

UFRRJ
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

DISSERTAÇÃO

**Desenvolvimento de um Modelo Empírico de Predição da
Atividade de Inibidores da Urease utilizando o Método
Semi-Empírico PM6**

SHEISI FONSECA LEITE DA SILVA ROCHA
2014



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**Desenvolvimento de um Modelo Empírico de Predição da
Atividade de Inibidores da Urease utilizando o Método
Semi-Empírico PM6**

SHEISI FONSECA LEITE DA SILVA ROCHA

Sob a Orientação do Professor
Dr. Carlos Mauricio Rabello de Sant'Anna

Dissertação submetida como requisito parcial para a obtenção do grau de **Mestre em Ciências**, no Curso de Pós-Graduação em Química, Área de Concentração em Química Orgânica.

Seropédica, RJ
Abril de 2014

547.47

R672d

T

Rocha, Sheisi Fonseca Leite da
Silva, 1987-

Desenvolvimento de um modelo
empírico de predição da atividade de
inibidores da uréase utilizando o
método semi-empírico PM6 /Sheisi
Fonseca Leite da Silva. - 2014.

88 f. : il.

Orientador: Carlos Mauricio
Rabello de Sant'Anna.

Dissertação (Mestrado) -
Universidade Federal Rural do Rio
de Janeiro, Curso de Pós-Graduação
em Química.

Bibliografia: f. 65-73.

1. Compostos organofosforados-
Teses. 2. Urease - Teses. 3.
Química orgânica -Teses. I.
Sant'Anna, Carlos Mauricio Rabello
de, 1965-. II. Universidade Federal
Rural do Rio de Janeiro. Curso de
Pós-Graduação em Química. III.
Título.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

SHEISI FONSECA LEITE DA SILVA ROCHA

Dissertação submetida como requisito parcial para a obtenção do grau de **Mestre em Ciências** no Curso de Pós Graduação em Química, Área de Concentração em Química Orgânica.

Dissertação aprovada em 15 de Abril de 2014

Prof. Dr. Carlos Mauricio Rabello de Sant'Anna, UFRRJ
(Orientador)

Prof. Dr. Glauco Favilla Bauerfeldt, UFRRJ

Prof. Dr. Sérgio de Paula Machado, UFRJ

“Seja bendito o nome de Deus para todo o
sempre, porque Dele é a sabedoria e a força.”
(Bíblia Sagrada, Daniel 2:20)

AGRADECIMENTOS

A Deus, por Sua misericórdia, amor e graça derramadas sobre minha vida, por ter me sustentado todos os dias e me cercado de bênçãos e alegrias, pelo presente mais precioso: Jesus, sem O qual a minha vida não teria sentido.

Ao meu esposo, José Geraldo, por estar sempre ao meu lado, nos bons e maus momentos, por toda ajuda, pelo grande incentivo, apoio e amizade.

Ao Gabriel, meu filhinho lindo, por tornar minha vida muito, muito mais feliz.

À minha mãe, pelo amor incondicional, por estar sempre pronta a me ajudar e a cuidar de mim, não importando o quanto eu já tenha crescido.

À minha irmã Shênia, pelo grande carinho que tem por mim e minha família, pelo amor posto em prática, por ter me socorrido muitas vezes e estar sempre disposta a isso.

Ao meu pai por sempre torcer por mim e querer o meu bem.

Aos meus cunhados Danielle e Rodrigo e aos meus sobrinhos, Aleph, Rebecca e Kaleb pela amizade, pelos momentos de descontração e de incentivo, pelo carinho que têm por mim e por estarem sempre dispostos a me ajudar.

Aos meus irmãos da Célula Peniel, pela amizade, força e comunhão.

Aos meus discipuladores Washington e Michele e aos meus pastores Elieser e Rosimere por todo cuidado, aconselhamento e orações.

Ao professor Carlos Maurício, por todo conhecimento transmitido, pela compreensão, paciência e orientação segura.

Aos colegas do LabMol/UFRRJ por todo apoio prestado durante esta jornada.

Aos docentes deste Departamento, pelo incentivo e por todo conhecimento transmitido.

Ao CNPq pela concessão da bolsa de estudos.

Ao PPGQ/UFRRJ por proporcionar este acontecimento.

RESUMO

ROCHA, F. L. S. Sheisi. **Desenvolvimento de um Modelo Empírico de Predição da Atividade de Inibidores da Urease utilizando o Método Semi-Empírico PM6**. 2014. 73p. Dissertação (Mestrado em Química, Química Orgânica). Instituto de Ciências Exatas, Departamento de Química, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2014.

A urease é uma enzima importante para as pesquisas relacionadas com a agricultura, meio ambiente e medicina. Ela catalisa a reação de hidrólise da uréia para formar amônia e carbamato, o qual se decompõe espontaneamente, produzindo uma segunda molécula de amônia e dióxido de carbono, provocando um significativo aumento do pH da solução. Com o objetivo de desenvolver modelos de predição da atividade de inibidores da urease, estudou-se inicialmente a multiplicidade de spin da enzima, que contém dois íons Ni(II), e o estado de protonação do oxigênio localizado entre estes íons. Os resultados indicaram que o sistema é melhor representado pelo estado tripleto ou quinteto e o oxigênio localizado entre os íons de níquel provavelmente é um íon hidroxila. A partir destes resultados, a construção dos modelos se baseou em propostas da literatura sobre o uso de ciclos termodinâmicos para se calcular a energia livre de interação entre ligantes e enzimas. No presente estudo, foram combinados termos referentes à entalpia de interação entre o inibidor e a enzima, a energia livre de Gibbs necessária para o inibidor passar do meio aquoso para o interior da enzima e as perdas entrópicas devido a restrições rotacionais após a interação do mesmo com a enzima para se obter funções de correlação com constantes inibitórias (K_i) obtidas experimentalmente. A quantificação destes parâmetros para alguns derivados do ácido fosfínico da literatura nos possibilitou o desenvolvimento de um modelo para determinação da atividade com boa correlação com dados experimentais ($r=0,92$). Este modelo foi utilizado na predição da atividade relativa de novas dialquilfosforilidrazonas, sintetizadas pelo grupo de síntese de organofosforados da UFRRJ. Foi possível identificar quais compostos são os mais promissores da série proposta e quais fatores devem ser alterados para otimizar o perfil de inibição da urease.

Palavras-chave: Urease. Organofosforados. Modelo de energia livre. Método semi-empírico.

ABSTRACT

Urease is an important enzyme for the research in agriculture, environment and medicine. This enzyme catalyzes the hydrolysis of urea to ammonia and carbamate, which decomposes spontaneously, yielding a second molecule of ammonia, causing a significant increase of pH solution. In order to develop theoretical models for the prediction of activities of urease inhibitors, we initially studied the enzyme's spin multiplicity, which contains two Ni(II) ions, and the state of protonation of the oxygen located between the nickel ions. The results indicate that the system is best represented by the triplet or quintet state and the oxygen atom located between the nickel ions, probably is a hydroxyl ion. Based on these results, the construction of the models was based on literature proposals about the use of thermodynamic cycles for the calculation of the free energy of binding between ligands and enzymes. In the present work, parameters such as the interaction enthalpy, the Gibbs free energy required for the inhibitor to go from the aqueous phase to the interior of the enzyme and the entropic losses associated to the freezing of bonds after the binding of the inhibitors to the enzyme were used to develop correlations with the measured experimental K_i values. The quantification of these parameters for some phosphinic acids derivatives from the literature allowed us to obtain a good empirical model for the correlation between experimental activity data and the theoretical parameters ($r=0.92$). The model was employed for the prediction of the relative activity of a series of new proposed compounds by the organophosphorous synthesis group of UFRRJ. It was possible to identify which compounds are the most promising and which are the main factors that should be modified in order to optimize the urease inhibition profile by these compounds.

Keywords: Urease. Organophosphorus compounds. Free energy models. Semi-empirical method.

ÍNDICE DE TABELAS

Tabela 1. Valores de K_i (μM) e $\log K_i$ para os compostos sintetizados por Vassiliou e colaboradores em 2008 (série 1) e em 2010 (série 2).	39
Tabela 2. Distâncias em Å entre os átomos doadores de elétrons e os átomos de Ni(II) no sítio ativo da urease 1EJW.	44
Tabela 3. Distâncias em Å entre os átomos doadores de elétrons e os átomos de Ni(II) no sítio ativo da urease 1FJW.	44
Tabela 4. Valores da raiz do desvio médio quadrático (RMSD) entre átomos doadores de elétrons e os cátions Ni^{+2}	45
Tabela 5. Distâncias em Å entre os átomos doadores de elétrons e os átomos de Ni(II) no sítio ativo da urease 1FWJ.	46
Tabela 6. Valores da raiz do desvio médio quadrático (RMSD) entre átomos doadores de elétrons e os cátions Ni^{+2}	46
Tabela 7. Valores da raiz do desvio médio quadrático (RMSD) para cada função testada no procedimento de <i>redocking</i> do ligante da estrutura 4UBP do PDB.	47
Tabela 8. Logaritmo dos valores de K_i experimentais, energia livre de Gibbs necessária para o transporte do inibidor da fase aquosa para o interior da enzima (ΔG_I , kcal/mol); número de ligações rotacionáveis congeladas durante a interação (N_{LR}), e entalpias de interação calculadas assumindo estado quineto e constante dielétrica 4,0 (ΔH_{int} , kcal/mol), para os inibidores da série 1.	49
Tabela 9. Contribuição de cada termo variável da equação 16, $\log K_i$ medido experimentalmente, $\log K_i$ calculado pela equação 16, e a diferença entre os valores de $\log K_i$ experimentais e calculados.	51
Tabela 10. Logaritmo dos valores de K_i experimentais, energia livre de Gibbs necessária para o transporte do inibidor da fase aquosa para o interior da enzima (ΔG_I , kcal/mol); número de ligações rotacionáveis congeladas durante a interação (N_{LR}), e entalpias de interação calculadas assumindo estado quinteto e constante dielétrica 4,0 (ΔH_{int} , kcal/mol), para os inibidores da série 2.	53
Tabela 11. Contribuição de cada termo variável da equação 17, $\log K_i$ medido experimentalmente, $\log K_i$ calculado pela equação 17, e a diferença entre os valores de $\log K_i$ experimentais e calculados.	55
Tabela 12. Energia livre de Gibbs necessária para o transporte dos inibidores da série 3 (ΔG_I , kcal/mol), número de ligações rotacionáveis congeladas (N_{LR}) e entalpia de interação inibidor/urease (ΔH_{int} , kcal/mol).....	59

Tabela 13. Contribuição de cada termo variável da equação 16 para os compostos da série 3.	62
Tabela 14. Contribuição de cada termo variável da equação 16 para a estrutura 34. ...	64

ÍNDICE DE FIGURAS

Figura 1. Reações de decomposição da uréia.	04
Figura 2. Sítio ativo da urease de <i>K. aerogenes</i> (código 1FWJ do PDB) determinada por cristalografia de raio-X	06
Figura 3. Dois mecanismos de reação propostos	07
Figura 4. Representação esquemática de dois mecanismos propostos para a urease, A e B. .	08
Figura 5. A urease quebra a uréia presente no suco gástrico em amônia que neutraliza o ambiente ácido do estômago.	10
Figura 6. Alguns inibidores de urease.	11
Figura 7. Isatina (1H-indol-2,3-diona).	12
Figura 8. Forma isomérica orto-quinonoídica e para-quinonoídica.	13
Figura 9. Número de publicações envolvendo <i>docking</i> , baseada em simples pesquisa da literatura.	24
Figura 10. Programas de <i>Docking</i> citados nos títulos de artigos ou resumos em 2010–2011.	28
Figura 11. Diagrama simplificado da distribuição dos elétrons nos orbitais d dos íons Ni ⁺² nas multiplicidades estudadas.	32
Figura 12. Série 1: Ácidos P-metil fosfínicos sintetizados por Vassiliou e colaboradores em 2008.	33
Figura 13. Série 2: Ácidos P-hidroximetil fosfínicos, sintetizados por Vassiliou e colaboradores em 2010.	34
Figura 14. Série 3: Compostos sintetizados pelo grupo do Dr. João Batista Neves da Costa.	35
Figura 15. Distâncias escolhidas entre os cátions Ni ⁺² e alguns átomos doadores de elétrons no sítio ativo da urease 1EJW.	43
Figura 16. Sobreposição das estruturas do ácido acetohidroxâmico co-cristalizado na urease (código 4UBP do PDB) e atracado com a função <i>ASP</i> , função <i>ChemPLP</i> , função <i>Chemscore</i> e função <i>Goldscore</i>	47
Figura 17. Gráfico de Lineweaver-Burk.	50
Figura 18. Gráfico de correlação entre log <i>K_i</i> experimental e calculado, para os inibidores da série 1.	52
Figura 19. Gráfico de correlação entre log <i>K_i</i> experimental e calculado, para os inibidores da série 2.	55
Figura 20. Esquema ilustrativo do comportamento de ΔG_{int} em função de ΔG_l e de suas componentes $\Delta G_{aq/enz}$ e $\Delta G_{enz/sítio}$ para a equação 16 e para a equação 17.....	57
Figura 21. Estrutura 26 complexada no sítio ativo da urease 4UBP (<i>Bacillus Pasteurii</i>) com o grupo nitro direcionado para os íons de níquel (II).	60

Figura 22. Estrutura 22 complexada no sítio ativo da urease 4UBP (<i>Bacillus Pasteurii</i>).....	61
Figura 23. Fosforilidrazona proposta, derivada da β -lapachona.....	63
Figura 24. Estrutura 34 complexada no sítio ativo da urease 4UBP (<i>Bacillus Pasteurii</i>).	63

ÍNDICE DE ABREVIACÕES

AG	Algoritmo Genético
ADF	Amsterdam Density Functional
AM1	Austin Model 1
AMBER	Assisted Model Building and Energy Refinement
COSMO	Conductor-like Screening Model
DFT	Teoria do Funcional de Densidade
FEP	Perturbação da energia livre
HF	Hartree-Fock
IT	Integração Termodinâmica
K_i	Constante inibitória
LCAO	Linear Combination of Atomic Orbitals
LIE	Linear Interaction Energy
LMO	Orbitais Moleculares Localizados
MM2	Molecular Mechanics 2
MMFF	Merck Molecular Force Field
MNDO	Modified Neglect of Diatomic Overlap
MNDO/d	Modified Neglect of Diatomic Overlap, incluindo orbitais d
MOPAC	Molecular Orbital Package
NBPT	Triamida N-butil-tiofosfórica
NBPTO	Triamida N-butil-tiofosfórica oxidada
NDDO	Neglect of Diatomic Differential Overlap
N_{LR}	Número de Ligações Rotacionáveis
OM	Orbital Molecular
PDB	Protein Data Bank
PM3	Parametric Method 3
PM5	Parametric Method 5
PM6	Parametric Method 6
PM7	Parametric Method 7
PMF	Potencial of Mean Force
PPD	Fenilfosforodiamidato
RM1	Recife Model 1
RMN	Ressonância Magnética Nuclear
RMSD	Root-Mean-Squared Deviation
SEP	Superfície de Energia Potencial
STO	Slater Type Orbitals
UHF	Unrestricted Hartree-Fock
ZDO	Zero Differential Overlap

SUMÁRIO

1 Introdução	01
2 Revisão de Literatura	03
2.1 Metaloenzimas	03
2.2 Urease	03
2.3 Ativação da Urease	04
2.4 O sítio ativo da urease ativada	05
2.5 Mecanismo catalítico da urease	06
2.6 infecções causadas por bactérias produtoras de amônia	09
2.7 Aumento da eficiência no uso de uréia na agricultura	10
2.8 Fosforamidatos derivados de isatinas e quinonas	11
2.8.1 Organofosforados	11
2.8.2 Isatina	12
2.8.3 Quinonas	13
2.8.4 Hidrazonas	14
2.9 A Química Computacional e a Modelagem Molecular	14
2.9.1 Métodos clássicos	16
2.9.2 Métodos quânticos	18
2.10 Métodos Semi-Empíricos	22
2.11 <i>Docking</i>	24
3 Objetivos	29
4 Metodologia	30
4.1 Estudo da multiplicidade de spin dos íons níquel(II) e do estado de protonação do átomo de oxigênio entre os dois íons níquel(II).	30
4.2 Otimização dos ligantes	32
4.3 <i>Docking</i>	36
4.4 Cálculo das entalpias de interação inibidor-urease	37
4.5 Construção do modelo Empírico para predição da atividade inibitória na urease.	38
5 Resultados e Discussão	43
5.1 Estudo da multiplicidade de spin dos íons níquel(II) e do estado de protonação do átomo de oxigênio entre os íons níquel(II).	43
5.2 Escolha da melhor função de pontuação para o <i>docking</i>	47

5.3 Construção do modelo de predição de atividade de inibidores da urease.	48
5.3.1. Modelo desenvolvido com base nos dados experimentais dos inibidores da série 1. ...	48
5.3.2. Modelo desenvolvido com base nos dados experimentais dos inibidores da série 2.	53
5.4 Análise dos termos envolvidos nos modelos da série 1 e 2.	56
5.4.1 Perda entrópica do inibidor durante a interação.	56
5.4.2 Energia livre de Gibbs para a passagem do inibidor do meio aquoso para o sítio da urease.	56
5.4.3 Entalpia de interação inibidor/urease	58
5.5 Predição da atividade relativa dos compostos da série 3.	58
5.6 Modificação proposta nas estruturas da série 3.	62
6. Conclusão	64
Referências Bibliográficas	65

1 INTRODUÇÃO

A urease (EC 3.5.1.5) é uma enzima de grande importância para pesquisas relacionadas com a agricultura, meio ambiente e medicina. Estratégias baseadas na inibição da urease têm sido consideradas tanto para os tratamentos de infecções causadas pelas bactérias produtoras de amônia, quanto para reduzir a poluição ambiental e alcançar maior eficiência no uso de uréia como fonte de nitrogênio na agricultura [HAQ & WADOOD, 2009]. Diversas classes de compostos são conhecidas por apresentarem atividade inibitória nesta enzima, sendo a classe dos fosforamidatos uma das mais ativas [VASSILIOU *et al.*, 2008].

Através do uso de métodos de química computacional em estudos envolvendo a interação entre uma enzima com algum ligante, é possível compreender os principais fatores envolvidos nesta interação. Se os resultados previstos por estes métodos para a eficiência da inibição enzimática puderem ser correlacionados com a atividade inibitória obtida experimentalmente através de um modelo matemático adequado, o modelo proposto pode ser explorado para se obter uma previsão sobre a inibição enzimática de novas moléculas.

O processo de reconhecimento molecular enzima-ligante é dirigido por uma combinação de efeitos entálpicos e entrópicos. Estes efeitos podem ser estimados através da energia livre de Gibbs da interação enzima-ligante que está diretamente relacionada à constante de inibição K_i , uma constante cinética, a qual pode ser medida experimentalmente. Diferentes métodos computacionais já foram desenvolvidos para estimar a energia livre de interação, tais como os métodos de integração termodinâmica (IT) [KOLLMAN, 1993], da perturbação da energia livre (FEP) [BROOKS *et al.*, 1988] e o método da energia de interação linear (LIE) [ASI *et al.*, 2004, HANSSON *et al.*, 1998], mas a obtenção de estimativas precisas de energia livre de Gibbs com estes métodos pode envolver um custo computacional muitas vezes inviável para sistemas moleculares com grande número de átomos.

Uma maneira mais simples que pode ser explorada para se solucionar este problema é o desenvolvimento de um modelo baseado em equações empíricas. Este modelo empírico é ajustado aos valores experimentais de K_i e associa termos de energia livre mais importantes para o fenômeno de interação enzima-ligante [WANG *et al.*, 1994].

Na primeira etapa deste trabalho buscou-se fazer um estudo do sítio ativo da urease, avaliando-se a multiplicidade de spin dos íons Ni^{2+} presentes neste sítio, o estado de protonação do átomo de oxigênio localizado entre os dois íons Ni^{2+} e o efeito da constante

dielétrica do meio.

Na segunda etapa foi feito um estudo de *docking* para a obtenção das estruturas de complexos iniciais entre a urease e inibidores conhecidos da literatura. Em razão das aproximações envolvidas no método de *docking*, aplicou-se em seguida um método quântico aos melhores resultados obtidos por *docking*, a fim de se obter resultados mais precisos tanto do ponto de vista estrutural quanto energético. A presença de um sítio catalítico bimetalico composto por dois íons Ni^{2+} com camada aberta pode trazer dificuldades na avaliação das estruturas dos complexos inibidor-enzima [MUSIANI *et al.*, 2001]. Utilizou-se nesta etapa o programa Mopac2009 (Stewart Computational Chemistry), que dispõe de um método quântico semi-empírico com parâmetros para níquel e que, graças ao uso de orbitais moleculares localizados (LMO), possibilita a modelagem molecular com muito mais rapidez do que os métodos de orbitais moleculares *ab initio* [STEWART, 1996].

A seguir, buscou-se desenvolver uma equação empírica de predição da atividade de substâncias com potencial inibidor da urease, baseado nos dados de atividade inibitória experimentais da literatura. Com esta equação, procurou-se combinar parâmetros teóricos importantes no fenômeno de interação enzima-ligante, como a entalpia de interação calculada na etapa anterior, e utilizá-los na predição da atividade relativa das dialquilfosforilidrazonas sintetizadas pelo grupo de síntese de organofosforados da UFRRJ. Essas informações foram usadas para se propor modificações nas estruturas desses organofosforados com o objetivo de se otimizar sua interação com a enzima e elevar a eficiência destes compostos.

2 REVISÃO DE LITERATURA

2.1 Metaloenzimas

Cerca de um terço de todas as proteínas conhecidas atualmente são metaloproteínas. A natureza selecionou os metais graças à suas propriedades especiais para realizar uma ampla variedade de funções associadas aos sistemas vivos. Os dois principais papéis dos íons metálicos em proteínas podem ser descritos como o estrutural e o funcional. No primeiro caso, o íon metálico ajuda a estabilizar a estrutura da proteína e no segundo, os íons metálicos estão envolvidos em processos reativos [FENTON, 1995].

Dentre as metaloproteínas que participam de processos reativos, muitas são consideradas enzimas, as quais são fundamentais na catálise de reações críticas à vida na Terra [ROSENZWEIG, 2006]. Metaloenzimas são enzimas que contêm no mínimo um íon metálico em seu sítio ativo. Estes sítios são constituídos por cadeias laterais de aminoácidos, algumas ligadas a um ou mais íons metálicos, e pontes intermetálicas, em alguns casos. As interações entre íons metálicos e biomoléculas são, geralmente, da mesma natureza das existentes em complexos e, por isso, são tratadas de acordo com os conceitos da química de coordenação [COWAN, 1993]. Sendo assim, as propriedades das biomoléculas que contêm metais dependem do número e distribuição de elétrons de valência nos orbitais d, da geometria, do campo e da multiplicidade de spin do sistema.

2.2 Urease

A urease é uma metaloenzima encontrada numa grande variedade de plantas, fungos e bactérias. Ela catalisa a reação de hidrólise da uréia para formar amônia e carbamato, que é a última etapa do metabolismo de nitrogênio dos organismos vivos. O carbamato se decompõe de forma rápida e espontânea, produzindo uma segunda molécula de amônia e uma molécula de dióxido de carbono (Figura 1) [PEARSON, 1997].

A uréia é produzida em grandes quantidades através do catabolismo de compostos que contêm nitrogênio. Em solução aquosa, a enzima urease é capaz de decompor a uréia em uma reação 10^{14} vezes mais rápida do que a reação não catalisada. [WOLFENDER & SNIDER, 2001].

A urease é uma enzima importante para as pesquisas relacionadas com a agricultura e medicina. Ela está diretamente envolvida com a formação de cálculos renais e contribui para patologias como encefalopatia hepática, coma hepático, incrustação de catéter urinário, gastrite e úlceras que podem progredir para o câncer. [SMOOT *et al.*, 1990]

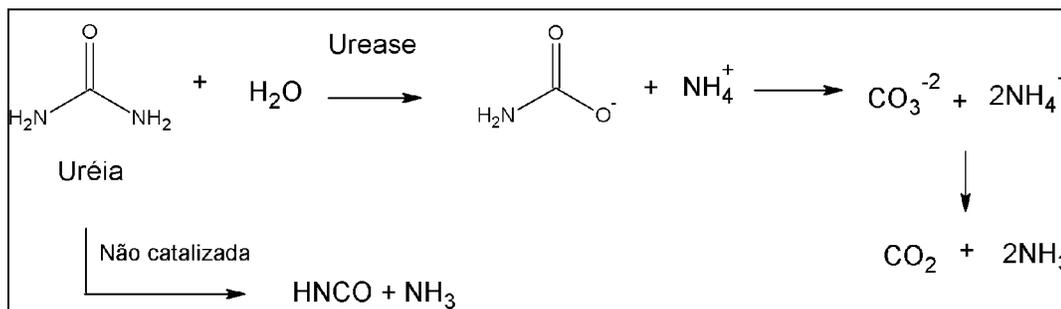


Figura 4. Reações de decomposição da uréia, adaptada de Estiu e Merz, 2007

Na agricultura, a elevada atividade desta enzima presente no solo causa problemas ambientais e econômicos por causa da liberação excessiva de amônia na atmosfera durante a fertilização com uréia [BREMNER, 1995]. Assim, a eficiência dos fertilizantes nitrogenados à base de uréia é diminuída, causando danos nas plantas que ficam privadas de um de seus principais nutrientes. No solo, o nitrogênio da uréia é perdido rapidamente na forma do gás NH_3 . Ao contrário do N_2O , o NH_3 não é um gás de efeito estufa, mas pode agir como uma fonte secundária de N_2O e, portanto, contribui indiretamente para o aquecimento global e destruição do ozônio estratosférico [SAGGAR *et al.*, 2012].

Desta forma, estratégias baseadas na inibição da urease têm sido consideradas tanto para os tratamentos de infecções causadas pelas bactérias produtoras de amônia, quanto para reduzir a poluição ambiental e alcançar maior eficiência no uso de uréia como fonte de nitrogênio na agricultura [HAQ & WADOOD, 2008].

2.3 Ativação da Urease

Diversas enzimas são produzidas na sua forma inativa e precisam sofrer alterações em sua estrutura para se tornarem ativas. A urease é assim, pois precisa receber dois íons de níquel e uma lisina deve ser carboxilada para se coordenar com o metal recebido. O processo de incorporação de níquel no sítio ativo das ureases requer a participação de várias proteínas

acessórias (ou ativadoras). Nas ureases de bactérias tem-se como proteínas ativadoras a ureH, ureF e ureG, e nas ureases de plantas e fungos existe também a UreE, que parece atuar como chaperona¹ específica (HAUSINGER *et al.*, 2001).

UreH, UreF, UreG provavelmente formam um complexo de ativação *in vivo*, que liga-se ao centro metálico e age em combinação com UreE (HAUSINGER *et al.*, 2001). Fong e colaboradores (2011) investigaram como ocorre a formação do complexo UreF-G-H. Eles concluíram que a UreH e UreF interagem formando um complexo binário. Nessa interação UreF sofre uma alteração conformacional, que permite o recrutamento e a ligação com UreG. UreE é uma chaperona de níquel que facilita a entrega de níquel para o sítio ativo. O complexo UreH-UreF-UreG insere CO₂ e níquel na apoenzima, em um processo mediado pela hidrólise de GTP pela UreG [ZAMBELLI *et al.*, 2005]. O CO₂ liga-se ao grupamento amino de um resíduo de lisina do sítio ativo da urease, formando um grupo carbamato que atua como ponte entre os dois átomos de níquel do sítio ativo [JABRI *et al.*, 1995].

2.4 O sítio ativo da urease ativada

A urease representa um marco na bioquímica porque ela foi a primeira enzima a ser cristalizada [SUMNER, 1926] e também a primeira metaloenzima caracterizada contendo átomos de níquel [DIXON *et al.*, 1975]. Atualmente, constam 74 estruturas cristalográficas da urease de diferentes microrganismos e plantas, com ou sem inibidores, depositadas no *Protein Data Bank* (PDB), um banco de dados de estruturas de proteínas de acesso público (www.rcsb.org).

No sítio ativo da urease de *Klebsiella aerogenes* (código do PDB: 1FWJ) é visualizado um complexo dinuclear de íons Ni²⁺ (Ni1 e Ni2) com uma distância interatômica de aproximadamente 3,5 Å entre Ni1 e Ni2 (Figura 2). Estes dois íons estão interconectados mediante uma molécula de água (WB) e o resíduo de lisina carbamoilada. O íon Ni1 está também coordenado a dois resíduos de histidina e a uma segunda molécula de água (W1). O íon Ni2 encontra-se coordenado com a cadeia lateral de um resíduo de aspartato, com dois resíduos de histidina e com outra molécula de água (W2) [SUAREZ *et al.*, 2003].

Na urease, a posição das moléculas de água no sítio ativo é de grande importância [BENINI *et al.*, 1999]. Uma quarta molécula de água (W3) interage fortemente com a ponte

¹ Chaperonas são proteínas responsáveis por auxiliar no enovelamento de outras proteínas e também por destruí-las caso não seja possível alcançar a configuração terciária correta.

formada pelas moléculas de água WB, W1 e W2 através de uma ligação hidrogênio curta [SUAREZ *et al.*, 2003].

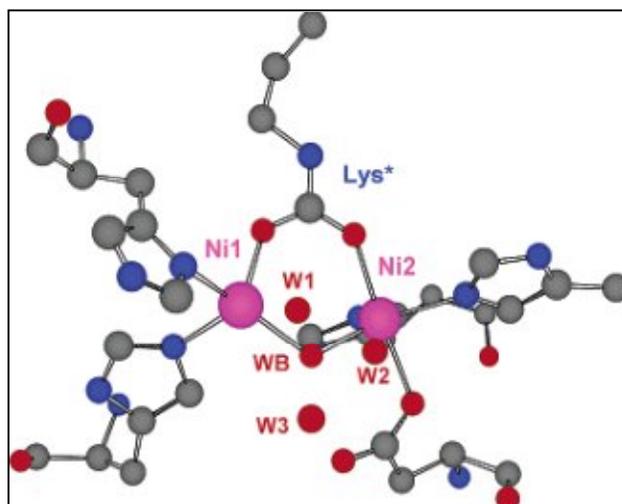


Figura 5. Sítio ativo da urease de *K. aerogenes* (código 1FWJ do PDB) determinada por cristalografia de raio-X [Suarez *et al.*, 2003].

2.5 Mecanismo catalítico da urease

Muitos estudos computacionais do mecanismo da urease foram publicados durante a última década, incluindo estudos de mecânica molecular por Zimer [ZIMER, 2000] e Smyj [SMYJ, 1997] e estudos com a Teoria do Funcional de Densidade (DFT) por Estiu e Merz [ESTIU & MERZ, 2007]. Diversos aspectos dessa reação têm sido alvo de interesse dos pesquisadores. Dentre eles, a competição entre o mecanismo hidrolítico e por eliminação tem sido um recente foco de interesse.

Baseado em estudos bioquímicos e experimentos mutagênicos, Karplus e colaboradores e Benini e colaboradores propuseram o mecanismo hidrolítico, no qual a molécula de água coordenada com o centro binuclear de níquel ataca o carbono carbonílico da molécula de uréia coordenada, resultando em um intermediário tetraédrico após a liberação de amônia e carbamato [BENINI *et al.*, 1999; KARPLUS *et al.*, 1997]. Conforme pode ser observado na figura 3, os dois mecanismos diferem basicamente na catalise ácido/base e no estado de protonação do resíduo de histidina 320. Ambos os mecanismos assumem que a uréia se liga aos íons níquel no lugar de W1 e WB e que o grupo carbonila da uréia se liga ao níquel mais eletrofílico Ni1, enquanto uma amina terminal da uréia se coordena ao Ni2. Este modo de ligação descrito pelos dois mecanismos estabiliza a uréia por ressonância facilitando

o ataque da hidroxila ao carbono carbonílico, formando assim, um intermediário tetraédrico. De acordo com a proposta de Benini e colaboradores a hidroxila em ponte com Ni1 e Ni2 poderia atuar como ácido doando seu átomo de hidrogênio, a molécula de amônia liberada interagiria com um resíduo de His320. De maneira alternativa, Karplus e colaboradores sugeriram que a protonação ocorreria através do resíduo ácido da His320, uma vez que a forma nativa da enzima a His320 encontra-se protonada. Karplus e colaboradores não estabeleceram nenhuma regra em relação à atuação nucleofílica da hidroxila em ponte com os dois metais.

Segundo Estiu e Merz (2007), o mecanismo de hidrólise precisa ser revisto, pois intermediários do ácido ciânico foram detectados em reações com modelos do sítio ativo enzimático contendo íons Ni^{2+} , apoiando a hipótese de um mecanismo de eliminação, que compete com o mecanismo hidrolítico, ainda que o segundo seja preferido devido à observação de liberação do ácido carbâmico na presença de urease. Estiu e Merz empregaram metodologias computacionais para calcular as energias envolvidas na decomposição catalítica da uréia via eliminação e via hidrolítica. O estudo computacional foi realizado com funcional de densidade B3LYP com o conjunto de bases 6-31G* para os não metais e a base LANL2DZ para os metais, através do programa Gaussian 03. Os autores descobriram que ambos os mecanismos são igualmente favorecidos e devem competir dentro do sítio ativo da enzima.

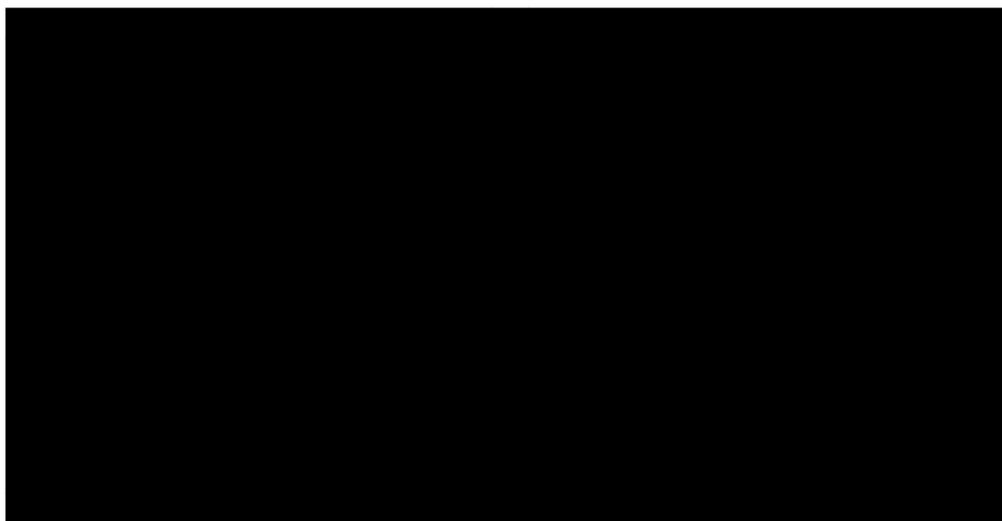


Figura 6. Dois mecanismos de reação propostos [Suarez *et al.*, 2003].

Um pouco mais tarde, Carlsson e Nordlander (2010), baseando-se nos estudos anteriores, realizaram cálculos DFT e identificaram que o oxigênio doador localizado entre os íons Ni^{2+} é um íon hidróxido. Os cálculos DFT de densidade foram realizados com o programa *Amsterdam Density Functional* (ADF), com o conjunto de base triple- ζ STO com uma função de polarização adicional. Carlsson e Nordlander propuseram dois mecanismos para a reação de hidrólise da uréia (A e B) (Figura 4).

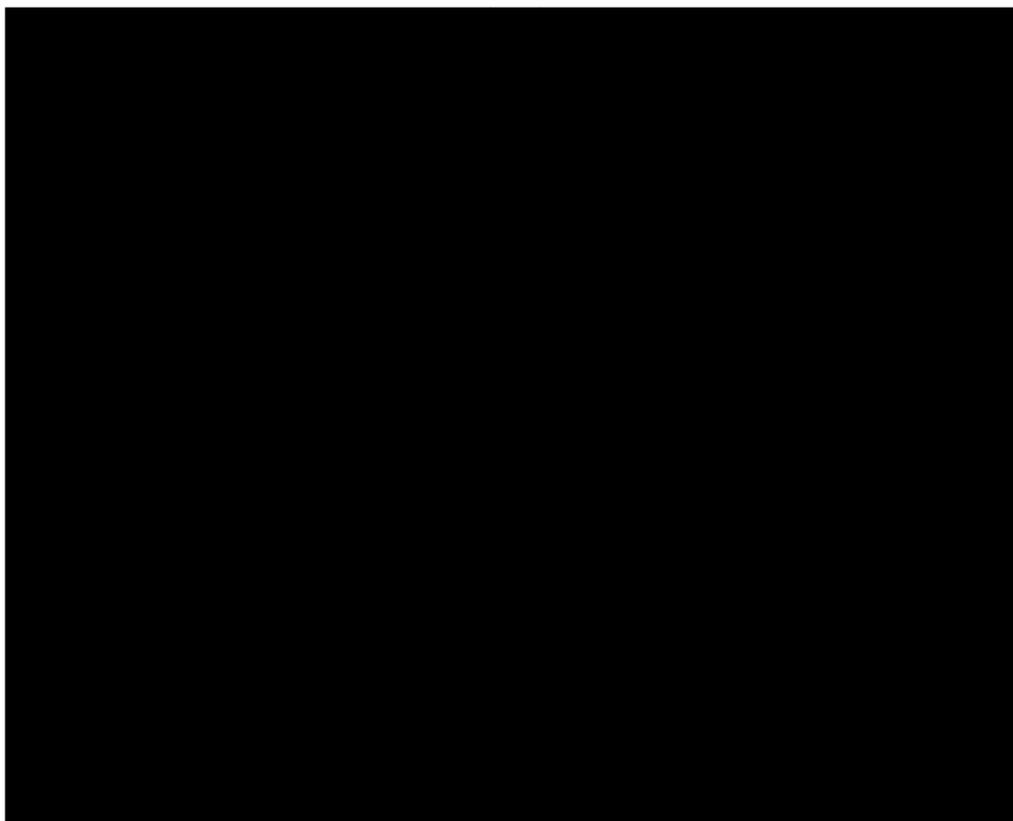


Figura 4. Representação esquemática de dois mecanismos propostos para a urease, A e B [CARLSSON & NORDLANDER, 2010].

O estudo computacional do mecanismo A proposto foi realizado e descobriu-se que é energeticamente possível coordenar um nitrogênio da uréia com o íon Ni^{2+} de maior número de coordenação, mas a etapa posterior, que consiste na formação do intermediário tetraédrico é energeticamente desfavorável pois, segundo a literatura [MOYO *et al.*, 1989], a energia de ativação é de cerca de 50 kJ/mol, enquanto os cálculos indicaram que 150 kJ/mol seriam necessários para se alcançar um intermediário tetraédrico. Segundo os autores, mecanismos alternativos envolvendo o ataque nucleofílico pelo hidróxido ligado ao íon Ni^{2+} de maior

coordenação ao carbono da uréia (mecanismo B) ainda precisam ser estudados detalhadamente [CARLSSON & NORDLANDER, 2010].

2.6 infecções causadas por bactérias produtoras de amônia

A bactéria *Helicobacter Pilory* foi descoberta a pouco mais de 30 anos, quando Robin Warren e Barry Marshall conseguiram identificar sua presença em biópsias de pacientes com gastrite crônica e úlcera péptica e hipotetizaram que a causa dessas doenças seria essa bactéria e não os fatores (como o estresse) até então tidos como causa [MARSHALL & WARREN, 1984]. A comunidade médica da época se recusou a ceder a esse estudo. Somente após Marshall ingerir uma cultura dessa bactéria, vindo a desenvolver um quadro de gastrite aguda e novos estudos serem realizados, a hipótese foi então confirmada. O impacto desse trabalho foi tão grande para a medicina, que esses pesquisadores foram laureados com um prêmio Nobel em 2005.

A infecção por *H. pylori* leva a um quadro de gastrite crônica em seus hospedeiros. A gastrite atinge cerca de 80% dos portadores da bactéria e é, na maioria dos casos, um quadro assintomático. Cerca de 10-15% dos portadores desenvolvem quadros de úlcera péptica e 1-3%, carcinoma gástrico. Cerca de 50% da população mundial tem essa bactéria, e sua prevalência é inversamente proporcional ao nível socioeconômico da população, ou seja, países pobres apresentam prevalência maior do que em países ricos [www.helico.com].

H. pylori é uma bactéria exclusivamente humana e habita o estômago, um ambiente muito ácido (pH~2). Esse pH muito baixo torna o estômago inóspito à vida, mas a *H. pylori* consegue sobreviver ali pois consegue produzir e liberar a urease. A enzima quebra a uréia presente no suco gástrico produzindo amônia, que neutraliza o ambiente ácido mantido pelo ácido clorídrico do suco gástrico (Figura 5) [www.helico.com].

São vários os antimicrobianos já usados no tratamento da infecção por *H. pylori*, em várias associações e com tempos diversos de tratamento, porém ainda não existe um esquema terapêutico ideal. Estratégias que impeçam a ativação da urease, interferindo na interação entre as unidades UreF, UreG e UreH, também estão sendo estudadas com o objetivo de se desenvolver um antibiótico contra a bactéria *H. pilory* [FONG *et al.*, 2011].

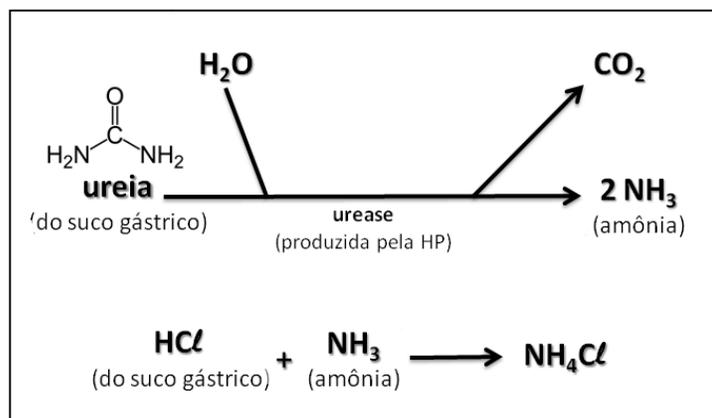


Figura 5. A urease quebra a uréia presente no suco gástrico em amônia que neutraliza o ambiente ácido do estômago.

2.7 Aumento da eficiência no uso de uréia na agricultura

A uréia é caracterizada como um dos fertilizantes sólidos granulados de maior concentração de nitrogênio, possuindo 45% deste elemento na forma amídica. Como vantagem da utilização da uréia, pode-se citar o baixo custo de transporte, alta solubilidade, baixa corrosividade e facilidade de mistura com outras fontes. Como desvantagem, possui elevada higroscopicidade e maior susceptibilidade à volatilização. A degradação e dissolução dos grânulos aplicados ao solo ocorrem na presença de umidade [RAIJ *et al.*, 2001].

A uréia, que responde por 60% dos fertilizantes nitrogenados empregados na agricultura brasileira, apresenta limitações quanto à aplicação superficial, devido à possibilidade de perdas por volatilização de NH₃. A reação inicial pode levar o pH do solo próximo de 9 nas imediações dos grânulos desse fertilizante, intensificando a volatilização de NH₃ [OVERREIN & MOE, 1967]. Como consequência, as perdas em aplicações de uréia na superfície podem alcançar 50%, dependendo das características do solo [DOMÍNGUEZ *et al.*, 2008]. Outro efeito negativo da hidrólise da uréia é o acúmulo de NO₂ devido à inibição de atividade de bactérias que oxidam NO₂ a nitrato. Esse acúmulo prejudica tanto as sementes quanto as plantas novas, além de aumentar a possibilidade do N ser liberado para a atmosfera como N₂O, um gás do efeito estufa e que também contribui para a redução da camada de ozônio [ZAMAN *et al.*, 2008].

O desenvolvimento de substâncias inibidoras de urease para serem aplicadas no solo junto com a uréia é uma estratégia usada para reduzir as perdas desse fertilizante [DOMÍNGUEZ *et*

al., 2008]. Entre os inibidores de urease mais ativos estão os derivados da classe das fosforodiamidas e fosfortriamidas. A triamida *N*-butil-tiofosfórica (NBPT) (Figura 6) é um inibidor efetivo de urease mesmo em baixas concentrações, em condições de laboratório [CARMONA *et al.*, 1990; GILL *et al.*, 1997]. Este inibidor forma complexos estáveis com a enzima após conversão à sua forma oxidada (NBPTO) no solo [McCarty *et al.*, 1989]. Outros exemplos de inibidores de urease são o fenilfosforodiamidato (PPD) e a *N*-diaminofosforil-4-fluorobenzamida (fluorofamida) (Figura 6). Recentemente, novos derivados tiofosfínicos foram desenvolvidos como inibidores de urease de *Bacillus pasteurii* e *Proteus vulgaris*, alguns apresentando valores de constantes inibitórias (K_i) na faixa nanomolar [VASSILIOU *et al.*, 2008; VASSILIOU *et al.*, 2010].

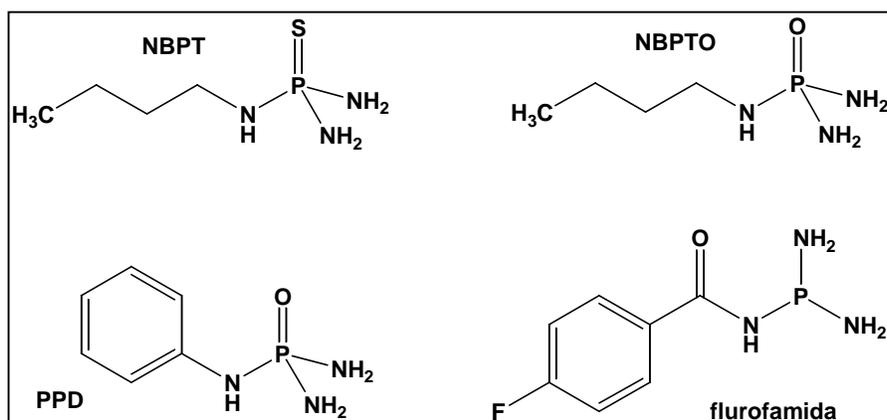


Figura 6. Alguns inibidores de urease.

2.8 Fosforamidatos derivados de isatinas e quinonas

2.8.1 Organofosforados

O átomo de fósforo é polarizável, de baixa a média eletronegatividade, possui distribuição eletrônica com orbitais d acessíveis e usualmente tem número de coordenação igual a três. No estado pentavalente possui três átomos ligados por ligação simples e uma ligação dupla com o átomo de oxigênio ou com outros átomos bivalentes. Os compostos de fósforo trivalentes são bons agentes redutores por causa da forte reatividade do par de elétrons não ligantes e da facilidade de formação da ligação P=O e, por isto, os compostos fosforados são rapidamente oxidados por oxigênio, ozônio, peróxidos e outros agentes oxidantes. A força

da ligação carbono-fósforo é similar à da ligação carbono-carbono e muito resistente à oxidação e hidrólise. As diferentes atividades dos compostos organofosforados estão relacionadas, evidentemente, com características estruturais, tais como o tipo de heteroátomo ou grupo funcional ligado ao átomo de fósforo e seu estado de oxidação [DACOSTA *et al.*, 2007].

2.8.2 Isatina

A isatina (1H-indol-2,3-diona) (figura 7) é um derivado metabólico da adrenalina e por isso está presente no cérebro, principalmente no hipocampo e no cerebelo. Encontra-se também distribuída na natureza como, por exemplo, em algumas plantas. Entre as várias funções da isatina, a função de modulador em diferentes processos bioquímicos tem servido de base a inúmeros estudos na última década. Sua descoberta deveu-se inicialmente a um possível agente inibidor da monoamina oxidase endógena. Outros estudos revelaram que a isatina atua também como inibidor da guanilato ciclase, estimulada pelo óxido nítrico [BACCHI *et al.*, 2005].

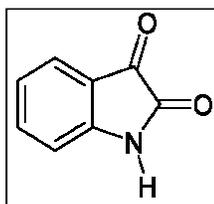


Figura 7. Isatina (1H-indol-2,3-diona)

Derivados da isatina constituem uma classe de compostos de amplo espectro de propriedades terapêuticas, destacando-se suas atividades bactericida, antiprotozoária, antitumoral e antiviral [BACCHI *et al.*, 2005]. A importância dessa substância em síntese orgânica deve-se ao fato de ser um composto heterocíclico muito versátil e, por isso, frequentemente utilizado como substrato na síntese de uma grande variedade de compostos como, por exemplo, de fosforamidatos sintetizados pelo grupo do Dr. João Batista Neves da Costa, estudados nesse trabalho como possíveis inibidores da urease.

2.8.3 Quinonas

Com base na sua estrutura molecular, as quinonas são divididas em diferentes grupos, utilizando-se como critério o tipo de sistema aromático do qual deriva o núcleo quinonoídico. Em decorrência de diferentes arranjos quinonoídicos (isomeria), podem-se ter diferentes quinonas, dependendo das disposições relativas das carbonilas. Por exemplo, a quinona β -lapachona é a forma isomérica orto-quinonoídica (carbonilas adjacentes), enquanto a α -lapachona é a forma para-quinonoídica (carbonilas estão separadas por dois carbonos) (Figura 8). Estas formas isoméricas diferem muito em suas propriedades físicas, químicas e biológicas [FERREIRA *et al.*, 2010].

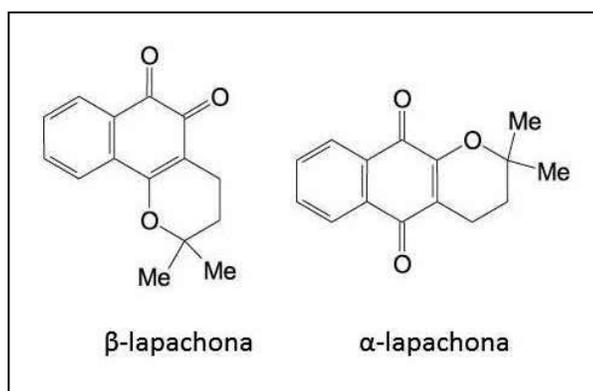


Figura 8. Forma isomérica orto-quinonoídica e para-quinonoídica.

Em estudos farmacológicos as lapachonas mostram variadas propriedades microbicidas, tripanossomicidas, viruscidas, antitumorais e inibidoras de sistemas celulares reparadores, processos nos quais atuam de diferentes formas. Como exemplo, destaca-se o estresse oxidativo que provocam, ao induzirem a formação deletéria endógena de espécies bioativas derivadas do oxigênio, como ocorre no *Trypanosoma cruzi*, agente causador da doença de Chagas [MONKS *et al.*, 1992]. Outra atividade marcante destas substâncias é a inibição do complexo das topoisomerasas, ação que provoca o desencadeamento da apoptose celular (suicídio celular) [SILVA *et al.*, 2003]. Neste trabalho, fosforamidatos derivados da α -lapachona e da β -lapachona sintetizados pelo grupo de síntese do Dr. João Batista Neves da Costa também foram estudados como possíveis inibidores da urease.

2.8.4 Hidrazonas

Na literatura Química, as hidrazonas não são usualmente reconhecidas como formadoras de uma classe diferenciada de compostos orgânicos, sendo frequentemente consideradas como substâncias derivadas de compostos carbonílicos. São consideradas hidrazonas todos os compostos cujas moléculas possuem em sua estrutura a porção triatômica C=N-N.

As hidrazonas têm uma vasta aplicação na síntese de moléculas com as mais variadas estruturas no isolamento de compostos carbonílicos e como ligantes para um grande número de cátions metálicos. São aplicadas como plastificantes e estabilizadores para polímeros, iniciadores de polimerização, antioxidantes, etc. Porém, a mais valiosa propriedade das hidrazonas é sua atividade biológica, entre elas, reguladores de crescimento de plantas, ação anti-inflamatória, antitrombótica, analgésica, hipotensiva, antitumoral (incluindo sarcomas e outros neoplasmas malignos) [BARREIRO *et al.*, 2002].

A função hidrazona está presente nos fosforamidatos derivados de isatinas e lapachonas, sintetizados pelo grupo de síntese do Dr. João Batista Neves da Costa, servindo de “ponte” entre o grupo fosforamida e a isatina ou a lapachona.

2.9 A Química Computacional e a Modelagem Molecular

Com o advento da Mecânica Quântica, no início do século XX, as aplicações da Matemática aos problemas da Química adentraram uma nova dimensão, expandindo suas possibilidades para além daquelas já existentes. Graças às poderosas ferramentas matemáticas surgidas durante o desenvolvimento da Química Quântica, tornou-se possível para os químicos prever algumas propriedades de moléculas antes mesmo de sua síntese.

Com o surgimento e o desenvolvimento da computação, grandes progressos começaram a ocorrer na área dos estudos teóricos em Química. Desde os primeiros “computadores” em Los Alamos na década de 40, até as modernas estações de trabalho, *clusters* e super-computadores dos centros de computação científica, a revolução promovida pela computação tem trazido grandes benefícios para essa área, permitindo que sistemas moleculares cada vez maiores sejam estudados. O uso de métodos computacionais para compreensão e previsão do comportamento de sistemas moleculares forma o ramo do conhecimento químico denominado Química Computacional [CLARK, 1985].

Com o desenvolvimento da Física Atômica, no início do século XX, foram dados os primeiros passos bem-sucedidos para a transição da representação de estruturas moleculares da esfera bidimensional das folhas de papel para o âmbito tridimensional. A cristalografia teve um papel decisivo neste processo, pois foi por meio desta técnica que foram obtidas as primeiras informações estruturais precisas sobre comprimentos e ângulos de ligação [HOLTJE & FOLKERS]. De posse destas informações estruturais, foram desenvolvidos os primeiros modelos de estruturas tridimensionais que, até hoje, desempenham papel essencial na pesquisa e no ensino. Com o desenvolvimento de programas capazes de determinar a estrutura molecular por meio de cálculos teóricos, surgiu a modelagem molecular, que, segundo a IUPAC, trata da investigação das estruturas e das propriedades moleculares pelo uso da química computacional e de técnicas de visualização gráfica, visando fornecer uma representação tridimensional, sob um dado conjunto de circunstâncias [SANT'ANNA, 2002].

Quando se trata de proteínas ou de seus complexos com moléculas pequenas (ligantes), técnicas experimentais como a cristalografia de raios-X e a ressonância magnética nuclear têm fornecido dados como estrutura e especificidade de uma enzima ou de um receptor, detalhando-se a estrutura do seu sítio ativo e seu modo de interação com um ligante. De posse de tais informações, a modelagem pode ser usada para o planejamento de novos ligantes, com propriedades estruturais otimizadas para um melhor reconhecimento molecular pela enzima, em se tratando de inibidores enzimáticos, ou pelo receptor, no caso de agonistas ou antagonistas.

A modelagem molecular pode contribuir em todos os estágios do planejamento de compostos candidatos a ensaios biológicos e testes clínicos, desde a descoberta de um protótipo até sua otimização, em termos de afinidade, especificidade, toxicidade e biodisponibilidade [OPREA, 2005]. Do ponto de vista do planejamento, é importante destacar que, quando se fala do uso da modelagem molecular, não se pretende chegar a uma molécula bioativa simplesmente através do uso de programas de computador. O processo de desenvolvimento dessas moléculas, devido à sua complexidade, envolve necessariamente o trabalho de uma equipe multidisciplinar, que emprega um grande conjunto de métodos computacionais de modo sistemático de forma a facilitar e otimizar o processo de desenvolvimento de compostos bioativos, em uma constante troca de informações com grupos de síntese química e avaliação da atividade destes compostos [SANT'ANNA, 2009].

Os dois métodos que têm sido usados predominantemente em estudos de modelagem molecular são o método clássico, que inclui os métodos da mecânica molecular e da dinâmica

molecular, e o método quântico, que inclui os métodos *ab initio* e semi-empíricos. De modo geral, a escolha entre estes métodos depende das propriedades que se deseja avaliar, da precisão desejada e da capacidade computacional disponível para a realização dos cálculos [SANT'ANNA, 2009].

2.9.1 Métodos clássicos

Os métodos de mecânica molecular baseiam-se no fato de que as moléculas são compostas de unidades estruturalmente similares e, por esta razão, algumas propriedades podem ser transferidas de uma molécula para a outra, tais como comprimentos de ligação, frequências vibracionais, entalpias de formação, entre outras [JENSEN, 1999].

O modelo da mecânica molecular é justificável porque os parâmetros associados a conjuntos de átomos permanecem razoavelmente constantes entre estruturas diferentes, desde que o tipo e a hibridação dos átomos envolvidos sejam os mesmos [BURKET, 1982]. De fato, tabelas de distâncias e ângulos “normais” ou “naturais” de ligação, determinados experimentalmente, são facilmente encontradas na literatura de química. O que se faz no modelo da mecânica molecular é desenvolver o chamado campo de força, um conjunto de funções de energia que determinam penalidades energéticas para o afastamento desses valores “normais” da estrutura [BURKET, 1982]. A minimização de energia é um processo no qual, através de um algoritmo matemático adequado, busca-se reduzir essas penalidades em conjunto, ou seja, reduzir a energia a um mínimo [SANT'ANNA, 2009].

Em um campo de força genérico da mecânica molecular, pode-se descrever, por exemplo, a ligação entre dois átomos por um oscilador harmônico simples,

$$E_{est} = \sum_{\text{ligações}} K_R^{(0)} (r - r^{(0)})^2 \quad (01)$$

onde a constante de força $K_R^{(0)}$ e a distância de equilíbrio do oscilador $r^{(0)}$ dependem dos átomos que estão participando da ligação e seus valores são obtidos de um conjunto de moléculas escolhido como padrão. O ângulo θ entre três átomos também pode ser descrito como um oscilador harmônico,

$$E_{ang} = \sum_{\text{ângulos}} K_R^{(0)} (\theta - \theta^{(0)})^2 \quad (02)$$

O termo de energia conformacional associado à torção de ângulos diedros ω é apropriadamente descrito por uma série truncada de Fourier,

$$E_{tors} = \sum_{n=1} V_n \cos(1 + n\omega) \quad (03)$$

onde n descreve a periodicidade de uma rotação.

O termo de energia eletrostática é dado pela expressão comum de Coulomb,

$$E_{el} = \sum_{i < j} \frac{q_i q_j}{\epsilon_0 R_{ij}} \quad (04)$$

onde q_i e q_j são as cargas de cada átomo, R_{ij} a distância entre eles e ϵ_0 é a permissividade no vácuo. O termo de energia de van der Waals pode ser descrito razoavelmente bem usando-se potenciais adequados, como o potencial de Leonard-Jones [HYPERCHEM, 1994].

Muitos campos de força estão disponíveis hoje em dia, a maioria deles parametrizados para sistemas orgânicos, fornecendo ótimos resultados qualitativos [ALLINGER, 1977; WEINER *et al.*, 1984; BROOKS *et al.*, 1983; JORGENSEN & TIRADORIVES, 1988]. Os resultados serão tanto melhores quanto mais próxima estiver a estrutura da molécula que está sendo estudada daquelas usadas como padrão. Um exemplo de campo de força muito difundido é o MM2 [ALLINGER, 1977], desenvolvido com o intuito de descrever pequenos sistemas orgânicos. Um outro exemplo, aplicável principalmente a macromoléculas de interesse biológico, como proteínas e ácidos nucleicos, é o AMBER [WEINER *et al.*, 1984].

Uma vez que bons parâmetros estejam disponíveis, será possível alcançar boa predição das geometrias, das energias relativas e barreiras conformacionais de uma grande quantidade de moléculas em um tempo computacional menor. Porém, os métodos de mecânica molecular consideram apenas a conformação de menor energia, o que trás limitações a este método uma vez que na temperatura da biofase, a molécula possui energia cinética suficiente para assumir muitas outras conformações.

Nos métodos de dinâmica molecular são exploradas muitas outras conformações, estes métodos utilizam função de busca determinística e fundamentos da mecânica molecular clássica para descrever o comportamento dinâmico do sistema ao longo do tempo; logo, descrevem as posições dos átomos em função do tempo, que tendem a se ajustar no sistema para diminuir a energia potencial [LEACH, 2006]. Neste método, a resolução das equações de movimento representa a evolução no tempo dos movimentos moleculares, chamada trajetória, que pode ser usada para o estudo de propriedades que dependem do tempo, tais como a difusão, o dobramento de cadeias moleculares (como acontece no enovelamento de proteínas) e a distribuição de moléculas de solvente ao redor de um soluto, entre outros. Em um cálculo de dinâmica molecular, as moléculas, que têm certa energia cinética, podem superar barreiras de energia potencial, o que permite a exploração da superfície de energia potencial (SEP) na busca de outras conformações estáveis [SANT'ANNA, 2009].

2.9.2 Métodos quânticos

O método quântico tem como objetivo central a obtenção de soluções da equação de Schrödinger independente do tempo (equação 05), para a determinação precisa de propriedades de sistemas atômicos e moleculares. Schrödinger concluiu que só seria possível determinar corretamente a energia em um sistema atômico com a inclusão do comportamento ondulatório proposto por de Broglie para todas as partículas (núcleos e elétrons) que constituem esse sistema. Inicialmente, Schrödinger definiu uma função de onda, Ψ , que depende das coordenadas de todos os núcleos e das coordenadas de todos os elétrons do sistema [SANT'ANNA, 2009]. Essa equação somente apresenta solução exata para problemas de um elétron.

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + V(x, y, z) \right] \psi(x, y, z) = E \psi(x, y, z) \quad (05)$$

Nesta equação, ∇^2 é o Laplaciano, \hbar é a constante de Planck dividida por 2π , m é a massa da partícula, $\Psi(x,y,z)$ é uma função de onda do sistema, $V(x,y,z)$ é um operador associado à energia potencial e E é a energia total da partícula. Esta é a equação de Schrödinger nas três dimensões (x,y,z) independente do tempo. A equação independente do tempo é suficiente para resolver os chamados problemas estacionários, que constituem muitos dos problemas de Química [ALCÁCER, 2007]. Todo o termo entre colchetes representa os

operadores de energia cinética e potencial que constituem um único operador matemático, chamado hamiltoniano H . Na mecânica quântica, o conceito de operador é de grande importância. Um operador representa uma instrução para se executar determinada ação (por exemplo, diferenciação ou integração) sobre um *operando*, que, neste caso, é a função de onda $\Psi(x,y,z)$. A ação do operador Hamiltoniano em uma função adequada (chamada autofunção) produz a própria função multiplicada pela energia total do sistema (chamada autovalor). Assim temos,

$$H \psi (r) = E \psi (r) \quad (06)$$

em que r representa o conjunto das três coordenadas do espaço.

Uma forma de simplificar a equação de Schrödinger é assumir a imobilidade dos núcleos atômicos. A massa nuclear é várias vezes maior que a massa dos elétrons, fazendo com que os elétrons se ajustem quase que instantaneamente a qualquer mudança de posição nuclear. Esta aproximação é conhecida como aproximação de Born-Oppenheimer, que permite separar a equação de Schrödinger em duas: uma para a parte eletrônica e outra para a nuclear [BORN & OPPENHEIMER, 1927]. A equação eletrônica tem a forma

$$H_{el} \psi_{el} (r) = E_{el} \psi_{el} (r) \quad (07)$$

onde o operador hamiltoniano é:

$$H = -\sum_{i=1}^N \frac{\nabla_i^2}{2m} - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (08)$$

Na última equação, Z_A é o número atômico do núcleo A , r_{ij} é a distância entre os elétrons i e j , r_{iA} é a distância entre o núcleo A e o elétron i ; N e M indicam, respectivamente, os números de elétrons e núcleos do sistema. O primeiro termo desta expressão é o termo da energia cinética dos N elétrons, o segundo é o de energia potencial atrativa dos elétrons para os núcleos e o terceiro termo corresponde às interações elétron-elétron. O termo que descreve a energia cinética nuclear é zero e o termo de interação coulombiano entre núcleos é um

parâmetro que deve ser adicionado à energia eletrônica E_{el} para produzir a energia total E do sistema.

As soluções da equação de Schrödinger estão sujeitas à restrições importantes. Em primeiro lugar, a função de onda Ψ deve ser normalizável e, em segundo lugar, tem que ser antissimétrica com respeito à permuta de dois elétrons quaisquer [ALCÁCER, 2007].

A equação de Schrödinger não pode ser resolvida exatamente para sistemas multieletrônicos, mesmo com a aproximação de Born-Oppenheimer [ALCÁCER, 2007]. Uma segunda complicação com espécies multieletrônicas é em relação ao spin. O spin é caracterizado pelo número quântico s , sendo um efeito puramente relativístico. Como a equação de Schrödinger é não relativista, inclusões adicionais que computem tal efeito na forma da função de onda devem ser feitas. A combinação da função de onda espacial do elétron com uma função de spin adequada produz a chamada função spin-orbital.

Diversos métodos têm sido desenvolvidos para a aproximação da solução da equação de Schrödinger para sistemas de muitos elétrons. Dentre estes métodos, o mais popular é o método de Hartree-Fock (HF) [HARTREE, 1923; HARTREE, 1924]. Além de ser capaz de fornecer uma solução bem aproximada para sistemas de muitos elétrons, o método HF tem a vantagem de servir como base para outros métodos: por um lado, os métodos semi-empíricos, onde aproximações adicionais são incluídas com o propósito de reduzir o custo computacional; por outro, métodos onde à solução HF são acrescentadas correções no sentido de aproximar a solução exata.

Se assumirmos que um sistema de N elétrons pode ser descrito como uma soma de N sistemas de um elétron que se move num campo gerado pelos núcleos estacionários e num campo médio resultante da distribuição espacial de todos os outros elétrons, o problema se reduz ao problema de N elétrons independentes. A teoria de HF, nesta aproximação, obriga a construção de funções spin-orbitais χ para cada elétron, que minimizem a energia do estado fundamental do sistema (método variacional) [MORGON & COUTINHO, 2007].

Estas funções spin-orbitais devem fazer parte de uma função determinantal escolhida de forma ótima, uma vez que uma função de onda Ψ_0 que descreve um sistema de muitos elétrons deve ser antissimétrica perante a troca de coordenadas de dois destes elétrons e a antissimetria é uma característica dos determinantes. Este determinante é conhecido como determinante de Slater [SLATER, 1929], onde as funções spin-orbitais moleculares são funções das coordenadas espaciais e de spin de um único elétron [MORGON & COUTINHO, 2007].

A maior diferença entre sistemas multieletrônicos e monoelétrônicos é a presença das interações entre elétrons. O que se deseja é obter uma função de onda de menor energia para um sistema multieletrônico mantendo a formulação spin-orbital individual. O problema é encontrar uma solução que simultaneamente englobe todos os movimentos eletrônicos, bem como o efeito que um elétron de um orbital atômico exerce no comportamento de um outro elétron em um outro orbital atômico devido a estes movimentos eletrônicos.

A solução para este problema pode ser encontrada na equação de Hartree-Fock onde se assume que cada elétron se move no campo médio produzido pelos demais elétrons e pelos núcleos [HOUSE, 2004]. Esta equação possui um operador (operador de Fock) que estabelece uma dependência entre a equação de cada orbital com os outros orbitais e os spin-orbitais moleculares também dependem das soluções da própria equação. Sendo assim, as equações devem ser resolvidas de forma acoplada, através de aproximações sucessivas (solução iterativa). As energias totais correspondentes a cada conjunto são comparadas até que a energia não varie entre duas estruturas consecutivas, dentro de um limite pré-determinado. A melhor função de onda é obtida de tal forma que a energia seja um mínimo.

Uma aproximação para representar a função de onda eletrônica em sistemas moleculares é usar uma combinação linear de funções de onda de um elétron estendidas por toda a molécula, estas funções são chamadas de orbitais moleculares (OM), e esta aproximação é chamada de LCAO (combinação linear dos orbitais atômicos). A distribuição eletrônica em uma molécula segue os mesmos princípios da distribuição em átomos, ou seja, cada OM pode conter no máximo 2 elétrons, com spins opostos (princípio de Pauli) [MORGON & COUTINHO, 2007].

Nos sistemas de camada fechada, os spins dos elétrons emparelhados se anulam uns aos outros por estarem ocupando o mesmo orbital molecular, logo, não é necessário considerar a questão do spin eletrônico na resolução da equação de Schrödinger. Para sistemas com elétrons desemparelhados, como é o caso dos radicais livres, é necessário tratar de alguma forma a questão do spin. A forma mais comum é dividir os OM em 2 grupos, chamados alfa e beta, que são ocupados separadamente por elétrons com apenas um determinado spin. Esse método é chamado de UHF (*Unrestricted Hartree-Fock*). Por exemplo, para um sistema com dois elétrons desemparelhados (estado tripleto), o número de orbitais alfa ocupados é duas unidades maior do que o número de orbitais beta. [SANT'ANNA, 2009].

2.10 Métodos Semi-Empíricos

Nos métodos semi-empíricos uma grande parte das integrais referentes à equação de Schrödinger são negligenciadas devido à introdução de parâmetros empíricos ou previamente calculados que permitem a eliminação do cálculo de algumas dessas integrais [SANT'ANNA, 2009]. Ou seja, as equações integro-diferenciais são trocadas por equações algébricas. Desta forma, o método semi-empírico é o mais indicado no estudo de grandes sistemas como as proteínas e seus complexos com outras moléculas, uma vez que métodos *ab initio* requerem maior capacidade computacional e tempo de máquina.

Em geral, os métodos semi-empíricos limitam-se ao tratamento dos elétrons de valência, não considerando explicitamente os elétrons do cerne, que contribuem muito menos, em geral, para o comportamento das moléculas [ALCÁCER, 2007]. Estes métodos procuram resolver, de forma autoconsistente, as equações de Hartree-Fock, utilizando parâmetros, obtidos por ajuste numérico ou derivados de resultados experimentais, para tornar mais rápida a resolução das equações. Na maioria das vezes, utilizam um conjunto de base mínimo de valência, formado por funções do tipo Slater (STO – *Slater type orbitals*), e fazem uso da teoria de orbitais moleculares para a construção da função de onda molecular, construindo estes orbitais moleculares a partir de da aproximação LCAO.

Existe uma grande variedade de métodos semi-empíricos, nos quais se fazem várias aproximações adicionais de modo a reduzir as dificuldades de cálculo. A resolução da equação de Schrödinger envolve o cálculo de uma série de integrais e a maneira mais fácil de simplificar esta resolução é o uso de aproximações que desconsideram algumas destas integrais. Dentre estas aproximações, uma das que têm impacto no custo computacional é conhecida como aproximação ZDO (*Zero Differential Overlap*). Esta aproximação ignora as integrais de sobreposição de dois elétrons que se encontram em orbitais atômicos diferentes, simplificando bastante as equações de Hartree-Fock, mas que pode levar a grandes erros nas várias propriedades moleculares calculadas, ficando então seu uso restrito ao tratamento de sistemas □ de moléculas planares [MORGON & COUTINHO, 2007]. Afim de que os métodos semi-empíricos pudessem tratar outros tipos de sistema, Pople e colaboradores propuseram novos métodos: CNDO (*Complete Neglect of Differential Overlap*), INDO (*Intermediate Neglect of Differential Overlap*) e NDDO (*Neglect of Diatomic Differential Overlap*). Estes métodos possuem níveis de aproximações diferentes e levam em consideração todos os elétrons de valência. A aproximação integral NDDO é a que considera um maior

número de integrais para serem calculadas e, desta forma é a melhor das três aproximações integrais [POPLE *et al.*, 1967]

Os métodos semi-empíricos paramétricos apoiam-se na proposta de Dewar, na qual a parametrização é feita de modo a reproduzir da melhor forma possível medidas experimentais. São selecionados conjuntos de átomos que pertencem a conjuntos de moléculas cujas entalpias de formação padrão, geometria molecular e momento dipolar sejam conhecidos experimentalmente. A determinação do melhor conjunto de parâmetros é feita por minimização dos erros nos cálculos de entalpia de formação, geometria e momento dipolar [ALCÁCER, 2007].

Como exemplos de métodos semi-empíricos, têm-se o MNDO (Modified Neglect of Diatomic Overlap) [DEWART & THIEL, 1977], AM1 (Austin Model 1) [DEWAR *et al.*, 1985], PM3 (Parametric Method 3) [STEWART, 1989], o MNDO/d (Modified Neglect of Diatomic Overlap, incluindo orbitais d) [THIEL & VOITYUK, 1992a; THIEL & VOITYUK, 1992b] e, mais recentemente, os métodos RM1 [ROCHA *et al.*, 2006], PM6 [STEWART, 2007] e PM7 [<http://openmopac.net/index.html>].

A extensão dos métodos semi-empíricos para metais de transição enfrentou algumas dificuldades [STEWART, 2007]: a escassez de dados de referência e o fato de que é comum que compostos de metais de transição apresentem camada aberta; é mais difícil resolver as equações do método auto consistente para sistemas de camada aberta porque o *spin* total é diferente de zero e precisa ser considerado de algum modo na resolução (como através do método UHF, descrito anteriormente). Finalmente, estudos preliminares indicam que cálculos *ab initio* de alto nível para metais de transição não são úteis para gerar dados de referência termoquímica, embora sejam úteis para gerar dados de geometrias [MORGON & COUTINHO, 2007].

A maioria dos grandes avanços obtidos na pesquisa de métodos semi-empíricos foi essencialmente incorporada em um único programa: o MOPAC (Molecular Orbital PACkage). Iniciado em 1983, no grupo do Prof. M. Dewar, com a participação do Prof. J. Stewart, que também é o responsável pelo desenvolvimento de muitos métodos paramétricos bem sucedidos [STEWART, 1996].

O método PM6 foi o primeiro disponibilizado no programa MOPAC que podia ser utilizado para tratar sistemas contendo metais de transição [STEWART, 2007]. Neste método, as entalpias de formação são mais bem representadas e as geometrias apresentam um significativo aumento da exatidão, quando comparados com métodos paramétricos mais

antigos. O PM6 introduz muitas modificações no termo de interação NDDO cerne-cerne e no método de otimização paramétrico. O erro médio para as entalpias de formação calculadas pelo PM6 de 1373 moléculas usadas como referência, contendo elementos relevantes biologicamente, é de apenas 4,4 kcal/mol, superando inclusive alguns métodos menos aproximados. O erro médio equivalente para outros métodos é: RM1, 5,0; B3LYP 6-31G(d), 5,2; PM5, 5,7; PM3, 6,3; HF 6-31G(d), 7,4; e AM1, 10,0 kcal/mol. Para o mesmo conjunto de dados, o método PM7, publicado recentemente [STEWART, 2012], tem erro médio de 4,0 kcal/mol. Nele, alguns erros na teoria NDDO que afetavam os resultados foram removidos e todos os parâmetros atômicos e diatômicos foram reotimizados [<http://openmopac.net/index.html>].

2.11 Docking

Os sítios ativos das enzimas e os sítios de interação presentes em receptores exibem alta seletividade no reconhecimento de seus bioligantes. Se de alguma forma estas interações pudessem ser explicadas e preditas por algum método a modelagem de moléculas inéditas seria realizada de forma muito mais eficiente. Com isso surgiu a técnica de atracamento molecular, também conhecida como *docking* (ou ancoramento) molecular, na década de 1980 [KUNTZ *et al.*, 1982], e atualmente tem sido bastante utilizada quando as estruturas dos sítios ativos e das moléculas estão disponíveis, ou ainda para o atracamento entre duas proteínas. A expansão do uso do *docking* na literatura tem crescido a cada ano (figura 9), tanto em trabalhos que tratam apenas do *docking*, quanto em trabalhos que usam esta técnica juntamente com a síntese e com testes biológicos, como demonstrado na pesquisa feita por Yuriev e Ramsland [YURIEV & RAMSLAND, 2013].

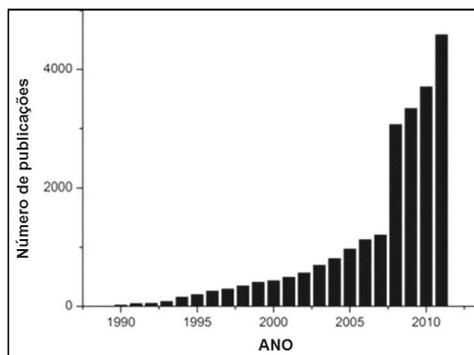


Figura 9. Número de publicações envolvendo *docking*, baseada em simples pesquisa da literatura [YURIEV & RAMSLAND, 2013].

Muitos programas de *docking* empregam métodos de inteligência artificial, como algoritmos genéticos (AG), que são particularmente úteis na resolução de problemas com um vasto número de variáveis, por possibilitarem uma amostragem eficiente [LINDEN, 2006]. Este algoritmo foi proposto no anos 60 por John H. Holland, pesquisador da Universidade de Michigan (EUA), tal metodologia foi primeiramente utilizada em Química por Lucasius e Kateman (1991), em um trabalho para seleção de comprimentos de onda em espectros de ultravioleta. O AG otimiza as soluções encontradas em várias situações, sendo considerado um dos melhores algoritmos de otimização estocástica, encontrando soluções muito próximas às verdadeiras (as encontradas em métodos experimentais, como a cristalografia de raios-X) [COSTA FILHO & POPPI, 1999; MORRIS *et al.*, 1998].

Os AG simulam matematicamente ideias genéticas, para a sua criação são necessárias 5 etapas básicas: a codificação das variáveis, a criação da população inicial, a avaliação da resposta (aptidão), o cruzamento (“*crossover*”) e a mutação. A codificação das variáveis consiste na criação dos “cromossomos”, formados por possíveis soluções (ou seja, pelos arranjos e orientações de um ligante frente a uma proteína/sítio ativo), que podem ser definidas por um grupo de valores (geralmente por codificação binária) descrevendo a translação, orientação e conformação do ligante atracado, ou seja, seu genótipo; enquanto as coordenadas atômicas de um ligante correspondem ao seu fenótipo. Com isto, temos a criação da população inicial, constituída por vários cromossomos. Quanto maior é a diversidade da população, maior o número de soluções diferentes. Se a população inicial foi gerada usando alguma técnica heurística para obtenção de indivíduos de boa qualidade, o AG pode chegar a soluções mais rapidamente quando comparado ao processo de geração aleatória de indivíduos [REEVES, 1995].

Sua aptidão é então calculada a partir dos valores descritos acima, baseados no mérito relativo das soluções encontradas, para saber quais cromossomos serão capazes de gerar a melhor resposta esperada, determinando com qual probabilidade este cromossomo irá cruzar. Pares randômicos de cromossomos são então cruzados, onde novas gerações de cromossomos herdarão características de seus “pais”. Alguns passarão pelo processo de mutação, onde ocorrem permutações nas informações contidas nos genes. A seleção da descendência de tais gerações é baseada na aptidão do cromossomo: aqueles que melhor atracam ao seu ambiente (proteína) são selecionados, e aqueles que não se adaptam bem são abandonados [COSTA FILHO & POPPI, 1999; MORRIS *et al.*, 1998; JONES *et al.*, 1997].

As funções de escore (“*scoring functions*”) ou funções de pontuação mensuram o grau de aptidão do indivíduo. Elas possuem termos que descrevem as interações estéricas, eletrostáticas, de ligações hidrogênio e hidrofóbicas e baseado nestes termos, classificam os melhores modos de ligação (orientação e conformação) do ligante no sítio receptor. Na prática, essas técnicas foram desenvolvidas para serem aplicadas na busca dos chamados posicionamentos, ou “*posing*”, que são combinações de orientações e conformações do ligante no interior da proteína [MORGON & COUTINHO, 2007]. Existem três tipos de funções de escore: as baseadas em campos de força (“*force-field-based scoring functions*”), as empíricas (“*empirical scoring functions*”) e as baseadas em conhecimentos (“*knowledge-based scoring functions*”). [KITCHEN *et al.*, 2004].

As funções de escore baseadas em campos de força quantificam a soma de duas energias: a de interação entre o receptor e o ligante (energia “externa”) e a energia “interna” do ligante. Por ser um método da Mecânica Molecular, são também investigadas as interações não-ligadas envolvidas no reconhecimento molecular. As interações entre o ligante e a proteína em jogo são quase sempre descritas utilizando termos de energia eletrostática e de van de Waals.

O termo eletrostático é uma formulação de Coulomb, com uma função dielétrica dependente de distância, para o tratamento de interações carga-carga. Funções de escore deste tipo possuem complicações, sendo necessária a definição de um “raio de corte” para o tratamento destas interações não-ligadas. Distâncias muito grandes complicam a acurácia do método, diminuindo a possibilidade de atracamento nestas posições muito distantes. Muitos destes escores foram originalmente feitos para modelar contribuições estruturais e entálpicas em fase gasosa, não incluindo termos entrópicos e de solvatação [KITCHEN *et al.*, 2004; SCHNEIDER & BÖHM, 2002]. Várias funções são baseadas em diferentes parâmetros de campos de força: por exemplo, o “*G-Score*” é baseado no campo de força Tripos [KRAMER *et al.*, 1999] e o “*AutoDock*” [MORRIS *et al.*, 1998] é baseado no campo de força AMBER [WEINER *et al.*, 1986]. Algumas funções possuem extensões dos termos de energia para englobar termos entrópicos e outros. Como exemplo, temos a inclusão de termos de entropia torsional ao “*GScore*” e “*Goldscore*” [VERDONK *et al.*, 2003], e a inclusão de termos para ligação hidrogênio entre ligante e proteína no “*Goldscore*” [VERDONK *et al.*, 2003] e “*AutoDock*” [MORRIS *et al.*, 1998].

As funções de escore empíricas reproduzem dados experimentais, como energias de interação e conformacionais, como uma soma de várias funções parametrizadas,

primeiramente propostas por Böhm (1992), baseado na aproximação das energias de interação por uma soma de termos individuais não-correlacionados. Tais termos são obtidos de uma análise por regressão, usando energias de interação determinadas experimentalmente, com informações estruturais (obtidas por cristalografia de raios-X). Suas formas funcionais são frequentemente mais simples que as funções de escore obtidas por campos de força, com termos de avaliação mais simples. Sua desvantagem reside na dependência dos conjuntos de dados moleculares, o que leva a diferentes fatores (quando comparados a funções de campos de força), que não podem ser facilmente recombinaados em uma nova função de aptidão [KITCHEN *et al.*, 2004]. Para evitar isso, termos que contabilizam interações não-ligadas podem ser implementados de várias formas. Na formulação de “*LUDI*” [BÖHM, 1992], os termos de ligação hidrogênio são separados em ligações hidrogênio neutras e ligações hidrogênio iônicas, enquanto o “*ChemScore*” [ELDRIDGE *et al.*, 1997] não diferencia os tipos de ligação hidrogênio. “*LUDI*” calcula as contribuições hidrofóbicas com base na representação da área superficial molecular, enquanto “*ChemScore*” avalia o contato entre pares de átomos hidrofóbicos [ELDRIDGE *et al.*, 1997].

E por último, também existem as funções de escore baseadas em conhecimento, que foram feitas para reproduzir estruturas experimentais ao invés de energias de ligação. Os complexos ligante-proteína são modelados usando potenciais de interação atômica relativamente simples. Assim como os métodos empíricos, incluem implicitamente efeitos de ligação difíceis de incluir explicitamente. Implementações destas funções incluem o “*PMF*” (“*Potencial of Mean Force*”) [MUEGGE & MARTIN, 1999] e o “*DrugScore*” [GOHLKE *et al.*, 2000], que também incluem correções como a acessibilidade ao solvente, dentre outras. Tais funções primam pela simplicidade, fazendo com que grandes conjuntos de moléculas possam ser avaliadas com rapidez. A desvantagem deste método reside no fato de que sua derivação é baseada na informação codificada implicitamente em conjuntos limitados de complexos ligante-proteínas.

No caso do *docking* ligante-proteína, a etapa de reconhecimento molecular entre o ligante e a macromolécula é um processo complexo, pois envolve fatores entálpicos e entrópicos e que incluem a flexibilidade do ligante e da proteína, o efeito do ambiente protéico na distribuição de cargas do ligante e possíveis interações com moléculas de água presentes no meio [ALONSO; BLIZNYUK; GREARY, 2006]. Assim, em razão desses fatores, é necessário avaliar qual função de escore se adequa melhor para cada sistema ligante-proteína a ser estudado.

O primeiro programa de *docking* foi criado em 1982, por Kuntz, sendo chamado de DOCK. Atualmente, existem inúmeros programas de *docking* comerciais e gratuitos amplamente empregados (Figura 10), tais como GOLD, FlexX, AutoDock, Glide, ICM, entre outros [KLEBE, 2006], sendo o mais recente o servidor DockThor, desenvolvido pelo Grupo de Modelagem Molecular de Sistemas Biológicos do LNCC (Laboratório Nacional de Computação Científica) para *docking* gratuito receptor-ligante (<http://www.dockthor.lncc.br>).

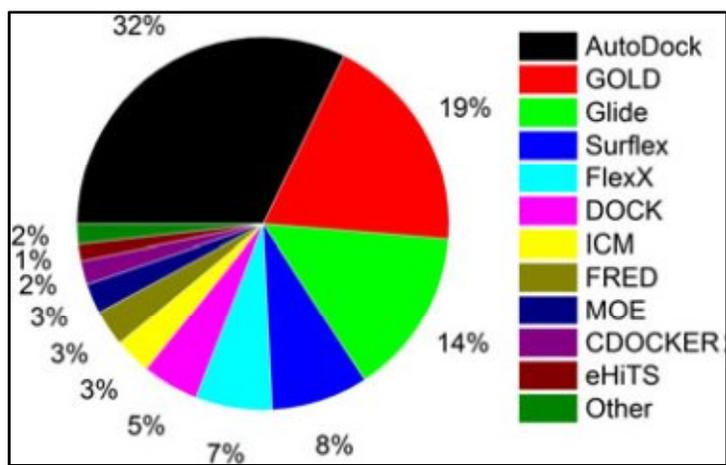


Figura 10. Programas de *Docking* citados nos títulos de artigos ou resumos em 2010–2011, pesquisa realizada na PubMed [YURIEV & RAMSLAND, 2013].

3. OBJETIVOS

A fim de se aperfeiçoar a síntese de novos organofosforados inibidores da urease fazendo uso da modelagem molecular, este trabalho buscou:

- Estudar o sítio ativo da urease, definindo o estado de spin dos íons Ni^{2+} e o estado de protonação do átomo de oxigênio localizado entre esses íons e o efeito da constante dielétrica nesta enzima.
- Identificar e calcular parâmetros teóricos que estejam associados com o processo de inibição da urease, como a entalpia de interação, perda entrópica e a energia livre de Gibbs necessária para o transporte do inibidor da fase aquosa para o interior da enzima, baseando-se na proposta de Wang e colaboradores (1994) de que tais parâmetros teóricos podem ser correlacionados com a atividade inibitória obtida experimentalmente, através de um modelo matemático adequado.
- Estabelecer uma relação quantitativa entre estes parâmetros com a atividade inibitória experimental dos ácidos *P*-metil fosfínicos [VASSILIOU et al. 2008] e dos *P*-hidroximetil fosfínicos [VASSILIOU et al. 2010], com o intuito de se propor um modelo de predição de atividade para novos inibidores.
- Utilizar o modelo proposto para a predição de atividade relativa de uma série de novos organofosforados, propostos pelo grupo de síntese do Dr. João Batista Neves da Costa (PPGQ/UFRRJ).
- Propor algumas modificações estruturais nesses compostos e avaliá-los segundo o modelo desenvolvido.

4. METODOLOGIA

4.1 Estudo da multiplicidade de spin dos íons níquel(II) e do estado de protonação do átomo de oxigênio entre os dois íons níquel(II).

Como primeira etapa do trabalho, foi necessário um estudo para definir a multiplicidade de spin dos íons níquel(II) do sítio ativo da urease e do estado de protonação do átomo de oxigênio entre esses íons. Foi realizada uma pesquisa no *Protein Data Bank* (PDB) a fim de selecionar algumas estruturas cristalográficas da urease que continham o átomo de oxigênio entre os átomos de níquel presentes no sítio ativo da enzima. Foram encontradas duas ureases com essa característica (códigos 1FWJ [PEARSON *et al.*, 1997] e 1EJW do PDB). O *Protein Data Bank* é um banco de dados da internet onde ficam depositadas as estruturas cristalográficas de diversas proteínas obtidas experimentalmente por difração de raios-X ou de nêutrons ou por ressonância nuclear magnética (RMN). Neste banco de dados os arquivos foram obtidos no formato *pdb*, onde são dadas as coordenadas cartesianas para os átomos. Como na estrutura cristalográfica, em geral, não são determinadas as posições dos átomos de hidrogênio e nem sua quantidade, foi necessário avaliar o estado de protonação do átomo de oxigênio do sítio ativo, ou seja, avaliar se o átomo de oxigênio correspondia a uma molécula de água ou um íon hidroxila.

Os arquivos fornecidos pelo PDB foram abertos no programa RasMol [BERNSTEIN *et al.*, 1999]. Os átomos presentes nestes sítios foram selecionados dentro de um raio de 6.0 Å ao redor dos átomos de níquel, utilizando-se o comando “*select within (6.0, Ni)*” e gravando-se a seleção no formato *pdb*. Para que todos os resíduos de aminoácidos selecionados estivessem completos no sítio ativo, editou-se essa primeira seleção para se encontrar os nomes e numerações dos aminoácidos e moléculas de água presentes no sítio. Com essa informação, retornou-se ao programa RasMol para ser feita a seleção nominal destes aminoácidos para a construção do sítio.

Para realizar os cálculos de otimização de energia, foram adicionados os átomos de hidrogênio na estrutura e o formato do arquivo foi convertido de *pdb* para o de coordenadas internas (distâncias de ligação, ângulos de ligação e ângulos diedros) do programa Mopac (extensão *mop*). Estas operações foram realizadas no programa Babel [<http://openbabel.org>], em ambiente DOS, utilizando a opção “-h” no comando de conversão para a inserção dos hidrogênios.

O arquivo fornecido foi aberto num editor de textos para avaliação e remoção dos átomos de hidrogênio que foram adicionados em excesso em resíduos do sítio ativo. Através deste procedimento, foram preparados os arquivos referentes ao sítio contendo água ou contendo íon hidroxila entre os átomos de níquel.

Os cálculos realizados para as minimizações foram feitos com o programa Mopac 2009 [Stewart Computational Chemistry], com o método semi-empírico PM6 [Stewart, 2007], em duas etapas: na primeira foram otimizados apenas os átomos de hidrogênio através das palavras-chave “noopt” e “opt-h”. Na segunda etapa, os átomos da cadeia peptídica foram fixados no espaço alterando as “*flags*” das três coordenadas internas do sistema de 1 para 0. Posteriormente, foi realizada a minimização de energia da estrutura, onde todos os átomos do sítio ativo foram otimizados, com exceção dos átomos da cadeia polipeptídica. Cálculos quânticos requerem a definição exata da multiplicidade do sistema. Isso é particularmente difícil no sistema em estudo, que contém dois íons Ni^{2+} com configuração eletrônica d^8 , mas em ambientes químicos não idênticos (geometrias piramidal-quadrada distorcida e octaédrica, conforme descrito no Item 2.4) e com possíveis desdobramentos diferentes. Através da análise das estruturas cristalográficas das três ureases selecionadas para esse teste, observou-se que o Ni1 descrito no Item 2.4 também se coordena a um resíduo de glicina, podendo-se então visualizar uma geometria octaédrica para ambos os íons Ni^{2+} . Dessa forma, a distribuição dos elétrons nos orbitais d dos íons de Ni^{2+} podem assumir três multiplicidades de spin distintas, conforme esquematizado na figura 11. Por isso, decidimos realizar os cálculos com a enzima nesses 3 estados: singlete (com o método RHF), tripleto e quinteto (ambos com o método UHF, usando-se as palavras-chave “UHF QUINTET” e “UHF TRIPLET”, respectivamente), contendo tanto água quanto um ânion OH^- entre os íons Ni^{2+} .

Após o término dos cálculos, os arquivos de saída contendo a estrutura já otimizada foram abertos com o programa Rasmol, onde se conferiu a estrutura e principalmente a posição dos átomos de hidrogênio. Os átomos de hidrogênio podem migrar de um aminoácido para outro durante a otimização, principalmente se esta for realizada no vácuo. O uso do método COSMO (*Conductor-like Screening Model*) [RZEPA *et al.*, 1991] com a constante dielétrica da água (palavra-chave EPS=78.4) e de uma constante que mimetiza o interior de proteínas (palavra-chave EPS=4) durante as otimizações evitou que isto acontecesse, além de aproximar mais o modelo das condições reais. O valor 4 é um dos mais utilizados na literatura para a constante dielétrica no interior de proteínas, embora muitos outros sejam sugeridos [LI *et al.*, 2013]

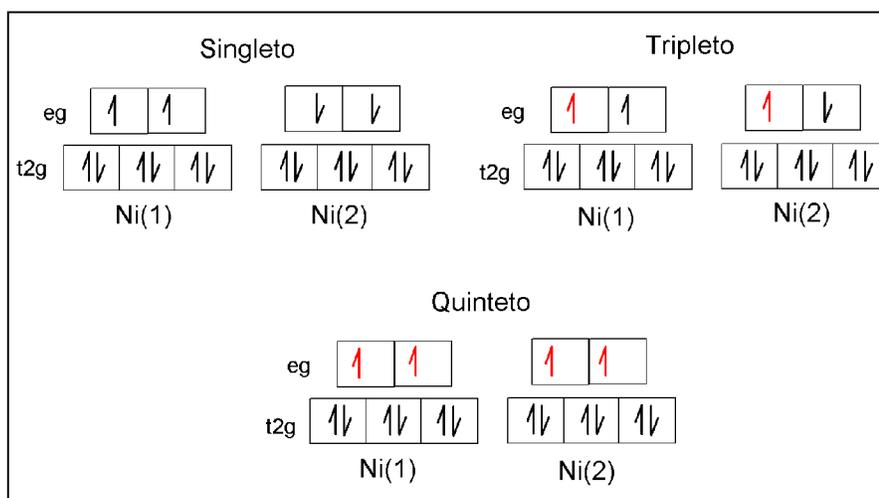


Figura 11. Diagrama simplificado da distribuição dos elétrons nos orbitais d dos íons Ni^{2+} nas multiplicidades estudadas. Os elétrons desemparelhados estão em vermelho para melhor visualização.

Ainda utilizando o programa Rasmol, na opção “*settings*” da barra de ferramentas do programa, foi selecionado o comando “*pick distance*”, que permite medir a distância entre dois átomos selecionados. Foram tomadas doze medidas entre alguns átomos localizados no sítio ativo da urease e os íons Ni^{2+} , tanto das estruturas otimizadas como das estruturas experimentais. Desta maneira, determinou-se qual modelo teórico havia apresentado uma menor variação da posição dos átomos em relação à estrutura experimental, ou seja, reproduzia melhor a estrutura experimental.

4.2 Otimização dos ligantes

Os ligantes da urease usados nesse trabalho podem ser divididos em três séries. Na primeira estão os ligantes sintetizados por Vassiliou e colaboradores em 2008 (figura 12). Na segunda estão os ligantes sintetizados por Vassiliou e colaboradores em 2010 (figura 13). Estas duas primeiras séries correspondem a ligantes que possuem dados de atividade inibitória da urease determinados experimentalmente, necessários para a construção do modelo. Todos possuem hidrogênios ácidos e foram otimizados nas suas formas aniônicas. Na série 3 estão os ligantes sintetizados pelo grupo de pesquisa coordenado pelo Dr. João Batista Neves da Costa (Dequim/UFRRJ), os quais tiveram suas atividades previstas pelo modelo desenvolvido nesse trabalho (figura 14).

Através do programa Spartan'08 [Wavefunction Inc.], os ligantes das três séries mencionadas foram desenhados e otimizados. A otimização dos ligantes referente à série 3 foi realizada pelo método de Monte Carlo com o campo de força *MMFF* (*Merck Molecular Force Field*) [Thomas, 1996], resultando em uma distribuição de conformeros. Após a seleção do conformero de menor energia de cada estrutura, sua geometria foi reotimizada com o método semi-empírico PM3 (*Parametric Method 3*) [STEWART, 1989], também através do programa Spartan'08. A reotimização dos ligantes das duas primeiras séries foi realizada através do método semi-empírico AM1 (*Austin Model 1*) [DEWAR *et al.*, 1985] ao invés do PM3, pois as estruturas otimizadas com este método apresentaram distâncias carbono-fósforo muito altas, o que pode ter sido causado por estas estruturas estarem na forma aniônica. As estruturas foram gravadas no formato “mol2” a fim de serem usadas como arquivos de entrada de dados no estudo de *docking*.

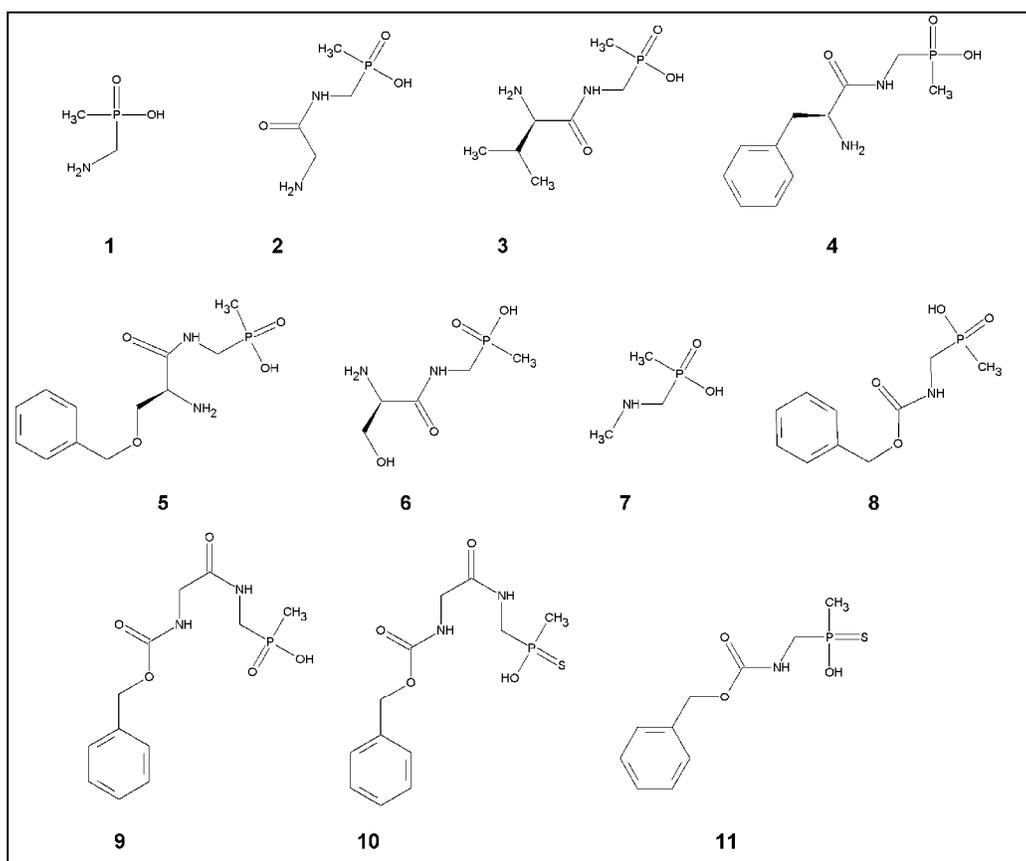


Figura 12. Série 1: Ácidos *P*-metil fosfínicos sintetizados por Vassiliou e colaboradores em 2008 [VASSILIOU *et al.*, 2008].

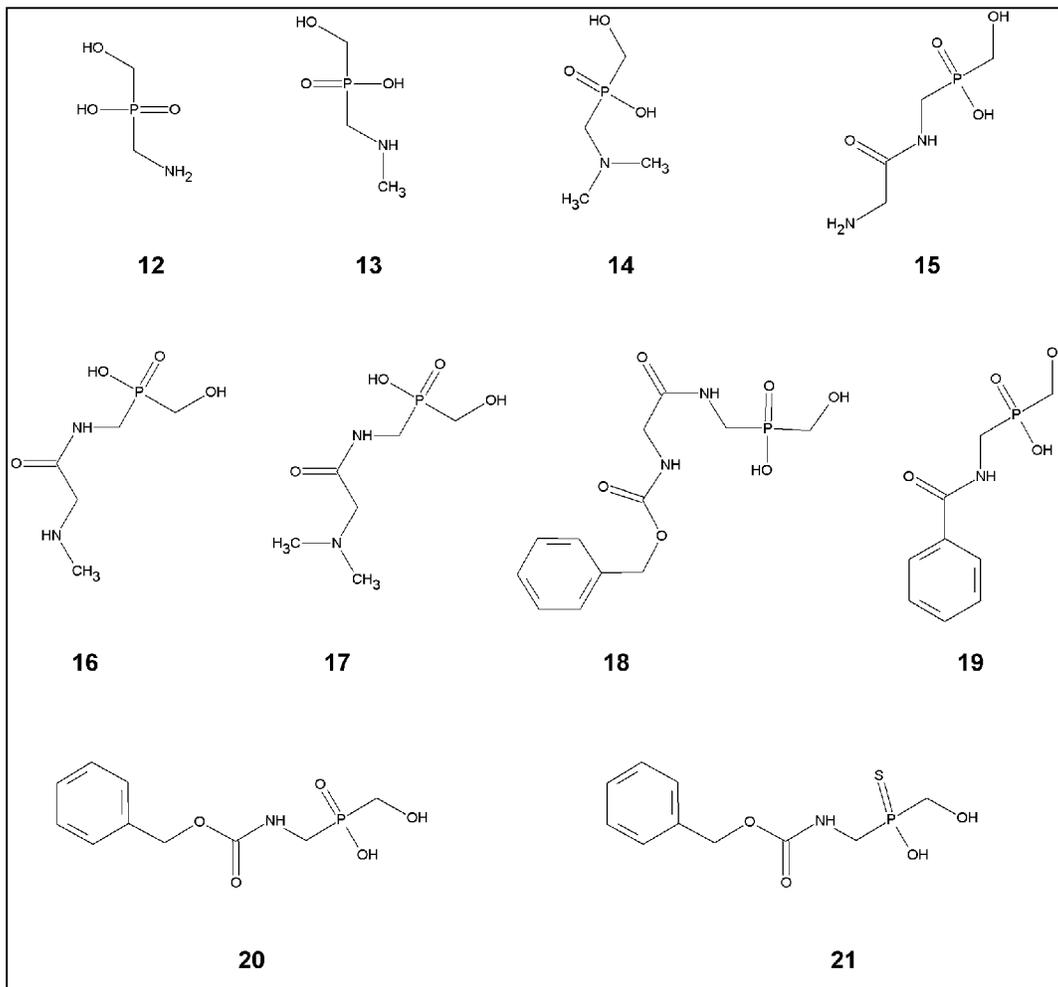


Figura 13. Série 2: Ácidos *P*-hidroximetil fosfínicos, sintetizados por Vassiliou e colaboradores em 2010 [VASSILIOU *et al.*, 2010].

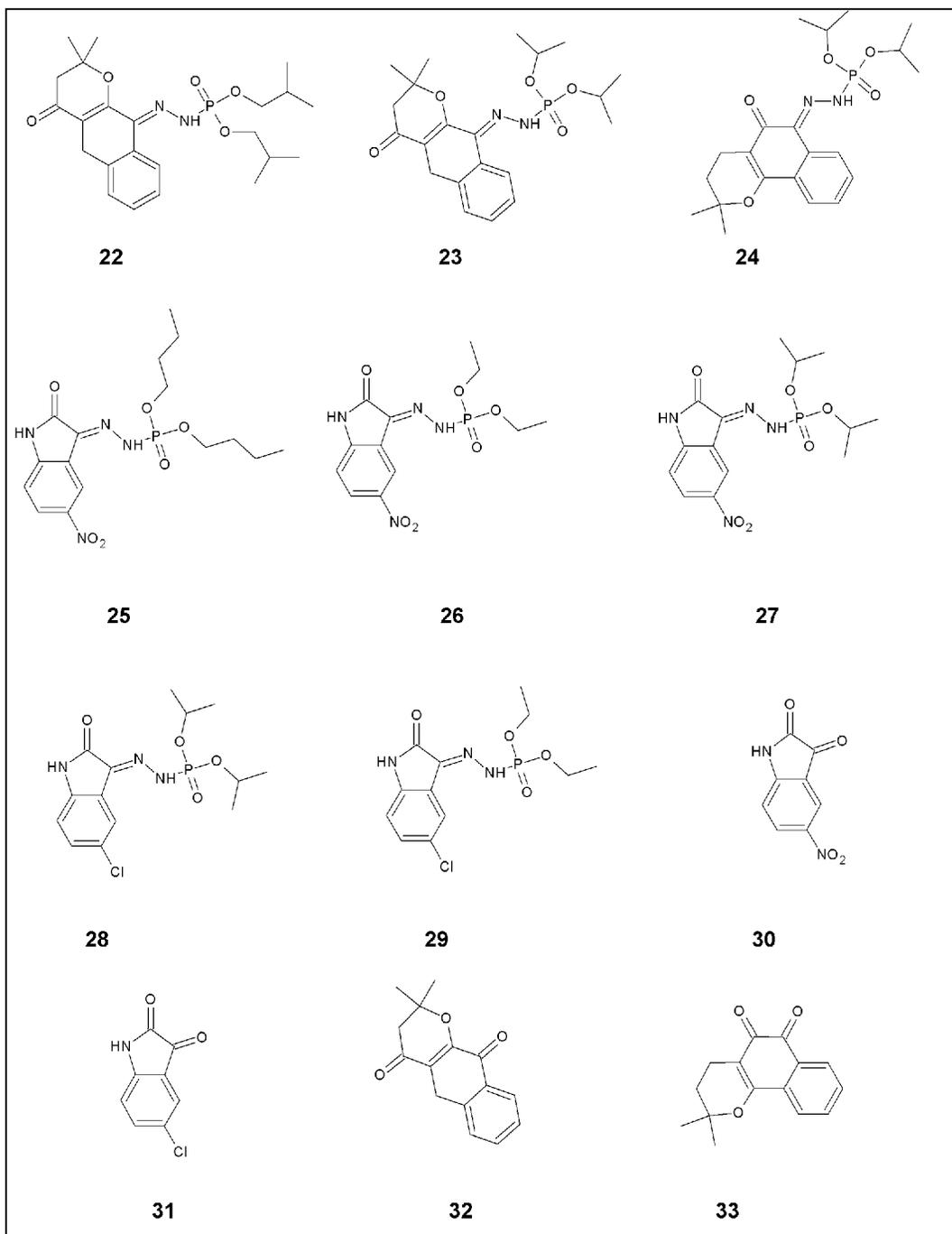


Figura 14. Série 3: Compostos sintetizados pelo grupo do Dr. João Batista Neves da Costa (Dequim, UFRRJ).

4.3 Docking

Segundo os artigos publicados por Vassiliou e colaboradores em 2008 e em 2010, os compostos sintetizados foram testados *in vitro* usando a urease de *B. pasteurii*, obtendo-se dados de CI_{50} e K_i correspondentes. Através da busca no PDB, foi obtida a estrutura cristalográfica da urease de *B. pasteurii* (código 4UBP do PDB) [BENINI *et al.*, 2000]. Esta estrutura possui 1,55 Å de resolução e contém o ligante ácido acetoidroxâmico co-cristalizado.

Antes do procedimento de *docking* dos ligantes começar a ser feito, foram testadas as quatro funções de pontuação disponíveis no programa de *docking* GOLD 5.1 (CCDC): *ChemPLP* [KORB *et al.*, 2007, 2009], *Chemscore* [ELDRIDGE *et al.*, 1997], *Goldscore* [VERDONK *et al.*, 2003], e *ASP* [BONET *et al.*, 1997]. Este programa visa identificar a melhor forma de interação (orientação e conformação) entre a molécula ligante e a proteína, através do uso de um AG. Os hidrogênios foram primeiramente adicionados em toda a estrutura da urease, fazendo-se as correções necessárias após inspeção visual. O ligante foi removido do sítio e foi realizado o *redocking* com cada uma das funções de pontuação.

Para o *docking* dos demais ligantes, o ligante co-cristalizado e as moléculas de água presentes na estrutura foram removidos. O sítio ativo foi definido com um raio 10 Å a partir do átomo de níquel 798. Por fim, fez-se a seleção da melhor interação, ou seja, foi selecionada para cada ligante a solução que apresentou o valor de pontuação mais alto. Através do comando “*export complex*”, cada complexo urease/ligante foi gravado em formato *pdb*.

As estruturas 10 e 11 da série 1 e a estrutura 21 da série 2, possuem um átomo de enxofre ligado ao grupamento P-terminal, segundo Vassiliou e colaboradores (2008), a boa atividade inibitória desses compostos pode ser explicada através da estrutura cristalográfica do β -mercaptoetanol complexado na urease, no qual o átomo de enxofre interage entre os dois íons Ni(II) presentes no sítio ativo. Nos artigos de Vassiliou e colaboradores (2008 e 2010), não havia menção sobre qual forma enantiomérica dessas estruturas foi usada no ensaio biológico de atividade, então o *docking* foi realizado com os dois enantiômeros de cada ligante. Porém no *docking* dos enantiômeros dessas estruturas, o enxofre não ficou localizado entre os íons Ni(II) e sim o átomo de oxigênio. Então se optou por usar o *docking* das estruturas 9, 8 e 20 para gerar os complexos das estruturas 10, 11 e 21, respectivamente, com a urease, substituindo o átomo de oxigênio localizado entre os íons Ni(II) pelo enxofre no arquivo *pdb* gerado após o *docking*. A distância de ligação fósforo-enxofre dessas estruturas

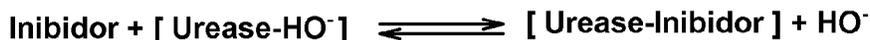
no complexo inibidor-enzima foi corrigida através do método semi-empírico PM6 (programa MOPAC2012) fixando-se toda a enzima e otimizando apenas o inibidor dentro do sítio.

4.4 Cálculo das entalpias de interação inibidor-urease

Os arquivos gravados na etapa do *docking*, contendo os complexos urease/ligante de maiores pontuações, foram abertos no programa RasMol. Os átomos presentes nestes sítios foram selecionados dentro de um raio de 4,5 Å ao redor do ligante e gravou-se a seleção no formato *pdb*. Para que todos os resíduos de aminoácidos selecionados fossem completos, editou-se essa primeira seleção para se encontrar os nomes e numerações dos aminoácidos, átomos de níquel e ligantes presentes nos sítios. Com essa informação, retornou-se ao programa RasMol para ser feita a seleção nominal para a construção do sítio. Os arquivos foram verificados para fazer quaisquer correções necessárias nos átomos de hidrogênio antes dos cálculos quânticos.

Os cálculos realizados para as otimizações foram realizados com o programa MOPAC 2012 [Stewart Computational Chemistry], através do método semi-empírico PM6, em duas etapas: Primeiro otimizou-se apenas os hidrogênios usando as palavras-chave “NOOPT” e “OPT-H”. Como segunda etapa, otimizou-se todo o resto (os hidrogênios, o ligante e as cadeias laterais dos aminoácidos).

Através do uso das palavras-chave “UHF QUINTET” e “EPS=4.0” foram obtidas as entalpias dos sistemas considerando o estado quinteto e a constante dielétrica igual a 4,0. Os valores de entalpia de interação (ΔH_{int}), ou seja, os valores das entalpias envolvidas para a formação complexo enzima-inibidor, foram calculadas com base na seguinte reação:



A presença do íon hidroxila foi usado para se manter o número e o tipo de pares de elétrons de ambos os lados da reação (reação isodêmica), minimizando-se os erros na avaliação da entalpia de reação pelo método semi-empírico. As entalpias de reação entre os inibidores e a urease foram calculadas através da seguinte equação:

$$\Delta H_{int} = \Delta H_{E-I} + \Delta H_{HO} - (\Delta H_{E-HO} + \Delta H_I) \quad (9)$$

onde ΔH_{int} é a entalpia envolvida na interação do inibidor com a enzima; ΔH_{E-I} é a entalpia de formação do complexo enzima-inibidor; ΔH_{HO} é a entalpia de formação do íon hidroxila; ΔH_{E-HO} é a entalpia de formação do complexo enzima-hidroxila; e ΔH_I é a entalpia de formação do inibidor.

4.5 Construção do modelo Empírico para predição da atividade inibitória na urease.

Em sistemas biológicos, com temperatura, pressão e número de partículas constantes, o processo de interação entre uma proteína (um receptor ou uma enzima) e um ligante formando um complexo proteína-ligante é determinado pela energia livre de Gibbs de interação (ΔG_{int}) [BROOIJMANS & KUNTZ, 2003]. Quanto mais negativo o valor de ΔG_{int} , mais favorável é a formação do complexo, que é constituída por termos entálpicos e entrópicos, conforme a equação 10:

$$\Delta G_{int} = \Delta H_{int} - T\Delta S_{int} \quad (10)$$

onde ΔH_{int} é a entalpia de interação, T é a temperatura absoluta do sistema e ΔS_{int} é a entropia de interação.

No sistema composto por solvente, receptor e ligante, antes da complexação, estes dois últimos estão solvatados e podem apresentar entropia translacional, rotacional, vibracional e conformacional elevadas devido à liberdade de movimento. Entretanto, ao se complexarem, têm as suas entropias diminuídas, enquanto a entropia do solvente aumenta de forma mais pronunciada, o que eleva a entropia do sistema [BROOIJMANS & KUNTZ, 2003].

No caso de complexos enzima-inibidor, o valor de ΔG_{int} pode ser correlacionado a medidas experimentais de afinidade como CI_{50} (concentração do inibidor necessária para inibir *in vitro* a atividade enzimática em 50%) e a constante de inibição (K_i) *in vitro*, que mantém uma correlação matemática com ΔG_{int} conforme a seguinte equação:

$$RT \ln K_i = 2,303 RT \log K_i = \Delta G_{int} \quad (11)$$

onde R é a constante geral dos gases ($1,987 \text{ cal.K}^{-1}.\text{mol}^{-1} = 0,001987 \text{ kcal.K}^{-1}.\text{mol}^{-1}$) e T é a temperatura absoluta (298,15 K, que corresponde à temperatura ambiente de 25 °C). É importante ressaltar que esta equação não apresenta o sinal negativo porque a constante de inibição K_i é definida para a reação de dissociação do complexo enzima/inibidor, enquanto o ΔG_{int} refere-se ao processo de associação entre a enzima e o inibidor [COPELAND, 2005].

Dados de K_i e CI_{50} estão disponíveis na literatura para uma série de inibidores da urease [VASSILIOU et al, 2008; VASSILIOU et al, 2010]. Alguns desses dados foram usados para a construção de um modelo para a predição de valores de atividade a partir de dados teóricos, similar às propostas desenvolvidas por Wang e colaboradores no estudo de inibidores de proteína quinase C [WANG et al., 1994] e posteriormente adaptadas por Oliveira e colaboradores no estudo de inibidores da fosfodiesterase 4 [OLIVEIRA et al., 2006] e por Rocha na predição da atividade de inibidores da esterol 14 α -desmetilase (CYP51) [ROCHA, 2009].

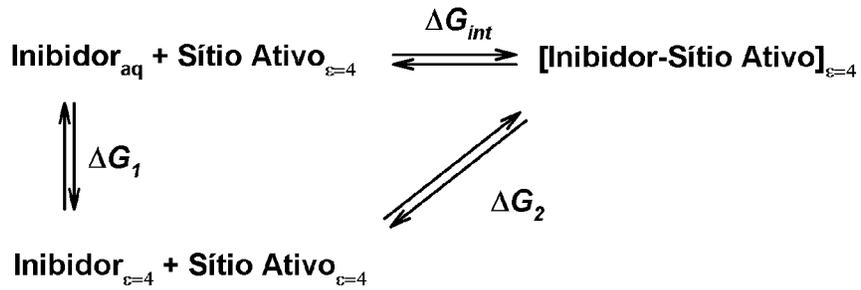
Vassiliou e colaboradores realizaram o ensaio biológico e os estudos de inibição com a urease purificada de *Bacillus pasteurii* CCM2056T, usando concentrações crescentes do inibidor na faixa de 2 a 100 mM. Cada ensaio foi realizado em triplicata. Os valores de K_i da série 1 foram obtidos pelo gráfico de Lineweaver-Burk com pelo menos cinco concentrações inibitórias. Os valores de K_i da série 2 foram calculados usando equações apropriadas do programa GraphPad Prism 5 software. Estes valores estão listados na tabela 1, juntamente com seus logaritmos.

Tabela 1. Valores de K_i (μM) e $\log(K_i)$ para os compostos sintetizados por Vassiliou e colaboradores em 2008 (série 1) e em 2010 (série 2).

Série 1	K_i	$\log(K_i)$	Série 2	K_i	$\log(K_i)$
1	340	2,53	12	95	1,98
2	21	1,32	13	0,43	-0,37
3	120	2,08	14	5,00	0,70
4	215	2,33	15	30,5	1,48
5	25	1,40	16	10,6	1,02
6	37	1,57	17	1,15	0,06
7	18	1,26	18	1,35	0,13
8	43	1,63	19	31,9	1,50
9	135	2,13	20	376	2,57
10	0,17	-0,77	21	428	2,63
11	17,5	1,24			

Os valores de $\log K_i$ listados na tabela 1 podem ser correlacionados com a energia livre de Gibbs de interação entre o inibidor e a enzima (ΔG_{int}) de acordo com a equação 11. Esta interação pode ser dividida em duas etapas: na primeira ocorre o transporte do inibidor da fase aquosa até o interior da proteína, onde se pressupõe que a constante dielétrica seja igual a 4, conforme será discutido posteriormente. Na segunda etapa, ocorre a interação entre o inibidor e o sítio ativo da urease.

Esquema 1



De acordo com o esquema 1, temos a seguinte equação:

$$\Delta G_{int} = \Delta G_1 + \Delta G_2 \quad (12)$$

Substituindo ΔG_{int} na equação 11, temos:

$$RT \ln K_i = 2,303 RT \log K_i = \Delta G_1 + \Delta G_2 \quad (13)$$

O termo ΔG_1 pode ser aproximado como a energia livre de Gibbs necessária para o transporte do inibidor da fase aquosa ($\varepsilon=78,4$) para o interior da enzima ($\varepsilon=4,0$). Este termo foi calculado através da equação 14.

$$\Delta G_1 \cong \Delta G_{78,4-4,0} = (H^{4,0} - TS^{4,0}) - (H^{78,4} - TS^{78,4}) \quad (14)$$

Os valores de entropia (S) e entalpia (H) da equação 14 foram obtidos através do programa MOPAC2012 usando-se a palavra-chave “THERMO” após a otimização dos inibidores com as constantes dielétricas 78,4 e 4,0. Através dessa palavra-chave, fatores termodinâmicos, como a capacidade de calorífica, entropia e energia interna, podem ser calculados usando-se as frequências vibracionais, momentos de inércia, o número de simetria da molécula e a temperatura.

Uma dependência quadrática entre ΔG_1 e o logaritmo de K_i foi proposta por Wang e colaboradores (1994) sob a justificativa de que, como a enzima analisada era uma enzima de membrana, valores extremos de solubilidade prejudicariam a atividade. Valores muito negativos para esse termo poderiam fazer com que o inibidor interagisse muito bem com a membrana ao invés de interagir com a enzima e valores muito positivos para este termo indicariam que a passagem do inibidor do meio aquoso para o interior da enzima não seria muito favorecida.

O termo ΔG_2 é composto de um termo entálpico, ΔH_2 , e um termo entrópico, $-T\Delta S_2$, este último associado a variações nas entropias translacional, rotacional e conformacional dos ligantes pela interação com a cavidade do sítio ativo da urease no interior da proteína. Assumindo-se perdas de entropias translacional e rotacional semelhantes para todos os compostos avaliados, que têm estruturas bastante semelhantes, o termo entrópico referente à ΔG_2 foi considerado como apenas proporcional ao número de ligações rotacionáveis N_{LR} que se tornaram “congeladas” como resultado da interação do inibidor com o sítio ativo, conforme proposto por Wang e colaboradores (1994). A variação da entalpia referente à ΔG_2 foi calculada conforme descrito no tópico anterior (4.3). Logo, pode-se reescrever a equação 13 da seguinte forma:

$$RT \log K_i = c_1(\Delta G_1 + c_2)^2 + c_3\Delta H_{\text{int}} + c_4N_{LR} + c_5 \quad (15)$$

na qual os coeficientes c_1 - c_5 podem ser determinados ajustando-se a equação aos valores experimentais do $\log K_i$ para a série de compostos da literatura, através de uma análise por regressão linear múltipla. O coeficiente c_5 representa o valor onde a reta intercepta o eixo da variável dependente.

Por fim, foram obtidas equações através da regressão múltipla linear dos dados referentes aos ácidos fosfínicos da série 1 e da série 2, separadamente. Avaliou-se, então, quais das duas séries de dados experimentais levavam a um melhor modelo de predição da

atividade. O melhor modelo foi usado para se predizer as constantes inibitórias dos ligantes da série 3 e sua análise possibilitou propor modificações nestas estruturas a fim de aumentar a atividade destes compostos.

5. RESULTADOS E DISCUSSÃO

5.1 Estudo da multiplicidade de spin dos íons níquel(II) e do estado de protonação do átomo de oxigênio entre os íons níquel(II).

Os sítios ativos otimizados das ureases 1FWJ, 1EJW e 1S3T em seus diferentes estados de multiplicidade de spin e protonação foram analisados com o programa RasMol, para se determinar qual estado possuía uma geometria otimizada mais parecida com a geometria da estrutura cristalográfica original. Para se avaliar os resultados, foram escolhidas as distâncias de ligação de alguns átomos capazes de se coordenarem aos íons Ni^{2+} (Figura 15).

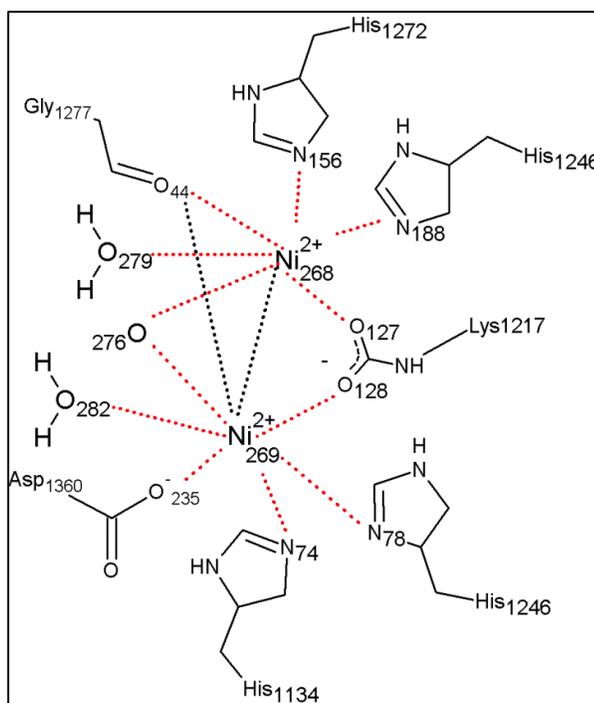


Figura 15. Distâncias escolhidas entre os cátions Ni^{2+} e alguns átomos doadores de elétrons representadas pelas linhas pontilhadas no sítio ativo da urease 1EJW. Cada cátion Ni^{2+} possui seis coordenações, representadas pelas linhas pontilhadas em vermelho.

As distâncias correspondentes entre estes átomos e os íons Ni^{2+} estão nas tabelas 2 e 3 onde d_{exp} é a distância experimental obtida a partir das estruturas cristalográficas, d_{s/HO^-} e $d_{\text{s}/\text{H}_2\text{O}}$ são as distâncias após a otimização, considerando-se o estado singleto, com HO^- e H_2O entre os íons Ni^{2+} , respectivamente; d_{v/HO^-} e $d_{\text{v}/\text{H}_2\text{O}}$ são as distâncias após a otimização, considerando-se o estado tripleto, com HO^- e H_2O entre os íons Ni^{2+} , respectivamente; d_{q/HO^-} e

d_{q/H_2O} são as distâncias após a otimização, considerando-se o estado quinteto, com HO^- e H_2O entre os íons Ni^{2+} , respectivamente.

Tabela 2. Distâncias em Å entre os átomos doadores de elétrons e os átomos de Ni(II) no sítio ativo da urease 1EJW.

Átomos	d_{exp}	d_{s/HO^-}	d_{s/H_2O}	d_{l/HO^-}	d_{l/H_2O}	d_{q/HO^-}	d_{q/H_2O}
Ni268-N188	2,13	1,85	1,86	1,96	1,86	1,97	1,88
Ni268-N156	2,08	1,87	1,88	1,94	1,87	1,93	1,88
Ni269-N78	2,11	1,86	1,87	1,91	1,86	1,94	1,94
Ni269-N74	2,13	1,87	1,88	1,96	1,87	1,94	1,94
Ni269-Ni268	3,69	4,25	4,65	3,62	4,88	3,64	3,82
Ni269-O276	2,18	2,71	3,07	1,96	3,59	2,01	2,20
Ni268-O276	2,1	1,84	1,92	1,98	1,93	1,97	2,02
Ni268-O127	1,98	2,37	3,00	2,00	2,98	2,01	2,17
Ni269-O128	2,06	1,88	1,88	2,05	1,86	2,04	2,04
Ni268-O44	3,33	3,11	2,83	3,34	2,5	3,31	3,10
Ni269-O44	5,92	6,41	6,64	5,91	6,52	5,94	5,94
Ni269-O235	2,02	1,87	1,86	2,01	1,86	2,01	2,00
Ni268-O279	2,11	1,97	1,94	2,26	2,09	2,40	2,03
Ni269-O282	2,25	3,17	3,08	3,22	2,16	3,50	2,13

Tabela 3. Distâncias em Å entre os átomos doadores de elétrons e os átomos de Ni(II) no sítio ativo da urease 1FJW.

Átomos	d_{exp}	d_{s/HO^-}	d_{s/H_2O}	d_{l/HO^-}	d_{l/H_2O}	d_{q/HO^-}	d_{q/H_2O}
Ni258-Ni259	3,59	4,18	4,65	3,66	3,88	3,64	3,83
Ni258-N157	2,02	1,88	1,88	1,94	1,89	1,94	1,9
Ni258-N189	2,28	1,85	1,86	1,96	1,88	1,96	1,91
Ni258-O263	2,15	1,83	1,92	1,97	2,03	1,97	2,04
Ni259-O263	1,97	2,62	3,07	1,95	2,23	1,96	2,2
Ni259-N87	2,14	1,88	1,87	1,91	1,94	1,91	1,94
Ni259-N74	2,26	1,88	1,88	1,98	1,94	1,97	1,95
Ni258-O127	2,06	2,29	3,00	2,00	2,18	2,00	2,04
Ni259-O128	3,34	1,88	1,88	2,04	2,04	2,05	2,05
Ni258-O44	3,36	3,21	2,83	3,42	3,17	3,44	3,32
Ni259-O44	5,92	6,47	6,64	6,06	6,06	6,04	6,05
Ni259-O232	2,17	1,86	1,86	1,99	1,98	1,99	1,98
Ni258-O266	2,09	2,00	1,94	2,41	2,05	2,47	2,12
Ni259-O269	2,12	3,09	3,48	3,76	2,16	3,78	2,18

Os valores da raiz do desvio médio quadrático (RMSD) são mostrados na tabela 4; os valores em negrito correspondem à condição de menor RMSD de cada urease. Pode-se observar que os resultados obtidos com a otimização estão mais próximos das estruturas experimentais nos estados tripleto e , porém o estado quinteto foi escolhido devido ao padrão hexacoordenado de cada íon Ni²⁺ observado nas estruturas cristalográficas.

Em relação à protonação do átomo de oxigênio entre os íons Ni²⁺, o mais provável é o estado aniônico (HO⁻).

Tabela 4. Valores da raiz do desvio médio quadrático (RMSD)* entre átomos doadores de elétrons e os cátions Ni²⁺.

Urease	singleto	singleto	tripleto	tripleto	quinteto	quinteto
	RHF/HO ⁻	RHF/H ₂ O	UHF/HO ⁻	UHF/H ₂ O	UHF/HO ⁻	UHF/H ₂ O
1EJW	0,34	0,56	0,12	0,66	0,13	0,14
1FWJ	0,55	0,72	0,39	0,41	0,39	0,39

$$*RMSD = \sqrt{\frac{1}{12} \sum_{i=1}^N (d_{\text{exp}} - d_{\text{teorica}})^2}$$

Nos cálculos semi-empíricos realizados para a definição da multiplicidade de spin e do estado de protonação foi considerado a constante dielétrica da água ($\epsilon=78,4$). Como a constante dielétrica pode ser um fator importante nos cálculos, foi realizado um estudo adicional otimizando-se o sítio da urease 1FWJ no estado quinteto e com OH⁻ com a constante dielétrica do vácuo ($\epsilon= 1,0$ a 298K) e com a constante dielétrica 4,0, um valor comumente usado para o interior de proteínas [FONG *et. al*, 2011; ESTIU & MERZ, 2007; CARLSSON & NORDLANDER, 2010].

As estruturas foram analisadas a fim de se determinar em qual caso haveria uma melhor reprodução da estrutura cristalográfica inicial. Para a análise dos resultados, foram escolhidos os mesmos átomos capazes de se coordenarem aos íons Ni²⁺, usados anteriormente (Figura 15). Através do programa RasMol, obteve-se o valor destas distâncias de ligação. As distâncias em Å entre estes átomos e os íons Ni²⁺ estão listadas na tabela 5, onde d_{exp} é a distância experimental obtida a partir da estrutura cristalográfica da urease 1FWJ; $d_{\epsilon=4,0}$ é a

distância obtida após a otimização considerando $\epsilon=4,0$; $d_{\epsilon=1,0}$ é a distância obtida após a otimização considerando $\epsilon=1,0$.

Tabela 5. Distâncias em Å entre os átomos doadores de elétrons e os átomos de Ni(II) no sítio ativo da urease 1FWJ.

1FWJ	d_{exp}	$d_{\epsilon=4,0}$	$d_{\epsilon=1,0}$
Ni258-Ni259	3,59	3,85	3,88
Ni258-N157	2,02	1,94	1,95
Ni258-N189	2,28	1,93	1,94
Ni258-O263	2,15	2,11	2,09
Ni259-O263	1,97	2,17	2,20
Ni259-N87	2,14	1,94	1,96
Ni259-N74	2,26	1,95	1,96
Ni258-O127	2,06	2,00	2,01
Ni259-O128	3,34	2,08	2,03
Ni258-O44	3,36	3,33	3,30
Ni259-O44	5,92	6,05	6,04
Ni259-O232	2,17	1,98	1,96
Ni258-O266	2,09	2,10	2,08
Ni259-O269	2,12	2,14	2,18

Os valores da raiz do desvio médio quadrático (RMSD) são mostrados na tabela 6, e como pode ser observado, a constante dielétrica parece não exercer nenhuma influência na estrutura do sítio da urease, visto que não houve grande diferença entre a estrutura nos dois casos ($\epsilon=4,0$ e $\epsilon=1,0$). No entanto, a constante dielétrica do meio no qual está inserida a enzima se faz relevante no processo de inibição enzimática, no qual, o ligante terá que ser transportado da solução aquosa para um meio dielétrico diferente.

Tabela 6. Valores da raiz do desvio médio quadrático (RMSD)* entre átomos doadores de elétrons e os cátions Ni^{+2} .

Constante dielétrica (ϵ)	4,0	1,0
<i>RMSD</i>	0,41	0,42

$$*RMSD = \sqrt{\frac{1}{12} \sum_{i=1}^N (d_{\text{exp}} - d_{\text{teorica}})^2}$$

5.2 Escolha da melhor função de pontuação para o *docking* de ligantes na urease.

Através da figura 16 podemos observar que todas as funções do programa GOLD conseguiram reproduzir muito bem a estrutura experimental do ácido acetohidroxâmico complexado no sítio ativo da urease. Porém, a função *ChemPLP* foi escolhida pois apresentou um menor valor de RMSD em relação às demais funções testadas, embora todas tenham tido um bom desempenho (tabela 7).

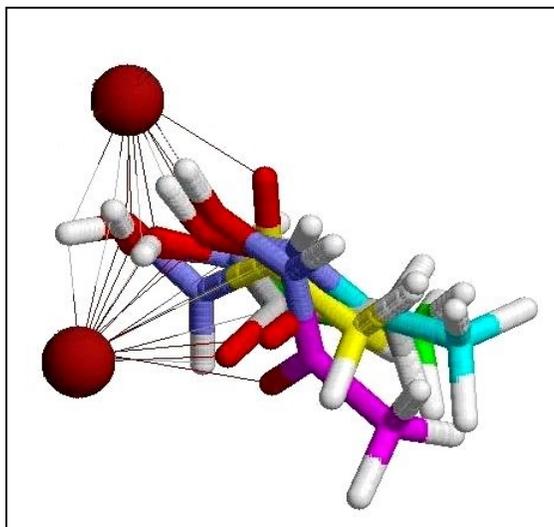


Figura 16. Sobreposição das estruturas do ácido acetohidroxâmico co-cristalizado na urease (código 4UBP do PDB) (padrão CPK de cores), e atracado com a função *ASP* (carbonos em verde), função *ChemPLP* (carbonos em ciano), função *ChemScore* (carbonos em magenta) e função *GoldScore* (carbonos em amarelo). Os íons de níquel (II) estão representados como esferas em marrom, e o ácido acetohidroxâmico no formato de bastões. Os aminoácidos do sítio foram omitidos para melhor visualização do inibidor.

Tabela 7. Valores da raiz do desvio médio quadrático (RMSD) para cada função testada no procedimento de *redocking* do ligante da estrutura 4UBP do PDB.

função	RMSD (Å)
<i>ChemPLP</i>	1,2611
<i>GoldScore</i>	1,4364
<i>ChemScore</i>	1,8398
<i>ASP</i>	1,8502

5.3 Construção do modelo de predição de atividade de inibidores da urease.

Para se construir um modelo teórico de predição de atividade relativa é mais apropriado que os dados de inibição experimentais sejam obtidos num mesmo experimento, uma vez que existem erros analíticos intrínsecos a cada conjunto de fatores envolvidos num determinado ensaio biológico, principalmente por parte do analista e da temperatura em que foi realizado o experimento, visto que a atividade das enzimas costuma ser muito sensível às variações desta. Ainda que tais erros sejam pequenos, dentro de um mesmo experimento os resultados obtidos podem ser comparados de forma mais segura, uma vez que estiveram submetidos às mesmas fontes de erros e nas mesmas proporções. Dessa forma, buscou-se explorar, separadamente, os dados referentes aos ácidos fosfínicos da série 1 e da série 2.

De acordo com a proposta de Wang e colaboradores (1994), desenvolvida no estudo de inibidores da proteína quinase C, é possível obter um modelo termodinamicamente consistente (equação 15) de predição de atividade através da correlação linear múltipla de termos importantes no processo de inibição enzimática, obtidos teoricamente. O mesmo modelo, aplicado a inibidores da fosfodiesterase 4 [OLIVEIRA *et al.*, 2006] e a inibidores da CYP51 [ROCHA, 2009] na predição da atividade de inibidores da esterol 14 α -desmetilase (CYP51), levou a bons modelos de predição de atividade.

5.3.1. Modelo desenvolvido com base nos dados experimentais dos inibidores da série 1.

Na tabela 8 estão listados os dados referentes aos ácidos *P*-metil fosfínicos, pertencentes à série 1 (figura 11): a energia livre de Gibbs necessária para o transporte da molécula do meio aquoso para o sítio ativo da enzima (ΔG_I) o número de ligações rotacionáveis que ficam “congeladas” (N_{LR}) após o encaixe desses inibidores na enzima, determinado diretamente através da observação da estrutura de cada ligante; e a entalpia de interação (ΔH_{int}) entre o inibidor e a urease, calculada através do método semi-empírico PM6, segundo a equação 9. Assumimos o estado quinteto para o sistema (cálculo UHF) devido ao padrão hexacoordenado de cada íon Ni²⁺ e à boa reprodução das estruturas experimentais (tabela 4) e constante dielétrica 4,0 por ser um valor que mimetiza o interior das proteínas e que também demonstrou boa reprodução das estruturas experimentais (tabela 6).

Wang e colaboradores (1994) consideraram o termo referente à ΔG_1 como sendo proporcional à solubilidade em água e fizeram uso de uma função quadrática para este termo (equação 15). Esta dependência quadrática entre K_i e a solubilidade foi justificada porque a proteína quinase C é uma enzima de membrana e os compostos devem apresentar alguma solubilidade na fase lipídica para conseguirem chegar à enzima. Como eles também devem ter solubilidade na fase aquosa, os melhores inibidores deveriam ter valores intermediários para esse termo. Embora a urease não seja uma enzima de membrana, a avaliação de um perfil quadrático para o termo de ΔG_1 é justificável também nesse caso se considerarmos a constante dielétrica bastante baixa no interior de proteínas (assumido como 4,0 neste estudo). Porém, a análise por regressão múltipla dos termos calculados na tabela 8 aplicados na equação 15 nos conduziu a um modelo com uma correlação fraca com os valores de K_i experimentais ($r^2=0,58$; $SD=0,74$).

Tabela 8. Logaritmo dos valores de K_i experimentais, energia livre de Gibbs necessária para o transporte do inibidor da fase aquosa para o interior da enzima (ΔG_1 , kcal/mol); número de ligações rotacionáveis congeladas durante a interação (N_{LR}), e entalpias de interação calculadas assumindo estado quinteto e constante dielétrica 4,0 (ΔH_{int} , kcal/mol), para os inibidores da série 1.

Estrutura	$\log(K_i)$	ΔG_1	N_{LR}	ΔH_{int}
1	2,53	-0,11	2	24,50
2	1,32	-0,70	4	31,98
3	2,08	0,20	5	-16,44
4	2,33	-0,33	6	-57,81
5	1,40	0,24	8	-22,61
6	1,57	-1,58	6	32,41
7	1,26	0,21	2	-19,21
8	1,63	0,96	5	71,00
9	2,13	-0,54	7	-49,44
10	-0,77	-1,90	7	11,02
11	1,24	-0,38	5	-4,67

Esta falta de correlação pode ter sido resultado da dificuldade em se estimar precisamente a entalpia de interação dos complexos inibidor-enzima, devido ao uso da aproximação UHF nos cálculos semi-empíricos, necessária para modelar o sistema no estado quinteto [MUSIANI *et al.*, 2001]. Nas estruturas otimizadas das séries 1 e 2, observou-se contaminação de spin em diferentes níveis (de 1 até 34%), tanto dos sítios, quanto dos

complexos, o que pode ter reduzido a precisão dos resultados obtidos. A contaminação de spin é causada pela mistura das funções de onda com o estado de spin esperado com outros estados de spin. Isto pode em alguns casos aumentar a energia, se houver mistura com um sistema de maior energia em comparação com a do sistema no estado de spin esperado.

Outra possibilidade para a falta de correlação refere-se aos valores do gradiente da energia em função dos parâmetros estruturais. Idealmente, como a energia é um mínimo em uma estrutura otimizada, o gradiente correspondente deve ser zero. O uso do modelo de contínuo COSMO (com a constante dielétrica 4,0 e 78,4) em geral aumenta a dificuldade em se alcançar a convergência, e os cálculos para a obtenção dos termos de ΔH_{int} e ΔG_I são muitas vezes terminados com valores de gradiente relativamente altos. Normas de gradientes altas afetam principalmente a precisão dos valores de ΔG_I , pois as variáveis termodinâmicas são calculadas a partir das frequências vibracionais, e por isso é necessário que o ligante esteja no seu mínimo de energia, ou seja, com uma norma de gradiente bastante baixa. O valor recomendado pelo programa MOPAC nos arquivos *output* gerados nos cálculos dos ligantes usando a palavra chave “THERMO” é de 0,2 kcal/(Å ou rad).

Por último, outra possível fonte de erro refere-se à forma com que foram obtidos os valores experimentais de K_i dos inibidores da série 1. Segundo Vassiliou e colaboradores (2008), os dados experimentais de K_i obtidos foram determinados através do gráfico de Lineweaver–Burk (figura 17).

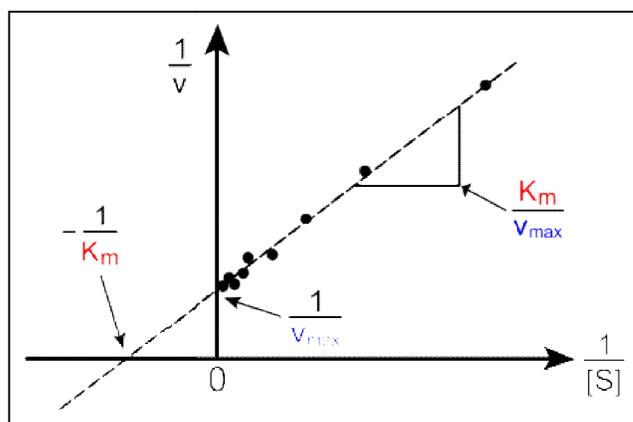


Figura 17. Gráfico de Lineweaver-Burk

Este gráfico, descrito por Hans Lineweaver e Burk Dear em 1934, tem origem na equação de Michaelis-Menten de cinética enzimática [LINEWEAVER & BURK, 1934], onde

V é a velocidade da reação, K_m é a constante de Michaelis-Menten, V_{max} é a velocidade máxima da reação e [S] é a concentração do substrato.

O gráfico de Lineweaver-Burk foi amplamente utilizado antes da disponibilidade dos computadores modernos e programas de regressão não-linear para se determinar termos importantes na cinética das enzimas, tais como K_m e V_{max} , mas está sujeito a erros: o eixo y tem o inverso da taxa de reação (1/V) o que faz com que pequenos erros de medição sejam maximizados. Além disso, a maioria dos pontos na equação é encontrada muito mais à direita do eixo y, por causa da limitação da solubilidade: como não existem grandes valores de [S], o inverso desse termo leva a números muito elevados, fazendo com que seja necessário uma grande extrapolação para obter o valor onde a reta intercepta o eixo y e x [HAYAKAWA *et al.*, 2006]. Um modelo linear muito mais exato é o diagrama de Eadie-Hofstee, mas nas investigações científicas atuais, linearizações tornaram-se obsoletas e foram substituídas por métodos mais fiáveis, baseados na análise de regressão não linear [BRAVO *et al.*, 2001].

Mesmo diante destas possíveis limitações, foi feita uma análise mais detalhada dos modelos obtidos. Ao se descartar a contribuição dos compostos **2** e **11**, foi obtida a equação (16), com uma boa correlação ($r^2=0,84$; $SD=0,55$).

$$\log K_i = -1,62(\Delta G_i + 0,38)^2 + 0,02\Delta H_{int} + 0,12N_{LR} + 2,18 \quad (16)$$

Apesar deste modelo (equação 16) não ter sido construído utilizando os dados dos termos calculados de todos os inibidores da série 1, ele ainda envolveu a participação de compostos com uma ampla faixa de valores experimentais de K_i (0,17 a 340 μM).

Uma análise das contribuições dos parâmetros utilizados neste modelo demonstrou que os três parâmetros contribuem de forma proporcional. Dentre os 9 compostos usados na construção deste modelo, 3 compostos tiveram sua atividade mais fortemente influenciada pela variação da entalpia de interação, outros 3 compostos tiveram sua atividade mais fortemente influenciada perda entrópica associada à perda de liberdade conformacional e por fim, outros 3 compostos tiveram o termo de ΔG_i como o fator principal para a redução do $\log K_i$ (tabela 9).

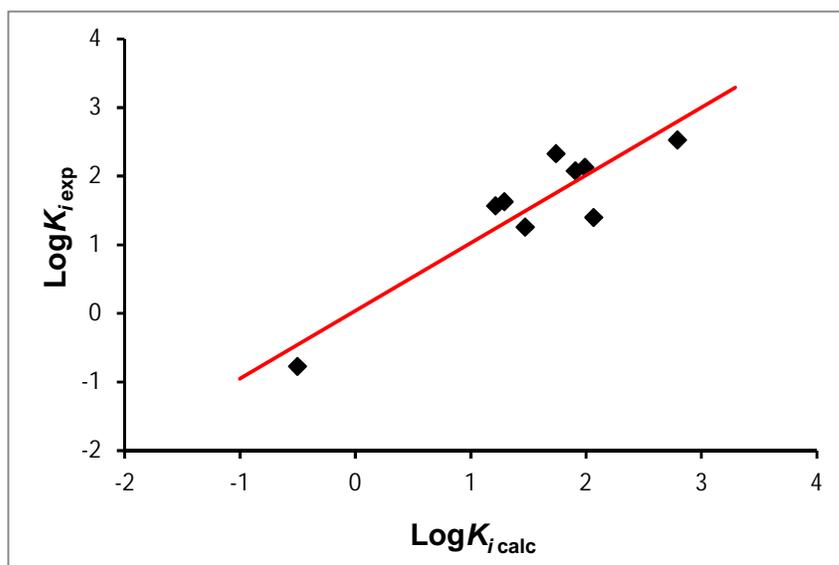
A aplicação da equação 16 na predição da atividade dos ácidos *P*-metil fosfínicos que foram utilizados neste estudo reproduziu bem os valores de atividade experimentais ($r=0,92$; $SD=0,39$), com exceção dos compostos que foram eliminados (tabela 9 e figura 18).

Tabela 9. Contribuição de cada termo variável da equação 16, $\log K_i$ medido experimentalmente, $\log K_i$ calculado pela equação 16, e a diferença entre os valores de $\log K_i$ experimentais e calculados.

Estrutura	$[-1,62(\Delta G_1 + 0,38)^2]$	$0,02\Delta H_{int}$	$0,12N_{LR}$	$\log K_{i \text{ exp}}$	$\log K_{i \text{ calc}}$	$\Delta \log K_i$
1	-0,12	0,49	0,24	2,53	2,79	-0,26
2^a	-0,17	0,64	0,48	1,32	3,13	-1,81
3	-0,54	-0,33	0,60	2,08	1,91	0,17
4	-0,00	-1,16	0,72	2,33	1,74	0,59
5	-0,62	-0,45	0,96	1,40	2,06	-0,66
6	-2,33	0,65	0,72	1,57	1,21	0,35
7	-0,56	-0,38	0,24	1,26	1,47	-0,21
8	-2,91	1,42	0,60	1,63	1,29	0,34
9	-0,04	-0,99	0,84	2,13	1,99	0,14
10	-3,74	0,22	0,84	-0,77	-0,50	-0,27
11^a	0,00	-0,09	0,60	1,24	2,69	-1,45

^a Estruturas eliminadas da série 1 para obtenção da correlação final (equação 16).

Figura 18. Gráfico de correlação entre $\log K_i$ experimental e calculado, para os inibidores da série 1.



5.3.2. Modelo desenvolvido com base nos dados experimentais dos inibidores da série 2.

Na tabela 10 estão listados os dados referentes aos ácidos *P*-metil fosfínicos, pertencentes à série 2 (figura 12): a energia livre de Gibbs necessária para o transporte da molécula do meio aquoso para o sítio ativo da enzima (ΔG_I) o número de ligações rotacionáveis que ficam “congeladas” (N_{LR}) após o encaixe desses inibidores na enzima, determinado diretamente através da observação da estrutura de cada ligante; e as entalpias de interação (ΔH_{int}) entre o inibidor e a urease, calculadas através do método semi-empírico PM6, segundo a equação 9. Assumimos o estado quinteto para o sistema (cálculo UHF) e constante dielétrica 4,0.

Tabela 10. Logaritmo dos valores de K_i experimentais, energia livre de Gibbs necessária para o transporte do inibidor da fase aquosa para o interior da enzima (ΔG_I , kcal/mol); número de ligações rotacionáveis congeladas durante a interação (N_{LR}), e entalpias de interação calculadas assumindo estado quinteto e constante dielétrica 4,0 (ΔH_{int} , kcal/mol), para os inibidores da série 2.

Estrutura	$\log(K_i)$	ΔG_I	N_{LR}	ΔH_{int}
12	1,98	0,36	4	-145,10
13	-0,37	-0,16	3	-9,00
14	0,70	-0,69	4	-69,75
15	1,48	0,25	6	28,99
16	1,02	0,12	6	-39,84
17	0,06	0,05	6	68,81
18	0,13	-1,57	9	-71,99
19	1,50	-0,01	5	74,13
20	2,57	-0,96	7	-20,20
21	2,63	0,47	7	-12,69

A análise por regressão múltipla dos termos calculados para as estruturas da série 2, utilizando os valores de ΔH_{int} nos conduziu a uma equação de correlação fraca com os valores de K_i experimentais ($r^2=0,42$; $SD=1,07$). Esta falta de correlação pode ser atribuída aos mesmos fatores discutidos na construção do modelo usando os compostos da série 1 (item 5.3.1) exceto pelo método usado para a obtenção dos valores experimentais de atividade. Segundo Vassiliou e colaboradores (2010) os dados experimentais de atividade dos ácidos *P*-hidroximetil fosfínicos (série 2, figura 12) foram obtidos através do programa GraphPad

Prism 5 software. Este programa possui a possibilidade de escolha entre os gráficos de Hanes-Woolf, Eadie-Hofstee, Linewear-Burk, entre outros, porém estes gráficos são usados apenas como uma forma de exibição dos dados, e não como a primeira etapa do tratamento dos dados. A primeira etapa é feita pelo programa através da regressão não linear dos dados para a obtenção de melhores resultados [<http://www.graphpad.com>].

Foi feita uma análise mais detalhada da equação obtida a fim de se identificar qual estrutura estaria contribuindo para a baixa correlação com os dados experimentais. Excluindo-se a contribuição de apenas um composto, **18**, obteve-se a equação 17 com uma boa correlação com os dados de atividade experimentais ($r^2=0,80$; $SD=0,66$).

$$\log K_i = 5,018(\Delta G_1 + 0,271)^2 + 0,002\Delta H_{\text{int}} + 0,029N_{LR} + 0,042 \quad (17)$$

Apesar do modelo não ter sido construído utilizando-se os dados dos termos calculados de todos os inibidores da série 2, ele ainda envolveu a participação de compostos com uma boa faixa de valores experimentais de K_i (0,43 a 428 μM).

Uma análise das contribuições dos três parâmetros utilizados neste segundo modelo demonstrou que, para esta série, a energia livre de Gibbs necessária para transportar o inibidor da fase aquosa para o sítio, é o fator que mais contribui para a determinação da atividade predita.

A aplicação da equação 17 na predição da atividade dos ácidos *P*-hidroximetil fosfínicos que foram utilizados neste estudo reproduziu bem os valores de atividade experimentais ($r=0,90$; $SD=0,44$), com exceção do composto que foi eliminado (tabela 11 e figura 19).

Tabela 11. Contribuição de cada termo variável da equação 17, $\log K_i$ medido experimentalmente, $\log K_i$ calculado pela equação 17, e a diferença entre os valores de $\log K_i$ experimentais e calculados.

Estrutura	$[5,018(\Delta G_1 + 0,271)^2]$	$0,002\Delta H_{int}$	$0,029N_{LR}$	$\log K_{i\text{ exp}}$	$\log K_{i\text{ calc}}$	$\Delta \log K_i$
12	1,998	-0,290	0,116	1,98	1,86	0,11
13	0,062	-0,018	0,087	-0,37	0,17	-0,54
14	0,881	-0,139	0,116	0,70	0,90	-0,20
15	1,362	0,058	0,174	1,48	1,64	-0,16
16	0,767	-0,079	0,174	1,02	0,90	0,12
17	0,517	0,1376	0,174	0,06	0,87	-0,81
18^a	8,467	-0,144	0,261	0,13	8,63	-8,49
19	0,341	0,148	0,145	1,50	0,68	0,82
20	2,382	-0,040	0,203	2,57	2,59	-0,02
21	2,755	-0,025	0,203	2,63	2,97	-0,34

^a Estrutura eliminada da série 2 para obtenção da correlação final (equação 17).

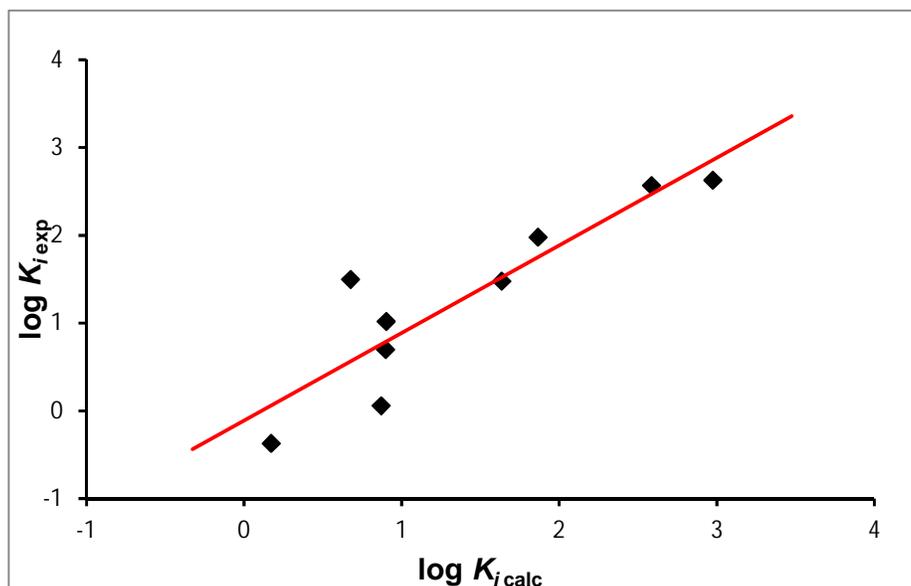


Figura 19. Gráfico de correlação entre $\log K_i$ experimental e calculado, para os inibidores da série 2

5.4 Análise dos termos envolvidos nos modelos da série 1 e 2.

5.4.1 Perda entrópica do inibidor durante a interação

Quando uma molécula deixa o meio aquoso e se liga à enzima, ocorrem significativas perdas entrópicas translacionais, rotacionais e conformacionais. As perdas entrópicas rotacionais e translacionais podem ser consideradas constantes [AMZEL, 1997]. A perda entrópica relacionada à nova conformação que o ligante passa a assumir no complexo com a enzima é dependente de características específicas associadas a cada estrutura. Esta perda entrópica conformacional pode ser mensurada a partir da avaliação do número de ligações rotacionáveis que o ligante possui. Como resultado da interação do inibidor com a enzima, todas ou grande parte das ligações rotacionáveis ficam “congeladas”. Quanto menor a perda entrópica que uma molécula sofre ao interagir com o sítio ativo da enzima, mais negativa (ou seja, mais favorável) será a energia livre de Gibbs da interação. Como ΔG_{int} é diretamente proporcional à constante inibitória ($\log K_i$), este termo experimental deve ser diretamente proporcional à perda entrópica (N_{LR}), o que pôde ser observado nos modelos obtidos a partir dos compostos das séries 1 e 2 (equação 16 e 17).

5.4.2 Energia livre de Gibbs para a passagem do inibidor do meio aquoso para o sítio da urease.

O valor negativo para o coeficiente do termo quadrático de ΔG_I no modelo da série 1 (equação 16) indica que a dependência da atividade no termo ΔG_I é representada por uma parábola com concavidade voltada para baixo. Assim, os valores mais extremos para esse termo contribuem para a obtenção de valores de K_i mais baixos (ou seja, de ΔG_{int} mais favoráveis). Esse comportamento poderia ser explicado considerando-se a passagem do inibidor para o sítio ativo da urease como ocorrendo em duas etapas de equilíbrio, com dependências distintas da solubilidade dos compostos. Desse modo, o termo ΔG_I não seria integralmente representado pela energia livre calculada pela equação 14 (igual a $\Delta G_{78,4-4,0}$), associada somente à passagem do ligante da água para o meio com $\varepsilon=4,0$. O termo ΔG_I real seria mais bem representado por uma soma de dois termos de energia livre, associados a cada etapa de equilíbrio (equação 18),

$$\Delta G_1 = \Delta G_{aq/enz} + \Delta G_{enz/sitio} \quad (18)$$

onde $\Delta G_{aq/enz}$ é a energia livre de Gibbs necessária para o ligante sair do meio aquoso e passar para o interior da enzima, um meio que na média apresenta baixa polaridade e $\Delta G_{enz/sitio}$ é a energia livre de Gibbs necessária para o inibidor entrar no sítio ativo da urease, um meio que deve ser bem mais polar devido à presença dos íons níquel (II) e hidroxila e de resíduos de aminoácidos carregados.

Claramente, os dois equilíbrios teriam dependências opostas do comportamento de solubilidade dos ligantes. Um composto que fosse mais solúvel em meio apolar teria um termo $\Delta G_{aq/enz}$ muito favorável (muito negativo) e um termo $\Delta G_{enz/sitio}$ desfavorável (positivo); o oposto seria observado para um composto muito solúvel em água e, em ambos os casos, a combinação dos termos $\Delta G_{aq/enz}$ e $\Delta G_{enz/sitio}$ levaria a um valor de ΔG_1 menor que zero (favorável) (figura 20). Já um composto com características de solubilidade intermediária apresentaria os termos $\Delta G_{aq/enz}$ e $\Delta G_{enz/sitio}$ com valores numericamente baixos, o que levaria a um valor de ΔG_1 próximo de zero.

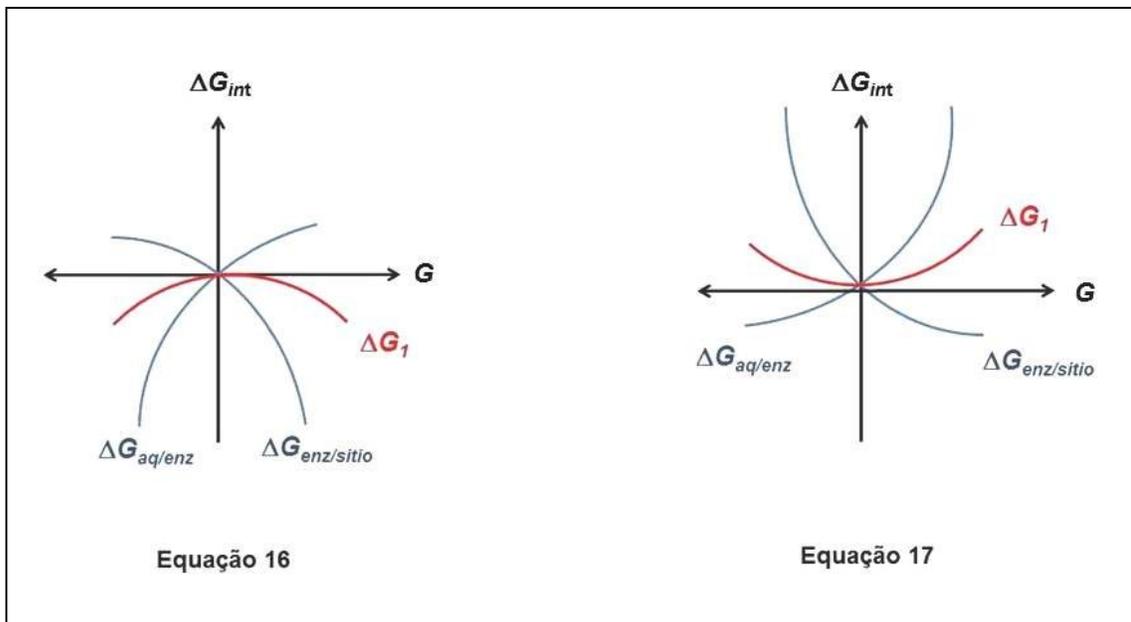


Figura 20. Esquema ilustrativo do comportamento de ΔG_{int} em função de ΔG_1 e de suas componentes $\Delta G_{aq/enz}$ e $\Delta G_{enz/sitio}$ para a equação 16 e para a equação 17.

A diferença entre o termo quadrático de ΔG_I da equação 16 e da equação 17, conforme pode ser observado no esquema ilustrativo da figura 20, seria a concavidade das curvas dos equilíbrios intermediários (em azul), o que faz com que a concavidade da parábola resultante de ΔG_I (em vermelho) fique voltada para baixo (equação 16) ou voltada para cima (equação 17).

O valor positivo para o coeficiente do termo quadrático de ΔG_I no modelo da série 2 (equação 17) indica que a dependência da atividade no termo ΔG_I é representada por uma parábola com concavidade voltada para cima. Assim, são os valores intermediários para ΔG_I que contribuem para a obtenção de valores de K_i mais baixos (e de ΔG_{int} mais favoráveis). O modelo 2 (equação 17) prediz que o valor ideal para o termo de ΔG_I é de aproximadamente -0,27 kcal/mol, pois este valor torna nulo o termo quadrático de ΔG_I , contribuindo para valores mais baixos de $\log K_i$, ou seja, aumentando a atividade dos inibidores.

5.4.3 Entalpia de interação inibidor/urease

Em ambos os modelos, o coeficiente do termo ΔH_{int} ficou positivo, o que está de acordo com o princípio de que quanto mais favorável for a interação entre um ligante e o sítio ativo de uma enzima (ou seja, quanto mais negativo for o valor da entalpia de interação), menor será do logaritmo de K_i (ou seja, maior a atividade).

5.5 Predição da atividade relativa dos compostos da série 3.

Alguns compostos organofosforados sintetizados pelo grupo de síntese da UFRRJ, coordenado pelo Dr. João Batista Neves da Costa, foram sintetizados como possíveis inibidores da urease (estruturas 22 a 29, da figura 13). Como são dialquilfosforilidrazonas, estes compostos apresentam alguns heteroátomos com pares solitários, capazes de interagirem com os íons de níquel (II) localizados no sítio ativo da urease, como os oxigênios ligados ao átomo de fósforo e os nitrogênios do grupo hidrazona. Porém, estes compostos também possuem cadeias alifáticas, que contém entre 2 a 4 carbonos, ligadas a dois dos átomos de oxigênio, dificultando a interação com os íons de níquel (II). Neste trabalho buscou-se avaliar também o perfil de atividade dos intermediários utilizados na síntese das dialquilfosforilidrazonas: nitroisatina, cloroisatina, α -lapachona e β -lapachona (estruturas 30, 31, 32 e 33 da figura 13, respectivamente).

Como a equação 16 apresentou correlação um pouco melhor do que a equação 17, ela foi escolhida para se estimar as atividades dos compostos da série 3. Conforme descrito por Wang e colaboradores (1994) o desenvolvimento da equação 15 (e, portanto da equação 16), baseia-se na suposição de que as moléculas sejam semelhantes. Como a série de compostos usada na obtenção da equação 16 (série 1) e a série 3 apresentam diferenças estruturais relevantes, a aplicação da equação aos inibidores da série 3 não deve levar a resultados quantitativos da atividade prevista. Contudo, como as estruturas da série 3 são semelhantes entre si, os erros associados com esta aplicação também devem ser proporcionais para estas estruturas. Assim, é razoável assumir que os resultados obtidos devem se manter paralelos aos valores da atividade real dos inibidores da série 3. Dessa forma, os resultados da aplicação direta da equação 16 aos compostos da série 3 devem ser considerados como um indicativo da atividade relativa destas estruturas.

Os três termos calculados para estes compostos estão listados na tabela 12: a energia livre de Gibbs necessária para transportar o ligante do meio aquoso para o interior da enzima (ΔG_1); o número de ligações rotacionáveis que ficam congeladas após a interação com a enzima (N_{LR}) e a entalpia de interação calculada pelo programa MOPAC2012 através do método semi-empírico PM6, considerando o estado quinteto e constante dielétrica 4,0 (ΔH_{int}).

Tabela 12. Energia livre de Gibbs necessária para o transporte dos inibidores da série 3 (ΔG_1 , kcal/mol), número de ligações rotacionáveis congeladas (N_{LR}) e entalpia de interação inibidor/urease (ΔH_{int} , kcal/mol).

Estrutura	ΔG_1	N_{LR}	ΔH_{int}
22	-1,54	5	66,30
23	-2,15	3	176,37
24	-1,59	3	168,12
25	0,26	7	90,91
26	-1,27	3	34,90
27	-0,96	3	62,15
28	-1,40	3	56,88
29	-0,68	3	93,68
30	0,20	0	-22,85
31	0,20	0	102,13
32	0,10	0	163,40
33	-0,10	0	85,94

Podemos observar pela tabela 12 que as entalpias de interação são previstas como desfavoráveis para todos compostos, exceto para a nitroisatina (estrutura **30**). Na interação entre os compostos da série 3 com a urease, pode-se observar a presença de algumas interações hidrofóbicas, mas estas não ocorreram em grandes proporções; houve também a presença de, em média, duas ligações de hidrogênio por complexo ligante/urease. Na maioria dos casos, a interação dos oxigênios ligados ao fósforo com os íons de níquel (II) não ocorreu. Dessa forma, podemos supor que a falta desse tipo de interação seja o fator principal pelas entalpias de interação desfavoráveis das dialquilfosforilidrazonas da série 3.

Observou-se que nas interações entre as dialquilfosforilidrazonas sintetizadas a partir da nitroisatina (estruturas **25**, **26** e **27**) com a urease, a extremidade contendo o átomo de fósforo ficou direcionada para fora do sítio, ficando, portanto, a extremidade contendo o grupo nitro direcionada para os íons de níquel (II).

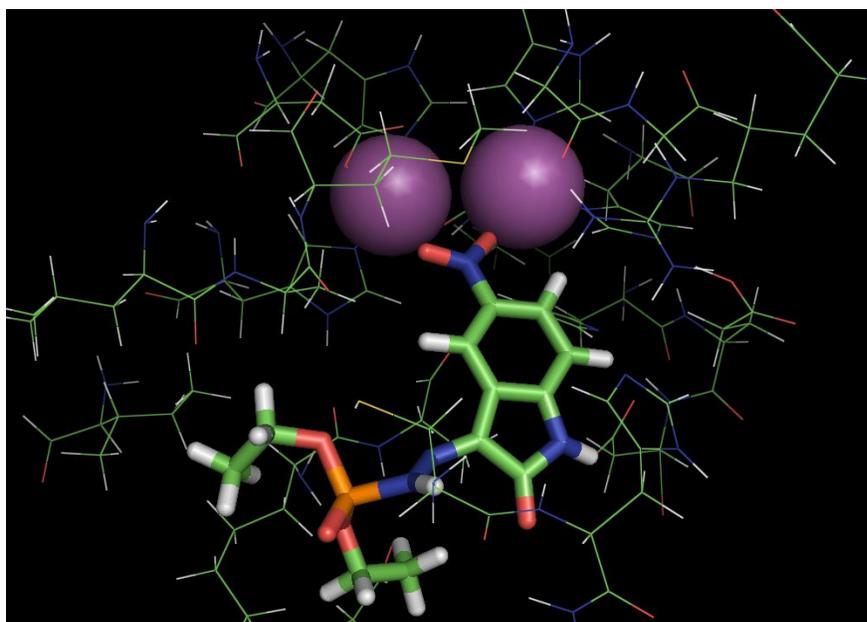


Figura 21. Estrutura **26** complexada no sítio ativo da urease 4UBP (*Bacillus Pasteurii*) com o grupo nitro direcionado para os íons de níquel (II). Inibidor no formato bastão e íons de níquel (II) no formato de esfera.

Observou-se também que, dentre as dialquilfosforilidrazonas sintetizadas a partir da nitroisatina, a estrutura **26** apresentou uma entalpia de interação menos desfavorável e isso pode estar relacionado ao fato dessa estrutura possuir menores cadeias alifáticas, fazendo com que o grupamento nitro ficasse mais próximo dos íons de níquel (II) devido ao menor volume

da molécula, facilitando o seu encaixe no sítio (figura 21). As distâncias entre cada átomo de oxigênio do grupo nitro dessa estrutura e os íons de Ni(II) foram de 2,27Å e 4,21Å.

Na interação entre o intermediário nitroisatina sem o grupamento *P*-terminal (estrutura **30**) com a urease, a entalpia de interação apresentou-se favorável (-22,85 kcal/mol) e as distâncias entre cada átomo de oxigênio do grupo nitro e os íons de Ni(II) foram bem menores: 1,88Å e 1,87Å, indicando que a interação entre átomos doadores de elétrons com os íons de níquel (II) é um fator muito importante para a obtenção de entalpias de interação mais favoráveis.

Nas demais dialquilfosforilidrazonas sintetizadas a partir da α -lapachona, β -lapachona e cloroisatina, o grupamento dialquilfosforila ficou direcionado aos íons de níquel (II). As entalpias de interação, nestes casos, provavelmente foram desfavoráveis devido às cadeias alifáticas ligadas aos oxigênios, que impediram a interação efetiva dos oxigênios com os íons de níquel (II) (figura 22).

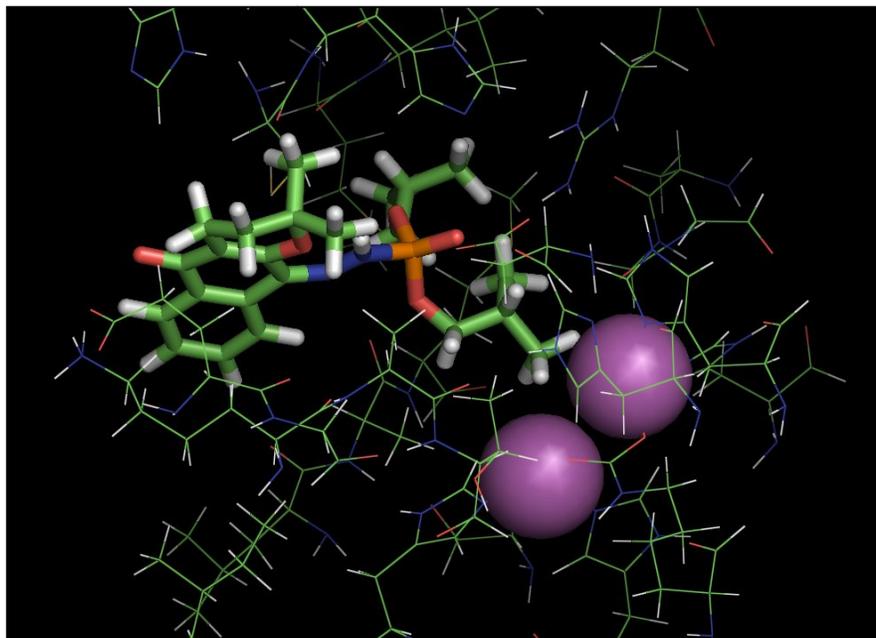


Figura 22. Estrutura **22** complexada no sítio ativo da urease 4UBP (*Bacillus Pasteurii*). Ligante no formato bastão e íons de níquel (II) no formato de esfera.

Tabela 13. Contribuição de cada termo variável da equação 16 para os compostos da série 3.

Estrutura	$[-1,62(\Delta G_1 + 0,38)^2]$	$0,02\Delta H_{int}$	$0,12N_{LR}$	$\log K_i \text{ calc}$
22	-2,20	1,33	0,48	1,79
23	-5,10	3,53	0,24	0,84
24	-2,40	3,36	0,24	3,38
25	-0,67	1,82	0,72	4,05
26	-1,30	0,70	0,24	1,82
27	-0,54	1,24	0,24	3,12
28	-1,69	1,14	0,24	1,87
29	-0,14	1,87	0,24	4,15
30	-0,55	-0,46	0,00	1,17
31	-0,54	2,04	0,00	3,68
32	-0,38	3,27	0,00	5,07
33	-0,12	1,72	0,00	3,77

Podemos observar na tabela 12 que os valores do termo de energia livre de Gibbs para o transporte dos inibidores do meio aquoso para o sítio da urease, são, em sua maioria, negativos, indicando que eles possuem mais afinidade pelo meio apolar. Na tabela 13 estão apresentadas as contribuições de cada termo variável da equação 16 para os compostos da série 3. Nesta tabela observa-se que, na maioria dos casos, a entalpia de interação foi o termo que mais contribuiu. A segunda maior contribuição deve-se ao termo ΔG_1 . O composto **23**, por ter uma alta afinidade pelo meio apolar, teve o termo ΔG_1 contribuindo fortemente para a redução do $\log K_i$ calculado.

5.6 Modificação proposta nas estruturas da série 3.

De maneira geral, a retirada das cadeias alifáticas das dialquilfosforilidrazonas da série 3 poderia contribuir bastante para o desenvolvimento de inibidores da urease mais promissores. Para avaliar o efeito disto na atividade, os três termos variáveis da equação 16 foram calculados para a fosforilidrazona sintetizada a partir da β -lapachona, na qual as cadeias alquílicas foram retiradas (estrutura **34**, figura 23).

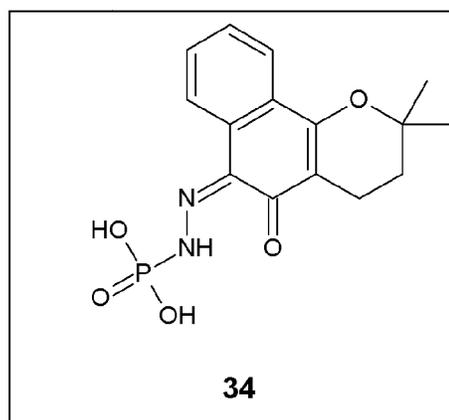


Figura 23. Fosforilidrazona proposta, derivada da β -lapachona.

Essa estrutura proposta apresentou entalpia de interação com a urease bastante favorável, provavelmente devido à interação de dois oxigênios do grupo fosforilidrazona (na forma desprotonada) com o íon de Ni(II) de menor coordenação, o Ni794 (distância de interação igual a 1,88 Å) e a interação do terceiro átomo de oxigênio do grupo fosforilidrazona com um resíduo de Arg562 (1,51 Å), conforme mostrado na figura 24.

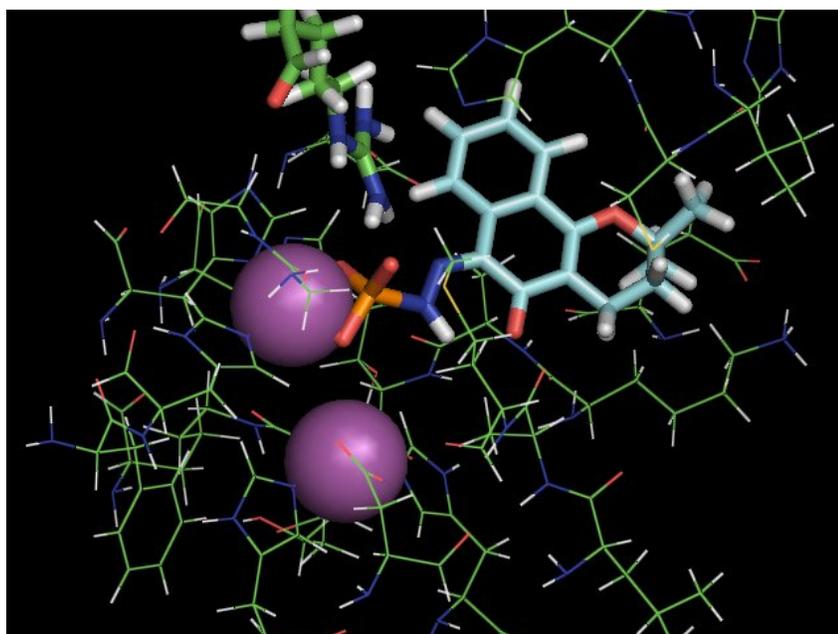


Figura 24. Estrutura **34** complexada no sítio ativo da urease 4UBP (*Bacillus Pasteurii*). Íons de Ni(II) no formato de esferas. Ligante no formato de bastões com os carbonos em ciano. Arginina 562 no formato de bastões com carbonos em verde.

A energia livre de Gibbs calculada para o transporte dessa estrutura do meio aquoso até o sítio foi de 0,22 kcal/mol, ou seja, a retirada das cadeias alifáticas, como esperado, resultou em uma molécula mais polar. Como esta estrutura não possui ligação rotacionável, o termo N_{LR} ficou nulo, o que contribuiu para um valor mais favorável da atividade. Na tabela 14 são apresentadas as contribuições de cada termo variável da equação 16 para a estrutura proposta. A estrutura **34** possui um valor de K_i predito na escala subnanomolar ($K_i=0,00034$ μ M ou 0,34 nM) e por isso é bastante promissora na inibição da urease.

Tabela 14. Contribuição de cada termo variável da equação 16 para a estrutura 34.

Estrutura	$[-1,62(\Delta G_1 + 0,38)^2]$	$0,02\Delta H_{int}$	$0,12N_{LR}$	$\log K_i \text{ calc}$
34	-0,59	-5,06	0,00	-3,47

Como uma proposta para futuras modificações nos compostos da série 3, tem-se que a retirada das cadeias alifáticas das dialquilfosforilidrazonas sintetizadas a partir da α -lapachona, nitroisatina e cloroisatina pode resultar em inibidores tão promissores quanto a estrutura **34**, derivado da β -lapachona.

6. CONCLUSÕES

Numa primeira etapa deste trabalho, buscou-se definir as características do sítio ativo da urease, avaliando a multiplicidade de spin dos íons de Ni(II) e o estado de protonação do átomo de oxigênio localizado entre estes íons, informações úteis nos cálculos de entalpia de interação realizados posteriormente. Estudos com o método semiempírico PM6 indicaram que o estado de spin mais provável para a urease pode ser tanto o tripleto quanto o quinteto, devido à melhor reprodução da geometria experimental de estruturas cristalográficas da enzima, porém o estado quinteto foi escolhido devido ao padrão hexacoordenado dos íons de Ni(II). Da mesma forma, os estudos indicaram que o átomo de oxigênio localizado entre os íons de Ni(II) é mais provavelmente um íon hidroxila

Numa segunda etapa, buscou-se construir um modelo de predição de atividade para inibidores da urease utilizando duas séries de compostos da literatura com valores de K_i experimentais conhecidos. O modelo desenvolvido nesse trabalho baseou-se na proposta original de Wang e colaboradores (1994), no qual o uso de uma equação resultante da avaliação dos fatores entrópicos e entálpicos envolvidos no processo de interação entre os

inibidores e a enzima permitiu o desenvolvimento de um modelo termodinamicamente consistente para a predição da energia livre de interação. Como demonstrado neste trabalho, este método demonstrou ter aplicabilidade geral para o desenvolvimento de modelos, podendo ser usado mesmo em sistemas como a urease, que possui fatores complicadores, como a presença dos íons Ni(II) no sítio catalítico, graças à disponibilidade do método PM6, um método quântico semi-empírico parametrizado para metais de transição.

Os modelos finais, que continham três termos independentes, o número de ligações rotacionáveis que ficam congeladas após o encaixe do inibidor no sítio (N_{LR}), a energia de livre de Gibbs necessária para o transporte do inibidor do meio aquoso para o interior do sítio (ΔG_I) e a entalpia de interação (ΔH_{int}), permitiram o cálculo de energias livres de interação, que, convertidas em dados de K_i , apresentaram boas correlações com os dados experimentais disponíveis.

Uma série de dialquilfosforilidrazonas sintetizadas pelo grupo de síntese da UFRRJ foi analisada segundo o modelo desenvolvido de melhor correlação com os dados experimentais (equação 16) e as atividades preditas indicaram que, provavelmente, estes compostos não são bons inibidores da urease, devido principalmente às entalpias de interação desfavoráveis.

Através do modelo desenvolvido, foi possível também estudar as modificações estruturais necessárias nas dialquilfosforilidrazonas a fim de que estas moléculas se tornassem inibidores mais promissores. A principal modificação observada refere-se à retirada dos grupamentos alquila, pois, além de diminuir o N_{LR} , faz com que o termo entálpico seja mais favorável por favorecer a interação dos ligantes como os íons Ni(II). A estrutura **34** proposta nesse trabalho teve sua atividade predita pela equação 16 e demonstrou ser bastante promissora, apresentando valor de K_i predito na escala subnanomolar

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHAM, D. J. Burger's: medicinal chemistry and drug discovery. Virginia: Wiley Interscience, p.1125. 2003.

ALCÁCER, L. Introdução à Química quântica computacional. Rio de Janeiro. Editora: IST Press, Lisboa, 2007. 305p. (Ensino da Ciência e da Tecnologia, 20)

ALLINGER, N.L. Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. J. Am. Chem. Soc., v.99, p.8127-8134, 1977.

ALONSO, H. *et al.* Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.*, v. 26, p. 531-568, 2006.

ASI, A. M. *et al.* Application of the linear interaction energy method (LIE) to estimate the binding free energy values of *Escherichia coli* wild-type and mutant arginine repressor C-terminal domain (ArgRc)–L-arginine and ArgRc–L-citrulline protein–ligand complexes. *J. Mol. Graphics Model.*, v.22, n.4, p.249-262, 2004.

BACCHI, A. *et al.* Antimicrobial and mutagenic properties of organotin(IV) complexes with isatin and N-alkylisatin bithiocarbonohydrazones. *J. Inorg. Biochem.* v.99, p.397. 2005.

BARREIRO, E. J. & FRAGA, C. A. M. *Química Medicinal*. Porto Alegre. Editora: Artmed, Porto Alegre, 2002.

BENINI, S. *et al.* A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels. *Struct. Fold. Des.* v.7, p.205. 1999.

BENINI, S. *et al.* The complex of *Bacillus pasteurii* urease with acetohydroxamate anion from X-ray data at 1.55 Å resolution. *J. Biol. Inorg. Chem.* v.1, n.5, p.110, 2000.

BENINI, S. *et al.* Molecular Details of Urease Inhibition by Boric Acid: □ Insights into the Catalytic Mechanism. *J. Am. Chem. Soc.*, v.12, n.126, p.3714–3715, 2004.

BÖHM, H.J.. LUDI: Rule-based automatic design of new substituents for enzyme inhibitor leads. *Journal of Computer-Aided Molecular Design*, v.8, p. 593-606. 1992

BORN, M. & OPPENHEIMER, J. R. Zur Quantentheorie der Molekeln. *Ann. Phys.*, v.84, p.457-484, 1927.

BONET, B. *et al.* A robust and fast selection mechanism for planning. *Proceedings of the 14th National Conference on Artificial Intelligence (American Association for Artificial Intelligence)*, p. 714-719, 1997.

BRAVO, I.G. *et al.* Kinetic properties of the Acylneuraminase Cytidylyltransferase from *Pasteurella haemolytica* A. *Biochem. J.*, v.358, p.585-598. 2001.

BREMNER, J. M. Recent Research on Problems in the Use of Urea as a Nitrogen Fertilizer. *Fert. Res.* v.42, p.321-329. 1995.

BROOKS, C. L. *et al.* A theoretical perspective of dynamics, structure, and thermodynamics. *Advances in chemical physics*; John Wiley: New York, v.LXXI. 1988.

- BROOKS, B.R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, v.4, p.187-217, 1983.
- BROOIJMANS, N. & KUNTZ, I.D.. Molecular recognition and docking algorithms. *Annual Reviews of Biophysical and Biomolecular Structures*, v.32, p.335-373. 2003.
- CLARK, T. A handbook of computational chemistry: a practical guide to chemical structure and energy calculations. New York: Wiley. p.352. 1985.
- CARLSSON, H. & NORDLANDER, E. Computacional Modeling of the Mechanism of Urease. *Bioinorganic Chem. And Applic.* 2010.
- CARMONA, G. *et al.* Temperature and low concentration effects of the urease inhibitor N-(n-butyl) thiophosphoric triamide (NBPT) on ammonia volatilisation from urea. *Soil Biol. Biochem.* v.22, p.933-937. 1990.
- COWAN, J. A. *Inorganic Biochemistry – An Introduction*. 2a Ed. New York: Ed. Wiley-VCH, 1993.
- DA COSTA, J. B. N. *et al.* Compostos Organofosforados Pentavalentes: Histórico, Métodos Sintéticos de Preparação e Aplicações como Inseticida e Agente Antitumorais. *Química Nova*, v.30, p.159. 2007.
- DA COSTA, P. A. & POPPI, R. J. Genetic algorithm in chemistry. *Química Nova*, v.22, p.405-411. 1999.
- DEWAR, M. J. S. *et al.* Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.*, v.107, p.3902-3909, 1985.
- DEWAR, M. J. S. & THIEL, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J. Am. Chem. Soc.*, v.99, p.4899-4907, 1977.
- DIXON, N. E. *et al.* Jack Bean Urease (EC 3.5.1.5) a Metalloenzyme. A simple Biological Role for Nickel? *J. Am. Chem. Soc.* 97: 4131-133, 1975.
- DOMÍNGUEZ, M. J. *et al.* Design, Synthesis, and Biological Evaluation of Phosphoramidate Derivatives as Urease Inhibitors *J. Agric. Food Chem.* v.56, p.3721-3731. 2008.
- ELDRIDGE, M.D., *et al.* Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design*, v.11, p. 425-445, 1997.

- ESTIU, G. & MERZ JR, K. M. Competitive Hydrolytic and Elimination Mechanisms in the Urease Catalyzed Decomposition of Urea. *J. Phys. Chem.* v.111, p.10263-10274. 2007.
- FENTON, D. E. *Biocoordination Chemistry*. New York: Oxford Chemistry Primers, 1995.
- FERREIRA, S. B. *et al.* B lapachona: sua importância em química medicinal e suas modificações estruturais. *Rev. Virtual de Química*, v.2, p.140-160. 2010.
- FONG, Y. *et al.* Assembly of Preactivation Complex for Urease Maturation in *Helicobacter pylori*: Crystal Structure of UreF-UreH Protein Complex. *Journal of Biological Chemistry*, v.50, p.286. 2011.
- GILL, J.S. *et al.* Efficiency of *N*-(*n*-butyl) thiophosphoric triamide in retarding hydrolysis of urea and ammonia volatilization losses in a flooded sandy loam soil amended with organic materials. *Nutr. Cycl. Agroecosyst.* v.53,p.203–207. 1997.
- GOHLKE, H., HENDLICH, M. & KLEBE, G.. Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, v.295, p.337-356. 2000.
- HAQ, Z. & WADOOD, A. Prediction of Binding Affinities for Hydroxamic Acid Derivatives as Urease Inhibitors by Molecular Docking and 3D-QSAR Studies. *Letters in Drug Design and Discovery*, v.6, p. 93-100. 2009.
- HANSSON T. *et al.* Ligand binding affinity prediction by linear interaction energy methods. *J. Comput.-Aided Mol. Des.*, v.12, n.1, p. 27-35, 1998.
- HAYAKAWA, K. *et al.* Determinação de atividades específicas e constantes cinéticas de biotinizase e lipoamidase em LEW rato e *Lactobacillus casei* (Shirota). *J Chrom Analyt Tec. Biomed.*, v.2, p.240-50. 2006.
- HARTREE, D. R.; *Proc. Cambridge. Phil. Soc.*, v.21, p.625, 1923.
- HARTREE, D. R.; *Proc. Cambridge. Phil. Soc.*, v.22, p.464, 1924.
- HÖLTJE, H. D. & FOLKERS, G. *Molecular modeling: basic principles and applications*. Weinheim: VCH. p.194. 1996.
- HOUSE, J. E. *Fundamentals of Quantum Chemistry*, Elsevier, San Diego, 2004.
- HYPERCHEM e (TM) *Computational chemistry manual*, 1994, (@c) hypercube inc.
- JABRI, E. *et al.* The crystal structure of urease from *Klebsiella aerogenes*. *Science*. v.268, p. 998-1004. 1995.
- JENSEN, F. *Introduction to computational chemistry*. Chichester: John Wiley & Sons, p.429. 1999.

JORGENSEN, W. L. & TIRADO-RIVES, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, v.110, p.1657-1666, 1988.

KARPLUS, P. A. *et al.* 70 Years of Crystalline Urease: □ What Have We Learned? *Acc. Chem. Res.*, v.30, p.330. 1997.

KITCHEN, D.B., *et al.* Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews in Drug Discovery*, v.3, p.935-949. 2004,

KLEBE, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today*, v. 11, p.580-594, 2006.

KOLLMAN P. A. Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.*, v.93, p.2395-23417, 1993.

KORB, O., *et al.* An ant colony optimization approach to flexible protein-ligand docking. *Swarm Intelligence*, v.1, 2, p.115-134, 2007.

KORB, O., *et al.* Empirical scoring function for advanced protein ligand docking with PLANTS. *Journal of Chemical Information and Modeling*, v.49, n.1,p.84-96, 2009.

KUNTZ, I.D. *et al.* Geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, v.161, p.269-288. 1982.

LINEWEAVER, H & BURK, D. Determinação de Constantes de inibição de enzimas. *J. A. Chem.*, v.3, p.53. 1934.

LEACH, A. R. *Molecular Modelling - Principles and Applications*. England: Person Prentice Hall, p.744. 2001.

Li, L. *et al.* On the Dielectric “Constant” of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi, *J. Chem. Theory Comput.*, v.9, 2126–2136. 2013

LINDEN, R. *Algoritmos Genéticos: Uma Importante Ferramenta da Inteligência Computacional*. Rio de Janeiro: Editora Brasport, p.428. 2006.

MCCARTY G. W. & BREMNER J. M., Laboratory evaluation of dicyandiamide as a soil nitrification inhibitor. *Commun. Soil Sci. Plant Anal.* V.20, p.2049–2065. 1989.

MARSHALL, B. J. & WARREN, J. R. Unidentified Curved Bacilli in the Stomach of Patients with Gastritis and Peptic Ulceration. *Lancet*, p. 1311-1315. 1984.

MITCHELL, M. *An Introduction a Genetic Algorithm*. London: MIT Press, 1998.

- MONKS, T. J. *et al.* Quinone, Chemistry and Toxicology. Appl. Pharmacol. v.2, p.112. 1992.
- MOYO, C.C. *et al.* Temperature Effects on Soil Urease Activity. Soil Biol. and Biochem. v.21, n.7, p.935-938. 1989.
- MORGON, N. H. & COUTINHO, K. Métodos de Química Teórica e Modelagem Molecular. São Paulo: Livraria da Física, p.539. 2007.
- MORRIS, G.M., *et al.* Automated docking using a Lamarckian Genetic Algorithm and an empirical binding free energy function. Journal of Computational Chemistry, v.19, p.1639-1662. 1998.
- MUEGGE, I. & MARTIN, Y.C.. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. Journal of Medicinal Chemistry, v.42, p.791-804. 1999.
- MUSIANI, F. *et al.* Structure-based computational study of the catalytic and inhibition mechanisms of urease. *J. Biol. Inorg. Chem.* v.3, p.300–314, 2001.
- OLIVEIRA, F. G. *et al.* Molecular docking study and development of an empirical binding free energy model for phosphodiesterase 4 inhibitors. Bioorg. Med. Chem., v.14, p.6001-6011, 2006.
- OVERREIN, L.N. & MOE, P.G. Factors affecting urea hydrolysis and ammonia volatilization in soil. Soil Science Society of America Proceedings, v.31, p. 57-61. 1967.
- OPREA, T. I. Chemoinformatics in drug discovery. Weinheim: Wiley-VCH, p.493. 2005.
- PEARSON, M. A. *et al.* Biochemistry, v.36, p.8164-8172. 1997.
- PEARSON, M. A. *et al.* Structures of Cys319 Variants and Acetohydroxamate-Inhibited *Klebsiella aerogenes* Urease. *Biochemistry*, v.36, n.26, p.8164–8172, 1997.
- POPLE, J. A. & NESBET, R. K. Self-consistent orbitals for radicals. J. Chem. Phys., v.22, n.3, p.571-572, 1954.
- POPLE, J. A. BEVERIDGE, D. L. DOBOSH, P. A. Approximate Self-consistent Molecular Orbital Theory V. Intermediate Neglect of Differential Overlap. J. Chem. Phys., v.6, n.47, p.2026, 1967.
- VASSILIOU, S. *et al.* Computer-Aided Optimization of Phosphinic Inhibitors of Bacterial Ureases. J. Med. Chem., v. 53 (15), p.5597–5606. 2010.

VASSILIOU, S. *et al.* Design, Synthesis and Evaluation of Novel Organophosphorus Inhibitors of Bacterial Ureases. *J. Med. Chem.* v.51, p.5736-5744. 2008.

RAIJ, B. V. *et al.* Análise química para avaliação da fertilidade de solos tropicais. Campinas, Instituto Agrônômico, p.285. 2001.

REEVES, C. R. *Modern Heuristic Techniques for Combinatorial Problems.* London: McGraw-Hill, 1995.

ROCHA JR., J. G. Desenvolvimento de um Modelo Empírico de Predição da Atividade de Inibidores da Esterol 14 α -Desmetilase (CYP51) utilizando o Método Semi-Empírico PM6. Dissertação. 2009.

ROSENZWEIG, A. C. & DOOLEY, D. M. Bioinorganic chemistry: Editorial overview. *Current Opinion in Chemical Biology*, v. 10, p.89-90, 2006.

ROOTHAAN, C.C.J. New developments in molecular orbital theory. *Rev. Mod. Phys.*, v.23, n.2, p.69-89, 1951.

ROOTHAAN, C. C. J. Self consistent field functions for the atomic configurations 1S², 1S²2S, e 1S²2S². *Rev. Mod. Phys.*, v.32, v.2, p.179-185, 1960.

SANT'ANNA, C. M. R. Glossário de Termos Usados no Planejamento de fármacos (Recomendações da IUPAC para 1997). *Quím. Nov.* v.25, n.3, p.505. 2002.

SANT'ANNA, C. M. R. Métodos de Modelagem Molecular para Estudo e Planejamento de compostos bioativos: Uma Introdução. *Ver. Virtual Quim.* v.1, n.1, p.49-57. 2009.

SCHNEIDER, G. & BÖHM, H-J. Virtual screening and fast automated docking methods. *Drug Discovery Today*, v.7, p.64-70. 2002.

SERGEEVA, M. V. & CATHERS, B. E., *Biochem. Pharmacol.* v.65, p.823. 2003.

SILVA, C. M. & BISCAIA, E. C. Genetic algorithm development for multi-objective optimization of batch free-radical polymerization reactors. *Comp. Chem. Eng.* v.27, p.1329-1344. 2003.

Sítio <http://www.dockthor.lncc.br>, acessado em novembro de 2013.

Sítio <http://www.graphpad.com>, acessado em janeiro de 2014.

Sítio <http://openmopac.net/index.html>, acessado em outubro de 2013.

Sítio <http://openbabel.org>, acessado em 2012.

SLATER, J. C. The theory of complex spectra. *Phys. Rev.*, v.34, p.1293-1322, 1929.

SMYJ, R. P. A Conformational Analysis Study of a Nickel(II) Enzyme: Urease. *J. Mol. Struct.* v.391, n.3, p. 207. 1997.

SMOOT, D.T. *et al.* Helicobacter Pylori Urease Activity is Toxic to Human Gastric Epithelial Cells. *Infect. Immun.* v.58, p1992-1994, 1990.

STEWART, J. J. P. Optimization of parameters for semiempirical methods I. *Method. J. Comput. Chem.*, v.10, n.2, p.209-220, 1989.

STEWART, J. J. P. Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations, *Int. J. Quantum Chem.* 58, 133-146, 1996.

STEWART, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.*, v.13, p.1173-1213, 2007.

STEWART, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to NDDO approximations and re-otimization of parameters. *J. Mol. Model.*, v. 19, p.1-32. 2013.

SUAREZ, D. *et al.* Ureases: Quantum Chemical Calculations on Cluster Models. *J. Am. Chem. Soc.* v.125, p.15325. 2003.

SUMNER, J.B. The Isolation and Crystallization of the Enzyme Urease. *J. Biol. Chem.* 69:435-441. 1926

TERAMOTO, R. & FUKUNISHI, H. Supervised consensus scoring for docking and virtual screening. *Journal of Chemical Information and Modeling*, v. 47, p. 526-534, 2007.

THIEL, W. & VOITYUK, A. A. Extension of the MNDO formalism to d-orbitals - Integral approximations and preliminary numerical results. *Theor. Chim. Acta*, v.81, p.391-404, 1992a.

THIEL, W. & VOITYUK, A. A. Extension of MNDO to d-orbitals - Parameters and results for the halogens. *Int. J. Quantum Chem.*, v. 44, p.807-829, 1992b.

VERDONK, M. *et al.* Improved protein-ligand docking using GOLD. *Proteins-Structure Function and Genetics*, v. 52, p. 609-623, 2003.

VERDONK, M.L., *et al.* Improved protein-ligand docking using GOLD. *Proteins*, v.52, p. 609-623. 2003.

VOTANO, J. R. *et al.* Prediction of Aqueous Solubility Based on Large Datasets Using Several QSPR Models Utilizing Topological Structure Representation. *J. Chem Info. Comput. Sci.*, v.1, n.11, p.1829-1841, 2004.

WANG, S. *et al.* Protein Kinase C. Modeling of the Binding Site and Prediction of Binding Constants. *J. Med. Chem.*, v.37, p.1326-1338, 1994.

WANG, R. X. *et al.* Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem.* v. 46, p. 2287-2303, 2003.

WEINER, S.J. *et al.* A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, v.106, p.765-784, 1984.

WOLFENDEN, R. & SNIDER, M. J. The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res.*, v.34, p.938, 2001.

YURIEV, E. & RAMSLAND, A. P. Latest developments in molecular docking: 2010-2011 in review. *J. Mol. Recognit*, v. 26, p.215–239. 2013.

ZAMAN, M. *et al.* Reducing NH₃, N₂O and NO₃⁻ - N losses from a pasture soil with urease or nitrification inhibitors and elemental S-amended nitrogenous fertilizers, *Biol. Fertil. Soils*. v.44, p.693–705. 2008.

ZAMBELLI, B. *et al.* UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn²⁺. *J Biol Chem*. v.280, p.4684. 2005.

ZIMMER, M. Are Classical Molecular Mechanics Calculations Still Useful in Bioinorganic Simulations. *Coord. Chem. Ver.* V.253, n5, p.817-826. 2009.