A COMPUTATIONAL APPROACH TO EVALUATE THE pK_a's OF QUINAZOLINE DERIVATIVES

by

Melisa Kıran

B.S., Chemistry, Boğaziçi University, 2020

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemistry
Boğaziçi University

2022

ACKNOWLEDGEMENTS

Foremost, I want to thank my thesis co-supervisor, Prof. Viktorya Aviyente, deserves special appreciation and respect for her valuable scientific assistance, support, and patience. I want to thank my adviser, Assist. Prof. Ahu Akın, for her unwavering support of my graduate studies and research, as well as her patience, encouragement, excitement, and vast expertise. Her advice was invaluable during the research and writing of this thesis.

I want to also thank my colleagues in the Department of Computational Chemistry at Bogazici University (CCBU) for their support throughout this period. Importantly, I am thankful to Evrim Arslan and Pınar Haşlak for their encouragement, support, and insightful comments.

I wish to extend my appreciation to my thesis committee members: Prof. İlknur Doğan and Prof. Aylin Sungur. I am thankful to them for their experience, guidance, and remarks during the study.

Computing resources used in this work were provided by the National Center for High Performance Computing of Turkey (UHeM) under grant number 5010142021.

Finally, I want to express my gratitude to my parents for their constant love, support, and understanding. My thesis is dedicated to my beloved family.

ABSTRACT

A COMPUTATIONAL APPROACH TO EVALUATE THE pK_a's OF QUINAZOLINE DERIVATIVES

One of the most important constants in chemistry is the ionization (dissociation) constant (pK_a). Estimating the pK_a value(s) for a potential drug is critical, especially since computations are considerably less expensive than obtaining pK_a values experimentally. Nearly 68 % of ionized medicines are said to have weak bases [1].

As a first step in developing an efficient estimation methodology for the pK_a of quinazoline and derivatives we studied three protocols. First one is based on the linear relationship between computed atomic charges of quinazoline and derivatives and the experimental pK_a. Based on our observations, the optimum method for reproducing observed pK_a's is to compute NPA (Natural Population Analysis) atomic charge using the CPCM (Conductor Like Polarizable Continuum Model) at the M06L/6-311++G** level ($R^2 = 0.93$). The experimental pK_a of a collection of quinazoline and derivatives were compared to several Conceptual Density Functional Theory descriptors computed. The highest approximations were observed when employing the M06L/6-311++G** with the CPCM solvation model using water as a solvent. In the final part of our study, M06L/6-311++G** have been used, in combination with CPCM continuum solvent model, to calculate the aqueous pK_a values of quinazoline derivatives by using an isodesmic reaction. Possible improvements to current methodology are suggested.

ÖZET

KİNAZOLİN TÜREVLERİNİN pK_a SABİTLERİNİN DEĞERLENDİRİLMESİNE HESAPSAL BİR YAKLAŞIM

Kimyadaki en önemli sabitlerden biri iyonlaşma (ayrışma) sabitidir (pK_a). Potansiyel bir ilaç için pK_a değerinin/değerlerinin güvenilir bir şekilde tahmin edilmesi, özellikle ilaç keşfi ve geliştirmesinin başlarında kritiktir. İyonize ilaçların yaklaşık % 68'inin zayıf bazlardan oluştuğu söylenmektedir [1]. Kinazolin ve türevlerinin iyonlaşma sabiti için verimli bir tahmin metodolojisi geliştirmede ilk adım olarak üç protokol üzerinde çalıştık. Birincisi, kuinazolin ve türevlerinin hesaplanan atomik yükleri ile deneysel pK_a arasındaki doğrusal ilişkiye dayanmaktadır. Gözlemlerimize dayanarak, gözlemlenen pK_a'ları yeniden üretmek için en uygun yöntem, M06L/6-311++G** düzeyinde ($R^2 = 0.93$) CPCM'yi (İletken Gibi Polarize Edilebilir Sürekli Modeli) kullanarak NPA (Doğal Nüfus Analizi) atom yükünü hesaplamaktır. Kinazolin ve türevlerinden oluşan bir koleksiyonun deneysel pK_a'sı, çeşitli yoğunluk fonksiyonelleri, temel setler ve çözme yöntemleri kullanılarak hesaplanan birkaç Kavramsal Yoğunluk Fonksiyonel Teorisi tanımlayıcısı ile karşılaştırıldı. En yüksek yaklaşımlar, çözücü olarak su kullanan CPCM solvasyon modeliyle M06L/6-311++G** kullanılırken elde edildi. Çalışmamızın son kısmında ise, kuinazolin türevlerinin sulu iyonlaşma sabiti değerlerini izodesmik reaksiyon kullanarak hesaplamak için metodoloji olarak M06L/6-311++G**//CPCM kullanıldı.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$\mathbf{A}$	Electron Affinity
E_{xc}	Exchange-correlation Energy Functional
E_0	Ground-state Energy
$F_{HK}[\rho]$	Universal Hohenberg-Kohn Functional
$\hat{\mathrm{H}}$	Hamiltonian
$\mathbf{I}$	Ionization Energy
$\mathbf{N}$	Number of Electrons
$\mathbf{r}$	Area
$\mathbf{T}$	Kinetic Energy Operator
$\mathbf{V}$	Potential Energy Operator
$\mathbf{V}(\mathbf{r})$	External Potential
V_{xc}	Exchange-correlation potential
$\mathbf{U}$	Interaction Energy Operator
$\mathbf{Z}$	Atomic Number
q	Atomic Charge
η	Chemical Hardness
ω	Global Electrophilicity Index
μ	Chemical Potential
ψ	Wavefunction
Θ	Normalized Wavefunction
ρ	density

LIST OF ACRONYMS/ABBREVIATIONS

ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicity
ATP	Adenosine Triphosphate
B3LYP	Becke-3-parameter Lee-Yang-Par functional
CAS	Chemical Abstract Service
COSMO	The Conductor-Like Solvation Model
CPCM	Conductor Like Polarizable Continuum Model
Def2TZVP	Valence Triple-Zeta With Two Sets Of Polarization Functions
DFT	Density Functional Theory
D-PCM	Dielectric Version of Polarizable Continuum Model
GB	The Generalized Born Approximation
GGA	Generalized Gradient Approximation
HF	Hartree-Fock Theory
HOMO	Higher Occupied Molecular Orbital Lower
IEF-PCM	Integral-Equation-Formalism Protocol
JDFT	Joint Density Functional Theory
LUMO	Unoccupied Molecular Orbital
LSDA	Local Spin Density Approximation
MAD	Mean Absolute Deviation
MD	Mean Deviation
MM	Molecular Mechanics
NPA	Natural Population Analysis
NPE	The Nonhomogeneous Poisson Equation
PCM	Polarizable Continuum Model
SMD	Universal Solvation Model
STO	Slater Type of Orbital
RMSE	Root Mean Square Error
QM	Quantum Mechanics

QSPR	Quantitative Structure Property Relationships

1. INTRODUCTION

1.1. Quinazoline Molecule

As a class, quinazoline and its derivatives became popular in industrial and research areas due to their diverse pharmacological activities. Quinazoline compounds are well known nitrogen containing heterocyclic compounds formed by two fused six-membered aromatic rings which are benzene and pyrimidine as shown in the Figure 1.1 The chemical formula of quinazoline is $C_8H_6N_2$ and it appears as a yellow crystalline substance. Any derivative of quinazoline is described as a quinazoline compound [4].

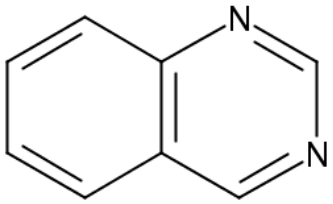


Figure 1.1. Quinazoline molecule.

Quinazoline derivatives are the privileged scaffolds in drug discovery due to their distinct and wide range of bioactivities. The functions of the quinazoline based compounds can be easily modified since highly delocalized π electrons in 3,4-double bond enhances the reactivity of quinazolines towards many types of nucleophiles [4].

Literature survey indicates that numerous quinazoline derivatives possess a broad spectrum of biological activities including anticancer [5–8], antiinflammatory [9, 10], anti-bacterial [11–14], analgesic [9–13], antituberculosis [15], antihypertensive [16], antidiabetic [17], antipsychotic [18] etc. Various quinazoline compounds are synthesized by the addition of numerous active substituents to the quinazoline molecule and potential applications have been explored [4].

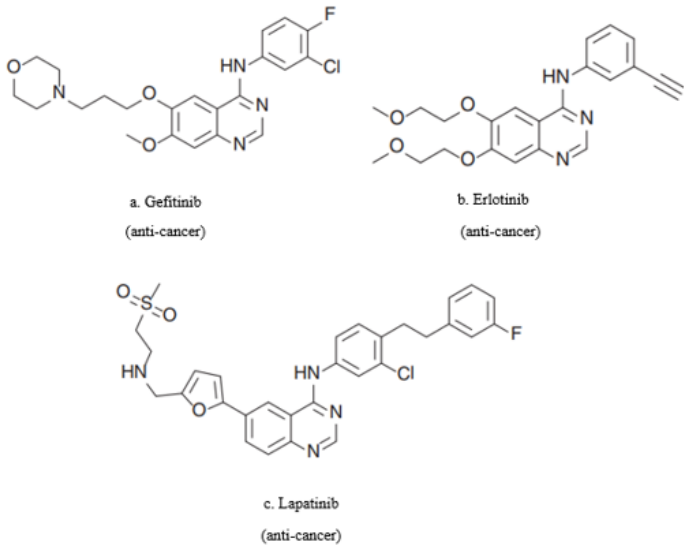


Figure 1.2. Schematic representations of the examples of quinazoline drugs [2].

Some of amino-quinazoline derivatives were found to be inhibitors of the tyrosine kinase, or dihydrofolate reductase enzymes so they work as potent anticancer agents [2]. Quinazoline compounds are used in the preparation of various functional materials for synthetic chemistry and, they are present in various drug molecules. Quinazoline nucleus can be considered as a significant scaffold of effective anticancer agents due to the advances in synthetic methodology and biological interactions [19]. If we need to give example on that, as shown in the Figure fig:Resim1.png, U.S. Food and Drug Administration has approved several quinazoline derivatives as anticancer drugs recently, such as gefitinib, erlotinib and lapatinib. They are 4-anilinoquinazolines. They inhibit kinase activity as they interact with the ATP-binding site, and they have led to a very high number of patents [2]. Quinazoline based pharmaceuticals are becoming more significant class of therapeutic agents and they are good bets to replace many other organic based pharmaceuticals anytime soon.

1.2. The Nature of pK_a Values

The concept of Brønsted-Lowry theory is the definition of acid and base ionization. Acids are substances that can ionize by giving a solvated hydrogen ion. On the

other hand, bases accept hydrogen ions. Ionization constants are small and demanding figures, and it has become more common to use their negative logarithms (known as pK_a values) [20].

Acidic and basic sites within a drug molecule governs the solubility, lipophilicity, and permeability in a cell membrane as well as ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties. Therefore, the biological activity of a ligand is dependent on the pK_a's of the ionizable functional groups [21]. The Brønsted-Lowry theory is the most used and accepted description of the ionization of both acids and bases. Brønsted (1923) showed the advantage of having the ionization of both acids and bases expressed on the same scale for the first time.

The pK_a of a Brønsted acid is perhaps one of the most significant physicochemical constant. It is a number that exposes a richness of chemical information about a material. It is defined as

$$pK_a \quad = -\log \mathrm{K}_a, \tag{1.1}$$

where K_a is the acid ionization constant (acidity constant) and can be described as

$$K_a = \frac{[H_3O^{+1}] \times [A^{-1}]}{[HA]} \tag{1.2}$$

for any acid species in water as represented by the equilibrium

$$HA_{(aq)} + H_2O_{(l)} \rightleftharpoons H_3O^{+1}_{(aq)} + A^{-1}_{(aq)} \tag{1.3}$$

displayed [22].

The Henderson-Hasselbach equation uses the pK_a and pH values to calculate the relative concentration of protonated and deprotonated forms [23]. The Henderson–Hasselbach equation defined as,

$$pH = pK_a + \log \frac{[A^{-1}]}{[HA]} \tag{1.4}$$

where K_{a} is the weak acid's dissociation constant, [HA] and $[A^{-1}]$ are the molarities of the weak acid and its conjugate base, respectively. It is used to calculate pH throughout a large portion of the titration range as represented in Equation (1.4)

Most medicines have at least one location that may reversibly protonate or deprotonate. The site is predominantly deprotonated if the pH is greater than the pK_a, otherwise, it is largely protonated. A molecule's binding and transport characteristics are substantially determined by the ratio of protonated and deprotonated versions [22].

Acid dissociation constant can be calculated by the thermodynamic cycle which considers the free energy of solvation, desolvation and deprotonation in the gas phase of the acid species. However, large uncertainties emerging from the solvation free energy of the proton H^+, instability of the species in the gas phase and large conformational differences between the gas and the solvent calculations make this method less applicable [24]. To improve the accuracy of estimated pK_{a} values, there is a tremendous effort based on computational approaches. Quantitative structure property relationship techniques, in addition to empirical methods such as PROPKA which is one of the most reliable pK_a predictors of [25] and methods based on Molecular Dynamics, Generalized Born equation, Poisson-Boltzmann equation, QM/MM, or a combination of one or more, provides faster and more accurate pK_a calculations by correlating molecular descriptors to the pK_a's of organic molecules. Topological state, atom type, group philicity, bond frequency, bond length, maximum surface potential, HOMO and LUMO energies, and atomic charges are all connected molecular descriptors. Atomic charges are one of them, and they have to do with a molecule's acidity and basicity.

1.3. Determining pK_a Values with Traditional Methods

The accurate determination of the acidity constants of organic reagents is often required in various chemical and biochemical areas and plays a fundamental role in many analytical procedures. The pK_a values affect the toxicity, chromatographic retention behavior, and pharmaceutical properties of organic acids and bases [26]. Experimental

determination of the acid dissociation constants can be made by various methodologies including H-NMR spectroscopy, capillary electrophoresis, FT-IR spectrometry, UV–VIS absorption and fluorescence spectrophotometry and potentiometry [27]. Traditionally, due to UV–VIS absorption spectrometry and potentiometry being simpler, easier for applications, low cost, more accurate and reproducible [27], they have been the method of choice for the determination of the acid dissociation constants [28].

Potentiometric titration requires the measurement of pH using a cell composed of two half cells known as electrodes. Potential changes as the H^+ concentration is changed. When potentiometry is inapplicable, ultraviolet spectrophotometry gets involved where both the ion and the molecule can be isolated in solution and observed independently. The pK_a values can be determined by potentiometric titration in 20 min, while the spectrometric methods may take as much time as half a working day. Different variety of methods can be used as alternatives to potentiometry and spectrophotometry, but each one has limited usefulness. The traditional conductimetry takes longer than potentiometry, and temperature control is painstaking, can also provide accurate results [20].

1.4. Determining pK_a Values with Computational Methods

The accuracy of the experimentally determined pK_a is affected when multiple tautomeric species are present in the medium, the concentration of the solution gets closer to the limits of quantification, the intermediates have short lifetimes, other solutes interfere, and the isolation cannot be performed properly [29–31]. Besides, experimental methods are less practical for the pK_a determination of large numbers of compounds. Therefore, the accurate pK_a calculations by theoretical approaches where different tautomeric contributions can be included in silico is of great interest.

Computational methods are also the choice for calculating pK_a values when the molecules can't be synthesized, or the experimental methods are not clear. For example, amino acids have pK_a values that vary according to their local environment, and

it makes the measurement troublesome [9]. In such circumstances, calculating acid dissociation constants computationally becomes significant and demanding. Countless studies are present in the literature that use numerous methods to obtain chemical accuracy. Recently, new developments have been made but still contradictions and discrepancies exist. There are different strategies for examining the individual pK_a calculations, but regardless of the reaction scheme, above all it is important to note that, the calculation of free energies in solution is generally carried out via the thermodynamic cycle.

1.4.1. Thermodynamic Cycle

Before heading over the diverse pK_a calculation methods, it's worth noting that, regardless of the reaction scheme used, the calculation of solution-phase reaction free energies is mainly handled using a thermodynamic cycle in which the solution-phase reaction free energies are calculated as the sum of the relating gas-phase free energy and the solvation free energy

$$\Delta G^*_{soln} = \Delta G^*_{gas} + \sum_{i=1}^{Nproducts} n_i \Delta G^*_{solv,i} - \sum_{j=1}^{Nreactants} n_j \Delta G^*_{solv,j}, \tag{1.5}$$

where * denotes a standard state of 1 mol/L [32].

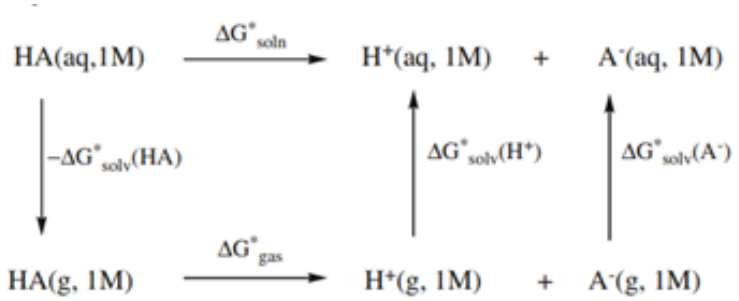
The key reason for using a thermodynamic cycle is that continuum solvation models are usually quantified to obtain reliable solvation energies, and the low levels of theory at which they are generally developed and built (such as small basis set HF or B3-LYP calculations) are rarely accurate enough to reproduce accurate total free energies in solution. Using a thermodynamic cycle, high-level ab initio calculations in the gas phase may be used to enhance the precision of the obtained reaction free energies [3].

Thermodynamic cycles are used to improve the effectiveness of solution-phase free energies, various types of proton exchange reactions are used to maximize error variance rescission, and/or water molecules are included in the different reaction schemes

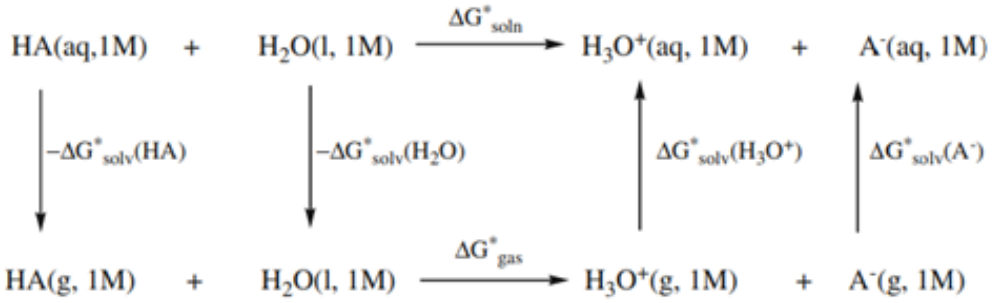
to enhance the analysis of explicit solute–solvent interactions. When you consider that these procedures may be used at different levels of theory and with alternative solvation techniques, the amount of potential pK_a calculation methodologies may even be intimidating [3]. Ho and Coote examined the performance of several thermodynamic cycles for the calculation of pK_a values in depth [3, 32–34].

1.4.1.1. Direct Method. The direct method, the most basic thermodynamic cycle, addresses the dissociation of an acid species HA into its corresponding base and proton. The first step is to desolvate HA in solution to gas phase, after which HA is deprotonated in gas phase to provide the conjugate base A^- and the isolated proton H^+. To complete the cycle, the resultant products A^- and H^+ are solvated [31].

All continuum solvent pK_a calculations, like those given in Figure 1.3, rely on a thermodynamic cycle that combines precise gas-phase acidity (experimental or computed using high-level ab initio techniques) with solvation free energies derived from different solvent models.



(a) Cycle A



(b) Cycle B

Figure 1.3. pK_a calculation via the direct method [3].

Despite the partial results claimed for this approach, it has a couple of major defects that restrict the application of Figure 1.3 as a comprehensive pK_a calculation technique. Both cycles produce two ionic compounds, and the accompanying inconsistencies in their solvation energies are substantially bigger, which might result in pK_a predictions with massive deviations.

1.4.1.2. Proton Exchange Method (Isodesmic Method). Isodesmic reaction is used to calculate the approximate free energies of deprotonation in solution. That method can be used when the case is about species with large conformational changes between gas and solution phases or when dealing with unstable species. Isodesmic reactions are also known as "proton exchange reactions" [1]. The method does not require explicit water and it takes the advantages of proton free energy of solvation being absent.

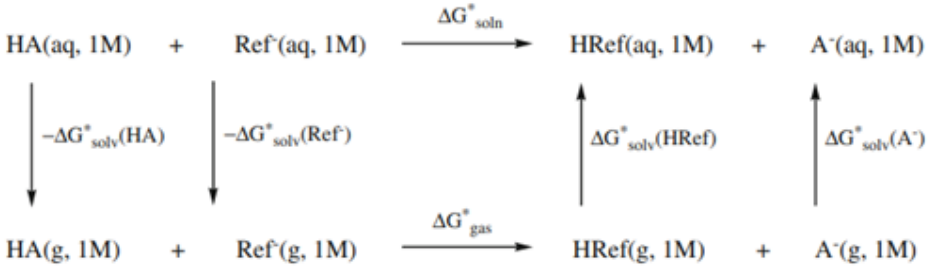


Figure 1.4. pK_a calculation via the proton exchange method [3].

The third thermodynamic cycle, known as the proton exchange approach, is based on an isodesmic reaction. In this cycle presented in Figure 1.4, the acid AH contributes a proton to a reference species Ref through an isodesmic reaction. The proton exchange method shown in Figure 1.4 takes a broader concept to computing pK_a by using a reference species to minimize the deviations caused by the solvent model and gas-phase free energy calculations. The method is ideal when the deviations in the computed pK_a s are predictable, which is most probable if the reference is structurally identical to the molecules under examination. Quinazoline is a strong base and to be able to precisely estimate the pK_a value it crucial to understand thermodynamic cycle in the light of literature.

1.4.1.3. Literature Background of Computational Methods. The effectiveness of various thermodynamic cycles for generating pK_a values was thoroughly investigated. The first and the simplest method, the directly calculated pK_a can be calculated through cycles A and B which are presented previously via Eqs. 1.6 and 1.7, respectively. The correction term in the Eq. 1.7 is required to account for the liquid water's standard condition, which is 55 mol/L. According to Bryantsev's study, the inappropriate assignment of standard state for water molecules has contributed to methodological problems in a series of studies [38]. They noted that generating a thermodynamic cycle correctly necessitates that each reactant or product in the upper and lower legs be in the exact same standard state and this is the most important aspect even though many authors had ignored or neglected that [38]. Pliego also addressed that issue as well. He claimed that cycle B produces a gas phase mixture of standard states. The HA, A^-, and H_3O^+

species all maintain a standard concentration of 1 mol/L, whereas H_2O maintains a pressure of 1 atm. According to his analysis, the improved the outcome achieved with the thermodynamic cycle B is just coincidental [39]. The misconception arises from the fact that the standard state for solutes in solution is 1 mol/L, while water is supposed to have a standard state of 1 mol/L when it functions as a co-reactant (e.g. cycle B in Figure 1.3).

Ho and Coote had calculated pK_a values for several acids by putting their experimental gas-phase acidities and solvation free energies into an equation defined as

$$pK_a = \frac{\Delta G^*_{soln}}{RT \ln 10}, \tag{1.6}$$

and to ensure that the equation is the proper formulation they defined it as

$$pK_a = \frac{\Delta G^*_{soln}}{RT \ln 10} - \log H_2O. \tag{1.7}$$

They had conducted their benchmarking analysis on a test set of 55 neutral species, which included alcohols, phenols, carboxylic acids, inorganic acids, and various carbon acids with different functionalities. They reported that, cycles A and B provide identical results that are quite close to the observed pK_a values. In practice, given that these "experimental" solvation free energies were derived using cycle A with experimental pK_a and gas-phase reaction energies, great consistency with experiment is almost surely confirmed by Ho and Coote [3] .

In fact, whenever a water molecule is present on the reactant or product part of the thermodynamic cycle, $\log[H_2O]$ can be simply removed or added to the pK_a value. If each quantity in cycles A and B is known with infinite precision, they should yield the identical results.

As an alternative to the direct cycle, is the proton exchange method. The computed ΔG_{aq} in proton exchange method cannot be directly connected to the pK_a of the acid HA since it is stated as the pK_a difference with the reference species HRef as

shown

$$pK_{aHA} = \frac{\Delta G^*_{soln}}{RT \ln 10} + pK_{a(HRef).} \tag{1.8}$$

This method had been successfully applied in the case of phenols [35–37], amines [35], pyridines [3,29,33,35–40], carboxylic acids [29,35–37,40–45], aliphatic alcohols [33], and amides [46] amongst others. Ho and Coote reported that while cycle A produces mistakes ranging from 5 to 12 pK_a units for carbon acids, the proton exchange method yields errors ranging from 2–4 pK_a units throughout most situations, and even 1 pK_a unit in the condition of neutral carbon acids [3].

However, the effectiveness of this strategy is strongly dependent on the reference acid selected [24]. For the validity of the pK_a computations, the reference species must always be carefully chosen in this methodology. It is confirmed, as in earlier studies, that identifying a reference species with a charge distribution around the acid group that is identical to the analyzed species is crucial (1, 25, 51, 52). This approach can be subjective since, if appropriate, a reference species with identical geometry, electrostatic distribution, solute–solvent interactions, intramolecular interactions and acidity to HA should be chosen [24].

Sastre and his co-workers analyzed the pK_a of different functional groups and they had computed pK_a values by using a structurally distinct species to have a better understanding of the effect of the reference species [1]. The pK_a of benzoic acids, phenols, and pyridines were computed employing acetic acid, ethanol, and ethylamine as reference species, respectively. They reported that when utilizing acetic acid as the reference species, MAD values of 0.5 pK_a units are observed, which is similar to the accuracy achieved when benzoic acid was picked as the reference species. When applying ethanol as the reference species, meanwhile, the deviations in phenol pK_a values are substantially greater (MAD values between 4.0 and 6.5 pK_a units). The MAD values in the pK_a of pyridines rise to 2.8 pK_a units when ethylamine is chosen as reference for CPCM calculations but remain closer to the MAD values when using pyridine as reference for SMD calculations [1]. As a result of their work, it is vital

to note that molecular geometry and conformational similarity should not be the sole factor to keep in mind when selecting a reference species. Even though, the reference species has an impact on the reliability of the estimated pK_a values, the correction of errors in the isodesmic reaction offers more flexibility in reference species selection. Looking at the whole picture, this approach is worthwhile to examine for theoretical pK_a estimates, particularly in circumstances when the thermodynamic cycles reveal issues with gas-phase calculations.

2. Aim of the Study

In this study, the aim is to expand the study of Roos [54] with the technique of substituted thiols and alcohols, including phenols and the work of Uğur [47] with the reproduction the pK_a's of alcohol and thiol, in this paper. The focus was on the linear regression step, and to look at how experimental pK_a values are connected to the atomic charge of the basic forms quinazoline and derivatives. Multiple issues are handled in this study: Is it possible to reproduce an experimental pK_a by using atomic charges, conceptual DFT descriptors, and isodesmic reactions? What effect would a basis set, solvation model or a DFT functional have on the fitting? What is the expected pK_a 's stability in consideration of the molecule's conformations and geometries? Considering the successful applications and findings of QM charges as descriptors, conceptual DFT descriptors and isodesmic reactions in the light of literature, the objective of this research was to produce an accurate protocol for estimating the pK_a values of quinazoline and its derivatives rapidly. We aimed to evaluate the robustness of the isodesmic reaction in the calculation of pK_a values of quinazoline and its derivatives and, to compare the results with those calculated with descriptors and atomic charges. The criteria determined and followed to pick the most ideal reference species HB was the resemblance of chemical structure with the studied species HA so; we selected quinazoline. We aimed to demonstrate the validity of this methodology for a group of 46 drug like quinazoline derivatives.

The calculations were performed on a total of 46 quinazoline derivatives (see Table 2.1). The ability to represent the broadest range of experimental pK_a 's was our primary criterion for selecting the molecules. The experimental pK_a's vary from 2.4 to 8.98. The 46 molecules are composed of moieties that have fragments like those of the kinase inhibitors. Drug-like compounds are often bigger and more complicated than the ones employed in this study. It is important to limit the size of the molecules and the number of rotatable bonds to have a more straightforward conformation search. Among all the molecules studied 30 of them were used to construct the training set

which will be represented in the next sections, and the remaining 16 molecules formed our test set. While we were separating training set from test set, we wanted to make sure that we had various types of molecules. For instance, the experimental pK_a's vary from 2.4 to 8.98 and our training set vary from 2.85 to 8.98. Our molecules have different molecular sizes, in our training set, we made sure that we have small-sized and large-sized molecules to have more reliable results. 2D representation of all compounds, training set and test set are displayed in Table 2.1

Table 2.1. 2D representations of 46 molecules.

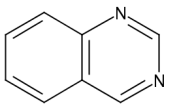
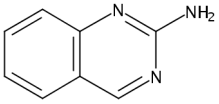
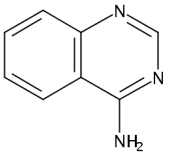
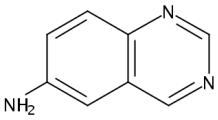
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
Quinazoline	253-82-7	1		3.31
2-Quinazolinamine	1687-51-0	2		4.43
4-Quinazolinamine	15018-66-3	3		5.73
6-Quinazolinamine	101421-72-1	4		3.2

Table 2.1. 2D representations of 46 molecules. (cont.)

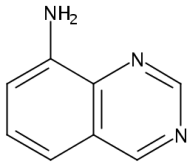
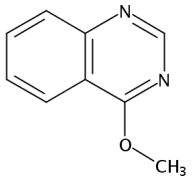
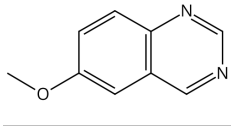
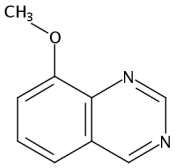
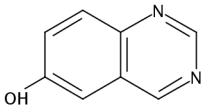
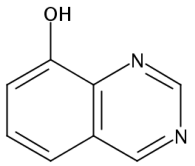
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
8-Quinazolinamine	101421-74-3	5		2.4
4-Methoxyquinazoline	16347-95-8	6		3.13
6-Methoxyquinazoline	7556-92-5	7		2.85
8-Methoxyquinazoline	7557-01- 9	8		3.51
6-Hydroxyquinazoline	7556-93-6	9		3.12
8-Hydroxyquinazoline	7557-02-0	10		3.41

Table 2.1. 2D representations of 46 molecules. (cont.)

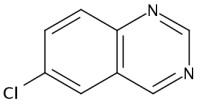
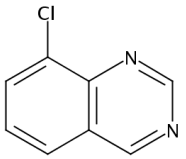
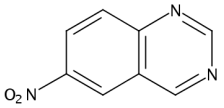
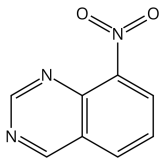
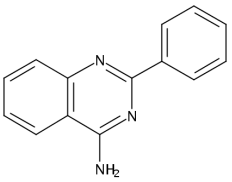
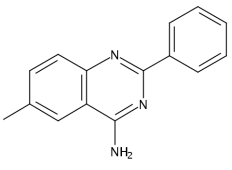
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
6-Chloroquinazoline	700-78-7	11		4.18
8-Chloroquinazoline	7557-04-2	12		4.00
6-Nitroquinazoline	7556-95-8	13		3.55
8-Nitroquinazoline	7557-05-3	14		3.30
2-Phenyl-4-quinazolinamine	1022-44-2	15		5.44
6-Methyl-2-pheny-l-4-quinazolinamine	310440-96-1	16		5.16

Table 2.1. 2D representations of 46 molecules. (cont.)

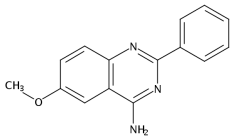
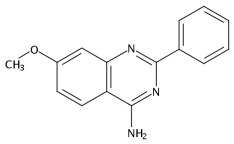
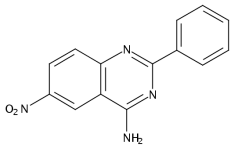
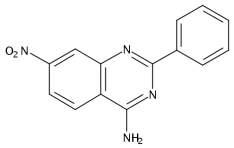
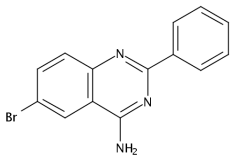
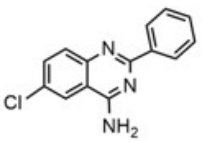
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
6-Methoxy-2-pheny-l-4-quinazolinamine	310440-98-3	17		5.33
7-Methoxy-2-pheny-l-4-quinazolinamine	310440-99-4	18		5.62
6-Nitro-2-pheny-l-4-quinazolinamine	310441-00-0	19		4.54
7-Nitro-2-pheny-l-4-quinazolinamine	310441-01-1	20		4.27
6-Bromo-2-pheny-l-4-quinazolinamine	93716-83-7	21		4.78
6-Chloro-2-pheny-l-4-quinazolinamine	310441-02-02	22		4.98

Table 2.1. 2D representations of 46 molecules. (cont.)

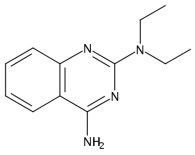
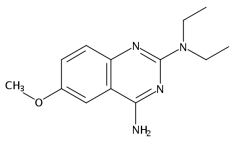
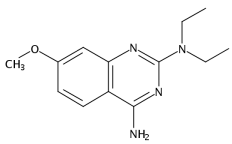
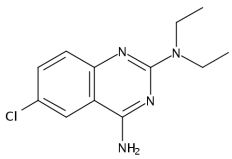
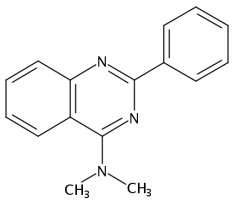
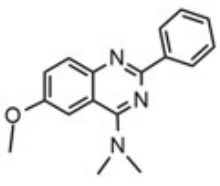
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
2,4-Quinazolinediamine-N2,N2-diethyl	7461-77-0	23		7.79
N2,N2-Diethyl-6-methoxy-2,4-quinazolinediamine	523979-74-0	24		7.82
N2,N2-Diethyl-7-methoxy-2,4-quinazolinediamine	523979-75-1	25		8.31
6-Chloro-N2,N2-Diethyl-2,4-quinazolinediamine	523979-76-2	26		6.98
4-(Dimethylamino)-2,4-quinazolinediamine	139474-19-4	27		6.31
6-Methoxy-N,N-dimethyl-2-phenyl-4-quinazolinamine	299196-54-6	28		6.61

Table 2.1. 2D representations of 46 molecules. (cont.)

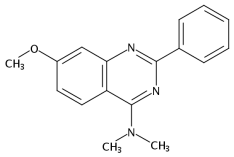
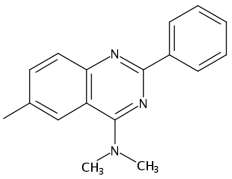
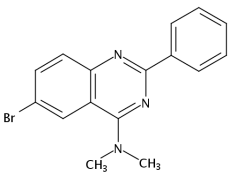
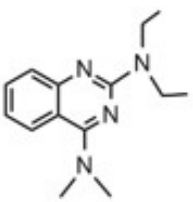
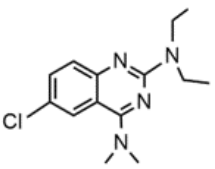
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
7-Methoxy-N,N-dimethyl-2-phenyl-4-quinazolinamine	158832-82-7	29		6.20
4-Quinazolinamine-N,N,6-trimethyl-2-phenyl	158832-77-0	30		6.52
6-Bromo-N,N-dimethyl-2-phenyl-4-quinazolinamine	158832-79-2	31		5.88
2,4-Quinazolinediamine-N2,N2-diethyl	95033-68-4	32		8.88
2,4-Quinazolinediamine-6-chloro-N2,N2-diethyl-N4,N4-dimethyl	299196-57-9	33		7.91

Table 2.1. 2D representations of 46 molecules. (cont.)

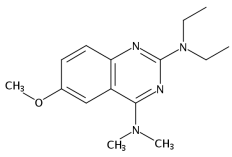
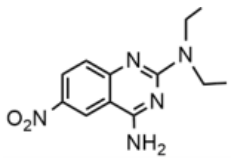
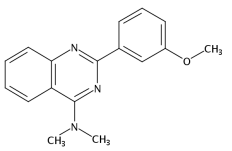
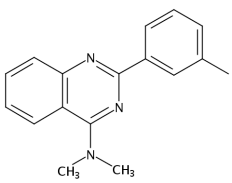
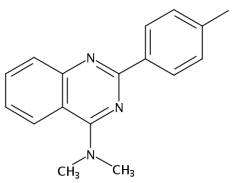
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
6-Methoxy-2-(diethylamino)-4-(dimethylamino)-quinazoline	299196-56-8	34		8.98
6-Nitro-2-(diethylamino)-4-(dimethylamino)-quinazoline	7154-34-9	35		6.63
2-(3-Methoxyphenyl)-N,N-dimethyl-4-quinazolinamine	959484-96-9	36		6.62
N,N-Dimethyl-2-(3-methylphenyl)-4-quinazolinamine	180906-20-1	37		6.40
N,N-Dimethyl-2-(4-methylphenyl)-4-quinazolinamine	180906-17-6	38		6.28

Table 2.1. 2D representations of 46 molecules. (cont.)

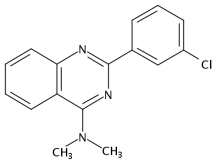
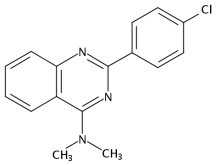
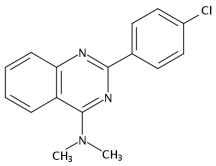
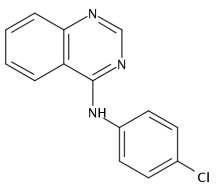
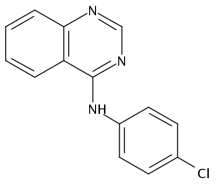
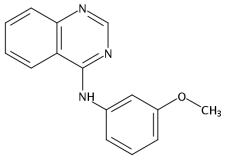
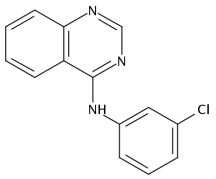
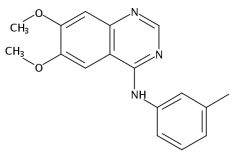
Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
2-(3-Chlorophenyl)-N,N-dimethyl-4-quinazolinamine	180906-21-2	39		5.88
2-(4-Chlorophenyl)-N,N-dimethyl-4-quinazolinamine	180906-18-7	40		5.54
N-(3-Trifluoromethyl-phenyl)-4-quinazolinamine	47155-57-7	41		5.03
N-(4-Chlorophenyl)-4-quinazolinamine	34923-97-2	42		6.02
N-Phenyl-4-quinazolinamine	34923-97-2	43		6.08

Table 2.1. 2D representations of 46 molecules. (cont.)

Name of the molecule	**CAS Number**	**Number**	**2D**	**exp** pK_a
N-(3-Methoxyphenyl)-4-quinazolinamine	146885-03-2	44		5.37
N-(3-Chlorophenyl)-4-quinazolinamine	88404-44-8	45		5.28
6,7-Dimethoxy-N-(3-methylphenyl)-4-quinazolinamine	666839-07-2	46		5.77

Quinazoline derivatives which are unsaturated heterocyclic bases are studied by Adrien Albert [48], Wojciech Zielinski [49], Aramarego [**?**] and Mehtap Işık [50]. They recorded the ionization constans for various quinazoline derivatives. The strengths of the heterocyclic derivatives generated from quinazoline have been examined and evaluated in their studies. Adrien Albert reported that the heterocyclic systems' ionization constants differ substantially from saturated systems because they contain a component not present in the latter. This component occurs from an alteration in the energetics of the resonating system of x electrons (due to ionization), a system in which the basic centre is deeply engaged [48].

We separated 46 selected quinazoline molecules to two groups which include thirty and sixteen molecules respectively. To reflect the widest range of experimental pK_a's was our primary consideration for putting the molecules into two different subgroups.

Our training set consists of 30 molecules as represented in Table 2.2. Our aim is to find a good correlation between the atomic charges of the training set and the experimental pK_a values, and for the next step, according to the equation we obtained from linear regressions, we aimed to predict the pK_a values of our test set molecules.

Table 2.2. Training Set molecules.

Number of the molecule	**Experimental** pK_a
1	3.31
3	5.73
4	3.2
7	2.85
8	3.51
9	3.12
10	3.41
11	3.55
12	3.3
13	4.18
15	5.44
16	5.16
17	5.33
18	5.62
19	4.54
21	4.78
22	4.98
23	7.79
24	7.82
27	6.31
28	6.61
29	6.2

Table 2.2.Training Set molecules. (cont.)

32	8.88
33	7.91
34	8.98
36	6.62
37	6.4
38	6.28
39	5.88
40	5.54

3. Theoretical Background

This section elaborates the fundamentals of the most broadly applied theoretical techniques in this thesis, including quantum mechanics.

3.1. Density Functional Theory

Density Functional Theory has a special place in modern quantum chemical methods and is a theory of electronic ground state structure based on the electronic density distribution $\rho(r)$ based on the electron density [51]. It has become much more effective for the comprehension and computation of the ground state density, $\rho(r)$, and energy of the studied substances, clusters, and solids, in a word, any system consisting of nuclei and electrons with or without applied static disturbances, since its start which was approximately three decades ago [51]. It's a different and complementary approach to classic quantum chemistry approaches, which are based on the many-electron wave function $\Psi(r)$ $(r_1,\ldots, r_N)$.

DFT attempts to address both HF and post-HF methods by replacing electronic wavefunction with the electronic density. Density Functional Theory rests on two main theorems by Kohn and Hohenberg. The first theorem establishes that the energy of an electron distribution is described by the electronic density function which is minimum for the ground state energy. The second theorem states that the ground state electron density determines all ground state properties of the electronic system [52]. Modern DFT may be attributed to the Thomas-Fermi and Hartree-Fock-Slater approaches. Modern DFT, on the other hand, is in theoretically precise, while those theorems are intrinsically approximated [51]. The fundamentals of DFT are sketched below with the kinetic, potential, and interaction energy operators T

$$T \equiv -\frac{1}{2}\sum_j \nabla_j^2, \tag{3.1}$$

V

$$V \equiv \sum_{j} \upsilon(r)_j, \tag{3.2}$$

and U

$$U \equiv \frac{1}{2} \sum_{i \neq j} \frac{1}{|r_i - r_j|} \tag{3.3}$$

respectively [51, 53, 54] and overall H is approximated as

$$H \equiv T + V + U. \tag{3.4}$$

3.1.1. Thomas-Fermi Theory

The use of the electron density rather than the wave function goes all the way back to Thomas and Fermi's pioneering studies [64], [65]. First, it is crucial to describe the electron density. Electron density, defined by $\rho(r)$ is the chance of detecting any of the N electrons with arbitrary spin within the area r, whereas the remaining N-1 electrons have flexible orientations and spin in the state denoted by ψ. The function of integrating to the number of electrons in total is described as

$$\rho(r) = N \int \cdots \int |\psi(x_1, x_2, \cdots, x_N)|^2 \, ds_1 x_2, \cdots, x_N \tag{3.5}$$

$$\int \rho(r) \, dr \; = N. \tag{3.6}$$

Quantum statistical theory based on a uniform electron gas is used to compute the kinetic energy of electrons, while the interactions between electrons and nuclei (electron-nucleus and electron-electron) are considered in a classical manner. The electrons' kinetic energy is defined as

$$T[\rho] = C_F \int \rho^{5/3}(r) \, dr. \tag{3.7}$$

and it is fulfilled by inserting the electron-nucleus and electron-electron interactions, yielding a total energy in terms of electron density shown as

$$E[\rho] = C_F \int \rho^{5/3}(r) \, dr \; - Z \int \frac{\rho(r)}{r} \, dr \; + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} \, dr_1 dr_2. \tag{3.8}$$

This basic Thomas-Fermi model in which the electron system is portrayed more like a standard liquid, is adequate not for how successfully it computes the ground state energy and density, but rather for demonstrating that the energy can be determined solely by the electron density [55, 56].

3.1.2. Hohenberg-Kohn Theorem

Hohenberg and his co-workers had constructed a precise variational method for the ground-state energy, wherein the density has been the variable function. A general functional is introduced into this theory, which applies to all electronic systems in their ground state, regardless of the external potential [57].

They considered a container consisting of a group of electrons that is affected by an external potential $\upsilon(r)$ We think that we know this system's electron density, that defines $\upsilon(r)$ and therefore all parameters. When there's another external potential $\upsilon'(r)$ that is more than a constant different from $\upsilon(r)$, it could also yield the same electron density $\rho(r)$ for the ground energy level. There will be two Hamiltonians ($\hat{\mathrm{H}}$ and $\hat{\mathrm{H}}$') whose ground state electron density is the same but the normalized wave functions would be again two (Φ and Φ') and the ground-state energies E_0 and E_0' are the for $\hat{\mathrm{H}}$ and $\hat{\mathrm{H}}$', respectively. The apparent disagreement is indicated as

$$\begin{aligned} E_0 < (\phi'|\hat{H}|\phi') &= (\phi'|\hat{H}'|\phi') + (\phi'|\hat{H} - \hat{H}'|\phi') \\ &= E_0' + \int \rho(r)[\upsilon(r) - \upsilon'(r)]\, dr \end{aligned} \tag{3.9}$$

$$\begin{aligned} E_0' < (\phi|\hat{H}|\phi) &= (\phi'|\hat{H}|\phi') + (\phi|\hat{H}' - \hat{H}'|\phi) \\ &= E_0' + \int \rho(r)[\upsilon(r) - \upsilon'(r)]\, dr \end{aligned} \tag{3.10}$$

$$E_0 + E_0' < E_0' + E_0 \tag{3.11}$$

and an outcome, no two external potentials can achieve the same output ρ(r). As a conclusion, ρ(r) determines υ(r) and all groundstate attributes individually. Hohenberg and his co-workers had expressed the energy E as a function of the electron density accurately described as

$$\begin{aligned} E[\rho] &= T[\rho] + T_{ne}[\rho] + V_{ee}[\rho] \\ &= \int \rho(r)\upsilon(r)\, dr + F_{HK}[\rho]. \end{aligned} \tag{3.12}$$

It is crucial to note that $\mathrm{F}_{HK}[\rho]$ is only dependent on and a universal functional of ρ and independent from any external potential υ(r). As a consequence, the energy will be at its lowest only if the electron density is at its smallest.

Even though Hohenberg-Kohn theorem demonstrated that the total energy can be determined from the ground state density, it remained unclear how to get the $\mathrm{F}_{HK}[\rho(\mathrm{r})]$ and ρ(r) coordinates. The inadequacy of Thomas-Fermi theory, according to Kohn and Sham, was mostly due to a poor definition of kinetic energy. To fix this issue, they intended to reintroduce the concept of one electron orbitals and use the kinetic energy of non-interacting electrons to estimate the kinetic energy of the system [58]. So the main equation in Kohn-Sham DFT is

$$(1/2\nabla^2 + \upsilon(r') + \int \frac{\rho(r)}{|r - r'|}\, dr' + \upsilon_{xc}(r))\phi_i = \epsilon\phi_i, \tag{3.13}$$

where ϕ are the Kohn-Sham orbitals. The terms from left side to right side in Equation (3.13) are the kinetic energy of the non-interacting reference system, the external potential, the Hartree potential, and exchange-correlation potential, respectively. With the final electron density, the total energy will be determined using the experssion

$$E = \sum_{i}^{N} \epsilon_i - \frac{1}{2} \iint \frac{\rho(r)\rho(r')}{|r - r'|}\, drdr' \; + E_{xc}[\rho] - \int \upsilon_{xc}(r)\rho(r)\, dr. \tag{3.14}$$

We need to know the nature of the exchange correlation energy functional in addition to applying the Kohn-Sham equations. The actual form of E_{xc}, however, is unknown and might not be discovered, that is why some approximations are neccessary which will be adressed in the next section.

3.1.3. Exchange-Correlation Functionals

We could compute the precise ground state density and total energy if each component in the Kohn-Sham energy functional was known. Sadly, one factor, the exchange-correlation (xc) functional, is undefined (E_{xc}). Perdew and his coworkers

demonstrated an efficient approach of classifying the numerous and diverse exchange-correlation functionals and that way is known as "Jacob's Ladder" [59]. On the rungs of a ladder, functionals are categorized according to their complication. The forms of exchange-correlation functionals used in current work are discussed briefly below.

The simplest fundamental approximation is LDA, (Local Density Approximation) which is an improvement over HF, where the energy is dependent on the density at the point where the functional is analyzed. A breakthrough came with the creation of functionals belonging to the generalized gradient approximation (GGA). GGA functionals can make a better description because they incorporate a dependence also on their gradient, not only on the electron density. Some examples for the GGA functionals are PBE [60] and BP86 [61] (Becke 1988). The next important progress was the introduction of hybrid functionals. Hybrid functionals mix GGA functionals with HF exchange. An example for hybrid functional is B3LYP [61, 62] which is commonly used because it yields good results and shows good performance for a wide variety of chemical systems. Even though some failures are identified, this method still is a dominant choice [52].

More recently, "meta-GGA" functionals are included with the methodological developments. With that step, GGA corrections had extended to higher derivatives and the "double-hybrid" functionals [63], [64] came into question [52]. In the future, we will likely see more functionals with wider range of applicability.

3.2. Basis Sets

Basis sets are set of functionals which build molecular orbitals and they enlarge as a linear combination. In the light of various number of publications that describe different basis sets, finding a set of function which is flexible enough to have good results and being economically and computationally tractable is not very easy [65].

The level of DFT functional and the basis set must be selected according to the behavior of the system. The first thing is to choose a good set of primitives. The number of primitives has an important effect on the cost and time to do the necessary integrals. Hence, smaller primitives can be considered. But for some molecular calculations, atomic basis sets require supplementation. There are some most widely used basis sets of all time developed by various chemists. STO-3G basis set is one of the most used minimal basis sets developed by Pople and his co-workers. And there are "split valence" basis sets such as 4-21G. And there are more extended Pople Basis Sets such as 6-311G** which provide better results. To obtain better accuracy and flexibility, polarization space is introduced by Pople and his co-workers. Polarization functions add higher angular momentum orbitals, and they are denoted by an asterisk (*) or "d". That addition is on heavy atoms but by a second asteriks or "d,p" polarization can also be added on lighter atoms. Additionally, they indicated "++" which are diffuse functions on all atoms. Diffuse functions are significant because they consider anions, and excited states. One "+" adds a set of diffuse s and p orbitals to the heavy atoms and second "+" add as a s function to hydrogen.

Another chemist, Dunning suggested high quality constructed basis functions that provide more flexibility [66]. Wiberg and his co-workers made a study about comparing the Pople basis sets and Dunning basis sets. According to their study, the aug-cc-pVDZ basis set generally provides less satisfactory results than 6-311++G**. They indicated that 6-311++G** basis set is smaller and significantly has higher computational efficiency [66].

3.3. Solvation Models

In quantum chemistry, the effect of the solvation model which mimic the solvent environment is also very significant. There are developments about the parametrization of atomic surface tensions. As shown in Figure 3.1, one can model first solvation shell by including or excluding first shell of solvent [67].

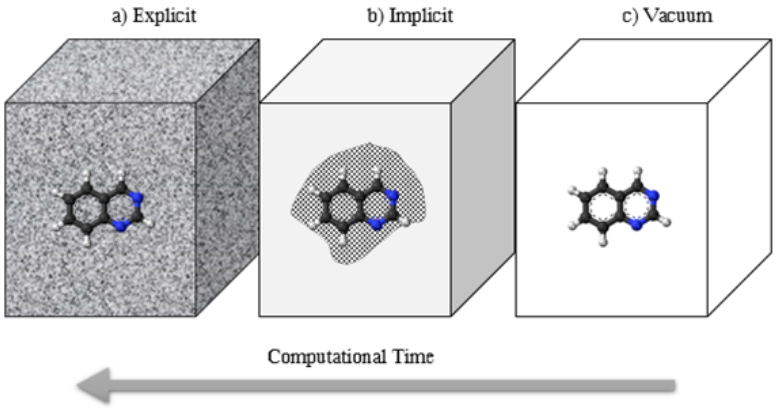


Figure 3.1. An illustration of the three solvent regimes. (a) explicit water solvent, (b) implicit solvent, and (c) vacuum.

Because of the diversity of the interface and the complexities of the solid and liquid materials involved, gaining a thorough knowledge of electrochemical interfaces through experiments is tricky. Computational studies offer alternative approaches to advance our fundamental knowledge of solid/liquid interfaces and forecast the features of new substances interfaces [67–69]. There are two basic approaches to treating solid/liquid interactions computationally which will be discussed.

Within experimental standard deviation of ± 1 kcal/mol, the solvation free energy of tiny neutral molecules may be estimated using different models [70, 71]. When dealing with ionic solutes with concentrated charge densities and strong local solute-solvent interactions, reliable prediction of solvation free energy for ions with continuous dielectric solvent models that are not parametrized for the same is a tough challenge [72–74]. Because continuum approaches can't correctly capture short-range intermolecular interactions like hydrogen bonding, explicit solvent molecules have been included to simulate ionic systems [75–77]. Because of the great number of solvent molecules essential to obtain the fundamental equilibrium features, this approach is extremely costly [68].

The electronic structure issue for a molecule in a liquid is minimized to the dimension of the solute of concern by using the continuum approximation for the solvent. Continuum solvation models (also known as implicit solvent models) depict a

solvated molecule at the atomic level inside an electrostatic cavity which is molecule-sized (and frequently molecule-shaped) and surrounded by a dielectric medium that symbolizes the solvent. Outside the solute cavity, the relative permittivity is commonly equated to the bulk solvent static dielectric constant, and inside the cavity, to a lower value. If the solute polarization is explicitly evaluated, the smaller value is generally accepted as unity (that is, the relative permittivity of vacuum) [78].

Another option is to consider the solute quantum mechanically and the solvent as a continuum, which implies that the solute is submerged in a solvent bath and the average of the solvent degrees of freedom is implicit in the solvent bath's characteristics. Fatteberg and Gyi pioneered implicit solvation models for plane wave DFT codes [79]. Aria advanced independently and put into the rigorous framework of JDFT (Joint Density Functional Theory) [80]. They had presented the first theoretically - based ab initio technique to ab initio computations in a continuous dielectric environment. Marzari extended the study by including a model for cavitation and dispersion [81]. All these developments and new approached presented more tractable ways for solvation models many approaches and have been used to understand these complex systems.

First, density-based solvation models must be considered. NPE (the nonhomogeneous Poisson equation) solvers which use continuous charge density instead of scattered point charges or multipoles to approximate it. The polarizable continuum model (PCM) [82, 83] is an example for density-based solvation models. By adding a charge distribution σ distributed across the cavity surface, the electrostatic challenge of assessing the interaction energy between solute and solvent, including mutual polarization effects, is handled in this solvation model [84]. There are numerous PCM formulations such as the integral-equation-formalism protocol (IEF-PCM) [84–87], the dielectric version of PCM (D-PCM) [82,83,88] and the conductor-like screening algorithm [71,89–94]. Due to the requirement of enhancing the prediction of the influence of bulk electrostatics with non-electrostatic and non-bulk contributions, both NPE solvers and Coulomb's law models lack validity (described below), as well as inaccuracies in the bulk electrostatic aspect itself. Due to the definition, size, and structure of the solute cavity, the

part of the solute charge that may lie outside the cavity, and the expected way in which the permittivity varies at and near the solute-solvent border, existing approaches based on the NPE contain limitations [95].

The generalized Born (GB) approximation is a continuum model as an alternative solvation model, which does not begin with the NPE and instead utilizes a beginning point based on Coulomb's law to describe the solute as a collection of point charges placed at the nuclear locations cite [78, 90, 95–99]. GB approach relies on partial atomic charges. Its precision for a given level of electronic structure theory may be contingent on the potential to generate valuable partial charges for that level of theory, but this possibility is not assured for all feasible theory levels and basis sets. As a result, density-based solvation models are considered to be less sensitive to the basis set employed.

Significant advances in recent years had been made and several SMx solvation models had been proposed and developed utilizing commonly accessible solvent descriptors that are applicable not just to water but to any organic solvent [66, 90, 100].

Most recently, SMD [101] solvation model is broadly used by chemists. The quantum mechanical charge density of a solute molecule interacts with a continuous description of the solvent in a continuum solvation model. The model is known as SMD, with the "D" denoting for "density" to indicate that the entire solute electron density is taken rather than partial atomic charges being described. The term "continuum" refers to a dielectric medium with surface tension at the solute-solvent boundary rather than an explicit representation of the solvent. SMD stands for "universal solvation model," meaning it may be used to any charged or uncharged solute in any solvent or liquid media for which a few fundamental elements including such dielectric constant, refractive index, bulk surface tension, and acidity and basicity parameters are provided [88].

For solvation energies of the neutral species, our study deals with two different solvation models, SMD [78] and CPCM [102] solvation model which are implemented in a Density Functional Theory. SMD is a universal solvation model which was proposed by the Cramer and Truhlar groups [113], and it can be applied to any charged or uncharged solute in any solvent or liquid medium. for which some descriptors are known (in particular, dielectric constant, refractive index, bulk surface tension, and acidity and basicity parameters) [78]. SMD solvation model is based on the quantum mechanical charge density of a solute that interacts with a solvent which is note represented explicitly [102].

Another solvation model that we had used is the CPCM i.e. Conductor like Polarizable Continuum Model. The conductor-like solvation model (COSMO) was proposed by Klamt and Schüürmann [71] for classical calculations and then extended to quantum mechanical systems. COSMO approach describes the solvent reaction by means of polarization charges distributed on the cavity surface [92]. According to CPCM solvation model the solute is placed in a cavity of roughly molecular shape [102].

The continual discovery of novel and accurate solvation models will undoubtedly be helpful in the future.

4. Literature Background on Prediction of pK_a

pK_a has been calculated using a range of techniques and methodologies, and the link between pK_a and a variety of theories has also been described in the literature, so it is crucial to consider the findings that we took into consideration.

Dixon and Jurs calculated the pK_a of oxyacids by using the empirical atomic charges and they developed a multiple linear regression with an accuracy of 0.5 units [103]. Numerous combinations of calculations for phenols and various computational strategies are tested by Varekova' [104]. They came up with the result that the partial atomic charges are good QSPR models for the estimation of pK_a. More than 25 percent of the models that they had analyzed (22 out of 83) have excellent quality and statistical criteria ($R^2 > 0.95$), and more than 50 percent (47 out of 83) had very good statistical criteria ($R^2 > 0.9$) [117].

Moreover, Uğur and her co-workers used a similar approach to predict amino acid pK_a values in thiols and alcohols recently [55]. They reproduced the pK_a values of alcohols and thiols in aqueous solution from the information of the experimental pK_a values of phenols, alcohols, and thiols. They devised a precise approach by establishing a link between the anionic form of the molecules' atomic charges and their experimental pK_a values. They managed to find that using NPA charges with the CPCM model at the B3LYP/3-21G and M062X/6-311G levels for alcohols ($R^2 = 0.995$) and thiols ($R^2 = 0.986$) levels is the best combination to regenerate the experimental pK_a among all tested DFT functionals, basis sets, semiempirical methods, solvation models, and charge models [40]. pK_a's of carboxylic acids [118]. The best combination of DFT functionals and basis sets were used in their study was M06L/6-311G(d,p) ($R^2 = 0.955$).

Theoretical chemistry aims to explore latest concepts and principles that can be used in the understanding of the chemical reactivity. One approach that has been

developed and used recently is density functional theory (DFT), which interprets and predicts reactive behavior of the molecules. DFT is based upon the Hohenberg–Kohn theorems. Descriptors are used immensely in the prediction of the site selectivity as well as the reactivity [29, 41–43].

The Fukui function and the Local Hyper Softness are called local reactivity descriptors, whereas chemical hardness (η), chemical potential (μ), polarizability (α) are global reactivity descriptors [105]. Parr and his co-workers [105] defined a novel descriptor as the global electrophilicity index (ω) which quantifies the global electrophilic power of the molecule.

Glossman-Mitnik and their co-workers studied a linear regression between the calculated Conceptual DFT descriptors and experimental pK_a values of various amines [106]. They tested 6 different functionals from the Minnesota family (M11, M11L, MN12L, MN12SX, N12 and N12SX) and tried to find the best fit. They constructed equations relating the pK_a values and the descriptors. From their study, they came up with the result that MN12SX functional with Def2TZVP basis set is the best methodology among other methods.

According to Yang and his coworkers, the electron density by itself does not explain everything and the change in electronic density under the influence of an approaching reagent is significant too [107]. The importance of the valence electrons in forming molecules is presented by Fukui and it is considered that the highest energy orbital electron density is the most significant for electrophilic attacks whereas the lowest energy is the most significant for nucleophilic attacks. HOMO and LUMO are the key principles for governing the chemical reactions.

Finally, Felipe and his co-workers [108] calculated pK_a of monoprotic acids and they tested the pK_a calculation's reliance on a set of reference molecules, a group of 22 monoprotic acids previously utilized by de Souza Silva and Custodio [109]. They used three training sets to show the sensitivity of average energy of the solvated proton with

the type and number of the reference. They evaluated the isodesmic reactions at three different levels of theory: two semiempirical levels (AM1 and PM6) and one composite method (G4CEP). They worked with seven functionals (LSDA, PBE0, B3LYP, M06-2X, CAM-B3LYP, WB97XD, and B2PLYP), and Hartree−Fock (HF). They tested two basis sets for HF and DFT: aug-cc-pVDZ and aug-cc-pVTZ. The solvent was described by the SMD model, including, and excluding explicit water molecules. They reported that pK_a results were acceptable and they had mean absolute error value less than 1 unit. They concluded that the isodesmic reactions are suitable for the theoretical calculations of pK_a values, especially when implying challenges for thermodynamic cycles.

5. Methodology

This section elaborates the methodology that we applied in this work.

5.1. Conformational Search

The characteristics of a series of conformations of a specific molecule are as crucial as the molecule's structure in determining a molecule's physical and biological features. The notion of reducing the strain energy of a molecule's conformation before using it in any analytical stage is now well established. However, the concept of investigating a particular flexible molecule's conformational space is less frequently applied due to the typical challenges of efficiently exploring conformational space for molecules with more than 10 or so rotatable bonds [110]. We had limited the size of the molecules and the number of rotatable bonds to minimize the computational time during conformation search and worked on small, relatively rigid molecules.

A conformational analysis study, in its widest sense, is putting up a series of conformations and then examining their energetic reasonableness in a separable procedure. Apparently, the global minimum might be obtained if all probable conformations could be constructed and studied. This is impossible since each rotatable bond has an unlimited amount (not 360) of distinct conformations. As a result, the task is simplified to testing the entirety of a given molecule's torsional space without being too practically infeasible to be practical [110].

We followed a protocol to prepare a Gaussian 16 computational input files for a set of conformers generated in Spartan Software. We started by drawing the structures in the Spartan workspace. After minimizing the structures, we performed a conformational search using the semi-empirical PM3 method [111]. We ensured that all expected conformations have been turned up. After exporting the conformations with (.xyz) files which are performed the work of forming Gaussian inputs for each conformation, we

ran the files in Gaussian and optimized the geometries with different methodologies which will be discussed in the next section.

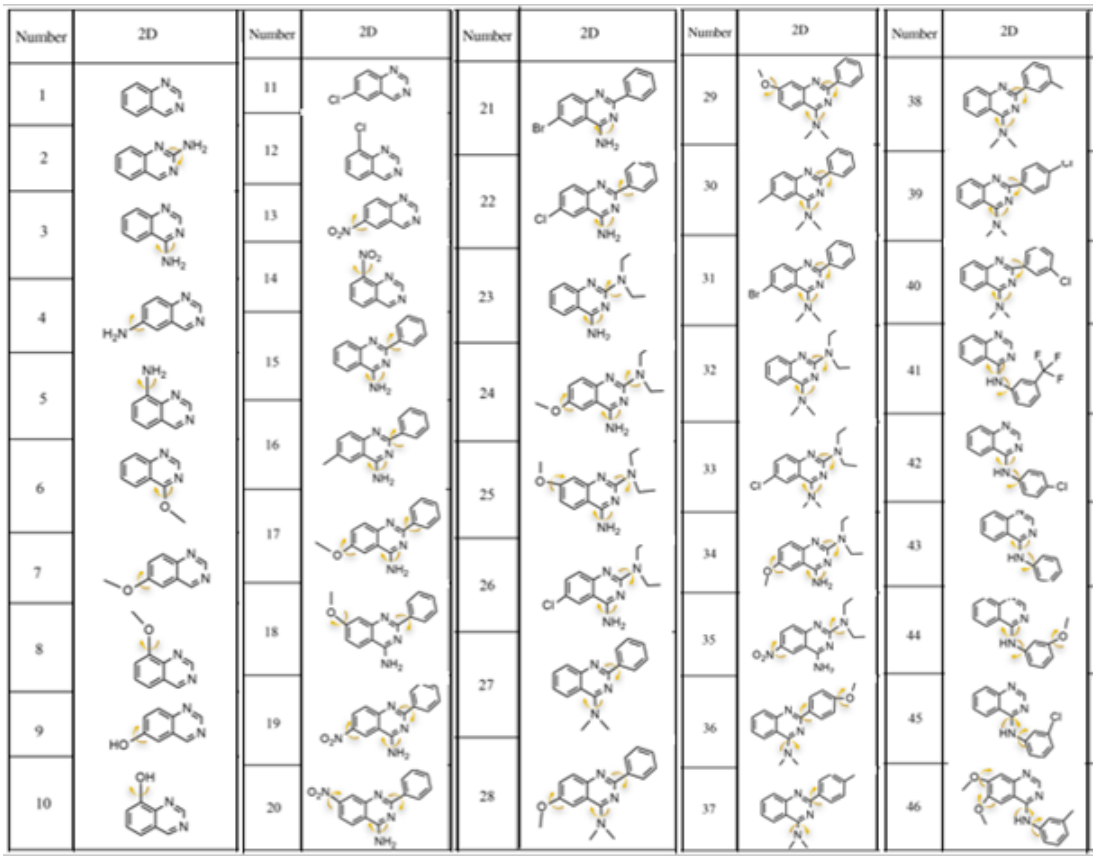


Figure 5.1. Representations of rotatable bonds in 46 molecules of quinazoline derivatives.

5.2. Choice of the Functionals and Basis Sets

In this study, all the quantum mechanical calculations are done by using Gaussian program package. A systematic conformation analysis was conducted for 46 molecules using the semi-empirical PM6 method. After that step, the structure localized as minima were further optimized by five different density functionals (B3LYP, M06L, M062X, wB97XD, and MN12SX). We considered M06L as meta-GGA XC functional, B3LYP as GGA exchange-correlation functional, M062X as hybrid meta-GGA XC functional, MN12SX as meta-GGA Hybrid Screened Exchange-Correlation Functional,

and wB97XD as Hyb-GGA Exchange Correlation Functional [131]. For our case of study, we did not examine LDA functionals, because the previous studies in literature show that models based on meta-GGA and GGA functionals are substantially better than LDA functionals for some properties [112].

Two different basis sets were used (6-31+G* and 6-311++G**). One charge model, NPA (Natural Population Analysis) was tested. We had optimized 46 molecules at five different levels of theory mentioned previously in water solvent environment by employing the SMD and CPCM solvation models.

With these methodologies, by using the resulting input files within Gaussian16®, we calculated DFT-optimized conformer geometries and the free energies. Additionally, we checked the frequencies using the vibrational frequency analysis and we made sure there were no imaginary frequencies. After fixing and eliminating the conformations for all methodologies, we worked to determine the "best" and the most stable conformers of each unique molecule as shown in the Figure 5.2.

5.3. Conceptual DFT Descriptors

For a bench of DFT functionals and basis sets tested for hundreds of conformers, a linear regression of experimental pK_a and the NPA atomic charges is evaluated. By looking at the strongest correlation from charge and experimental pK_a correlation, we analyzed conceptual DFT descriptor and compared the performance of several current conceptual DFT functionals in generating chemical reactivity descriptors.

First, for energy related descriptors, we studied chemical potential (μ), global hardness (η), LHS and the electrophilicity index (ω). The chemical potential and it equals to negative value of electronegativity which is described as

$$\mu = \frac{\partial E}{\partial N_{v(r)}} \qquad\qquad \mu \approx \frac{1}{2}(I + A). \tag{5.1}$$

Chemical potential measures the tendency of an electron escaping from the equilibrium [113]. The global hardness measures the resistance to charge transfer by using Koopman's theorem, these expressions can be written as

$$\eta = (\frac{\partial E^2}{\partial N^2})_{v(r)} \qquad \eta \approx \frac{1}{2}(I - A), \tag{5.2}$$

where I is the ionization energy and the A is the electron affinity [114–117]. To calculate I and A, we used HOMO and LUMO energies presented as

$$I \approx -HOMO, \tag{5.3}$$

$$A \approx -LUMO. \tag{5.4}$$

The electrophilicity index, shown as ω measures the stabilization energy of the systems when the electron density increases and it can be expressed as

$$\omega = \frac{\mu^2}{2\eta}. \tag{5.5}$$

Electron accepting power is described as ω^+ and the electron donating power is described as ω^-. If the value of ω^+ is large, that means the system is capable to accept charge well.

The chemical potential which is the first derivative of the energy relative to the number of electrons is constant within the molecule whereas the global hardness, the softness are functions of position [118]. The number of electrons in a molecule change as it contacts and exchanges electrons with its environment. Consequently, in solution-phase chemistry, the quantity of electrons is not a natural parameter. When describing compounds as open systems, the electronic chemical potential is more effective in regulating the mean number of electrons on a segment. The local hyper softness [119] which is a local reactivity descriptor [106], is the dual descriptor's equivalent which is defined as

$$LHS = \frac{\triangle f(r)}{\eta^2}. \tag{5.6}$$

The Fukui function is a descriptor that has two values for inserting and withdrawing electrons in isolated systems; these are accurate extensions of the well-known LUMO and HOMO orbital theories [105, 120, 121]. It helps in the comprehension and

prediction of relative reactivities of various locations within a particular molecule. The corresponding condensed functions are (see the Eqs. 5.7 and 5.8) $f_k^+ = q_k(\mathrm{N}+1) - q_k(\mathrm{N})$ (for nucleophilic attack) which measures the intramolecular reactivity at site r toward a nucleophilic reagent, $f_k^- = q_k\ (\mathrm{N}) - q_k\ (\mathrm{N}-1)$ (for electrophilic attack) which measures the intramolecular reactivity at site r toward a electrophilic reagent. The Fukui functions are the very important reactivity indexes for examining the attack of a nucleophile or electrophile [122] and they are defined as

$$f_k^- = q_k(N) - q_k(N-1) \tag{5.7}$$

$$f_k^+ = q_k(N+1) - q_k(N) \tag{5.8}$$

$$\Delta f_k = f_k^+ - f_k^-, \tag{5.9}$$

where q_k is the gross charge of atom k in the molecule.

Morell and his coworkers proposed a new descriptor and denoted it with Δf_k which is the condensed dual function [144]. In order to emphasize that this is a Fukui function of second order, it has been substituted with the current notation $f_{(k)}^2$ [123]. The condensed Fukui functions as shown in Eq 5.9 are used to determine the atoms' reactivity, and they are significant to characterize the reactivity of a site of the molecule. Because of this reason the dual descriptor is able to distinguish those sites of nucleophilic and electrophilic behavior, it has also been demonstrated to be a strong tool to predict specific sites of nucleophilic and electrophilic attacks in a far more efficient manner than the Fukui function can demonstrate on its own. The mathematical difference between nucleophilic and electrophilic Fukui functions is the basic equation for generating the condensed dual function, Δf_k. If the original system is neutral and contains N number of electrons, the approximation requires subtracting the sum of electronic densities of the system with one more electron (N+1 electrons, which means an anion) and one less electron (N-1 electrons, which means a cation).

When $\Delta \mathrm{f}_k > 0$, a nucleophilic attack controls the reaction on atom k and then that atom acts an electrophilic species; on the contrary, when $\Delta \mathrm{f}_k < 0$, the process is controlled by an electrophilic attack over atom k and for this reason atom k acts as a

nucleophilic species [123].

The condensed Fukui functions, global electrophilicity ω, chemical potential μ, total chemical hardness η, condensed dual descriptor Δf_k and local hyper softness LHS of the thirty training set molecules (See Table 2.2) are calculated with single point calculations on the previous optimized structures. Fortunately, single-point calculations of the lowest energy conformers are used to calculate the descriptors with more precision without having to do further computations involving systems with N-1 and N+1 electrons. We employed M06L and B3LYP density functionals and the 6-31+G* and 6-311++G** Pople basis sets with the SMD and CPCM solvation models because these methodologies had consistently produced acceptable results for a variety of structural and thermodynamic characteristics of quinazoline and derivatives. All the computations were done in the presence of water as a solvent. We verified which density functional is the best suited for the calculation of the conceptual DFT descriptors for this system.

5.4. Isodesmic Reactions

Finally, isodesmic reactions are part of a larger effort that is currently underway. The method does not require explicit water and it takes the advantages of proton free energy of solvation being absent.

As illustrated in Figure 5.2 in the Theoretical Background section, one of the pK_a prediction approach employed in this work is based on an isodesmic reactions between a base and its conjugate acid of a reference molecule, whose experimental pK_a is known. Eq 5.10 demonstrates the proton transfer reaction's equilibrium constant, which may be written in terms of the acid dissociation constants of the species HA and HRef.

Equation (5.11) may be used to compute the pK_a value of HA in relation to HRef. Because solvation energies are not needed in the isodesmic scheme, inaccuracies from gas phase calculations are eliminated, resulting in more accurate pK_a values. The pK_a

of the molecule AH can be calculated from

$$K_{eq} = \frac{[A][HRef^+]}{[HA^+][Ref]} = \frac{K_a[HA^+]}{K_b[HRef^+]} \tag{5.10}$$

$$\Delta G_{soln} = G_{soln}(A^{q-1}) + G_{soln}(HRef^q) - G(HA^q) - G(Ref^{q-1}) \tag{5.11}$$

$$pK_{a_{HA}} = \frac{G_{soln}}{RT \ln 10} + pK_{a_{HRef}}. \tag{5.12}$$

We carried out benchmark studies to evaluate isodesmic reactions on the 46 molecules with the use of M06L and 6-311++G** basis set with CPCM solvation model and water as solvent during geometry optimization. We used the best conformers that we had obtained from the conformational search and optimized the protonated geometries which we draw with Spartan software in the light of our conformational search analysis. We protonated N2 (which we will discuss later) which is away from the ring for all molecules. The protonated forms of our molecules can exist in multiple conformations as neutral forms but in this study a conformational analysis for the protonated forms is not studied. That may cause to not to achieve the best results in our study.

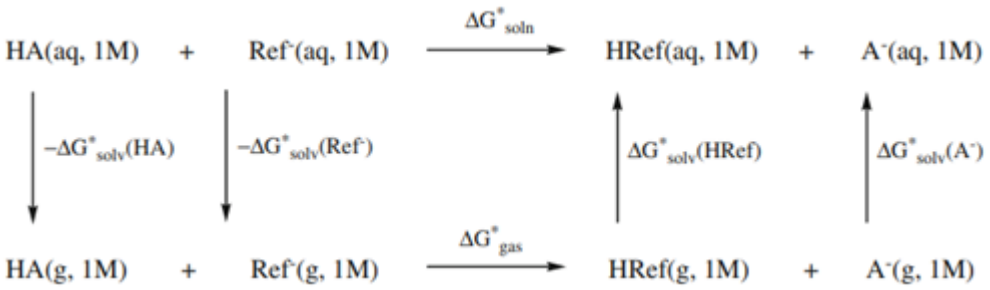


Figure 5.2. pK_a calculation via the proton exchange method [3].

Our reference molecules were 2-N,2-N-diethylquinazoline-2,4-diamine (23), and 4-(N,N-dimethylamino)-2-phenylquinazoline (27) for this approach as shown in Figure 5.3.

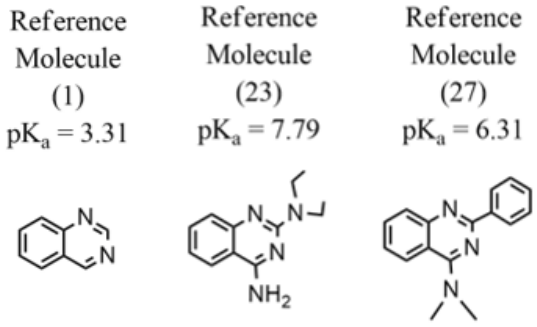


Figure 5.3. Reference species.

6. Results and Discussion

6.1. Analysis of the Atomic Charges of Quinazoline Derivatives

We demonstrated a procedure for computing low-energy conformations of 46 small to drug-like molecules (1-4 rotatable bonds) and aimed to discover conformations which are more "fit" and have a higher probability of being present shown in Figure 6.1

After we investigated the conformations of our molecules, we made a correlation between atomic charges and experimental pK_a values of the training set composed of 30 quinazoline derivatives.

We found that the choice of computational framework is significant in the accuracy of the pK_a prediction. From the overall work, we present the most reliable methodology for the pK_a prediction of quinazoline molecules. The combination of charge method and solvent models showing the best result is NPA with CPCM. M06L/6-311++G** is the DFT method which is the most assuring. In the following sections, the correlation between the experimental pK_a values and the atomic charges is computed with various theoretical frameworks discussed in the methodology section.

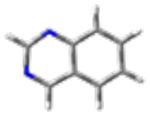
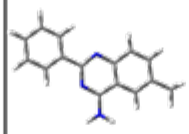
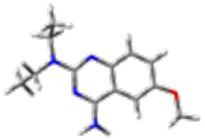
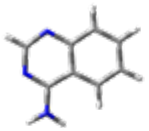
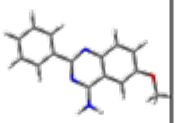
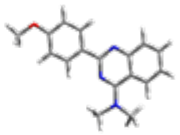
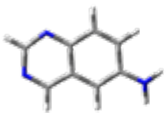
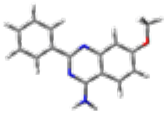
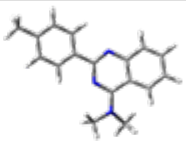
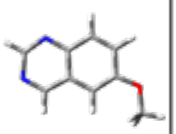
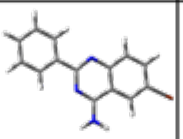
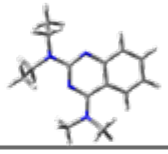
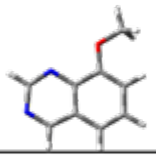
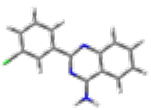
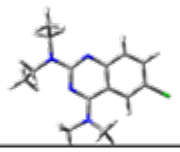
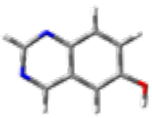
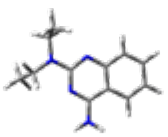
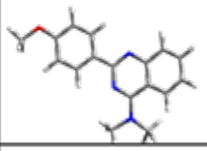
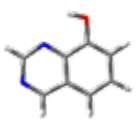
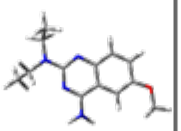
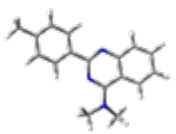
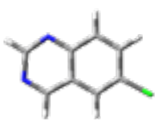
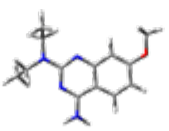
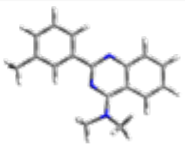
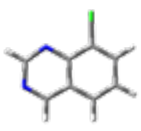
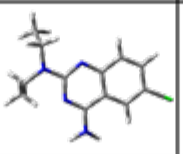
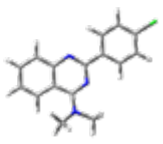
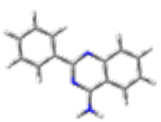
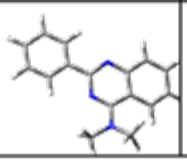
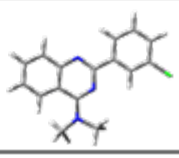
#		pK_a	#		pK_a	#		pK_a
1		3.31	16		5.16	28		6.61
3		5.73	17		5.33	29		6.2
4		3.2	18		5.62	32		8.88
7		2.85	21		4.78	33		7.91
8		3.51	22		4.98	34		8.98
9		3.12	23		7.79	36		6.62
10		3.41	24		7.82	37		6.4
11		3.55	25		8.31	38		6.28
12		3.3	26		7.82	39		5.88
15		5.44	27		6.31	40		5.54

Figure 6.1. 3D representations of the best conformers of training set molecules.

It is reasonable to expect distinction in the acidity constants because electronic alterations at atomic positions are present. In the quinazoline molecule there are two nitrogen atoms and pK_a varies according to the substituent. NPA atomic charges can be calculated for all the atoms present in the molecule, but nitrogen atoms are more negative than other atoms. The average atomic charges of N are approximately -0.5 for the derivatives. The highest electron density region are the sites where electrophiles may attack and therefore N atoms are generally the most active centers in this set [124]. As seen in the electrostatic potential map of the molecule 15 presented in Figure 6.2, quinazoline derivatives have two main basic high electron density sites. Red color corresponds to electron-rich regions and the blue color correspond to electron-poor regions. Map color legend is shown at the side of the Figure 6.2.

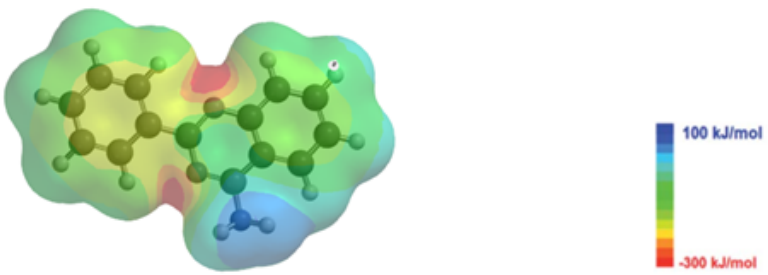


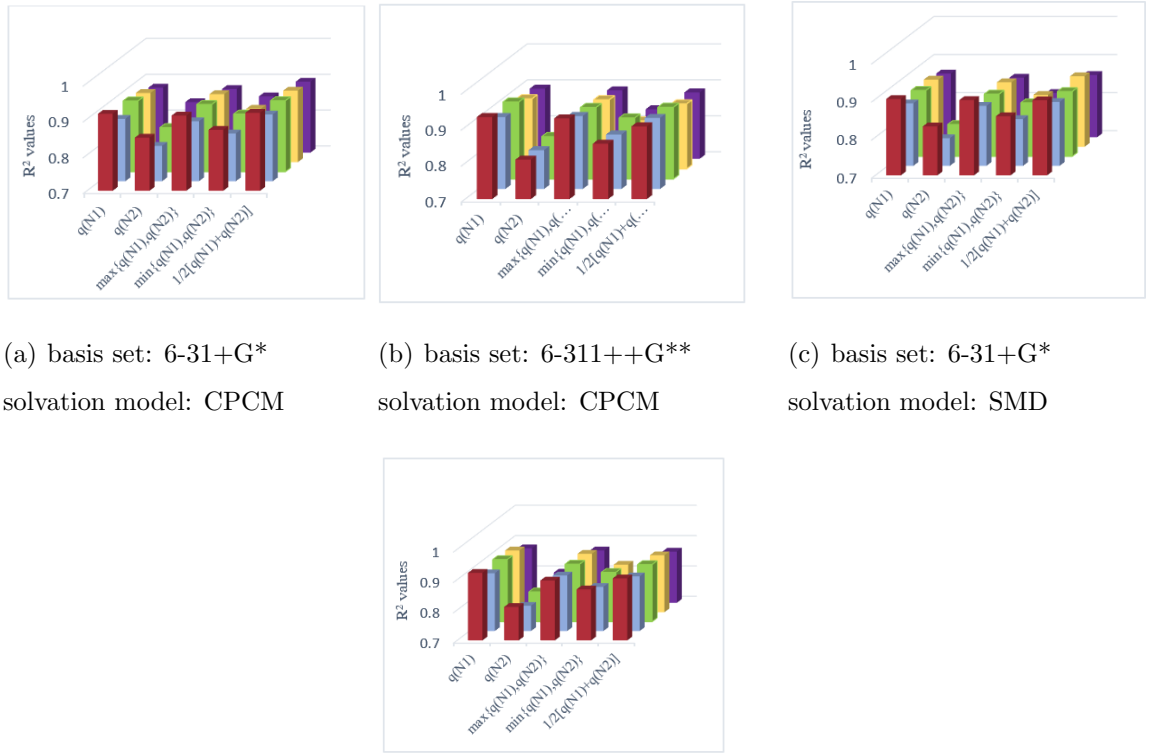
Figure 6.2. Spartan calculated electrostatic potential map of the quinazoline derivative (molecule number 15 is presented as an example).

In our training set, we have 30 quinazoline derivatives and two nitrogen atoms present in the fused ring are common and they are the most electronegative ones. To find a correlation between the atomic charges of nitrogen atoms and the experimental pK_a values, we examined both nitrogen atoms in five different ways. We investigated N_1 (the one close to the ring), N_2 (the one far from the ring), $q_{(N_{min})}$, $q_{(N_{max})}$, and $q_{(N_{average})}$ described as

$$q(N_{min}) = min\{q(N_1), q(N_2)\}, \tag{6.1}$$

$$q(N_{max}) = max\{q(N_1), q(N_2)\}, \tag{6.2}$$

$$q(N_{average}) = average\{(q(N_1) + q(N_2))/2\}. \tag{6.3}$$



(a) basis set: 6-31+G*
solvation model: CPCM

(b) basis set: 6-311++G**
solvation model: CPCM

(c) basis set: 6-31+G*
solvation model: SMD

(d) basis set: 6-311++G**
solvation model: SMD

Figure 6.3. Linear Regression between calculated atomic charges and the experimental pK_a Calculations are made with 5 different methods (M06L, M062X, B3LYP, wb97XD, MN12SX) respectively.

Figure 6.3 represents an illustration that help us to compare the results graphically. In the graphs a, b, c, and d, basis sets and solvation models are changed to benchmark the effect of nitrogen choice. In graph a, calculations are made with 5 different methods (M06L, M062X, B3LYP, wb97XD, MN12SX with 6-31+G* basis set and CPCM solvation model. $q_{(N_1)}$, q(Nmax), and $q_{(N_{average})}$ values provided better results than $q_{(N_2)}$ and $q_{(N_{min})}$. The effect of different substituents on the quinazoline derivatives must be considered at that point. We searched for a relationship between the atomic charges on nitrogen atoms considering that alteration. This relationship is more consistent with $q_{(N_1)}$, $q_{(N_{max})}$, and $q_{(N_{average})}$ value. This conclusion is supported by the graphs b, c, and the d presented in Figure 6.3. In all methodologies, the highest R^2 are offered with the same three. Among this three, we proceeded with the $q_{(N_1)}$,

because highest score is offered. From the overall atomic charge vs experimental pK_a study, we concluded that we will be analyzing particularly $q_{(N_1)}$.

We also investigated how the correlation depends on the basis set and DFT functionals used. First of all, all the DFT methods that we employed produced good correlation between atomic charges and experimental pKa's with acceptable R^2 values ($0.93 \geq R^2 \geq 0.86$). As reported in Table 4, it is possible to distinguish are most reliable methods by looking at the R^2 values.

The results are better compared to the SMD model using CPCM with the methods. The R^2 values generally gave higher results when the methods combined with CPCM model. For example, R^2 value of SMD model with M06L/6-311++G** method is 0.9255, whereas R^2 value of CPCM model is 0.9285 with the same combination. This trend is valid for all functionals except for the wB97XD method. R2 value of SMD model with wB97XD/6-311++G** method is 0.9035 whereas R^2 value of CPCM model with wB97XD/6-311++G** method is 0.8948. In that combination, SMD solvation model provided a better result but still the scores are very close to each other.

We can also investigate the effect of basis set by looking at R^2 values. For all combinations, higher basis set (6-311++G**) provided higher scores than the small basis set (6-31+G*) which was expected. We decided to move on only higher basis sets in the next sections to have faster findings without analyzing the small basis set. Among the DFT functionals, the most accurate methods from the work are M06L and B3LYP methods. All M06L combinations produced R^2 value higher than 0.9 and B3LYP also produced R^2 value higher than 0.87 and two of the combinations produced R^2 value higher than 0.9. MN12SX, M06L, and wB97XD gave also acceptable correlations, but they are not the most suitable among twenty combination we had tested on our training sets. The best findings include M06L/6-31+G*/CPCM, M06L/6-311++G**/SMD, M06L/6-311++G**/CPCM, and B3LYP/6-311++G**/CPCM. ($R^2 > 0.91$) for $q_{(N_1)}$.

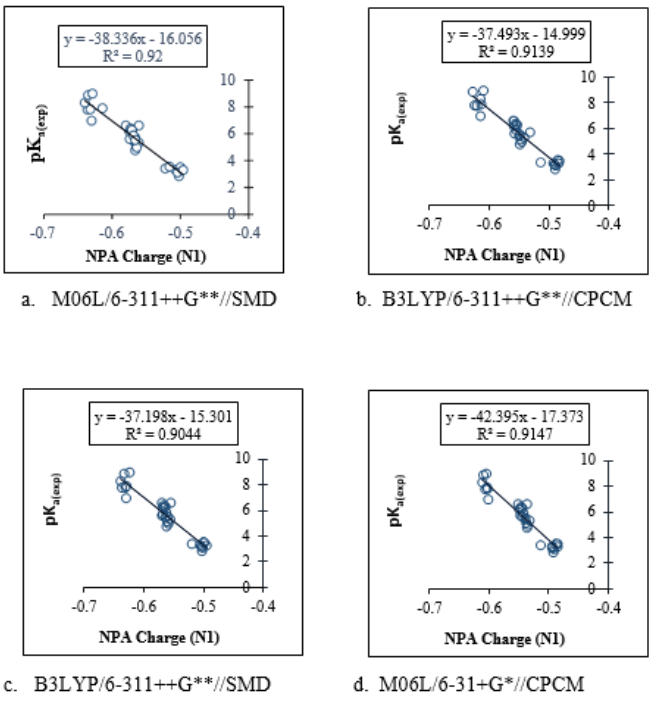


Figure 6.4. Linear Regression between calculated atomic charges and the experimental pK_a.

Figure 6.4 shows the best relationships between the experimental pK_a and the predicted NPA charges for the training set. There are groupings in the graphs. It seems the molecules are grouped into three, but it does not affect the R^2 value. That grouping is because of the charge values and also underlines the broadness of the pK_a range that we worked on. Multiple linear equations are obtained by a least-square fit

$$pK_a = m \cdot q + n \ \textbf{with} \ q = q(N_1), \tag{6.4}$$

where m and n are the fitted parameters and as q we use N1 atom. The R^2 was determined to be 0.9285 for the training set. We can report that no notable outlier components were detected. The maximum deviation between estimated and measured pK_a will be discussed through all compounds (see Tables 6.2 and 6.3). These findings reveal a tight relationship between experimental pK_a and $q_{(N_1)}$.

The method excelled in accuracy across all solvents, basis sets, charge models, atom choices, and solvation models tested, the M06L/6-311++G**/CPCM with using

N1 atom and water as solvent gave the most reasonably accurate R^2 value (R^2 = 0.9285).

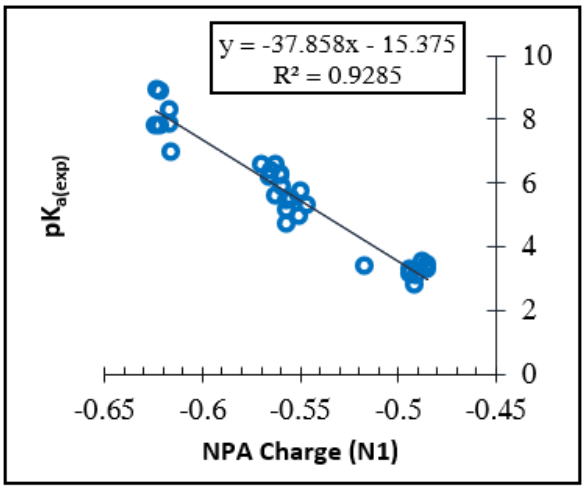


Figure 6.5. Linear Regression between calculated atomic charges and the experimental pK_a. Calculations were done with M06L/6-311++G**/ CPCM in water and NPA atomic charge model is used.

$$pK_a = -37.858x - 15.375 \quad \text{(from M06L/6-311++G**//CPCM)} \tag{6.5}$$

$$pK_a = -38.336x - 16.056 \quad \text{(from M06L/6-311++G**//SMD)} \tag{6.6}$$

$$pK_a = -37.493x - 14.999 \quad \text{(from B3LYP/6-311++G**//CPCM)} \tag{6.7}$$

$$pK_a = -37.493x - 14.999 \quad \text{(from B3LYP/6-311++G**//SMD)} \tag{6.8}$$

will be the four main equations that we will employ through the evaluation of our test set. If x is known, the y can be calculated. "x" is charge of N1 atom and the "y" is predicted pK_a.

Table 6.1. R^2 values of the correlation between charges vs pK_a (exp).

Method	**Basis Set**	**Solvation model**	R^2 N_1	R^2 N_2	R^2 N_{max}	R^2 N_{min}	R^2 N_{av}
M06L	6-31+G*	SMD	0.90	0.83	0.90	0.86	0.90
M06L	6-311++G**	SMD	0.92	0.81	0.90	0.87	0.90

Table 6.1. R^2 values of the correlation between charges vs pK_a(exp). (cont.)

M06L	6-31+G*	CPCM	0.92	0.85	0.91	0.87	0.92
M06L	6-311++G**	CPCM	0.93	0.82	0.92	0.86	0.90
M062X	6-31+G*	SMD	0.86	0.77	0.86	0.82	0.87
M062X	6-311++G**	SMD	0.89	0.78	0.88	0.84	0.88
M062X	6-31+G*	CPCM	0.87	0.80	0.87	0.83	0.89
M062X	6-311++G**	CPCM	0.90	0.81	0.90	0.85	0.90
B3LYP	6-31+G*	SMD	0.87	0.78	0.86	0.84	0.87
B3LYP	6-311++G**	SMD	0.90	0.80	0.89	0.86	0.89
B3LYP	6-31+G*	CPCM	0.90	0.82	0.90	0.86	0.90
B3LYP	6-311++G**	CPCM	0.91	0.82	0.89	0.86	0.90
wB97XD	6-31+G*	SMD	0.91	0.78	0.87	0.84	0.89
wB97XD	6-311++G**	SMD	0.90	0.77	0.89	0.84	0.88
wB97XD	6-31+G*	CPCM	0.89	0.79	0.89	0.83	0.89
wB97XD	6-311++G**	CPCM	0.90	0.78	0.89	0.83	0.88
MN12SX	6-31+G*	SMD	0.87	0.78	0.86	0.82	0.86
MN12SX	6-311++G**	SMD	0.89	0.81	0.88	0.82	0.84
MN12SX	6-31+G*	CPCM	0.86	0.84	0.86	0.85	0.89
MN12SX	6-311++G**	CPCM	0.89	0.78	0.87	0.89	0.82

For a bench of DFT functionals, solvation models, and basis sets we obtained regressions to apply on our test set composed of sixteen quinazoline derivative. The methodology we proposed in this paper aims at predicting accurate $q_{(N_1)}$, not only for our training set but as a final goal, for pK_a our test set. According to all the regressions we had obtained, we calculated the pK_a values of the test set by using the Eqs 6.5, 6.6, 6.7, 6.8. In Table 6.2, it is shown the calculated pK_a's, deviations from the experimental pK_a's (ΔpK_a) for M06L method. We applied that protocol for all method which will be discussed later. ΔpK_a value alter between 0.15 – 1.81 pK_a unit for M06L/6-311++G**//CPCM and 0.20 – 1.79 pK_a unit for M06L/6-311++G**//SMD as shown in Table 6.2. ΔpK_a value changes between 0.10 – 2.07 pK_a unit for B3LYP/6-

311++G**//CPCM and 0.05 – 1.83 pK_a unit for B3LYP /6-311++G**//SMD as shown in Table 6.3.

Table 6.2. Test Set: Differences between Experimental and Predicted pK_a Values.

		M06L/6-311++G**//CPCM		M06L/6-311++G**//SMD	
	pK_a (exp)	pK_a (calc)	ΔpK_a	pK_a (calc)	ΔpK_a
2	4.43	6.235	1.805	6.222	1.792
5	2.4	3.791	1.391	3.626	1.226
6	3.13	4.451	1.321	4.529	1.399
13	4.18	3.016	-1.164	3.062	-1.118
14	4.00	2.744	-1.256	2.874	-1.126
19	4.54	5.187	0.647	5.244	0.703
20	4.27	5.653	1.383	5.544	1.274
30	6.52	5.769	-0.751	5.620	-0.900
31	5.88	5.730	-0.150	5.657	-0.223
35	6.63	7.437	0.807	7.087	0.457
41	5.03	4.877	-0.153	4.980	-0.05
42	6.02	5.032	-0.988	5.093	-0.927
43	6.08	5.110	-0.970	5.168	-0.912
44	5.37	5.110	-0.260	5.168	-0.202
45	5.28	4.916	-0.364	5.178	-0.262
46	5.77	5.265	-0.505	5.319	-0.451

Table 6.3. Test Set: Differences between Experimental and Predicted pK_a Values.

		B3LYP/6-311++G**//CPCM		B3LYP/6-311++G**//SMD	
	pK_a (exp)	pK_a (calc)	ΔpK_a	pK_a (calc)	ΔpK_a
2	4.43	6.497	2.067	6.260	1.830
5	2.4	3.707	1.307	3.860	1.460
6	3.13	4.563	1.433	4.610	1.480
13	4.18	3.261	-0.919	3.148	-1.032
14	4.00	3.075	-0.925	2.998	-1.002
19	4.54	5.641	1.101	5.660	1.120
20	4.27	5.381	1.111	5.435	1.165
30	6.52	5.567	-0.953	5.660	-0.860
31	5.88	5.604	-0.276	5.660	-0.220
35	6.63	8.394	1.764	7.534	0.904
41	5.03	5.121	0.091	4.985	-0.045
42	6.02	5.195	-0.825	5.097	-0.923
43	6.08	5.307	-0.773	5.210	-0.870
44	5.37	5.269	-0.101	5.172	-0.198
45	5.28	5.121	-0.159	5.022	-0.258
46	5.77	5.381	-0.389	5.285	-0.485

As a conclusion and summary for Tables 6.2 and 6.3, the majority of the compounds in the test sets had good estimations, however few molecules in the test set have a deviation higher than 1.5 pK_a unit. For instance, for almost all calculations, 2-aminoquinazoline was detected to be above the $\pm$ 1.5 range limit. We have grounds to believe that the experimental pK_a for 2-aminoquinazoline seems to be incorrect. In the article that we found the acidity constant of 2-aminoquinazoline, the ionization of heterocyclic bases is searched [125]. They reported that the addition of a second nitrogen atom to a six-membered ring which already has one diminishes the basic strength

considerably (generally as much of 4 pK_a units). They also observed the amino derivatives present extra ionic resonance with one exception which is 2-aminoquinazoline. The increase was over 2 pK_a units than expected in their study. That is why we have doubts about 2-aminoquinazoline which is an outlier and can even be eliminated from the test set.

Another point that must be discussed is that some ΔpK_a values are negative values and some are positive. For instance, for all subsets, molecules number 13, 14, 30, 31, 42, 43, 44, 45, 46 have $\Delta pK_a < 0$. Molecules 42, 43, 44, 45, and 46 have very similar structures, the common feauture is that they all have similar substituents on 4^{th} position as shown in Figure 6.6. They are predicted more acidic than it should be.

$\Delta pK_a < 0$ is true also for 13 and 14 which have nitro group as the substituent. Nitro is an electron withdrawing group and it alters the charge distribution and weakens the basicity. It is hard to properly match our pK_a predictions with the previously published in the literature.

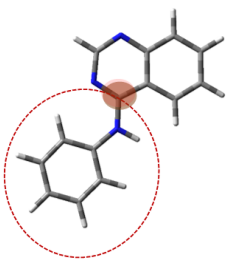


Figure 6.6. 3D representation molecule number 43.

As confirmed by Settimo and his co-workers, predicting the pK_a values of bases are much more challenging than acids because pK_a values of basic compounds are spread over a broader range [126]. This methodology produced poor results while prediction of derivatives with strong electron withdrawing groups. Our test was various enough to understand which derivatives are not responding properly to the methodologies. To examine the efficacy of our estimations, we evaluated MD (mean deviation) which is defined as

$$\mathrm{MD} = \frac{\sum(x_i - y_i)}{N}, \tag{6.9}$$

MAD (mean absolute deviation) which is defined as

$$\text{MAD} = \frac{\sum |x_i - y_i|}{N}, \tag{6.10}$$

and RMSE (root mean square error) which is defined as

$$\text{RMSE} = \sqrt{\frac{\sum (x_i - y_i)^2}{N}}. \tag{6.11}$$

x_i represents the calculated value, y_i represents the experimental value, and N indicates the total number of molecules which is 16 in this section.

The deviations helped us to pick the most appropriate functional and the basis set for our data set in this study. We analyzed MD, MAD and RMSE values. We observed that when diffuse functions and polarization functions are added to the basis set the strength of predictivity for any of the DFT functionals is marginally increased. 6-31+G* has higher MAD than 6-311++G** basis sets as presented in Table 6.5. Based on the functional performances, the lowest MAD values are achieved using either M06L or B3LYP functionals in all subsets.

Table 6.5. MD, MAD and RMSE values for all methods and basis sets.

	M06L				B3LYP			
	6-31+G*		6-311++G**		6-31+G*		6-311++G**	
	SMD	CPCM	SMD	CPCM	SMD	CPCM	SMD	CPCM
MD	0.17	0.24	0.04	0.01	0.04	0.26	0.22	0.13
MAD	0.71	0.75	0.81	0.87	0.73	0.81	0.89	0.86
RMSE	0.67	0.96	0.17	0.04	0.15	1.05	0.89	0.52

	M062X				MN12SX			
	6-31+G*		6-311++G**		6-31+G*		6-311++G**	
	SMD	CPCM	SMD	CPCM	SMD	CPCM	SMD	CPCM
MD	0.32	0.16	0.19	0.196	0.34	0.37	0.08	0.39
MAD	0.84	0.81	0.86	0.89	0.83	0.84	0.99	0.89

Table 6.5 MD, MAD and RMSE values for all methods (cont.)

	M06L				B3LYP			
RMSE	1.26	0.66	0.76	0.75	1.35	1.50	0.30	1.57

	wB97XD			
	6-31+G*		6-311++G**	
	SMD	CPCM	SMD	CPCM
MD	0.36	0.44	0.24	0.2
MAD	0.83	0.86	0.87	0.89
RMSE	1.42	1.76	0.94	0.78

The least accurate predictions were obtained when the calculations were with wB97XD functional for all subsets except MN12SX with 6-311++G**/SMD basis set. We can report that MN12SX and wB97XD are not the best choices to obtain accurate and fast pK_a predictions. Overall, the M06L functional combined with higher basis set did a better job to estimate pK_a values more accurately with relatively low RMSE values with respect to other methods and low computational cost for the data set included in this study.

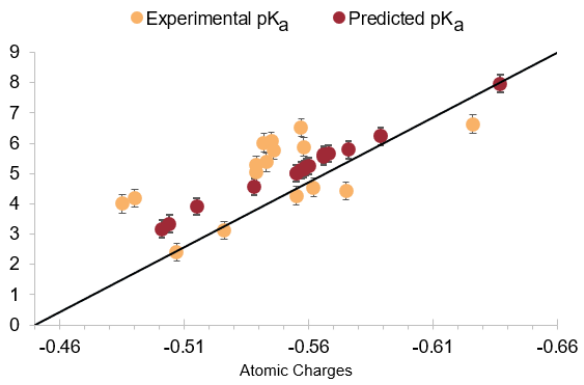


Figure 6.7. Correlation plot of the calculated (M06L / 6-311++G** // CPCM) and experimental pK_a values of 16 molecules.

In Figure 6.7, the correlation plot of the calculated (M06L/6-311++G**//CPCM) and experimental pK_a values of 16 molecules is presented. We gave a fit with R^2 equal to 1 to observe the deviations in a clear way. We tested and achieved applicable findings according to the protocol that we had suggested. The next step will involve the study of the conceptual DFT descriptors with the best findings from atomic charge vs experimental pK_a correlation studies.

6.2. Analysis of the Conceptual DFT Descriptors

Two density functionals were selected to evaluate the pK_a values by using conceptual DFT descriptors considered here. The selected functionals are M06L and B3LYP and they offer satisfactory results consistently in atomic charge vs experimental pK_a study. We computed our studiesused the 6-311++G** basis set. Also, since the diffuse functions usually achieve better results for anions, higher basis set was a way better option [127]. We had selected the most stable conformers for this section. The goal was to ensure that the findings were accurate for evaluating the pK_a value with the conceptual DFT Descriptors. We used the experimental pK_a's of the quinazoline and derivatives to conduct a regression analysis of the global and local conceptual DFT descriptors. The correlation coefficients pK_a for each of the fits of the different descriptors calculated with the two density functionals are presented in the Table 6.6.

Table 6.6. R^2 values of the DFT descriptors with the experimental pK_a. Calculations are made with 6-311++G** basis set.

	$R^2(f_k^+)$	$R^2(f_k^-)$	$R^2(\mu)$	$R^2((\omega)$	$R^2(\eta)$	$R^2(\Delta f_k)$	R^2(LHS)
M06L//CPCM	0.817	0.813	0.723	0.192	0.441	0.839	0.850
M06L//SMD	0.792	0.761	0.710	0.076	0.773	0.794	0.805
B3LYP//CPCM	0.836	0.794	0.718	0.259	0.722	0.837	0.837
B3LYP//SMD	0.789	0.770	0.641	0.168	0.711	0.803	0.803

The best candidate was determined for the case in which the pK_a was correlated to Local Hyper Softness. However, the condensed dual function, (Δf_k, and nucleophilic and electrophilic Fukui functions can also be suggested to predict pK_a values of the test set. We centered the analysis on the descriptors which provide more accurate results. M06L density functional and the 6-311++G** basis set using water as solvent simulated with the CPCM parametrization provided the highest score ($R^2 = 0.846$) for LHS and 0.839 for (Δf_k). As a result, we developed a set of equations

$$\textbf{Electrophilic Fukui:} \quad \mathrm{pK}_{acalc} = (-34.812)(f_k^-) + 1.3422 \tag{6.12}$$

$$\textbf{Nucleophilic Fukui:} \quad \mathrm{pK}_{acalc} = (-60,784)(f_k^+) + 11.501 \tag{6.13}$$

$$\textbf{Condensed Dual:} \quad \mathrm{pK}_{acalc} = (-22.85)(\Delta f_k) + 5.0278 \tag{6.14}$$

$$\textbf{Local Hyper Softness:} \quad \mathrm{pK}_{acalc} = (0.2604)(LHS) + 4.964 \tag{6.15}$$

correlating the pK_a of quinazoline and derivatives to fukui functions and LHS in the form of Equation (6.4), where m and n are obtained from the previously identified linear regression study presented in Figure 6.8.

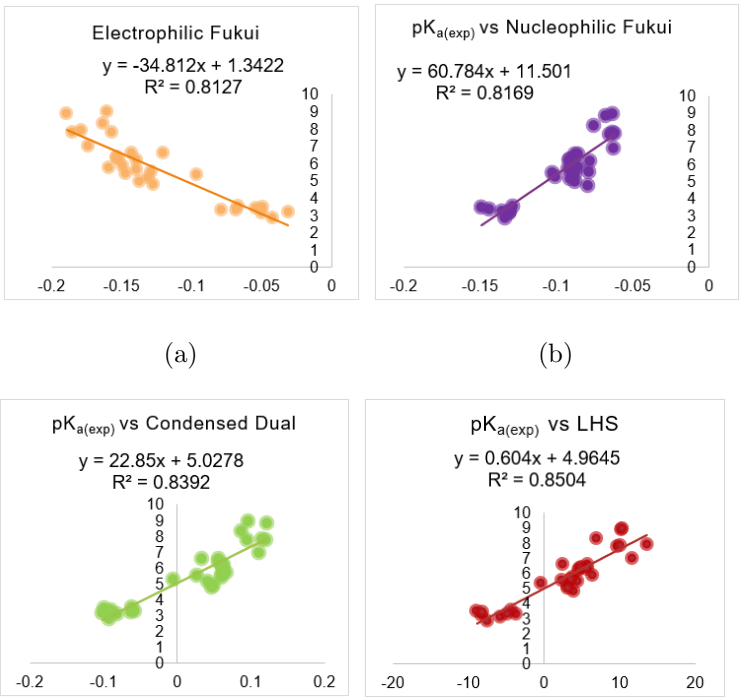


Figure 6.8. Experimental pK_a versus Conceptual DFT Descriptors. Calculations are carried out with M06L/6-311++G** methodology using water as solvent with the CPCM solvent model.

Table 6.7. ΔpK_a values computed from the descriptors. Calculations are made with M06L/6-311++G**//CPCM.

Number of the molecule	pK_a	f_k^+	f_k^-	Δf_k	**LHS**
2	4.43	0.142	2.099	1.398	1.267
5	2.4	0.105	0.161	0.046	0.091
6	3.13	1.746	2.494	2.218	2.057
13	4.18	4.343	5.622	5.281	5.580
14	4	4.583	3.713	4.113	4.040
19	4.54	5.259	1.989	3.253	4.610
20	4.27	5.651	2.747	3.888	5.558
30	6.52	-2.435	-0.443	-1.172	-1.282

Table 6.7. ΔpK_a values computed from the descriptors (cont.)

31	5.88	0.454	0.788	0.702	0.457
35	6.63	3.351	1.152	2.054	5.653
41	5.03	1.183	0.037	0.455	0.279
42	6.02	-0.293	-1.336	-0.969	-1.037
43	6.08	-0.961	-8.045	-5.622	-5.709
44	5.37	-0.251	-7.335	-4.912	-5.184
45	5.28	-0.283	-0.422	-0.389	-0.435
46	5.77	-0.834	-2.165	-1.725	-1.721

As we can see from the Table 6.7, the application has some outliers which has significantly higher Δf_k values. For example, for 6-nitro quinazoline evaluated with LHS equation, pK_a prediction is 5.58 pK_a units over the experimental value, which is significantly high. Similar overestimation is true for molecules numbered 13, 14, 19, 20, and 35. One argument we can suggest is these compunds differ from the other compounds in the substitution of nitro group. The most reasonable explanation for the deviation is having nitro groups attached. Even though, the correlations are acceptable with R^2 values, when our test set is evaluated, we can say that the experimental pK_a values do not fit well with any of the DFT descriptors. This might be because the molecule's structure considering the substituents, has a significant impact on the proton giving or electron receiving potential. These systems are notoriously tricky to converge and produce reliable results but still [**?**], it's also worth noting that our suggested Fukui function, distinguishes between distinct active sites in a molecule better than existing approaches that include information regarding set of molecular orbitals.

For the test set, the smallest MDs were found for nucleophilic fukui function (MD = 0.07). Nucleophilic fukui function did not produce acceptable results. Considering other methods may improve the results to have a better score in evaluation of pK_a values.

Table 6.8. MD, MAD and RMSE values for all M06L/6-311++G**//CPCM.

	f_k^+	f_k^-	Δf_k	**LHS**
MD	0.07	1.36	0.54	0.89
MAD	2.53	1.99	2.39	2.81
RMSE	0.26	5.44	2.15	3.56

6.3. Analysis of the Isodesmic Reactions

In this part of the study, we applied isodesmic reaction scheme to the 46 drug-like quinazoline derivatives to evaluate their pK_a values by using the neutral and protonated states of quinazoline (1), 2-N,2-N-diethylquinazoline-2,4-diamine (23), and 4-(N,N-dimethylamino)-2-phenylquinazoline (27) as reference species. The predicted pK_a values are calculated at M06L/6-311++G**//CPCM level of theory.

38 molecules out of 46 were classified into 3 main groups with 3 reference molecules according to their structural similarities and based on their backbone structures as illustrated in Figure 6.9. We referred that while using the isodesmic reaction method, picking of a good reference molecule is crucial in pK_a predictions. We tried tp pick a handy reference species that have nearly similar groups, solute–solvent interactions, electronic structure as the species under examination. For the calculation of the pK_a's, reference species were determined for each group which are shown on the right-hand side of Figure 6.9.

The reason we neglected 5 molecules which are 4-methoxyquinazolineamine, 6-quinazolineamine, 8-quinazolineamine, 6-nitroquinazolineamine, 8- nitroquinazolineamine is they are exemplary of reference molecules. They are small-sized and backbones of the remaining derivatives. To evaluate pK_a values, we used neutral states and N2 protonated states. Figure 6.10 illustrates the isodesmic reaction we applied.

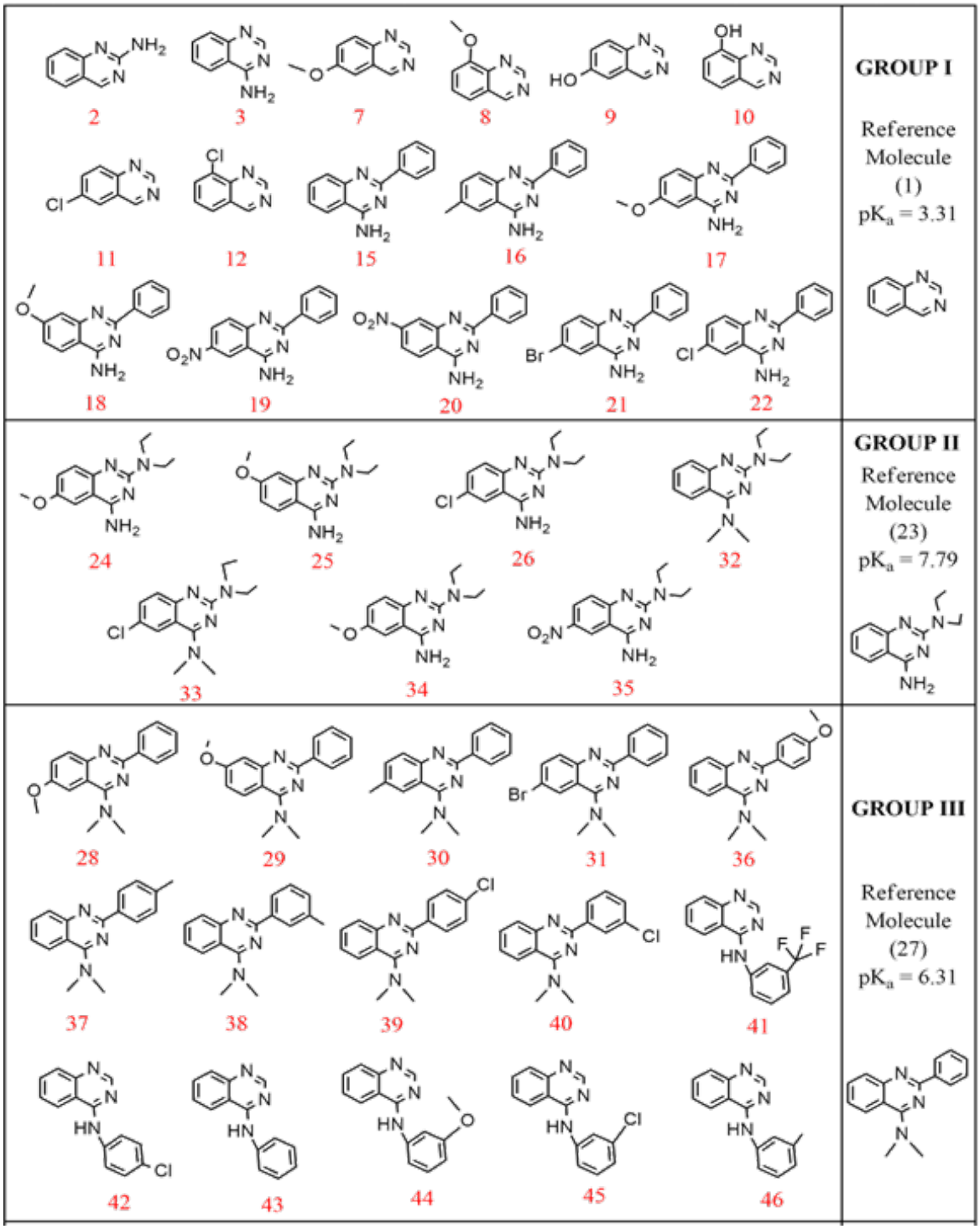


Figure 6.9. 2D representations of 38 molecules (left) in three main groups and the corresponding reference molecules with the experimental pK_a values (right) for pK_a prediction.

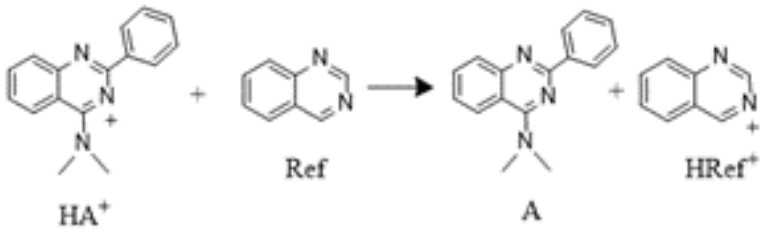


Figure 6.10. Isodesmic reaction between a protonated quinazoline derivative and a reference species, quinazoline.

Quinazoline has two basic sites, but distinguishing which nitrogen atom takes the proton in aqueous solution isn't vital because addition of water (whether N1 or N2) to the anhydrous cation will result in the same stabilized structure. This is valid for all hydrated cations, although the value of the ionization constant for anhydrous cations depends on whether N1or N2 is protonated. As a result, evaluating the ionization constants of anhydrous cations should be approached with caution [128].

There is no clear scientific proof to demonstrate which of the nitrogens in quinazoline is more basic, but we may assume that N2 is the basic center by adopting the reasoning that isoquinoline is more basic than quinoline [151]. As presented in Figure 6.13, if protonation happens on N1, water addition at position 3,4 will produce the resonance-stabilized hydrated cation (3), and if protonation occurs on N2, water addition at position 1,4 will give the same cation.

In cases there are more than two basic centers in quinazoline such as when there are nitrogen containing substituents, protonation site becomes significant. Substituents behave in many ways in heterocyclic compounds. Armarego reported that protonation happens on the ring nitrogen atoms and the high basic strengths of the 2-, 4-, and 7-amino isomers is attributed to resonance in the cations [**?**].

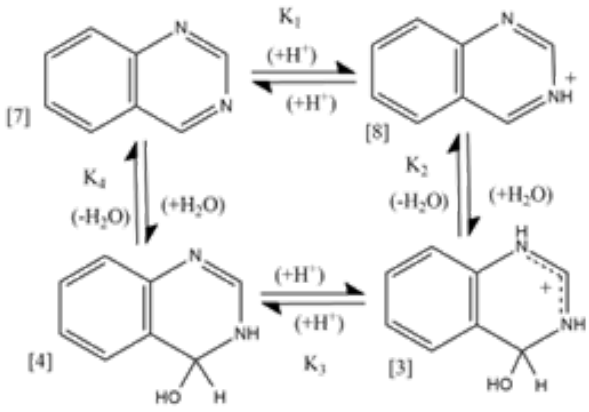


Figure 6.11. An overall and composite ionization scheme of quinazoline molecule involving the equilibria K_1, K_2, K_3, and K_4 [**?**].

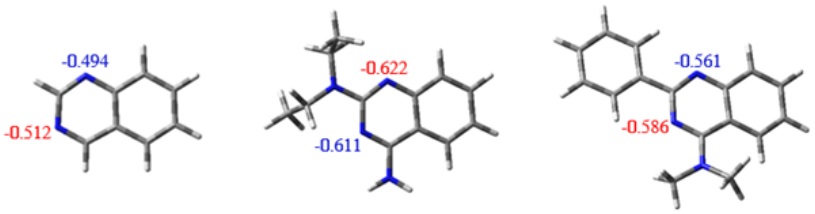


Figure 6.12. 3D representation of three reference molecules with N1 and N2 charges (NPA) on them.

In this study, we protonated only N2 and applied isodesmic reactions accordingly. As presented in Figure 6.12, $q(N_1)$ and $q(N_2)$ are similar so protonation site can not be determined exactly. Further analysis can be done to have a deeper understanding and evaluation of predicting pK_a values. Predicted pK_a values' deviations, mean absolute deviation, mean deviation and root mean square deviation are given below in Table 6.9. According to the literature, computed pK_a values are acceptable if the mean absolute error is less than one pK_a unit [129]. Predicted pK_a's of 27 over 38 molecules are within the range of ± 1 compared to the experimental values: 8 of them are within the ≤ 0.5 pK_a confidence interval, 19 are within the $0.5 \leq pK_a \leq 1.0$ confidence and finally, 11 predictions show a deviation of > 1.0 pK_a units. That finding shows that although the isodesmic reaction approach for pK_a evaluation has a strong predictive power for these molecules, selected references are not adequate for molecules numbered 7, 11, 12, 19,

20, 22, 25, 35, 41, 42, and 46. In Group I, molecules resemble with their substituents on quinazoline backbone. Molecules 19 and 20 differ because they have nitro group as substituent. Because there are two protonation sites, overestimation is the point in question with a deviation of almost > 2.0 pK_a units. In Group III, we achieved a 0 ΔpK_a with molecule number 28, when molecule 27 is picked as a reference which is a highly strong prediction.

Table 6.9. ΔpK_a values for 3 groups and MD, MAD, RMSD values.

Ref 1	pK_a (exp)	ΔpK_a	Ref 23	pK_a (exp)	ΔpK_a	Ref 27	pK_a (exp)	ΔpK_a
Number			Number			Number		
2	3.88	-0.55	24	8.70	0.88	28	6.61	0.00
3	5.74	0.01	25	11.79	3.48	29	7.02	0.82
7	4.23	1.38	26	6.35	-0.63	30	6.68	0.16
8	3.58	0.07	32	8.36	-0.52	31	4.86	-1.02
9	4.12	1,00	33	7.17	-0.74	36	6.55	-0.07
10	2.74	-0.67	34	9.82	0.84	37	7.20	0.80
11	2.27	-1.28	35	4.83	-1.80	38	6.80	0.52
12	2.11	-1.19	MD	0.10		39	5.02	-0.86
13	5.11	-0.33	MAD	0.55		40	4.74	-0.80
16	5.70	0.54	RMSD	1.17		41	3.92	-1.11
17	5.68	0.35				42	4.43	-1.59
18	6.42	0.80				43	5.61	-0.47
19	1.77	-2.77				44	5.93	0.56
20	2.29	-1.98				45	4.73	-0.55
21	3.98	-0.80				46	7.87	2.10
22	3.80	-1.18				MD	-0.09	
MD	-0.41					MAD	0.71	
MAD	0.93					RMSD	0.90	
RMSD	1.36							

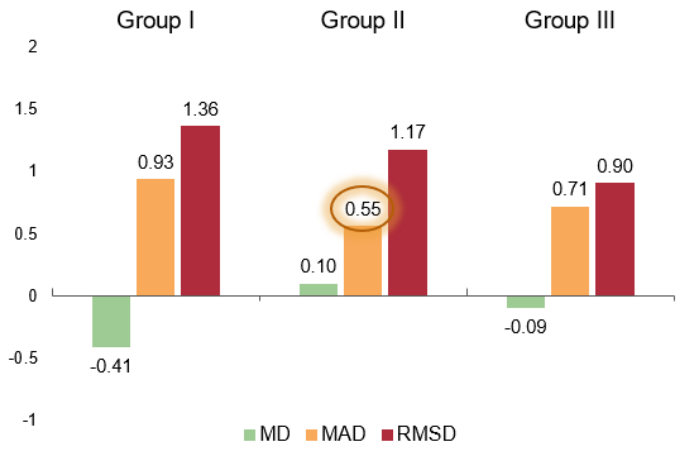


Figure 6.13. MD, MAD, and RMSE values of 3 groups.

4-aminoquinazolines and derivatives have three potential protonation sites and it is crucial to have an idea of which one is the most favorable to the proton attack. Zielinski and his coworkers has suggested two distinct methods to describe the susceptible protonation site [130]. The first method in the correlation of pK_a with Hammett constants, σ That approach is an indirect strategy whereas the second suggestion represents a direct approach. They constituted X-ray analysis on 4-(N,N-dimethylamino)-2-(p-methoxyphenyl)quinazoline hydrochloride which is molecule number 37. They reported that the protonation site is susceptible on N1. There was no hydrogen bonding interaction on N2 or N3.

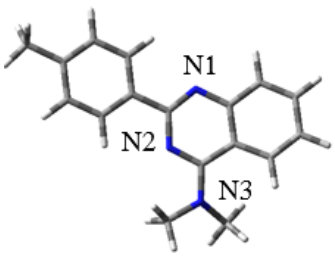


Figure 6.14. 3D represenation of 4-(N,N-dimethylamino)- 2-(p-methoxyphenyl) quinazoline with N numbers on them.

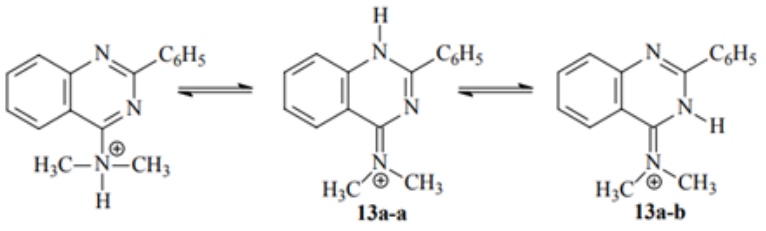


Figure 6.15. Fast hydrogen exchange between the two endocyclic nitrogen atoms and that a possitive charge is concentrated on the exocyclic nitrogen atom of the dimethylamino group [153].

There is a strong resonance stabilization of protonated Group III molecules and N1 produces extremely stable paraquinoid structure this limits the rotation of dimethylamino groups from the plane. Electron transfer through carbon bonds are limited and that affects the parameters. It means that it is hard to say there is one particularly specified protonation site in such arrangements.

7. Conclusion

In the present study, we have suggested a protocol to achieve an accurate and fast pK_a prediction nmethod for quinazoline and derivatives. We suggested a protocol which is based on the linear regression of the experimental pK_a value with the atomic charges, conceptual DFT descriptors and isodesmic reactions. We used our training set to produce an equation by employing various DFT methods (combination of five DFT-functionals and 2 basis sets), two solvent models with water as the solvent. In the atomic charge study, all these methods produced acceptable linearities having R^2 higher than 0.75 (and more than 0.85 for most of the cases). The best combination of DFT functionals and basis sets are found to be M06L/6-311++G**//CPCM, M06L/6-311++G**//SMD, B3LYP/6-311++G**//CPCM, B3LYP/6-311++G**// SMD. Using the best DFT combinations, we predicted pK_a values of our test set molecules and calculated the difference between experimental and predicted pK_a. [U+F044]pKa value varied between 0.150 – 1.805 pK_a unit for M06L/6-311++G**//CPCM, 0.202 – 1.792 pK_a unit for M06L/6-311++G**//SMD, 0.101 – 2.067 pK_a unit for B3LYP/6-311++G**//CPCM and 0.045 – 1.830 pK_a unit for B3LYP /6-311++G**//SMD. MM-MD-RMSD calculations showed that the estimated pK_a's deviate within ± 1 unit with respect to different conformational changes. The average predicted pK_a yielded precise predictions.

The next step was to design a protocol that accurately and efficiently can predict the pK_a of quinazoline and derivatives by combining these findings and using the conceptual DFT descriptors. The goal of the study's analysis was to ensure that the findings were accurate for evaluating the pK_a value considering the conceptual DFT Descriptors. We picked the most stable conformers for this section and aimed to find a correlation with the descriptors and the experimental pK_a. M06L/6-311++G** methodology using water as solvent with the CPCM solvation model was the best combination among overall scores. Single point calculations with the best methodologies were performed on the optimized geometries for anion and cation forms. We employed M06L/6-

311++G**//CPCM, M06L/6-311++G**//SMD, B3LYP/ 6-311++G**// CPCM and B3LYP/ 6-311++G**// SMD for conceptual DFT descriptors. The best descriptor was determined to LHS. However, the condensed dual function, Δf_K, and nucleophilic and electrophilic Fukui functions can also be suggested to predict pK_a values of the test set in a limited fashion. We centered the analysis on the descriptors which provide more accurate results. From the whole of the results presented in this work, the best predictions for the pK_a in comparison with the experimental values were for the M06L density functional, for which a MD 0.07 for nucleophilic fukui descriptors, 1.36 for electrophilic fukui descriptors, 0.54 for condensed dual descriptor, and 0.89 for local hyper softness was obtained for the quinazoline and derivatives.

In the last step of this study, we calculated the acid dissociation constants of quinazoline derivatives by applying the isodesmic reaction scheme. In this context of pK_a calculations, the combination that works best for our molecule set we employed M06L/6-311++G**. We classified our molecules in 3 main groups based on their structural similarities and geometrical backbones. Each group has a reference molecule which is picked by considering whether they include similar functional groups and identical charge distributions. According to the literature, computed pK_a values are acceptable if the mean absolute error is less than one pK_a unit [129]. Predicted pK_a's of 27 over 38 molecules are within the range of ± 1 compared to the experimental values: 8 of them are within the ≤ 0.5 pK_a confidence interval, 19 are within the $0.5 \leq pK_a \leq 1.0$ confidence and finally, 11 predictions show a deviation of >1.0 pK_a units. That points out that the isodesmic method can be applied successfully for the pK_a estimation of quinazoline and derivatives. Outliers are investigated to understand whether the reference molecules are susceptible. Multiple molecules gave high deviations, and the reasons must be investigated in a deeper way. In further studies, new computational methods can be tested. A deeper conformational analysis can be performed for all the parts. Training set and the test can be developed and be more diversified in terms of both geometrical structure and the pK_a range. For isodesmic reaction part, other nitrogen atoms can be protonated and search for an average value. Also, conformational analysis can be done for the protonated forms of the derivatives.

REFERENCES

1. Sastre, S., R. Casasnovas, F. Muñoz and J. Frau, "Isodesmic reaction for pK_a calculations of common organic molecules", *Theoretical Chemistry Accounts*, Vol. 132, No. 2, p. 1310, 2013.

2. Ravez, S., O. Castillo-Aguilera, P. Depreux and L. Goossens, "Quinazoline derivatives as anticancer drugs: A patent review (2011-present)", , 2015.

3. Ho, J. and M. L. Coote, "pK_a Calculation of Some Biologically Important Carbon Acids - An Assessment of Contemporary Theoretical Procedures", *Journal of Chemical Theory and Computation*, Vol. 5, No. 2, pp. 295–306, 2009.

4. Dangi, R. R., N. S. Chundawat and K. L. Ameta, "Synthesis and biological evaluation of some quinazoline heterocyclic derivatives", *Green Chemistry: Synthesis of Bioactive Heterocycles*, pp. 393–412, Springer India, 2014.

5. Weissberger, A., *A Series Of Monographs*, Tech. rep., New York, 1967.

6. Chandregowda, V., A. K. Kush and G. Chandrasekara Reddy, "Synthesis and in vitro antitumor activities of novel 4-anilinoquinazoline derivatives", *European Journal of Medicinal Chemistry*, Vol. 44, No. 7, pp. 3046–3055, 2009.

7. Al-Rashood, S. T., I. A. Aboldahab, M. N. Nagi, L. A. Abouzeid, A. A. Abdel-Aziz, S. G. Abdel-hamide, K. M. Youssef, A. M. Al-Obaid and H. I. El-Subbagh, "Synthesis, dihydrofolate reductase inhibition, antitumor testing, and molecular modeling study of some new 4(3H)-quinazolinone analogs", *Bioorganic and Medicinal Chemistry*, Vol. 14, No. 24, pp. 8608–8621, 2006.

8. Vasdev, N., P. N. Dorff, A. R. Gibbs, E. Nandanan, L. M. Reid, J. P. O'Neil and H. F. VanBrocklin, "Synthesis of 6-acrylamido-4-(2-[18F]fluoroanilino)quinazoline: A prospective irreversible EGFR binding probe",

Journal of Labelled Compounds and Radiopharmaceuticals, Vol. 48, No. 2, pp. 109–115, 2005.

9. Baba, A., N. Kawamura, H. Makino, Y. Ohta, S. Taketomi and T. Sohda, "Studies on Disease-Modifying Antirheumatic Drugs: Synthesis of Novel Quinoline and Quinazoline Derivatives and Their Anti-inflammatory Effect", *Journal of Medicinal Chemistry*, Vol. 39, No. 26, pp. 5176–5182, 1996.

10. Alagarsamy, V., V. Raja Solomon and K. Dhanabal, "Synthesis and pharmacological evaluation of some 3-phenyl-2-substituted-3H-quinazolin-4-one as analgesic, anti-inflammatory agents", *Bioorganic & Medicinal Chemistry*, Vol. 15, No. 1, pp. 235–241, 2007.

11. Rohini, R., P. Muralidhar Reddy, K. Shanker, A. Hu and V. Ravinder, "Antimicrobial study of newly synthesized 6-substituted indolo[1,2-c]quinazolines", *European Journal of Medicinal Chemistry*, Vol. 45, No. 3, pp. 1200–1205, 2010.

12. Gupta, K., S. Giri and P. K. Chattaraj, "Acidity of meta- and para-substituted aromatic acids: a conceptual DFT study", *New Journal of Chemistry*, Vol. 32, No. 11, p. 1945, 2008.

13. Antipenko, L., A. Karpenko, S. Kovalenko, A. Katsev, E. Komarovska-Porokhnyavets, V. Novikov and A. Chekotilo, "Synthesis of New 2-Thio-[1,2,4]triazolo[1,5-c]quinazoline Derivatives and Its Antimicrobial Activity", *Chemical and Pharmaceutical Bulletin*, Vol. 57, No. 6, pp. 580–585, 2009.

14. Aly, A. A., "Synthesis of Novel Quinazoline Derivatives as Antimicrobial Agents", *Chinese Journal of Chemistry*, Vol. 21, No. 3, pp. 339–346, 2010.

15. Nandy, P., M. T. Vishalakshi and A. R. Bhat, "Synthesis and antitubercular activity of Mannich bases of 2-methyl-3H-quinazolin-4-ones", *Indian Journal of Heterocyclic Chemistry*, Vol. 15, pp. 293–294, 2006.

16. Hess, H. J., T. H. Cronin and A. Scriabine, "Antihypertensive 2-amino-4(3H)-quinazolinones", *Journal of Medicinal Chemistry*, Vol. 11, No. 1, pp. 130–136, 1968.

17. Malamas, M. S. and J. Millen, "Quinazolineacetic acids and related analogs as aldose reductase inhibitors", *Journal of Medicinal Chemistry*, Vol. 34, No. 4, pp. 1492–1503, 1991.

18. Alvarado, M., M. Barceló, L. Carro, C. Masaguer and E. Raviña, "Synthesis and Biological Evaluation of New Quinazoline and Cinnoline Derivatives as Potential Atypical Antipsychotics", *Chemistry & Biodiversity*, Vol. 3, No. 1, pp. 106–117, 2006.

19. Asif, M., "Chemical Characteristics, Synthetic Methods, and Biological Potential of Quinazoline and Quinazolinone Derivatives", *International Journal of Medicinal Chemistry*, Vol. 2014, pp. 1–27, 2014.

20. Albert, A. and E. P. Serjeant, *The Determination of Ionization Constants*, Springer Netherlands, 1984.

21. Davies, J. E., "The pharmacological basis of therapeutics", *Occupational and Environmental Medicine*, Vol. 64, No. 8, pp. e2–e2, 2007.

22. Zhang, S., J. Baker and P. Pulay, "A Reliable and Efficient First Principles-Based Method for Predicting pK_a Values. 1. Methodology", *The Journal of Physical Chemistry A*, Vol. 114, No. 1, pp. 425–431, 2010.

23. Po, H. N. and N. M. Senozan, "The Henderson-Hasselbalch Equation: Its History and Limitations", *Journal of Chemical Education*, Vol. 78, No. 11, p. 1499, 2001.

24. Sastre, S., R. Casasnovas, F. Muñoz and J. Frau, "Isodesmic reaction for accurate theoretical pK_a calculations of amino acids and peptides", *Physical Chemistry Chemical Physics*, Vol. 18, No. 16, pp. 11202–11212, 2016.

25. Olsson, M. H. M., C. R. Sôndergaard, M. Rostkowski and J. H. Jensen, "PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions", *Journal of Chemical Theory and Computation*, Vol. 7, No. 2, pp. 525–537, 2011.

26. Goudarzi, N. and M. Goodarzi, "Prediction of the acidic dissociation constant (pK_a) of some organic compounds using linear and nonlinear QSPR methods", *Molecular Physics*, Vol. 107, No. 14, pp. 1495–1503, 2009.

27. Alizadeh, K., A. R. Ghiasvand, M. Borzoei, S. Zohrevand, B. Rezaei, P. Hashemi, M. Shamsipur, B. Maddah, A. Morsali, K. Akhbari and I. Yavari, "Experimental and computational study on the aqueous acidity constants of some new aminobenzoic acid compounds", *Journal of Molecular Liquids*, Vol. 149, No. 3, pp. 60–65, 2009.

28. Milan Meloun, J. Havel and Erik Hôgfeldt, *Computation of Solution Equilibria*, Vol. 111, 1988.

29. Pliego, J. R. and J. M. Riveros, "Theoretical Calculation of pK_a Using the Cluster−Continuum Model", *The Journal of Physical Chemistry A*, Vol. 106, No. 32, pp. 7434–7439, 2002.

30. Thapa, B. and H. B. Schlegel, "Density Functional Theory Calculation of pK_a's of Thiols in Aqueous Solution Using Explicit Water Molecules and the Polarizable Continuum Model", *The Journal of Physical Chemistry A*, Vol. 120, No. 28, pp. 5726–5735, 2016.

31. Casasnovas, R., J. Ortega-Castro, J. Frau, J. Donoso and F. Muñoz, "Theoretical pK_a calculations with continuum model solvents, alternative protocols to thermodynamic cycles", *International Journal of Quantum Chemistry*, Vol. 114, No. 20, pp. 1350–1363, 2014.

32. Ho, J., "Are thermodynamic cycles necessary for continuum solvent calculation of pK_as and reduction potentials?", *Physical Chemistry Chemical Physics*, Vol. 17, No. 4, pp. 2859–2868, 2015.

33. Ho, J. and M. L. Coote, "A universal approach for continuum solvent pK_a calculations: Are we there yet?", *Theoretical Chemistry Accounts*, Vol. 125, No. 1-2, pp. 3–21, 2010.

34. Ho, J. and M. L. Coote, "First-principles prediction of acidities in the gas and solution phase", *Wiley Interdisciplinary Reviews: Computational Molecular Science*, Vol. 1, No. 5, pp. 649–660, 2011.

35. Rebollar-Zepeda, A. M. and A. Galano, "First principles calculations of pK_a values of amines in aqueous solution: Application to neurotransmitters", *International Journal of Quantum Chemistry*, Vol. 112, No. 21, pp. 3449–3460, 2012.

36. Rebollar-Zepeda, A. M., T. Campos-Hernández, M. T. Ramírez-Silva, A. Rojas-Hernández and A. Galano, "Searching for Computational Strategies to Accurately Predict p ¡i¿K¡/i¿ ¡sub¿a¡/sub¿ s of Large Phenolic Derivatives", *Journal of Chemical Theory and Computation*, Vol. 7, No. 8, pp. 2528–2538, 2011.

37. Galano, A., J. R. Alvarez-Idaboy and M. Francisco-Márquez, "Physicochemical Insights on the Free Radical Scavenging Activity of Sesamol: Importance of the Acid/Base Equilibrium", *The Journal of Physical Chemistry B*, Vol. 115, No. 44, pp. 13101–13109, 2011.

38. Bryantsev, V. S., M. S. Diallo and W. A. Goddard III, "Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models", *The Journal of Physical Chemistry B*, Vol. 112, No. 32, pp. 9709–9719, 2008.

39. Pliego, J. R., "Thermodynamic cycles and the calculation of pK_a", *Chemical Physics Letters*, Vol. 367, No. 1-2, pp. 145–149, 2003.

40. Casasnovas, R., J. Frau, J. Ortega-Castro, A. Salva, J. Donoso and F. Muñoz, "Absolute and relative pK_a calculations of mono and diprotic pyridines by quantum methods", *Journal of Molecular Structure: THEOCHEM*, Vol. 912, No. 1-3, pp. 5–12, 2009.

41. Caballero, N., F. Melendez, C. Muñoz-Caro and A. Niño, "Theoretical prediction of relative and absolute pK_a values of aminopyridines", *Biophysical Chemistry*, Vol. 124, No. 2, pp. 155–160, 2006.

42. Namazian, M. and S. Halvani, "Calculations of pK_a values of carboxylic acids in aqueous solution using density functional theory", *The Journal of Chemical Thermodynamics*, Vol. 38, No. 12, pp. 1495–1502, 2006.

43. Liptak, M. D., K. C. Gross, P. G. Seybold, S. Feldgus and G. C. Shields, "Absolute Determinations for Substituted Phenols", *Journal of the American Chemical Society*, Vol. 124, No. 22, pp. 6421–6427, 2002.

44. Camaioni, D. M. and C. A. Schwerdtfeger, "Comment on "Accurate Experimental Values for the Free Energies of Hydration of H^+ , OH^-, and H_3O^+ """, *The Journal of Physical Chemistry A*, Vol. 109, No. 47, pp. 10795–10797, 2005.

45. Donald A. McQuarrie, *Statistical Mechanics*, Harper & Row, 1970.

46. Mujika, J. I., J. M. Mercero and X. Lopez, "A Theoretical Evaluation of the pK_a for Twisted Amides Using Density Functional Theory and Dielectric Continuum Methods", *The Journal of Physical Chemistry A*, Vol. 107, No. 31, pp. 6099–6107, 8 2003.

47. Ugur, I., A. Marion, S. Parant, J. H. Jensen and G. Monard, "Rationalization of the pK_a Values of Alcohols and Thiols Using Atomic Charge Descriptors and Its Application to the Prediction of Amino Acid pK_a", *Journal of Chemical Information and Modeling*, Vol. 54, No. 8, pp. 2200–2213, 2014.

48. Albert, A. and J. N. Phillips, "264. Ionization constants of heterocyclic substances. Part II. Hydroxy-derivatives of nitrogenous six-membered ring-compounds", *Journal of the Chemical Society (Resumed)*, p. 1294, 1956.

49. Kudelko, A. and W. Zielinski, "The synthesis of 8-hydroxyquinazoline derivatives and their acid-base interactions", *Journal of Heterocyclic Chemistry*, Vol. 41, No. 2, pp. 247–251, 2004.

50. Işık, M., A. S. Rustenburg, A. Rizzi, M. R. Gunner, D. L. Mobley and J. D. Chodera, "Overview of the SAMPL6 pK_a challenge: evaluating small molecule microscopic and macroscopic pK_a predictions", *Journal of Computer-Aided Molecular Design*, Vol. 35, No. 2, pp. 131–166, 2021.

51. Kohn, W., A. D. Becke and R. G. Parr, "Density Functional Theory of Electronic Structure", *The Journal of Physical Chemistry*, Vol. 100, No. 31, pp. 12974–12980, 1996.

52. Orio, M., D. A. Pantazis and F. Neese, "Density functional theory", *Photosynthesis Research*, Vol. 102, No. 2-3, pp. 443–453, 2009.

53. Parr, R. G., "Density Functional Theory of Atoms and Molecules", *Horizons of Quantum Chemistry*, pp. 5–15, Springer Netherlands, Dordrecht, 1980.

54. Dreizler, R. M. and E. K. U. Gross, *Density Functional Theory*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1990.

55. Thomas, L. H., "The calculation of atomic fields", *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 23, No. 5, pp. 542–548, 1927.

56. Lieb, E. H. and B. Simon, "The Thomas-Fermi theory of atoms, molecules and solids", *Advances in Mathematics*, Vol. 23, No. 1, pp. 22–116, 1977.

57. Hohenberg, P. and W. Kohn, "Inhomogeneous Electron Gas", *Physical Review*,

Vol. 136, No. 3B, pp. B864–B871, 1964.

58. Kohn, W. and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Physical Review*, Vol. 140, No. 4A, pp. A1133–A1138, 1965.

59. Perdew, J. P., "Jacob's ladder of density functional approximations for the exchange-correlation energy", *AIP Conference Proceedings*, pp. 1–20, AIP, 2001.

60. Perdew, J. P., K. Burke and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", *Physical Review Letters*, Vol. 77, No. 18, pp. 3865–3868, 1996.

61. Becke, A. D., "Density-functional exchange-energy approximation with correct asymptotic behavior", *Physical Review A*, Vol. 38, No. 6, pp. 3098–3100, 1988.

62. Lee, C., W. Yang and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Physical Review B*, Vol. 37, No. 2, pp. 785–789, 1988.

63. Neese, F., T. Schwabe and S. Grimme, "Analytic derivatives for perturbatively corrected "double hybrid" density functionals: Theory, implementation, and applications", *The Journal of Chemical Physics*, Vol. 126, No. 12, p. 124115, 2007.

64. Grimme, S., "Semiempirical hybrid density functional with perturbative second-order correlation", *The Journal of Chemical Physics*, Vol. 124, No. 3, p. 034108, 2006.

65. Davidson, E. R. and D. Feller, "Basis set selection for molecular calculations", *Chemical Reviews*, Vol. 86, No. 4, pp. 681–696, 1986.

66. Wiberg, K. B., "Basis set effects on calculated geometries: 6-311++G** vs. aug-cc-pVDZ", *Journal of Computational Chemistry*, Vol. 25, No. 11, pp. 1342–1346, 2004.

67. Cramer, C. J. and D. G. Truhlar, "Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics", *Chemical Reviews*, Vol. 99, No. 8, pp. 2161–2200, 1999.

68. Mathew, K., R. Sundararaman, K. Letchworth-Weaver, T. A. Arias and R. G. Hennig, "Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways", *The Journal of Chemical Physics*, Vol. 140, No. 8, p. 084106, 2014.

69. Letchworth-Weaver, K. and T. A. Arias, "Joint density functional theory of the electrode-electrolyte interface: Application to fixed electrode potentials, interfacial capacitances, and potentials of zero charge", *Physical Review B*, Vol. 86, No. 7, p. 075140, 2012.

70. Tomasi, J., B. Mennucci and R. Cammi, "Quantum Mechanical Continuum Solvation Models", *Chemical Reviews*, Vol. 105, No. 8, pp. 2999–3094, 2005.

71. Klamt, A. and G. Schûurmann, "COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient", *J. Chem. Soc., Perkin Trans. 2*, , No. 5, pp. 799–805, 1993.

72. Pomogaeva, A. and D. M. Chipman, "Hydration Energy from a Composite Method for Implicit Representation of Solvent", *Journal of Chemical Theory and Computation*, Vol. 10, No. 1, pp. 211–219, 2014.

73. Chipman, D. M., "New formulation and implementation for volume polarization in dielectric continuum theory", *The Journal of Chemical Physics*, Vol. 124, No. 22, p. 224111, 2006.

74. Autrey, T., A. K. Brown, D. M. Camaioni, M. Dupuis, N. S. Foster and A. Getty, "Thermochemistry of Aqueous Hydroxyl Radical from Advances in Photoacoustic Calorimetry and ab Initio Continuum Solvation Theory", *Journal of the American*

Chemical Society, Vol. 126, No. 12, pp. 3680–3681, 2004.

75. Wåhlin, P., V. Vallet, U. Wahlgren and I. Grenthe, "Water Exchange Mechanism in the First Excited State of Hydrated Uranyl(VI)", *Inorganic Chemistry*, Vol. 48, No. 23, pp. 11310–11313, 2009.

76. Asthagiri, D., L. R. Pratt, M. E. Paulaitis and S. B. Rempe, "Hydration Structure and Free Energy of Biomolecularly Specific Aqueous Dications, Including Zn^+ and First Transition Row Metals", *Journal of the American Chemical Society*, Vol. 126, No. 4, pp. 1285–1289, 2004.

77. Ashbaugh, H. S. and D. Asthagiri, "Single ion hydration free energies: A consistent comparison between experiment and classical molecular simulation", *The Journal of Chemical Physics*, Vol. 129, No. 20, p. 204501, 2008.

78. Marenich, A. V., C. J. Cramer and D. G. Truhlar, "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*, Vol. 113, No. 18, pp. 6378–6396, 2009.

79. Fattebert, J.-L. and F. Gygi, "First-principles molecular dynamics simulations in a continuum solvent", *International Journal of Quantum Chemistry*, Vol. 93, No. 2, pp. 139–147, 2003.

80. Petrosyan, S. A., A. A. Rigos and T. A. Arias, "Joint Density-Functional Theory: Ab Initio Study of Cr_2O_3 Surface Chemistry in Solution", *The Journal of Physical Chemistry B*, Vol. 109, No. 32, pp. 15436–15444, 2005.

81. Andreussi, O., I. Dabo and N. Marzari, "Revised self-consistent continuum solvation in electronic-structure calculations", *The Journal of Chemical Physics*, Vol. 136, No. 6, p. 064102, 2012.

82. Miertuš, S., E. Scrocco and J. Tomasi, "Electrostatic interaction of a solute with a

continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects", *Chemical Physics*, Vol. 55, No. 1, pp. 117–129, 1981.

83. Miertu*tildes*, S. and J. Tomasi, "Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes", *Chemical Physics*, Vol. 65, No. 2, pp. 239–245, 1982.

84. Cancês, E., B. Mennucci and J. Tomasi, "A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics", *The Journal of Chemical Physics*, Vol. 107, pp. 3032–3041, 1997.

85. Mennucci, B. and J. Tomasi, "Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries", *The Journal of Chemical Physics*, Vol. 106, No. 12, pp. 5151–5158, 1997.

86. Pecul, M., D. Marchesan, K. Ruud and S. Coriani, "Polarizable continuum model study of solvent effects on electronic circular dichroism parameters", *The Journal of Chemical Physics*, Vol. 122, No. 2, p. 024106, 2005.

87. Tomasi, J., B. Mennucci and E. Cancès, "The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level", *Journal of Molecular Structure: THEOCHEM*, Vol. 464, No. 1-3, pp. 211–226, 1999.

88. Cossi, M., V. Barone, R. Cammi and J. Tomasi, "Ab initio study of solvated molecules: A new implementation of the polarizable continuum model", *Chemical Physics Letters*, Vol. 255, No. 4-6, pp. 327–335, 1996.

89. Zimmermann, T., J. Leszczynski and J. V. Burda, "Activation of the cisplatin and transplatin complexes in solution with constant pH and concentration of chloride anions; quantum chemical study", *Journal of Molecular Modeling*, Vol. 17, No. 9,

pp. 2385–2393, 2011.

90. Berstis, L., G. T. Beckham and M. F. Crowley, "Electronic coupling through natural amino acids", *The Journal of Chemical Physics*, Vol. 143, No. 22, p. 225102, 2015.

91. Cossi, M., N. Rega, G. Scalmani and V. Barone, "Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model", *Journal of Computational Chemistry*, Vol. 24, No. 6, pp. 669–681, 2003.

92. Barone, V. and M. Cossi, "Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model", *The Journal of Physical Chemistry A*, Vol. 102, No. 11, pp. 1995–2001, 1998.

93. Baldridge, K. and A. Klamt, "First principles implementation of solvent effects without outlying charge error", *The Journal of Chemical Physics*, Vol. 106, No. 16, pp. 6622–6633, 1997.

94. Truong, T. N. and E. V. Stefanovich, "A new method for incorporating solvent effect into the classical, ab initio molecular orbital and density functional theory frameworks for arbitrary shape cavity", *Chemical Physics Letters*, Vol. 240, No. 4, pp. 253–260, 1995.

95. Marenich, A. V., C. J. Cramer and D. G. Truhlar, "Perspective on Foundations of Solvation Modeling: The Electrostatic Contribution to the Free Energy of Solvation", *Journal of Chemical Theory and Computation*, Vol. 4, No. 6, pp. 877–887, 2008.

96. McDouall, J. J. W., "Chapter 3. Theoretical organic chemistry", *Annual Reports Section "B" (Organic Chemistry)*, Vol. 88, p. 39, 1991.

97. Tsui, V. and D. A. Case, "Molecular Dynamics Simulations of Nucleic Acids with a Generalized Born Solvation Model", *Journal of the American Chemical Society*,

Vol. 122, No. 11, pp. 2489–2498, 2000.

98. Hawkins, G. D., C. J. Cramer and D. G. Truhlar, "Pairwise solute descreening of solute charges from a dielectric medium", *Chemical Physics Letters*, Vol. 246, No. 1-2, pp. 122–129, 1995.

99. Constanciel, R., O. Chalvet and J. C. Rayez, "Comparative study of the pK_a of acridine, thionine and phenazine molecules in their first excited singlet and triplet states", *Theoretica Chimica Acta*, Vol. 37, No. 4, pp. 305–318, 1975.

100. Zhu, T., J. Li, D. A. Liotard, C. J. Cramer and D. G. Truhlar, "Analytical energy gradients of a self-consistent reaction-field solvation model based on CM2 atomic charges", *The Journal of Chemical Physics*, Vol. 110, No. 12, pp. 5503–5513, 1999.

101. Sutton, C. C. R., G. V. Franks and G. da Silva, "First Principles pK_a Calculations on Carboxylic Acids Using the SMD Solvation Model: Effect of Thermodynamic Cycle, Model Chemistry, and Explicit Solvent Molecules", *The Journal of Physical Chemistry B*, Vol. 116, No. 39, pp. 11999–12006, 2012.

102. Shende, A. T., N. Tayade and M. P. Tirpude, "Aqueous L-alanine molecular interaction from Gibb's free energy: CPCM and SMD in DFT and ultrasonic studies", *Journal of Physics: Conference Series*, Vol. 1913, No. 1, p. 012007, 2021.

103. Dixon, S. L. and P. C. Jurs, "Estimation of pK_a for organic oxyacids using calculated atomic charges", *Journal of Computational Chemistry*, Vol. 14, No. 12, pp. 1460–1467, 1993.

104. Svobodova Varekova, R., S. Geidl, C.-M. Ionescu, O. Skrehota, M. Kudera, D. Sehnal, T. Bouchal, R. Abagyan, H. J. Huber and J. Koca, "Predicting pK_a Values of Substituted Phenols from Atomic Charges: Comparison of Different

Quantum Mechanical Methods and Charge Distribution Schemes", *Journal of Chemical Information and Modeling*, Vol. 51, No. 8, pp. 1795–1806, 2011.

105. Parr, R. G. and Y. Weitao, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1995.

106. Frau, J., N. Hernandez-Haro and D. Glossman-Mitnik, "Computational prediction of the pK_as of small peptides through Conceptual DFT descriptors", *Chemical Physics Letters*, Vol. 671, pp. 138–141, 2017.

107. Yang, W. and W. J. Mortier, "The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines", *Journal of the American Chemical Society*, Vol. 108, No. 19, pp. 5708–5711, 1986.

108. Dutra, F. R., C. d. S. Silva and R. Custodio, "On the Accuracy of the Direct Method to Calculate pK_a from Electronic Structure Calculations", *The Journal of Physical Chemistry A*, Vol. 125, No. 1, pp. 65–73, 2021.

109. de Souza Silva, C. and R. Custodio, "Assessment of pK_a Determination for Monocarboxylic Acids with an Accurate Theoretical Composite Method: G4CEP", *The Journal of Physical Chemistry A*, Vol. 123, No. 38, pp. 8314–8320, 2019.

110. Judson, R., E. Jaeger, A. Treasurywala and M. Peterson, "Conformational searching methods for small molecules. II. Genetic algorithm approach", *Journal of Computational Chemistry*, Vol. 14, No. 11, pp. 1407–1414, 1993.

111. Stewart, J. J. P., "Optimization of parameters for semiempirical methods I. Method", *Journal of Computational Chemistry*, Vol. 10, No. 2, pp. 209–220, 1989.

112. Giese, T. J. and D. M. York, "Density-functional expansion methods: Evaluation of LDA, GGA, and meta-GGA functionals and different integral approximations", *The Journal of Chemical Physics*, Vol. 133, No. 24, p. 244107, 2010.

113. Glossman-Mitnik, D., "A comparison of the chemical reactivity of naringenin calculated with the M06 family of density functionals", *Chemistry Central Journal*, Vol. 7, No. 1, p. 155, 2013.

114. CJ Cramer, *Essentials of computational chemistry: theories and models*, Chicester, 2 edn., 2013.

115. Frank Jensen, *Introduction to Computational Chemistry, 2nd Edition*, 2 edn., 2007.

116. David Young, *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*, 2 2001.

117. E.G. Lewars, *Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, Vol. 3, Kluwer Academy Publishers, New York, 1st edn., 2003.

118. Mineva, T., E. Sicilia and N. Russo, "Density-Functional Approach to Hardness Evaluation and Its Use in the Study of the Maximum Hardness Principle", *Journal of the American Chemical Society*, Vol. 120, No. 35, pp. 9053–9058, 1998.

119. C*hata*rdenas, C., N. Rabi, P. W. Ayers, C. Morell, P. Jaramillo and P. Fuentealba, "Chemical Reactivity Descriptors for Ambiphilic Reagents: Dual Descriptor, Local Hypersoftness, and Electrostatic Potential", *The Journal of Physical Chemistry A*, Vol. 113, No. 30, pp. 8660–8667, 2009.

120. Parr, R. G. and W. Yang, "Density functional approach to the frontier-electron theory of chemical reactivity", *Journal of the American Chemical Society*, Vol. 106, No. 14, pp. 4049–4050, 1984.

121. Morell, C., A. Grand and A. Toro-Labbê, "Theoretical support for using the Δf(r) descriptor", *Chemical Physics Letters*, Vol. 425, No. 4-6, pp. 342–346, 2006.

122. Tozer, D. J. and F. De Proft, "Modeling temporary anions in density functional

theory: Calculation of the Fukui function", *The Journal of Chemical Physics*, Vol. 127, No. 3, p. 034108, 2007.

123. Martînez-Araya, J. I., G. Salgado-Morôn and D. Glossman-Mitnik, "Computational Nanochemistry Report on the Oxicams—Conceptual DFT Indices and Chemical Reactivity", *The Journal of Physical Chemistry B*, Vol. 117, No. 21, pp. 6339–6351, 2013.

124. Musa, A. Y., A. A. H. Kadhum, A. B. Mohamad, A. A. B. Rahoma and H. Mesmari, "Electrochemical and quantum chemical calculations on 4,4-dimethyloxazolidine-2-thione as inhibitor for mild steel corrosion in hydrochloric acid", *Journal of Molecular Structure*, Vol. 969, No. 1-3, pp. 233–237, 2010.

125. Albert, A., R. Goldacre and J. Phillips, "455. The strength of heterocyclic bases", *Journal of the Chemical Society (Resumed)*, p. 2240, 1948.

126. Settimo, L., K. Bellman and R. M. A. Knegtel, "Comparison of the Accuracy of Experimental and Predicted pK_a Values of Basic and Acidic Compounds", *Pharmaceutical Research*, Vol. 31, No. 4, pp. 1082–1095, 2014.

127. Contreras, R. R., P. Fuentealba, M. Galvân and P. Pêrez, "A direct evaluation of regional Fukui functions in molecules", *Chemical Physics Letters*, Vol. 304, No. 5-6, pp. 405–413, 1999.

128. Armarego, W. L. F., "107. Quinazolines. Part IV. Covalent hydration in the cations of substituted quinazolines", *Journal of the Chemical Society (Resumed)*, p. 561, 1962.

129. "The Application of Resonance Light Scattering Technique for the Determination of Tinidazole in Drugs", *Dyes and Drugs*, pp. 50–63, Apple Academic Press, 2016.

130. Zielinski, W. and A. Kudelko, "Acid-base interactions in some isoquinoline and quinazoline amino derivatives", *Arkivoc*, Vol. 2005, No. 5, pp. 66–82, 2004.