below each slope, a 95% confidence interval is indicated in parentheses.

Density Functional	Slope	R^2	MAE	RMSE	MSE
B3LYP	0.67 (0.52,0.82)	0.60	0.58	0.79	0.63
M062-X	0.71 (0.57,0.86)	0.65	0.64	0.79	0.63

A strong linear correlation was observed when comparing the log $P_{O/W}$ (partition coefficient between octanol and water) with experimental values using both the B3LYP and M06-2X functionals. The confidence intervals for the regression coefficients (as shown in Table 6) indicate that the predictions obtained using both functions are statistically equivalent. Generally, similar results were obtained for both methods. The M06-2X functional exhibited a slightly better Pearson correlation coefficient, but the comparison of mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) showed similar results for both functionals.

Overall, the predictions achieved with both methods were highly accurate, enabling meaningful comparisons between calculated values using the B3LYP and M06-2X functionals, along with the SMD solvation model and the 6-311++G** basis set. The MAE indicated a difference of approximately half a logarithmic unit between the calculated and experimental partition coefficients for octanol/water. Therefore, the density functional theory (DFT)-based SMD solvation model appears to be suitable for predicting the partition coefficient between octanol and water for this set of molecules.

It is reasonable to expect a similar level of prediction for partition coefficients in other solvents; however, experimental validation would be necessary to confirm this.
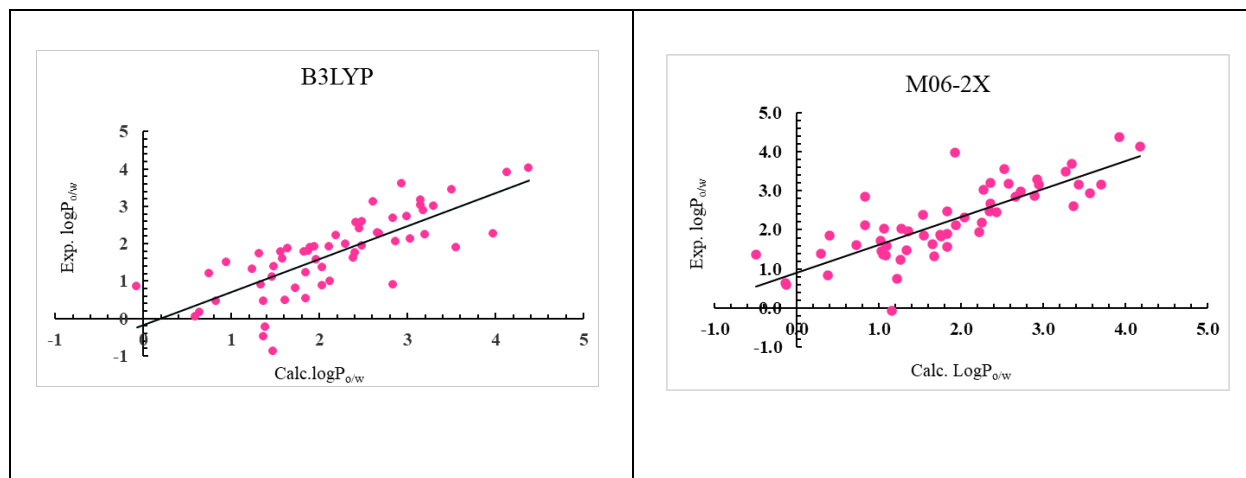
Figure 6. Comparison of the calculated log P in octanol/water with the experimental log P in (a) octanol/water for B3LYP and (b) M06-2X DFT calculations.

Table 7. List of calculation (B3LYP and M06-2X) and experimental $LogP_{o/w}$ used in this study.

Molecule	Calculation log $P_{O/W}$ (B3LYP)	Calculation log $P_{O/W}$ (M06-2X)	Experimental log $P_{O/W}$
Ethylbenzene	3.03	2.95	3.15 [49]
Propylbenzene	NA	3.35	3.69 [50]
Butylbenzene	4.02	3.92	4.38 [51]
1-phenylethanone	1.58	1.13	1.58 [52]
1-phenylpropan-1-one	2.22	2.25	2.19 [52]
1-phenylbutan-1-one	2.28	NA	2.66 [52]

1-phenylpentan-1-one	NA	3.42	3.15 [53]
1-phenylheptan-1-one	3.91	4.17	4.13 [52]
Furan	0.89	0.90	1.34 [54]
2-nitroaniline	0.54	0.40	1.85 [50]
2,3-benzofuran	1.90	1.83	1.90 [50]
Diphenylmethanone	2.88	2.57	3.18 [55]
Benzamide	0.16	-0.12	0.64 [56]
4-chloroaniline	1.78	1.75	1.83 [57]
2,3-dimethylphenol	2.60	2.34	2.48 [58]
Naphthalen-2-ol	1.94	1.83	2.48 [59]
3-methylphenol	1.57	1.42	1.96 [60]
2,4-dimethylphenol	1.98	NA	2.30 [61]
Naphtalene	3.01	2.92	3.30 [62]
Pyrimidine	-0.47	-0.49	1.37 [63]
Benzaldehyde	1.38	1.34	1.48
3-chloroaniline	1.81	1.74	1.88 [64]
1H-pyrrole	1.20	1.21	0.75
3-nitroaniline	0.47	1.05	1.37[50]
4-chlorophenol	1.61	1.53	2.39
Phenol	1.10	1.02	1.46 [65]
Methylbenzoate	1.003	0.82	2.12 [51]
Bromobenzene	2.73	2.75	2.99
1,4-xylene	3.17	3.70	3.15 [66]
2-methylaniline	1.73	1.67	1.32 [67]
1-methoxy-2-nitrobenzen	0.80	1.01	1.73 [68]
N-4-chlorophenylacetamide	1.76	NA	2.41 [69]
Aniline	1.32	1.26	1.24 [67]
Nitrobenzene	1.22	1.54	1.85 [70]
Chlorobenzene	2.68	2.66	2.84 [71]
N-phenylacetamide	0.89	0.82	2.84 [57]
4-nitroaniline	-0.22	0.28	1.39 [50]
Anisole	1.92	1.93	2.11 [70]
Benzonitrile	1.77	1.83	1.56 [70]
1-ethyl-4-nitrobenzene	2.13	2.27	3.03 [72]
1-methoxy-4-nitrobenzene	0.88	1.06	2.03 [73]
N, N-diethyl-4-nitroaniline	1.90	2.52	3.55[74]
Benzyl benzoate	2.26	1.92	3.97 [75]
Caffeine	0.86	1.15	-0.07 [76]
Corticosterone	1.49	NA	0.94 [77]
Cortisone	-0.88	NA	1.47 [78]
b-estradiol	3.44	3.27	3.50 [79]

Estriol	2.40	2.43	2.45 [77]
Cortisol	0.48	0.72	1.61 [80]
Hydroquinone	0.05	-0.12	0.59
Quinoline	1.36	1.27	2.03 [70]
Atrazine	3.12	3.37	2.61 [81]
Diuron	2.28	2.35	2.68 [82]
Fluometuron	2.56	NA	2.42 [83]
Isoproturon	NA	2.89	2.87 [84]
Linuron	2.25	2.35	3.20 [85]
Metobromuron	NA	2.04	2.32 [35]
Monuron	1.91	2.22	1.94 [86]
Metoxuron	1.88	1.65	1.64 [86]
Phenyl urea	0.47	0.38	0.83 [87]
Propazine	3.61	3.57	2.93 [88]

4.4   References

[1] Hubbe, M. A., Rojas, O. J., & Lucia, L. A. (2015). Green modification of surface characteristics of cellulosic materials at the molecular or nano scale: A review. BioResources, 10(3), 6095–6206. doi: 10.15376/biores.10.3.Hubbe.

[2] Ghosh, S., Ray, A., & Pramanik, N. (2020). Self-assembly of surfactants: An overview of general aspects of amphiphiles. Biophysical Chemistry, 265(May), 106429. doi: 10.1016/j.bpc.2020.106429.

[3] Elsabee, M. Z., Morsi, R. E., & Al-Sabagh, A. M. (2009). Surface active properties of chitosan and its derivatives. Colloids and Surfaces B: Biointerfaces, 74(1), 1–16. doi: 10.1016/j.colsurfb.2009.06.021.

[4] Marrink, S. J., & Mark, A. E. (2002). Molecular dynamics simulations of mixed micelles modeling human bile. Biochemistry, 41(17), 5375–5382. doi:10.1021/bi015613i.

[5] Nagarajan, R. (1991). Theory of Surfactant Self-Assembly: A Predictive Molecular Thermodynamic Approach. Journal Name, 7(12), 2934–2969. doi:10.1021/la00060a012.

[6] Malik, M. A., Hashim, M. A., Nabi, F., Al-Thabaiti, S. A., & Khan, Z. (2011). Anti-corrosion ability of surfactants: a review. Int. J. Electrochem. Sci, 6(6), 1927-1948.

[7] Venturoli, M., Sperotto, M. M., Kranenburg, M., & Smit, B. (2006). Mesoscopic models of biological membranes. Physics Reports, 437(1-2), 1-54. doi: 10.1016/j.physrep.2006.07.006.

[8] Perinelli, D. R., Cespi, M., Lorusso, N., Palmieri, G. F., Bonacucina, G., & Blasi, P. (2020). Surfactant self-assembling and critical micelle concentration: one approach fits all. Langmuir, 36(21), 5745-5753. doi: 10.1021/acs.langmuir.0c00420.

[9] Priev, A., Zalipsky, S., Cohen, R., & Barenholz, Y. (2002). Determination of critical micelle concentration of lipopolymers and other amphiphiles: comparison of sound velocity and fluorescent measurements. Langmuir, 18(3), 612-617.doi: 10.1021/la0110085.

[10] Correa, N. M., Silber, J. J., Riter, R. E., & Levinger, N. E. (2012). Nonaqueous polar solvents in reverse micelle systems. Chemical Reviews, 112(8), 4569-4602.doi: 10.1021/cr200254q.

[11] Wu, Z., Gao, R., Zhou, G., Huang, Y., Zhao, X., Ye, F., & Zhao, G. (2021). Effect of temperature and pH on the encapsulation and release of β-carotene from octenylsuccinated oat β-glucan micelles. Carbohydrate Polymers, 255, 117368. doi: 10.1016/j.carbpol.2020.117368.

[12] Lukyanov, A. N., & Torchilin, V. P. (2004). Micelles from lipid derivatives of water-soluble polymers as delivery systems for poorly soluble drugs. Advanced drug delivery reviews, 56(9), 1273-1289. doi: 10.1016/j.addr.2003.12.004.

[13] Liu, M., Kono, K., & Fréchet, J. M. (2000). Water-soluble dendritic unimolecular micelles: Their potential as drug delivery agents. Journal of Controlled Release, 65(1-2), 121-131. doi:10.1016/S0168-3659(99)00245-X.

[14] Torchilin, V. P. (2001). Structure and design of polymeric surfactant-based drug delivery systems. Journal of controlled release, 73(2-3), 137-172. doi: 10.1016/S0168-3659(01)00299-1.

[15] Venkataraman, S., Hedrick, J. L., Ong, Z. Y., Yang, C., Ee, P. L. R., Hammond, P. T., & Yang, Y. Y. (2011). The effects of polymeric nanostructure shape on drug delivery. Advanced drug delivery reviews, 63(14-15), 1228-1246. doi: 10.1016/j.addr.2011.06.016.

[16] Chiappetta, D. A., & Sosnik, A. (2007). Poly (ethylene oxide)–poly (propylene oxide) block copolymer micelles as drug delivery agents: improved hydrosolubility, stability and bioavailability of drugs. European Journal of Pharmaceutics and Biopharmaceutics, 66(3), 303-317. doi: 10.1016/j.ejpb.2007.03.022.

[17] Palazzolo, S., Bayda, S., Hadla, M., Caligiuri, I., Corona, G., Toffoli, G., & Rizzolio, F. (2018). The clinical translation of organic nanomaterials for cancer therapy: a focus on polymeric nanoparticles, micelles, liposomes and exosomes. Current medicinal chemistry, 25(34), 4224-4268. doi: 10.2174/0929867324666170830113755.

[18] Seynhaeve, A. L. B., Amin, M., Haemmerich, D., Van Rhoon, G. C., & Ten Hagen, T. L. M. (2020). Hyperthermia and smart drug delivery systems for solid tumor therapy. Advanced Drug Delivery Reviews, 163, 125-144. doi: 10.1016/j.addr.2020.02.004.

[19] Movassaghian, S., Merkel, O. M., & Torchilin, V. P. (2015). Applications of polymer micelles for imaging and drug delivery. Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 7(5), 691-707. doi: 10.1002/wnan.1332.

[20] Torchilin, V. P. (2010). Passive and active drug targeting: drug delivery to tumors as an example. Drug delivery, 3-53. doi: 10.1007/978-3-642-00477-3.

[21] Rai, P., Mallidi, S., Zheng, X., Rahmanzadeh, R., Mir, Y., Elrington, S., ... & Hasan, T. (2010). Development and applications of photo-triggered theranostic agents. Advanced drug delivery reviews, 62(11), 1094-1124. doi: 10.1016/j.addr.2010.09.002.

[22] Gao, M., Deng, J., Liu, F., Fan, A., Wang, Y., Wu, H., ... & Zhao, Y. (2019). Triggered ferroptotic polymer micelles for reversing multidrug resistance to chemotherapy. Biomaterials, 223, 119486. doi: 10.1016/j.biomaterials.2019.119486.

[23] Jain, K. K. (2008). Drug delivery systems-an overview. Drug delivery systems, 1-50. doi:10.1007/978-1-59745-210-6_1.

[24] Saranjam, L., Fuguet, E., Nedyalkova, M., Simeonov, V., Mas, F., & Madurga, S. (2021). Prediction of Partition Coefficients in SDS Micelles by DFT Calculations. Symmetry, 13(9), 1750. doi: 10.3390/sym13091750.

[25] Allen, C., Maysinger, D., & Eisenberg, A. (1999). Nano-engineering block copolymer aggregates for drug delivery. Colloids and Surfaces B: Biointerfaces, 16(1-4), 3-27. doi:10.1016/S0927-7765(99)00058-2.

[26] Noble, A. (1993). Partition coefficients (n-octanol—water) for pesticides. Journal of Chromatography A, 642(1-2), 3-14. doi: 10.1016/0021-9673(93)80072-G.

[27] Magenheim, B., Levy, M. Y., & Benita, S. (1993). A new in vitro technique for the evaluation of drug release profile from colloidal carriers-ultrafiltration technique at low pressure. International journal of pharmaceutics, 94(1-3), 115-123. doi:10.1016/0378-5173(93)90015-8.

[28] Fuguet, E., Ràfols, C., Bosch, E., Abraham, M. H., & Roses, M. (2002). Solute–solvent interactions in micellar electrokinetic chromatography: III. Characterization of the selectivity of micellar electrokinetic chromatography systems. Journal of Chromatography A, 942(1-2), 237-248. doi:10.1016/S0021-9673(01)01383-8.

[29] Deeb, S. E., Iriban, M. A., & Gust, R. (2011). MEKC as a powerful growing analytical technique. Electrophoresis, 32(1), 166-183. doi: 10.1002/elps.201000398.

[30] Ingram, T., Storm, S., Kloss, L., Mehling, T., Jakobtorweihen, S., & Smirnova, I. (2013). Prediction of micelle/water and liposome/water partition coefficients based on molecular dynamics simulations, COSMO-RS, and COSMOmic. Langmuir, 29(11), 3527-3537. doi: 10.1021/la305035b.

[31] Ritter, E., Yordanova, D., Gerlach, T., Smirnova, I., & Jakobtorweihen, S. (2016). Molecular dynamics simulations of various micelles to predict micelle water partition equilibria with COSMOmic: Influence of micelle size and structure. Fluid Phase Equilibria, 422, 43-55. doi: 10.1016/j.fluid.2016.03.006.

[32] Yordanova, D., Ritter, E., Gerlach, T., Jensen, J. H., Smirnova, I., & Jakobtorweihen, S. (2017). Solute partitioning in micelles: Combining molecular dynamics simulations, COSMOmic, and experiments. The Journal of Physical Chemistry B, 121(23), 5794-5809. doi: 10.1021/acs.jpcb.7b03147.

[33] Turchi, M., Kognole, A. A., Kumar, A., Cai, Q., Lian, G., & MacKerell Jr, A. D. (2020). Predicting Partition Coefficients of Neutral and Charged Solutes in the Mixed SLES–Fatty Acid Micellar System. The Journal of Physical Chemistry B, 124(9), 1653-1664. doi: 10.1021/acs.jpcb.9b11199.

[34] Nedyalkova, M. A., Madurga, S., Tobiszewski, M., & Simeonov, V. (2019). Calculating the partition coefficients of organic solvents in octanol/water and octanol/air. Journal of chemical information and modeling, 59(5), 2257-2263. doi: 10.1021/acs.jcim.9b00212.

[35] Jones, M. R., Brooks, B. R., & Wilson, A. K. (2016). Partition coefficients for the SAMPL5 challenge using transfer free energies. Journal of computer-aided molecular design, 30, 1129-1138. doi: 10.1007/s10822-016-9964-6.

[36] Chou, J. T., & Jurs, P. C. (1979). Computer-assisted computation of partition coefficients from molecular structures using fragment constants. Journal of Chemical Information and Computer Sciences, 19(3), 172-178. doi: 10.1021/ci60019a013.

[37] Rustenburg, A. S., Dancer, J., Lin, B., Feng, J. A., Ortwine, D. F., Mobley, D. L., & Chodera, J. D. (2016). Measuring experimental cyclohexane-water distribution coefficients for the SAMPL5 challenge. Journal of computer-aided molecular design, 30, 945-958. doi: 10.1007/s10822-016-9971-7.

[38] Marenich, A. V., Cramer, C. J., & Truhlar, D. G. (2008). Perspective on foundations of solvation modeling: The electrostatic contribution to the free energy of solvation. Journal of Chemical Theory and Computation, 4(6), 877-887. doi: 10.1021/ct800029c.

[39] Marenich, A. V., Cramer, C. J., & Truhlar, D. G. (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. The Journal of Physical Chemistry B, 113(18), 6378-6396. doi: 10.1021/jp810292n.

[40] Stephens, P. J., Devlin, F. J., Chabalowski, C. F., & Frisch, M. J. (1994). Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. The Journal of physical chemistry, 98(45), 11623-11627. doi: 10.1021/j100096a001.

[41] Zhao, Y., & Truhlar, D. G. (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. Theoretical chemistry accounts, 120, 215-241. doi: 10.1007/s00214-007-0310-x.

[42] Fuguet, E., Ràfols, C., Bosch, E., & Rosés, M. (2003). Characterization of the solvation properties of sodium n-dodecyl sulfate micelles in buffered and unbuffered aqueous phases by solvatochromic indicators. Langmuir, 19(1), 55-62. doi: 10.1021/la026307o.

[43] Fuguet, E., Ràfols, C., Rosés, M., & Bosch, E. (2005). Critical micelle concentration of surfactants in aqueous buffered and unbuffered systems. Analytica Chimica Acta, 548(1-2), 95-100. doi: 10.1016/j.aca.2005.05.069.

[44] Mennucci, B., Cances, E., & Tomasi, J. (1997). Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: theoretical bases, computational implementation, and numerical applications. The Journal of Physical Chemistry B, 101(49), 10506-10517. doi: 10.1021/jp971959k.

[45] Baldwin, R. L. (2007). Energetics of protein folding. Journal of molecular biology, 371(2), 283-301. doi: 10.1016/j.jmb.2007.05.078.

[46] Buhmann, S. Y., & Welsch, D. G. (2007). Dispersion forces in macroscopic quantum electrodynamics. Progress in quantum electronics, 31(2), 51-130. doi: 10.1016/j.pquantelec.2007.03.001.

[47] Cramer, C. J., & Truhlar, D. G. (1999). Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. Chemical Reviews, 99(8), 2161–2200. doi:10.1021/cr960149m.

[48] Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. Journal of Cheminformatics, 4(8), 1–17. doi:10.1186/1758-2946-4-17.

[49] Cytochrome, P., Lewis, D. F. V., Sams, C., & Loizou, G. D. (2003). A Quantitative Structure – Activity Relationship Analysis on a Series of Alkyl Benzenes Metabolized by Human. Journal of Biochemical and Molecular Toxicology, 17(1), 47–52. doi:10.1002/jbt.10055.

[50] Cramer, C. J., & Truhlar, D. G. (1999). Implicit solvation models: equilibria, structure, spectra, and dynamics. Chemical Reviews, 99, 2161-2200. doi: 10.1039/C8CC06621C.

[51] Ogata, K., Hatakeyama, M., & Nakamura, S. (2018). Effect of atomic charges on octanol–water partition coefficient using alchemical free energy calculation. Molecules, 23(2), 425. doi: 10.3390/molecules23020425.

[52] Valko, K., Du, C. M. Y., Bevan, C. D., Reynolds, D. P., & Abraham, M. H. (2000). Rapid-Gradient HPLC Method for Measuring Drug Interactions with Immobilized Artificial Membrane: Comparison with Other Lipophilicity Measures. Journal of Chemistry, 89(8), 1085–1096. doi:10.1002/1520-6017(200008)89:8<1085: AID-JPS13>3.0.CO;2-N.

[53] Kibbey, C. E., Poole, S. K., Robinson, B. E. N., Jackson, J. D., & Durham, D. (2001). An Integrated Process for Measuring the Physicochemical Properties of Drug Candidates in a Preclinical Discovery Environment. Journal of Chemistry, 90(8), 1164–1175. doi: 10.1002/jps.1070.

[54] Smith, C. J., Perfetti, T. A., Garg, R., & Hansch, C. (2003). IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. Journal of Chemistry, 41, 807–817. doi:10.1016/S0278-6915(03)00021-8.

[55] Van Stee, L. L. P., et al. (2002). Use of semi-permeable membrane devices and solid-phase extraction for the wide-range screening of microcontaminants in surface water by GC-AED/MS. Journal of Chemistry, 36, 4455–4470. doi:10.1016/S0043-1354(02)00177-X.

[56] Bas, D., Dorison-Duval, D., Moreau, S., Bruneau, P., & Chipot, C. (2002). Rational determination of transfer free energies of small drugs across the water− oil interface. Journal of medicinal chemistry, 45(1), 151-159. doi:10.1021/jm010289a.

[57] Stéen, E. J. L., Nyberg, N., Lehel, S., Andersen, V. L., Di Pilato, P., Knudsen, G. M., ... & Herth, M. M. (2017). Development of a simple proton nuclear magnetic resonance-based procedure to estimate the approximate distribution coefficient at physiological pH (log D7. 4): Evaluation and comparison to existing practices. Bioorganic & Medicinal Chemistry Letters, 27(2), 319-322. doi: 10.1016/j.bmcl.2016.11.048.

[58] Miyake, Y., Yumoto, T., Kitamura, H., & Sugimoto, T. (2002). Solubilization of organic compounds into as-synthesized spherical mesoporous silica. Physical Chemistry Chemical Physics, 4(12), 2680-2684. doi:10.1039/B200074C.

[59] Huang, H., Wang, X., Shao, Y., Chen, D., Dai, X., & Wang, L. (2003). QSAR for prediction of joint toxicity of substituted phenols to tadpoles (Rana japonica). Bulletin of environmental contamination and toxicology, 71, 1124-1130. doi: 10.1007/s00128-003-8790-4.

[60] Xie, Y. J., Liu, H., Liu, H. X., Zhai, Z. C., & Wang, Z. Y. (2008). Determination of solubilities and n-octanol/water partition coefficients and QSPR study for substituted phenols. Bulletin of environmental contamination and toxicology, 80, 319-323. doi: 10.1007/s00128-008-9369-x.

[61] Poerschmann, J., Trommler, U., Nyplova, P., Morgenstern, P., & Górecki, T. (2008). Complexation–flocculation of organic contaminants by the application of oxyhumolite-based humic organic matter. Chemosphere, 70(7), 1228-1237. doi: 10.1016/j.chemosphere.2007.08.004.

[62] Brown, R. S., Akhtar, P., Åkerman, J., Hampel, L., Kozin, I. S., Villerius, L. A., & Klamer, H. J. (2001). Partition controlled delivery of hydrophobic substances in toxicity tests using poly (dimethylsiloxane)(PDMS) films. Environmental science & technology, 35(20), 4097-4102. doi:10.1021/es010708t.

[63] Liu, J., Nile, S. H., Xu, G., Wang, Y., & Kai, G. (2021). Systematic exploration of Astragalus membranaceus and Panax ginseng as immune regulators: insights from the comparative biological and computational analysis. Phytomedicine, 86, 153077. doi: 10.1016/j.phymed.2019.153077.

[64] Schultz, T. W. (1999). Structure− toxicity relationships for benzenes evaluated with tetrahymena pyriformis. Chemical research in toxicology, 12(12), 1262-1267.doi:10.1021/tx9900730.

[65] Chang, Y. T., Chang, F. Y., Chen, Y. K., Lee, C. J., & Yang, C. S. (2009). Water-Dragging Ability of Aromatic Compounds in Octanol-Water Systems: A Quantitative Approach by Spectra Deconvolution. Journal of the Chinese Chemical Society, 56(2), 279-288. doi:10.1002/jccs.200900041.

[66] Lodge, K. B., & Egyepong, E. J. (2010). Evidence for self-association of nonionic and other organic solutes in liquid phases comprising 1-octanol and water. The Journal of Physical Chemistry A, 114(15), 5132-5140. doi:10.1021/jp907752w.

[67] Poduval, R., Kurzątkowska, K., Stobiecka, M., Dehaen, W. F. A., Dehaen, W., Radecka, H., & Radecki, J. (2010). Systematic study of interaction of the neutral form of anilines with undecylcalix [4] resorcinarene derivatives by means of potentiometry. Supramolecular Chemistry, 22(7-8), 413-419. doi: 10.1080/10610278.2010.486437.

[68] Chen, Z., & Weber, S. G. (2007). High-throughput method for lipophilicity measurement. Analytical chemistry, 79(3), 1043-1049. doi:10.1021/ac061649a.

[69] Ruelle, P., & Kesselring, U. W. (1998). The hydrophobic effect. 3. A key ingredient in predicting n-octanol–water partition coefficients. Journal of pharmaceutical sciences, 87(8), 1015-1024. doi: 10.1021/js9703030.

[70] Toulmin, A., Wood, J. M., & Kenny, P. W. (2008). Toward prediction of alkane/water partition coefficients. Journal of medicinal chemistry, 51(13), 3720-3730. doi: 10.1021/jm701549s.

[71] Byrns, G. (2001). The fate of xenobiotic organic compounds in wastewater treatment plants. Water Research, 35(10), 2523-2533. doi:10.1016/S0043-1354(00)00529-7.

[72] Burgess, D. A., & Rae, I. D. (1977). Oxidation of alkyl groups accompanying the Zinin reduction of nitroarenes. Australian Journal of Chemistry, 30(4), 927-931. doi:10.1071/CH9770927.

[73] Zafrani, Y., Yeffet, D., Sod-Moriah, G., Berliner, A., Amir, D., Marciano, D., ... & Saphier, S. (2017). Difluoromethyl bioisostere: examining the "lipophilic hydrogen bond donor" concept. Journal of Medicinal Chemistry, 60(2), 797-804. doi: 10.1021/acs.jmedchem.6b01691.

[74] Schneider, G. M., Kautz, C. B., & Tuma, D. (2000). Physico-chemical principles of supercritical fluid science. In Supercritical Fluids: Fundamentals and Applications (pp. 31-68). Dordrecht: Springer Netherlands. doi:10.1007/978-94-011-3929-8_2.

[75] Fagundez, C., Sellanes, D., Peña, S., Scarone, L., Aguiar, A. C., de Souza, J. O., ... & Serra, G. L. (2018). Synthesis, profiling, and in vivo evaluation of cyclopeptides containing N-methyl amino acids as antiplasmodial agents. ACS Medicinal Chemistry Letters, 10(1), 137-141.doi:10.1021/acsmedchemlett.8b00543.

[76] Soulsby, D. (2019). Band-selective excitation NMR spectroscopy and quantitative time-domain analysis using Complete Reduction to Amplitude-Frequency Table (CRAFT) to determine distribution coefficients during drug development. Magnetic Resonance in Chemistry, 57(11), 953-960. doi:10.1002/mrc.4888.

[77] Chen, C. P., Chen, C. C., Huang, C. W., & Chang, Y. C. (2018). Evaluating molecular properties involved in transport of small molecules in stratum corneum: A quantitative structure-activity relationship for skin permeability. Molecules, 23(4), 911. doi:10.3390/molecules23040911.

[78] Tartaglia, A., Locatelli, M., Kabir, A., Furton, K. G., Macerola, D., Sperandio, E., ... & Samanidou, V. F. (2019). Comparison between exhaustive and equilibrium extraction using different SPE sorbents and sol-gel carbowax 20M coated FPSE media. Molecules, 24(3), 382. doi: 10.3390/molecules24030382.

[79] Nguyen, A., Top, S., Pigeon, P., Vessières, A., Hillard, E. A., Plamont, M. A., ... & Jaouen, G. (2009). Synthesis and structure–activity relationships of ferrocenyl tamoxifen derivatives with modified side chains. Chemistry–A European Journal, 15(3), 684-696. doi:10.1002/chem.200801108.

[80] Haraguchi, T., Uchida, T., Yoshida, M., Kojima, H., Habara, M., & Ikezaki, H. (2018). The utility of the artificial taste sensor in evaluating the bitterness of drugs: correlation with responses of human TASTE2 receptors (hTAS2Rs). Chemical and Pharmaceutical Bulletin, 66(1), 71-77. doi:10.1248/cpb.c17-00619.

[81] Jin, Y., Qi, Y., Tang, C., & Shao, B. (2021). Hierarchical micro-and mesoporous metal–organic framework-based magnetic nanospheres for the nontargeted analysis of chemical hazards in vegetables. Journal of Materials Chemistry A, 9(14), 9056-9065. doi:10.1039/D1TA00120E.

[82] Gonec, T., Kralova, K., Pesko, M., & Jampilek, J. (2017). Antimycobacterial N-alkoxyphenylhydroxynaphthalenecarboxamides affecting photosystem II. Bioorganic & Medicinal Chemistry Letters, 27(9), 1881-1885. doi: 10.1016/j.bmcl.2017.03.050.

[83] Locke, M. A., Zablotowicz, R. M., Steinriede, R. W., & Kingery, W. L. (2007). Degradation and sorption of fluometuron and metabolites in conservation tillage soils. Journal of agricultural and food chemistry, 55(3), 844-851. doi:10.1021/jf062070g.

[84] Guardo, A. D., Williams, R. J., Matthiessen, P., Brooke, D. N., & Calamari, D. (1994). Simulation of pesticide runoff at Rosemaund Farm (UK) using the SoilFug model. Environmental Science and Pollution Research, 1, 151-160. doi: 10.1007/BF02986938.

[85] Hu, J. Y., Aizawa, T., Ookubo, Y., Morita, T., & Magara, Y. (1998). Adsorptive characteristics of ionogenic aromatic pesticides in water on powdered activated carbon. Water Research, 32(9), 2593-2600. doi:10.1016/S0043-1354(98)00014-1.

[86] Liu, J., & Qian, C. (1995). Hydrophobic coefficients of s-triazine and phenylurea herbicides. Chemosphere, 31(8), 3951-3959. doi:10.1016/0045-6535(95)00267-C.

[87] Dorosti, N., Delfan, B., Gholivand, K., & Valmoozi, A. A. E. (2016). Synthesis, crystal structure, biological evaluation, electronic aspects of hydrogen bonds, and QSAR studies of some new N- (substituted phenylurea) diazaphosphore derivatives as anticancer agents. Medicinal Chemistry Research, 25, 769-789. doi: 10.1007/s00044-016-1527-9.

[88] Mibu, N., Yokomizo, K., Yuzuriha, A., Otsubo, M., Kawaguchi, Y., Sano, M., ... & Sumoto, K. (2017). Antiviral activities of some new 2, 4, 6-trisubstituted 1, 3, 5-triazines having alkoxy and/or alkylamino groups. Heterocycles, 94, 1653-1677. doi: 10.3987/COM-17-13735.

CHAPTER
V

PREDICTION OF THE PARTITION COEFFICIENT FOR FLEXIBLE MOLECULES

Chapter V

Prediction of the partition coefficient for flexible molecules

---

5.1 Flexible molecules

The flexibility of molecules refers to their ability to undergo conformational changes or deformations while maintaining their overall chemical structure. This flexibility is primarily determined by the types of chemical bonds present within the molecule and the presence of certain functional groups or substituents.

Typically, organic molecules predominantly contain two types of chemical bonds: sigma (σ) bonds and pi (π) bonds. Sigma bonds are created when atomic orbitals overlap along the bond axis and permit unrestricted rotation around that axis. Consequently, molecules linked by sigma bonds can internally rotate without necessitating bond breakage [1,2].

π bonds, on the other hand, are formed by the sideways overlap of atomic orbitals above and below the bond axis. Pi bonds restrict rotation around the bond axis since it would involve breaking the pi bond. However, pi bonds can be relatively easily broken and reformed, allowing for conformational changes in the molecule [3].

In addition, pi bonds, and other factors can contribute to the flexibility of molecules. These include the presence of single bonds (allowing rotation), the presence of flexible linkers or chains, and the absence of rigid structures or bulky substituents [4,5].

On the other hand, certain factors can restrict the flexibility of molecules. These include the presence of multiple bonds (such as double or triple bonds), the presence of rigid structures or bulky substituents that hinder rotation or conformational changes, and the formation of stable ring structures [6,7].

It's important to note that the flexibility of a molecule can have significant implications for its behavior, properties, and interactions. For example, flexible molecules may exhibit different

conformations or shapes in different environments, which can influence their reactivity, binding affinity, and biological activity [8]. Understanding and predicting the flexibility of molecules is therefore crucial in various fields, including chemistry, biochemistry, and drug design [9-11].

5.1.1 Effect of flexibility on molecular Properties

Molecules have a significant impact on the properties of substances. The specific arrangement, composition, and interactions of molecules dictate various characteristics and behaviors of materials. Here are some key effects of molecules on properties:

- **Physical Properties:** Molecules determine the physical properties of substances, such as melting point, boiling point, density, viscosity, and optical properties (color, transparency, refractive index). For example, molecules' size, shape, and polarity influence their interactions with electromagnetic radiation, leading to variations in the absorption, transmission, and reflection of light [12].

- ✓ Melting and boiling point: The melting and boiling points of flexible molecules are often lower than those of rigid molecules of the same size and composition. This is because of flexible molecules' freedom of motion and ability to modify their conformation, which makes it simpler for them to overcome intermolecular interactions and change states from solid to liquid or liquid to gas [13].

- **Chemical Reactivity:** The structure and composition of molecules determine their chemical reactivity. Molecules with functional groups or specific bonding patterns exhibit unique chemical reactions and can participate in various chemical transformations. The presence of specific atoms or groups can confer acidity, basicity, or reactivity toward specific reagents or environmental conditions [14,15].

  The flexibility of molecules can have a profound effect on their chemical reactivity in several ways:

- ✓ Conformational Effects: Flexible molecules can adopt different conformations due to rotation around single bonds or other internal motions. These different conformations can impact the accessibility of reactive sites within the molecule. Certain conformations may expose reactive functional groups, making them more accessible for chemical reactions,

while other conformations may shield or hinder these groups, reducing their reactivity [16,17].

✓ Steric Effects: Flexibility can lead to steric hindrance, which refers to the spatial interference between atoms or groups within a molecule. In flexible molecules, different conformations can result in varying degrees of steric hindrance. High steric hindrance can impede the approach of other molecules or reactants, reducing the reactivity of the molecule. Conversely, low steric hindrance can enhance reactivity by allowing easier access to reactive sites [18].

✓ Reaction Pathways: Flexibility can influence the preferred reaction pathways of a molecule. Different conformations or arrangements can result in distinct energy landscapes, leading to different transition states and reaction mechanisms. This can result in changes in reaction rates, selectivity, and the formation of different reaction products [19].

To summarize, the flexibility of molecules plays a vital role in determining their chemical reactivity. It affects reactivity by influencing the accessibility of reactive sites, introducing steric hindrance, modulating reaction pathways, interacting with solvents, and exhibiting dynamic behavior during reactions. Understanding the flexibility of molecules is essential for accurately predicting and explaining their reactivity in different chemical processes [20].

- **Solubility:** Molecules play a role in determining the solubility of substances in diverse solvents. The flexibility of a molecule can indeed affect its solubility [21,22]. Here's how flexibility influences solubility:

✓ Molecular Size: Flexible molecules can adopt different conformations, leading to changes in their overall size and shape. The size of a molecule directly impacts its solubility in a solvent. In general, smaller and more compact conformations of a flexible molecule tend to increase solubility, as they can more easily fit into the solvent's molecular arrangement [23] .

✓ Solvation Energy: The energy required to solvate a flexible molecule can vary depending on its conformation. Some conformations may experience favorable solvation interactions, resulting in lower solvation energy and higher solubility. Conversely, other conformations

may experience less favorable solvation interactions, leading to higher solvation energy and lower solubility [24,25].

✓ Molecular Surface Area: The flexibility of a molecule can influence its surface area exposed to the solvent. Different conformations may expose larger or smaller surface areas to the solvent molecules. Generally, larger surface areas increase the opportunities for solvent-solute interactions and can enhance solubility [26].

- **Electrical Conductivity:** The effect of flexibility on the electrical conductivity of molecules depends on their specific characteristics and the mechanisms through which they conduct electricity. Here are a few ways in which the flexibility of molecules can impact electrical conductivity [27]:

✓ π-Conjugated Systems: In organic molecules with extended π-conjugated systems, such as conjugated polymers or aromatic compounds, flexibility can affect the delocalization of π-electrons. Rigid, planar conformations often promote efficient π-electron delocalization, leading to enhanced electrical conductivity. On the other hand, increased molecular flexibility can disrupt the conjugation and reduce the overall electrical conductivity [28].

✓ Charge Transport Pathways: The flexibility of molecules can influence the pathways through which charge carriers (e.g., electrons or ions) move within the material. In some cases, flexible molecules can allow for the formation of well-defined pathways that facilitate charge transport. Conversely, excessive flexibility can lead to disordered molecular arrangements, hindering the efficient movement of charge carriers and reducing electrical conductivity [29].

✓ Interactions and Packing: Flexible molecules can exhibit different types of intermolecular interactions and packing arrangements in the solid state. The strength and orientation of these interactions can significantly impact the efficiency of charge transfer between neighboring molecules. Optimal packing arrangements with close intermolecular contacts may enhance electrical conductivity, while excessive flexibility can disrupt these contacts and reduce conductivity [30].

- **Mechanical Properties:** Flexible molecules can have a significant impact on the mechanical properties of materials. The flexibility of a molecule refers to its ability to

undergo conformational changes or deformations in response to external forces. These conformational changes can affect various mechanical properties, including elasticity, strength, and toughness [19]. Here are a few ways in which flexible molecules can influence mechanical properties:

✓ Elasticity: Flexible molecules can enhance the elasticity of materials. When subjected to an external force, the flexible molecules can deform, stretch or bend to accommodate the stress. This deformation allows the material to absorb energy and return to its original shape when the force is removed, resulting in high elasticity. Examples of flexible molecules that contribute to elasticity include rubber polymers, which can stretch and recoil [31,32].

✓ Strength: The presence of flexible molecules can improve the strength of materials. When a force is applied to a material, flexible molecules can undergo conformational changes, redistributing the stress and preventing the propagation of cracks or fractures. This ability to dissipate stress and resist fracture can enhance the overall strength of the material [33].

✓ Toughness: Flexible molecules can also increase the toughness of materials. Toughness refers to a material's ability to withstand deformation and absorb energy without fracturing. Flexible molecules can deform and reorient in response to stress, allowing the material to absorb energy and resist fracture. This property is particularly important in applications where impact resistance is necessary [33].

✓ Ductility: Flexible molecules can contribute to the ductility of materials, which is the ability to undergo plastic deformation without breaking. When a material contains flexible molecules, it can undergo extensive plastic deformation, allowing it to be drawn into wires or shaped without fracture. Metals, for example, exhibit high ductility due to the presence of flexible metallic bonds [34].

• **Magnetic Properties:** Flexible molecules themselves do not directly affect the magnetic properties of materials. Magnetic properties are primarily determined by the arrangement and behavior of magnetic moments within a material, which are influenced by factors such as the presence of magnetic elements, crystal structure, and external magnetic fields [35]. However, the presence of flexible molecules can indirectly impact the magnetic properties in the following ways:

- ✓ Encapsulation and stabilization: Flexible molecules can act as ligands or complexing agents, surrounding and stabilizing magnetic species or nanoparticles. This encapsulation can prevent aggregation or oxidation of magnetic particles, maintaining their magnetic properties and enhancing their stability [36,37].

- ✓ Alignment and orientation: Flexible molecules with anisotropic properties, such as liquid crystals, can align or orient magnetic particles or domains within a material. This alignment can affect the overall magnetic behavior of the material, such as enhancing the magnetic anisotropy or changing the magnetic susceptibility in specific directions [38].

- ✓ Hybrid materials: Flexible molecules can be combined with magnetic materials to form hybrid systems. For example, flexible polymers or organic ligands can be functionalized with magnetic nanoparticles or paramagnetic centers. The interactions between the flexible molecules and the magnetic components can influence the overall magnetic behavior, such as magnetic ordering, magnetic transition temperatures, or magnetoresistance [39,40].

- **Biological Activity:** Flexible molecules can have a significant impact on biological activity, particularly in the context of drug design and molecular interactions with biological targets. The flexibility of molecules plays a crucial role in their ability to bind to target proteins, enzymes, receptors, or other biomolecules. Here are some ways in which flexible molecules can affect biological activity:

- ✓ Binding affinity: The flexibility of a molecule can influence its binding affinity to a biological target. Flexible molecules can adopt different conformations or undergo conformational changes, allowing them to interact with target molecules in a complementary manner [41,42]. This flexibility enables the molecule to optimize its interactions with the target, enhancing binding affinity and potentially leading to stronger biological activity.

- ✓ Selectivity: Flexible molecules can exhibit selectivity in their interactions with biological targets. Different conformations or conformer populations of a flexible molecule may interact differently with various target proteins or receptors. This selectivity can be harnessed in drug design to specifically target a particular biological process or pathway while minimizing off-target effects [43].

✓ Induced fit: Flexible molecules can undergo conformational changes upon binding to a biological target. This phenomenon, known as induced fit, allows the molecule to adapt its shape to better fit the target's binding site. Induced fit can improve the binding affinity and specificity of the molecule, enhancing its biological activity by optimizing the molecular interactions between the ligand and the target [44].

✓ Solubility and bioavailability: The flexibility of molecules can affect their solubility and bioavailability, influencing their ability to reach and interact with their biological targets. Flexible molecules can exhibit improved solubility in biological fluids, facilitating their absorption and distribution in the body. This improved solubility can enhance their bioavailability and increase the likelihood of achieving therapeutic effects.

It's important to note that the specific effects of flexible molecules on biological activity will depend on the nature of the molecule, the target, and the surrounding biological environment. Molecular flexibility is a key consideration in drug design and optimization, as it allows for the development of molecules with improved binding characteristics, selectivity, and therapeutic potential [45,46].

Comprehending the connection between molecular structure and properties is crucial across multiple scientific fields, such as chemistry, materials science, biology, and pharmacology [47]. This understanding facilitates the design and refinement of molecules and materials to possess desired properties tailored for specific applications, leading to advancements in various technological domains.

5.1.2 Relationship between partition coefficient and conformational properties

The conformation of a molecule refers to its spatial arrangement or shape, which can vary due to the rotation around single bonds or the flexibility of certain parts of the molecule [48]. The conformational flexibility can affect the partition coefficient of a molecule in the following ways:

- **Surface Area:** The conformation of a molecule can alter its overall surface area exposed to the solvent phases. Generally, a more extended conformation with a larger

surface area tends to have a higher partition coefficient, as it can interact more favorably with the nonpolar solvent [49].

- **Molecular Interactions:** The conformational changes in a molecule can affect its ability to form specific intermolecular interactions. For example, the presence of hydrogen bonding or hydrophobic interactions can influence the partition coefficient. Changes in conformation can alter the accessibility or strength of these interactions, leading to variations in partition coefficients [50,51].

- **Solvent-Preferring Groups:** Certain regions of a molecule may exhibit a preference for a specific solvent phase. For example, polar functional groups tend to favor polar solvents, while nonpolar groups have a greater affinity for nonpolar solvents [52]. The conformation of the molecule can expose or shield these regions, thereby impacting the partition coefficient.

It is important to note that predicting the exact relationship between the conformation of a molecule and its partition coefficient is challenging due to the complexity of intermolecular interactions and solvent effects. Experimental measurements and computational simulations are often employed to study and understand the specific conformational preferences and their impact on the partitioning behavior of molecules.

Overall, the conformation of a molecule can influence its partition coefficient through various mechanisms, including changes in surface area, molecular interactions, and the exposure of solvent-preferring groups. Understanding these relationships is valuable for predicting and optimizing the distribution and behavior of molecules in different solvent environments.

5.2 Determination of Partition Coefficients for a set of flexible molecules

The motivation behind this study was to investigate the impact of molecular flexibility on physicochemical properties and analyze the conformational space and properties dependent on conformers to explore the lipophilicity space of a diverse database of molecules with varying levels of flexibility [53,54].

The partition coefficient plays the role of a crucial factor among the numerous physicochemical properties that characterize a specific molecule. This significance is underscored by the basic

requirement in drug development, where candidates must exhibit favorable lipophilic properties to satisfy the measures for clinical suitability.

As the field of chemistry advances, our comprehension of molecular structure develops, containing the intricate relationship between geometric attributes and physicochemical properties [2]. This indicates that accurately describing the flexibility of molecules and considering how it affects property variations can improve the accuracy of predicting those properties. By including this previously ignored information, property predictions can be more reliable [55].

The concept of property space is based on the understanding that a flexible molecule, capable of taking on multiple conformations, will show various values for a specific physicochemical property that depends on its conformation. As a result, each conformation corresponds to a specific property value, and the combination of these values defines the property space. In this dynamic viewpoint, an average value can represent a molecular property.

The average value of a property, particularly when weighted, provides a more significant indication compared to a value specific to a particular conformer, even if that conformer is the most likely (lowest-energy) or the one with bioactivity [56].

Recent assessments of methods used for predicting log P values have indicated that 3D conformer-based systems are less effective compared to simpler fragmental methods. However, it is important to acknowledge that 3D structures encompass more comprehensive molecular information in comparison to methods that rely on 2D structures as input [57]. In this work, compounds with 3D structures include flexible molecules, while fragment methods do contain not flexible compounds.

5.2.1 Generation of a set of molecular conformation

Gabedit is a software program used for drawing and visualizing molecular structures. It is free and open-source software that provides a user-friendly interface for constructing and manipulating molecular models. Gabedit supports a wide range of molecular file formats, allowing users to import and export structures from various computational chemistry software packages. It also offers additional features such as displaying molecular orbitals, calculating molecular properties, and performing basic molecular simulations.

Gabedit software offers a notable advantage in its capacity to depict various structures of a flexible compound. This feature proves particularly valuable when working with molecules that can adopt multiple conformations or configurations. Through Gabedit, users can effectively explore the conformational space of flexible compounds by visually representing different structures. This functionality greatly assists in analyzing molecular behavior, comprehending the interplay between structure and properties, and investigating the impacts of diverse conformations on molecular interactions and reactions. By facilitating the representation of a flexible compound's diverse structures, Gabedit significantly enhances the software's versatility and usefulness in numerous computational chemistry and research endeavors.

The process of preparing the molecule in Gabedit software is as follows:

- The molecule is first drawn using the Gabedit molecule editor.
- The molecular structure is then optimized using an optimization algorithm, likely within a quantum chemistry software or molecular modeling program.
- After optimization, the Amber potential (a force field commonly used in molecular dynamics simulations) is selected to describe the interatomic interactions.
- Select the molecular dynamics conformationally.
- After, set the number of selected geometries to 10.
- The simulation temperature is set to 1000 Kelvin, and the simulation time is set to 50 picoseconds (ps), indicating the duration for which the molecular dynamics simulation will be run.
- The temperature is controlled using the Andersen method.
- Other simulation parameters, not mentioned explicitly, are left at their default values.

Finally, the initial file, referred to the input file contains all the necessary information for the molecular dynamic simulation.

5.2.2 Computational conditions

Like the previous chapter, partition coefficients were estimated by using Density Functional Theory (DFT) with the B3LYP method and a 6-31++G** basis set to optimize the geometries of all compounds. Additionally, the calculations involved employing the continuum solvation model

based on density (SMD) and were conducted using the electronic structure program Gaussian 16 [58].
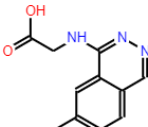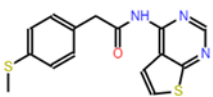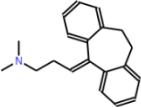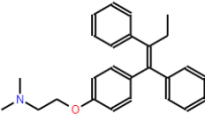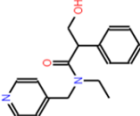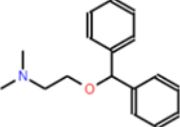
The initial molecular structures were generated using Gabedit, a molecule editor that is freely available and compatible across different platforms. Solvation energies were considered only from minimizations that resulted in all positive frequencies, ensuring stability. Harmonic vibrational frequencies were calculated for all compounds to characterize their vibrational behavior. Thermochemical values, such as enthalpy and entropy, were determined at a standard temperature of 298.15 K and pressure of 1.0 atm offering a comprehensive suite for predicting various molecular properties.

5.2.3 *Comparison of calculated and experimental Octanol/Water partition coefficient of flexible molecules*
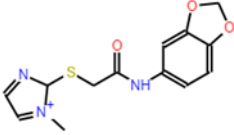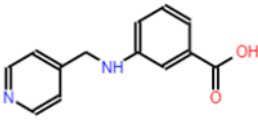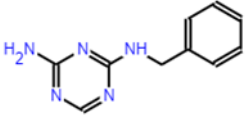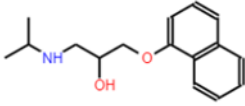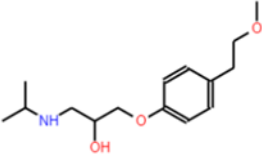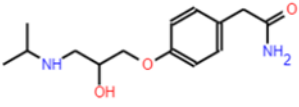
To evaluate the reliability of the computed partition coefficients in different solvents, a comparative analysis was conducted between the predicted values and the corresponding experimental data for the octanol and cyclohexane systems. The compounds being investigated are presented in Table 1. The experimental partition coefficient values are shown for octanol and cyclohexane solvents in Table 2 and Table 3, respectively. These experimental results serve as reference points for validating the accuracy and predictive ability of the computed partition coefficients. By comparing the predicted and experimental values, insights can be gained into the suitability and effectiveness of the computational methods employed in estimating partition coefficients in various solvents.
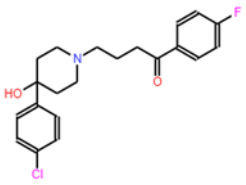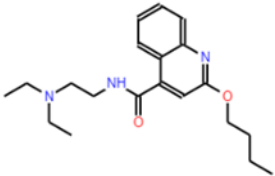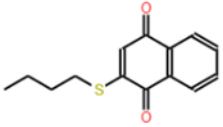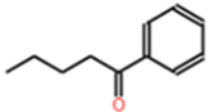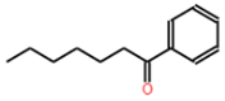
Table 1. Structure, formula, and number of the conformations of molecules employed.

Molecules	Structure	Formula	Number of Conformations
2-[(6-methylquinazolin-4-yl) amino] acetic acid		$C_{11}H_{11}N_3O_2$	7
2-(4-methylsulfanylphenyl)-N-thieno[2,3-d] pyrimidin-4-ylacetamide		$C_{15}H_{13}N_3OS_2$	2
Amitriptyline		$C_{20}H_{23}N$	4
Tamoxifen		$C_{26}H_{29}ON$	5
Tropicamide		$C_{17}H_{20}N_2O_2$	7
2-diphenyl methoxy-N, N-dimethyl ethanamine		$C_{17}H_{21}NO$	4
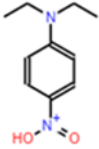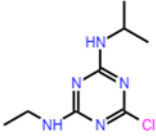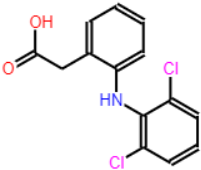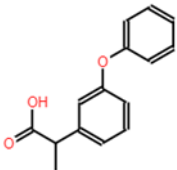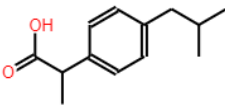
N-(1,3-benzodioxol-5-yl)- 2-(1-methylimidazol-2-yl) sulfanyl acetamide		$C_{13}H_{13}N_3O_3S$	4
3-((4-pyridinylmethyl) amino) benzoic acid		$C_{13}H_{12}N_2O_2$	3
2-Benzylamino-6- aminotriazine		$C_{10}H_{11}N_5$	2
Propranolol		$C_{16}H_{21}O_2N$	10
(RS)-metoprolol		$C_{15}H_{25}NO_3$	10
(RS)-atenolol		$C_{14}H_{22}N_2O_3$	10
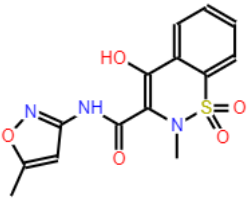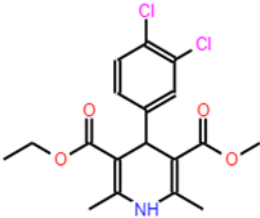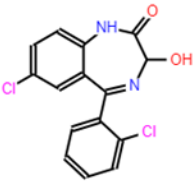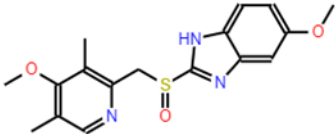
Haloperidol		$C_{21}H_{23}ClFNO_2$	6
Dibucain		$C_{20}H_{29}N_3O_2$	10
2-butylmercapto-1,4-naphthoquinone		$C_{14}H_{14}O_2S$	7
1-phenylpentan-1-one		$C_{11}H_{14}O$	7
1-phenylheptan-1-one		$C_{13}H_{18}O$	10

N, N-diethyl-4-nitroaniline		$C_{10}H_{14}N_2O_2$	2
Atrazine		$C_8H_{14}ClN_5$	5
Diclofenac		$C_{14}H_{11}Cl_2NO_2$	8
Fenoprofen		$C_{15}H_{14}O_3$	4
Ibuprofen		$C_{13}H_{18}O_2$	8

Isoxicam		$C_{14}H_{13}N_3O_5S$	5
Felodipine		$C_{18}H_{19}Cl_2NO_4$	7
Lorazepam		$C_{15}H_{10}Cl_2N_2O_2$	5
Omeprazole		$C_{17}H_{19}N_3O_3S$	7
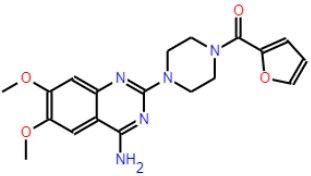
| Prazosin | | $C_{19}H_{21}N_5O_4$ | 5 |

Table 2. List of experimental partition coefficients of compounds in octanol solvent.

Molecule	log $P_{octanol/water}$
2-[(6-methylquinazolin-4-yl) amino] acetic acid	-1.80 [59]
Amitriptyline	4.90 [60]
Tamoxifen	6.20 [61]
Tropicamide	1.28 [61]
2-diphenyl methoxy-N, N-dimethyl ethanamine	3.38
Propranolol	3.3 [62]
(RS)-metoprolol	2.2 [63]
(RS)-atenolol	0.43 [64]
Haloperidol	4.3 [65]
Dibucain	3.7 [66]
1-phenylpentan-1-one	3.15
1-phenylheptan-1-one	4.13
N, N-diethyl-4-nitroaniline	3.55 [67]
Atrazine	2.61 [68]
Diclofenac	4.51 [69]
Fenoprofen	3.1 [70]
Ibuprofen	3.97 [69]

| Glipizide | 1.64[71] |
| Isoxicam | 1.38[71] |

Table 3. List of experimental partition coefficients of compounds in cyclohexane solvent.

Molecule	log $P_{cyclohexane/water}$
2-[(6-methylquinazolin-4-yl) amino] acetic acid	-2.2 [72]
2-(4-methylsulfanylphenyl)-N-thieno[2,3-d] pyrimidin-4-ylacetamide	0.2 [72]
Amitriptyline	1.6 [72]
Tamoxifen	2.5 [72]
2-diphenylmethoxy-N, N-dimethyl ethanamine	0.6 [72]
N-(1,3-benzodioxol-5-yl)-2-(1-methylimidazol-2-yl) sulfanyl acetamide	-0.9 [73]
3-((4-pyridinylmethyl) amino) benzoic acid	-1.7 ± 0.6 [73]
2-Benzylamino-6-aminotriazine	-1.9 ± 0.8 [72]
Propranolol	-1.3 ±0.3 [72]
(RS)-metoprolol	-2.8 ± 0.3 [73]
(RS)-atenolol	-2.2 ± 0.3 [73]
Tamoxifen	2.5 ± 0.3 [73]
Haloperidol	-0.0 ± 0.3 [73]
Dibucain	0.7 ±0.3 [73]
2-butylmercapto-1,4-naphthoquinone	3.43 [74]
Atrazine	1.4 [73]

Figure 1(a) shows the correlation between the average partition coefficient data and the experimental partition coefficient in the octanol solvent (Table 2) also, Figure 1(b) illustrates the relationship between the minimum partition coefficient data and the corresponding experimental partition coefficients obtained [73-75]. The linear correlation equation, R-squared (R^2), Root mean squared error (RMSE), Mean absolute error (MAE), and Average signed error (ASE) for average and minimum partition coefficients are shown in Table 5.

As seen in Table 5, it is clear that the highest correlation is observed between the experimental data and the average partition coefficient. The $R^2 \geq 0.78$, indicates a strong correlation between the two variables.

Additionally, the slope of the linear regression line is determined to be 0.89, suggesting a nearly one-to-one correspondence between the log $P_{Experimental}$ and average log $P_{Calculation}$ values.

Moreover, MAE is calculated to be 0.64. The MAE provides an estimate of the average absolute difference between the experimental and average partition coefficient values. In this case, the relatively low MAE value indicates that the average partition coefficient is generally in close agreement with the experimental data.

In a comparison between Figures 1(a) and 1(b), it becomes obvious that the average partition coefficient value shows a higher degree of correlation than the minimum partition coefficient value. This means that the average partition coefficient offers a more precise depiction of the relationship with the experimental data when compared to the minimum log P value. Thus, it seems that the consideration of molecular flexibility must be considered for calculations of partition coefficients.

This happens because flexible molecules can change their structure over time, leading to dynamic behaviour within the partitioning medium. These structure changes result in different relations between the molecule and the solvent. By evaluating the average partition coefficient, which takes into account the contributions from these various conformations, we gain a comprehensive understanding of the molecule's behaviour in the solvent. This integration of different conformations allows us to better understanding how the molecule moves and interacts with the solvent, resulting in a stronger correlation between the experimental data and its behaviour.

It's important to note that the specific factors influencing the correlation between the average and minimum partition coefficients can depend on the characteristics of the flexible molecule, the solvent system, and the specific interactions involved.

Table 4. List of average and minimum log P calculations for B3LYP in octanol solvent.

Molecule	log $P_{o/w}$	Average	Minimum
2-[(6-methylquinazolin-4-yl) amino] acetic acid	1.8	0.79	1.43
Amitriptyline	4.9	4.83	4.50
Tamoxifen	6.2	5.71	5.42
2-(4-methylsulfanylphenyl)-N-thieno[2,3-d] pyrimidin-4-ylacetamide	NA	0.86	0.24
Tropicamide	1.34	1.59	3.61
2-diphenylmethoxy-N,N-dimethylethanamine	3.38	3.33	3.16
N-(1,3-benzodioxol-5-yl)-2-(1-methylimidazol-2-yl) sulfanyl acetamide	NA	0.43	0.85
NO NAME	NA	3.50	3.14
propranolol	3.30	3.40	2.50
(RS)-metoprolol	2.20	2.21	1.53
(RS)-atenolol	0.43	0.28	1.09
tamoxifen	6.20	5.15	3.47
haloperidol	4.30	3.90	3.35
Dibucain	3.70	2.99	2.03
2-butylmercapto-1,4-naphthoquinone	NA	3.24	2.88
3-((4-pyridinylmethyl) amino) benzoic acid	NA	0.04	0.61
2-Benzylamino-6-aminotriazine	NA	1.82	1.69

Figure 1. Comparison of the Experimental partition coefficient of Octanol solvent with **a)** the average log P calculated for B3LYP and **b)** the minimum log P calculated for B3LYP and **c)** compassion between average and minimum partition coefficient calculated.
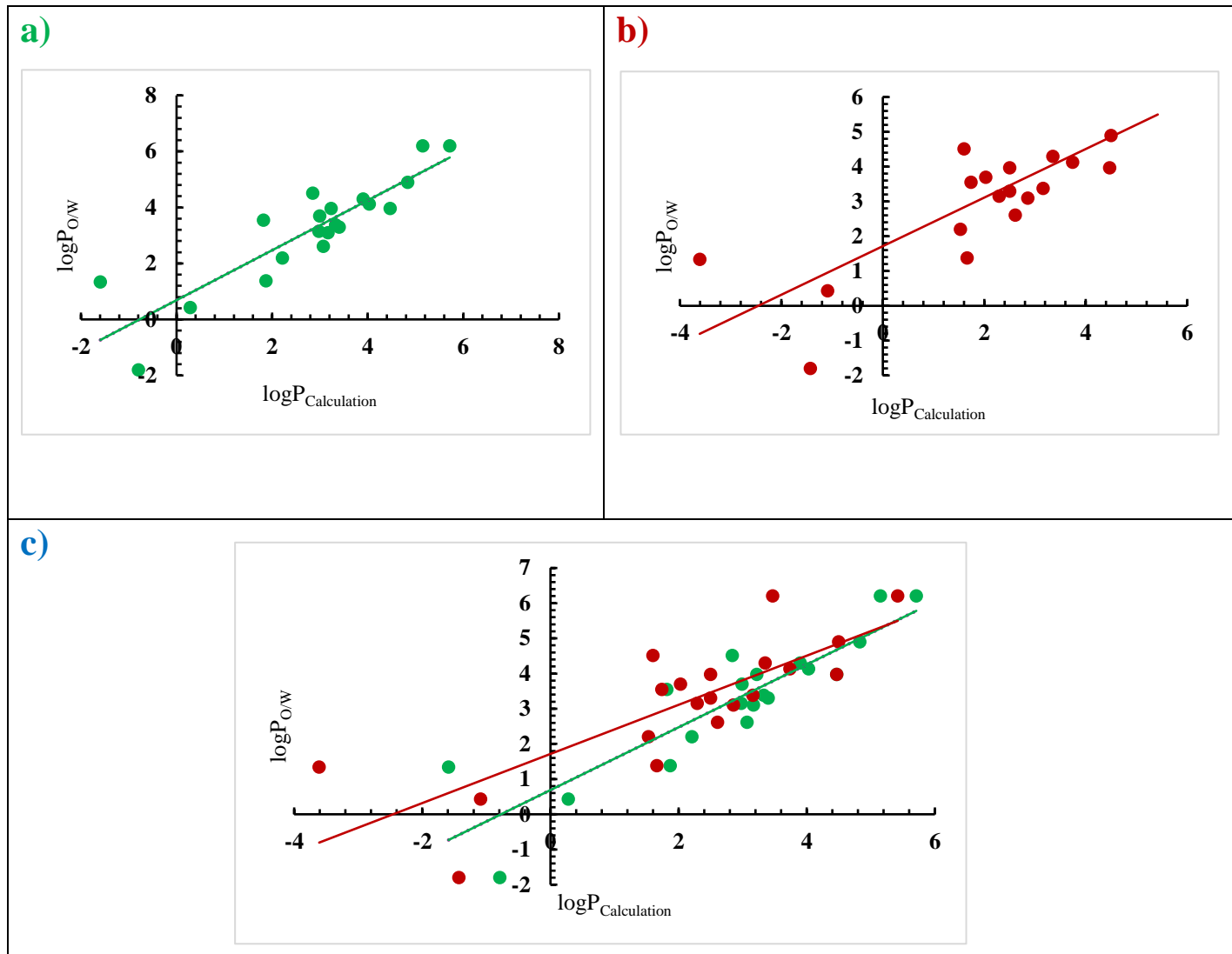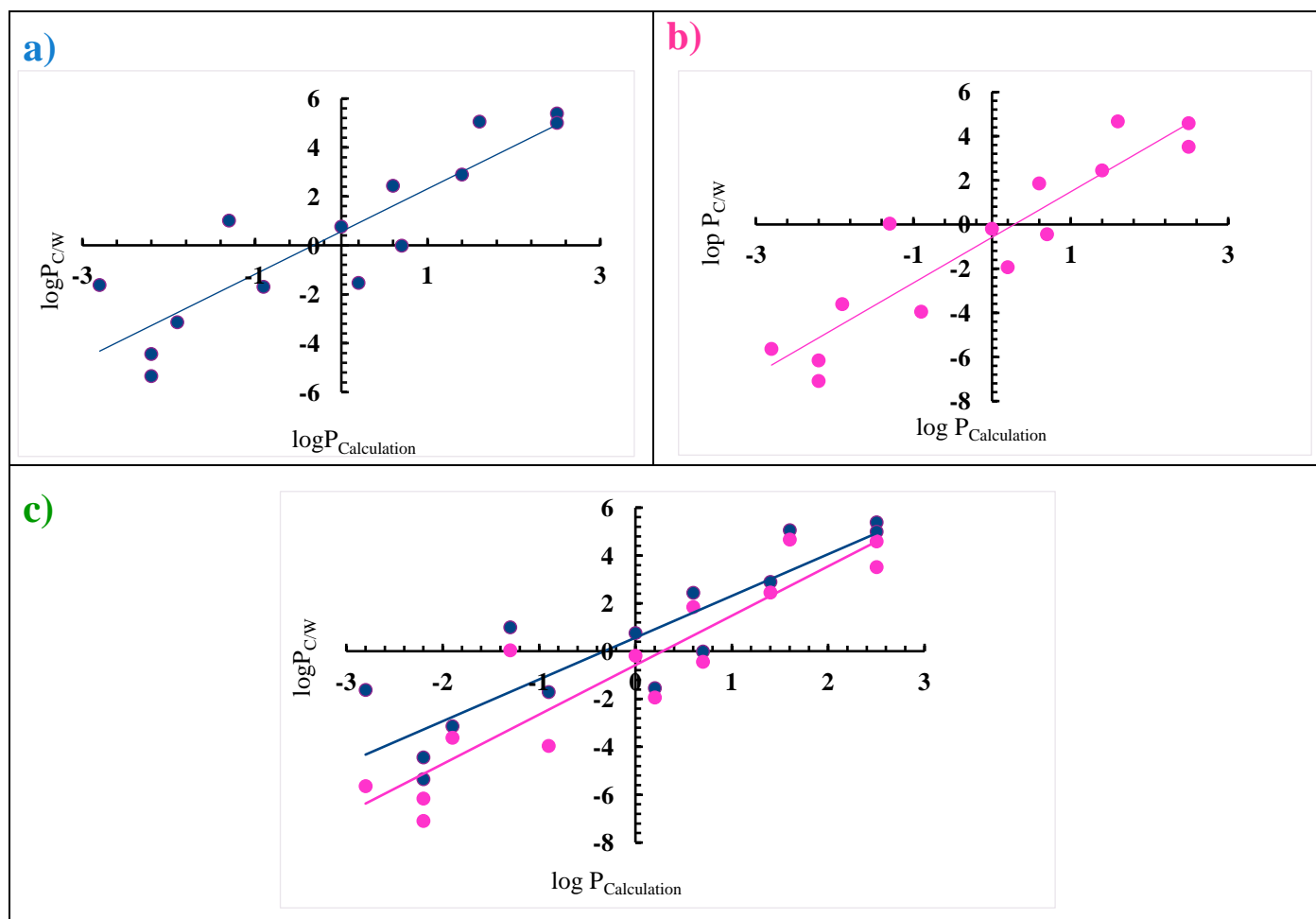
Table 5. Linear Regression Parameters Obtained for a) average log $P_{O/W}$ (octanol/water) and

b) minimum log $P_{O/W}$ with respect to the experimental partition coefficients in octanol/water.

	Solvent	R^2	Equation	RMSE	ASE	MAE
a	Octanol	0.78	y= 0.89x + 0.68	0.96	0.93	0.64
b		0.63	y=0.69x+1.71	1.66	2.75	1.17

5.2.4 Comparison with experimental Cyclohexane/Water partition coefficient with calculation partition coefficients in a different conformation

Figure 2 shows the relationship between the experimental log P values and the average log P and minimum log P values of 16 compounds dissolved in cyclohexane solvent (as described in Table 3).

The R^2 for Figure 2(a), which presents the correlation between the calculated average partition coefficient in a cyclohexane solvent and the experimental data, was determined to be 0.78. However, it is worth noting that (the slope of the Figure), 1.74, does not align with the expected value (1).

The R-squared (R^2) is 0.85 for the minimum calculated partition coefficient, displaying a stronger correlation compared to the average partition coefficient. However, it's important to note that the slope obtained from the equation is far from the expected value, and not good values of MAE for both methodologies (average and minimum) are obtained.

Also, it has to be noted that the number of studied molecules in the cyclohexane solvent is less than in the octanol solvent.

The minimum partition coefficient can provide a better correlation with experimental data compared to the average partition coefficient due to two main reasons:

Firstly, the minimum partition coefficient often corresponds to a specific conformation of the molecule that interacts strongly with the partitioning medium. This dominant conformation

contributes significantly to the overall partitioning behaviour, resulting in a better correlation with experimental data.

Secondly, certain solvents may have a preference for stabilizing or interacting favorably with specific conformations of the molecule. The minimum partition coefficient captures these conformations that have a strong affinity for the solvent, leading to improved correlation with the experimental data compared to the average partition coefficient.

Figure 2. Comparison of the Experimental partition coefficient in Cyclohexane solvent with **a)** the average log P calculated for B3LYP and **b)** the minimum log P calculated for B3LYP and **c)** compassion between average and minimum partition coefficient calculated.

Table 6. Linear Regression Parameters Obtained for a) average Log $P_{C/W}$ (Cyclohexane/water) and b) minimum log $P_{C/W}$ with Respect to the Experimental partition coefficients in Cyclohexane /water.

	Solvent	**R^2**	**Solvent**	**RMSE**	**ASE**	**MAE**
a	Cyclohexane	0.78	y=1.74x+0.56	2.12	4.49	1.89
b		0.85	y=2.06x-0.58	2.41	5.84	2.09

As a result, according to the obtained data, log $P_{Cyclohexane}$ has a greater error in their estimation than log $P_{octanol}$ and the methodology of calculation of partition coefficients calculated in cyclohexane should be improved.

The obtained results indicate the utility to investigate the partition coefficient of various conformations of flexible compounds in different solvents and then compare the results with experimental data. More analysis should be performed to identify the appropriate methodology for each solvent.

5.3 References

[1] Kertesz, M. (2019). Pancake Bonding: An Unusual Pi-Stacking Interaction. Chem. - A Eur. J., 25(2), 400–416. doi:10.1002/chem.201802385.

[2] Adhikary, C., Sana, S., & Chattopadhyay, K. N. (2020). Teaching Sigma (σ) And Pi (π) Bonds: An Orbital. Ijsart, 3(10), 429–434. ISSN: 2454-2415.

[3] Pauling, L. (1931). The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. J. Am. Chem. Soc., 53(4), 1367–1400. doi:10.1021/ja01355a027.

[4] Aparicio, F., Mayoral, M. J., Montoro-García, C., & González-Rodríguez, D. (2019). Guidelines for the assembly of hydrogen-bonded macrocycles. Chem. Commun., 55(51), 7277–7299. doi:10.1039/c9cc03166a.

[5] Lee, J. H., Jeoung, S., Chung, Y. G., & Moon, H. R. (2019). Elucidation of flexible metal-organic frameworks: Research progresses and recent developments. Coord. Chem. Rev., 389, 161–188. doi:10.1016/j.ccr.2019.03.008.

[6] Toyota, S. (2010). Rotational isomerism involving acetylene carbon. Chem. Rev., 110(9), 5398–5424. doi:10.1021/cr1000628.

[7] Appavoo, S. D., Huh, S., Diaz, D. B., & Yudin, A. K. (2019). Conformational Control of Macrocycles by Remote Structural Modification. Chem. Rev., 119(17), 9724–9752. doi: 10.1021/acs.chemrev.8b00742.

[8] Gunasekaran, K., Ma, B., & Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? Proteins Struct. Funct. Genet., 57(3), 433–443. doi:10.1002/prot.20232.

[9] Arslan, E., Findik, B. K., & Aviyente, V. (2020). A blind SAMPL6 challenge: insight into the octanol-water partition coefficients of drug-like molecules via a DFT approach. J. Comput. Aided. Mol. Des., 34(4), 463–470. doi:10.1007/s10822-020-00284-3.

[10] NISHI, H. (2021). Development of Fast and Selective Analytical Methods of Pharmaceuticals and Herbal Medicines by High-Performance Liquid Chromatography and Capillary Electrophoresis. CHROMATOGRAPHY, 42(1), 1–16. doi:10.15583/jpchrom.2020.026.

[11] Ganesan, A., Coote, M. L., & Barakat, K. (2017). Molecular dynamics-driven drug discovery: leaping forward with confidence. Drug Discov. Today, 22(2), 249–269. doi: 10.1016/j.drudis.2016.11.001.

[12] Hird, M. (2007). Fluorinated liquid crystals – properties and applications. Chem. Soc. Rev., 36(12), 2070–2095. doi:10.1039/b610738a.

[13] Yalkowsky, S. H., & Alantary, D. (2018). Estimation of Melting Points of Organics. J. Pharm. Sci., 107(5), 1211–1227. doi:10.1016/j.xphs.2017.12.013.

[14] Lehn, J.-M. (1988). Supramolecular Chemistry-Scope and Perspectives Molecules, Supermolecules, and Molecular Devices (Nobel Lecture). Angew. Chemie Int. Ed. English, 27(1), 89–112. doi:10.1002/anie.198800891.

[15] Etter, M. C. (1991). Hydrogen bonds as design elements in organic chemistry. J. Phys. Chem., 95(12), 4601–4610. doi:10.1021/j100165a007.

[16] Lorber, D. M., & Shoichet, B. K. (1998). Flexible ligand docking using conformational ensembles Despite important successes. Protein Sci., 7, 938–950. doi:10.1002/pro.5560070411.

[17] Pike, A. C. W., Brew, K., & Acharya, K. R. (1996). Crystal structures of guinea-pig, goat and bovine α-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. Structure, 4(6), 691–703. doi:10.1016/S0969-2126(96)00075-5.

[18] Sharif, K., et al. (2021). The Putative Adverse Effects of Bisphenol A on Autoimmune Diseases. Endocrine, Metab. Immune Disord. - Drug Targets, 22(7), 665–676. doi: 10.2174/1871530321666210210154309.

[19] Bryngelson, J. D., Onuchic, J. N., Socci, N. D., & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: A synthesis. Proteins Struct. Funct. Bioinforma., 21(3), 167–195. doi:10.1002/prot.340210302.

[20] Dai, J., & Zhang, H. (2021). Recent Advances in Catalytic Confinement Effect within Micro/Meso-Porous Crystalline Materials. Small, 17(22), 1–25. doi:10.1002/smll.202005334.

[21] Guo, Z., Lue, B. M., Thomasen, K., Meyer, A. S., & Xu, X. (2007). Predictions of flavonoid solubility in ionic liquids by COSMO-RS: Experimental verification, structural elucidation, and solvation characterization. Green Chem., 9(12), 1362–1373. doi:10.1039/b709786g.

[22] Lan, Y., et al. (2014). Comparing and correlating solubility parameters governing the self-assembly of molecular gels using 1,3:2,4-dibenzylidene sorbitol as the gelator. Langmuir, 30(47), 14128–14142. doi:10.1021/la5008389.

[23] Dong, X., Zhang, Z., & Tian, Z. (2022). Precambrian metamorphic basement of the southern Lhasa terrane, Tibet. Precambrian Research, 368, 106478. doi:10.1016/j.precamres.2021.106478.

[24] Vermaas, J. V., Crowley, M. F., & Beckham, G. T. (2020). Molecular Lignin Solubility and Structure in Organic Solvents. ACS Sustainable Chemistry & Engineering, 8(48), 17839–17850. doi:10.1021/acssuschemeng.0c07156.

[25] Wang, C., Rosbottom, I., Turner, T. D., Laing, S., Maloney, A. G., Sheikh, A. Y., ... & Roberts, K. J. (2021). Molecular, solid-state and surface structures of the conformational polymorphic forms of ritonavir in relation to their physicochemical properties. Pharmaceutical Research, 38(6), 971-990. doi:10.1007/s11095-021-03048-2.

[26] Rossi Sebastiano, M., et al. (2018). Impact of Dynamically Exposed Polarity on Permeability and Solubility of Chameleonic Drugs beyond the Rule of 5. Journal of Medicinal Chemistry, 61(9), 4189–4202. doi:10.1021/acs.jmedchem.8b00347.

[27] Kato, T., Yoshio, M., Ichikawa, T., Soberats, B., Ohno, H., & Funahashi, M. (2017). Transport of ions and electrons in nanostructured liquid crystals. Nature Reviews Materials, 2(4). doi: 10.1038/natrevmats.2017.1.

[28] Mateo-Alonso, A. (2023). π-Conjugated Materials: Here, there, and everywhere. Chemistry of Materials, 35(4), 1467–1469. doi:10.1021/acs.chemmater.2c03567.

[29] Bässler, H., & Köhler, A. (2012). Charge transport in organic semiconductors. Topics in Current Chemistry, 312, 1–65. doi:10.1007/128_2011_218.

[30] Hunter, C. A., & Sanders, J. K. M. (1990). The Nature of Interactions. Journal of the American Chemical Society, 112(9), 5525–5534. doi:10.1021/ja00170a016.

[31] Bao, G., & Suresh, S. (2003). Cell and molecular mechanics of biological materials. Nature materials, 2(11), 715-725. doi:10.1038/nmat1001.

[32] Sheiko, S. S., Sumerlin, B. S., & Matyjaszewski, K. (2008). Progress in Polymer Science: Cylindrical molecular brushes: Synthesis, characterization, and properties, 33, 759–785. doi: 10.1016/j.progpolymsci.2008.05.001.

[33] Huang, W., et al. (2019). Multiscale Toughening Mechanisms in Biological Materials and Bioinspired Designs, 1901561, 1–37. doi:10.1002/adma.201901561.

[34] Meredith, H. J., & Wilker, J. J. (2015). The Interplay of Modulus, Strength, and Ductility in Adhesive Design Using Biomimetic Polymer Chemistry, 5057–5065. doi: 10.1002/adfm.201501880.

[35] Miller, J. S., & Epstein, A. J. (1994). Organic and organometallic molecular magnetic materials—designer magnets. Angewandte Chemie International Edition in English, 33(4), 385-415. doi:10.1002/anie.199403851.

[36] Issa, B., Obaidat, I. M., Albiss, B. A., & Haik, Y. (2013). Magnetic nanoparticles: surface effects and properties related to biomedicine applications. International journal of molecular sciences, 14(11), 21266-21305. doi:10.3390/ijms141121266.

[37] Markad, G. B., Padma, N., Chadha, R., Gupta, K. C., Rajarajan, A. K., Deb, P., & Kapoor, S. (2020). Mutual influence on aggregation and magnetic properties of graphene oxide and copper phthalocyanine through non-covalent, charge transfer interaction. Applied Surface Science, 505, 144624. doi:10.1016/j.apsusc.2019.144624.

[38] Herbert, K. M., Fowler, H. E., McCracken, J. M., Schlafmann, K. R., Koch, J. A., & White, T. J. (2022). Synthesis and alignment of liquid crystalline elastomers. Nature Reviews Materials, 7(1), 23-38. doi:10.1038/s41578-021-00359-z.

[39] Descalzo, A. B., Martínez-Máñez, R., Sancenón, F., Hoffmann, K., & Rurack, K. (2006). The supramolecular chemistry of organic–inorganic hybrid materials. Angewandte Chemie International Edition, 45(36), 5924-5948. doi:10.1002/anie.200600734.

[40] Fu, S., Cai, Z., & Ai, H. (2021). Stimulus-responsive nanoparticle magnetic resonance imaging contrast agents: design considerations and applications. Advanced Healthcare Materials, 10(5), 2001091. doi:10.1002/adhm.202001091.

[41] Andricopulo, A. D., Guido, R. V., & Oliva, G. (2008). Virtual screening and its integration with modern drug design technologies. Current medicinal chemistry, 15(1), 37-46. doi:10.2174/092986708783330683.

[42] Gohlke, H., & Klebe, G. (2002). Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Angewandte Chemie International Edition, 41(15), 2644-2676. doi:10.1002/1521-3773(20020802)41:15<2644: AID-ANIE2644>3.0.CO;2-OCitations: 620.

[43] Craik, D. J., Fairlie, D. P., Liras, S., & Price, D. (2013). The future of peptide-based drugs. Chemical biology & drug design, 81(1), 136-147. doi:10.1111/cbdd.12055.

[44] Yuriev, E., Holien, J., & Ramsland, P. A. (2015). Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. Journal of Molecular Recognition, 28(10), 581-604. doi:10.1002/jmr.2471.

[45] Del Hoyo, C. (2007). Layered double hydroxides and human health: an overview. Applied Clay Science, 36(1-3), 103-121. doi:10.1016/j.clay.2006.06.010.

[46] Mascarenhas-Melo, F., Carvalho, A., Gonçalves, M. B. S., Paiva-Santos, A. C., & Veiga, F. (2022). Nanocarriers for the topical treatment of psoriasis-pathophysiology, conventional treatments, nanotechnology, regulatory and toxicology. European Journal of Pharmaceutics and Biopharmaceutics, 176, 95-107. doi:10.1016/j.ejpb.2022.05.012.

[47] Workman, P., & Collins, I. (2010). Probing the probes: fitness factors for small molecule tools. Chemistry & biology, 17(6), 561-577.doi:10.1016/j.chembiol.2010.05.013.

[48] Buggert, M., Cadena, C., Mokrushina, L., Smirnova, I., Maginn, E. J., & Arlt, W. (2009). COSMO-RS calculations of partition coefficients: different tools for conformation search. Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology, 32(6), 977-986. doi:10.1002/ceat.200800654.

[49] Shaytan, A. K., Shaitan, K. V., & Khokhlov, A. R. (2009). Solvent accessible surface area of amino acid residues in globular proteins: correlation of apparent transfer free energies with experimental hydrophobicity scales. Biomacromolecules, 10(5), 1224-1237. doi:10.1021/bm8015169.

[50] Leodidis, E. B., & Hatton, T. A. (1990). Amino acids in AOT reversed micelles. 2. The hydrophobic effect and hydrogen bonding as driving forces for interfacial solubilization. Journal of Physical Chemistry, 94(16), 6411-6420. doi:10.1021/j100379a047.

[51] Li, Y., & Yang, L. (2015). Driving forces for drug loading in drug carriers. Journal of microencapsulation, 32(3), 255-272. doi:10.3109/02652048.2015.1010459.

[52] Nita, S., Cann, N. M., & Horton, J. H. (2004). Structure, selectivity, and solvation of a model chiral stationary phase. The Journal of Physical Chemistry B, 108(11), 3512-3522. doi:10.1021/jp0365615.

[53] Vistoli, G., Pedretti, A., & Testa, B. (2009). Partition coefficient and molecular flexibility: the concept of lipophilicity space. Chemistry & Biodiversity, 6(8), 1152-1169. doi:10.1002/cbdv.200900072.

[54] Pagliara, A., Carrupt, P. A., Caron, G., Gaillard, P., & Testa, B. (1997). Lipophilicity profiles of ampholytes. Chemical reviews, 97(8), 3385-3400. doi:10.1021/cr9601019.

[55] Caron, G., Vallaro, M., & Ermondi, G. (2013). The Block Relevance (BR) analysis to aid medicinal chemists to determine and interpret lipophilicity. MedChemComm, 4(10), 1376-1381. doi:10.1039/C3MD00140G.

[56] Vistoli, G., Pedretti, A., & Testa, B. (2008). Assessing drug-likeness–what are we missing. Drug discovery today, 13(7-8), 285-294. doi:10.1016/j.drudis.2007.11.007.

[57] Mohsin, K., Long, M. A., & Pouton, C. W. (2009). Design of lipid-based formulations for oral administration of poorly water-soluble drugs: precipitation of drug after dispersion of formulations in aqueous solution. Journal of pharmaceutical sciences, 98(10), 3582-3595. doi:10.1002/jps.21659.

[58] Qian, H. L., Yang, C., & Yan, X. P. (2018). Layer-by-layer preparation of 3D covalent organic framework/silica composites for chromatographic separation of position isomers. Chemical communications, 54(83), 11765-11768. doi:10.1039/C8CC06621C.

[59] van Muijlwijk-Koezen, J. E., Timmerman, H., van der Goot, H., Menge, W. M., Frijtag von Drabbe Künzel, J., de Groote, M., & IJzerman, A. P. (2000). Isoquinoline and quinazoline urea analogues as antagonists for the human adenosine A3 receptor. Journal of medicinal chemistry, 43(11), 2227-2238. doi:10.1021/jm000002u.

[60] Kopsky, D. J., & Hesselink, J. M. K. (2021). U.S. Patent No. 11,147,799. Washington, DC: U.S. Patent and Trademark Office.

[61] Saettone, M. F., Giannaccini, B., Delmonte, G., Campigli, V., Tota, G., & La Marca, F. (1988). Solubilization of tropicamide by poloxamers: physicochemical data and activity data in rabbits and humans. International journal of pharmaceutics, 43(1-2), 67-76. doi:10.1016/0378-5173(88)90060-9.

[62] Sahoo, D. K., Martinez, M. N., Dao, K., Gabriel, V., Zdyrski, C., Jergens, A. E., ... & Mochel, J. P. (2023). Canine Intestinal Organoids as a Novel In Vitro Model of Intestinal Drug Permeability: A Proof-of-Concept Study. Cells, 12(9), 1269. doi:10.3390/cells12091269.

[63] Markovic, M., Zur, M., Ragatsky, I., Cvijić, S., & Dahan, A. (2020). BCS Class IV oral drugs and absorption windows: Regional-dependent intestinal permeability of furosemide. Pharmaceutics, 12(12), 1175. doi:10.3390/pharmaceutics12121175.

[64] Rankovic, Z. (2017). CNS physicochemical property space shaped by a diverse set of molecules with experimentally determined exposure in the mouse brain: miniperspective. Journal of Medicinal Chemistry, 60(14), 5943-5954. doi:10.1021/acs.jmedchem.6b01469.

[65] Punetha, A., Green, K. D., Garzan, A., Chandrika, N. T., Willby, M. J., Pang, A. H., ... & Garneau-Tsodikova, S. (2021). Structure-based design of haloperidol analogues as inhibitors of acetyltransferase Eis from Mycobacterium tuberculosis to overcome kanamycin resistance. RSC Medicinal Chemistry, 12(11), 1894-1909. doi:10.1039/D1MD00239B.

[66] Schump, M. D., Fox, D. M., Bertozzi, C. R., & Riley, L. W. (2017). Subcellular partitioning and intramacrophage selectivity of antimicrobial compounds against Mycobacterium tuberculosis. Antimicrobial Agents and Chemotherapy, 61(3), 10-1128. doi:10.1128/aac.01639-16.

[67] Fuguet, E., Ràfols, C., Bosch, E., & Rosés, M. (2003). Characterization of the solvation properties of sodium n-dodecyl sulfate micelles in buffered and unbuffered aqueous phases by solvatochromic indicators. Langmuir, 19(1), 55-62. doi:10.1021/la026307o.

[68] Vieira, Y., Silveira, J. P., Dotto, G. L., Knani, S., Vieillard, J., Georgin, J., ... & Lima, E. C. (2022). Mechanistic insights and steric interpretations through statistical physics modelling and density functional theory calculations for the adsorption of the pesticides atrazine and diuron by Hovenia dulcis biochar. Journal of Molecular Liquids, 367, 120418. doi:10.1016/j.molliq.2022.120418.

[69] Wiegel, S., Aulinger, A., Brockmeyer, R., Harms, H., Löffler, J., Reincke, H., ... & Wanke, A. (2004). Pharmaceuticals in the river Elbe and its tributaries. Chemosphere, 57(2), 107-126. doi:10.1016/j.chemosphere.2004.05.017.

[70] Ye, M., Nagar, S., & Korzekwa, K. (2016). A physiologically based pharmacokinetic model to predict the pharmacokinetics of highly protein-bound drugs and the impact of errors in plasma protein binding. Biopharmaceutics & drug disposition, 37(3), 123-141. doi:10.1002/bdd.1996.

[71] Ye, M., Nagar, S., & Korzekwa, K. (2016). A physiologically based pharmacokinetic model to predict the pharmacokinetics of highly protein-bound drugs and the impact of errors in plasma protein binding. Biopharmaceutics & drug disposition, 37(3), 123-141. doi:10.1002/bdd.1996.

[72] Jones, M. R., Brooks, B. R., & Wilson, A. K. (2016). Partition coefficients for the SAMPL5 challenge using transfer free energies. Journal of computer-aided molecular design, 30, 1129-1138. doi:10.1007/s10822-016-9964-6.

[73] Rustenburg, A. S., Dancer, J., Lin, B., Feng, J. A., Ortwine, D. F., Mobley, D. L., & Chodera, J. D. (2016). Measuring experimental cyclohexane-water distribution coefficients for the SAMPL5 challenge. Journal of computer-aided molecular design, 30, 945-958. doi:10.1007/s10822-016-9971-7.

[74] Bannan, C. C., Calabró, G., Kyu, D. Y., & Mobley, D. L. (2016). Calculating partition coefficients of small molecules in octanol/water and cyclohexane/water. Journal of chemical theory and computation, 12(8), 4015-4024. doi:10.1021/acs.jctc.6b00449.

[75] Saranjam, L., Fuguet, E., Nedyalkova, M., Simeonov, V., Mas, F., & Madurga, S. (2021). Prediction of Partition Coefficients in SDS Micelles by DFT Calculations. Symmetry, 13(9), 1750. doi:10.3390/sym13091750.

CHAPTER
VI

CONCLUSION

Chapter VI
Conclusions

- The DFT method was employed to study the partition coefficient of alcoholic compounds (ethanol, n-propanol, i-propanol, and n-butanol) in various organic solvents. Three different basis sets (6-31G(d), 6-311+G*, and 6-311++G**) were used, and the results were compared with experimental data. Interestingly, it was found that alcohol compounds using the 6-31G(d) set exhibited the strongest correlation with the experimental data.

- The partition coefficient of the THF molecule in carbon tetrachloride, heptane, hexane, and trichloroethane solvents was studied using 6.31G(d), 6-311+G*, and 6-311++G**. The results were then compared to the experimental value, showing that the partition coefficient calculated with the 6-311++G** basis set displayed a stronger correlation with the experimental partition coefficient.

- The partition coefficient of the Xylene, Benzene, and Ethylbenzene compounds in octanol and heptane solvents were studied using 6-31G(d), 6-311+G*, and 6-311++G** basis sets. The study showed that the partition coefficient of Xylene in octanol solvent computed with a 6-31G(d) basis set shows a high correlation with the experimental partition coefficient. Also, the partition coefficient of Benzene and Ethyl Benzene using 6-31G(d) and 6-311+G* and 6-311++G** in octanol and heptane displayed a strong correlation with the experimental partition coefficient.

- The partition coefficient of 63 molecules in 15 different solvents was determined using the DFT method, employing B3LYP and M06-2X functionals with a 6-311++G** basis set. These calculated values were then compared to the experimental partition coefficient for the SDS micelle. The log $P_{Calculation}$ in 2-propanol, 1-propanol, and methanol solvents exhibits the highest correlation with the log P_{SDS}.

- The log P Calculated for 63 molecules in alcoholic solvents (1-propanol, 2-propanol, and methanol) calculated with B3LYP and M06-2X and a high basis set shows the highest correlation with the experimental partition coefficient of SC micelles.

- The B3LYP and M06-2X methods made similar slopes for the FPLOS micelle, presenting a comparable prediction of the partition coefficient. A better correlation was found between the experimental log P_{LPFOS} and log $P_{1\text{-propanol/water}}$.

- We calculated the partition coefficients of 63 molecules using two methods, B3LYP and M06-2X, with a 6.31++G** basis set. These calculations were conducted across 15 different solvents. When we compared these computed values to the experimental partition coefficients of HTAB micelles, we observed that there wasn't a strong correlation with the experimental data. As a result, our attention shifted towards determining the partition coefficients of molecules that lacked aromatic nitrogen atoms or amide groups. When we compared these newly computed values to the experimental data, they exhibited a significant correlation with the experimental values.

- A strong linear correlation was shown while comparing the $\log P_{o/w}$ (partition coefficient between octanol and water) for 63 molecules with experimental values using both the B3LYP and M06-2X functionals.

- The study involved calculating the minimum and average partition coefficients of 17 flexible compounds in octanol solvent and 16 flexible molecules in cyclohexane solvent in various conformations. These computed values were then compared to experimental data. Notably, the average partition coefficient calculated in octanol solvent showed a high correlation with the experimental partition coefficients. Additionally, the partition coefficients calculated in cyclohexane did not display specific results.
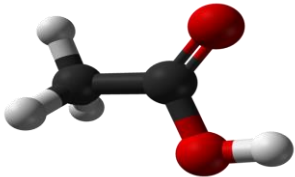
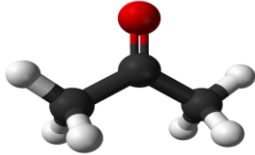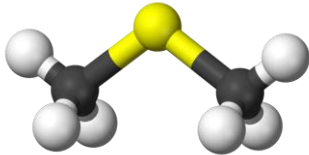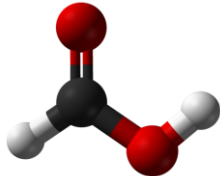# CHAPTER
VII
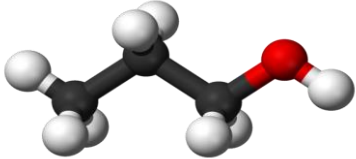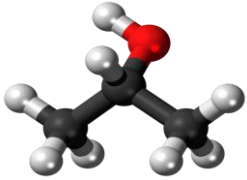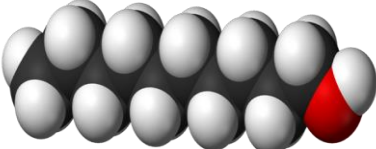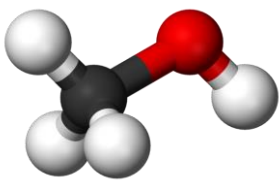
PUBLICATIONS
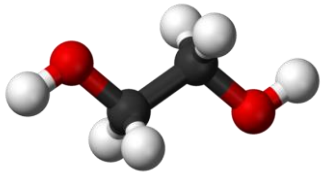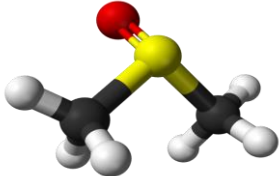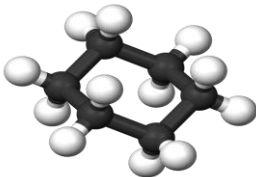
Publication

[1] Saranjam, L.; Fuguet, E.; Nedyalkova, M.; Simeonov, V.; Mas, F.; Madurga, S. Prediction of Partition Coefficients in SDS Micelles by DFT Calculations, Symmetry 2021, 13, 1750. https://doi.org/10.3390/sym13091750.

[2] Saranjam, L.; Nedyalkova, M.; Fuguet, E.; Simeonov, V.; Mas, F.; Madurga, S. Collection of Partition Coefficients in Hexadecyltrimethylammonium Bromide, Sodium Cholate, and Lithium Perfluorooctanesulfonate Micellar Solutions: Experimental Determination and Computational Predictions, Molecules 2023, 28, 5729, https://doi.org/10.3390/molecules28155729.

Iupac name	Cas number	Formula	Structure
Propan-1-ol	71-23-8	C_3H_8O	
Propan-2-ol	67-63-0	C_3H_8O	
Decan-1-ol	112-30-1	$C_{10}H_{21}OH$	
Methanol	67-56-1	CH_4O	
Octan-1-ol	111-87-5	$C_8H_{18}O$	

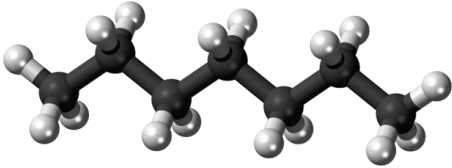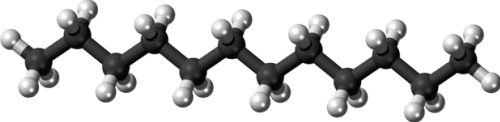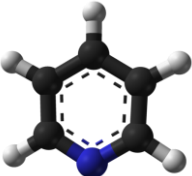|			
Ethane-1,2-diol	107-21-1	$C_2H_6O_2$	
(Methanesulfinyl)met hane	67-68-5	C_2H_6OS	
Cyclohexane	110-82-7	C_6H_{12}	
Acetic acid	64-19-7	$C_2H_4O_2$	
Acetone	67-64-1	C_3H_6O	
(Ethylsulfanyl)ethane	352-93-2	$C_4H_{10}S$	
Formic acid	64-18-6	CH_2O_2	

Heptane	142-82-5	C_7H_{16}	
Dodecane	112-40-3	$C_{12}H_{26}$	
Pyridine	110-86-1	C_5H_5N	

Figure 1. Chemical structures of the solvents used in DFT calculations.

	Coefficient	Standard Error	P-value	Lower 95%	Upper 95 %	F	Signific ance F	R²
Calculated logP propan-1-ol/water vs Exp. logP of SDS micelles with B3LYP								
Intercept	-0.21	0.24	0.38	-0.68	0.27			
x	1.19	0.09	1.2E-18	1.01	1.38	163.6	1.2E-18	0.73
Calculated logP propan-1-ol/water vs Exp. logP of SDS micelles with M06-2X								
Intercept	-0.19	0.28	0.50	-0.74	0.37			
x	1.20	0.11	3.0E-15	0.98	1.42	120.3	3.0E-15	0.69
Calculated logP propan-2-ol/water vs Exp. logP of SDS micelles with B3LYP								
Intercept	0.04	0.27	0.87	-0.49	0.58			
x	1.14	0.10	2.1E-15	0.93	1.35	119.4	2.1E-15	0.68
Calculated logP propan-2-ol/water vs Exp. logP of SDS micelles with M06-2X								
Intercept	-0.28	0.30	0.35	-0.89	0.32			
x	1.27	0.12	2.9E-14	1.03	1.52	109.0	2.9E-14	0.68
Calculated logP methanol/water vs Exp. logP of SDS micelles with B3LYP								
Intercept	-0.21	0.25	0.41	-0.72	0.30			
x	1.20	0.10	4.1E-17	0.99	1.40	142.1	4.1E-17	0.71
Calculated logP methanol/water vs Exp. logP of SDS micelles with M06-2X								
Intercept	-0.39	0.35	0.27	-1.10	0.31			
x	1.31	0.14	2.6E-12	1.02	1.60	83.5	2.6E-12	0.62
Calculated logP heptane/water vs Exp. logP of SDS micelles with B3LYP								
Intercept	-1.32	1.15	0.26	-3.62	0.98			
x	0.67	0.45	0.15	-0.24	1.57	2.18	1.5E-01	0.04
Calculated logP heptane/water vs Exp. logP of SDS micelles with M06-2X								
Intercept	-1.99	1.21	0.11	-4.41	0.44			
x	0.92	0.47	0.06	-0.02	1.87	3.83	5.6E-02	0.07
Calculated logP cyclohexane/water vs Exp. logP of SDS micelles with B3LYP								
Intercept	-1.33	1.16	0.26	-3.66	1.00			
x	0.64	0.46	0.16	-0.27	1.55	1.99	1.6E-01	0.03
Calculated logP cyclohexane/water vs Exp. logP of SDS micelles with M06-2X								
Intercept	-1.78	1.19	0.14	-4.18	0.61			
x	0.85	0.46	0.07	-0.08	1.77	3.40	7.1E-02	0.06
Calculated logP N-dodecane/water vs Exp. logP of SDS micelles with B3LYP								
Intercept	-1.86	1.01	0.07	-3.87	0.15			
x	0.90	0.40	0.03	0.10	1.70	5.12	2.7E-02	0.08
Calculated logP N-dodecane/water vs Exp. logP of SDS micelles with M06-2X								

Intercept	-1.65	1.22	0.18	-4.11	0.81			
x	0.73	0.48	0.14	-0.24	1.70	2.28	1.4E-01	0.04

Calculated logP pyridine/water vs Exp. logP of SDS micelles with B3LYP

Intercept	0.45	0.65	0.50	-0.85	1.75			
x	0.97	0.25	3.2E-04	0.46	1.48	14.57	3.2E-04	0.20

Calculated logP pyridine/water vs Exp. logP of SDS micelles with M06-2X

Intercept	0.16	0.63	0.80	-1.10	1.43			
x	1.13	0.24	2.4E-05	0.64	1.62	21.65	2.4E-05	0.30

Calculated logP diethyl sulfide/water vs Exp. logP of SDS micelles with B3LYP

Intercept	0.14	0.78	0.86	-1.41	1.70			
x	0.88	0.30	0.01	0.28	1.49	8.45	5.1E-03	0.12

Calculated logP diethyl sulfide/water vs Exp. logP of SDS micelles with M06-2SX

Intercept	-0.48	0.81	0.56	-2.09	1.14			
x	1.18	0.31	3.7E-04	0.56	1.80	14.54	3.7E-04	0.22

Calculated logP acetic acid/water vs Exp. logP of SDS micelles with B3LYP

Intercept	-1.80	0.86	0.04	-3.52	-0.08			
x	0.45	0.34	0.18	-0.22	1.13	1.84	1.8E-01	0.03

Calculated logP acetic acid/water vs Exp. logP of SDS micelles with M06-2X

Intercept	-2.46	0.84	4.9E-03	-4.14	-0.78			
x	0.74	0.32	0.03	0.09	1.38	5.18	2.7E-02	0.09

Calculated logP decan-1-ol/water vs Exp. logP of SDS micelles with B3LYP

Intercept	-0.87	0.47	0.07	-1.81	0.08			
x	0.83	0.18	3.1E-05	0.46	1.20	20.39	3.1E-05	0.26

Calculated logP decan-1-ol/water vs Exp. logP of SDS micelles with M06-2X

Intercept	-1.19	0.48	0.02	-2.16	-0.23			
x	0.99	0.19	5.3E-06	0.60	1.38	25.85	5.3E-06	0.34

Calculated logP octanol/water vs Exp. logP of SDS micelles with B3LYP

Intercept	-0.53	0.37	0.17	-1.28	0.22			
x	0.91	0.15	8.4E-08	0.61	1.20	37.72	8.4E-08	0.40

Calculated logP octanol/water vs Exp. logP of SDS micelles with M06-2X

Intercept	-1.00	0.33	4.1E-03	-1.67	-0.33			
x	1.16	0.13	5.0E-12	0.90	1.42	77.57	5.0E-12	0.48

Calculated logP acetone/water vs Exp. logP of SDS micelles with B3LYP

Intercept	0.94	0.60	0.13	-0.27	2.14			
x	0.88	0.24	4.1E-04	0.41	1.35	14.01	4.1E-04	0.19

Calculated logP acetone/water vs Exp. logP of SDS micelles with M06-2X

Intercept	0.38	0.57	0.51	-0.77	1.52			
x	1.15	0.22	3.7E-06	0.71	1.59	27.07	3.7E-06	0.35

Calculated logP 1,2-ethane diol/water vs Exp. logP of SDS micelles with B3LYP

Intercept	-1.33	0.38	9.4E-04	-2.10	-0.57			
x	0.41	0.15	0.01	0.12	0.71	7.71	7.4E-03	0.12

Calculated logP 1,2-ethane diol/water vs Exp. logP of SDS micelles with M06-2X

Intercept	-1.99	0.43	3.1E-05	-2.87	-1.12			
x	0.73	0.18	1.4E-04	0.37	1.09	17.02	1.4E-04	0.26

Calculated logP dimethyl sulfoxide/water vs Exp. logP of SDS micelles with B3LYP

Intercept	0.13	0.55	0.81	-0.97	1.24			
x	0.64	0.22	4.4E-03	0.21	1.07	8.79	4.4E-03	0.13

Calculated logP dimethyl sulfoxide/water vs Exp. logP of SDS micelles with M06-2X

Intercept	-0.57	0.62	0.37	-1.82	0.68			
x	1.00	0.26	2.8E-04	0.49	1.52	15.31	2.8E-04	0.23

Calculated logP formic acid/water vs Exp. logP of SDS micelles with B3LYP

Intercept	-1.29	0.62	0.04	-2.53	-0.05			
x	0.46	0.24	0.06	-0.03	0.94	3.57	6.4E-02	0.06

Calculated logP formic acid/water vs Exp. logP of SDS micelles with M06-2X

Intercept	-1.85	0.71	0.01	-3.27	-0.44			
x	0.64	0.27	0.02	0.09	1.19	5.54	2.2E-02	0.10

Experimental logP octanol/water vs calculated logP octanol/water with B3LYP

Intercept	1.00	0.15	1.4E-08	0.70	1.30			
x	0.67	0.07	1.3E-12	0.52	0.82	82.48	1.3E-12	0.60

Experimental logP octanol/water vs calculated logP octanol/water with M06-2X

Intercept	0.90	0.15	3.3E-07	0.59	1.22			
x	0.71	0.07	1.3E-13	0.57	0.86	97.54	1.3E-13	0,65

Table 2. Statistical parameters were obtained from the linear regressions for B3LYP and M06-2X calculations [1].
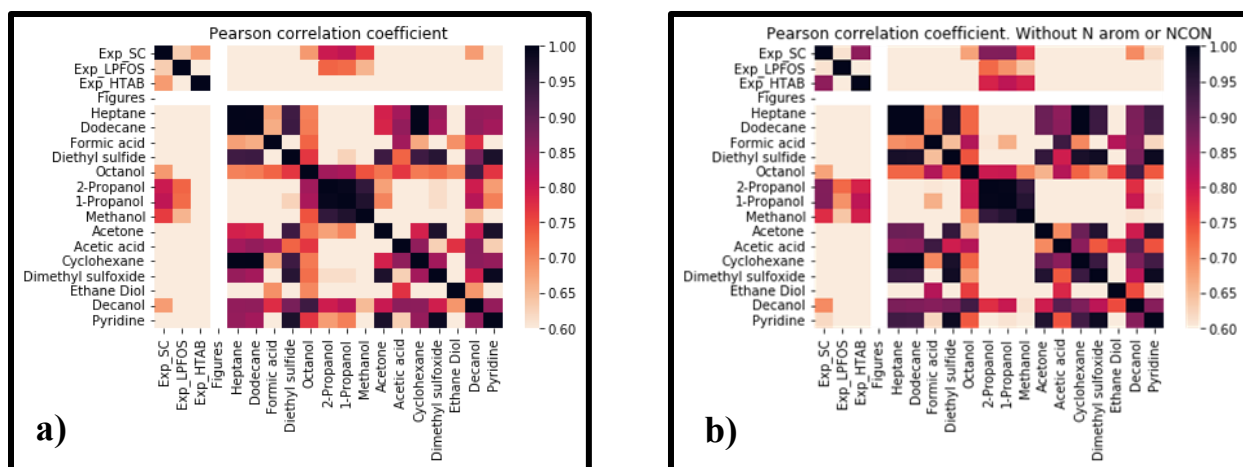
Figure 2. Heatmap of pairwise correlation of logP values for experimental and B3LYP calculated predictions. The first logP values are the experimental values in SC, LPFOS, and HTAB micelles respectively. The heat map is colored by the significance of the Pearson coefficient, where a darker red indicates a higher degree of correlation. In a) all compounds are used, in b) molecules having Nitrogen in an aromatic ring or having the urea (carbamide) group are excluded

References

[1] Saranjam, L.; Fuguet, E.; Nedyalkova, M.; Simeonov, V.; Mas, F.; Madurga, S. Prediction of Partition Coefficients in SDS Micelles by DFT Calculations, Symmetry, 2021, 13, 1750. https://doi.org/10.3390/sym13091750.