

UFRRJ
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
MATEMÁTICA E COMPUTACIONAL

DISSERTAÇÃO

MODELAGEM MOLECULAR DE NOVOS COMPOSTOS
DERIVADOS DA RIBAVIRINA COMO CANDIDATOS A FÁRMACOS
PARA O CONTROLE DA DENGUE

CAROLINA GURSKI OLANDA

2016



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
MATEMÁTICA E COMPUTACIONAL

MODELAGEM MOLECULAR DE NOVOS COMPOSTOS
DERIVADOS DA RIBAVIRINA COMO CANDIDATOS A FÁRMACOS
PARA O CONTROLE DA DENGUE

CAROLINA GURSKI OLANDA

Sob a Orientação do Professor

Carlos Mauricio Rabello de Sant'Anna

e Co-orientação do Professor

Arthur Eugen Kümmerle

Dissertação submetida como
requisito parcial para obtenção do grau
de **Mestre em Ciências**, no Curso de
Programa de Pós-Graduação em
Modelagem Matemática e
Computacional, Área de Concentração
em Matemática

Seropédica, RJ

Setembro de 2016

511.8

O42m

T

Olanda, Carolina Gurski, 1990-

Modelagem molecular de novos compostos derivados da ribavirina como candidatos a fármacos para o controle da dengue / Carolina Gurski Olanda. - 2016.

72 f.: il.

Orientador: Carlos Mauricio Rabello de Sant'Anna.

Dissertação (mestrado) - Universidade Federal Rural do Rio de Janeiro, Curso de Pós-Graduação em Modelagem Matemática e Computacional, 2016.

Bibliografia: f. 48-50.

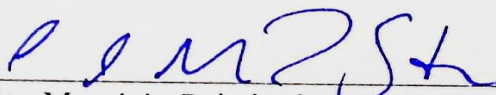
1. Modelos matemáticos - Teses. 2. Moléculas - Modelos - Teses. 3. Dengue - Tratamento - Teses. 4. Ribavirina - Uso terapêutico - Teses. I. Sant'Anna, Carlos Mauricio Rabello de, 1965- II. Universidade Federal Rural do Rio de Janeiro. Curso de Pós-Graduação em Modelagem Matemática e Computacional. III. Título.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
CURSO DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E
COMPUTACIONAL

CAROLINA GURSKI OLANDA

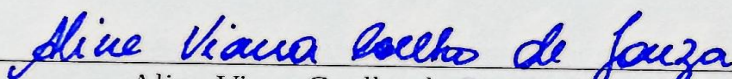
Dissertação submetida como requisito parcial para obtenção do grau de **Mestre em Ciências**, no Curso de Pós-Graduação Modelagem Matemática e Computacional, área de Concentração em Matemática.

DISSERTAÇÃO APROVADA EM ____/____/____

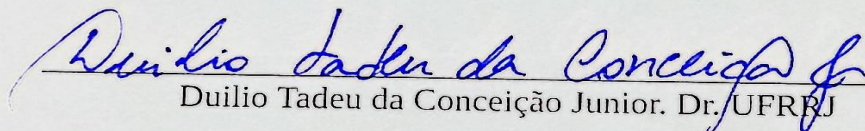


Carlos Mauricio Rabelo de Sant'Anna. Dr. UFRRJ
(Orientador – Presidente da banca)

Arthur Eugen Kümmerle. Dr. UFRRJ
(Co-orientador)



Aline Viana Coelho de Souza. Dr. CEFET



Duilio Tadeu da Conceição Junior. Dr. UFRRJ

Maria Letícia de Castro Barbosa. Dr. UFRJ (suplente)

Robson Mariano da Silva. Dr. UFRRJ (suplente)

DEDICATÓRIA

*“Porque dEle e por Ele, e para Ele,
são todas as coisas; glória, pois, a Ele
eternamente. Amém”*

Romanos 11:36

AGRADECIMENTOS

- Exalto ao meu Deus, verdadeiro e Santo, pelo infinito amor e misericórdia, que cuidadosa e carinhosamente planejou cada mínimo detalhe da minha vida, guiando os meus passos até aqui. Obrigada Senhor por ouvir minhas orações e me sustentar a cada dia, obrigada por tudo.
- Agradeço à minha mãe Carla, por manter sua fé em Deus inabalável, sendo meu alicerce forte durante toda a minha vida. Amo você.
- Aos meus tios Beth e Kiko, pelo suporte a todo e qualquer momento. Aos meus familiares pelo apoio ao longo do caminho.
- Ao meu querido namorado, Johnny, pela atenção, amizade, carinho, companheirismo a mim dedicados. Seu amor e sua ajuda foram, e são fundamentais para mim. Você é um grande presente de Deus na minha vida.
- À Mariane, por sua ajuda, incentivo, “puxadas de orelha” e orações. Obrigada por me fazer entender o verdadeiro significado da palavra amizade. Você é uma bênção na minha vida.
- À minha *sis* Vanessa, presente todos os dias dessa jornada, que se alegrou e chorou comigo a cada passo dado; sempre disposta a me ajudar. Deus me abençoou grandemente quando me deu você como irmã.
- Aos amigos: Anninha, Carol “Marmuth”, Felipe, Felipe Santos e Thuany, obrigada por fazer os meus dias mais felizes.
- Aos amigos da turma 2014 do Mestrado: Bruno, Felipe, Keiti, Marcelle, Marlon, Maurício, Pablo, Rúpila e Ulisses, que demonstraram com louvor o conceito de “a união faz a força”. Serei sempre grata pelo companheirismo de vocês a todo momento.
- Aos irmãos da Igreja Batista Manancial, pelas orações e incentivo. Em especial, à Pastora Ellen, que foi instrumento usado por Deus, para renovar minhas forças e reconstruir minha esperança.
- Ao meu Professor, orientador e amigo Carlos Maurício, pela infinita paciência, compreensão, presença e transmissão de conhecimentos ao longo desse projeto, serei eternamente grata.
- Aos companheiros de trabalho, Bruno, Catarina, Larissa e Sheisi, pelo auxílio e descontração nas conversas no laboratório.
- Aos docentes e funcionários do Programa de Pós-Graduação em Modelagem Matemática e Computacional (PPGMMC), em especial Janaína, pelo convívio e crescimento profissional.
- À Universidade Federal Rural do Rio de Janeiro (UFRRJ), minha segunda casa, por me encantar e surpreender a cada dia com sua inegável beleza e conforto.
- A CAPES pelo auxílio concedido.
- Em especial, agradeço àquela que foi e sempre será meu exemplo de filha, irmã, amiga, mulher, profissional, esposa e guerreira: Juliana Gurski (*in memoriam*). Àquela que me ensinou a lutar e vencer cada segundo e não desistir nunca. À minha irmã, pela qual meu amor e saudades serão eternos. Eu consegui Ju, e a vitória, é nossa.

RESUMO

OLANDA, Carolina Gurski. Modelagem molecular de novos compostos derivados Da Ribavirina Como Candidatos A Fármacos para o Controle Da Dengue. 2016. Dissertação (Mestrado em Modelagem Matemática e Computacional, Matemática). Instituto de Ciências Exatas, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2016.

A dengue é uma doença infecciosa causada por um arbovírus, que é transmitido pelo mosquito *Aedes aegypti*. No Brasil, devido aos períodos chuvosos no verão, é comum a ocorrência de epidemias. Nos últimos anos, o número de casos da doença no país vem crescendo de forma alarmante, aumentando não só a quantidade de registros, mas também o número de óbitos. Até o momento, não há tratamento antiviral que seja efetivo, apenas medidas para a amenização dos sintomas. A enzima viral RNA polimerase dependente de RNA (NS5) exerce papel de grande importância para a replicação viral. Este projeto tem como objetivo o planejamento racional de C-nucleosídeos 1,2,3 triazólicos derivados da ribavirina (1-[(2R,3R,4S,5R)-3,4-dihidroxi-5-(hidroximetil)oxolan-2-il]-1H-1,2,4-triazol-3-carboxamida), um potente inibidor da replicação dos flavivírus, inibidores propostos das enzimas NS5. Para se alcançar este objetivo, serão usados os métodos de *docking* molecular e de cálculos semi-empíricos para o estudo das interações entre os derivados propostos por nosso grupo e as duas enzimas. A realização desse estudo servirá como base para a síntese dos derivados mais promissores, que futuramente poderão levar à criação de um protótipo de um fármaco antiviral, que contribua para o tratamento de todas as formas da dengue, assim como na prevenção da mesma.

Palavras-chave: Modelagem molecular. *Docking*. Ribavirina. Dengue.

ABSTRACT

OLANDA, Carolina Gurski. Molecular Modeling Of New Compounds Derivatives From Ribavirin As Candidates to Drugs For Dengue Control. 2016. Dissertation (Masters in Mathematics and Computational Modeling, Mathematics). Exact Sciences Institute, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2016.

Dengue is an infectious disease caused by an arbovirus which is transmitted by the mosquito *Aedes aegypti*. In Brazil, due to the rainy season in the summer, it is common the occurrence of epidemics. In recent years, the number of cases in the country is growing at an alarming rate, increasing not only the number of records, but also the number of deaths. Until this date, there is no antiviral treatment that is effective, only measures for the sweetening of symptoms. The viral enzyme RNA-dependent RNA polymerase (NS5) plays a role of great importance for viral replication. This project aims to rational planning of C-nucleosides 1,2,3 triazole derivatives of ribavirin (1 - [(2R, 3R, 4S, 5R) -3,4-dihydroxy-5- (hydroxymethyl) dioxolan-2- yl] -1H-1,2,4-triazole-3-carboxamide), a potent inhibitor of the replication of the flavivirus, proposed inhibitors. To achieve this, was used molecular *docking* methods and semi-empirical calculations for the study of interactions between the derivatives proposed by our group and the two enzymes. The realization of this study will serve as basis for the synthesis of most promising products, which in the future may lead to the creation of a prototype of an antiviral drug that contributes to the treatment of all forms of dengue, as well as in preventing it.

Keywords: Molecular Modeling. *Docking*. Ribavirin. Dengue

LISTA DE FIGURAS

Figura 1.1 Tipos de Vírus: (a) Helicoidal; (b) Poliédrico; (c) Helicoidal Envelopado; (d) Poliédrico Envelopado; (e) Complexo	5
Figura 1.2 <i>Aedes aegypti</i> (esquerda) e <i>aedes albopictus</i>	7
Figura 1.3 Sintomas da dengue clássica e dengue hemorrágica	8
Figura 1.4 Estrutura do vírus da dengue	8
Figura 1.5 Esquematização do procedimento de replicação do vírus da dengue	9
Figura 1.6 Fórmulas estruturais da ribavirina (à esquerda) e guanósina	12
Figura 1.7 Fórmula Estrutural dos Inibidores Derivados da Ribavirina Propostos	13
Figura 1.8 Fórmula Estrutural dos Substituintes Propostos	14
Figura 1.9 Funcionamento do AG (Adaptado de LUCAS, 2000)	20
Figura 3.1 Fluxograma de planejamento proposto para o desenvolvimento da dissertação.....	29
Figura 3.2 Ribavirina 5'-monofosfato , estrutura construída usando o programa Spartan'14	30
Figura 3.3 Fluxograma das Etapas da Modelagem do Ribavirina 5'-monofosfato	30
Figura 3.4 Posição de átomos nos grupos do fosfato de ribavirina	31
Figura 3.5 Substituição 2,4 com substituinte R4 após a otimização do método MMFF	33
Figura 3.6 Proteína 1R6A com a Ribavirina destacada por um círculo branco	34
Figura 3.7 Sobreposição da solução da função ChemPLP no <i>redocking</i> da ribavirina 5'-monofosfato (ciano) com a estrutura no complexo 1R6A (carbonos em verde)	35
Figura 3.8 Proteína 1R6A com sítio selecionado com raio de 15 Å (padrão de cor CPK para os átomos). O círculo branco destaca a ribavirina, enquanto a seta branca aponta para o átomo C β do resíduo	36
Figura 3.9 Docagem do inibidor substituição 1,4 com substituinte R3 (cores CPK, representação “bola e bastão”) destacado por um círculo vermelho, no sítio ativo (pdb: 1R6A). Enzima em representação “fitas”. Código de cores: Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja. (software RasMol)	37

- Figura 4.1** Gráfico de barras de ΔH_{int} (kcal/mol) da série posição. Barras em vermelho representam os cálculos feitos em meio aquoso e as barras em azul, os cálculos feitos no interior da proteína 41
- Figura 4.2** Gráfico de barras de ΔH_{int} (kcal/mol) da série pontuação. Barras em vermelho representam os cálculos feitos em meio aquoso e as barras em azul, os cálculos feitos no interior da proteína 42
- Figura 4.3** Pontos de interação (ligações de hidrogênio) entre o ligante e resíduos da estrutura substituição 1,4 com o substituinte R4 da série posição com $\square = 4,0$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja. (Programa PyMOL) 43
- Figura 4.4** Pontos de interação (ligações de hidrogênio) entre o ligante e resíduos da estrutura substituição 1,4 com o substituinte R3 da série pontuação com $\square = 4,0$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja. (Programa PyMOL) 44
- Figura 4.5** Pontos de interação (ligações de hidrogênio) entre o ligante e resíduos da estrutura substituição 4,5 com o substituinte R3 da série pontuação com $\square = 4,0$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja. (Programa PyMOL) 45
- Figura 4.6** Pontos de interação (ligações de hidrogênio) entre o ligante e resíduos da estrutura substituição 4,5 com o substituinte R4 da série pontuação com $\square = 78,4$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja. (Programa PyMOL) 46

LISTA DE TABELAS

Tabela 1 Problemas de otimização classificados quanto ao comportamento das funções objetivo e restrições	17
Tabela 2 RMSD da otimização do ribavirina 5'-monofosfato através do Método semi-empírico PM6	31
Tabela 3 RMSD da otimização do ribavirina 5'-monofosfato através do do método DFT B3LYP/6-31G*	32
Tabela 4 RMSD otimização do ribavirina 5'-monofosfato através do método <i>ab initio</i> Hartree-Fock 6-31G*	32
Tabela 5 RMSD otimização do ribavirina 5'-monofosfato através do método MMFF	32
Tabela 6 RMSD médio das funções de escore disponibilizadas no Gold	36
Tabela 7 Inibidores propostos relacionados com a quantidade de poses geradas	38
Tabela 8 Valores de pontuação da função ChemPLP, que mostram os melhores resultados gerais de cada ligante no sítio ativo (1R6A). As linhas em negrito da mostram quais ligantes obtiveram melhor desempenho no modelo em comparação com a pontuação obtida no <i>redocking</i> do ribavirina 5' monofosfato	38
Tabela 9 Valores de entalpia dos três melhores resultados das séries posição e pontuação para $\epsilon = 4,0$	42

LISTA DE ABREVIACÕES E SÍMBOLOS

ΔH	Termo entálpico (entalpia)
ΔH_c	Entalpia do complexo enzima-ligante
ΔH_{int}	Entalpia de formação
ΔH_L	Entalpia do ligante
ΔH_s	Entalpia do sítio ativo
ϵ	Constante Dielétrica
Ψ	Função de onda
Å	Ângstrom(s)
1R6A	Código PDB para a 2'O metiltransferase proveniente do vírus da dengue
AG.....	Algoritmo Genético
CN	Computação Natural
DFT	Teoria do Funcional de Densidade
DNA	Ácido desoxirribonucléico
GOLD.....	Programa de atracamento molecular (“ <i>Genetic Optimisation for Ligand Docking</i> ”) – CCDC Ltd.
HF	Hartree-Fock
ICTV.....	International Committee on Taxonomy of Virus)
MetCap	Atividade S-adenosil metionina transferase
MM.....	Mecânica Molecular
MMFF.....	<i>Merck Molecular Force Field</i>
MOPAC2016..	Programa de cálculos quânticos (“ <i>Molecular Orbital Package</i> ”) - Stewart Computational Chemistry
NS5.....	Enzima
OMS.....	Organização Mundial de Saúde
PDB.....	Banco de dados de proteínas (do inglês “ <i>Protein Data Bank</i> ”)
PM6.....	Método paramétrico número 6 (do inglês “ <i>Parametric method number 6</i> ”)
PM7.....	Método paramétrico número 6 (do inglês “ <i>Parametric method number 7</i> ”)
R1.....	Substituinte Proposto. Fórmula Molecular: CH_2NH^+_3
R2.....	Substituinte Proposto. Fórmula Molecular: CH_2OH
R3.....	Substituinte Proposto. Fórmula Molecular: $\text{CH}_2\text{C}(\text{O})\text{-NHNH}^+_3$
R4.....	Substituinte Proposto. Fórmula Molecular: $\text{C}(\text{O})\text{NHC}(\text{=NH}^+_2)\text{-NH}_2$
RdRp.....	RNA polimerase dependente de RNA
RMSD.....	Desvio de mínimos quadrados, desvio quadrático médio (do inglês “ <i>root mean square deviation</i> ”)
RNA.....	Ácido ribonucléico

SUMÁRIO

1. INTRODUÇÃO	1
1.1 Modelagem Matemática	1
1.2 Vírus	5
1.2.1 Dengue: Patologia, Estrutura Viral e Replicação	6
1.3 Inibição Enzimática	9
1.3.1 Inibição Reversível	10
1.3.2 Inibição Enzimática Irreversível.....	11
1.4 NS5	11
1.5 Ribavirina e Inibidores Propostos	12
1.6 Otimização	14
1.6.1 Conceitos de Otimização	15
1.6.2 Classificação de Problemas	16
1.6.3 Metodologia de Otimização	18
1.6.4 Algoritmos Genéticos	18
1.6.5 Operadores Genéticos	21
1.6.6 Atualização e Finalização	21
1.7 Modelagem Molecular	21
1.7.1 Métodos Clássicos - Mecânica Molecular	22
1.7.2 Métodos Quânticos	22
1.7.2.1 Métodos <i>Ab Initio</i>	23
1.7.2.2 Métodos Semi-Empíricos	24
1.7.2.3 Funcional de Densidade	24
1.8 Docking	24
1.8.1 GOLD	25
1.9 Cálculos Quânticos	25
1.9.1 MOPAC	26
2. OBJETIVOS	27
3. METODOLOGIA	28

3.1 Modelagem da Ribavirina	29
3.2 Modelagem dos Inibidores Propostos	33
3.3 Escolha da Função de Pontuação	33
3.4 <i>Docking</i> dos Inibidores	37
3.5 Preparação de Arquivos de Entrada	39
4. RESULTADOS	41
5. CONCLUSÕES	47
6. REFERÊNCIAS BIBLIOGRÁFICAS	48
ANEXOS	51
Anexo A	52
Anexo B	54
Anexo C	56
Anexo D	57

1. INTRODUÇÃO

Apesar de grandes campanhas promovidas pelo governo brasileiro para a prevenção da proliferação do mosquito *Aedes aegypti*, vetor da dengue no Brasil, o número de casos registrados da doença permanece elevado, sendo o verão o período do ano que apresenta a maior taxa de ocorrências, devido às suas condições climáticas favoráveis à proliferação do inseto. Em 2016, foram registrados 1.399.480 casos prováveis de dengue no país até a Semana Epidemiológica (SE) 27 (3/1/2016 a 09/07/2016). Nesse período, a região Sudeste registrou o maior número de casos prováveis (837.400 casos; 59,8%) em relação ao total do país, seguida das regiões Nordeste (295.036 casos; 21,1%), Centro-Oeste (154.359 casos; 11,0%), Sul (76.465 casos; 5,5%) e Norte (36.220 casos; 2,6%) (PORTAL DA SAÚDE, 2016).

O vírus da dengue pertence à mesma classe (*flaviviridae*) dos vírus da hepatite C, da febre amarela e da febre do Oeste do Nilo (PORTAL DA SAÚDE, 2016). A ribavirina (1-[(2*R*,3*R*,4*S*,5*R*)-3,4-diidroxi-5-(hidroxi-metil)oxolan-2-il]-1*H*-1,2,4-triazol-3-carboxamida), um análogo da guanossina, vem sendo usada no tratamento da hepatite C (GRANDI, 2012). Entretanto, alguns de seus efeitos colaterais são considerados graves, como a anemia hemolítica (LOPES, 2004), que causa a destruição dos glóbulos vermelhos do sangue mais rapidamente que a produção do organismo para a sua substituição.

No entanto, ainda que conhecida há bastante tempo - enciclopédias chinesas datadas do período de 221 a.C. a 206 a.C., ano da dinastia Chin, relatam uma doença associada a insetos voadores e sintomas semelhantes aos da dengue - a patologia possui como tratamento apenas formas para se amenizar os sintomas que são desenvolvidos pelo indivíduo infectado. Até o momento não há nenhuma terapia antiviral disponível que possa efetivamente controlar a ação do vírus no organismo humano.

1.1 Modelagem Matemática

Quando há dificuldade em compreender alguns fenômenos ou situações que cercam o homem, ocorre a necessidade de se ter um modelo para que se torne possível o estudo de seu comportamento. Dá-se o nome de modelagem matemática quando um fenômeno ou situação é descrito matematicamente, mesmo que as formas usadas para explicá-lo não sejam

exatamente fiéis à realidade, mas uma aproximação da mesma. No campo científico, as pesquisas envolvendo modelagem, nos mais variados campos de atuação, se mostram cada vez mais eficazes, tornando-se uma das ferramentas matemáticas mais utilizadas e importantes. Os modelos matemáticos, aliados aos recursos tecnológicos atuais, contribuem efetivamente a favor da sociedade, com nos modelos que preveem o comportamento de doenças, predição de fenômenos climáticos e pesquisas de novos candidatos a fármacos, assunto proposto neste trabalho. A seguir, são descritos os principais conceitos e métodos da modelagem matemática, nos campos da educação e ciência.

A modelagem matemática engloba uma variedade de problemas motivacionais, que vão dos mais simples aos mais complexos e é utilizada em diversos campos, inclusive na educação, facilitando a aprendizagem, desenvolvendo habilidades gerais de exploração e mostrando a aplicabilidade da matemática e importância do papel sociocultural da mesma (BARBOSA, 2003). São sistematizados cinco perspectivas sobre Modelagem (KAISER & SRIRAMAN, 2006):

- **Realística ou Modelagem Aplicada:** as situações-problema são autênticas e retiradas do mundo real, utilizadas para a compreensão do mesmo;
- **Epistemológica:** as situações-problemas são estruturadas para gerarem o desenvolvimento da teoria matemática;
- **Educacional:** metas relacionadas à área pedagógica ou relativas à área de estudos matemáticas. Divide-se em:
 - **Didática:** estruturação de processos de aprendizagem e desenvolvimento;
 - **Conceitual:** introdução do conceito e desenvolvimento;
- **Sócio-crítica:** objetivos pedagógicos devem proporcionar a compreensão da natureza dos modelos matemáticos e seu papel na sociedade;
- **Contextual ou Modelagem Teórica:** os objetivos são todos voltados para a construção e desenvolvimento da teoria matemática.

Os Parâmetros Curriculares Nacionais (BRASIL, 2008) incentivam os professores a desenvolverem com seus alunos novos formatos de ensino, para que haja o aprimoramento de comunicação, resolução de problemas, tomada de decisões, criação, noção de

interdisciplinaridade, construção do pensamento crítico e abstração, importantes as suas experiências em sociedade. A modelagem matemática dentro das salas de aula proporciona aos discentes uma nova visão sobre a matemática, mostrando que a mesma é aplicável em inúmeras áreas do conhecimento, o que a torna mais um objeto de estudo de maior interesse e maior compreensão.

Para que um modelo possa ser escrito é necessário reunir informações acerca do problema motivacional a ser tratado e retirar dessas, dados que sejam relevantes para que estes possam se tornar variáveis que colaborem efetivamente para a modelagem. Duas metodologias podem ser aplicadas à modelagem matemática: a teórica e empírica. Na metodologia teórica, busca-se discutir e comprovar as conclusões teóricas acerca de um assunto, assim como discutir seu alcance, apresentado em literaturas anteriores. Não carece de pesquisas experimentais para se tornar legítima. É utilizada de forma complementar às pesquisas empíricas que procuram mostrar e avaliar os conhecidos teóricos junto à prática.

O empirismo provém de pesquisas de campo, onde o objetivo é a comprovação prática de algo, seja através de experimentos ou coleta de dados no campo. Assiste a teoria, fornecendo a comprovação prática dos conceitos encontrados nas literaturas, o que é fundamental para algumas ciências, como a biológica. Diferentemente da pesquisa teórica, a pesquisa empírica não é autossuficiente, é preciso sempre da formalização teórica, fundamental na validação da pesquisa prática. Existem distintos métodos de se modelar e a seguir, temos uma breve explicação sobre cada um.

- **Autômatas celulares:** São sistemas dinâmicos discretos, onde as variáveis de estado mantêm-se inalteradas ao longo do tempo, mudando seus valores apenas em momentos que são bem determinados. Possui o objetivo de simular sistemas onde as interações obedecem à regras simples e pré-determinadas, deixando a complexidade do sistema aparecer conforme as interações são realizadas. John Von Neumann e Stanislaw Ulam introduziram a teoria por volta dos anos 40, mas somente no final dos anos 60, os autômatos celulares conseguiram mais destaque com uma visão mais prática, através do matemático Conway, que desenvolveu o Jogo da vida, o exemplo mais conhecido de automata celular (PASCOAL, 2005).
- **Equações Diferenciais - parciais e/ou ordinárias:** Equações e sistemas de equações são usados para descrever a variação de fenômenos, que normalmente possuem um grau de complexidade maior. Resolvendo-se a equação, ou sistema de equações, pode-se

possivelmente prever o comportamento do fenômeno em questão, assim como extrair dados relevantes.

Equações diferenciais (parciais ou ordinárias) compõe a maioria dos estudos na modelagem matemática, sendo empregadas desde o século XVII. Em 1798, o economista inglês Thomas Malthus exibiu um dos modelos de crescimento populacional mais conhecido, onde a taxa que um país cresce é proporcional à população daquele país no momento. Definido $P(t)$ como população total no momento t e k uma taxa de proporcionalidade, obtemos a equação diferencial $\frac{dP(t)}{dt} = kP$. Ainda no século XVII, o matemático belga Pierre Verhulst, fez um modelo de crescimento logístico. Sua equação é basicamente o modelo de Malthus modificado:

$$\begin{cases} \frac{dP}{dt} = rP \left(1 - \frac{P}{P_\infty}\right) \\ P(0) = P_0, \quad r > 0 \end{cases}$$

onde a população tende à sua capacidade máxima $P \rightarrow P_\infty$, quando $t \rightarrow \infty$.

Um modelo matemático de um fenômeno com equações diferenciais é estruturado com dados sobre as taxas de variação do fenômeno (derivadas), o que permite a escrita de uma equação, que relacionam tais taxas e a função, que por sua vez resolvida, trará informações pertinentes ao comportamento do fenômeno estudado. Parâmetros (constantes) que interferem no fenômeno, variáveis (independentes) que não são alvo do estudo e variáveis (dependentes) que são o objetivo, também são incluídos na formulação das equações diferenciais. Modelos que procuram retratar dois ou mais aspectos de um fenômeno são descritos através de sistemas de equações diferenciais. Um exemplo para isto é o modelo de Lotka-Volterra (presa-predador), desenvolvido em 1920, dado por:

$$\frac{dx}{dt} = f(x) - h(x, y)$$

$$\frac{dy}{dt} = -g(y) + kh(x, y)$$

onde $f(x)$ é o crescimento da presa, $g(y)$ a mortalidade do predador, $h(x, y)$ a predação e k o coeficiente da conversão de presas em biomassa do predador.

Devido à complexidade de alguns casos, a resolução de alguns sistemas de equações se torna impossível, sendo necessárias ferramentas computacionais. Programas são

desenvolvidos especialmente para servir à modelagem matemática, como o privado MATLAB (THE MATHWORKS, INC. 2016) e o gratuito Scilab (SCILAB ENTERPRISES, 2015).

1.2 Vírus

Vírus - do latim *virus*, veneno ou toxina - são seres submicroscópicos acelulares que não pertencem a nenhum reino, formados por proteínas (alguns apresentam também envelopes formados por lipídeos, proteínas e carboidratos), um único tipo de ácido nucléico e que não possuem a capacidade de reprodução e metabolismo. São estruturados em cápsulas predominantemente proteicas (capsídeos) que protegem o genoma, apresentado em cadeias simples ou duplas de RNA ou DNA. Cada capsídeo é composto por capsômeros, que são subunidades proteicas, podendo ter seus tipos de proteínas variados. Pode haver ou não a presença de um envelope, o que classifica o vírus em envelopado ou não envelopado. A estrutura do capsídeo estará diretamente relacionada ao formato do vírus, que pode ser classificado como: helicoidal, poliédrico, envelopado (helicoidal envelopado ou poliédrico envelopado) e complexo (TORTORA, 2010)

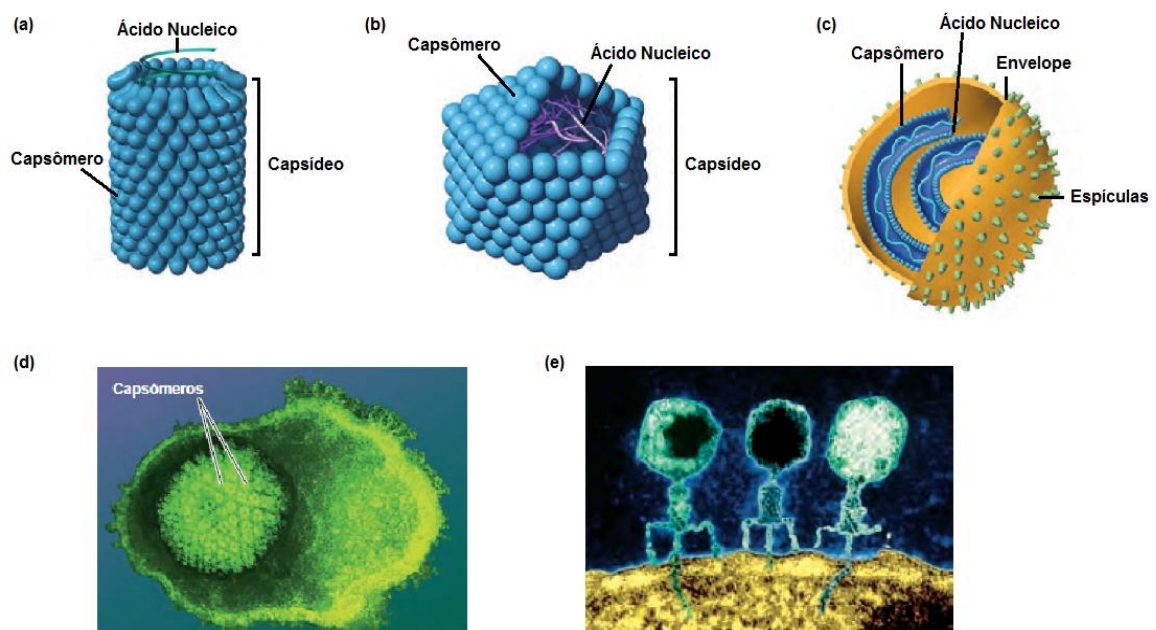


Figura 1.1 Tipos de Vírus: (a) Helicoidal; (b) Poliédrico; (c) Helicoidal Envelopado; (d) Poliédrico Envelopado; (e) Complexo (Tortora, 2010).

Em 1966, para a taxonomia dos vírus, o ICTV (*International Committee on Taxonomy of Viruses*) os agrupou em famílias com base no tipo de ácido nucléico viral, na estratégia de

replicação e na morfologia. O sufixo *virus* é usado para os gêneros, enquanto as famílias de vírus recebem o sufixo *viridae*, e as ordens, o sufixo *ales* (TORTORA, 2010). Seu cultivo é feito somente em culturas de células e sua replicação se dá através da síntese realizada no interior de células hospedeiras (intracelular obrigatório). A replicação dos vírus pode ser descrita em sete etapas (CARTER **Erro! Fonte de referência não encontrada.**, 2007):

- Ancoramento do vírion (partícula viral completa) na superfície da célula hospedeira;
- Entrada na célula;
- Transcrição de genes do vírus em moléculas de RNA mensageiro (mRNA);
- Tradução de mRNAs em proteínas virais;
- Replicação do genoma;
- Montagem de proteínas e genomas em vírions;
- Saída dos vírions da célula.

Também segundo o ICTV, uma espécie viral é uma população de vírus com características similares (incluindo morfologia, genes e enzimas) que ocupam um nicho ecológico específico (TORTORA, 2010) e o seu alcance de infecção (virose) consiste na especificidade do vírus para um determinado tipo de célula hospedeira, na qual estão incluídos bactérias, fungos, plantas, invertebrados e vertebrados.

1.2.1 Dengue: Patologia, Estrutura Viral e Replicação

Dentre os variados tipos de virose, há a arbovirose, que, de acordo com a OMS, é a infecção transmitida para humanos por artrópodes, como mosquitos e carrapatos. A dengue é considerada a mais importante arbovirose que afeta o ser humano, devido às suas complicações clínicas e por ser um problema de saúde de alcance mundial. O *Aedes aegypti* é o principal vetor da dengue, que é também transmitida pelo mosquito *Aedes albopictus*. A contaminação se dá através da picada da fêmea do mosquito, que pode transmitir o vírus a uma pessoa ou ser infectada por ela. Apesar dos casos diagnosticados erroneamente e de casos não relatados, a incidência de dengue no mundo está aumentando consideravelmente (WHO, 2016).



Figura 1.2 *Aedes aegypti* (esquerda) e *aedes albopictus*.

Segundo uma estimativa recente relatada pela OMS, 390 milhões de casos são registrados a cada ano, dos quais 96 milhões se manifestam clinicamente (independente da gravidade dos sintomas). Em 2012, somente no Brasil foram notificados 1,5 milhões de casos, o que representa uma taxa de aproximadamente 64% dos casos registrados na América Latina (WHO, 2016).

A doença pode ser classificada em dois tipos (PORTAL DA SAÚDE, 2016): clássica, que é uma forma mais amena da doença, e hemorrágica, que agrava o quadro clínico do paciente por causar hemorragias internas e que pode levar à morte.

Após o período de incubação, tempo entre a picada e a manifestação da doença, que pode variar de 2 a 10 dias, os sintomas começam a ser apresentados (PORTAL DA SAÚDE, 2016), sendo eles:

- Dengue Clássica: febre alta com início súbito, forte dor de cabeça, dor atrás dos olhos, perda do paladar e apetite, manchas e erupções na pele semelhantes ao sarampo (principalmente no tórax e membros superiores), náuseas e vômitos, tonturas, extremo cansaço, moleza e dor no corpo, muitas dores nos ossos e articulações.
- Dengue Hemorrágica: dores abdominais fortes e contínuas, vômitos persistentes, pele pálida, fria e úmida, sangramento pelo nariz, boca e gengivas, manchas vermelhas na pele, sonolência, agitação e confusão mental, sede excessiva e boca seca, pulso rápido e fraco, dificuldade respiratória, perda de consciência.

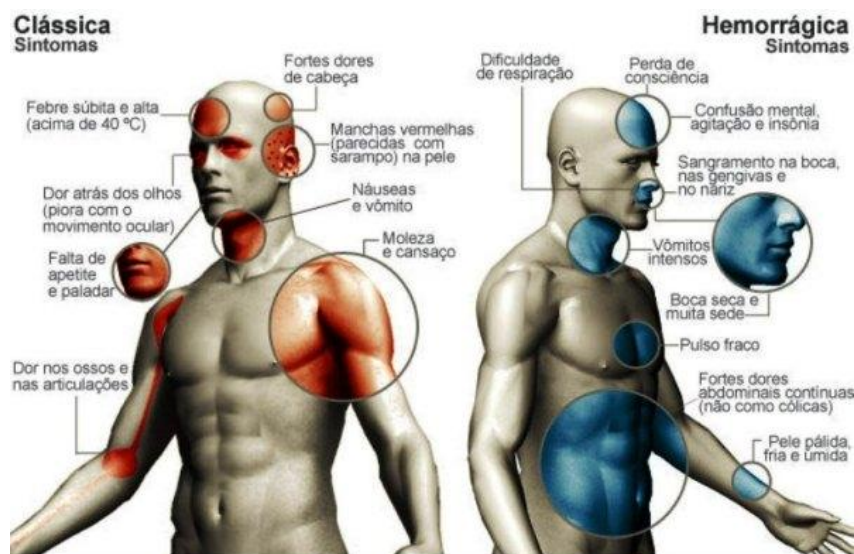


Figura 1.3 Sintomas da dengue clássica e dengue hemorrágica.

O vírus da dengue possui um envelope de proteínas com formato icosaédrico e contém uma fita negativa única de RNA.

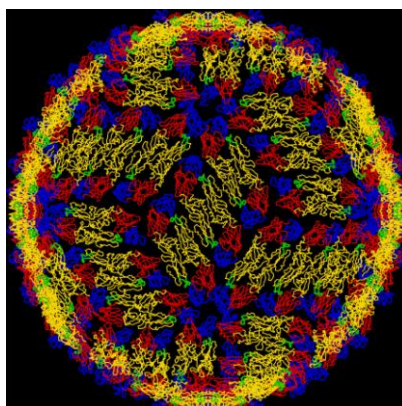


Figura 1.4 Estrutura do vírus da dengue. Fonte: Purdue University

O processo de replicação é esquematizado na **Figura 1.5**. Após o atracamento, o vírus penetra na célula hospedeira, onde ocorre a queda do pH no interior do envelope proteico que reage mudando a estrutura do capsídeo, fazendo com que as proteínas se agrupem de modo que tomem o formato de "pinças". Essas "pinças" torcem as membranas do envelope viral até que ele se rompa, ocasionando a liberação do capsídeo para o citoplasma, que se quebra, liberando o RNA do vírus. O RNA percorre um trajeto até o retículo endoplasmático (bolsas e tubos com estruturas membranosas que formam uma rede de canais interligados no citoplasma das células) e, com o auxílio de proteínas, é traduzido diretamente como uma poliproteína. O RNA se une a um complexo de proteínas, formando um círculo, para o começo da tradução e todo o genoma viral é traduzido como uma longa cadeia polipeptídica.

No citoplasma, enzimas da célula hospedeira e viral quebram ligações peptídicas entre os aminoácidos da poliproteína, isto é, quebram uma molécula complexa em moléculas mais simples. (HHMI, 2016)

O viral RNA é sintetizado da seguinte forma (HHMI, 2016):

- O complexo formado se atraca ao RNA, fazendo uma cópia positiva a partir da fita negativa, deixando o RNA com dupla hélice. Novamente, o RNA se torna um círculo e a partir disto, o cordão negativo passa a servir como modelo para se produzir uma fita positiva, em um processo de replicação contínua até que eventualmente haja uma quantidade suficiente para se formar um novo vírus;

- O envelope de proteínas se agrupa na cavidade do retículo endoplasmático, o capsídeo se agrupa do lado citoplásmico e o vírus RNA se vincula à cápsula de proteína e forma um novo “carregamento” de vírus, ainda inativo, que é expelido da célula hospedeira. Antes de atingir a superfície da célula, a pré-membrana é processada e envolvida por um novo envelope, se tornando um vírus ativo e pronto para infectar outras células.

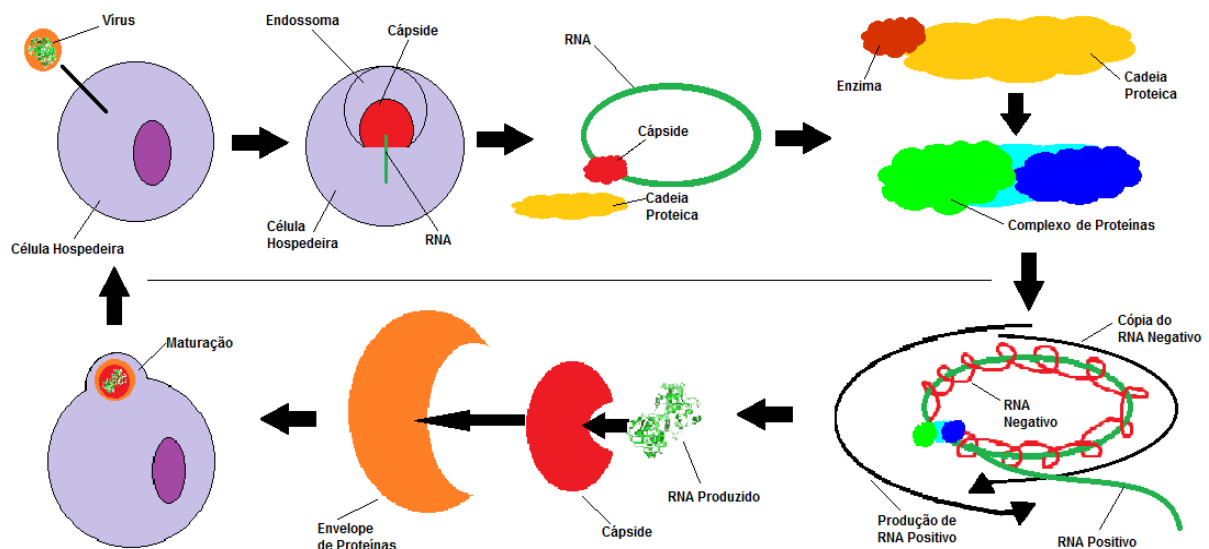


Figura 1.5 Esquematização do processo de replicação do vírus da dengue

1.3 Inibição Enzimática

A elevação da temperatura e a adição de um catalisador são dois fatores que contribuem para o aumento da reação química. O processo químico onde ocorre o aumento da velocidade de uma reação, ocasionada pela presença de espécies químicas, é denominado

catálise. Tais espécies, os catalisadores, ainda possuem a característica de não alterar a reação, com relação à quantidade e composição química (NELSON & COX, 2005).

As enzimas são proteínas especializadas na catálise de reações biológicas. Quando uma parcela das moléculas atinge certa quantidade de energia, ligações químicas podem se romper ou se formar, formando o estado de transição, a partir do qual é formado o produto. A energia necessária para se atingir o estado de transição é denominada energia livre de ativação.

As enzimas são proteínas especializadas na catálise de reações biológicas. As reações catalisadas por enzimas possuem os mesmos princípios gerais da cinética para reações, porém, com a particularidade em relação à saturação com seu substrato. Para o caso mais simples, onde há somente a presença de um substrato, a teoria desenvolvida por L. Michaelis e M. L. Menten em 1913 considera que a enzima E combina-se primeiramente com o substrato S para formar o complexo ES (enzima-substrato), que se rompe e forma a enzima livre e produto P . A equação de Michaelis - Menten expressa as relações entre a velocidade inicial V_0 de uma reação catalisada por enzimas, a concentração do substrato $[S]$ e as características da enzima, é definida matematicamente por

$$V_0 = \frac{V_{max} [S]}{K_m + [S]} \quad (\text{Eq. 1})$$

onde K_m é a constante de Michaelis - Menten, que é igual à concentração de substrato no qual a velocidade inicial e a metade da velocidade máxima (NELSON & COX, 2005).

A constante K_m da maioria das enzimas se altera com o pH, pois cada uma apresenta um pH característico em que sua atividade é máxima. Inibidores enzimáticos são agentes moleculares capazes de interferir na atividade catalisadora das enzimas, retardando ou neutralizando a reação química. Por possuírem tal propriedade, os inibidores enzimáticos são de grande importância em pesquisas da área farmacológica, por exemplo, quando a ação da enzima inibida é essencial para a sobrevivência de algum microrganismo responsável pelo processo patológico.

1.3.1 Inibição Reversível

A inibição enzimática reversível está dividida em três tipos:

- **Competitiva** - o inibidor compete com o substrato pelo sítio ativo da enzima, exigindo que a concentração do substrato para atingir a velocidade máxima

seja elevada, não ocorrendo a catálise e reduzindo a eficiência da enzima.

- **Incompetitiva** - o inibidor se combina com o complexo *ES* para um complexo *ES* que não será capaz de realizar a reação necessária para gerar o produto.
- **Não - competitiva** - Pode se combinar ao complexo *ES* ou a enzima livre. Ligam-se a um sítio da enzima, diferente do sítio ativo, causando uma mudança conformacional na molécula, que terá sua velocidade de reação diminuída e conseqüentemente, não resultará no produto esperado.

1.3.2 Inibição Enzimática Irreversível

Algumas enzimas podem ter suas funções catalisadoras completamente inibidas quando são submetidas a agentes que são capazes de modificar, através da ligação covalente e permanente, um grupo funcional necessário para a catálise. (NELSON & COX, 2005).

1.4 NS5

Existem diferentes abordagens para a realização de uma inibição antiviral, onde busca-se impedir que o vírus prossiga de uma fase para outra em seu processo de replicação. A inibição de enzimas virais tem se demonstrado a forma mais eficaz (NOBLE, 2013). A enzima NS5, presente nos quatro sorotipos do vírus da dengue, contém cerca de 900 aminoácidos em sua estrutura e exerce atividade S-adenosil metionina transferase (MetCap) e RNA polimerase dependente de RNA (RdRp) em seus domínios N-terminal e C-terminal, respectivamente.

O domínio RdRp catalisa a biossíntese da fita negativa do RNA, que funciona como um molde para a tradução de uma fita positiva, que posteriormente será utilizada como molde para novas fitas negativas, que por sua vez irão originar novas proteínas e RNAs virais e que formarão novas partículas virais. O domínio MetCap transfere um grupo metil da S-adenosilmetionina para a guanosina, etapa que completa o capeamento do RNA viral (SHATKIN, 1976). Como o capeamento do RNA é essencial para a replicação do vírus, inibidores da NS5 apresentam-se como possíveis fármacos ou quimioterápicos antivirais contra a dengue.

1.5 Ribavirina e Inibidores Propostos

A ribavirina (1-[(2R,3R,4S,5R)-3,4-dihidroxi-5-(hidroximetil)oxolan-2-il]-1H-1,2,4-triazol-3-carboxamida, **Figura 1.6**), sintetizada pela primeira vez em 1972, é um análogo sintético da guanossina (**Figura 1.6**). Como a maioria dos análogos de nucleosídeos, a ribavirina é fosforilada na posição 5' por cinases celulares após sua entrada na célula. Apresenta ação anti-viral, mas seu mecanismo de ação ainda não está completamente esclarecido.

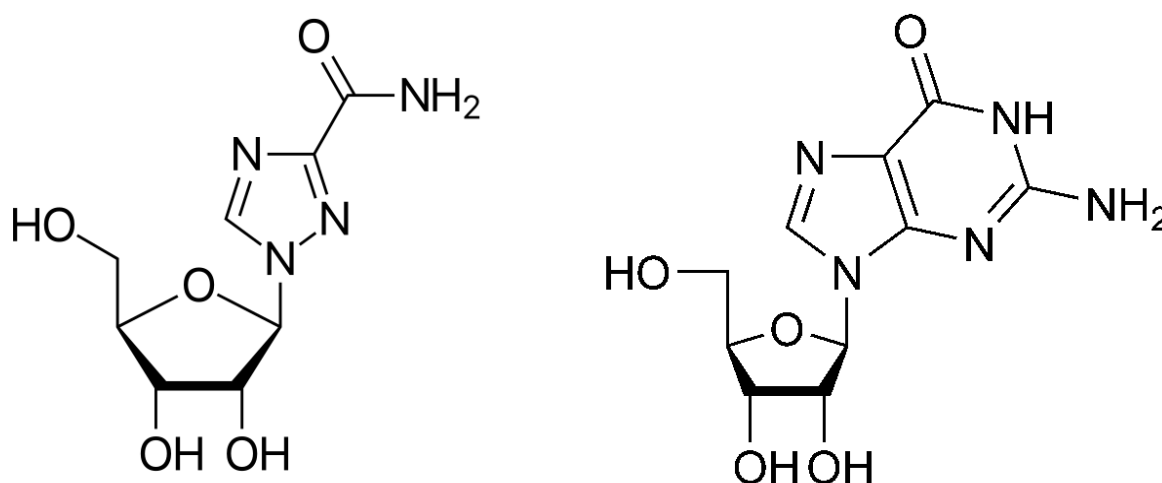


Figura 1.6 Fórmulas estruturais da ribavirina (à esquerda) e da guanossina.

Uma das possíveis razões para a ação antiviral vem do fato de que a ribavirina 5'-monofosfato é um inibidor potente da enzima celular inosina 5'-monofosfato desidrogenase, cuja inibição resulta na depleção da fonte intracelular do nucleotídeo guanossina para enzimas tanto de origem viral quanto celular.

Uma segunda possível razão estaria no fato de que o domínio N-terminal da NS5, referido como NS5MTase_{DV}, é capaz de ligar uma molécula de GTP que pode mimetizar a estrutura do RNA capeado antes da metilação (EGLOFF *et al.*, 2002). Assim, buscou-se determinar se análogos da guanossina poderiam se ligar ao sítio do GTP na NS5MTase_{DV}. Se assim fosse, esses análogos da guanossina poderiam atuar como potenciais inibidores competitivos da ligação do RNA capeado. De fato, Banarroch e colaboradores (2004) demonstraram que a ribavirina 5'-trifosfato inibe a atividade NS5MTase_{DV} do vírus da dengue. Além disso, os autores apresentaram uma visão estrutural da ligação de ribavirina 5'-trifosfato a esta enzima através da determinação da estrutura cristalina de um complexo

ternário constituído por NS5MTase_{DV}, ribavirina 5'-trifosfato e *S*-adenosil-*L*-homocisteína, com uma resolução de 2,6 Å (BANARROCH *et al.*, 2004).

Esta descrição das interações no nível atômico forneceram os primeiros detalhes estruturais da inibição de uma enzima viral pela ribavirina 5'-trifosfato, sendo sugerido pelos autores como uma base para o desenho racional de ligantes com melhor especificidade contra flavivírus.

Este projeto visa ao planejamento e preparação de *C*-nucleosídeos 1,2,3-triazólicos similares à ribavirina (**Figura 1.7**), como potenciais inibidores da NS5. Neste trabalho, são propostos inibidores com quatro grupos (R1-R4) substituídos em quatro posições diferentes do anel triazólico, totalizando 16 estruturas modeladas. Como em pH fisiológico os inibidores estariam protonados nos átomos de nitrogênio básicos das estruturas, consideramos como protonados os anéis heteroaromáticos e os substituintes (**Figura 1.8**).

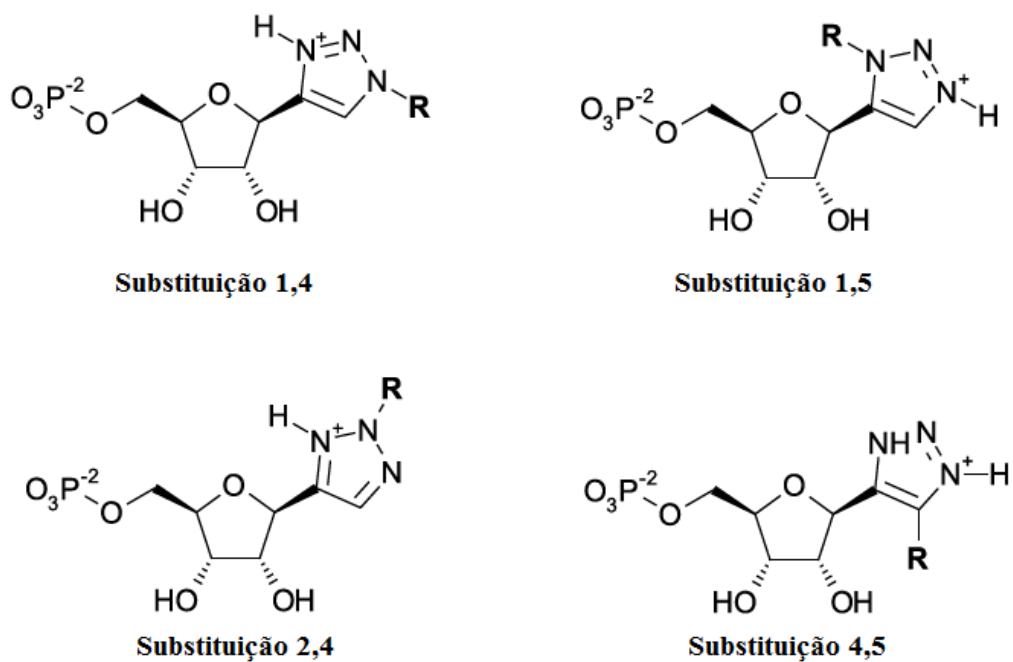


Figura 1.7 Fórmula Estrutural dos Inibidores Propostos

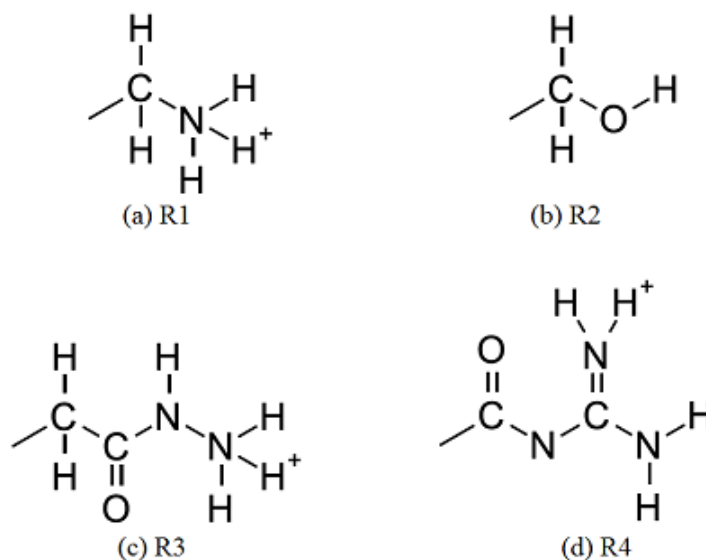


Figura 1.8 Fórmula Estrutural dos Substituintes Propostos

1.6 Otimização

Segundo o dicionário da língua portuguesa (MICHAELIS, 2016), otimizar significa: proceder à otimização de; tornar ótimo. Aceitar ou admitir como ótimo. Determinar o valor ótimo de uma grandeza. Ou seja, é uma busca dos valores que maximizam ou minimizam algo, como por exemplo, o lucro de uma empresa ou o desperdício de água em residências. A otimização pode ser aplicada em diversas áreas de conhecimento (engenharias, computação, tecnologia, medicina, ciências humanas, biológicas, exatas, da terra e outros) e é uma ferramenta muito importante para a resolução, ou ao menos a amenização, de problemas reais.

Quando um problema de otimização é matematicamente formulado, busca-se encontrar os valores ótimos de uma função, denominada função objetivo, sujeitos ou não, à certas condições. Portanto, procura-se determinar o conjunto de variáveis que fazem tal função atingir seus valores máximos e mínimos, que pode ser descrito da forma:

$$\text{Sujeito à } \begin{cases} c_i(x) \leq 0, & i=1, \dots, l; \\ g_k(x) \leq 0, & k=1, \dots, m; \\ x_i^L \leq x_i \leq x_i^U, & i=1, \dots, n. \end{cases}$$

Maximizar ou minimizar $f(x_1, x_2, \dots, x_n); x_n \in R^n$.

Onde f é a função objetivo, x_i é um ponto em R^n , $c_i(x)$ e $g_k(x)$ são funções que representam as restrições de comportamento e as funções x_i^L e x_i^U são as restrições laterais aplicadas nas x_i variáveis. Assume-se que a função objetivo e as restrições devem ser funções contínuas em R_n .

1.6.1 Conceitos de Otimização

Na otimização, faz-se uso de alguns conceitos, definidos a seguir (RITA, 2015).

- **Ponto Viável:** Ponto que satisfaz todas as restrições;
- **Região Viável:** Conjunto de pontos que satisfazem todas as restrições;
- **Função Objetivo:** Representação matemática do critério de eficiência adotado no problema de otimização. É influenciada pelas variáveis de projeto, conhecidas também como variáveis de controle do problema. Quando o problema de otimização apresenta apenas uma função objetivo, é denominado mono-objetivo, quando apresentam duas ou mais, multiobjetivos. Quando uma variável influencia o critério de eficiência, é chamada unidimensional; quando duas ou mais influenciam, define-se como multidimensional. Se a função tem um único mínimo (ou máximo), ela é chamada de função unimodal; e quando a função tem vários mínimos (ou máximos), ela é denominada multimodal.

- **Variáveis de Projeto ou Incógnitas:** Parâmetros que se deseja determinar. Requerem estar contidas dentro de um intervalo pré-definido e para as quais o problema é resolvido, pois podem adotar qualquer valor, o que pode gerar soluções inviáveis para o problema. Um limite inferior e um limite superior delimitam a variação das mesmas. Se subdivide em três tipos:

- **Contínuas:** Variáveis que podem assumir qualquer valor dentro de um intervalo de variação, variáveis reais;

- **Discretas:** Variáveis que assumem valores específicos;
- **Inteiras:** Variáveis inteiras.
- **Restrições:** São equações de igualdade ou de desigualdade que determinam os limites de disponíveis do projeto. Usualmente expressam as condições impostas pelo sistema, como a geometria, os esforços admissíveis, os recursos disponíveis, os custos envolvidos, dentre outros.
- **Espaço de Busca:** Conhecido como domínio viável do problema, é a região do domínio que satisfaz às restrições do problema. É delimitado pelas condições desejáveis do sistema e pelo intervalo de variação das variáveis do projeto.
- **Ponto Ótimo:** É o ponto, dado pelo vetor $x^* = (x_1, x_2, \dots, x_n)$, que satisfaz as restrições e são dados pelas variáveis de projeto que minimiza (ou maximiza) a função objetivo
- **Valor Ótimo:** Valor da função objetivo correspondente a $f(x^*)$.
- **Solução ótima:** Par ordenado formado pelo ponto e valor ótimos, $(x^*, f(x^*))$. Pode ser categorizado em:
 - **Local:** quando o valor ótimo é localizado numa vizinhança;
 - **Global:** quando o valor ótimo é global na região viável;
 - **Restringida:** quando atende a todas as restrições;
 - **Não restringida:** caso não atenda a alguma das restrições.
- **Programação Matemática:** Área da matemática que estuda os problemas de otimização.

1.6.2 Classificações de Problemas

Os problemas de otimização são divididos em dois grupos, caracterizados segundo o comportamento das funções do sistema e quanto às restrições impostas. Novamente, dois agrupamentos são feitos: a programação linear e a não linear. A programação linear tem como objetivo encontrar a solução ótima em problemas nos quais a função objetivo e todas as restrições são representadas por funções lineares das variáveis do projeto. Na Engenharia, são poucos os casos que podem ser reproduzidos por funções lineares das variáveis de projetos.

Sua representação é dada por uma “formulação padrão” (OLIVIERI, 2004), ou seja:

$$\text{Minimizar } f = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

Sujeita a:

$$b_{11} x_1 + b_{12} x_2 + \dots + b_{1n} x_n = c_1$$

$$b_{21} x_1 + b_{22} x_2 + \dots + b_{2n} x_n = c_2$$

. . . .

$$b_{m1} x_1 + b_{m2} x_2 + \dots + b_{mn} x_n = c_m$$

$$\text{e } x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0,$$

Onde f é a função objetivo, x_i são as variáveis de projeto, a_i , b_i e c_i são constantes do problema. A programação não linear é definida quando a função objetivo é não linear e ao menos uma das restrições é uma função não linear das variáveis do projeto. A metodologia para a solução de problemas de programação não linear se divide em dois subgrupos: determinísticos e não determinísticos. A programação não linear possui ainda um caso particular, denominado programação quadrática, em que a função objetivo é quadrática e as restrições são funções lineares das variáveis de projeto.

Abaixo encontra-se na **Tabela 1** que resume as características das classificações dos problemas de otimização.

Tabela 1 Problemas de otimização classificados quanto ao comportamento das funções objetivo e restrições.

<i>Função objetivo</i>	<i>Restrições</i>	<i>Classificação</i>
Linear	Linear	Programação Linear
Não Linear	Não Linear	Programação Não Linear
Quadrática	Linear	Programação Quadrática

A existência de restrições impostas pelo sistema, separam os problemas de otimização em duas categorias: sem restrições e com restrições.

1.6.3 Metodologias de Otimização

Muitos autores apontam diferentes propostas para a classificação de métodos de otimização, porém como cada problema tem a sua complexidade (programação linear ou não linear, problemas restritos ou irrestritos, variáveis contínuas ou discretas), particularidades do tipo de resposta esperada (otimização local ou global, projeto ótimo ou controle ótimo), técnicas empregadas (determinística, estocástica ou híbrida), número de variáveis (pequenos ou grandes), segundo a suavidade das funções (diferenciáveis ou não diferenciáveis), nenhum método é unanimamente aceito (RITA, 2015).

Entretanto, podemos classificar os métodos em duas grandes áreas: métodos determinísticos (clássicos), os quais incluem técnicas de programação linear geralmente baseados em derivadas de primeira ou segunda ordem ou aproximação destas, e não determinísticos (probabilísticos), que avaliam a função objetivo e incorporam dados estocásticos durante o processo de otimização (BASTOS, 2014).

1.6.4 Algoritmos Genéticos

Em 1858, Charles Darwin e Alfred Russel Wallace apresentaram a base de uma teoria que mostra a evolução através de seleção natural, o que contrariava a teoria criacionista, aceita pela comunidade científica à época. Com os dados recolhidos durante suas viagens e experimentos realizados, Darwin ganha fundamentação para sua obra, “*On the Origin of Species by Means of Natural Selection*”, publicada em 1859. A teoria Darwiana não influenciou somente a área biológica e ecológica, mas também os pensamentos religiosos, políticos e econômicos da época.

No mesmo período, surgia também a computação moderna, fundamentada por Charles Babbage, amigo pessoal de Darwin.

Desenvolvida em 1865 por Gregor Mendel, a teoria sobre os princípios básicos de herança genética, foi redescoberta, anos mais tarde, exercendo grande influência sobre os futuros trabalhos relacionados à evolução. A teoria da evolução moderna combina a genética e a seleção natural, criando o princípio básico de Genética Populacional: a variabilidade entre indivíduos em uma população de organismos que se reproduzem sexualmente é produzida pela mutação e pela recombinação genética. Tratando-se de algoritmos genéticos (AG), John Holland foi o primeiro pesquisador tido como referência, quando em 1975 publicou o seu livro “*Adaptation in Natural and Artificial Systems*”, ainda muito conceituado. Desde então,

AG vêm sendo aplicados com sucesso nos mais diversos problemas de otimização e aprendizado de máquina.

Nesta seção, aborda-se como os AG funcionam e, para uma maior compreensão do assunto, alguns conceitos básicos utilizados no procedimento são explicados a seguir (RITA, 2015).

- **Cromossomo:** Cadeia de símbolos representando alguma informação relativa às variáveis do problema. Cada cromossomo representa deste modo um indivíduo da população, e conseqüentemente, uma solução do problema.
- **Gene:** É a unidade básica do cromossomo. Cada cromossomo tem um certo número de genes, cada um descrevendo uma certa variável do problema.
- **Alelo:** Cada símbolo utilizado para codificação do cromossomo.
- **Locus:** Posição de um gene no cromossomo.
- **População:** Conjunto de cromossomos ou soluções.
- **Geração:** Cada passo do processo evolutivo, isto é, é o ciclo de criação e transformação de uma população, representada pelo número da iteração executada pelo algoritmo genético.
- **Operações genéticas:** Operações realizadas sobre a estrutura dos cromossomos que simulam sua recombinação e mutação. Tais operações têm por objetivo promover a evolução do indivíduo.
- **Genótipo:** Conjunto de genes de um indivíduo.
- **Fenótipo:** Conjunto das características de um indivíduo determinadas pelo genótipo.
- **Função Objetivo ou Função de aptidão:** É a função que se quer otimizar. Ela contém a informação numérica do desempenho de cada cromossomo na população. Nela estão representadas as características do problema que o AG necessita para realizar seu objetivo. É através dela que se seleciona os melhores indivíduos de cada população para a aplicação das operações genéticas, e assim, promover a evolução da solução.

O funcionamento, esquematizado na **Figura 1.9**, do AG levam em

consideração a evolução e seleção natural dos indivíduos, assim como os processos ocorridos na natureza, operando da seguinte forma (ROSA, 2009):

- A população inicial de um AG é produzida aleatoriamente e de um número (*semente*), os valores iniciais são gerados. Como os dados gerados por um AG dependem do ponto de partida, a reprodução dos resultados só será a mesma se for usada novamente a mesma *semente*.
- A função de aptidão avalia a população mensurando a capacidade de sobrevivência de cada indivíduo durante o processo evolutivo, atribuindo-os uma nota. Em problemas que com restrições, uma função denominada função de penalização avalia se a solução dada pela função de aptidão viola os limites impostos (OLIVIERI, 2004).
- Uma porcentagem dos indivíduos mais adaptados é mantida, enquanto os outros são descartados.
- Os membros mantidos passam por novas combinações, determinadas pelos operadores genéticos, podendo ser modificados pelo processo e gerando uma população descendente. O processo é repetido até que uma solução aceitável seja alcançada.

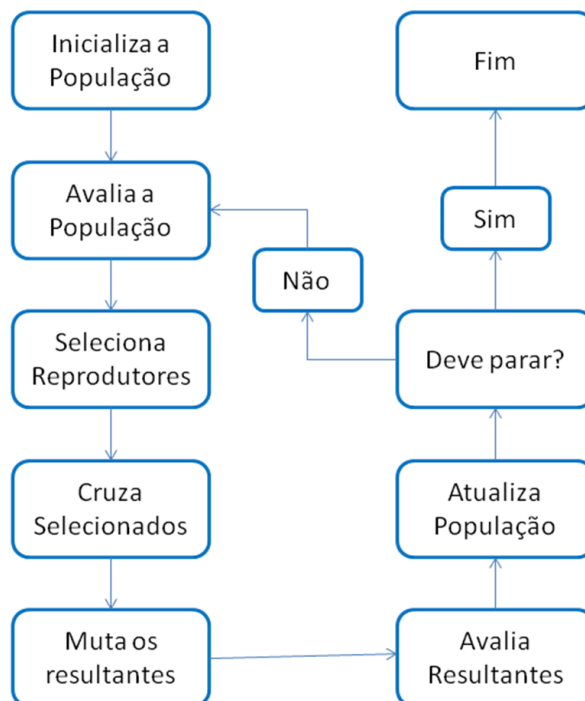


Figura 1.9 Funcionamento do AG (Adaptado de LUCAS, 2000)

1.6.5 Operadores Genéticos

Para aprimorar os resultados obtidos em cada geração durante o processo evolutivo, o AG utiliza operadores que modificam os indivíduos para que tenham uma diferenciação de indivíduos, sem que eles percam a capacidade de sobrevivência adquirida durante o processo de evolução anterior. São de extrema importância, pois fazem o AG buscar soluções em novos espaços de busca, permitindo a exploração nas regiões de domínio do problema. A seguir, dois operadores genéticos são destacados: *crossover* (cruzamento) e *mutação* (OLIVIERI, 2004).

- **Cruzamento:** Procedimento que mistura os cromossomos progenitores gerando um descendente que herde as características dos pais. Antes do cruzamento de material genético, observa-se a probabilidade de cruzamento (uma taxa de variação entre 0 e 1) entre os indivíduos, que determina se há ou não, a ocorrência do cruzamento. Caso o cruzamento não ocorra, os “pais” são copiados para a nova geração sem nenhuma modificação.
- **Mutação:** Ocorre após o cruzamento, dificultando uma convergência precipitada do problema. Troca aleatoriamente um gene de um cromossomo. Como na genética natural, a taxa de mutação deve se manter baixa.

1.6.6 Atualização e Finalização

Depois de passarem pelos operadores genéticos, os indivíduos resultantes serão inseridos na população para novamente passarem pelo processo de evolução. A população é formada segundo os critérios determinados para a atualização da população, como definir um número máximo de indivíduos, substituição de pais pelos filhos ou manter apenas certo número dos melhores indivíduos. Como critério de parada, o AG realiza um teste pré-estabelecido, como por exemplo o número de iterações feitas pelo programa.

1.7 Modelagem Molecular

Através dos métodos matemáticos utilizados na modelagem molecular, pode-se calcular a energia molecular, uma informação fundamental para determinar a estabilidade das estruturas modeladas. Há diversos modelos usados para se calcular a energia e entre os métodos mais utilizados encontram-se os baseados na mecânica clássica, como a mecânica molecular (MM), e os baseados na mecânica quântica, como os métodos de Hartree-Fock (HF), semi-empírico e Funcional de Densidade (DFT).

Métodos computacionais são empregados para o planejamento racional de compostos bioativos, baseando-se em suposições sobre o mecanismo de ação desses compostos. Os métodos computacionais não substituem os métodos convencionais para o desenvolvimento de compostos bioativos, mas são de grande auxílio, facilitando e otimizando o desenvolvimento desses compostos. Antes de discorrer sobre os resultados, um resumo sobre a metodologia da modelagem molecular e dos modelos computacionais mais comuns é feito neste capítulo.

1.7.1 Métodos Clássicos - Mecânica Molecular

A mecânica molecular (MM) descreve as moléculas como um conjunto de esferas ligadas por molas com constantes de forças específicas. Essa descrição é fundamentada porque se baseia no princípio no qual 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. Para calcular a energia, a MM considera a soma de termos de energia relacionados às posições de equilíbrio do sistema (distâncias de ligação, ângulos de ligação, ângulos diedros), que usam constantes de força para avaliar as penalidades energéticas ocorridas durante o processo, além de termos que envolvem átomos não ligados (interações de Van der Waals, ligações hidrogênio, interações eletrostáticas, etc.).

A MM é bastante abrangente, podendo ser aplicada desde moléculas de baixa massa molecular até biomacromoléculas, e é eficiente ao tratar sistemas moleculares complexos, pois possui rapidez na avaliação e, conseqüentemente, demanda computacional baixa (SANT'ANNA, 2009). O programa Spartan'14 possui disponível o método do campo de força MMFF (*Merck Molecular Force Field*), usado neste trabalho.

1.7.2 Métodos Quânticos

A mecânica quântica analisa a interação entre os elétrons e núcleos de uma molécula, utilizando as equações diferenciais da física quântica que modelam o comportamento ondulatório dos elétrons e núcleos para calcular as propriedades moleculares. Na equação definida por Schrödinger, o comportamento ondulatório dos elétrons e núcleos é descrito através da função de onda Ψ , que depende das coordenadas (R) dos núcleos e das coordenadas (r) de todos os elétrons do sistema. A integração da equação, que contém dois

operadores, energia cinética T e energia potencial V , permite o cálculo da energia total E de um sistema:

$$[T + V]\Psi(R, r) = E\Psi(R, r)$$

Como os elétrons possuem uma movimentação muito mais rápida do que os núcleos, é possível descrever a energia eletrônica separadamente da energia nuclear (Aproximação de Born-Oppenheimer).

Os métodos da mecânica quântica podem ser subdivididos em duas categorias: *ab initio* e semi-empírico. Cálculos *ab initio* usam equações exatas, sem aproximações, e envolvem a população eletrônica total da molécula (SANT'ANNA, 2002), o que demanda tempo e custo computacionais grandes. Já os cálculos semi-empíricos resolvem a equação de Schrödinger negligenciando algumas das integrais envolvidas na resolução e compensando essa negligência com o auxílio de dados empíricos provenientes de parâmetros armazenados; além disso, incluem nos cálculos somente os elétrons de valência. Portanto, seus resultados são menos exatos, porém são obtidos muito mais rapidamente e exigem menos esforço computacional, podendo ser aplicados a sistemas com número elevado de átomos.

1.7.2.1 Métodos *Ab initio*

Nestes métodos, os orbitais são representados através dos chamados conjuntos de base, que são constituídos por funções que acomodam todos os elétrons de cada átomo. O conjunto de base mínimo usa o menor número possível de funções capazes de acomodar os elétrons. Porém, o conjunto de base mínimo reduz a flexibilidade disponível para se distribuir os elétrons em cada ambiente molecular específico. Para contornar essa dificuldade, há os conjuntos de base de valência dividida, em que divide a função que representa os elétrons da camada de valência em pelo menos dois componentes, um mais interno e compacto e outro mais externo e difuso. Os conjuntos de base de polarização possibilitam deslocamentos do centro de carga eletrônica em relação às posições nucleares. (CARVALHO *et al.* 2003)

O método de Hartree-Fock (HF), onde cada elétron é considerado imerso no campo médio dos outros elétrons, foi um dos métodos escolhidos para a modelagem do ribavirina 5'-monofosfato, composto do qual os inibidores propostos neste trabalho são derivados. O conjunto de base de valência dividida 6-31G* foi o utilizado, em que os elétrons da camada interna são representados por um conjunto de seis funções gaussianas e os elétrons da camada

de valência são representados pela soma de um conjunto de três gaussianas, mais interna e compacta, e por mais uma gaussiana, mais externa e difusa, com o asterisco (*) indicando a introdução de funções de polarização para os átomos diferentes de hidrogênio.

1.7.2.2 Métodos Semi-Empíricos

Estes métodos implicam na resolução da equação de Schrödinger com o uso parcial de parâmetros originados de dados experimentais. O método PM6 - *Parametric Method 6* – (STEWART, 2007), também usado aqui na modelagem do monofosfato da ribavirina, disponibiliza parâmetros para 70 elementos da tabela periódica, incluindo metais de transição.

1.7.2.3 Funcional de Densidade

O método da Teoria do Funcional de Densidade (Density Functional Theory - DFT) surge como uma opção além dos métodos *ab initio* HF e semi-empíricos. Pode ser classificado como um método *ab initio*, por possuir a característica de não fazer uso de parâmetros provenientes de aproximações e cálculos empíricos (MORGON & CUSTODIO, 1994).

Este método faz uso das propriedades da densidade eletrônica (ρ) e não das funções de onda dos elétrons, consideradas nos métodos quânticos já descritos. De forma simplificada, a densidade ρ mede a probabilidade de encontrarmos um elétron em um ponto do espaço. As propriedades de interesse são obtidas através de um funcional, ou seja, uma função de outra função (neste caso, a densidade ρ), que contém toda a informação relevante sobre o sistema. Por exemplo, a energia E do sistema pode ser encontrada a partir de ρ como $E[\rho]$ (MARQUES & BOTTI, 2006).

1.8 Docking

Descrever a forma como estruturas se ligam a sítios ativos constitui o chamado “problema de atracamento” (*docking problem*). Deste problema surgiu a técnica de atracamento molecular, ou *docking*, que é a predição da conformação e orientação de uma pequena molécula (ligante) no sítio ativo de ligação de uma macromolécula, seguido de avaliação e classificação do modo de ligação composto. A técnica faz várias tentativas de posicionamentos (poses) do ligante dentro do sítio ativo, a fim de buscar a melhor solução.

Como há inúmeras possibilidades para se posicionar o ligante, o processo de *docking* usa um algoritmo de busca de soluções e as chamadas funções de pontuação ou funções de pontuação, que quantificam a interação do sistema ligante-proteína, permitindo um ranqueamento dos posicionamentos do ligante na proteína. Um atracamento pode ser avaliado com diferentes funções de pontuação, que são distintas entre si, pois se baseiam em cálculos de energia de interação através de diferentes fatores (DEL CISTIA, 2010).

1.8.1 GOLD

GOLD é um programa desenvolvido pelo *Cambridge Crystallographic Data Centre* – CCDC (Centro de Dados Cristalográficos de Cambridge), que permite a visualização e o *docking* entre moléculas, tendo disponíveis parâmetros que avaliam a estabilidade da interação entre as estruturas. O programa tem ainda opções de configuração que ficam a cargo do usuário, o que permite uma variedade maior de possibilidades para se realizar o *docking*, assim como um maior controle do mesmo. Para a realização do *docking* é necessária a escolha de uma função de pontuação ou escore usada para classificar as soluções encontradas através do uso de um algoritmo genético. O GOLD tem quatro funções de pontuação pré-definidas: ChemPLP (*Piecewise Linear Potential*), ChemScore, ASP (*Astex Statistical Potential*) e GoldScore, sendo possível ainda a implementação e inclusão da sua própria função de pontuação.

1.9 Cálculos Quânticos

Para indicar a estabilidade da interação entre o inibidor (I) e enzima (E), pode ser usado o cálculo da entalpia de interação (ΔH_{int}), determinada pela equação:

$$\Delta H_{int} = \Delta H_C - (\Delta H_S + \Delta H_L), \quad \text{(Equação 2)}$$

onde ΔH_C é a entalpia de formação do complexo enzima-ligante, ΔH_S é a entalpia de formação do sítio ativo e ΔH_L é a entalpia de formação do ligante. A fim de se obter um melhor resultado nos cálculos de interação, as 32 estruturas provenientes do *docking* passaram por uma nova otimização de suas geometrias feita no programa MOPAC2012, através do método semi-empírico PM7 (STEWART, 2013).

1.9.1 MOPAC

MOPAC (*Molecular Orbital Package*, Stewart Computational Chemistry) é um programa de acesso livre de química quântica semiempírico, baseado na aproximação NDDO. O hamiltoniano mais recente disponível, PM7 (STEWART, 2013), foi usado para o cálculo da parte eletrônica para se obter orbitais moleculares e calor da formação. O programa é do tipo “input” (não necessita de interface gráfica, é do tipo executável) e trabalha com as extensões para arquivos de entrada “.mop”, “.dat”, ou “.pdb”, como, por exemplo, “*nome.mop*”.

Os arquivos executados de forma bem sucedida geram um arquivo de saída “.out” que contém informações sobre o sistema, o que aconteceu durante o procedimento, entre outros resultados. Há ainda o arquivo de saída “.arc”, que resume o cálculo e contém a geometria resultante do trabalho. O MOPAC executa os cálculos através de palavras-chave inseridas no arquivo de entrada. Um exemplo de arquivo de entrada está representado no **Anexo B** (MOPAC2016).

2. OBJETIVOS

A enzima viral NS5, presente nos quatro sorotipos do vírus, desempenha papel relevante para a produção da fita positiva de RNA (OLIVEIRA *et al*, 2014), processo fundamental para a replicação do vírus.

Este trabalho apresenta o estudo da interação entre novos compostos similares à ribavirina, planejados pelo grupo de pesquisa do Professor Dr. Arthur Eugen Kümmerle (Departamento de Química, Universidade Federal Rural do Rio Janeiro) e a enzima viral NS5, com a finalidade de se obter informações para a síntese de derivados mais promissores. Serão utilizados os métodos de *docking* molecular e cálculos semi-empíricos para a análise das interações entre os inibidores propostos e a enzima NS5.

3. METODOLOGIA

A pesquisa será dividida em três momentos: modelar e otimizar os ligantes propostos, determinar os modos de interação entre os ligantes e a enzima pelo método de *docking* e otimizar as estruturas dos complexos para determinar a entalpia da interação enzima-ligante pelo método semi-empírico. A primeira etapa se baseia em construir os inibidores propostos, utilizando-se programas que permitem a avaliação por meio de modelos teóricos das interações entre os átomos da estrutura, para que as estruturas possam ser o mais próximo possível da realidade. Na segunda parte, será realizado o *docking* das moléculas nas proteínas a serem inibidas, obtendo-se soluções com diferentes orientações e conformações (poses) geradas por um AG, as quais serão ranqueadas de acordo com uma função de pontuação adequada. Destas soluções, as melhores serão selecionadas para a terceira fase, na qual serão realizados cálculos quânticos, reduzindo-se a quantidade de átomos para algumas centenas, a fim de se calcular a entalpia de interação enzima-ligante, a qual nos informará as condições de estabilidade dos complexos entre a enzima e as estruturas analisadas.

Após a realização das etapas do estudo, espera-se obter informações para se propor modificações nas estruturas e melhorar o perfil de interação com a enzima, otimizando-se o seu potencial de inibição. As novas estruturas serão sugeridas ao grupo de síntese, para que seja realizada a síntese das novas estruturas propostas, que poderão ser utilizadas para a criação de um protótipo de um fármaco antiviral, auxiliando no combate à dengue. A importância desse projeto se deve ao fato dele ser um primeiro passo para se obter compostos que poderão ser fármacos eficazes no tratamento da dengue. O fluxograma na **Figura 3.1**, mostra de uma forma concisa, o planejamento deste projeto.

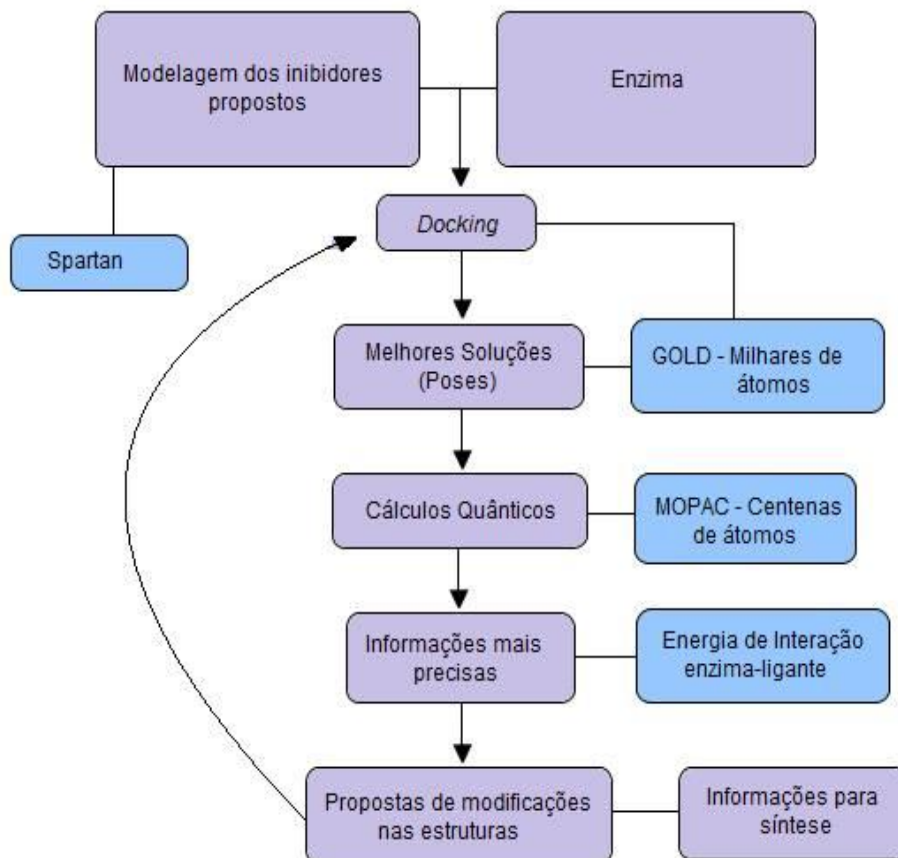


Figura 3.1 Fluxograma de planejamento proposto para o desenvolvimento da dissertação

3.1 Modelagem da Ribavirina

Com a finalidade de se determinar qual o melhor método de modelagem para a otimização das estruturas dos inibidores propostos, considerando a velocidade do método e qualidade dos resultados após os cálculos, foi realizada uma comparação da estrutura da ribavirina 5'-monofosfato modelada com os métodos de mecânica molecular MMFF, semi-empírico PM6, *ab initio* HF/6-31G* e DFT B3LYP/6-31G* com a estrutura da ribavirina 5'-monofosfato co-cristalizada na enzima inosina monofosfato desidrogenase de *Tritrichomonas foetus* (código 1ME8 no *Protein Data Bank* – PDB) (PROSISE, 2011). Para a construção do modelo da estrutura ribavirina 5'-monofosfato e cálculos de minimização de energia, foi usado o programa Spartan'14 (WAVEFUNCTION, INC.). A **Figura 3.2** mostra a estrutura construída.

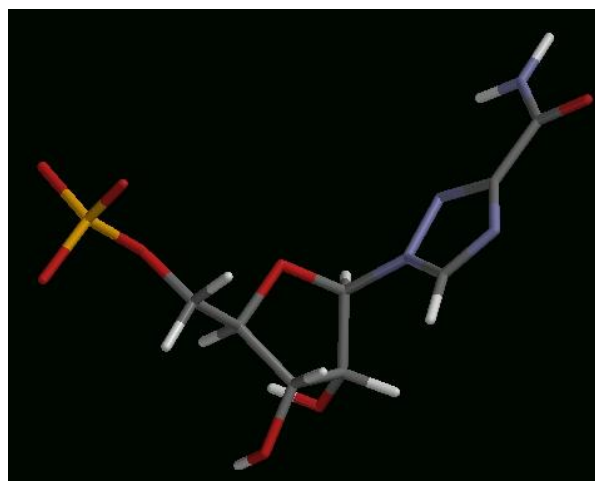


Figura 3.2 Ribavirina 5'-monofosfato , estrutura construída usando o programa Spartan'14

Foi feita a análise conformacional com o método de Monte Carlo com campo de força MMFF. Foram geradas sessenta conformações da molécula, que tiveram suas energias (em *kcal/mol*) comparadas, sendo selecionado o confômero que apresentou o mais baixo valor de energia, $-113,492$ *kcal/mol*. Todos os resultados obtidos pelo método de busca de Monte Carlo podem ser encontrados no **Anexo A**. Após a realização dos cálculos de análise conformacional, foram feitas três cópias da estrutura mais estável, para que cada uma passasse por um diferente método de otimização. A **Figura 3.3** retrata de forma mais clara os procedimentos aos quais as estruturas do ribavirina 5'-monofosfato foram submetidos.

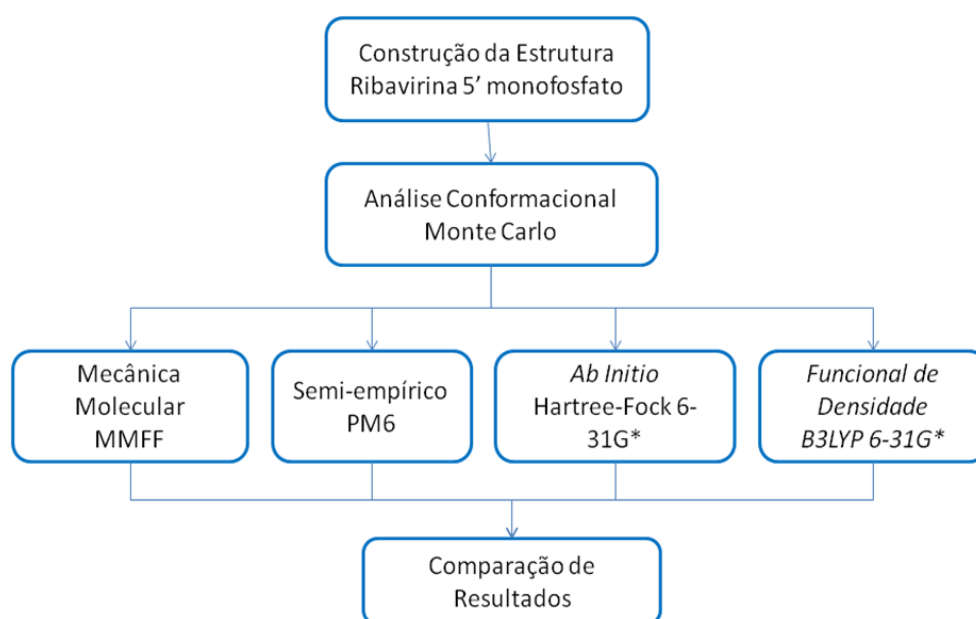


Figura 3.3 Fluxograma das Etapas da Modelagem do Ribavirina 5'-monofosfato

Foram analisadas as medidas dos ângulos e distâncias entre os átomos do grupo fosfato, do anel da pentose (anel 1) e do anel triazólico (anel 2), também presentes nas estruturas dos inibidores propostos.

A **Figura 3.4** mostra como estão dispostos os átomos utilizados na análise do grupo fosfato, anel da pentose e anel triazólico, respectivamente, do ribavirina 5'-monofosfato.

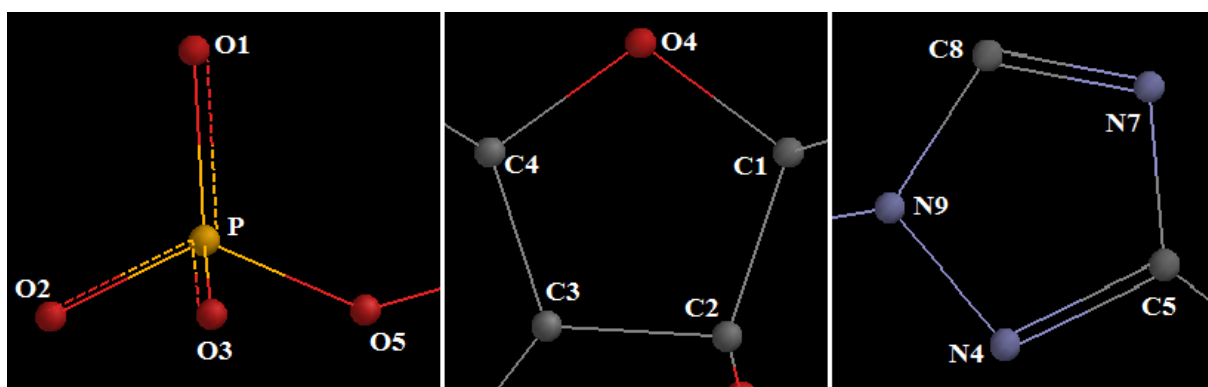


Figura 3.4 Posição de átomos nos grupos do fosfato de ribavirina

Nas **tabelas 2, 3, 4 e 5**, apresentamos os valores do RMSD retirados das estruturas resultantes de cada método de modelagem, dado por $\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta^2}$, onde δ é a diferença (valor absoluto) entre as medidas (distâncias e ângulos) da estrutura modelada e da estrutura co-cristalizada, N é o número de medidas feita no grupo em questão.

Tabela 2 RMSD após otimização da ribavirina 5'-monofosfato através do método MMFF

Grupo	Distâncias (Å)	Ângulos de ligação (°)	Ângulos Diedros (°)
Fosfato	0,03	7,58	-
Anel 1	0,021	2,3	5,71
Anel 2	0,03	3,63	-

Tabela 3 RMSD após otimização da ribavirina 5'-monofosfato através do método semi-empírico PM6

Grupo	Distâncias (Å)	Ângulos de ligação (°)	Ângulos Diedros (°)
Fosfato	0,033	8,196	-
Anel 1	0,01	2,493	34,477
Anel 2	0,049	3,681	-

Tabela 4 RMSD após otimização da ribavirina 5'-monofosfato através do método *ab initio* Hartree-Fock 6-31G*

Grupo	Distâncias (Å)	Ângulos de ligação (°)	Ângulos Diedros (°)
Fosfato	0,031	8,591	-
Anel 1	0,012	1,521	15,779
Anel 2	0,048	3,937	-

Tabela 5 RMSD após otimização da ribavirina 5'-monofosfato através do método DFT B3LYP/6-31G*

Grupo	Distâncias (Å)	Ângulos de ligação (°)	Ângulos Diedros (°)
Fosfato	0,032	7,625	-
Anel 1	0,01	1,034	8,635
Anel 2	0,05	3,551	-

Os valores de RMSD apresentados nas distâncias entre os átomos dos grupos avaliados em todos os métodos foram bons, abaixo de 0,05 Å. Para os ângulos de ligação, os valores de RMSD obtidos ficaram abaixo de 4° para os anéis 1 e 2, mas ficaram mais altos para o fosfato, acima de 7,5°. Para os ângulos diedros, os valores estão mais altos para os métodos quânticos. A interação entre o grupo fosfato e átomos de hidrogênio ligados em átomos de nitrogênio interferiram na planarização do anel triazólico em todos os métodos quânticos, o que não ocorreu para o método de mecânica molecular, MMFF. Sendo o mais rápido, sem prejudicar a qualidade, o método MMFF foi o escolhido para a modelagem dos inibidores.

3.2 Modelagem dos Inibidores Propostos

Para a modelagem dos inibidores propostos foi utilizado o programa Spartan'14, que permite a construção de estruturas e a otimização da geometria molecular pela realização de cálculos quânticos e de mecânica molecular. Após a construção das estruturas, foi empregada a mecânica molecular (campo de força MMFF) para o cálculo de análise conformacional das estruturas, aplicando-se o método de busca de Monte Carlo (SANT'ANNA, 2002), que é um procedimento de simulação que consiste em amostrar aleatoriamente o espaço conformacional de uma molécula. Para cada estrutura, foi escolhido o conformero de menor energia gerado através deste procedimento. Após esta seleção, foi feita a reotimização da geometria molecular, para se obter uma melhor interação entre os átomos que compõem o inibidor em questão (Figura 3.5).

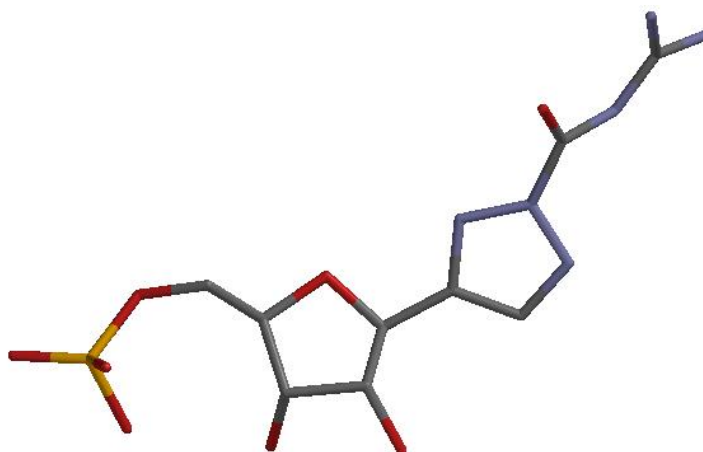


Figura 3.5 Exemplo de estrutura otimizada com o método MMFF: substituição 2,4 do substituinte R4.

3.3 Escolha da Função de Pontuação

Com a possibilidade de se utilizar mais de uma função para avaliar um *docking* realizado, surge um problema em definir qual função será a melhor opção para se escolher. Para auxiliar nessa escolha, usa-se o procedimento de “*redocking*”, que consiste em extrair de uma estrutura cristalográfica de um complexo proteína-ligante o seu ligante co-cristalizado, geralmente com estrutura parecida com os compostos a serem docados, e realizar o reancoramento em seu sítio. Para cada *pose* é disponibilizado um RMSD, obtido através da comparação do posicionamento da solução com a posição original do ligante co-cristalizado.

Assim, a função que apresentar os melhores resultados referentes ao RMSD, será a mais adequada para realizar o *docking* dos novos compostos.

Para a realização do *docking* dos inibidores propostos é usado um algoritmo genético, cujas soluções são classificadas pela função de pontuação, com o objetivo de se chegar a soluções ótimas, ou seja, quanto maior a pontuação, melhores as interações previstas entre o ligante e a proteína. Para selecionar a função a ser utilizada no projeto, foi feito o *redocking* do ligante co-cristalizado no sítio ativo da enzima com código 1R6A no PDB (**Figura 3.6**), utilizando-se as quatro funções de pontuação disponíveis no GOLD.

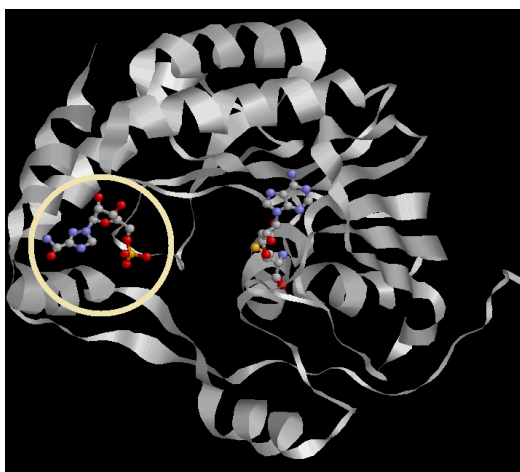


Figura 3.6 Proteína 1R6A com a ribavirina 5'-monofosfato destacada por um círculo branco. (figura gerada com o programa RasMol).

O programa GOLD 5.2 dispõe de quatro funções de pontuação pré-definidas: ChemPLP (atual função padrão), ASP, ChemScore e GoldScore. As pontuações são adimensionais e cada uma se baseia em cálculos que levam em consideração diferentes fatores, o que não permite a comparação direta entre os valores pontuados. Para contornar essa questão, avaliou-se as médias do RMSD apresentadas nas soluções obtidas do *redocking* (**Figura 3.7**) da estrutura ribavirina 5'-monofosfatoco-cristalizada com a enzima NS5 (pdb: 1R6A).

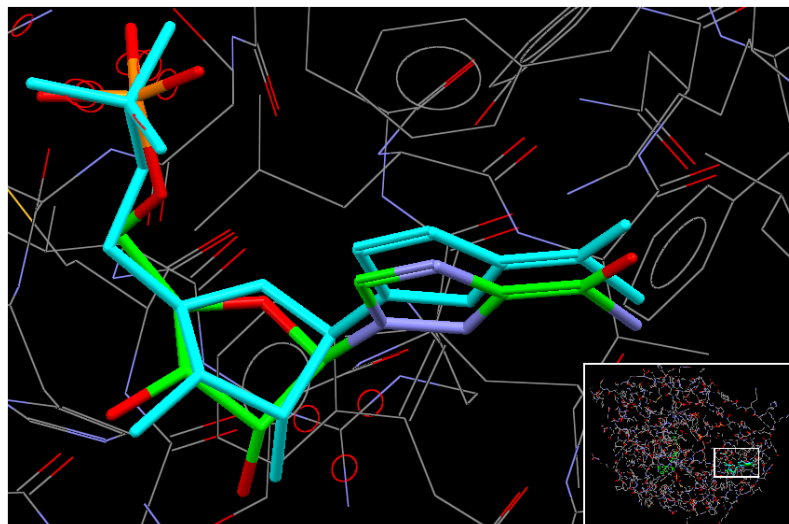


Figura 3.7 Sobreposição da solução da função ChemPLP no *redocking* da ribavirina 5'-monofosfato (ciano) com a estrutura no complexo 1R6A (carbonos em verde) (figura gerada com o programa GOLD).

Foram adicionados, através do programa GOLD, os átomos de hidrogênio para que a minimização de energia e a predição correta do ancoramento do ligante à estrutura proteica fossem possíveis. O ligante ribavirina 5'-monofosfato foi extraído da proteína, juntamente com as moléculas de água presentes. Foi selecionado um raio de 15 Å para o sítio de interação, utilizando como origem do sítio o átomo C β do resíduo Ser214 (**Figura 3.8**). A média dos RMSD, para cada função de pontuação, foi calculada utilizando o somatório dos valores dos RMSD apresentados em cada solução, que foi dividido pelo número correspondente à quantidade de soluções obtidas.

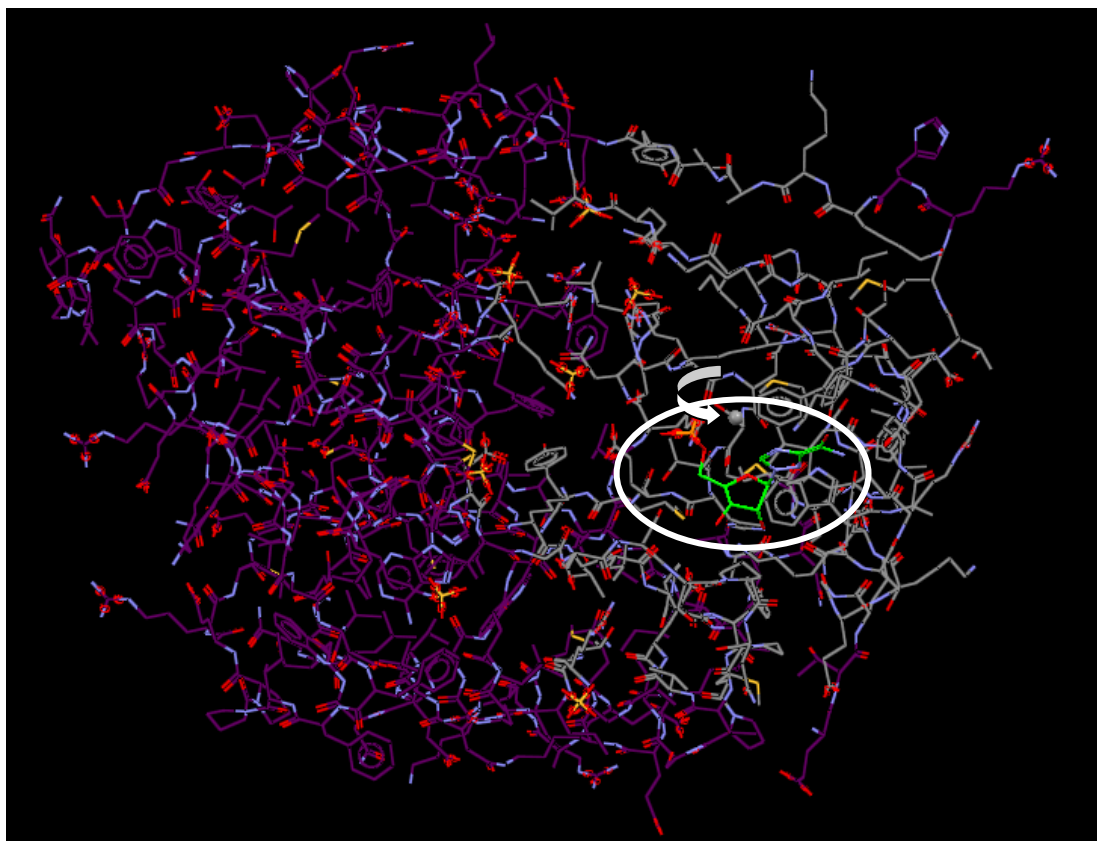


Figura 3.8 Proteína 1R6A com sítio selecionado com raio de 15 Å (átomos de C em cinza). O círculo branco destaca a ribavirina, enquanto a seta branca aponta para o átomo C β do resíduo Ser214. (figura gerada com o programa GOLD).

Após as comparações entre os resultados (**Tabela 6**), as funções ChemScore e GoldScore foram desconsideradas por apresentarem médias acima de 2 Å, valor limite considerado satisfatório pela literatura. Dentre as duas funções restantes, escolheu-se a função ChemPLP, que, somado ao fato de apresentar o melhor RMSD médio, também obteve maior velocidade dos cálculos para a geração das soluções.

Tabela 6 RMSD médio das funções de pontuação disponibilizadas no GOLD

Funções de Pontuação	RMSD médio (Å)
ASP	1,83
ChemPLP	1,22
ChemScore	4,82
GoldScore	8,37

3.4 Docking dos Inibidores Propostos

O *docking* dos inibidores (**Figura 3.9**) foi realizado no programa GOLD, com a função de pontuação ChemPLP, no sítio ativo da estrutura co-cristalizada (1R6A), escolhida como proteína-alvo por conter a ribavirina 5'-monofosfato como um inibidor da atividade NS5MTase_{DV} do vírus da dengue. Juntamente com o *docking* dos inibidores, foi realizado novamente o *redocking* do ligante ribavirina 5'-monofosfato para fins de comparação dos resultados de pontuação.



Figura 3.9. Resultado de *docking* do inibidor substituição 1,4 com substituinte R3 (cores CPK, representação “esfera e bastão”) destacado por um círculo vermelho, no sítio ativo (código pdb: 1R6A). Enzima em representação “fitas”. Padrão de cores CPK. (figura gerada com o programa RasMol).

Tomando-se o átomo C β do resíduo Ser214, o mesmo utilizado no *redocking* como centro, definiu-se um raio de 15 Å para a área de exploração das tentativas de *docking* dos inibidores propostos. Foram analisadas todas as poses geradas (**Tabela 7**) para cada um desses inibidores, sendo selecionadas em duas séries:

- **Pontuação:** Poses que obtiveram a melhor pontuação, ranqueadas de acordo com a função ChemPLP;

- **Posição:** Poses que mais se assemelhavam à posição da ribavirina 5'-monofosfato, presente na estrutura co-cristalizada.

Tabela 7 Inibidores propostos relacionados com a quantidade de poses geradas

Estrutura	Quantidade de Soluções Geradas
1,4 com R1	6
1,4 com R2	10
1,4 com R3	10
1,4 com R4	4
1,5 com R1	10
1,5 com R2	10
1,5 com R3	3
1,5 com R4	10
2,4 com R1	10
2,4 com R2	10
2,4 com R3	10
2,4 com R4	10
4,5 com R1	3
4,5 com R2	10
4,5 com R3	10
4,5 com R4	10
Ribavirina 5'-monofosfato	5

Os melhores resultados gerais de acordo com os valores de pontuação (ChemPLP) são apresentados na **Tabela 8**.

Tabela 8. Valores de pontuação da função ChemPLP, que mostram os melhores resultados gerais de cada ligante no sítio ativo (1R6A). As linhas em negrito da mostram quais ligantes obtiveram melhor desempenho no modelo em comparação com a pontuação obtida no *redocking* do ribavirina 5' monofosfato.

Estruturas	Pontuação
14 com R1	62,98
14 com R2	57,15
14 com R3	60,71
14 com R4	55,8
15 com R1	56,8
15 com R2	47,81
15 com R3	62,82
15 com R4	50,31
24 com R1	59,58
24 com R2	53,5
24 com R3	70,7

24 com R4	52,41
45 com R1	52,71
45 com R2	54,34
45 com R3	56,5
45 com R4	49,55
Ribavirina 5'-monofosfato	63,76

Assim, os inibidores foram separados em dois grupos: selecionados por pontuação e posição, em um total de 32 soluções, que passaram por uma otimização através do método semi-empírico PM7 (STEWART, 2013) no programa MOPAC2016 (Stewart Computational Chemistry), para se obter a entalpia de interação (ΔH_{int}) do complexo enzima-ligante calculada, assunto discutido nas seções 3.5 e 4.

3.5 Preparação de Arquivos de Entrada

Um recorte com 15 Å a partir do ligante foi efetuado no sítio ativo da estrutura, utilizando o programa Swiss PDB Viewer 4.1 (GUEX & PEITSCH, 1997), com o objetivo de diminuir o tamanho da estrutura com a qual seriam efetuados os cálculos quânticos. Para analisar o comportamento das estruturas sob diferentes condições, foram adicionados aos cálculos de otimização as constantes dielétricas para o interior de proteínas ($\epsilon = 4,0$) e em meio aquoso ($\epsilon = 78,4$). Esse método é chamado de aproximação do contínuo. A introdução da constante dielétrica igual a 4 no cálculo é uma maneira de compensar parcialmente o efeito dos aminoácidos mais distantes que foram excluídos no momento do recorte. Utilizando o valor 4, estamos supondo que a maior parte do espaço excluído era composto por outros aminoácidos da proteína, que estavam mais distantes. Quando utilizado o valor 78,4, estamos considerando que haveria principalmente a presença de água próxima ao sítio.

Todos os recortes foram salvos no formato “.pdb”, extensão que descreve a estrutura molecular por uma matriz, e tais arquivos foram editados com a inserção das palavras-chave necessárias para os cálculos. Para o processo de otimização, as seguintes palavras-chaves foram usadas (<<http://openmopac.net/manual/index.html>>, acessado em 01/09/2016.):

- MOZYME: Usa o método do Orbital Molecular Localizado para acelerar os cálculos executados;
- PL: Monitora a convergência da matriz densidade;
- MMOK: Usa a correção da mecânica molecular para ligações CONH;

- EPS: Constante dielétrica a ser usada no cálculo;
- CHARGE: Carga do sistema;
- PDBOUT: Gera um arquivo “.pdb” ao final do procedimento.

Cada estrutura teve sua matriz de coordenadas atômicas editada para fixação no espaço das ligações peptídicas, através da inserção do marcador (“*flag*”) +0 em substituição ao marcador +1, localizado ao lado de cada uma das três coordenadas cartesianas da matriz. Desta forma, o sítio não perde sua estrutura terciária original durante a minimização da energia (DEL CISTIA, 2010). O mesmo processo de fixação foi utilizado no sítio ativo vazio. Um exemplo de uma matriz editada está representado no **anexo B**. Todos os demais átomos da proteína e dos ligantes foram otimizados sem restrições.

4. RESULTADOS

Com os resultados gerados pela otimização semi-empírica, foi possível realizar o cálculo de entalpia de interação (Equação 2). Na figura 4.1 são comparados os valores de ΔH_{int} para a série posição, definida em 5.3, com $\varepsilon = 4,0$ e $\varepsilon = 78,4$.

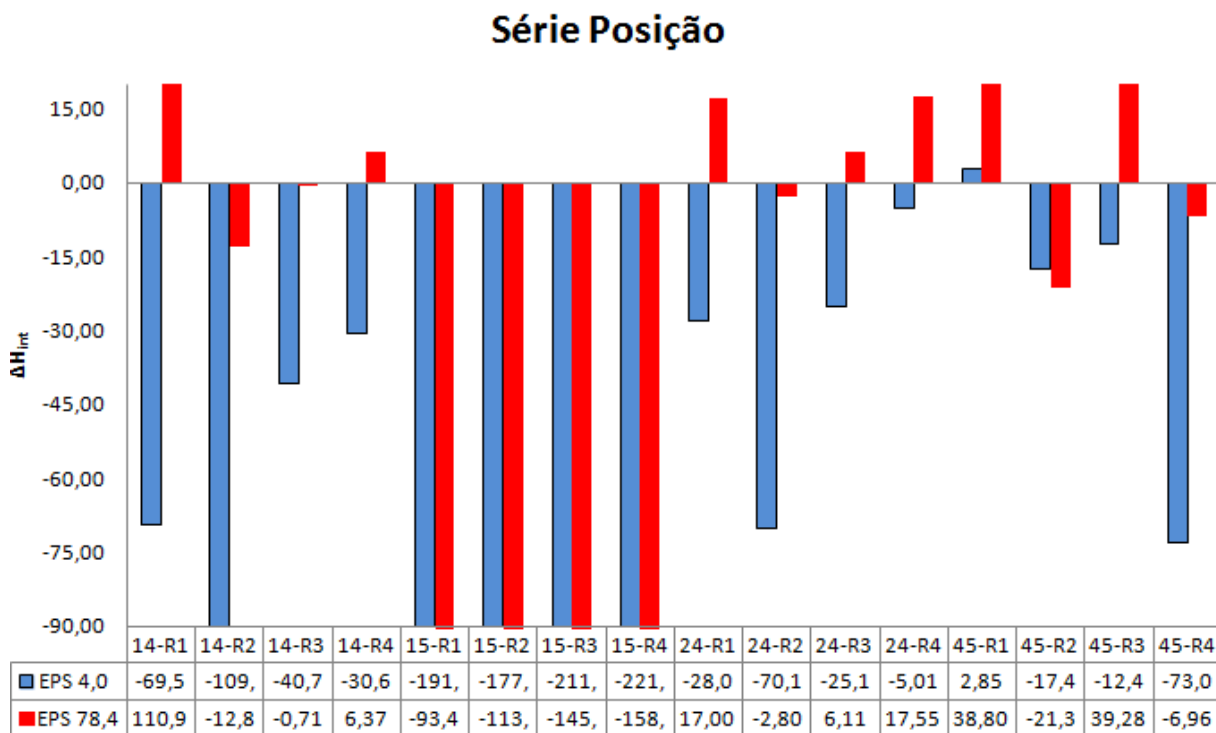


Figura 4.1: Gráfico de barras de ΔH_{int} (kcal/mol) da série posição. Barras em vermelho representam os cálculos feitos em meio aquoso e as barras em azul, os cálculos feitos no interior da proteína.

Na figura 4.2 são comparados os valores de ΔH_{int} para a série pontuação, definida em 5.3, com $\varepsilon = 4,0$ e $\varepsilon = 78,4$.

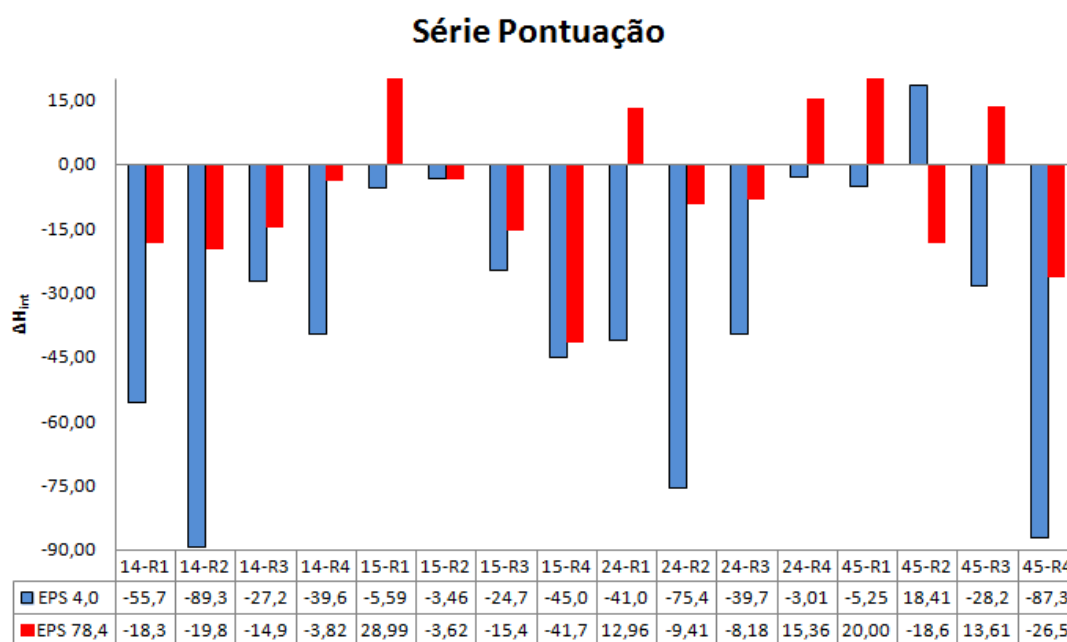


Figura 4.2: Gráfico de barras de ΔH_{int} (kcal/mol) da série pontuação. Barras em vermelho representam os cálculos feitos em meio aquoso e as barras em azul, os cálculos feitos no interior da proteína.

Utilizando-se a constante dielétrica da água para os cálculos ($\epsilon = 78,4$), observou-se que os valores de ΔH_{int} se tornam em alguns casos positivos, indicando que, nesses casos, o transporte dos ligantes para o interior da proteína não seria entalpicamente favorável no “contínuo aquoso”. A situação é mais favorável quando $\epsilon = 4$, uma vez que os valores de ΔH_{int} são quase todos negativos, nas duas séries. Todos os valores da entalpia de interação dos complexos, sítios e ligantes encontram-se no **anexo C**.

Considerando os valores de entalpia apresentados nas duas séries, foram selecionados os melhores resultados de cada uma para cada constante dielétrica. A **Tabela 9** destaca os quatro melhores valores de entalpia de interação, sendo todos com padrão 1,5 de substituição e da série posição para $\epsilon = 4,0$.

Tabela 9. Valores de entalpia dos melhores resultados das séries posição para $\epsilon = 4,0$.

Estrutura	ΔH_{int} (kcal/mol)
1,5 com R1	-191
1,5 com R2	-177
1,5 com R3	-211
1,5 com R4	-221

Os resultados da entalpia de formação para $\epsilon=78,4$ apresentaram em geral valores negativos baixos. No entanto, a substituição 1,5 com substituinte R4 para $\epsilon=78,4$ foi uma exceção a esses resultados, obtendo $\Delta H_{\text{int}} = -41,7$ kcal/mol. A situação é mais favorável quando $\epsilon=4,0$, uma vez que os valores de ΔH_{int} são quase todos negativos, nas duas séries.

A **Figura 4.3** destaca os pontos de interação por ligação de hidrogênio das estruturas com substituição 1,5 com o substituinte R3, da série posição com $\epsilon=4$.

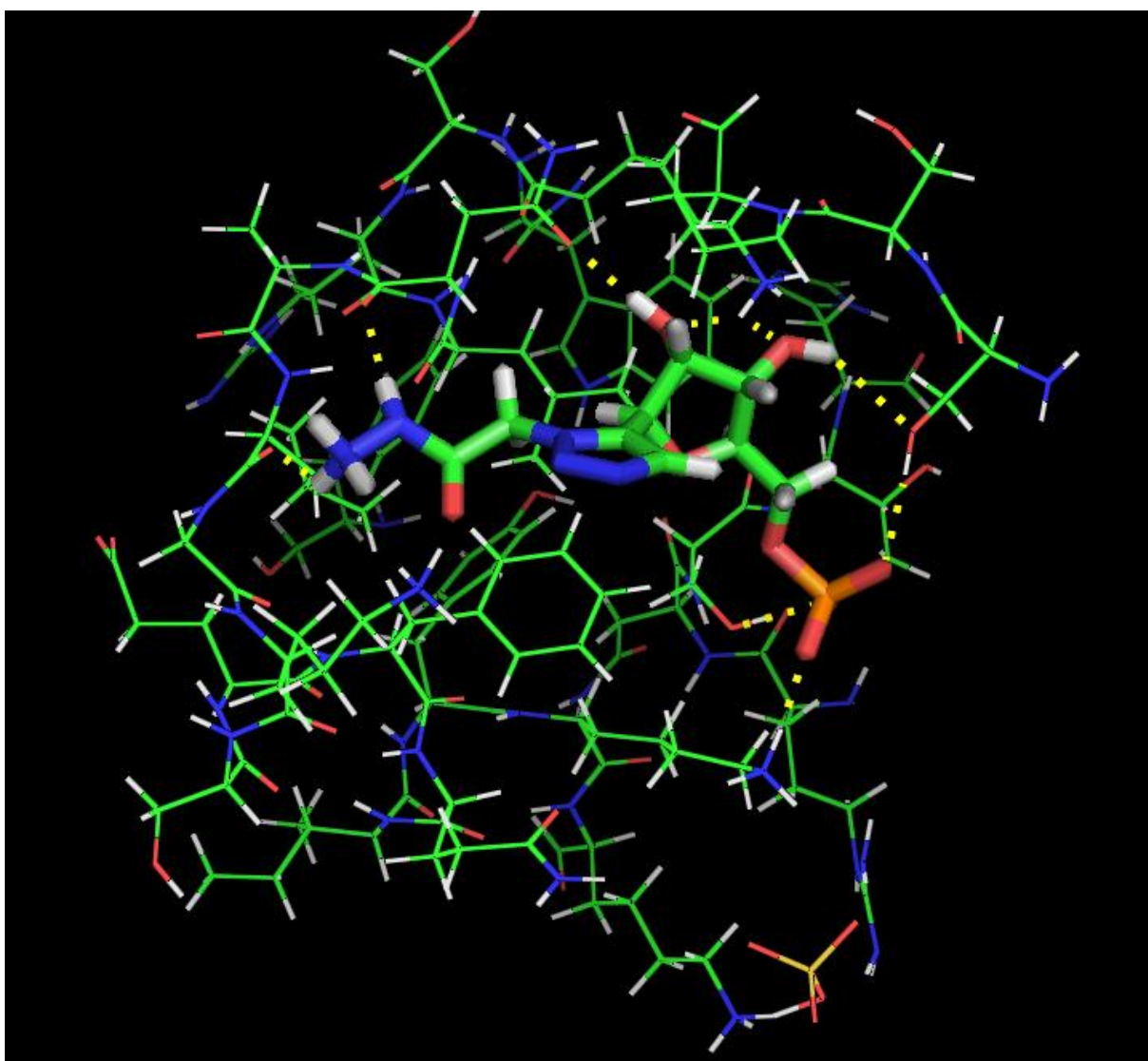


Figura 4.3 Pontos de interação por ligação de hidrogênio entre o ligante e resíduos da estrutura substituição 1,5 com o substituinte R3 da série posição com $\epsilon=4,0$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja (figura gerada com o programa PyMOL).

A **figura 4.4** destaca os pontos de interação por ligação de hidrogênio da estrutura com substituição 1,5 com o substituinte R4, com $\epsilon=4,0$; na **Figura 4.5**, a substituição 1,5 com

o substituinte R3, com $\epsilon= 78,4$; na **Figura 4.6** a substituição 1,5 com o substituinte R4 com $\epsilon= 78,4$, sendo todas da série posição.

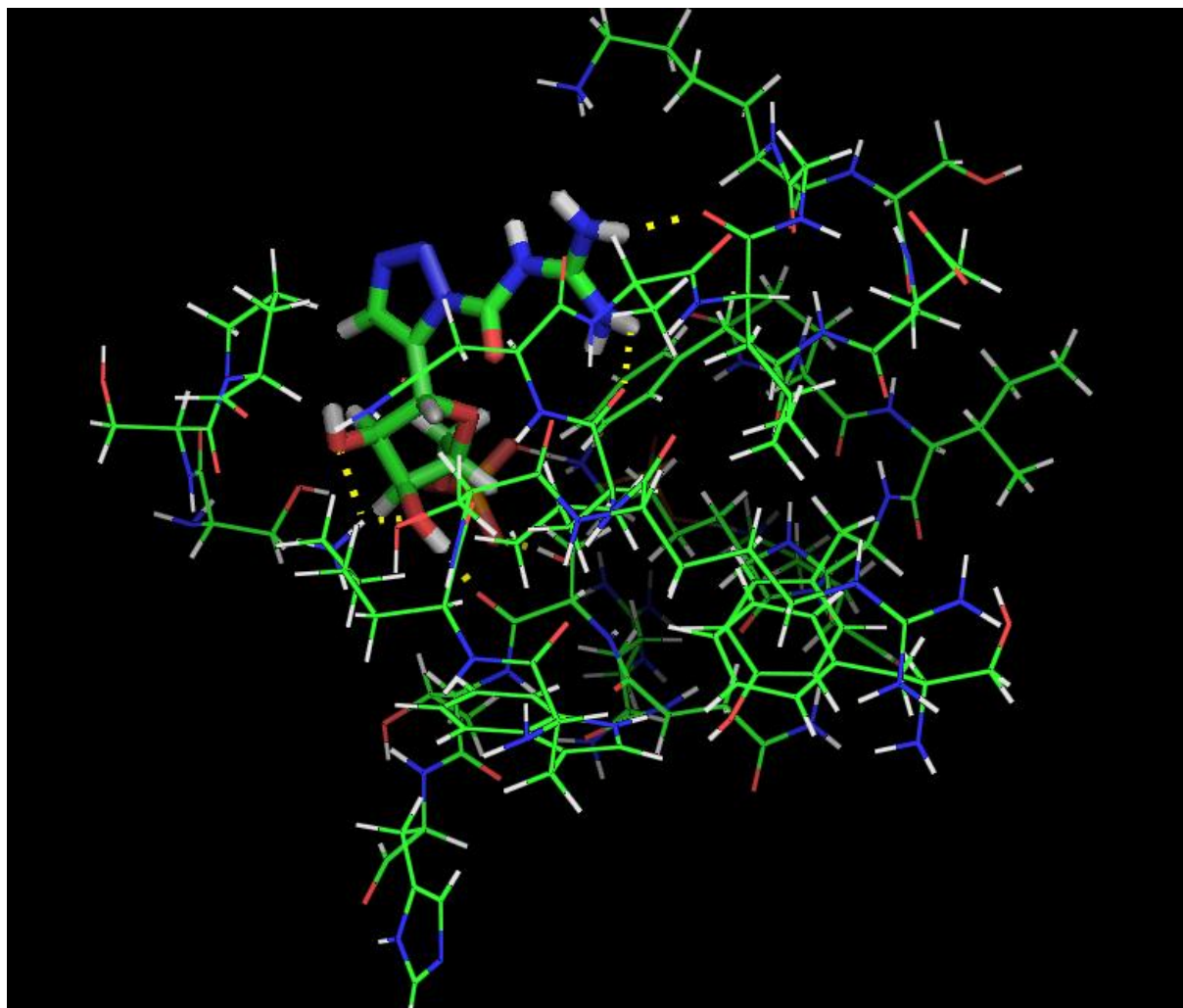


Figura 4.4 Pontos de interação por ligação de hidrogênio entre o ligante e resíduos da estrutura substituição 1,5 com o substituinte R4 da série pontuação com $\epsilon= 4,0$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja (figura gerada com o programa PyMOL).

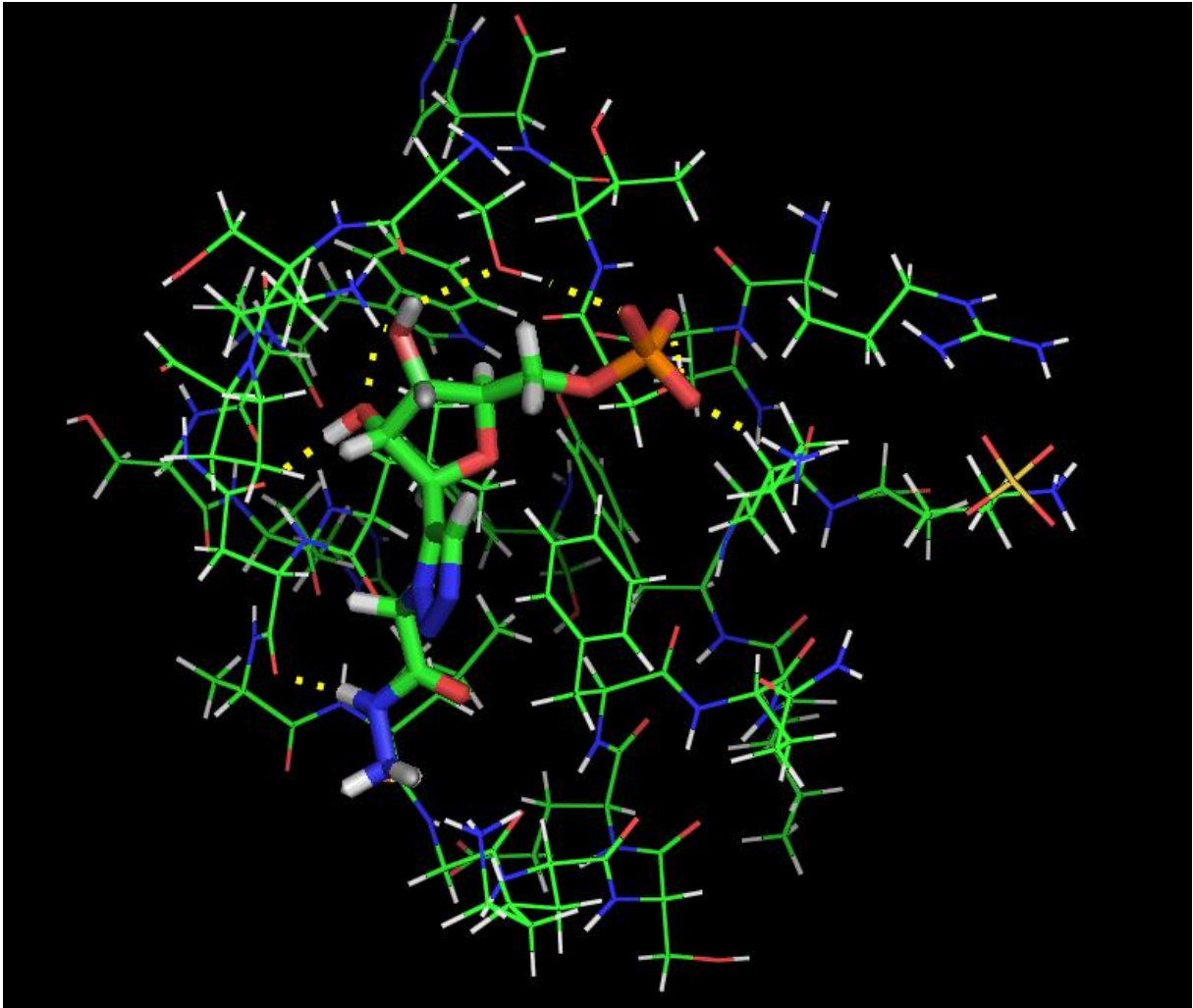


Figura 4.5 Pontos de interação por ligação de hidrogênio entre o ligante e resíduos da estrutura substituição 1,5 com o substituinte R3 da série pontuação com $\epsilon=78,4$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja(figura gerada com o programa PyMOL).

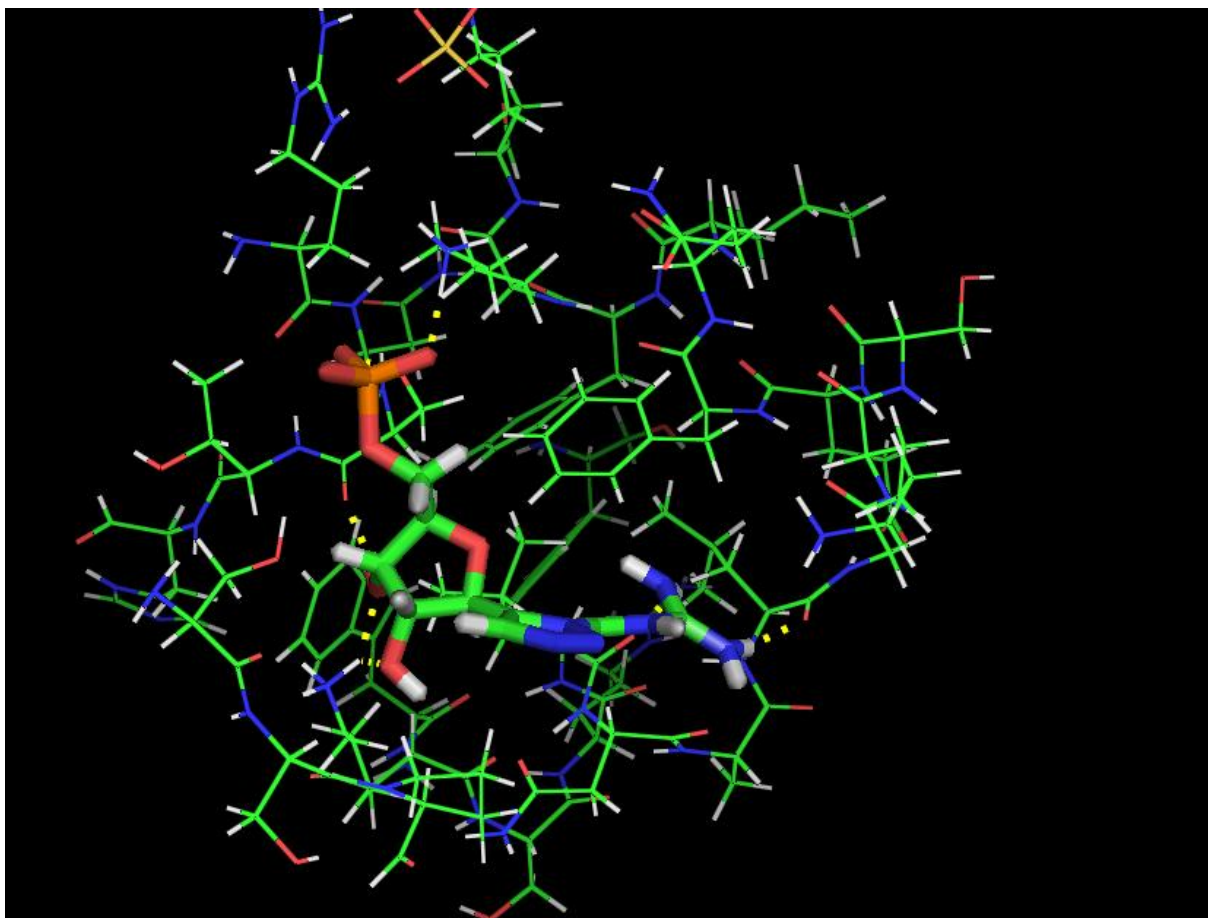


Figura 4.6 Pontos de interação por ligação de hidrogênio entre o ligante e resíduos da estrutura substituição 4,5 com o substituinte R4 da série pontuação com $\epsilon = 78,4$. São consideradas as ligações com comprimentos menores ou iguais a 3 Å. Código de cores: Carbono verde; Oxigênio vermelho; Nitrogênio azul; Hidrogênio branco; Fósforo laranja (figura gerada com o programa PyMOL).

5. CONCLUSÕES

Neste trabalho foi feito um estudo de *docking* com o método ChemPLP de dezesseis ligantes propostos, derivados da ribavirina 5'-monofosfato, na enzima NS5MTaseDV do vírus da dengue (código 1R6A no PDB), seguido da determinação da entalpia de interação pelo método semi-empírico PM7, usando-se a aproximação do contínuo em duas condições: constante dielétrica igual a 4 (interior de proteínas) e constante dielétrica igual a 78,4 (meio aquoso).

Um critério para seleção das soluções do *docking* para a etapa posterior de cálculos com o método semi-empírico PM7 foi a maior pontuação obtida no *docking*, mas, como observou-se uma grande variação nas posições dos ligantes no sítio usando esse critério, selecionou-se também um segundo grupo de soluções, usando como critério a maior semelhança com a posição da ribavirina 5'-monofosfato na enzima NS5MTaseDV.

Para os 2 grupos de soluções, foram obtidos valores em geral favoráveis para a entalpia de interação. Os valores mais favoráveis, em geral, foram encontrados para o grupo selecionado de acordo com o critério de maior semelhança com a posição da ribavirina 5'-monofosfato na estrutura cristalográfica.

Para este grupo, houve uma clara tendência favorecendo o padrão de substituição 1,5, que apresentou valores bem mais favoráveis de entalpia de interação em relação aos demais padrões de substituição.

Para o padrão de substituição 1,5, as melhores entalpias de interação seguiram a ordem dos substituintes R4, R3, R1 e R2, quando os cálculos foram feitos com constante dielétrica igual a 4 (interior de proteínas); para constante dielétrica igual a 78,4 (meio aquoso), este também foi o grupo com valores mais favoráveis de entalpia de interação (embora menos favoráveis) e a ordem se alterou para R4, R3, R2 e R1.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BARBOSA, J. C., SANTOS, M. A. Modelagem Matemática, Perspectivas e Discussões. 2003.
- BASTOS, E. A. Otimização de Seções Retangulares de Concreto Armado Submetidas à FlexoCompressão Oblíqua Utilizando Algoritmos Genéticos. Tese de Mestrado, UFRJ, Rio de Janeiro, 2014.
- BENARROCH, D., EGLOFF, M.-P., MULARD, L., GUERREIRO, C., ROMETTE, J.-L., CANARD, B. J. Biol. Chem. 279, 35638-35643. 2004.
- CARTER, J. B; SAUNDERS, V. Virology – Principles and Applications. ISBN: 978-0-470-02386-0. 2007.
- CARVALHO, I., PUPO, M. T., BORGES, A. D. L., BERNARDES, L. S. C. Quim. Nova, Vol. 26, No. 3, 428-438, 2003.
- DEL CISTIA, C. N. Modelagem molecular aplicada ao estudo de reações de inibição enzimática com aplicação potencial no controle de Leishmania amazonensis. Tese de Doutorado. UFRRJ. 2010.
- EGLOFF, M. P., BENARROCH, D., SELISKO, B., ROMETTE, J. L., CANARD, B. EMBO J. 21, 2757–2768. 2002.
- GRANDI, TARCIANA. Estudos de Polimorfismos Genéticos e resposta ao Tratamento com Interferon-Alfa/Ribavirina em Pacientes com o Vírus da Hepatite C. p. 19-22, 2012.
- GUEX, N. AND PEITSCH, M.C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18, 2714-2723. 1997.
- HHMI. Disponível em: <<http://www.hhmi.org/biointeractive/dengue-virus-life-cycle>>. 2016.
- KAISER, Gabriele, SRIRAMAN, Bharath. *A global survey of international perspectives on modelling in mathematics education*. ZDM Vol. 38(3). 2006.
- LOPES, SUSANA I. O. Caracterização clínica e virológica do efeito do interferon peguilado e ribavirina no retratamento da hepatite C crônica em doentes récidivantes e não respondedores a tratamentos prévios de combinação. Porto, Portugal. p. 36-38, 2004.
- MARQUES, M. L., BOTTI, S. O que é e para que serve a Teoria dos Funcionais de Densidade? Disponível em: <http://nautilus.fis.uc.pt/gazeta/revistas/29_4/vol29_4_Art02.pdf>. 2007.
- MICHAELIS. Disponível em: <<http://michaelis.uol.com.br/>>. 2016.

- PORTAL DA SAÚDE. Disponível em <<http://portalsaude.saude.gov.br/index.php/o-ministerio/principal/secretarias/svs/dengue>>. 2016.
- MOPAC2016, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA. 2016.
- NELSON, D.L. & COX, M.M. Lehninger's Principles of Biochemistry. 4^aed. New York: W. ISBN-10: 0716743396. ISBN-13: 978-0716743392. 2004.
- OLIVEIRA, André S., SILVA, Milene L. da, OLIVEIRA, Ana Flávia C. S. SILVA, Cynthia C. da, TEIXEIRA, Róbson R., DE PAULA, Sérgio O. NS3 and NS5 Proteins: Important Targets for Anti-Dengue Drug Design. J. Braz. Chem. Soc., Vol. 25, No. 10, 1759-1769, 2014.
- OLIVIERI, B. P. Otimização do Projeto de Pontes Protendidas Pré-Moldadas pelo Método dos Algoritmos Genéticos. Tese de Mestrado, UFRJ, Rio de Janeiro, 2004.
- PASCOAL, F. S., Sociedade Artificial Fight4life: Autômato Celular Modelando Vida Artificial. P. 11-37. 2005.
- PROSISE, G. L., WU, J., LUECKE, H. PDB ID: 1ME8. Journal: J.Biol.Chem. 277: 50654-50659. 2002. PDB ID: 1ME8.
- RITA, Mariane R., Otimização da Fase Construtiva de Estruturas de Concreto Massa em Ambiente Paralelo. 2015.
- ROSA, T. O., LUZ, H. S. Conceitos Básicos de Algoritmos Genéticos: Teoria e Prática. In: XI Encontro de Estudantes de Informática do Tocantins, 2009, Palmas. Anais do XI Encontro de Estudantes de Informática do Tocantins. Palmas: Centro Universitário Luterano de Palmas, 2009. p. 27-37.
- SANT'ANNA, C. M. R., Glossário De Termos Usados No Planejamento De Fármacos (Recomendações Da Iupac Para 1997). Quim. Nova, Vol. 25, No. 3, 505-512, 2002.
- SANT'ANNA, C. M. R., Métodos de modelagem molecular para estudo e planejamento de compostos bioativos: Uma introdução. Rev. Virtual Quim., 2009, 1 (1), 49-57. Data de publicação na Web: 2 de Fevereiro de 2009.
- SCILAB ENTERPRISES. Disponível em: <<http://www.scilab.org/>>. 2015.
- SHATKIN, A. J. Capping of eucaryotic mRNAs. Cell, v. 9, n. 4, p. 645-653, 1976.
- STEWART, James J. P., *Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. Journal of Molecular Modeling*. Volume 13, Issue 12, pp 1173-1213. 2007.

- STEWART, James J. P., *Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. Journal of Molecular Modeling*. Volume 19, Issue 1, pp 01-32. 2013.
- THE MATHWORKS, INC. Disponível em: <<https://www.mathworks.com/?requestedDomain=www.mathworks.com>>. 1994-2016.
- TORTORA, G. J FUNKE, B. R; CASE, C. L. *Microbiologia*. 10ª Ed. p. 367-395, 2010. University, USA, 1995.
- WAVEFUNCTION, INC. Disponível em < <https://www.wavefun.com/>>. 2016.
- WORLD HEALTH ORGANIZATION – WHO (2012). Dengue e Dengue hemorrágico. Nota descritiva, n. 117. Disponível em: <www.who.int/mediacentre/factsheets/fs117/es/>. 2016.

ANEXOS

Anexo A Valores de energia dos confôrmeros obtidos através do método de busca de Monte Carlo	52
Anexo B Demonstração da matriz que descreve o ligante substituição 1,4 com substituinte R1. Em amarelo, encontra-se destacada a linha de comando do programa MOPAC2016.....	54
Anexo C Demonstração da matriz que descreve uma parte do complexo substituição 1,4 com substituinte R1. Em amarelo, encontra-se destacada a linha de comando do programa MOPAC2016. As colunas encontram-se destacadas pela presença de chaves, que demonstram seu significado perante a matriz. As referentes às “ <i>flags</i> ”, se modificadas para +0, fazem com que os átomos fiquem fixos; se permanecem com o valor +1, possuem posicionamento livre, de acordo com a minimização do sistema	56
Anexo D Tabelas com valores da entalpia de interação dos complexos, sítios e ligantes para a série pontuação para $\epsilon = 4,0$ (Tabela D.1), série pontuação para $\epsilon = 78,4$ (Tabela D.2), série posição para $\epsilon = 4,0$ (Tabela D.3) e série posição para $\epsilon = 4,0$ (Tabela D.4)	57

Anexo A Valores de energia dos confôrmeros obtidos através do método de busca de Monte Carlo.

Nome	E (kcal/mol)
Molécula 1	-113.492064
Molécula 2	-109.679436
Molécula 3	-109.641813
Molécula 4	-109.498410
Molécula 5	-109.308892
Molécula 6	-108.969732
Molécula 7	-108.954862
Molécula 8	-108.953636
Molécula 9	-108.925965
Molécula 10	-108.923237
Molécula 11	-108.910918
Molécula 12	-113.488089
Molécula 13	-108.834205
Molécula 14	-108.417510
Molécula 15	-108.416327
Molécula 16	-108.257558
Molécula 17	-108.238577
Molécula 18	-108.139720
Molécula 19	-108.041235
Molécula 20	-107.949457
Molécula 21	-107.948037
Molécula 22	-107.947786
Molécula 23	-113.485977
Molécula 24	-107.946363
Molécula 25	-107.944426
Molécula 26	-107.943676
Molécula 27	-107.929928
Molécula 28	-107.919521
Molécula 29	-107.840710
Molécula 30	-107.837231
Molécula 31	-107.822902
Molécula 32	-107.797724
Molécula 33	-107.333622
Molécula 34	-113.485302
Molécula 35	-107.308083
Molécula 36	-107.287702
Molécula 37	-106.964152
Molécula 38	-106.304840
Molécula 39	-106.294198
Molécula 40	-106.291459
Molécula 41	-105.823705
Molécula 42	-105.773477
Molécula 43	-105.773021
Molécula 44	-105.671279
Molécula 45	-113.482881

Molécula 46	-105.588464
Molécula 47	-105.587985
Molécula 48	-105.587574
Molécula 49	-105.047055
Molécula 50	-104.778614
Molécula 51	-104.559325
Molécula 52	-104.528387
Molécula 53	-104.290266
Molécula 54	-104.289275
Molécula 55	-104.281361
Molécula 56	-113.475427
Molécula 57	-103.911314
Molécula 58	-109.696084
Molécula 59	-109.692726
Molécula 60	-109.692710

Anexo B Demonstração da matriz que descreve o ligante substituição 1,4 com substituinte R1. Em destaque, encontra-se sublinhada a linha de comando do programa MOPAC2016.

MOZYME PL MMOK EPS=4.0 PDBOUT

C(HETATM 1 C UNK 1) -1.49943777 +1 -0.96018551 +1 -2.41789172 +1
C(HETATM 2 C UNK 1) -0.06249653 +1 -1.36745308 +1 -2.78192938 +1
C(HETATM 3 C UNK 1) 0.76401259 +1 -0.42223252 +1 -1.87302548 +1
C(HETATM 4 C UNK 1) -0.14209941 +1 0.84877509 +1 -1.84005420 +1
H(HETATM 5 H UNK 1) 0.15931287 +1 -2.44939276 +1 -2.63775004 +1
H(HETATM 6 H UNK 1) 0.94374436 +1 -0.86408500 +1 -0.87012698 +1
H(HETATM 7 H UNK 1) 0.18242547 +1 1.57319924 +1 -2.63612368 +1
O(HETATM 8 O UNK 1) 0.25741106 +1 -0.96914782 +1 -4.09390651 +1
O(HETATM 9 O UNK 1) 2.02757684 +1 -0.07037138 +1 -2.33992452 +1
O(HETATM 10 O UNK 1) -1.46334053 +1 0.46141865 +1 -2.15058318 +1
C(HETATM 11 C UNK 1) -0.14359861 +1 1.51590753 +1 -0.50836643 +1
C(HETATM 12 C UNK 1) 0.78327167 +1 1.42250813 +1 0.54632083 +1
N(HETATM 13 N UNK 1) 0.34615301 +1 2.33003471 +1 1.49301853 +1
N(HETATM 14 N UNK 1) -0.78660840 +1 2.94610685 +1 1.09933926 +1
N(HETATM 15 N UNK 1) -1.08048887 +1 2.45508449 +1 -0.10423687 +1
H(HETATM 16 H UNK 1) 1.64827057 +1 0.78317573 +1 0.62506498 +1
C(HETATM 17 C UNK 1) 0.91000915 +1 2.68038751 +1 2.79970358 +1
N(HETATM 18 N UNK 1) 2.13772844 +1 1.85447045 +1 3.08403910 +1
H(HETATM 19 H UNK 1) 0.13119341 +1 2.50773680 +1 3.61193843 +1
H(HETATM 20 H UNK 1) 1.17594019 +1 3.78406127 +1 2.82382304 +1
H(HETATM 21 H UNK 1) 1.93920632 +1 0.82853037 +1 3.06762903 +1
H(HETATM 22 H UNK 1) -0.15843816 +1 -1.56683972 +1 -4.75016425 +1
H(HETATM 23 H UNK 1) 2.00759528 +1 0.07437639 +1 -3.32936860 +1
H(HETATM 24 H UNK 1) -1.93675023 +1 2.76702919 +1 -0.59020004 +1
H(HETATM 25 H UNK 1) -2.21823767 +1 -1.05888133 +1 -3.24819590 +1
C(HETATM 26 C UNK 1) -1.97038037 +1 -1.67065826 +1 -1.12561014 +1
H(HETATM 27 H UNK 1) -3.02330119 +1 -1.41498524 +1 -0.90305918 +1
H(HETATM 28 H UNK 1) -1.35027683 +1 -1.35190969 +1 -0.25664506 +1
O(HETATM 29 O UNK 1) -1.76987839 +1 -3.02695314 +1 -1.46583722 +1

P(HETATM 30 P UNK 1) -2.18451289 +1 -4.33318777 +1 -0.48314418 +1
O(HETATM 31 O UNK 1) -1.85851603 +1 -5.39649439 +1 -1.51629259 +1
H(HETATM 32 H UNK 1) 2.90924266 +1 2.05510036 +1 2.41887073 +1
O(HETATM 33 O UNK 1) -1.25188759 +1 -4.27708320 +1 0.70917767 +1
O(HETATM 34 O UNK 1) -3.65173259 +1 -4.18037627 +1 -0.13971745 +1
H(HETATM 35 H UNK 1) 2.49487304 +1 2.07823441 +1 4.04493891 +1

Anexo C Demonstração da matriz que descreve uma parte da estrutura substituição 1,4 com substituinte R1. Em amarelo, encontra-se destacada a linha de comando do programa MOPAC2016. As colunas encontram-se destacadas pela presença de chaves, que demonstram seu significado perante a matriz. As referentes às “flags”, se modificadas para +0, fazem com que os átomos fiquem fixos; se permanecem com o valor +1, possuem posicionamento livre, de acordo com a minimização do sistema.

MOZYME PL CHARGE=+2 EPS=4.0 PDBOUT

```
N(ATOM  1 N  GLY A  4) 3.81800000 +0 -44.79400000 +0 23.66900000 +0
C(ATOM  2 CA GLY A  4) 4.63100000 +0 -45.47500000 +0 22.67400000 +0
C(ATOM  3 C  GLY A  4) 6.06400000 +0 -45.50300000 +0 23.16100000 +0
O(ATOM  4 O  GLY A  4) 6.90000000 +0 -46.27100000 +0 22.68000000 +0
H(ATOM  5 HA1 GLY A  4) 4.56504947 +0 -44.96372551 +0 21.67989344 +0
H(ATOM  6 HA2 GLY A  4) 4.25415798 +0 -46.51006640 +0 22.49725748 +0
H(ATOM  7 H  GLY A  4) 3.89879590 +0 -43.78906395 +0 23.68038408 +0
H(ATOM  8 H  GLY A  4) 3.83039349 +0 -45.20217074 +0 24.58194776 +0
N(ATOM  9 N  GLU A  5) 6.34200000 +0 -44.65400000 +0 24.14000000 +0
C(ATOM 10 CA  GLU A  5) 7.66500000 +0 -44.58000000 +0 24.72100000 +0
C(ATOM 11 C  GLU A  5) 7.94000000 +0 -45.80700000 +0 25.57800000 +0
O(ATOM 12 O  GLU A  5) 9.09600000 +0 -46.23300000 +0 25.71500000 +0
C(ATOM 13 CB  GLU A  5) 7.77200000 +1 -43.32800000 +1 25.55200000 +1
C(ATOM 14 CG  GLU A  5) 7.70700000 +1 -42.07100000 +1 24.73900000 +1
C(ATOM 15 CD  GLU A  5) 7.80000000 +1 -40.86700000 +1 25.62500000 +1
O(ATOM 16 OE1 GLU A  5) 6.76600000 +1 -40.52700000 +1 26.23700000 +1
O(ATOM 17 OE2 GLU A  5) 8.90300000 +1 -40.28800000 +1 25.73300000 +1
H(ATOM 18 HA  GLU A  5) 8.44628151 +0 -44.57782479 +0 23.89319128 +0
H(ATOM 19 HB1 GLU A  5) 8.73392025 +0 -43.34803934 +0 26.12915390 +0
H(ATOM 20 HB2 GLU A  5) 6.99301071 +0 -43.30716977 +0 26.35254485 +0
H(ATOM 21 H  GLU A  5) 5.68260724 +0 -43.95165091 +0 24.49764040 +0
H(ATOM 22 HG1 GLU A  5) 6.77008970 +0 -42.00512813 +0 24.14697094 +0
H(ATOM 23 HG2 GLU A  5) 8.53233234 +0 -42.03945886 +0 23.99172959 +0
N(ATOM 24 N  LYS A  6) 6.87500000 +0 -46.36400000 +0 26.15600000 +0
C(ATOM 25 CA  LYS A  6) 6.98500000 +0 -47.57600000 +0 26.96900000 +0
```

Anexo D Tabelas com valores da entalpia de interação dos complexos, sítios e ligantes para a série pontuação para $\epsilon = 4,0$ (Tabela D.1), série pontuação para $\epsilon = 78,4$ (Tabela D.2), série posição para $\epsilon = 4,0$ (Tabela D.3) e série posição para $\epsilon = 4,0$ (Tabela D.4)

Tabela C.1 Valores da entalpia de interação dos complexos, sítios e ligantes para a série pontuação para $\epsilon = 4,0$.

Molécula	Complexo	Sítio	Ligante	Interação
14 com R1	-1660,68793	-1375,76014	-229,14629	-55,7815
14 com R2	-1826,69669	-1375,76014	-361,55687	-89,37968
14 com R3	-1667,55526	-1375,76014	-264,50855	-27,28657
14 com R4	-1666,58185	-1375,76014	-251,18074	-39,64097
15 com R1	-1625,74315	-1375,76014	-244,38919	-5,59382
15 com R2	-1778,58264	-1375,76014	-399,36353	-3,45897
15 com R3	-1667,17496	-1375,76014	-266,7025	-24,71232
15 com R4	-1688,33078	-1375,76014	-267,56407	-45,00657
24 com R1	-1633,23455	-1375,76014	-216,40060	-41,07381
24 com R2	-1835,11999	-1375,76014	-383,93228	-75,42757
24 com R3	-1668,89795	-1375,76014	-253,41221	-39,7256
24 com R4	-1636,04986	-1375,76014	-257,28336	-3,00636
45 com R1	-1635,67822	-1375,76014	-254,67202	-5,24606
45 com R2	-1816,96053	-1375,76014	-459,61453	18,41414
45 com R3	-1674,34855	-1375,76014	-270,305	-28,28341
45 com R4	-1752,28323	-1375,76014	-289,19209	-87,331

Tabela C.2 Valores da entalpia de interação dos complexos, sítios e ligantes para a série pontuação para $\epsilon = 78,4$.

Molécula	Complexo	Sítio	Ligante	Interação
14 com R1	-1866,47508	-1530,62673	-317,51795	-18,3304
14 com R2	-2020,00522	-1530,62673	-469,54143	-19,83706
14 com R3	-1882,48391	-1530,62673	-336,92395	-14,93323
14 com R4	-1872,33113	-1530,62673	-337,88672	-3,81768
15 com R1	-1826,06947	-1530,62673	-324,43320	28,99046
15 com R2	-2003,37661	-1530,62673	-469,13458	-3,6153
15 com R3	-1889,48855	-1530,62673	-343,39116	-15,47066
15 com R4	-1911,04401	-1530,62673	-338,62556	-41,79172
24 com R1	-1828,17544	-1530,62673	-310,51315	12,96444

24 com R2	-1992,74567	-1530,62673	-452,70876	-9,41018
24 com R3	-1872,92656	-1530,62673	-334,11993	-8,1799
24 com R4	-1848,02149	-1530,62673	-332,75886	15,3641
45 com R1	-1848,35478	-1530,62673	-337,73137	20,00332
45 com R2	-2016,67714	-1530,62673	-467,45295	-18,59746
45 com R3	-1879,77218	-1530,62673	-362,75346	13,60801
45 com R4	-1915,73389	-1530,62673	-358,60914	-26,49802

Tabela C.3 Valores da entalpia de interação dos complexos, sítios e ligantes para a série posição para $\epsilon = 4,0$.

Molécula	Complexo	Sítio	Ligante	Interação
14 com R1	-2144,5605	-1845,85491	-229,14629	-69,5593
14 com R2	-2316,939	-1845,85491	-361,55687	-109,52722
14 com R3	2151,06859	-1845,85491	-264,50855	-40,70513
14 com R4	-2127,67875	-1845,85491	-251,18074	-30,6431
15 com R1	-2253,13537	-1845,85491	-215,83983	-191,44063
15 com R2	-2422,72459	-1845,85491	-399,36353	-177,50615
15 com R3	-2323,74513	-1845,85491	-266,7025	-211,18772
15 com R4	-2335,30669	-1845,85491	-267,56407	-221,88771
24 com R1	-2090,31937	-1845,85491	-216,40060	-28,06386
24 com R2	-2299,93662	-1845,85491	-383,93228	-70,14943
24 com R3	-2124,45219	-1845,85491	-253,41221	-25,18507
24 com R4	-2108,15123	-1845,85491	-257,28336	-5,01296
45 com R1	-2097,67525	-1845,85491	-254,67202	2,85168
45 com R2	-2322,87686	-1845,85491	-459,61453	-17,40742
45 com R3	-2128,59003	-1845,85491	-270,305	-12,43012
45 com R4	-2208,12536	-1845,85491	-289,19209	-73,07836

Tabela C.4 Valores da entalpia de interação dos complexos, sítios e ligantes para a série posição para $\epsilon = 78,4$.

Molécula	Complexo	Sítio	Ligante	Interação
14 com R1	-2208,12536	-2001,58939	-317,51795	110,98198
14 com R2	-2483,96233	-2001,58939	-469,54143	-12,83151
14 com R3	-2339,22471	-2001,58939	-336,92395	-0,71137
14 com R4	-2333,104	-2001,58939	-337,88672	6,37211

15 com R1	-2419,43326	-2001,58939	-324,43320	-93,41067
15 com R2	-2583,89494	-2001,58939	-469,13458	-113,17097
15 com R3	-2490,52731	-2001,58939	-343,39116	-145,54676
15 com R4	-2498,90358	-2001,58939	-338,62556	-158,68863
24 com R1	-2295,10184	-2001,58939	-310,51315	17,0007
24 com R2	-2457,10231	-2001,58939	-452,70876	-2,80416
24 com R3	-2329,59469	-2001,58939	-334,11993	6,11463
24 com R4	-2316,79995	-2001,58939	-332,75886	17,5483
45 com R1	-2300,51997	-2001,58939	-337,73137	38,80079
45 com R2	-2490,40315	-2001,58939	-467,45295	-21,36081
45 com R3	-2325,05944	-2001,58939	-362,75346	39,28341
45 com R4	-2367,1601	-2001,58939	-358,60914	-6,96157
