



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

DISSERTAÇÃO DE MESTRADO

**Otimização de Protocolo Computacional para
Predição de Valores de pK_a em
Solução de Dimetilsulfóxido com
Modelos Implícito e Híbrido**

SAYONARA MARIA FERREIRA DE ARAÚJO

**JOÃO PESSOA – PB – Brasil
NOVEMBRO / 2020**

SAYONARA MARIA FERREIRA DE ARAÚJO

**Otimização de Protocolo Computacional para
Predição de Valores de pK_a em
Solução de Dimetilsulfóxido com
Modelos Implícito e Híbrido**

Dissertação apresentada como
requisito para obtenção do
título de Mestre em Química pela
Universidade Federal da Paraíba.

Orientador: Prof. Dr. Otávio L. Santana

***Bolsista CAPES**

**JOÃO PESSOA – PB – Brasil
NOVEMBRO / 2020**

Catálogo na publicação
Seção de Catalogação e Classificação

A663o Araújo, Sayonara Maria Ferreira de.
Otimização de protocolo computacional para predição
de
valores de pKa em solução de dimetilsulfóxido com
modelos implícito e híbrido / Sayonara Maria Ferreira
de Araújo. - João Pessoa, 2021.
55 f. : il.

Orientação: Otávio Luís de Santana.
Coorientação: Railton Barbosa Andrade.
Dissertação (Mestrado) - UFPB/CCEN.

1. Solvatação. 2. Dinâmica molecular. 3. Protocolo
computacional. I. Santana, Otávio Luís de. II. Andrade,
Railton Barbosa. III. Título.

UFPB/BC

CDU 544.353.2(043)

Dissertação de Mestrado submetida ao Corpo Docente do Programa de Pós-Graduação em Química do Departamento de Química do Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba como parte dos requisitos para a obtenção do grau de Mestre em Química.

Banca examinadora:



Prof. Dr. Otávio L. Santana
Departamento de Química da UFPB
(Orientador)



Dr. Railton Barbosa Andrade
Departamento de Química da UFPB
(Co-orientador)



Prof. Dr. Gerd Bruno da Rocha
Departamento de Química da UFPB
(Membro Titular Interno)



Prof. Dr. Sidney Ramos de Santana
Departamento de Química da UFPB
(Membro Titular Externo)

AGRADECIMENTOS

Minhas primeiras palavras de agradecimento são direcionadas a Aquele que me proporcionou o sopro da vida e de uma existência maravilhosa, cheia de desafios, mas repleta de vitórias e momentos únicos. A DEUS, agradeço sem medidas.

Em seguida, agradeço aos meus PAIS por todo o esforço em proporcionar-me as melhores condições possíveis durante toda essa jornada de crescimento e por SEMPRE acreditarem em mim. Sem vocês nada disso teria sido possível.

Meus sinceros agradecimentos a minha família: às minhas irmãs, Samara e Suerda, que sempre estiveram ao meu lado, me fortalecendo com palavras e exemplo na lida em seus cursos de pós-graduação também; à minha tia, Cristina, pelo grande exemplo de garra em todos os campos de sua vida, especialmente, o profissional.

Agradeço a todas as minhas amigas e amigos que sempre acreditaram em mim: Verlúcia Amanda, Yane Lisset, Thayana Maria, Maria Vânia, Jefferson Cristiano, Diego Gomes, Anne Karinny, e todos os seu familiares. Amigos, vocês são preciosos.

Meus agradecimentos ao meu orientador, prof. Dr. Otávio Luís de Santana, pela condução espetacular de todo esse processo de construção de conhecimento, sempre acessível e disposto a encontrar soluções para os problemas que foram surgindo ao longo do caminho e pelo exemplo de profissional a ser seguido e por ter acreditado no meu potencial, mesmo sabendo que não tinha tido nenhuma experiência na área de modelagem molecular na graduação.

Igualmente, agradeço ao Dr. Railton Barbosa, meu co-orientador, pela ajuda sem medidas. Não tinha dia nem hora para me ajudar quando fosse necessário. Você é um ser raro. Muito obrigada por todas as vezes que me estendeu a mão e por todo o conhecimento compartilhado comigo.

Agradeço, ainda, a todos os colegas com os quais tive a oportunidade de compartilhar a “bancada” do laboratório durante esse tempo e aprender coisas que foram fundamentais para o meu crescimento.

Agradeço ao Programa de Pós-Graduação em Química da UFPB, a todos os professores pelo aprendizado compartilhado e a Marcos por toda a ajuda dispendida.

Agradeço à CAPES, pelo financiamento concedido.

RESUMO

A acidez ou basicidade de um composto consiste em uma propriedade que é reflexo direto de sua propensão em doar ou aceitar um próton e seu conhecimento é de fundamental importância para o entendimento de diversos processos químicos e bioquímicos. Esta propriedade pode ser expressa por meio do pK_a da substância. Quando o sistema de interesse é modelado em solução os resultados obtidos frequentemente não são tão satisfatórios quanto em fase gasosa, visto que há dificuldades em mensurar os impactos provocados pelo efeito da atuação do solvente na reação, como avaliados em alguns estudos de propriedades físico-químicas de sistemas solvatados. Poucos trabalhos têm sido realizados sobre a predição de pK_a em solventes apróticos como o dimetilsulfóxido (DMSO). O presente estudo propõe a otimização de um protocolo computacional de modo a ser possível a previsão com erros absolutos inferiores a 1,8 unidades de pK_a em relação aos valores experimentais. Os cálculos neste trabalho foram realizados no nível M06-2X 6-31+G(d) com o modelo de solvatação implícita C-PCM visando a predição de valores de pK_a em DMSO para 105 compostos orgânicos não correlacionados através de um procedimento direto. Foram realizadas duas etapas neste estudo. A primeira etapa consistiu no cálculo das energias livres dos estados inicial e final da reação de desprotonação dos 105 ácidos em DMSO utilizando apenas o modelo de solvatação C-PCM. A segunda etapa do estudo teve como finalidade o cálculo do pK_a absoluto das 10 primeiras estruturas em uma abordagem híbrida, na qual se incluiu uma molécula explícita de DMSO, tratando as demais moléculas de solvente implicitamente com o modelo contínuo de solvatação. As moléculas explícitas de DMSO foram posicionadas a partir da análise de trajetórias de simulações de dinâmica molecular *ab initio* realizadas no nível GFN2-xTB e, posteriormente, foram realizados cálculos no mesmo nível da primeira etapa. Os resultados obtidos, para o cálculo do pK_a apenas com solvatação implícita, apresentaram um significativo coeficiente de correlação ($r^2 = 0,950$), indicando um elevado poder de previsão. Na abordagem híbrida foi possível obter um erro absoluto médio um pouco acima dos obtidos pela abordagem implícita para a previsão de valores absolutos e significativamente menores que os obtidos para os valores relativos.

Palavras-Chave: solvatação, dinâmica molecular, protocolo computacional.

ABSTRACT

The acidity or basicity of a compound consists of a property that is a direct reflection of its propensity to donate or accept a proton and its knowledge is of fundamental importance for the understanding of several chemical and biochemical processes. This property can be expressed through the substance's pK_a . When the system of interest is modeled in solution, the results obtained are often not satisfactory as in the gas phase, as there are difficulties in measuring the impacts caused by the effect of the solvent acting on the reaction, as evaluated in some studies of physico-chemical properties of solvated systems. Few studies have been carried out on the prediction of pK_a in aprotic solvents such as dimethyl sulfoxide (DMSO). The present study proposes the optimization of a computational protocol in order to make prediction possible with absolute errors below 1.8 units of pK_a in relation to the experimental values. The calculations in this work were performed at the level M06-2X/6-31+G(d) with the implicit solvation model C-PCM aiming at the prediction of pK_a values in DMSO for 105 organic compounds not correlated through a direct procedure. Two steps were taken in this study. The first step consisted of calculating the free energies of the initial and final states of the deprotonation reaction of 105 acids in DMSO using only the C-PCM solvation model. The second stage of the study aimed to calculate the absolute pK_a of the first 10 structures in a hybrid approach, which included an explicit DMSO molecule, treating the other solvent molecules implicitly with the continuous solvation model. The explicit DMSO molecules were positioned from the trajectory analysis of *ab initio* molecular dynamics simulations performed at the GFN2-xTB level and, subsequently, calculations were performed at the same level as the first stage. The results obtained, for the calculation of pK_a only with implicit solvation, showed a significant correlation coefficient ($r^2 = 0.950$), indicating a high predictive power. In the hybrid approach it was possible to obtain an average absolute error slightly above those obtained by the implicit approach for the prediction of absolute values and significantly less than those obtained for the relative values.

Keywords: pK_a , chemical accuracy, hybrid method.

LISTA DE ILUSTRAÇÕES

Figura 1: Estados de protonação da espécie de interesse (A) e do solvente (B, DMSO), empregados no cálculo das energias livres. Fonte: própria.....	15
Figura 2: Ciclo termodinâmico para o cálculo do pKa a partir de suas diversas contribuições para a composição do valor do ΔG_r . Fonte: Yue (2010).....	22
Figura 3: Descrição macroscópica do dielétrico. Fonte: Pliego Jr. (2006).....	25
Figura 4: Cavidades VdW (verde), SES (vermelho) e SAS (azul). Fonte: Quan (2018).....	26
Figura 5: Cavidades VdW, SES e SAS. Fonte: site do ORCA.....	26
Figura 6: Correlação entre resultados experimentais e teóricos, obtidos com o método M06-2X/6-31+G(d), utilizando o procedimento direto/absoluto para os 105 ácidos em DMSO.....	36
Figura 7: Correlação entre resultados experimentais e teóricos, obtidos com o método M06-2X/6-31+G(d), utilizando o procedimento direto/relativo para os 105 ácidos em DMSO.....	36

LISTA DE TABELAS

Tabela 1: Estruturas com os menores erros absolutos de pK_a relativos (até 1,8 unidades) previstos pelo procedimento direto com o método M06-2X/6-31+G(d) e modelo de solvatação C-PCM.....	37
Tabela 2: Valores de pK_a absolutos previstos pelo procedimento direto para as 10 primeiras estruturas com o método M06-2X/6-31+G(d) pelo protocolo de solvatação implícita e híbrida (erros absolutos em parêntesis).....	40

LISTA DE ABREVIATURAS E SIGLAS

C-PCM: *Modelo de Cálculo pelo Condutor Polarizável*

DFT: *Teoria do Funcional de Densidade*

DMSO: *Dimetilsulfóxido*

HK: *Hohenberg e Kohn*

IPCM: *Modelo do Contínuo Polarizável de Isodensidade*

KS-SCF: *Método de Campo Auto-Consistente com os orbitais de Kohn-Sham*

PCM: *Método Contínuo Polarizável*

SAS: *Superfície Acessível ao Solvente*

SCF: *Método de Campo Auto-Consistente*

SCIPCM: *Modelo do Contínuo Polarizável de Isodensidade Auto-Consistente*

SES: *Superfície Excluída do Solvente*

SMD: *Método de Solvatação Universal baseado na densidade eletrônica*

TFD: *Thomas-Fermi-Dirac*

VdW: *Superfície de Van der Waals*

SUMÁRIO

Agradecimentos.....	5
Resumo.....	6
Abstract.....	7
Lista de Ilustrações.....	8
Lista de Tabelas.....	9
Lista de Abreviaturas e Siglas.....	10
1. INTRODUÇÃO.....	14
2. OBJETIVOS.....	17
2.1 Objetivo Geral.....	17
2.2 Objetivos Específicos.....	17
3. FUNDAMENTAÇÃO TEÓRICA.....	19
3.1 Transferência de Prótons: Definições e Aspectos Termodinâmicos.....	19
3.2 Cálculo Teórico do pK_a : Procedimentos Computacionais.....	20
3.2.1 Procedimento Direto.....	20
3.2.2 Procedimento Indireto.....	22
3.3 Métodos Computacionais: Aspectos Gerais no Cálculo de pK_a	23
3.3.1 Teoria do Funcional da Densidade (DFT).....	23
3.3.2 Métodos de Solvatação.....	24
3.3.2a Modelos de Solvente Implícito.....	24
3.3.2b Abordagem Híbrida: Implícito-Explícito.....	27
3.3.3 Tendências para o Desenvolvimento de Metodologias para o Cálculo de pK_a	29
4. METODOLOGIA.....	31
5. RESULTADOS E DISCUSSÃO.....	35
5.1 Cálculo de pK_a com Modelo Implícito de Solvatação C-PCM.....	35
5.2 Cálculo de pK_a com Modelo Híbrido de Solvatação.....	38
6. CONCLUSÃO.....	43
7. PERSPECTIVAS.....	45
8. APÊNDICES.....	46
8.1 Modelo de Input para Otimização de Geometria e Cálculo de Frequências.....	46

8.2 Estruturas Investigadas (Forma Ácida).....	47
8.3 Energias Livres após Otimização com Solvatação Implícita (ácidos & Bases).....	48
8.4 Dados das Simulações Moleculares.....	49
9. REFERÊNCIAS.....	51

CAPÍTULO 1

INTRODUÇÃO

1. INTRODUÇÃO

A acidez ou basicidade de um composto, expressa por meio do seu pK_a , consiste em uma propriedade que é reflexo direto de sua propensão em doar ou aceitar prótons e seu conhecimento é de fundamental importância para o entendimento de diversos processos químicos e bioquímicos [1]. A determinação do pK_a envolve o cálculo da diferença de energia livre da reação de transferência de prótons entre soluto e solvente. O tratamento teórico deste processo tem sido discutido em diversos trabalhos recentes [2-14]. A modelagem molecular para a predição de valores teóricos envolve a determinação das energias livres de duas estruturas, uma de referência (o ácido propriamente dito) e uma desprotonada (a sua base conjugada, obtida após a liberação de um próton).

A crescente evolução da Química Computacional tem contribuído para a determinação cada vez mais precisa de propriedades físico-químicas quando a modelagem envolve moléculas pequenas em fase gasosa, utilizando-se abordagens que incluam conjuntos de bases extensas e que consigam recuperar efeitos de correlação eletrônica, como as abordagens pós-Hartree-Fock ou as baseadas na *Teoria do Funcional da Densidade* (DFT, do inglês *Density Functional Theory*) [15-17]. Esforços também vem sendo empregados na utilização de abordagens teóricas para a determinação da acidez/basicidade de moléculas em solução, visto que diversas espécies não são passíveis de caracterização experimental como, por exemplo, intermediários de reação, que, de forma geral, são espécies de existência transitória, nem sempre detectáveis experimentalmente, comuns de serem formados em reações orgânicas [1]. Todavia, frequentemente a modelagem dos sistemas de interesse em solução não é tão satisfatória quanto em fase gasosa, visto que há dificuldades em mensurar os impactos provocados pelo efeito da atuação do solvente na reação [8,9]. Dessa forma, são necessárias abordagens que consigam abarcar esses efeitos, incluindo um método de solvatação apropriado ao escopo do estudo a ser realizado.

Os métodos de solvatação mais empregados em abordagens teóricas para o tratamento de sistemas condensados são os denominados *contínuos*, que são de fácil utilização. Um bom exemplo desta classe de métodos é o *Modelo Contínuo Polarizável* (PCM, do inglês *Polarizable Continuum Model*), desenvolvido por Tomasi e colaboradores [18,19]. Este modelo trata o solvente como um contínuo dielétrico, com suas moléculas tratadas de forma implícita, as quais envolvem uma molécula de soluto que possui um certo valor de momento de dipolo μ , inserido em uma cavidade de modo a gerar uma polarização no meio dielétrico.

Diversos estudos reportaram bons resultados na predição de valores de pK_a de ácidos orgânicos em água utilizando o modelo PCM quando utilizado para o cálculo dessa propriedade de ácidos carboxílicos e outros ácidos orgânicos em água [2,3], mas relativamente poucos trabalhos têm sido realizados sobre a predição em solventes apróticos, como o dimetilsulfóxido (DMSO), um solvente de grande relevância na condução de diversas reações químicas em laboratório. Contudo, Pliego e Riveros [20-22] observaram que o modelo PCM pode levar a erros de até 7 unidades de pK_a quando os ácidos apresentam uma faixa maior nos valores dessa propriedade. Isto corresponde a diferenças muito elevadas entre os valores teóricos e experimentais para as energias livres de Gibbs da reação de transferência de prótons. Visando a minimização de erros no cálculo teórico, Pliego e Riveros desenvolveram uma abordagem híbrida, no que se refere ao tratamento da solvatação, através de um modelo *cluster*-contínuo, combinando a inclusão de moléculas explícitas de solvente com o modelo contínuo de solvatação no cálculo das energias livres, tendo conseguido uma redução do erro de 7 para cerca de 2 unidades de pK_a de 17 ácidos em solução aquosa. Essa redução do erro é de grande relevância para áreas como a Bioquímica, que necessitam dos valores de pK_a para correlacionar a estrutura e a reatividade de diversos compostos. Todavia, nota-se que poucos trabalhos têm sido desenvolvidos com a finalidade de contribuir na predição de valores de pK_a em soluções orgânicas, mesmo esse tipo de conhecimento sendo muito importante para áreas como a Química Orgânica [20,21].

O presente estudo propõe a otimização de um protocolo computacional para a determinação de valores de pK_a com erros absolutos inferiores a 1,8 unidades de pK_a em relação aos valores experimentais, considerando um banco de dados contendo 105 compostos orgânicos não relacionados em DMSO. O protocolo consiste em um procedimento de cálculo direto das energias livres de Gibbs de dois estados de protonação, como descrito na Figura 1.



Figura 1: Estados de protonação da espécie de interesse (A) e do solvente (B, DMSO), empregados no cálculo das energias livres. Fonte: própria.

CAPÍTULO 2

OBJETIVOS

2. OBJETIVOS

2.1 OBJETIVO GERAL

Otimizar um protocolo computacional para a predição de valores absolutos e relativos de pK_a visando a obtenção da exatidão química (erros absolutos inferiores a 1,8 unidades da propriedade), considerando um conjunto de 105 moléculas orgânicas.

2.2 OBJETIVOS ESPECÍFICOS

1. Para cada uma das 105 moléculas do conjunto calcular os valores das energias livres da estrutura de referência (ácido) e da estrutura desprotonada (base) em DMSO com um modelo de solvatação implícita (abordagem implícita);
2. Calcular os valores absolutos e relativos de pK_a em DMSO pelo procedimento direto a partir das energias livres obtidas com modelo de solvatação implícita;
3. Para um subconjunto de 10 moléculas proceder a simulação computacional de dinâmica molecular visando identificar estruturas compatíveis com o processo de transferência de próton para cada par ácido/solvente;
4. Para cada *cluster* ácido/solvente identificado da trajetória da simulação proceder a otimização de geometria e cálculo de frequências para a obtenção das energias livres com modelo de solvatação implícita (abordagem híbrida);
5. Avaliar o efeito da inclusão de uma molécula de solvente nos valores calculados de pK_a em relação aos valores experimentais.

CAPÍTULO 3

FUNDAMENTAÇÃO TEÓRICA

3. FUNDAMENTAÇÃO TEÓRICA

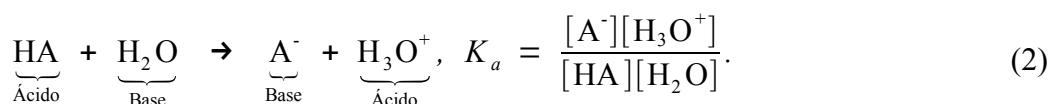
3.1 TRANSFERÊNCIA DE PRÓTONS: DEFINIÇÕES E ASPECTOS TERMODINÂMICOS

Existem algumas definições fundamentais para compreendermos as reações de transferências de prótons como as de ácidos e bases. Uma muito utilizada, que descreve muitas reações ácido-base em química orgânica, considera que uma espécie será considerada um ácido se for capaz de doar um próton [23,24], estabelecendo o seguinte equilíbrio:

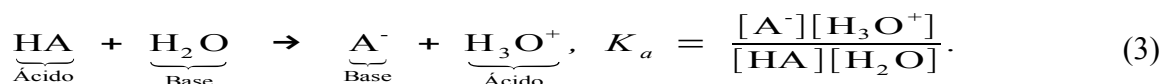


A perda de um próton (H^+) pelo ácido gera sua base conjugada e, em contrapartida, a protonação da base gera seu ácido conjugado. Neste caso a reação consiste em uma transferência protônica, de modo que esta definição está centrada na migração de um próton.

Como os ácidos possuem habilidades diferentes ao doar próton, pode-se dizer, de forma geral, que a sua força, de acordo com Bronsted-Lowry, pode ser descrita através da sua constante de equilíbrio da reação de sua dissociação em água (eq. 2), determinada pela estabilidade relativa do ácido e da base conjugada, e pode ser determinada por:



Ácidos fortes apresentam o equilíbrio descrito acima deslocado para a direita, levando a valores altos da constante ácida, e o inverso ocorre para ácidos considerados fracos. Por conveniência, expressa-se a força dos ácidos em uma escala logarítmica ($\text{p}K_a$) [23,24], definida por:



Em Química Orgânica, vale ressaltar, os ácidos carboxílicos são ácidos orgânicos amplamente utilizados, apresentando um caráter ácido maior do que os álcoois devido a máxima estabilização do íon carboxilato (base conjugada) por ressonância quando há a saída de um próton de uma estrutura que possua esse grupo [23,24]. A energia livre da base conjugada é menor do que a

do ácido, aumentando a constante de equilíbrio, o que não ocorre nos álcoois. No caso das estruturas que possuem esse último grupo funcional, onde os valores de pK_a estão em torno de 16, a ausência de estabilização por ressonância verificada no íon alcóxi (base conjugada) impacta as energias livres e, portanto, os valores de pK_a . A estabilização por ressonância também explica o fato dos fenóis serem mais ácidos (pK_a menores) do que os álcoois alifáticos e menos ácidos do que os ácidos carboxílicos.

3.2 CÁLCULO TEÓRICO DO pK_a : PROCEDIMENTOS COMPUTACIONAIS

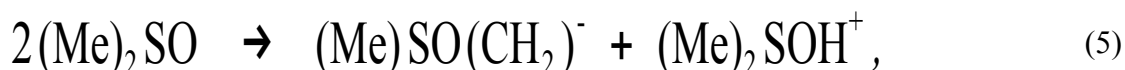
O cálculo computacional de valores de pK_a pode ser realizado através de dois procedimentos gerais [1-21]. O procedimento mais utilizado é o *indireto*, que leva em consideração um ciclo termodinâmico para descrever separadamente a reação de desprotonação em fase gasosa (tratada em um nível de cálculo mais alto) e os processos de solvatação das espécies envolvidas na reação (frequentemente empregando métodos de solvatação implícita, em combinação com um método de estrutura eletrônica compatível com a calibração do método de solvatação). Um segundo procedimento geral, denominado *direto*, consiste no cálculo das energias livres diretamente em fase solvatada empregando-se um único método de estrutura eletrônica.

3.2.1 Procedimento Direto

Com menor custo computacional (e maior simplicidade) que o indireto, este procedimento é baseado na reação de transferência protônica, considerando a estrutura inicial, o ácido, que se pretende determinar o valor do pK_a , e uma estrutura que atua como base [1], a qual faz parte do solvente. No caso específico da reação de desprotonação de um ácido HA na presença do DMSO, a reação a ser considerada é a seguinte [1]:

$$pK_a = -\log K_a, \quad K_a = e^{-\Delta G_f/RT} \Rightarrow pK_a = \frac{\Delta G_f}{RT} \log e = \frac{\Delta G_f}{RT \ln 10}. \quad (4)$$

Com base nas eqs. 2 e 3, obtém-se:



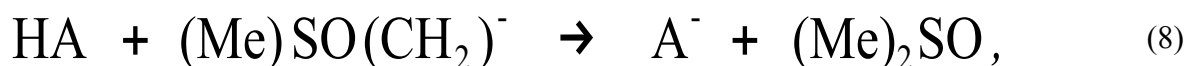
O valor experimental do pK_a do DMSO em DMSO ($pK_a \approx 35$, a 298 K) [1] está relacionado a seguinte reação de autoionização:

$$pK_a(\text{DMSO}) = \frac{[G((\text{Me})\text{SO}(\text{CH}_2)^-) + G((\text{Me})_2\text{SOH}^+)] - [2G((\text{Me})_2\text{SO})]}{RT \ln 10} = 35, \quad (6)$$

de modo:

$$pK_a(\text{HA}) = \frac{[G(\text{A}^-) + G((\text{Me})_2\text{SO})] - [G(\text{HA}) + G((\text{Me})\text{SO}(\text{CH}_2)^-)]}{RT \ln 10} + 35. \quad (7)$$

o que possibilita escrever uma expressão para o pK_a relativo (do ácido HA em relação ao DMSO):



Esta última expressão é consistente ao cálculo do pK_a segundo a reação:

$$\Delta G_r^{(\text{solv})} = \Delta G_r^{(\text{gas})} + \Delta \Delta G_{\text{solv}}, \quad (9)$$

corrigida pelo pK_a experimental do solvente [1]. Esta última estratégia é útil para eliminar a solvatação do próton (especialmente no caso do tratamento indireto, via ciclo termodinâmico). No entanto, este problema também é contornado no tratamento direto quando a reação de equilíbrio é tratada na forma expressa pela eq. 4, na qual o próton não aparece livre.

O custo computacional do procedimento direto (tanto do pK_a absoluto, eq. 5, quanto o relativo, eq. 8) é muito menor que o relacionado ao procedimento indireto, visto que há um número menor de cálculos de estrutura eletrônica a serem realizados, onde todas as etapas do procedimento são realizados no mesmo nível (método e conjunto de base) [25] e calcula-se a energia livre das espécies reagente e produto diretamente em solução ($G^0(\text{J}, \text{svt})$, onde “J” se refere a espécie reagente ou produto e “svt” se refere a fase solvatada).

Em alguns estudos computacionais realizados por Junming Ho foi possível obter uma boa concordância entre valores teóricos e experimentais de pK_a e de potenciais de redução de aminoácidos em água e DMSO através da abordagem via procedimento direto utilizando o método de solvatação implícita SMD [26,27]. Este modelo de solvatação, baseado na densidade eletrônica do soluto, é recomendado para sistemas nos quais são observadas pequenas variações na geometria induzidas pela solvatação.

3.2.2 Procedimento Indireto

Uma outra possibilidade para obtenção de valores de pK_a a partir das energias livres de espécies envolvidas em uma reação de transferência de próton é o procedimento indireto, que combina, a partir de um ciclo termodinâmico, as energias livres de Gibbs envolvidas no processo em fase gasosa com as energias livres de solvatação, como pode ser observado na Figura 2.

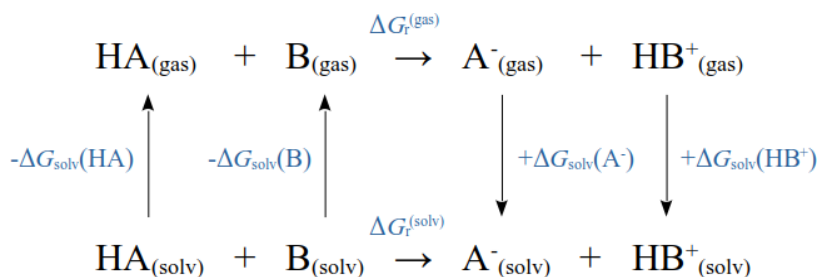


Figura 2: Ciclo termodinâmico para o cálculo do pK_a a partir de suas diversas contribuições para a composição do valor do ΔG_r .
Fonte: Yue (2010).

A descrição do ciclo termodinâmico apresentado na Figura 2 possibilita obter o pK_a de um ácido HA a partir da decomposição da variação da energia livre da reação de desprotonação ácida em determinado solvente ($\Delta G_r^{(\text{solv})}$), possibilitando empregar níveis de cálculo diferentes para cada uma das contribuições. Em geral, emprega-se um método de maior custo computacional para descrever a reação em fase gasosa ($\Delta G_r^{(\text{gas})}$) e um método compatível com o modelo de solvatação (ΔG_{solv}) [8]. No procedimento indireto o custo computacional é significativamente superior em relação ao direto devido ao fato de que há mais cálculos a serem realizados, além de ser empregado um método de maior custo na descrição da fase gasosa. A energia livre da reação em determinado solvente é calculada como:

$$\Delta \Delta G_{\text{solv}} = \Delta G_{\text{solv}}(\text{A}^{-}) + \Delta G_{\text{solv}}(\text{HB}^{+}) - \Delta G_{\text{solv}}(\text{HA}) - \Delta G_{\text{solv}}(\text{B}), \quad (10)$$

e o pK_a é calculado a partir da eq. 3.

3.3 MÉTODOS COMPUTACIONAIS: ASPECTOS GERAIS NO CÁLCULO DE pK_a

Os métodos computacionais mais empregados no cálculo teórico de propriedades termodinâmicas como o pK_a precisam considerar a sensibilidade da propriedade às pequenas mudanças na determinação dos valores da variação da energia livre de Gibbs, que sofre a influência dos efeitos provocados pela solvatação. Por esta razão, é necessário o emprego de estratégias computacionais que descrevam o ambiente da solução o mais realista possível, usando tanto modelos de solvente implícito, como o modelo contínuo utilizado no presente estudo, quanto modelos que busquem a inserção de moléculas explícitas do solvente no tratamento do cálculo. Idealmente, busca-se a composição de metodologias computacionais rápidas e de baixo custo, que descartem a necessidade de um grande número de etapas de cálculos (como preconizado no procedimento indireto) ou da utilização de métodos *ab initio* de alta demanda computacional (empregando-se, por exemplo, a DFT), como apontado por Abramson [28].

A viabilidade da utilização do procedimento via ciclo termodinâmico é bastante questionada na literatura para o cálculo de pK_a devido, entre outros fatores, ao custo computacional e a mistura de métodos utilizados nas etapas do cálculo, uma vez que estudos apontam que a exatidão na predição de valores de pK_a está intimamente relacionada a uma descrição do ambiente da solução, como discutido em diversos estudos [28-31].

3.3.1 Teoria do Funcional da Densidade (DFT)

Apesar das falhas apresentadas pelo modelo Thomas-Fermi-Dirac (TFD), não se pode negar a sua contribuição para a *Teoria do Funcional de Densidade* (DFT), considerada uma alternativa computacional mais econômica aos métodos que usam a função de onda multieletrônica [32,33], pois foi um modelo que estava sustentado na aplicação de um funcional de energia desenvolvido por Thomas-Fermi, mas com o diferencial de adicionar o termo de energia de troca para um gás de elétrons proposta por Dirac [34].

Os fundamentos da DFT foram desenvolvidos em 1964, quando da publicação de dois teoremas por Hohenberg e Kohn, os quais demonstraram que existe um funcional de energia exato da densidade eletrônica, $E[\rho]$ [35]. Na descrição do seu primeiro teorema, Kohn e Hohenberg definem que o potencial externo em relação aos elétrons, V_e , é um funcional único da densidade

eletrônica $\rho(\mathbf{r})$ (sendo $\mathbf{r}$ um vetor) somado a uma constante, o que levaria à seguinte expressão para a energia total do sistema:

$$E_0[\rho] = \langle \Psi | \hat{T} + \hat{V} | \Psi \rangle = \int \hat{V}_e[\rho(\mathbf{r})] d\mathbf{r}. \quad (11)$$

Com a elaboração do seu segundo teorema, Hohenberg e Kohn conseguiram esclarecer que a densidade eletrônica pode ser usada como uma variável para determinação da energia do estado fundamental de um sistema [35].

Além dos teoremas de Hohenberg e Kohn, Sham forneceu contribuições essenciais à teoria, através das chamadas equações de Kohn-Sham (KS) [36,37], resolvidas por meio de um procedimento auto-consistente (KS-SCF), levando a obtenção das funções referentes aos orbitais de Kohn-Sham, autofunções do Hamiltoniano efetivo KS que, por sua vez, depende da densidade eletrônica $\rho(\mathbf{r})$.

Para a realização de um cálculo DFT é preciso proceder a escolha de um funcional [32,33]. Por exemplo, o funcional M06-2X foi desenvolvido para a realização de cálculos que visam a obtenção de propriedades termoquímicas e cinéticas [38-40]. Em um estudo realizado por Silva [41], verificou-se que a utilização deste funcional, com o conjunto de base 6-31+G(d) (*duplo-zeta* com um conjunto de difusas e de polarização *d* em átomos pesados) [42,43], possibilitou a obtenção de potenciais de redução com um erro absoluto médio inferior a 25 mV (equivalente a cerca de 0,6 kcal·mol⁻¹ de erro na energia livre) em relação aos valores experimentais.

3.3.2 MÉTODOS de Solvatação

3.3.2a Modelos de Solvente Implícito

Os métodos de solvatação implícita [20-22,44], também conhecidos como modelos contínuos de solvente, são amplamente empregados na modelagem de diversos sistemas devido ao seu baixo custo computacional em relação aos métodos de solvatação explícita (que necessitam de um procedimento para a descrição de um número elevado de configurações nas quais o solvente pode se organizar em torno do soluto). Os modelos de solvatação (tanto implícitos quanto explícitos) desempenham papel relevante para a compreensão do efeito do meio condensado nas propriedade de interesse, tais como potenciais de redução e pK_a [26,27].

O modelo implícito do solvente possibilita descrever as interações de longo alcance entre o soluto e o solvente, tratando este último como um contínuo dielétrico [45]. Para uma descrição macroscópica do dielétrico, pode-se considerar que a contribuição eletrostática de cada molécula de solvente é proveniente da sua interação com a distribuição de carga do soluto (Figura 3).

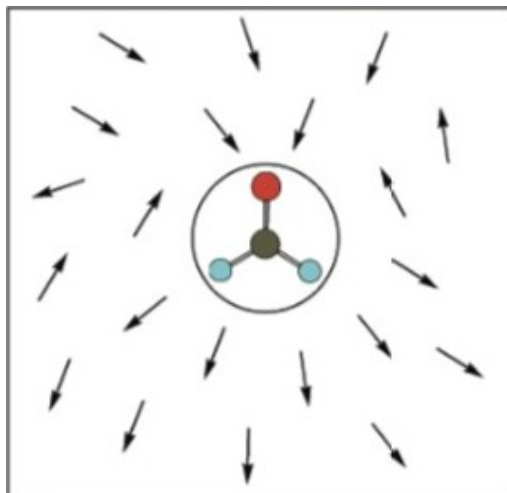


Figura 3: Descrição macroscópica do dielétrico.
Fonte: Pliego Jr. (2006).

Nesse modelo, o soluto é descrito de forma mecânico-quântica, levando-se em conta a polarização do solvente, que polariza de volta o soluto. Esta perturbação mútua é realizada até que a auto-consistência do modelo seja alcançada, razão pela qual o método também é denominado de *Campo de Reação Autoconsistente* (SCRF, do inglês *Self-Consistent Reaction Field*) [19,45].

Para explicitar o funcionamento dos modelos contínuos de solvatação é essencial definir a cavidade na qual o soluto será acomodado. Nesse modelo, o soluto é inserido em uma cavidade não acessível ao solvente [19-22,44,45]. O formato da cavidade considerada no modelo contínuo deve reproduzir da melhor maneira possível a forma molecular do soluto. Cavidades que não observem essa condição podem gerar deformações na distribuição de carga após a polarização do solvente, com grandes efeitos que não são reais nos resultados. De acordo com Tomassi e colaboradores (2005) [19], outro requisito para a cavidade é conseguir comportar a maior parte possível da distribuição de carga da molécula de soluto, sendo importante perceber que haverá uma sobreposição entre a distribuição de carga do soluto que está inserido na cavidade e a distribuição de carga do solvente.

Para uma descrição acurada da cavidade nos modelos contínuos, pode-se utilizar uma superfície de densidade eletrônica constante (superfície de isodensidade), através de uma definição *ab initio* da cavidade. No programa *Gaussian* o valor definido como padrão para a isodensidade corresponde a faixa 0,0004-0,001 u.a. [19].

É possível gerar cavidades para o soluto com diferentes critérios para a utilização em modelos contínuos. As cavidades mais usuais são VdW, SAS e SES (Figuras 4 e 5) [46]. A cavidade VdW é construída a partir dos raios atômicos do soluto (raios de van der Waals). A SAS (do inglês *Solvent Accessible Surface*) é definida pelo centro de uma sonda esférica que representa o solvente, posicionada no espaço molecular externo, a partir da qual a superfície acessível ao solvente é gerada rolando-a sobre a molécula de soluto (técnica empregada por Betley, 1998) [47]. A SES (do inglês *Solvent Excluded Surface*) é obtida pelo “método da esfera rolante” (assim como a SAS), a partir da qual a região excludente ao solvente é definida pela cavidade gerada pelos pontos da superfície da sonda (e não pelo centro, como na SAS), o que gera uma cavidade de volume intermediário (entre o da VdW e o da SAS) [48,49].

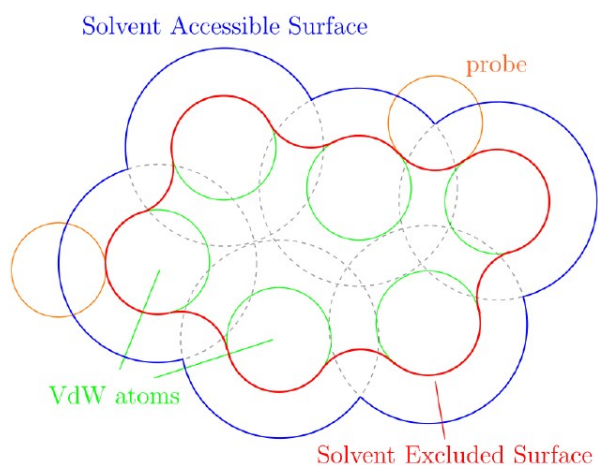


Figura 4: Cavidades VdW (verde), SES (vermelho) e SAS (azul). Fonte: Quan (2018).

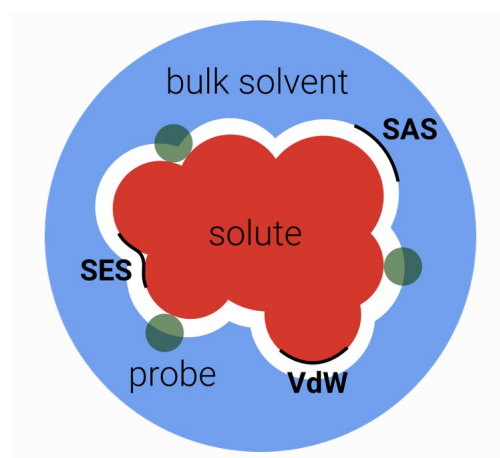


Figura 5: Cavidades VdW, SES e SAS. Fonte: site do ORCA.

Um modelo de solvente implícito largamente utilizado na modelagem de diversos sistemas é o do *contínuo polarizável* (PCM) [50-54]. Este modelo utiliza o método das cargas aparentes para obter o potencial eletrostático formado pelo dielétrico e cálculos de estrutura eletrônica *ab initio* determinam a densidade eletrônica, incluindo a polarização do soluto [22].

Modelos implícitos de solvente estão implementados em diversos programas, tais como *Gaussian* e *ORCA*. Uma versão recente do PCM, implementada no *Gaussian*, é geralmente conhecida como o modelo *isodensity*-PCM (IPCM) [55]. O mesmo modelo foi estendido para permitir que a superfície de isodensidade em cada iteração SCF seja variada, ou seja, a cavidade não é fixada, mesmo depois que a geometria é fixada (como é o caso no modelo PCM padrão), mas é relaxada para a isodensidade da molécula solvatada em cada iteração SCF. Este desenvolvimento adicional é geralmente conhecido como isodensidade autoconsistente (SCIPCM) [56]. Os dois modelos (IPCM e SCIPCM) estão implementados no *Gaussian*.

Apesar do seu uso ser bastante popular, os modelos que tratam o solvente implicitamente não conseguem descrever totalmente o ambiente da solução, levando a erros médios na energia de solvatação entre 0,6 a 1,0 kcal/mol para estruturas neutras, e de 4,0 kcal para íons [57]. Desse modo, é mais do que necessário abordagens computacionais mais rigorosas em relação à descrição dos efeitos do solvente, que incluam moléculas de solvente explicitamente interagindo com o soluto, melhorando, possivelmente, a acurácia dos valores de energia [58-62].

3.3.2b Abordagem Híbrida: Implícito-Explícito

A dificuldade no cálculo teórico do pK_a reside na elevada sensibilidade à variações na energia livre da reação de desprotonação. A acuracidade computacional é difícil de ser alcançada uma vez que um erro de $0,325 \text{ kcal} \cdot \text{mol}^{-1}$ na variação de energia livre da reação de transferência de um próton leva a um erro de 1 unidade de pK_a , sendo um grande desafio determiná-lo teoricamente. Em sua maioria, as abordagens teóricas utilizam ciclos termodinâmicos, aumentando o número de cálculos, o que pode levar a erros computacionais, principalmente quando o solvente é prótico e não se consegue calcular a energia livre do próton por métodos de Química Quântica [3,4]. Além disso, ainda existem poucos estudos para predição de pK_a em solventes orgânicos, como o DMSO.

Para uma determinação mais criteriosa de valores de pK_a , estudos tem sido conduzidos utilizando abordagem híbrida, também chamada de modelo do *cluster*-contínuo. Nesta abordagem, busca-se descrever com mais fidelidade os fenômenos que ocorrem na fase solvatada combinando-se o modelo de solvatação implícita (que visa descrever o efeito médio e de longo alcance do meio condensado) com a inclusão de algumas poucas moléculas de solvente no cálculo (visando descrever as interações soluto-solvente mais relevantes). Este procedimento faz uso simulações de dinâmica molecular [1,63-65].

Utilizar a dinâmica molecular como método para simular líquidos é vantajoso devido à consideração da evolução temporal da estrutura do meio condensado, o que possibilita o estudo de propriedades que tenham dependência temporal [66]. A dinâmica molecular permite avaliar como as configurações (que diz respeito às posições e orientações relativas das espécies moleculares) dos constituintes de um sistema evoluem no tempo. A partir dessa técnica computacional pode-se obter os movimentos das partículas de qualquer sistema desde que se conheça o potencial de interação entre elas e as equações que comandam seu movimento. Utilizando-se as equações de movimento newtonianas e as posições e velocidades de todas as partículas do sistema em um dado instante t_0 é possível determinar as forças resultantes em cada partícula em instantes subsequentes, possibilitando a produção sequencial de um número elevado de configurações ao longo do tempo, nas quais as posições e velocidades de cada partícula do sistema são conhecidas [67].

A quantidade de partículas do sistema a ser simulado é dependente da natureza do próprio sistema, das propriedades desejadas e da capacidade computacional que se tem disponível, devendo, sobretudo, representar o sistema macroscópico real. Para iniciar a simulação é necessário especificar as posições iniciais das partículas que compõem o sistema. As interações em simulações de dinâmica molecular, na maioria das vezes, são tratadas classicamente, assumindo que o sistema estudado apresente estados nos quais os efeitos quânticos possam ser desconsiderados. Em alguns casos os potenciais clássicos que descrevem as interações podem ser obtidos a partir de aproximações derivadas de cálculos quânticos [68,69]. Alternativamente, é possível realizar uma dinâmica baseada na aproximação de Born-Oppenheimer (método BOMD, do inglês *Born-Oppenheimer Molecular Dynamics*), que assume o desacoplamento do movimento dos núcleos (que é propagado no tempo segundo as equações newtonianas) e dos elétrons (cuja densidade eletrônica, determinada a partir de cálculos quânticos, é utilizada para derivar as forças que atuam nos núcleos; neste método a estrutura eletrônica é obtida a cada novo posicionamento dos núcleos) [70-72].

O procedimento mais comum para a construção da configuração inicial do sistema consiste na criação de uma caixa de simulação (geralmente cúbica) onde estão as N moléculas do sistema e sua dimensão é escolhida de modo que a densidade numérica (N/V) corresponda a do sistema real, de acordo com as condições termodinâmicas desejadas. Geralmente são utilizadas condições periódicas de contorno, na qual a caixa original é reproduzida espacialmente em todas as direções e, quando um átomo se move na caixa original, sua imagem periódica se move da mesma maneira nas demais caixas.

Pliego e Riveros avaliaram a aplicação do modelo do *cluster*-contínuo, comparando os resultados obtidos com os de modelos de solvatação implícita, obtendo boa concordância em relação aos valores experimentais (erro de 2,2 unidades de pK_a usando a abordagem híbrida contra 7 unidades de pK_a usando apenas o método implícito) [20-22]. Fu e colaboradores utilizaram a referida abordagem híbrida, em um procedimento indireto, para calcular o pK_a dos 105 ácidos orgânicos considerados neste trabalho, obtendo altos coeficientes de correlação (0,99) e desvios padrão de 1,8 e 1,7 unidades de pK_a [1].

3.3.3 Tendências para o Desenvolvimento de Metodologias para o Cálculo de pK_a

O desenvolvimento e aprimoramento de estratégias computacionais para a predição de propriedades como pK_a , solubilidade, lipofilicidade e a capacidade de uma substância atravessar a membrana plasmática são de grande importância para a Biologia e o mercado de planejamento de fármacos [73].

Uma das abordagens muito utilizadas para a previsão de tais propriedades físico-químicas consiste na utilização de modelos quantitativos de relação estrutura-atividade (QSAR) ou estrutura-propriedade (QSPR), desenvolvidos pela primeira vez em 1960 [74]. Modelos QSAR vem sendo desenvolvidos em associação com abordagens de aprendizado de máquina (*machine learning*) para a predição do pK_a . No entanto, essa abordagem depende demais da qualidade e quantidade de dados que são utilizadas nos cálculos, existindo uma escassez de conjunto de dados extensos, o que afeta a robustez do modelo [75-77].

Cálculos mais sofisticados incluem modelos treinados com rede neural para predição de pK_a , incluindo descritores de parâmetros orgânicos estruturais e físicos (SPOC) para representar as características eletrônicas e estruturais das moléculas [78]. Todavia, o custo computacional é elevado, o que limita a sua aplicabilidade.

Alternativamente, neste trabalho optou-se por um protocolo com menos etapas, visando a identificação das interações soluto-solvente relevantes para o tratamento computacional, considerando um banco de dados consolidado constituído por 105 moléculas não correlacionadas.

CAPÍTULO 4

METODOLOGIA

4. METODOLOGIA

As geometrias das 105 estruturas moleculares de referência (ácidos) investigadas no presente trabalho foram obtidas de um banco de dados de um estudo realizado em DMSO [1]. Os dados experimentais das referidas estruturas foram obtidos desde 1975 por Bordwell e diversos colaboradores na área de Química Orgânica, os quais foram compilados ao longo dos últimos 45 anos [79], tendo sido utilizados pela comunidade química há anos. Para as primeiras determinações experimentais foi desenvolvido um método espectrofotométrico com vistas a obter os valores de pK_a de ácidos orgânicos em DMSO. A escala de pK_a em DMSO foi ancorada por comparações de valores obtidos pelo método espectrofotométrico com aqueles obtidos por potenciometria na faixa de 8 a 11 pK_a . Como resultado, diversos ácidos orgânicos apresentaram aumento nos valores de suas constantes, arbitradas anteriormente ao método, inclusive o DMSO, solvente utilizado no presente estudo, cujo valor de pK_a foi determinado por Bordwell e colaboradores como sendo igual a 35,1 [80].

A primeira etapa do estudo consistiu no análise conformacional das estruturas dos 105 ácidos e do solvente DMSO em seus dois estados de protonação, nos respectivos estados eletrônicos fundamentais, utilizando-se o programa *CREST* (acrônimo de *Conformer-Rotamer Ensemble Sampling Tool*) [81,82] para a determinação das estruturas de mínima energia global.

Após determinadas as geometrias de mínima energia foram preparados os arquivos de entrada com as configurações dos cálculos de otimização de geometria e análise vibracional das estruturas ácida (referência) e de sua base conjugada (estrutura desprotonada), totalizando quatro cálculos a partir de um único arquivo de entrada (*input* duplo, Apêndice 8.1). Do total de estruturas de referência investigadas 6 são catiônicas e as demais neutras. As cargas das estruturas e suas multiplicidades foram cuidadosamente avaliadas e o mesmo nível de cálculo foi empregado para as duas estruturas. Todos os cálculos de estrutura eletrônica foram realizados com o programa *Gaussian 09* [83,84], com o método DFT [34-37] utilizando o funcional M06-2X [38-40], conjunto de funções de base 6-31+G(d) [42] e modelo de solvatação C-PCM [85] para a descrição do efeito do solvente DMSO.

Os cálculos de frequências vibracionais realizados possibilitaram confirmar a obtenção de estruturas de mínima energia, bem como determinar as energias livres de Gibbs, necessárias para a determinação dos valores teóricos de pK_a . Todos os cálculos foram realizados na configuração

padrão do *Gaussian 09* (1 bar, 298K). Não se realizou o ajuste para o estado padrão em solução (1 M) pois ocorre o cancelamento do trabalho de compressão ($\sim 1,9$ kcal/mol) [86] devido ao igual número de termos nos dois lados das eqs. 4 e 9. Os valores das energias livres retiradas dos arquivos de saída foram convertidos de hartree para kJ/mol.¹ Os valores de pK_a absoluto foram calculados segundo as eqs. 4 e 5, e os valores de pK_a relativo a partir das eqs. 8 e 9, utilizando, nos dois casos, o procedimento de cálculo direto. Os valores obtidos foram avaliados a partir do cálculo do erro absoluto médio (e_{abs}), do erro absoluto máximo (e_{max}), do desvio padrão dos erros (σ_{abs}) e do coeficiente de correlação linear (r^2), calculados em relação aos valores experimentais.

A segunda etapa do estudo teve como finalidade incluir moléculas explícitas de solvente no cálculo do pK_a das 10 primeiras estruturas, executando, assim, uma abordagem híbrida, ou seja, fazendo uso do modelo do contínuo dielétrico e moléculas explícitas de DMSO. Com o programa *Solvate* [87] foram geradas caixas esféricas de simulação que contém o número mínimo de moléculas de solvente para envolver completamente o soluto (este número é determinado iterativamente). Com as caixas geradas foram realizadas simulações de dinâmica molecular clássica com potenciais quânticos (BOMD), utilizando-se o programa *ORCA 4.2* [88] com o hamiltoniano semiempírico GFN2-xTB, em um *ensemble* NVT na temperatura de 298 K, com 6000 passos de 0,5 fs. Como critério inicial, foram descartados os primeiros 1000 passos (tratados aqui como *passos de equilíbrio*) e analisados os últimos 5000 passos (tratados aqui como *passos de produção*), no qual foram buscadas estruturas compatíveis com o processo de transferência de próton.² O protocolo da referida etapa contém os seguintes passos, de forma mais detalhada:

- **Etapa 1: Empacotamento**

Objetivo: Geração da caixa de simulação com o programa *Solvate*. O número de moléculas de solvente é determinado iterativamente pelo programa. Após a criação da caixa de simulação o programa prepara o *input* para a execução da simulação de dinâmica molecular.

Comando:

```
solvate <soluto>.com <solvente>.com [-npck=I -shape=box -density=N]  
[I=inteiro; N=real]
```

¹ Fator de conversão utilizado: 2625,5 kJ·mol⁻¹/hartree.

² É importante destacar que, apesar desta denominação, não afirmamos que a equilíbrio foi atingida com apenas 1000 passos de 0,5 fs (equivalente a 0,5 ps de simulação). No entanto, o critério aqui adotado se baseou em testes realizados com um dos sistemas investigados. A estratégia visa identificar estruturas compatíveis com o processo de transferência de prótons a partir da identificação da distância e orientação mais favorável para a molécula de solvente. A estrutura obtida da trajetória da simulação é, posteriormente, otimizada, de modo que, a partir deste ponto, a simulação é descartada. A estratégia adotada visa a redução do custo computacional da simulação.

- **Etapa 2: Simulação**

Objetivo: Realização da simulação com o programa *ORCA/xTB* a partir da caixa de simulação gerada na etapa anterior. Reexecuta-se o *Solvate* acrescentando a informação do número de moléculas de solvente (*nmol*) na caixa de simulação, acrescentando-se a opção de inicialização da dinâmica (*mdsim*) e o número de passos de simulação (*nsteps* = 6000). Na versão atual tamanho do passo (*tsteps*) está ajustado para 0,5 fs por padrão.

Comando:

```
solvate <soluto>.com <solvente>.com -nmol=I1 -mdsim -nsteps=I2
```

- **Etapa 3: Análise da Trajetória**

Objetivo: Identificação de estruturas compatíveis com o processo de transferência de prótons. Aqui o interesse é reconhecer apenas as moléculas de solvente que interagem mais efetivamente com o soluto. Considerou-se apenas uma molécula de solvente mais próxima ao soluto. Para isto, o *Solvate* foi reexecutado retirando a opção de execução de simulação e acrescentando o parâmetro para o número de moléculas de solvente a selecionar (*nsvt*) e o intervalo da simulação a ser analisado (*cfgtrj*; o primeiro parâmetro especifica o número de passos a serem desconsiderados, enquanto o segundo indica o número de passos a serem analisados). Foi utilizado um *script* para realizar automaticamente este procedimento de modo a gerar 10 estruturas a partir da dinâmica, as quais foram otimizadas a nível GFN2-xTB, com o modelo de solvatação GBSA para representar o DMSO. Ao final as estruturas obtidas foram comparadas.

Comando:

```
solvate <soluto>.com <solvente>.com -nmol=I1 -nsvt=I2 -cfgtrj=[I3:I4]
```

- **Etapa 4: Submissão**

Objetivo: Realização da otimização da estrutura selecionada na etapa de análise da trajetória. Com todos os arquivos gerados até aqui pode-se proceder aos cálculos de otimização de geometria e cálculo de propriedades termodinâmicas. Esta etapa foi realizada com o programa *Gaussian 09* em separado do programa *Solvate*, de modo que outras análises e ajustes sejam realizados de maneira independente.

CAPÍTULO 5

RESULTADOS E DISCUSSÃO

5. RESULTADOS E DISCUSSÃO

5.1 CÁLCULO DE pK_a COM MODELO IMPLÍCITO DE SOLVATAÇÃO C-PCM

Nesta etapa do estudo, a determinação teórica dos valores de pK_a das 105 estruturas foi conduzida utilizando-se o procedimento direto para, primeiramente, obter os valores das energias livres do estado de referência (estrutura do ácido) e do estado desprotonado (estrutura da base conjugada) em DMSO, com o método implícito de solvatação C-PCM. Os valores de pK_a foram determinados conforme as eqs. 4 e 5 (valores absolutos) e eqs. 8 e 9 (valores relativos). Os resultados obtidos para os valores de pK_a absoluto e relativo para as 105 estruturas são mostrados nas Figuras 6 e 7 (os pontos em vermelho foram descartados da análise e necessitam de maior investigação).

Em relação aos resultados obtidos nesta primeira etapa, os valores de pK_a absolutos apresentaram erro absoluto médio de 7,2 unidades de pK_a , enquanto os valores de pK_a relativos resultaram em um erro absoluto médio de 4,7 unidades, indicando que a inclusão do fator de correção referente a autoionização do DMSO é fundamental para a predição dos valores dessa propriedade. Nos dois casos, obteve-se um coeficiente de correlação $r^2 = 0,95$. Os resultados indicam que, apesar do alto coeficiente de correlação (o que indica que é possível prever o valor experimental a partir da equação de ajuste com 95 % de poder preditivo), ainda há erros elevados nos valores previstos em relação aos valores experimentais.

As 19 estruturas que apresentaram os menores valores de erro absoluto (Tabela 1), no procedimento direto de cálculo do pK_a relativo, são todas neutras e ácidos fracos, sugerindo que esta combinação do procedimento direto com o modelo de solvatação implícita C-PCM pode fornecer valores em boa concordância com os dados experimentais para este grupo de ácidos. Desse modo, pode ser vantajosa a utilização de uma abordagem computacional menos custosa, como o procedimento direto combinado com o modelo de solvatação implícita, para o grupo de moléculas com essas características.

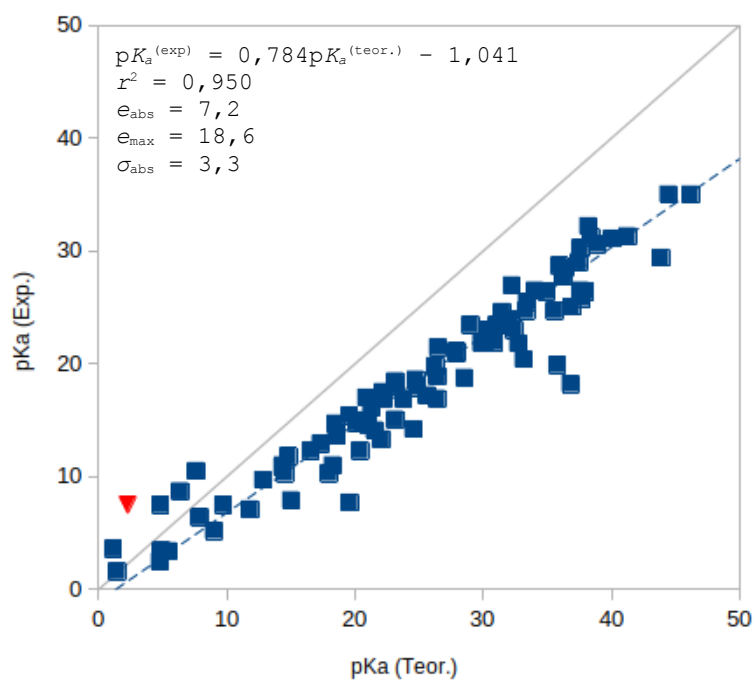


Figura 6: Correlação entre resultados experimentais e teóricos, obtidos com o método M06-2X/6-31+G(d), utilizando o procedimento direto/absoluto para os 105 ácidos em DMSO.

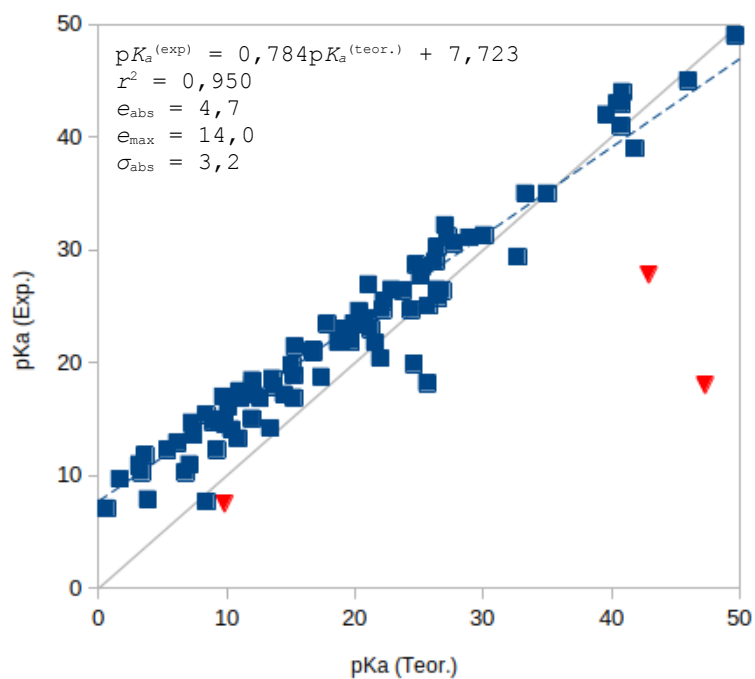


Figura 7: Correlação entre resultados experimentais e teóricos, obtidos com o método M06-2X/6-31+G(d), utilizando o procedimento direto/relativo para os 105 ácidos em DMSO.

Tabela 1: Estruturas com os menores erros absolutos de pK_a relativos (até 1,8 unidades) previstos pelo procedimento direto com o método M06-2X/6-31+G(d) e modelo de solvatação C-PCM.

Cód.	Estrutura	Calc.	Exp.	E_{abs}
012	$\text{CH}_2=\text{CHCH}(\text{H})\text{NO}_2$	8,4	7,7	0,7
030	$\text{CH}_3\text{COCH}_2\text{COOEt}$	13,4	14,2	0,8
042	$\text{NO}_2\text{C}(\text{CH}_3)_2\text{-H}$	15,2	16,9	1,7
053	$\text{CF}_3\text{SO}_2\text{CH}_2\text{-H}$	17,4	18,8	1,4
057	$p\text{-NO}_2\text{-C}_6\text{H}_4\text{CH}_2\text{-H}$	22,0	20,4	1,6
062	4-Acetilpiridina	21,6	21,8	0,2
066	PhCHMeCN	21,2	23,0	1,8
073	$\text{PhCOCH}_2\text{-H}$	24,4	24,7	0,3
074	Ciclobutanona	25,7	25,1	0,7
075	Ciclopentanona	26,5	25,8	0,7
077	Ciclohexanona	26,8	26,4	0,4
079	$\text{CH}_3\text{COCH}_2\text{-H}$	26,4	26,5	0,1
092	$\text{NCCH}_2\text{-H}$	30,1	31,3	1,2
094	$\text{CH}_3\text{SOCH}_2\text{-H}$	35,0	35,0	0,0
095	4-Metilpiridina	33,3	35,0	1,7
097	$\text{NH}_2\text{-H}$	40,8	41,0	0,2
103	$\text{CH}_3\text{SCH}_2\text{-H}$	46,0	45,0	1,0
104	$\text{PhOCH}_2\text{-H}$	49,7	49,0	0,7
105	$\text{CH}_3\text{-H}$	54,3	56,0	1,7

5.2 CÁLCULO DE pK_a COM MODELO HÍBRIDO DE SOLVATAÇÃO

Nesta segunda etapa do estudo foram avaliadas as 10 primeiras estruturas considerando a inclusão de uma molécula explícita de DMSO no cálculo do pK_a absoluto. Esta molécula foi incluída após o procedimento de simulação de dinâmica molecular e análise da trajetória, com a finalidade de identificar a molécula de solvente com maior interação com o soluto.

Para a inclusão de moléculas de solvente no cálculo foi utilizado o programa *Solvate*, que inicialmente gerou uma caixa esférica de simulação contendo em seu centro uma molécula de soluto e um número mínimo de moléculas de DMSO, determinado iterativamente, de modo a atender a dois critérios: envolver completamente a molécula do soluto com ao menos uma molécula de solvente em qualquer direção e atingir a densidade ajustada em $2,0 \text{ g/cm}^3$ (a densidade elevada visa garantir que o soluto permaneça centro da esfera durante toda a simulação). Após a construção da caixa de simulação o *Solvate* gerou o arquivo de entrada para o cálculo de dinâmica molecular BOMD com o programa *ORCA*, utilizando o hamiltoniano semiempírico GFN2-xTB. Foram realizadas simulações em um *ensemble* NVT na temperatura de 298 K com 6000 passos de 0,5 fs, nos quais se considerou os primeiros 1000 passos como de *equilíbrio* e os 5000 passos seguintes de *produção* (ver Apêndice 8.4). Após a obtenção do arquivo de trajetória foi possível analisar as interações mais relevantes considerando o intervalo de produção dividido em 10 partes iguais: em cada uma destas a análise realizada pelo *Solvate* identificou a molécula de solvente que interage mais diretamente com o soluto; cada estrutura identificada foi otimizada a nível GFN2-xTB, com o modelo de solvatação implícita GBSA para descrever o efeito do DMSO, e armazenada em um arquivo para visualização posterior. É importante salientar que nesta etapa nenhum critério para identificar a transferência de próton foi imposta na análise da trajetória, embora as estruturas registradas na análise tenham sido, em sua quase totalidade, compatíveis com a reação sob investigação. As estruturas identificadas compatíveis com o processo de transferência de próton foram, então, otimizadas no mesmo nível de cálculo previamente utilizado (M06-2X/6-31+G(d)), empregando-se o modelo de solvatação implícita C-PCM. As estruturas foram caracterizadas como estruturas de mínima energia a partir do cálculo de frequências. Esta última etapa foi realizada com o programa *Gaussian 09*.

Para os 10 sistemas investigados a otimização de geometria retornou apenas um estado de protonação, correspondente ao estado ácido. Deste modo, considerou-se a utilização do estado otimizado em um cálculo híbrido, e o outro estado (desprotonado) foi descrito a partir dos fragmentos separados. No caso dos *clusters* (mostrados na Tabela 2) a energia livre de Gibbs foi obtida diretamente do arquivo de saída, e no caso dos fragmentos separados as energias de cada fragmento desprotonado foi somada a do DMSO protonado. Os resultados obtidos são apresentados na Tabela 2.

Tabela 2: Valores de pK_a absolutos previstos pelo procedimento direto para as 10 primeiras estruturas com o método M06-2X/6-31+G(d) pelo protocolo de solvatação implícita e híbrida (erros absolutos em parêntesis).

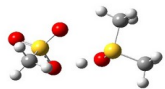
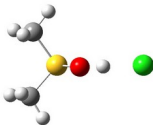
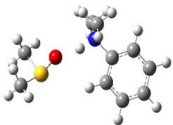
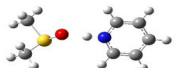
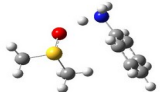
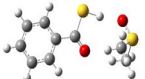
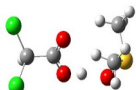
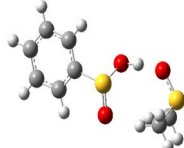
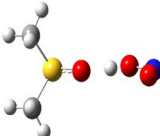
Cód.	Estrutura	Implícito Absoluto	Implícito Relativo	Híbrido	Exp.
001		1,5 (0,1)	-9,7 (11,3)	6,2 (4,6)	1,6
002		-0,4 (2,2)	-11,5 (13,3)	6,1 (4,3)	1,8
003		4,8 (2,3)	-6,4 (8,8)	7,2 (4,8)	2,5
004		4,9 (1,5)	-6,3 (9,7)	7,6 (4,1)	3,5

Tabela 2: Continuação.

Cód.	Estrutura	Implícito Absoluto	Implícito Relativo	Híbrido	Exp.
005		5,5 (2,1)	-5,7 (9,1)	7,3 (3,9)	3,4
006		1,2 (2,4)	-10,0 (13,6)	3,6 (0,0)	3,6
007		9,0 (3,8)	-2,2 (7,4)	5,8 (0,6)	5,2
008		7,9 (1,5)	-3,3 (9,7)	9,8 (3,4)	6,4
009		11,8 (4,7)	0,6 (6,5)	12,5 (5,4)	7,1
010		9,7 (2,2)	-1,4 (8,9)	9,8 (2,3)	7,5
e_{abs}		2,3	9,8	3,3	

Com o método híbrido empregado obteve-se um valor de erro absoluto médio ($e_{\text{abs}} = 3,3$) ligeiramente superior ao obtido para os valores de pK_a absoluto com o método de solvatação implícita ($e_{\text{abs}} = 2,3$), e significativamente inferior aos obtidos para os valores relativos ($e_{\text{abs}} = 9,8$). Na série investigada, apenas para as estruturas 006 e 007 houve uma redução significativa no erro absoluto. As duas moléculas apresentam um anel aromático em suas estruturas e átomos bastante eletronegativos (O, S e N).

CAPÍTULO 6

CONCLUSÃO

6. CONCLUSÃO

As previsões teóricas de pK_a em solventes orgânicos ainda fazem parte de um campo que precisa ser melhor entendido. Inúmeros estudos tratam apenas esse tipo de cálculo usando níveis altos de cálculo com ciclo termodinâmico (procedimento indireto), refinando os cálculos em fase gasosa e calculando as energias de solvatação em separado, o que torna o cálculo computacionalmente custoso pois utiliza muitas etapas para sua determinação [1,7-9].

O presente estudo pretendeu otimizar um protocolo para predição de valores de pK_a para 105 ácidos não correlacionados em DMSO através otimização de um procedimento direto de cálculo. As energias livre de Gibbs dos estados inicial e final foram calculadas no nível da teoria M06-2X com o conjunto de base 6-31+G(d), adotando duas estratégias para a descrição do efeito do solvente, com a utilização do modelo implícito de solvatação C-PCM. Mesmo com todo esforço na descrição da solvatação envolvida nesses sistemas, o poder preditivo do protocolo direto incluindo uma molécula explícita de solvente ainda está distante dos valores experimentais.

O protocolo direto com apenas o método de solvatação implícita apresentou um poder preditivo de 95 % dos valores de pK_a para os 105 ácidos não correlacionados. Os menores valores dos erros absolutos foram verificados para ácidos com valores de pK_a entre 7,7 e 49. Para estas 19 estruturas os erros absolutos obtidos pelo protocolo adotado permitem concluir que o protocolo atingiu a exatidão química (até 1,8 unidades de pK_a). Este grupo de moléculas apresenta em suas estruturas átomos muito eletronegativos, tais como F, O, N e S, e são ácidos fracos (pK_a altos), majoritariamente.

O *Solvate* contribuiu para a identificação da molécula de solvente mais relevante para a descrição do pK_a na abordagem híbrida. No entanto, a inclusão de uma molécula explícita no protocolo direto para o cálculo de pK_a das 10 estruturas selecionadas não foi capaz de reduzir os valores dos erros absolutos para a maioria das estruturas, quando comparadas aos da primeira etapa do estudo, na qual as estruturas foram modeladas apenas com a utilização do modelo implícito de solvatação C-PCM. Para 2 estruturas (006 e 007) a abordagem híbrida possibilitou uma boa concordância com os valores experimentais.

CAPÍTULO 7

PERSPECTIVAS

7. PERSPECTIVAS

- Testar outros parâmetros da solvatação, em especial a topologia da cavitação e o modelo de solvatação implícita;
- Avaliar os arquivos de trajetória para identificar interações soluto-solvente significativas adicionais;
- Concluir a determinação teórica dos valores de pK_a para os demais ácidos com procedimento direto e modelo híbrido de solvatação, avaliando seus efeitos para grupos específicos de ácidos.

Espera-se contribuir para a modelagem de substâncias ácidas neutras, envolvidas em diversos processos químicos e bioquímicos que ocorram em meio solvatado por DMSO, independente da faixa de valores de pK_a .

8. APÊNDICES

8.1 MODELO DE INPUT PARA OTIMIZAÇÃO DE GEOMETRIA E CÁLCULO DE FREQUÊNCIAS

```
#M062X/6-31+G(d) Opt=(Cartesian) Freq=(NoRaman) SCF=XQ
  SCRF=(CPCM,Solvent=DMSO)
¶
Comentário: [Estado Ácido/Fase Solvatada/Otimização+Frequência]
¶
0 1
Cl          0.09466664   -0.00000000   -0.00000000
H          -1.17874583    0.00000000    0.00000000
¶
--link1--
#M062X/6-31+G(d) Opt=(Cartesian) Freq=(NoRaman) SCF=XQ
  SCRF=(CPCM,Solvent=DMSO)
¶
Comentário: [Estado Básico/Fase Solvatada/Otimização+Frequência]
¶
-1 1
Cl          0.00000000    0.00000000    0.00000000
¶
¶
```

Nota: ¶ significa linha em branco.

8.2 ESTRUTURAS INVESTIGADAS (FORMA ÁCIDA)

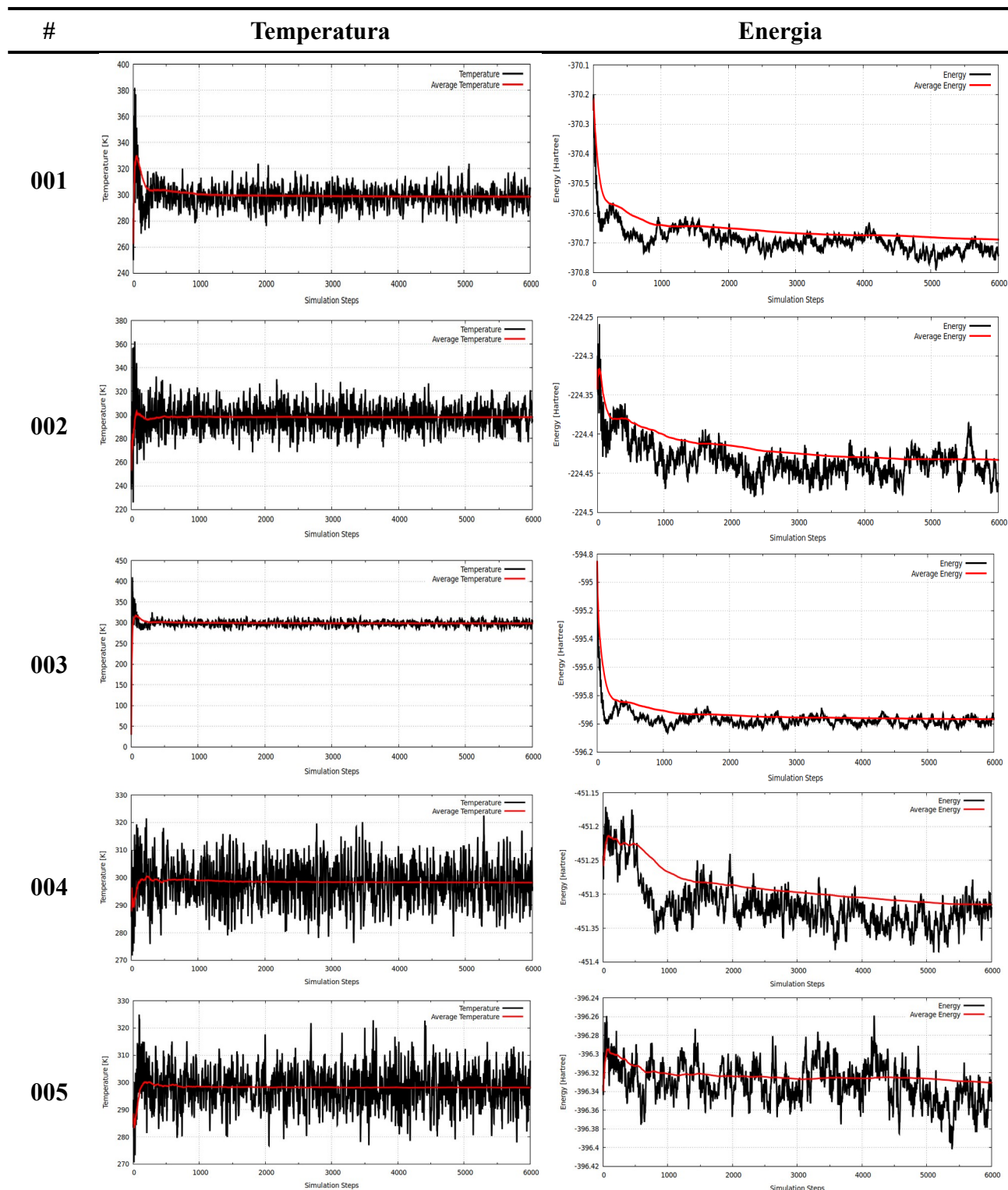
Cód.	Fórmula	Cód.	Fórmula	Cód.	Fórmula
001	CH ₃ SO ₃ -H	041	PhC(=S)NH-H	081	NH ₂ CONH-H
002	H-Cl	042	NO ₂ C(CH ₃) ₂ -H	082	2-Aminopiridina
003	PhN ⁺ HMe ₂	043	CH ₃ CH ₂ CH ₂ CH ₂ S-H	083	Cicloheptanona
004	CF ₃ COO-H	044	CF ₃ CONH-H	084	3-Aminopiridina
005	Piridina-H ⁺	045	NO ₂ CH ₂ -H	085	PhC≡C-H
006	PhNH ₂ -H ⁺	046	CH ₃ SO ₂ NH-H	086	CH ₃ O-H
007	PhCOS-H	047	(CH ₃ CO) ₂ N-H	087	Tiazol
008	Cl ₂ CHCOO-H	048	PhO-H	088	(CH ₃) ₂ CHO-H
009	PhSO ₂ -H	049	Ciclopentadieno	089	PhNH-H
010	ONO-H	050	(CH ₃) ₃ S ⁺	090	CH ₃ SO ₂ CH ₂ -H
011	H ⁺ -NH ₂ CH ₂ COOH	051	CH ₃ C(=S)-NH-H	091	HO-H
012	CH ₂ =CHCH(H)NO ₂	052	Imidazol	092	NCCH ₂ -H
013	H-N ₃	053	CF ₃ SO ₂ CH ₂ -H	093	(CH ₃) ₂ CO-H
014	H ⁺ -NH ₂ CH ₂ COOEt	054	PhCONHNH ₂	094	CH ₃ SOCH ₂ -H
015	CF ₃ SO ₂ NH-H	055	Pirazol	095	4-Metilpiridina
016	1,3-Ciclohexanodiona	056	PhCH ₂ COCH ₃	096	H-Φ(SS)
017	PhS-H	057	<i>p</i> -NO ₂ -C ₆ H ₄ CH ₂ -H	097	NH ₂ -H
018	NH ₃ -H ⁺	058	Indol	098	PhSCH ₂ -H
019	<i>p</i> -NO ₂ -C ₆ H ₄ O-H	059	(H ₂ N) ₂ C=S	099	2-Metiltiofeno
020	PhCOO-H	060	PhNHCOCH ₃	100	2-Metilfurano
021	(NC) ₂ CH-H	061	CH ₃ CONHNH ₂	101	CH ₂ =CH-CH ₂ -H
022	S=ΦN-H	062	4-Acetilpiridina	102	PhCH ₂ -H
023	CH ₃ COO-H	063	PhCH ₂ -CN	103	CH ₃ SCH ₂ -H
024	PhCH ₂ NO ₂	064	Nicotinamida	104	PhOCH ₂ -H
025	H-CN	065	Pirrol	105	CH ₃ -H
026	(CH ₃ CO) ₂ CH-H	066	PhCHMeCN		
027	PhCONHOH	067	PhCONH-H		
028	1,2,3-Triazol	068	HCONH-H		
029	Uracil	069	F ₃ CCH ₂ CONH-H		
030	CH ₃ COCH ₂ COOEt	070	MeOCH ₂ CONH-H		
031	Succinimida	071	H-NHCOOEt		
032	CH ₃ C(=S)NHPh	072	NH ₂ CH ₂ CONH-H		
033	1,2,4-Triazol	073	PhCOCH ₂ -H		
034	O=Φ=O,N-Me	074	Ciclobutanona		
035	O=Φ=N-H	075	Ciclopentanona		
036	H-F	076	CH ₃ CONH-H		
037	CH ₃ CH(COCH ₃) ₂	077	Ciclohexanona		
038	PhCH ₂ S-H	078	2-Piperidinona		
039	PhSONH-H	079	CH ₃ COCH ₂ -H		
040	NC-NH-H	080	4-Aminopiridina		

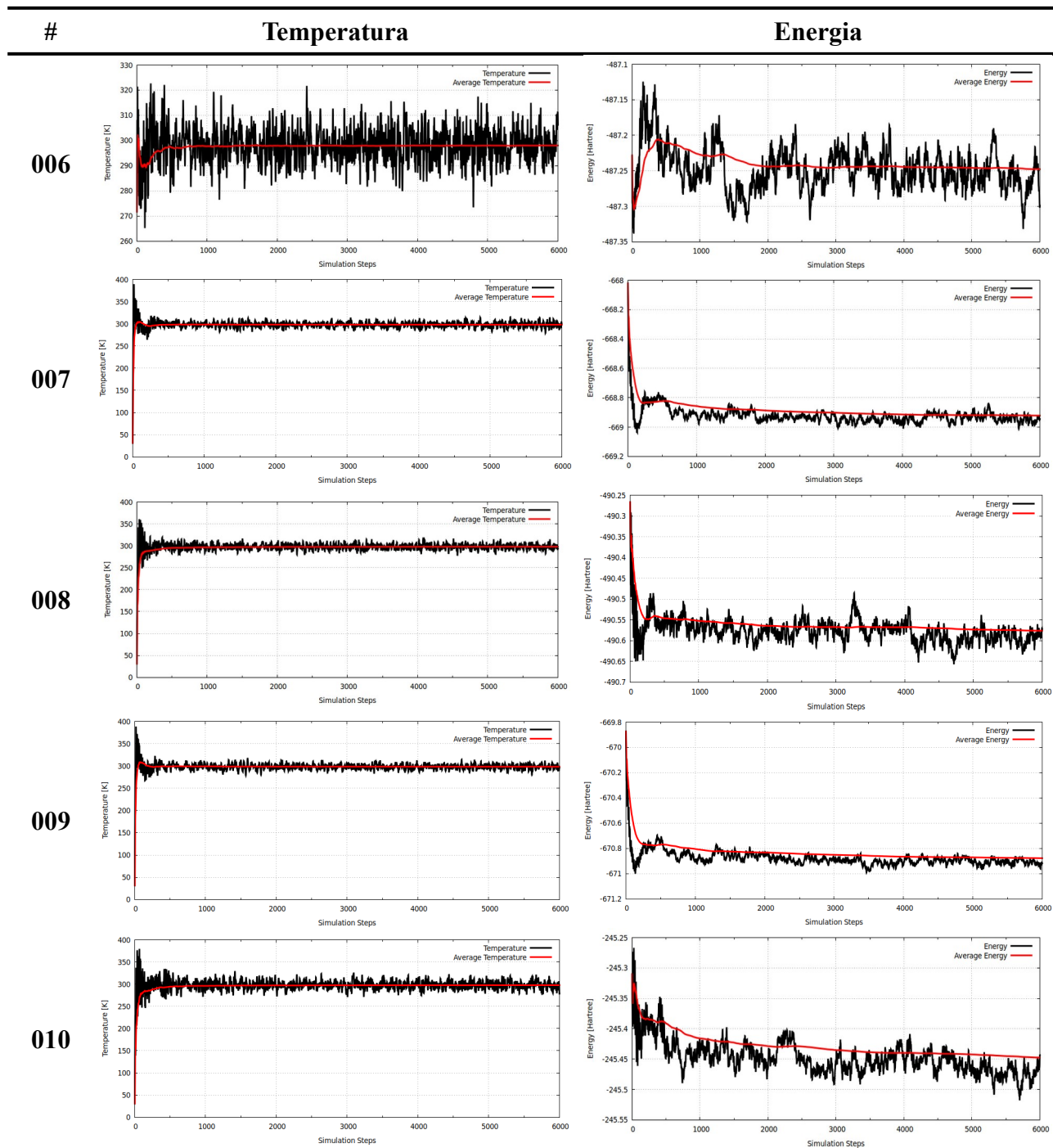
8.3 ENERGIAS LIVRES APÓS OTIMIZAÇÃO COM SOLVATAÇÃO IMPLÍCITA (ÁCIDOS & BASES)

Cód.	Ácido	Base	Cód.	Ácido	Base	Cód.	Ácido	Base
001	-664,152634	-663,737062	041	-723,65519	-723,191116	081	-225,155245	-224,672712
002	-460,771998	-460,360389	042	-323,439583	-322,969741	082	-303,457217	-302,96594
003	-366,338886	-365,91594	043	-556,414531	-555,954784	083	-348,900098	-348,403379
004	-526,631638	-526,208561	044	-506,76372	-506,302797	084	-303,448255	-302,956813
005	-248,547903	-248,123552	045	-244,896745	-244,428365	085	-308,19107	-307,700498
006	-287,814003	-287,398825	046	-644,27531	-643,814609	086	-115,639905	-115,146009
007	-743,526799	-743,094784	047	-361,65659	-361,190162	087	-568,922162	-568,414404
008	-1148,106815	-1147,677263	048	-307,272757	-306,809893	088	-194,18517	-193,691047
009	-780,580076	-780,141957	049	-193,947613	-193,407983	089	-287,398825	-286,901879
010	-205,63017	-205,198184	050	-517,549172	-517,056691	090	-628,224436	-627,724861
011	-284,700081	-284,277141	051	-532,036096	-531,57329	091	-76,38889	-75,892984
012	-322,225794	-321,770787	052	-226,095545	-225,629308	092	-132,683942	-132,181811
013	-164,724407	-164,279304	053	-925,880563	-925,406075	093	-233,45874	-232,963237
014	-363,230648	-362,80434	054	-455,991165	-455,521263	094	-553,037856	-552,525061
015	-941,932698	-941,492349	055	-226,074535	-225,605024	095	-287,397508	-286,888441
016	-383,658336	-383,206862	056	-423,902729	-423,412512	096	-953,393788	-952,866131
017	-630,231607	-629,787564	057	-475,789631	-475,303408	097	-56,511159	-55,985817
018	-56,940138	-56,511159	058	-363,579436	-363,106305	098	-669,499181	-669,045732
019	-511,716288	-511,272673	059	-548,104458	-547,63146	099	-592,137689	-591,614896
020	-420,586104	-420,142426	060	-439,971716	-439,501717	100	-269,165946	-268,641231
021	-224,891445	-224,439289	061	-264,373518	-263,896175	101	-117,785842	-117,260281
022	-646,288745	-645,84415	062	-400,682737	-400,19914	102	-271,345857	-270,820363
023	-228,96559	-228,51717	063	-363,568425	-363,088863	103	-477,878462	-477,341792
024	-475,785166	-475,3284	064	-416,75727	-416,279374	104	-346,526909	-345,982252
025	-93,398551	-92,945406	065	-210,031487	-209,553087	105	-40,458233	-39,903556
026	-345,569548	-345,109145	066	-402,836574	-402,35368	000	-553,038369	-552,525062
027	-475,839526	-475,386637	067	-400,709401	-400,228532			
028	-242,112692	-241,659965	068	-169,812949	-169,332957			
029	-414,628682	-414,169412	069	-452,602703	-452,127217			
030	-460,046431	-459,580596	070	-323,534345	-323,052057			
031	-360,484164	-360,026037	071	-323,559703	-323,078812			
032	-762,916982	-762,464356	072	-264,403634	-263,918708			
033	-242,139484	-241,683343	073	-384,635928	-384,146163			
034	-415,807112	-415,349302	074	-231,069974	-230,577335			
035	-323,330437	-322,873108	075	-270,367133	-269,872736			
036	-100,416444	-99,96027	076	-209,090513	-208,605443			
037	-384,836648	-384,373861	077	-309,63801	-309,143026			
038	-669,500727	-669,045731	078	-325,709015	-325,220663			
039	-835,890361	-835,431686	079	-193,015715	-192,52155			
040	-148,728463	-148,267654	080	-303,453924	-302,967459			

Nota: 1 Hartree = 2625,5 kJ.mol⁻¹.

8.4 DADOS DAS SIMULAÇÕES MOLECULARES





9. REFERÊNCIAS

- [1] Fu, Yao et al.; *First-Principle Predictions of Absolute pKa's of Organic Acids in Dimethyl Sulfoxide Solution*. **J. Am. Chem. Soc.**, v. 126, p. 814-822 (2004).
- [2] Toth, A. M.; Liptak, M. D.; Phillips, D. L.; Shields, G. C.; *Accurate relative pKa calculations for carboxylic acids using complete basis set and Gaussian-n models combined with continuum solvation methods*. **J. Chem. Phys.**, v. 114(10), p. 4595 (2001).
- [3] Liptak, M. D.; Shields, G. C.; *Accurate pKa Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with CPCM Continuum Solvation Methods*. **J. Am. Chem. Soc.**, v. 123, p. 7314-7319 (2001).
- [4] Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields, G. C.; *Absolute pKa Determinations for Substituted Phenols*. **J. Am. Chem. Soc.**, v. 124, p. 6421 (2002).
- [5] Brown, T. N.; Mora-Diez, N.; *Computacional Determination of Aqueous pKa Values of Protonated Benzimidazoles (Part 2)*. **J. Phys. Chem. B**, v. 110(41), p. 20546–20554 (2006).
- [6] Namazian, M. et al.; *Møller–Plesset perturbation theory calculations of the pKa values for a range of carboxylic acids*. **J. Mol. Struct.: THEOCHEM**, v. 758, p. 275-278 (2006).
- [7] Ho, J.; Coote, M. L.; *First-Principles Prediction of the pKas of Anti-inflammatory Oxicams*. **J. Phys. Chem. A**, v. 114, p. 11992-12003 (2010).
- [8] Yue, Zeng et al.; *Estimate of pKa Values of the Carboxylic Acids in Aqueous Solution with the Density Functional Theory*. **Chin. J. Chem.**, v. 28, p. 727-733 (2010).
- [9] Sutton, C. C. R.; *First Principles pKa Calculations on Carboxylic Acids Using the SMD Solvation Model: Effect of Thermodynamic Cycle, Model Chemistry, and Explicit Solvent Molecules*. **J. Phys. Chem. B**, v. 116, p. 11999-12006 (2012).
- [10] Zhang, J., Zhang, H., Wu, T., Wang, Q., Spoel, D.; *Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents*. **J. Chem. Theory Comput.**, v. 13, p. 1034-1043 (2017).
- [11] Thapa, B., Raghavachari, K.; *Accurate pKa Evaluations for Complex Bio-Organic Molecules in Aqueous Media*. **J. Chem. Theor. Comput.**, v. 15(11), p. 6025-6035 (2019).
- [12] Geremia, K. L., Seybold, P. G.; *Computational Estimation of the Acidities of Purines and Indoles*. **J. Mol. Model.**, v. 25, p. 12 (2019).
- [13] Mirzaei, S., Ivanov, M. V., Timerghazin, Q. K.; *Improving Performance of the SMD Solvation Model: Bondi Radii Improve Predicted Aqueous Solvation Free Energies of Ions and pKa Values of Thiols*. **J. Phys. Chem. A**, v. 123(44), p. 9498-9504 (2019).
- [14] Ristić, M. M., Petković, M., Milovanović, B., Belić, J., Etinski, M.; *New hybrid cluster-continuum model for pKa values calculations: Case study of neurotransmitters' amino group acidity*. **Chem. Phys.**, v. 516, p. 55-62 (2019).

- [15] Becke, A. D.; *Density-Functional Thermochemistry. III. The Role of Exact Exchange*. **J. Chem. Phys.**, v. 98, p. 5648-5652 (1993).
- [16] Zhao, Y.; Truhlar, D. G.; *The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function*. **Theor. Chem. Acc.**, v. 120, p. 215-241 (2008).
- [17] Dutra, F. R.; Silva, C. S.; Custodio, R.; *On the Accuracy of the Direct Method to Calculate pKa from Electronic Structure Calculations*. **J. Phys. Chem. A**, v. 125(1), p. 65-73 (2021).
- [18] Miertuš, S.; Scrocco, E.; Tomasi, J.; *Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects*. **Chem. Phys.**, v. 55(1), p. 117-129 (1981).
- [19] Tomasi, J.; Mennucci, B.; Cammi, R.; *Quantum Mechanical Continuum Solvation Models*. **Chem. Rev.**, v. 105, p. 2999 (2005).
- [20] Pliego, J. R., Jr.; Riveros, J. M.; *The Cluster-Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species*. **J. Phys. Chem. A**, v. 105, p. 7241 (2001).
- [21] Pliego, J. R., Jr.; Riveros, J. M.; *Calculation of pKa Using the Cluster-Continuum Model*. **J. Phys. Chem. A**, v. 106, p. 7434 (2002).
- [22] Pliego Jr., J. R.; *Modelos Contínuos do Solvente: Fundamentos*. **Quím. Nova**, v. 29(3), p. 535-542 (2006).
- [23] McMurry, J.; *Química Orgânica*. Ed. 7ª, São Paulo (2011).
- [24] Atkins, P. W.; Paula, J. de; *Físico-Química*. Ed. 9ª, Rio de Janeiro (2012).
- [25] Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; *Use of Solution-Phase Vibrational Frequencies in Continuum Models for the Free Energy of Solvation*. **J. Phys. Chem. B**, v. 115, p. 14556-14562 (2011).
- [26] Ho, J.; *Predicting pKa in Implicit Solvents: Current Status and Future Directions*. **Aust. J. Chem.**, v. 67(10), p. 1441-1460 (2014).
- [27] Ho, J.; *Are thermodynamic cycles necessary for continuum solvent calculation of pKas and reduction potentials?*. **Phys. Chem. Chem. Phys.**, v. 17(4), p. 2859-2868 (2015).
- [28] Abramson, R. A.; Baldridge, K. K.; *Defined-Sector Explicit Solvent in Continuum Cluster Model for Computational Prediction of pKa: Consideration of Secondary Functionality and Higher Degree of Solvation*. **J. Chem. Theory Comput.**, v. 9(2), p. 1027-1035 (2013).
- [29] Alongi, K. S.; Shields, G. C.; *Theoretical Calculations of Acid Dissociation Constants: A Review Article*. **Annu. Rep. Comput. Chem.**, v. 6, p. 113-138 (2010).
- [30] Zevatskii, Y. E.; Samoilov, D. V.; *Modern Methods for Estimation of Ionization Constants of Organic Compounds in Solution*. **Russ. J. Org. Chem.**, v. 47(10), p. 1445-1467 (2011).

- [31] Thapa, B.; Schlegel, H. B.; *DFT calculation of pK_a 's of Tiois in Aqueous Solution using Explicit Water Molecules and Polarizable Continuum Model*. **J. Phys. Chem. A**, v. 120(28), p. 5726-5735 (2016).
- [32] Cramer, C. J.; *Essentials of Computational Chemistry: Theories and Models*. Ed. 2^a, Minnesota - USA (2004).
- [33] Jensen, F.; *Introduction to Computational Chemistry*. Ed. 1^a, Odense - Denmark (2007).
- [34] Dirac, P. A. M.; *Note on Exchange Phenomena in the Thomas Atom*. **Math. Proc. Camb. Phil. Soc.**, v. 26(3), p. 376 (1930).
- [35] Hohenberg, P.; Kohn, W.; *Inhomogeneous Electron Gas*. **Phys. Rev. B**, v. 136(3), p. 864-871 (1964).
- [36] Parr, R. G.; Yang, W.; *Density Functional Theory of Atoms and Molecules*. Ed. , Oxford (1989).
- [37] Dreizler, R. M.; Gross, E. K. U.; *Density Functional Theory. An Approach to the Quantum Many-Body Problem*. Ed. , Berlin (1990).
- [38] Zhao, Y.; Truhlar, D. G.; *A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions*. **J. Chem. Phys.**, v. 125(19), p. 194101 (2006).
- [39] Zhao, Y.; Truhlar, D. G.; *Density functional for spectroscopy: no long-range selfinteraction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states*. **J. Chem. Phys.**, v. 125(19), p. 194101 (2006).
- [40] Zhao, Y.; Truhlar, D. G.; *The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals*. **Theor. Chem. Acc.**, v. 120, p. 215-241 (2008).
- [41] Silva, A. F.; *Cálculo de Potenciais de Redução em Meio Aprótico (DMF) de Adutos da Reacao de Morita-Baylis-Hillman com Potencialidades Biologicas anti-Leishmania*. *Dissertação - Universidade Federal da Paraíba. João Pessoa/PB*, p. 53 (2015).
- [42] Ditchfield, R.; Hehre, W. J.; Pople, J. A.; *Self-Consistent Molecular-Orbital Organic Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules*. **J. Chem. Phys.**, v. 54(2), p. 724 (1971).
- [43] Silva, A. F. da, Filho, A. J. da S., Vasconcellos, M., Santana, O. L. de.; *One-Electron Reduction Potentials: Calibration of Theoretical Protocols for Morita-Baylis-Hillman Nitroaromatic Compounds in Aprotic Media*. **Molecules**, v. 23(9), p. 2129 (2018).
- [44] Decherchi, S; Masetti, M; Vyalov, I; Rocchia, W.; *Implicit Solvent Methods for Free Energy Estimation*. **Eur. J. Med. Chem**, v. 91, p. 27-42 (2015).
- [45] Klamt, A.; *Cosmo-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*. Ed. , Netherlands (2005).

- [46] Pascual-ahuir, J. L.; Silla, E.; Tuñón, I.; *The solvent-Excluding Surface (SES) as a descriptor of ionic channels: Gramicidin-A*. **J. Mol. Struct. (Theochem)**, v. 426(1-3), p. 331-338 (1998).
- [47] Quan, C.; Stamm, B.; Maday, Y.; *A domain decomposition method for the polarizable continuum model based on the solvent excluded surface*. **Math. Models Methods Appl. Sci.**, v. 28(7), p. 1233-1266 (2018).
- [48] Bentley, J.; *Behavior of Electron Density Functions in Molecular Interactions*. **J. Phys.Chem. A**, v. 102(29), p. 6043-6051 (1998).
- [49] Bentley, J.; *Electron Density as a Descriptor of Thermal Molecular Size*. **J. Phys. Chem. A**, v. 104(42), p. 9630-9635 (2000).
- [50] Miertus, S.; Tomasi, J.; *Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes*. **Chem. Phys.**, v. 65, p. 239-245 (1982).
- [51] Miertus, S.; Scrocco, E.; Tomasi, J.; *Electrostatic interaction of a solute with a continuum. A direct utilizaion of ab initio molecular potentials for the prevision of solvent effects*. **Chem. Phys.**, v. 55, p. 117-129 (1981).
- [52] Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J.; *Ab initio study of solvated molecules: a new implementation of the polarizable continuum model*. **Chem. Phys. Lett.**, v. 255, p. 327-335 (1996).
- [53] Tomasi, J.; Cammi, R.; Mennucci, B.; *Medium effects on the properties of chemical systems: An overview of recent formulations in the polarizable continuum model (PCM)*. **Int. J. Quant. Chem.**, v. 75, p. 783-803 (1999).
- [54] Tomasi, J.; Cammi, R.; Mennucci, B.; Cappelì, C.; Corni, S.; *Molecular properties in solution described with a continuum solvation model*. **Phys. Chem. Chem. Phys.**, v. 4, p. 5697-5712 (2002).
- [55] Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J.; *Solvent Effects. 5. Influence of Cavity Shape, Truncation of Electrostatics, and Electron Correlation on ab Initio Reaction Field Calculations*. **J. Phys. Chem.**, v. 100, p. 16098-16104 (1996).
- [56] Wiberg, K. B.; Keith, T. A.; Frisch, M. J.; Murcko, M.; *Solvent Effects on 1,2-Dihaloethane Gauche/Trans Ratios*. **J. Phys. Chem.**, v. 99, p. 9072-9079 (1995).
- [57] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; *Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions*. **J. Phys. Chem.**, v. 113(18), p. 6378-6396 (2009).
- [58] Jorgensen, W. L.; *Free energy calculations: a breakthrough for modeling organic chemistry in solution*. **Acc. Chem. Res.**, v. 22, p. 184-189 (1989).
- [59] Kollman, P. A.; *Free energy calculations: Applications to chemical and biochemical phenomena*. **Chem. Rev.**, v. 93, p. 2395-2417 (1993).
- [60] Pliego Jr., J. R.; de Almeida, W. B.; *A theoretical abinitio and Monte Carlo simulation study of the pyridine+CCl₂ reaction kinetics in the gas phase and in carbon tetrachloride solution*

- using canonical flexible transition state theory. **Phys. Chem. Chem. Phys.**, v. 1, p. 1031 (1999).
- [61] Pliego Jr., J. R.; Riveros, J. M.; *On the Calculation of the Absolute Solvation Free Energy of Ionic Species: Application of the Extrapolation Method to the Hydroxide Ion in Aqueous Solution.* **J. Phys. Chem.**, v. 104, p. 5155-5160 (2000).
- [62] Rocha, W. R.; Martins, V. M.; Coutinho, K.; Canuto, S.; *Solvent effects on the electronic absorption spectrum of formamide studied by a sequential Monte Carlo/quantum mechanical approach.* **Theor. Chem. Acc.**, v. 108, p. 31-37 (2002).
- [63] Muddana, H. S., Sapra, N. V., Fenley, A. T., Gilson, M. K.; *The SAMPL4 hydration challenge: evaluation of partial charge sets with explicit-water molecular dynamics simulations.* **J. Comput. Aided Mol. Des.**, v. 28, p. 277287 (2014).
- [64] Gerhard, K., Frank, C. P. Y. M., Bernard, R. B.; *Predicting hydration free energies with a hybrid QM/MM approach: an evaluation of implicit and explicit solvation models in SAMPL4.* **J. Comput. Aided Mol. Des.**, v. 28, p. 245257 (2014).
- [65] Mobley, D. L., Bayly, C. I., Cooper, M. D., Shirts, M. R., and Dill, K. A.; *Correction to Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations.* **J. Chem. Theory Comput.**, v. 11(3), p. 1347 (2015).
- [66] Georg, H. C.; Canuto, S.; *Métodos de Química Teórica e Modelagem Molecular: Métodos Híbridos para Modelagem do Ambiente Molecular.* Ed. 1ª, São Paulo (2007).
- [67] Martinez, L., Borin, I. A., Skaf, M. S.; *Métodos de Química Teórica e Modelagem Molecular: Fundamentos de Simulação por Dinâmica Molecular.* Ed. 1ª, São Paulo (2007).
- [68] Grimme, S.; *A General Quantum Mechanically Derived Force Field (QMDF) for Molecules and Condensed Phase Simulations.* **J. Chem. Theory Comput.**, v. 10(10), p. 4497-4514 (2014).
- [69] Silveira, L. G. da, Jacobs, M., Prampolini, G., Livotto, P. R., Cacelli, I.; *Development and Validation of Quantum Mechanically Derived Force-Fields: Thermodynamic, Structural, and Vibrational Properties of Aromatic Heterocycles.* **J. Chem. Theory Comput.**, v. 14(9), p. 4884-4900 (2018).
- [70] Helgaker, T.; Uggerud, E.; Jensen, H. J. A.; *Integration of the Classical Equations of Motion on ab initio Molecular-Potential Energy Surfaces Using Gradients and Hessians – Application to Translational Energy-Release Upon Fragmentation.* **Chem. Phys. Lett.**, v. 173, p. 145-150 (1990).
- [71] Grimme, S.; Bannwarth, C.; Shushkov, P.; *A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z =1–86).* **J. Comput. Chem.**, v. 13, p. 1989-2009 (2017).
- [72] Bannwarth, C.; Ehlert, S.; Grimme, S.; *GFN2-xTB: An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and*

- Density-Dependent Dispersion Contributions*. **J. Chem. Theory Comput.**, v. 15(3), p. 1652-1671 (2019).
- [73] Silva, D. G.; *Potenciais de Redução em Meio Aquoso: Aplicação e Desenvolvimento de Protocolos de Cálculo Baseados em Modelos Implícitos e Métodos Híbridos*. Dissertação - Universidade Federal da Paraíba. , p. 97 (2020).
- [74] Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M.; *Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients*. **Nature**, v. 194, p. 178-180 (1962).
- [75] Mansouri, K.; Grulke, C. M.; Richard, A. M.; et al.; *An automated curation procedure for addressing chemical and inconsistencies in public datasets used in QSAR modelling*. **SAR QSAR Environ. Res.**, v. 27, p. 911-937 (2016).
- [76] Mansouri, K.; Cariello, N. F.; Korotcov, A. et al.; *Modelos QSAR de código aberto para previsão de pKa usando abordagens de aprendizado de máquina*. **J. Cheminform**, v. 11, p. 60 (2019).
- [77] Yang, Q.; Li, Y.; Yang, J.; Liu, Y.; Zhang, L.; Luo, S.; Cheng, J.; *Holistic Prediction of the pKa in Diverse Solvents Based on a Machine-Learning Approach*. **GDCh**, v. 132, p. 19444-19453 (2020).
- [78] Fourches, D.; Muratov, E.; Tropsha, A.; *Trust, but verify II: a practical guide to chemogenomics data curation*. **J. Chem. Inf. Model.**, v. 56, p. 1243-1252 (2016).
- [79] Bordwell, F. G.; *Equilibrium acidities in dimethyl sulfoxide solution*. **Acc. Chem. Res.**, v. 21, p. 456-463 (1988).
- [80] Matthews, W. S., Bares, J. E., Bartmess, J. E., Bordwell, F.G., Cornforth, F. J., Drucker, G. E., Margolin, Z., McCallum, R. J., McCollum, G. J., Vanier, N. R.; *Equilibrium acidities of carbon acids. VI. Establishment of an absolute scale of acidities in dimethyl sulfoxide solution*. **J. Am. Chem. Soc.**, v. 97(24), p. 7006-7014 (1975).
- [81] Grimme, S.; *Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations*. **J. Chem. Theory Comput.**, v. 15, p. 2847-2862 (2019).
- [82] Pracht, P.; Bohle, F.; Grimme, S.; *Automated Exploration of the low-energy Chemical Space with fast Quantum Chemical Methods*. **Phys. Chem. Chem. Phys.**, v. 22(14), p. 7169-7192 (2020).
- [83] Frisch, M. J.; et. al; *Gaussian 09, Revision D.01* (2013).
- [84] Frisch, M. J.; et. al; *Manual do Gaussian 09*. Ed. 2009, Gaussian, Inc., Wallingford CT (2009).
- [85] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V.; *Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model*. **J. Comput. Chem.**, v. 24, p. 669-81 (2003).

- [86] Kelly, C. P., Cramer, C. J., & Truhlar, D. G.; *Single-ion solvation free energies and the normal hydrogen electrode potential in methanol, acetonitrile, and dimethyl sulfoxide*. **J. Phys. Chem. B**, v. 111(2), p. 408-422 (2007).
- [87] Santana, O. L.; *Solvate 2020* (2019).
- [88] Neese, F.; Wennmohs, F.; *ORCA 4.2.0: An Ab-Initio, DFT and Semiempirical SCF-MO Package* (2019).