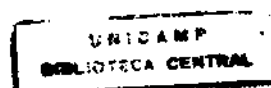


ANDERSON COSER GAUDIO

**RELAÇÕES ENTRE ESTRUTURA QUÍMICA E ATIVIDADE
BIOLÓGICA DE INIBIDORES DA TIMIDINA-CINASE DO
VÍRUS HERPES SIMPLES**



UNIDADE	IQ
N.º CHAMADA:	
G	233R
V.	EL
TOMBO BC	35.242
PROC.	3.95/95
C	☐
D	☒
PREÇO	R\$ 11,00
DATA	30/06/98
N.º CPD	40002446-9

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE QUÍMICA
UNICAMP**

G233R Gaudio, Anderson Coser
Relações entre estrutura química e atividade biológica
de inibidores da timidina-cinase do vírus herpes simples/
Anderson Coser Gaudio.-- Campinas, [SP:s.n.], 1998.

Orientador: Yuji Takahata.
Co-orientador: Antônia Tavares do Amaral.

1. *Modelagem molecular. 2. *Modo de ligação.
3. Análise conformacional. I. Takahata, Yuji. II. Amaral,
Antonia Tavares do. III. Universidade Estadual de Campinas.
Instituto de Química. IV. Título.

Universidade Estadual de Campinas

Instituto de Química

**RELAÇÕES ENTRE ESTRUTURA QUÍMICA E ATIVIDADE
BIOLÓGICA DE INIBIDORES DA TIMIDINA-CINASE DO
VÍRUS HERPES SIMPLES**

Tese de Doutorado

ANDERSON COSER GAUDIO

Orientador: Prof. Dr. Yuji Takahata

Co-orientadora: Profa. Dra. Antônia Tavares do Amaral

CAMPINAS

1998

ANDERSON COSER GAUDIO

**RELAÇÕES ENTRE ESTRUTURA QUÍMICA E ATIVIDADE
BIOLÓGICA DE INIBIDORES DA TIMIDINA-CINASE DO
VÍRUS HERPES SIMPLES**

BANCA EXAMINADORA

TESE PARA OBTENÇÃO DE GRAU DE DOUTOR

Presidente :	Prof. Dr. Yuji Takahata
2º Examinador:	Profa. Dra. Fulvia Maria Luisa Gravina Stamato (DQ-UFSCar)
3º Examinador:	Prof. Dr. Nelson Henrique Morgon (IQ-Unicamp)
4º Examinador:	Prof. Dr. Lauro Euclides Soares Barata (IQ-Unicamp)
5º Examinador:	Profa. Dra. Márcia Miguel Castro Ferreira (IQ-Unicamp)

Campinas (SP), 08 de Junho de 1998

GAUDIO, A.C. Relações entre Estrutura Química e Atividade Biológica de Inibidores da Timidina-Cinase do Vírus Herpes Simples. Campinas, 1998. pp. 286 *Tese* (Doutorado em Química) – Instituto de Química, Universidade Estadual de Campinas.

RESUMO

Investigou-se a atividade inibitória dos derivados da N^2 -fenilguanina (PHG), substituídos nas posições meta e para do anel fenila, sobre a enzima timidina-cinase do vírus herpes simples dos tipos 1 e 2 (HSV1 TK e HSV2 TK, respectivamente). A investigação foi precedida por análise conformacional das PHG, que revelou a existência de quatro conformações estáveis, C1, C2, C3 e C4, que são definidas pelos valores dos ângulos torsionais entre os anéis guanina e fenila. Utilizaram-se a estrutura tridimensional de HSV1 TK em complexo com 2'-deoxitimidina e ganciclovir, determinada experimentalmente por cristalografia de raios-X, e o modelo tridimensional de HSV2 TK, construído a partir de HSV1 TK, para determinar o modo de ligação das PHG no sítio ativo de HSV TK. O estudo indicou ser C3 a conformação bioativa. O estudo das relações entre estrutura química e atividade biológica indicou a influência de efeito estérico sobre a atividade inibitória dos derivados meta das PHG sobre HSV2 TK. A atividade dos derivados meta sobre HSV1 TK não é influenciada por fatores estéricos. O modo de ligação proposto indicou que o resíduo de aminoácido Leu 97 de HSV2 TK é o responsável pela queda de atividade dos derivados meta-substituídos sobre HSV2 TK. A presença do resíduo Ile 97 em HSV1 TK faz com que o efeito estérico não seja observado.

Palavras-chave: modelagem molecular; modo de ligação; análise conformacional; QSAR.

GAUDIO, A.C. Structure-activity relationships of inhibitors of herpes simplex virus thymidine kinase. Campinas, 1998. pp. 286 *Tese* (Doutorado em Química) – Instituto de Química, Universidade Estadual de Campinas.

ABSTRACT

The inhibitory activity of the N²-phenylguanine (PHG) derivatives, substituted on meta and para positions of phenyl ring, against herpes simplex virus of type 1 and 2 thymidine kinases (HSV1 TK and HSV2 TK, respectively) has been investigated. The investigation was preceded by a conformational analysis of PHG derivatives, that revealed the existence of four stable conformations, C1, C2, C3 and C4, defined by specific torsion angle values between the guanine and phenyl rings. The three-dimensional structure of HSV1 TK, in complex with 2'-deoxythymidine and ganciclovir, determined by X-ray crystallography, and a three-dimensional model of HSV2 TK, built from HSV1 TK as template, have been used to determine the binding mode of the PHG derivatives in the HSV1 TK active site. The results indicated C3 as the bioactive conformation. The quantitative structure-activity relationship study indicated that the inhibitory activity of meta derivatives on HSV2 TK is influenced by a steric effect. The activity of the meta PHG derivatives on HSV1 TK seems not to be influenced by such effect. The proposed binding mode indicated that the aminoacid residue Leu 97 of HSV2 TK is the cause of the decrease in the inhibitory activity of the meta derivatives on HSV2 TK. The presence of Ile 97 in the same position of HSV1 TK makes the steric effect not to be observable.

Keywords: molecular modelling; binding mode; conformational analysis; QSAR.

DADOS CURRICULARES

ANDERSON COSER GAUDIO

NASCIMENTO: 29.11.1959 – Vila Velha, ES

FILIAÇÃO: José Gaudio Sobrinho e Maria Yolanda Coser Gaudio

1980/1982 Graduação em Farmácia, na Faculdade de Farmácia e Bioquímica do Espírito Santo – Vitória, ES

1981/1983 Professor em nível de 2º Grau, no Colégio Americano Batista – Vitória, ES

1983 Especialização em Bioquímica, na Faculdade de Farmácia e Bioquímica do Espírito Santo – Vitória, ES

1984/1985 Professor em nível de 2º Grau na Associação de Ensino e Cultura "Dom Bosco" – Vila Velha, ES

1985/1989 Professor em nível de 2º, Grau na Escola de 1º e 2º Graus "Vasco Coutinho" – Vila Velha, ES

1986/1988 Professor Auxiliar de Ensino, no Departamento de Físico-Química, Faculdade de Farmácia e Bioquímica do Espírito Santo – Vitória, ES

1987/1990 Professor em nível de 2º Grau, no Colégio Sagrado Coração de Maria – Vitória, ES

1988/1990 Professor em nível de 2º Grau, no Centro Educacional Objetivo – Vitória, ES

1990/1992 Mestrado em Físico-Química no Instituto de Química da Unicamp – Campinas, SP

1993/Professor Assistente 3, no Departamento de Física, Centro de Ciências Exatas, Universidade Federal do Espírito Santo – Vitória, ES

TRABALHOS PUBLICADOS

GAUDIO, A.C.; TAKAHATA, Y. Calculation of molecular surface area with numerical factors. *Comput. Chem.*, v.16, p.277-284, 1992.

GAUDIO, A.C.; KOROLKOVAS, A.; TAKAHATA, Y. Conformational analysis of the 1,4-dihydropyridines linking the structural aspects to the biological binding event: a study of the receptor-site conformation. *J. Mol. Struct. (Theochem)*, v.303, p.255-263, 1994.

GAUDIO, A.C.; KOROLKOVAS, A.; TAKAHATA, Y. Quantitative structure-activity relationships for 1,4-dihydropyridine calcium channel antagonists (nifedipine analogues): a quantum chemical/classical approach. *J. Pharm. Sci.*, v.83, p.1110-1115, 1994.

GAUDIO, A.C.; FERRACIOLI, L. The "Folk Balloon". *The Physics Teacher*, v.33, p.582-583, 1995.

GAUDIO, A.C. Modelos de estudo quantitativo das relações entre estrutura química e atividade biológica. *Química Nova*, v.19, p.278-289, 1996.

PIRES, J.M.; FLORIANO, W.B.; GAUDIO, A.C. Extension of the frontier reactivity indices to groups of atoms and application to quantitative structure-activity relationship studies. *J. Mol. Struct. (Theochem)*, v.389, p.159-167, 1997.

COSTA, M.C.A.; GAUDIO, A.C.; TAKAHATA, Y. A comparative study of principal component and linear multiple regression analysis in SAR and QSAR applied to 1,4-dihydropyridine calcium channel antagonists (nifedipine analogues). *J. Mol. Struct. (Theochem)*, v.394, p.291-300, 1997.

GAUDIO, A.C.; TAKAHATA, Y.; RICHARDS, W.G. Prediction of the binding mode of N²-phenylguanine derivative inhibitors to herpes simplex virus type 1 thymidine kinase. *J. Comp.-Aid. Mol. Des.*, v.12, p.15-25, 1998.

Dedico este trabalho à mulher que eu amo. Os anos que vivi longe de ti, Claudia, em virtude deste trabalho, são irrecuperáveis. Nosso amor, no entanto, foi sólido bastante para suportar o longo tempo de provação.

AGRADECIMENTOS

Sou eternamente grato ao Prof. Yuji Takahata pelo árduo e paciente serviço de orientação deste trabalho.

Agradeço à Profa. Dra. Antônia Tavares do Amaral por ter-me acolhido em seu laboratório e pela grande ajuda prestada, principalmente no início deste trabalho.

Agradeço ao Dr. William Graham Richards, New Chemistry Laboratory – University of Oxford, pela forma calorosa com que me acolheu em Oxford e pela orientação prestada na parte de modelagem molecular. Também agradeço ao Dr. Peter Goodford, Institute of Biophysics – University of Oxford, pelas discussões sobre modelagem de enzimas em meio aquoso.

Agradeço ao Dr. Mark R. Sanderson e ao Dr. John Champness, Kings College – Londres, pela gentil cessão da estrutura tridimensional da enzima timidina-cinase do vírus herpes simples do tipo 1.

Sou grato ao Dr. Corwin Hansch, Pomona College – Claremont, Califórnia - pela ajuda prestada na definição dos valores de algumas constantes de substituintes e ao Dr. Hugo Kubinyi, Basf - Ludwigshafen, Alemanha - pelo esclarecimento de diversas dúvidas sobre os aspectos teóricos da análise de QSAR.

Sou grato aos professores Márcia Miguel Castro Ferreira, Nelson Henrique Morgon, Pedro Luiz Onofre Volpe, Rogério Custódio e Watson Loh do Instituto de Química - Unicamp e ao professor Marcelo Giordan Santos do Instituto de Educação – USP.

Agradeço aos amigos e colegas do Instituto de Química – Unicamp Aline Thais Bruni, Anselmo Elcana de Oliveira, Edgar Alberto de Brito, Edilson Grünheidt Borges, Ednalva Dantas da Silva Duarte, Hermes Fernandes de Souza, Júlio do Nascimento Trevas,

Maria Cristina Andreazza Costa, Maria Eugênia Porto Garcia, Maria Inês de Camargo Harris, Paulo Henrique Guadagnini, Rosana Vendrame e Sandra Gomes de Moraes.

Agradeço aos amigos e colegas do Physical and Theoretical Chemistry Laboratory – University of Oxford Ana Castro, Annabel Todd, Barry Hardy, Charlotte Dean, Daniel Robinson, Martin Parretti, Massoud Mahmoudian, Péter Várnai, Peter Winn, Romano Kroemer, Steve Doughty and Thomas Barlow.

Agradeço a todos os funcionários do Instituto de Química da Unicamp, que direta ou indiretamente contribuíram para a realização deste trabalho. Em particular agradeço aos funcionários André Luiz Camargo, Daniel Machado, Denise Capocci, Eliane Aparecida Araújo Pimentel, Iara de Assis Dezorde, Izabel Aquino Calasso, Iveraldo Rodrigues, Maria Paula Nogueira de Carvalho, Nelson Luiz Saldanha e Neusa Maria Lopes Lima.

Também agradeço ao laboratório de computação do Centro de Ciências Exatas – Universidade Federal do Espírito Santo, ao laboratório de computação do Physical and Theoretical Chemistry Laboratory – University of Oxford e ao Centro Nacional de Computação de Alto Desempenho - CENAPAD - Unicamp pela utilização de seus ambientes computacionais.

Agradeço ao Departamento de Física do Centro de Ciências Exatas da Universidade Federal do Espírito Santo por ter-me concedido o afastamento de minhas atribuições didáticas, que possibilitou o término deste trabalho.

Também agradeço à secretária do Departamento de Física do Centro de Ciências Exatas da Universidade Federal do Espírito Santo, Ana Maria Caldeira, pela imensa ajuda prestada durante o tempo que passei afastado do departamento.

Finalmente agradeço à Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior – CAPES e ao Plano Institucional de Capacitação Docente e Técnico – PICDT da Universidade Federal do Espírito Santo pelo suporte financeiro ao presente projeto.

ÍNDICE ANALÍTICO

RESUMO	4
ABSTRACT	5
DADOS CURRICULARES	6
TRABALHOS PUBLICADOS	7
AGRADECIMENTOS	9
ÍNDICE ANALÍTICO	11
ÍNDICE DE FIGURAS	15
ÍNDICE DE TABELAS	21
ÍNDICE DE QUADROS	25
ÍNDICE DE DIAGRAMAS	25
ABREVIACÕES UTILIZADAS	27
CAPÍTULO 1 INTRODUÇÃO	33
1.1 HERPESVÍRUS	34
1.1.1 Classificação	35
1.1.2 Morfologia	36
1.1.3 Estrutura do genoma	37
1.1.4 Polipeptídios	40
1.1.5 Propriedades biológicas	40
1.2 VÍRUS HERPES SIMPLES	41
1.2.1 Propriedades biológicas	43
1.2.2 Patogenia	46
1.2.3 Combate ao vírus herpes simples	47
1.2.4 Derivados da N^2 -fenil-(substituído)-guanina	54
CAPÍTULO 2 OBJETIVOS E MÉTODOS	57
2.1 OBJETIVOS	57
2.2 SELEÇÃO DOS COMPOSTOS	58
2.3 ANÁLISE CONFORMACIONAL	59
2.3.1 Fundamentos do método AM1	60
2.3.2 Critérios de convergência do cálculo SCF em MOPAC	63
2.3.3 Critérios de otimização de geometria molecular	64
2.3.4 Cálculo do calor de formação	64
2.4 CÁLCULO DAS PROPRIEDADES FÍSICO-QUÍMICAS	65
2.4.1 Energias do HOMO (ϵ_{HOMO}) e LUMO (ϵ_{LUMO})	66
2.4.2 Dureza molecular (η)	66

2.4.3 Volume de van der Waals (<i>VvdW</i>)	67
2.4.4 Cargas atômicas eletrostáticas (<i>Qn</i>)	67
2.4.5 Momento dipolar (μ)	70
2.4.6 Polarizabilidade molecular (α)	71
2.4.7 Índices de fronteira	73
2.5 RELAÇÕES ENTRE ESTRUTURA QUÍMICA E ATIVIDADE BIOLÓGICA	74
2.5.1 Regressão linear múltipla	75
2.5.2 Seleção de variáveis por busca sistemática	80
2.6 ALINHAMENTO DAS SEQUÊNCIAS DE AMINOÁCIDOS	82
2.7 MODELAGEM MOLECULAR	83
2.7.1 Campo de força CHARMM	84
2.7.2 Simulação da dinâmica molecular	85
2.7.3 Cargas atômicas	86
CAPÍTULO 3 ANÁLISE CONFORMACIONAL	89
3.1 INTRODUÇÃO	89
3.2 TAUTOMERIA NA GUANINA	89
3.2.1 Tautomeria	89
3.2.2 Tautomeria nos processos bioquímicos	91
3.2.3 Tautomeria na guanina	91
3.2.4 Forma tautomérica preferencial da guanina	96
3.3 GEOMETRIA EXPERIMENTAL DA GUANINA	97
3.3.1 Comprimentos de ligação	97
3.3.2 Ângulos de ligação	99
3.4 PADRONIZAÇÃO DA METODOLOGIA DE CÁLCULO MOLECULAR	101
3.4.1 Comprimentos de ligação	104
3.4.2 Ângulos de ligação	104
3.4.3 Método de escolha	107
3.5 MAPAS CONFORMACIONAIS DAS N^2 -FENILGUANINAS	108
3.5.1 Introdução	108
3.5.2 Detalhes metodológicos	110
3.5.3 Resultados	111
3.6 GEOMETRIA MOLECULAR DAS N^2 -FENILGUANINAS	117
3.6.1 Principais conformações de equilíbrio das N^2 -fenilguaninas	117
3.7 CONSIDERAÇÕES FINAIS	123
3.8 CONCLUSÕES DA ANÁLISE CONFORMACIONAL	124
CAPÍTULO 4 QSAR COM PARÂMETROS CLÁSSICOS	125
4.1 INTRODUÇÃO	125
4.1.1 Estudo QSAR de Gambino e colaboradores	126
4.1.2 Estudo QSAR de Hadjipavlou-Litina	129

4.2 ESTRATÉGIA DA ANÁLISE	131
4.3 PARÂMETROS CLÁSSICOS UTILIZADOS.....	132
4.4 ANÁLISE DOS COMPOSTOS META-SUBSTITUÍDOS	133
4.4.1 <i>Atividade inibitória sobre HSV1 TK.....</i>	<i>133</i>
4.4.2 <i>Atividade inibitória sobre HSV2 TK.....</i>	<i>133</i>
4.5 ANÁLISE DOS COMPOSTOS PARA-SUBSTITUÍDOS	141
4.5.1 <i>Atividade inibitória sobre HSV1 TK.....</i>	<i>141</i>
4.5.2 <i>Atividade inibitória sobre HSV2 TK.....</i>	<i>141</i>
4.6 ANÁLISE DOS COMPOSTOS META E PARA-SUBSTITUÍDOS.....	141
4.6.1 <i>Atividade inibitória sobre HSV1 TK.....</i>	<i>141</i>
4.6.2 <i>Atividade inibitória sobre HSV2 TK.....</i>	<i>141</i>
4.7 ANÁLISE DOS COMPOSTOS DI-SUBSTITUÍDOS	141
4.7.1 <i>Atividade inibitória sobre HSV1 TK.....</i>	<i>141</i>
4.7.2 <i>Atividade inibitória sobre HSV2 TK.....</i>	<i>141</i>
4.8 ANÁLISE GLOBAL.....	142
4.8.1 <i>Atividade inibitória sobre HSV1 TK.....</i>	<i>142</i>
4.8.2 <i>Atividade inibitória sobre HSV2 TK.....</i>	<i>145</i>
4.9 CONCLUSÕES DO ESTUDO QSAR COM PARÂMETROS CLÁSSICOS	148
4.10 LISTAGEM DOS PARÂMETROS CLÁSSICOS.....	150
 CAPÍTULO 5 MODELAGEM MOLECULAR E CONFORMAÇÃO BIOATIVA	 154
5.1 INTRODUÇÃO	154
5.2 ANÁLISE ESTRUTURAL DE HSV1 TK.....	155
5.2.1 <i>Estrutura primária de HSV1 TK.....</i>	<i>155</i>
5.2.2 <i>Estrutura secundária de HSV1 TK.....</i>	<i>158</i>
5.2.3 <i>Estruturas terciária e quaternária de HSV1 TK.....</i>	<i>160</i>
5.2.4 <i>Sítio de ligação da 2'-deoxitimidina.....</i>	<i>164</i>
5.2.5 <i>Sítio do ATP.....</i>	<i>171</i>
5.2.6 <i>Sítio do Mg²⁺</i>	<i>172</i>
5.2.7 <i>Comparação das seqüências de outras timidinas cinases de vírus herpes</i>	<i>176</i>
5.2.8 <i>Comparação das seqüências de timidina-cinase de HSV1 e humana.....</i>	<i>178</i>
5.3 ANÁLISE ESTRUTURAL DE HSV2 TK.....	180
5.3.1 <i>Estrutura primária de HSV2 TK.....</i>	<i>180</i>
5.4 MODELAGEM DE HSV1 TK E HSV2 TK.....	183
5.4.1 <i>Preparação da molécula protéica para cálculos moleculares.....</i>	<i>183</i>
5.4.2 <i>Reconstrução das seqüências ausentes de HSV1_TK_dT_2.....</i>	<i>185</i>
5.4.3 <i>Reconstrução das seqüências ausentes de tk_ganc_nov8g.....</i>	<i>187</i>
5.4.4 <i>Modelagem das estruturas terciária e quaternária de HSV2 TK</i>	<i>188</i>
5.4.5 <i>Relaxamento da estrutura cristalina de HSV1 TK e do modelo de HSV2 TK</i>	<i>191</i>
5.5 MODO DE LIGAÇÃO DAS N²-FENILGUANINAS	196
5.5.1 <i>Possibilidades</i>	<i>198</i>

5.5.2 Experimento computacional.....	203
5.5.3 Detalhes do experimento.....	207
5.5.4 Resultados.....	211
5.5.5 Indicação experimental do modo de ligação das N^2 -fenilguaninas.....	214
5.5.6 Natureza das regiões do sítio ativo próximas às posições meta e para.....	218
5.5.7 Ângulos torsionais θ_1 e θ_2 na conformação bioativa.....	221
5.6 CONCLUSÕES DA ANÁLISE CONFORMACIONAL.....	222
CAPÍTULO 6 QSAR COM PARÂMETROS CLÁSSICOS E CALCULADOS.....	225
6.1 INTRODUÇÃO	225
6.1.1 Geometria Molecular	225
6.1.2 Parâmetros incluídos na análise	226
6.1.3 Correlação entre os parâmetros calculados.....	227
6.2 ANÁLISE DOS COMPOSTOS META-SUBSTITUÍDOS	228
6.2.1 Atividade inibitória sobre HSV1 TK.....	228
6.2.2 Atividade inibitória sobre HSV2 TK.....	235
6.3 ANÁLISE DOS COMPOSTOS PARA-SUBSTITUÍDOS	245
6.3.1 Atividade inibitória sobre HSV1 TK.....	245
6.3.2 Atividade inibitória sobre HSV2 TK.....	245
6.4 ANÁLISE DOS COMPOSTOS DI-SUBSTITUÍDOS	246
6.4.1 Atividade inibitória sobre HSV1 TK.....	246
6.4.2 Atividade inibitória sobre HSV2 TK.....	246
6.5 ANÁLISE DOS COMPOSTOS META E PARA-SUBSTITUÍDOS.....	246
6.5.1 Atividade inibitória sobre HSV1 TK.....	246
6.5.2 Atividade inibitória sobre HSV2 TK.....	249
6.6 ANÁLISE GLOBAL.....	253
6.6.1 Atividade inibitória sobre HSV1 TK.....	253
6.6.2 Atividade inibitória sobre HSV2 TK.....	253
6.7 CONCLUSÕES DO ESTUDO QSAR COM PARÂMETROS CLÁSSICOS E CALCULADOS.....	254
6.8 LISTAGEM DOS PARÂMETROS CALCULADOS.....	255
CONCLUSÕES.....	266
ANÁLISE CONFORMACIONAL	266
QSAR COM PARÂMETROS CLÁSSICOS	266
MODELAGEM MOLECULAR E CONFORMAÇÃO BIOATIVA.....	267
QSAR COM PARÂMETROS CLÁSSICOS E CALCULADOS	268
REFERÊNCIAS BIBLIOGRÁFICAS.....	270

ÍNDICE DE FIGURAS

FIGURA 1 - (A) REPRESENTAÇÃO ESQUEMÁTICA DA ESTRUTURA DO VÍRUS HERPES SIMPLES; (B) MICROGRAFIA ELETRÔNICA DO VÍRUS HERPES SIMPLES. OS HERPESVÍRUS POSSUEM ENVELOPE ENVOLVENDO O CAPSÍDEO ICOSAÉDRICO DE APROXIMADAMENTE 100 NM DE DIÂMETRO, QUE CONTÉM A MOLÉCULA DE DNA. (C) MICROGRAFIA ELETRÔNICA DO VÍRUS HERPES SIMPLES. GLICOPROTEÍNAS CODIFICADAS PELO HSV SÃO INCORPORADAS NO ENVELOPE VIRAL E SÃO VISÍVEIS COMO PICOS LOCALIZADOS NA SUPERFÍCIE DO VÍRION. NESTA MICROGRAFIA, A GLICOPROTEÍNA B (gB) É VISTA COMO PICOS AGRUPADOS NA SUPERFÍCIE EXTERNA DO ENVELOPE, TENDO APROXIMADAMENTE 10 NM DE COMPRIMENTO; (D) MICROGRAFIA ELETRÔNICA DO CAPSÍDEO DO HSV. AS MICROGRAFIAS ELETRÔNICAS FORAM OBTIDAS NO ENDEREÇO HTTP://WWW.TULANE.EDU/~DMSANDER/ Big_VIROLOGY/BVDNAHERPES.HTML.....	38
FIGURA 2 - IMAGENS COMPUTACIONAIS DO CAPSÍDEO DO HSV CONSTRUÍDAS A PARTIR DE INFORMAÇÕES OBTIDAS ATRAVÉS DE EXPERIMENTOS DE CRIOMICROSCOPIA ELETRÔNICA. (A) A UNIDADE ESTRUTURAL DO CAPSÍDEO É O CAPSÔMERO, QUE POSSUI ESTRUTURA PROTÉICA. NO HSV, O CAPSÍDEO É FORMADO POR 162 CAPSÔMEROS, QUE NO CONJUNTO APRESENTAM O FORMATO ICOSAÉDRICO (IMAGEM OBTIDA NO ENDEREÇO HTTP://WWW.BOCKLABS.WISC.EDU/IMAGES/HERPESCAPSID.GIF); (B) CAPSÍDEO DO HSV1; (C) CORTE DO CAPSÍDEO DE HSV1 (IMAGENS (B) E (C) OBTIDAS EM HTTP://WWW.MED.VIRGINIA.EDU/ED-PROGRAMS/BIOPHYSICS/ATTRIB.HTML. A ANIMAÇÃO DA CONSTRUÇÃO DO CAPSÍDEO PODE SER VISTA EM HTTP://RSB.INFO.NIH.GOV/NIH-IMAGE/JAVA/ROTATOR/HSV.HTML).	39
FIGURA 3 - ALGUMAS DAS MANIFESTAÇÕES CLÍNICAS DO HERPES SIMPLES. (A) HERPES NA REGIÃO ENTRE A BOCA E O NARIZ (HTTP://WWW.MEDIC.MIE-U.AC.IP/DERMA/BILddb/BILDER/CD2/IMG0020.HTM); (B) HERPES GENITAL (HTTP://MED-WWW.BU.EDU/PEOPLE/SYCAMORE/STD/HERPES.HTM); (C) HERPES NASAL (HTTP://WWW.MEDIC.MIE-U.AC.IP/DERMA/BILddb/BILDER/CD18/IMG0053.HTM); (D) CONJUNTIVITE HERPÉTICA (HTTP://WWW.DERMA.MED.UNI-ERLANGEN.DE/BILddb/DIAGNOSE/ENGLISCH/I054900.HTM); (E) HERPES LABIAL (HTTP://WWW.MEDIC.MIE-U.AC.IP/DERMA/BILddb/BILDER/CD18/IMG0045.HTM); (F) HERPES NEONATAL (HTTP://TRAY.DERMATOLOGY.UIOWA.EDU/HSVCON01.HTM).	42
FIGURA 4 - REPLICAÇÃO VIRAL (VER DETALHES NO TEXTO).	45
FIGURA 5 - ESQUEMA DO CICLO EVOLUTIVO DO VÍRUS HERPES, ILUSTRADO PARA O CASO DE INFECÇÕES CUTÂNEAS.....	47
FIGURA 6 - REPLICAÇÃO DA MOLÉCULA DE DNA DO HSV. NESTE ESQUEMA, O CRESCIMENTO DA CADEIA DÁ-SE NO SENTIDO 5'→3'. A DNA-POLIMERASE CATALISA O PROCESSO DE INCORPORAÇÃO DAS BASES NITROGENADAS (G, C, T, A) NA NOVA CADEIA.....	50
FIGURA 7 - REPLICAÇÃO DA MOLÉCULA DE DNA DO HSV. NESTE ESQUEMA, O CRESCIMENTO DA CADEIA DÁ-SE NO SENTIDO 5'→3'. A INCLUSÃO DA MOLÉCULA DE ACICLOVIR NA CADEIA DE DNA PROVOCA A INTERRUPÇÃO DO CRESCIMENTO DA CADEIA DEVIDO À AUSÊNCIA DO GRUPO 3'-OH.....	52
FIGURA 8 - (A) TAUTOMERIA ALDO-ENÓLICA; (B) TAUTOMERIA CETO-ENÓLICA.	90
FIGURA 9 - (A) TAUTOMERIA DE PRÓTON; (B) TAUTOMERIA DE VALÊNCIA.	90
FIGURA 10 - TAUTÔMEROS MAIS ESTÁVEIS DAS BASES NITROGENADAS DOS ÁCIDOS NUCLEÍCOS.	92
FIGURA 11 - TAUTÔMEROS DA GUANINA MAIS FREQUENTEMENTE CITADOS COMO SENDO OS MAIS ESTÁVEIS.	93

- FIGURA 12 - COMPRIMENTOS DE LIGAÇÃO (EM Å) DOS DERIVADOS DA GUANINA, OBTIDOS ATRAVÉS DE CRISTALOGRAFIA DE RAIOS-X. 9-EG/1-MC E 9-EG/1-M-5-FC, 9-ETILGUANINA EM COMPLEXO DE 1:1 COM 1-METILCITOSINA, E COM 1-METIL-5-FLUOROCITOSINA (O'BRIEN, 1967); GUANOSINA (THEWALT ET AL, 1970); G.H₂O, GUANINA MONOHIDRATADA (THEWALT ET AL, 1971); 9-EG, 9-ETILGUANINA (DESTRO ET AL, 1974).**..... 98
- FIGURA 13 - MÉDIA ARITMÉTICA DOS COMPRIMENTOS DE LIGAÇÃO (EM Å) DOS DERIVADOS DA GUANINA, APRESENTADOS NA FIGURA 12. OS NÚMEROS ENTRE PARÊNTESES CORRESPONDEM AOS MILÉSIMOS DE UNIDADE DO DESVIO-PADRÃO DA MÉDIA.** 99
- FIGURA 14 - ÂNGULOS DE LIGAÇÃO, EM GRAUS, DOS DERIVADOS DA GUANINA, OBTIDOS ATRAVÉS DE CRISTALOGRAFIA DE RAIOS-X. 9-EG/1-MC E 9-EG/1-M-5-FC, 9-ETILGUANINA EM COMPLEXO DE 1:1 COM 1-METILCITOSINA, E COM 1-METIL-5-FLUOROCITOSINA (O'BRIEN, 1967); GUANOSINA (THEWALT ET AL, 1970); G.H₂O, GUANINA MONOHIDRATADA (THEWALT ET AL, 1971); 9-EG, 9-ETILGUANINA (DESTRO ET AL, 1974).**..... 100
- FIGURA 15 - MÉDIA ARITMÉTICA DOS ÂNGULOS DE LIGAÇÃO DOS DERIVADOS DA GUANINA, APRESENTADOS NA FIGURA 14. OS NÚMEROS ENTRE PARÊNTESES CORRESPONDEM AOS DÉCIMOS DE UNIDADE DO DESVIO-PADRÃO DA MÉDIA.** 101
- FIGURA 16 - ESTRUTURA MOLECULAR DOS DERIVADOS DA N²-FENILGUANINA. X E Y SÃO SUBSTITUÍNTES DIVERSOS (VER TABELA 1, PAG. 57). OS ÂNGULOS TORSIONAIS θ_1 E θ_2 FORNECEM GRAUS DE LIBERDADE ROTACIONAIS CAPAZES DE GERAR GRANDE NÚMERO DE CONFORMEROS.**..... 109
- FIGURA 17 - MAPAS CONFORMACIONAIS BIDIMENSIONAL E TRIDIMENSIONAL CORRESPONDENTES AOS VALORES RELATIVOS DOS CALORES DE FORMAÇÃO APRESENTADOS (E) NA TABELA 8, REFERENTES AO COMPOSTO 3'-CF₃, 4'-H (COMPOSTO 1, TABELA 1, PAG. 57). O GRÁFICO TRIDIMENSIONAL SOFREU ROTAÇÃO DE APROXIMADAMENTE.**..... 112
- FIGURA 18 - MAPAS CONFORMACIONAIS BIDIMENSIONAL E TRIDIMENSIONAL CORRESPONDENTES AOS VALORES RELATIVOS DOS CALORES DE FORMAÇÃO (E) APRESENTADOS NA TABELA 9, REFERENTES AO COMPOSTO 3'-Br, 4'-Br (COMPOSTO 10, TABELA 1, PAG. 57). O GRÁFICO TRIDIMENSIONAL SOFREU ROTAÇÃO DE APROXIMADAMENTE.**..... 113
- FIGURA 19 - MAPAS CONFORMACIONAIS BIDIMENSIONAL E TRIDIMENSIONAL CORRESPONDENTES AOS VALORES RELATIVOS DOS CALORES DE FORMAÇÃO (E) APRESENTADOS NA TABELA 10, REFERENTES AO COMPOSTO 3'-N-PROP, 4'-H (COMPOSTO 11, TABELA 1, PAG. 57). O GRÁFICO TRIDIMENSIONAL SOFREU ROTAÇÃO DE** 114
- FIGURA 20 - MAPAS CONFORMACIONAIS BIDIMENSIONAL E TRIDIMENSIONAL CORRESPONDENTES AOS VALORES RELATIVOS DOS CALORES DE FORMAÇÃO (E) APRESENTADOS NA TABELA 11, REFERENTES AO COMPOSTO 3'-H, 4'-H (COMPOSTO 15, TABELA 1, PAG. 57). O GRÁFICO TRIDIMENSIONAL SOFREU ROTAÇÃO DE** 115
- FIGURA 21 - MAPA CONFORMACIONAL DAS N²-FENILGUANINAS. AS ÁREAS DESTACADAS CORRESPONDEM ÀS REGIÕES DE BAIXA ENERGIA CONFORMACIONAL.** 116
- FIGURA 22 - VALORES INICIAIS DE θ_1 E θ_2 UTILIZADOS NA LOCALIZAÇÃO DAS PRINCIPAIS CONFORMAÇÕES DE EQUILÍBRIO DAS N²-FENILGUANINAS. CADA PONTO SIGNIFICA UM CONJUNTO PARTICULAR DE VALORES DE θ_1 E θ_2 , QUE DEFINEM A GEOMETRIA MOLECULAR INICIAL A SER SUBMETIDA À OTIMIZAÇÃO.** 118

- FIGURA 23 - MAPAS CONFORMACIONAIS DOS COMPOSTOS 1 (3'-CF₃, 4'-H), 10 (3'-Br, 4'-Br), 11 (3'-N-Prop, 4'-H) e 15 (3'-H, 4'-H), MOSTRANDO A POSIÇÃO DAS PRINCIPAIS CONFORMAÇÕES DE EQUILÍBRIO, C1, C2, C3 e C4 (VER TAMBÉM A TABELA 12). O SÍMBOLO "X" INDICA AS COORDENADAS DE PARTIDA (θ_1 e θ_2 INICIAIS) DAS DIVERSAS OTIMIZAÇÕES DA GEOMETRIA MOLECULAR..... 120
- FIGURA 24 - ESTEREOGRAMA DA CONFORMAÇÃO C1 DO COMPOSTO 1 (3'-CF₃, 4'-H, TABELA 1, PAG. 57). $\theta_1 = 184^\circ$ e $\theta_2 = 58^\circ$ 121
- FIGURA 25 - ESTEREOGRAMA DA CONFORMAÇÃO C2 DO COMPOSTO 1 (3'-CF₃, 4'-H, TABELA 1, PAG. 57). $\theta_1 = 174^\circ$ e $\theta_2 = 127^\circ$ 122
- FIGURA 26 - ESTEREOGRAMA DA CONFORMAÇÃO C3 DO COMPOSTO 1 (3'-CF₃, 4'-H, TABELA 1, PAG. 57). $\theta_1 = 186^\circ$ e $\theta_2 = 233^\circ$ 122
- FIGURA 27 - ESTEREOGRAMA DA CONFORMAÇÃO C4 DO COMPOSTO 1 (3'-CF₃, 4'-H, TABELA 1, PAG. 57). $\theta_1 = 176^\circ$ e $\theta_2 = 303^\circ$ 123
- FIGURA 28 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N²-FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (101). AS LINHAS PONTILHADAS SÃO HIPÉRBOLES QUE DEMARCAM A REGIÃO CORRESPONDENTE A 95% DE CONFIANÇA DOS VALORES PREVISTOS..... 135
- FIGURA 29 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N²-FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (102). 137
- FIGURA 30 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N²-FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (103). 140
- FIGURA 31 - ATIVIDADE INIBITÓRIA OBSERVADA DO CONJUNTO GLOBAL DOS DERIVADOS DA N²-FENILGUANINA SOBRE HSV1 TK EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (104). 145
- FIGURA 32 - ATIVIDADE INIBITÓRIA OBSERVADA DO CONJUNTO GLOBAL DOS DERIVADOS DA N²-FENILGUANINA SOBRE HSV2 TK EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (105). 148
- FIGURA 33 - CONCLUSÕES GERAIS DO ESTUDO QSAR COM PARÂMETROS CLÁSSICOS. LIPOFÍLICO+, SUBSTITUINTES MAIS LIPOFÍLICOS AUMENTAM A ATIVIDADE; HIDROFÍLICO+, SUBSTITUINTES MAIS HIDROFÍLICOS AUMENTAM A ATIVIDADE; RESSONÂNCIA+, SUBSTITUINTES CAPAZES DE ATRAIR ELÉTRONS POR MECANISMO DE RESSONÂNCIA AUMENTAM A ATIVIDADE; INDUTIVO+, SUBSTITUINTES CAPAZES DE ATRAIR ELÉTRONS POR MECANISMO DE INDUÇÃO AUMENTAM A ATIVIDADE. 150
- FIGURA 34 - ESQUEMA DA ESTRUTURA TERCIÁRIA DO MONÔMERO DE HSV1 TK, PRESENTE NO ARQUIVO HSV1_TK_dT_2. CÍRCULOS REPRESENTAM α -HÉLICES E QUADRADOS REPRESENTAM FITAS β . LINHAS TRACEJADAS INDICAM SEQUÊNCIAS AUSENTES. A ORIENTAÇÃO ESPACIAL DOS ELEMENTOS PRESENTES É INDICADA PELA DISPOSIÇÃO DAS SETAS. NOTA-SE QUE TODAS AS FITAS β QUE FORMAM A FOLHA β MAIOR ESTÃO PROJETADAS PARA FORA DO PLANO DO PAPEL. A POSIÇÃO DE CADA ELEMENTO AO LONGO DA SEQUÊNCIA DE AMINOÁCIDOS PODE SER OBSERVADA NO DIAGRAMA 3 (PAG. 163). OS ELEMENTOS DE ESTRUTURA SECUNDÁRIA PEQUENOS OU INCOMPLETOS NÃO FORAM INCLUÍDOS NESTE ESQUEMA (WILD ET AL, 1995). 161
- FIGURA 35 - REPRESENTAÇÃO GRÁFICA DA ESTRUTURA CRISTALOGRAFICA DE HSV1 TK, PRESENTE NO ARQUIVO HSV1_TK_dT_2. A REPRESENTAÇÃO MOSTRA A ESTRUTURA NA FORMA ORIGINAL E SEM MODIFICAÇÕES. A ENZIMA APRESENTA-SE NA FORMA DE DÍMERO E COM SIMETRIA C₂ (O EIXO DE SIMETRIA

PASSA PELO CENTRO DO DÍMERO, NA DIREÇÃO VERTICAL). PODE-SE NOTAR A PRESENÇA DO SUBSTRATO 2'-DEOXTIMIDINA E DO GRUPO SULFATO NOS SÍTIOS ATIVOS DE CADA MONÔMERO. AS BOLAS VERMELHAS ISOLADAS REPRESENTAM OS ÁTOMOS DE OXIGÊNIO DAS MOLÉCULAS DE ÁGUA (45 ÁTOMOS DE OXIGÊNIO). AS α -HÉLICES SÃO MOSTRADAS EM VERMELHO, AS FITAS β EM AZUL E AS VOLTAS E DEMAIS SEGMENTOS EM CINZA.	162
FIGURA 36 - REPRESENTAÇÃO GRÁFICA DA ESTRUTURA CRISTALOGRÁFICA DE HSV1 TK, PRESENTE NO ARQUIVO TK_GANG_NOV8G, CEDIDO PELO DR. M.R. SANDERSON. A REPRESENTAÇÃO MOSTRA A ESTRUTURA NA FORMA ORIGINAL E SEM MODIFICAÇÕES. AS DIFERENÇAS EM RELAÇÃO À FIGURA 35 SÃO A PRESENÇA DE GANCICLOVIR NO LUGAR DE 2'-DEOXTIMIDINA, E A MAIOR QUANTIDADE DE MOLÉCULAS DE ÁGUA (173).	163
FIGURA 37 - ESTEREOGRAMA DO DÍMERO DE HSV1 TK, CONSTRUÍDO A PARTIR DO ARQUIVO HSV1_TK_DT_2.	163
FIGURA 38 - ESTEREOGRAMA DO MONÔMERO DE HSV1 TK, CONSTRUÍDO A PARTIR DO ARQUIVO HSV1_TK_DT_2.	164
FIGURA 39 - REPRESENTAÇÃO DO SÍTIO DE LIGAÇÃO DE 2'-DEOXTIMIDINA EM HSV1 TK. ALGUNS ÁTOMOS E LIGAÇÕES DA MOLÉCULA DE 2'-DEOXTIMIDINA FORAM OMITIDOS PARA MAIOR CLAREZA. O FRAGMENTO TYR 172 ENCONTRA-SE ATRÁS DO PLANO DO PAPEL, ENQUANTO QUE MET 128 ESTÁ ACIMA DO PLANO DO PAPEL. AS LINHAS RETAS TRACEJADAS, ENTRE 2'-DEOXTIMIDINA E OS DIVERSOS FRAGMENTOS, REPRESENTAM INTERAÇÕES INTERMOLECULARES.	166
FIGURA 40 - REPRESENTAÇÃO GRÁFICA DO MODO DE LIGAÇÃO DE 2'-DEOXTIMIDINA. VER DETALHES NA LEGENDA DA FIGURA 39.	167
FIGURA 41 - ESTEREOGRAMA DO MODO DE LIGAÇÃO DE 2'-DEOXTIMIDINA NO SÍTIO ATIVO DE HSV1 TK..	167
FIGURA 42 - REPRESENTAÇÃO DO SÍTIO DE LIGAÇÃO DE GANCICLOVIR EM HSV1 TK. OBSERVA-SE A ORIENTAÇÃO ESPACIAL INVERTIDA DO GRUPO AMIDA DE GLN 125, EM RELAÇÃO À FIGURA 39. VER EXPLICAÇÃO DOS DETALHES NA FIGURA 39.	169
FIGURA 43 - REPRESENTAÇÃO GRÁFICA DO MODO DE LIGAÇÃO DE GANCICLOVIR. VER DETALHES NA LEGENDA DA FIGURA 39.	170
FIGURA 44 - ESTEREOGRAMA DO MODO DE LIGAÇÃO DE GANCICLOVIR NO SÍTIO ATIVO DE HSV1 TK.	170
FIGURA 45 - SÍTIO DE LIGAÇÃO DO Mg^{2+} EM HSV TK. DE ACORDO COM BROWN E COLABORADORES (BROWN ET AL, 1995), O RESÍDUO RESPONSÁVEL PELA COORDENAÇÃO DE Mg^{2+} EM HSV1 TK É ASP 162. NO ENTANTO, HÁ INDICAÇÕES DE QUE GLU 83 TAMBÉM PARTICIPA NA COORDENAÇÃO DESSE ÍON.	174
FIGURA 46 - COMPLEXO ATP- Mg^{2+} NA POSIÇÃO DE TRANSFERÊNCIA DO GRUPO FOSFATO PARA A 2'-DEOXTIMIDINA. A POSIÇÃO ESPACIAL DE GLU 83 É COMPATÍVEL PARA QUE POSSA PARTICIPAR DA COORDENAÇÃO DE Mg^{2+}	175
FIGURA 47 - ESTEREOGRAMA COMPLEXO ATP- Mg^{2+} NA POSIÇÃO DE TRANSFERÊNCIA DO GRUPO FOSFATO PARA A 2'-DEOXTIMIDINA.	176
FIGURA 48 - PRIMEIRA POSSÍVEL CONFIGURAÇÃO BIOATIVA DO ANEL PURÍNICO DAS N^2 -FENILGUANINAS. (A) A CONFIGURAÇÃO DO ANEL PURÍNICO DE GANCICLOVIR FOI UTILIZADO COMO MOLDE (BROWN ET AL, 1995). (B) A PONTE DE HIDROGÊNIO DUPLA COM GLN 125 É USADA COMO REFERÊNCIA PARA O ESTABELECIMENTO DA CONFIGURAÇÃO BIOATIVA DO ANEL PURÍNICO DAS N^2 -FENILGUANINAS.	199

- FIGURA 49 - POSSÍVEIS CONFORMAÇÕES DO ANEL FENILA DE N^2 -FENILGUANINA, CONSIDERANDO-SE A CONFIGURAÇÃO APRESENTADA NA FIGURA 48B. NA CONFORMAÇÃO (A) O ÂNGULO TORSIONAL θ_1 É CERCA DE 0° E NA CONFORMAÇÃO (B) θ_1 É APROXIMADAMENTE 180° . DEVIDO ÀS INTERAÇÕES INTRAMOLECULARES, O ANEL FENILA NÃO SE ENCONTRA NO MESMO PLANO DO ANEL PURÍNICO. (C) O MODO DE LIGAÇÃO DE GANCICLOVIR TAMBÉM É APRESENTADO PARA EFEITO DE COMPARAÇÃO..... 200**
- FIGURA 50 - SEGUNDA POSSÍVEL CONFIGURAÇÃO BIOATIVA DO ANEL PURÍNICO DAS N^2 -FENILGUANINAS. (A) A CONFIGURAÇÃO DO ANEL PIRIMIDÍNICO DE 2'-DEOXTIMIDINA FOI UTILIZADA COMO MOLDE (BROWN ET AL., 1995); (B) A PONTE DE HIDROGÊNIO DUPLA COM GLN 125 É USADA COMO REFERÊNCIA PARA O ESTABELECIMENTO DA CONFIGURAÇÃO BIOATIVA DO ANEL PURÍNICO DAS N^2 -FENILGUANINAS. 202**
- FIGURA 51 - ÚNICA CONFORMAÇÃO POSSÍVEL DO ANEL FENILA DAS N^2 -FENILGUANINAS, CONSIDERANDO-SE A CONFIGURAÇÃO APRESENTADA NA FIGURA 50B. (A) NESTA CONFORMAÇÃO, O ÂNGULO TORSIONAL θ_1 É CERCA DE 180° . A CONFORMAÇÃO EM QUE θ_1 É CERCA DE 0° NÃO É PERMITIDA DEVIDO AO IMPEDIMENTO ESTÉRICO ENTRE O ANEL FENILA E RESÍDUOS DE AMINOÁCIDOS DO RECEPTOR. DEVIDO ÀS INTERAÇÕES INTRAMOLECULARES, O ANEL FENILA NÃO SE ENCONTRA NO MESMO PLANO DO ANEL PURÍNICO. (B) O MODO DE LIGAÇÃO DE 2'-DEOXTIMIDINA É APRESENTADO PARA EFEITO DE COMPARAÇÃO..... 203**
- FIGURA 52 - COMPARAÇÃO DA POSIÇÃO DOS ÁTOMOS PASSÍVEIS DE FOSFORILAÇÃO DE GANCICLOVIR E N^2 -FENIL-2'-DEOXI-GUANOSINA. (A) MODO DE LIGAÇÃO DE GANCICLOVIR NO SÍTIO ATIVO DE HSV1 TK E POSIÇÃO ESPACIAL DO ÁTOMO DE OXIGÊNIO QUE É FOSFORILADO; (B) SUPERPOSIÇÃO DAS MOLÉCULAS DE GANCICLOVIR E N^2 -FENIL-2'-DEOXI-GUANOSINA. A DISTÂNCIA ENTRE OS ÁTOMOS DE OXIGÊNIO DAS HIDROXILAS PASSÍVEIS DE FOSFORILAÇÃO DAS DUAS MOLÉCULAS É CERCA DE 0,5 Å. SE AS N^2 -FENILGUANOSINAS ADOTASSEM O MODO DE LIGAÇÃO DE GANCICLOVIR, O ÁTOMO O5' FICARIA EM POSIÇÃO ESPACIAL FAVORÁVEL PARA SER FOSFORILADO..... 215**
- FIGURA 53 - SOBREPOSIÇÃO DA MOLÉCULA DE GANCICLOVIR (AMARELO), EM SUA CONFORMAÇÃO BIOATIVA, E N^2 -FENIL-2'-DEOXIGUANOSINA (AZUL). A DISTÂNCIA ENTRE OS ÁTOMOS DE OXIGÊNIO DAS HIDROXILAS DESTACADAS É CERCA DE 0,5 Å. 216**
- FIGURA 54 - COMPARAÇÃO DA POSIÇÃO DOS ÁTOMOS O5' DE 2'-DEOXTIMIDINA E N^2 -FENIL-2'-DEOXI-GUANOSINA. (A) MODO DE LIGAÇÃO DE 2'-DEOXTIMIDINA AO SÍTIO ATIVO DE HSV1 TK E POSIÇÃO ESPACIAL DO ÁTOMO O5'. (B) SUPERPOSIÇÃO DAS MOLÉCULAS DE 2'-DEOXTIMIDINA E N^2 -FENIL-2'-DEOXI-GUANOSINA. A MENOR DISTÂNCIA ENTRE OS ÁTOMOS O5' DAS DUAS MOLÉCULAS QUE PODE SER OBTIDA POR MANIPULAÇÃO DOS ÂNGULOS TORSIONAIS ϕ_1 E ϕ_2 É CERCA DE 2,8 Å..... 217**
- FIGURA 55 - SOBREPOSIÇÃO DA MOLÉCULA DE 2'-DEOXTIMIDINA (AMARELO), EM SUA CONFORMAÇÃO BIOATIVA, E N^2 -FENIL-2'-DEOXIGUANOSINA (AZUL). A DISTÂNCIA ENTRE OS ÁTOMOS DE OXIGÊNIO DAS HIDROXILAS DESTACADAS É SUPERIOR A 2,8 Å. 218**
- FIGURA 56 - REGIÃO DO SÍTIO ATIVO DE HSV1 TK NAS VIZINHANÇAS DA POSIÇÃO META DO ANEL FENILA DAS N^2 -FENILGUANINAS. 219**
- FIGURA 57 - REGIÃO DO SÍTIO ATIVO DE HSV2 TK NAS VIZINHANÇAS DA POSIÇÃO META DO ANEL FENILA DAS N^2 -FENILGUANINAS. 220**
- FIGURA 58 - ESTEREOGRAMA DA CONFORMAÇÃO BIOATIVA DAS N^2 -FENILGUANINAS, DE ACORDO COM AS CONCLUSÕES DO PRESENTE ESTUDO, EXEMPLIFICADA PELO DERIVADO 3'-CF₃ (COMPOSTO 1, TABELA 1,**

PAG. 57). ESTA CONFORMAÇÃO É SEMELHANTE À CONFORMAÇÃO C3, APRESENTADA NA SEÇÃO 3.6.1 (“PRINCIPAIS CONFORMAÇÕES DE EQUILÍBRIO DAS N^2 -FENILGUANINAS”, PAG. 119).....	222
FIGURA 59 - REPRESENTAÇÃO GRÁFICA DO MODO DE LIGAÇÃO DAS N^2 -FENILGUANINAS. OS SUBSTITUÍNTES LIGADOS À POSIÇÃO META DO ANEL FENILA ORIENTAM-SE EM DIREÇÃO AO RESÍDUO TYR 101.	223
FIGURA 60 - ESTEREOGRAMA DO MODO DE LIGAÇÃO PROPOSTO PARA AS N^2 -FENILGUANINAS.	224
FIGURA 61 - NUMERAÇÃO DOS ÁTOMOS NA MOLÉCULA DA N^2 -FENILGUANINA.....	227
FIGURA 62 - DENSIDADE ELETRÔNICA DO HOMO (MODELO DE CIMA) E LUMO DO COMPOSTO 15 (3'-H, 4'-H). ESSES ORBITAIS NÃO APRESENTAM DENSIDADE ELETRÔNICA SOBRE O ÁTOMO H(12) (SETA).....	231
FIGURA 63 - ESTEREOGRAMA DA SITUAÇÃO FINAL TÍPICA DO MODO DE LIGAÇÃO DAS N^2 -FENILGUANINAS APÓS A SIMULAÇÃO DE DINÂMICA MOLECULAR EXECUTADA NA SEÇÃO 5.5.3 (“DETALHES DO EXPERIMENTO”, PAG. 211). NESTE EXEMPLO, APENAS UMA PONTE DE HIDROGÊNIO FOI MANTIDA ENTRE O COMPOSTO 1 E GLN 125 (LINHA TRACEJADA, DISTÂNCIA DE 1,8 Å). A OUTRA PONTE NÃO PÔDE SER FORMADA DEVIDO À LONGA DISTÂNCIA ENTRE O SEGUNDO PAR DE ÁTOMOS (3,8 Å).	232
FIGURA 64 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (113).	234
FIGURA 65 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (122).	236
FIGURA 66 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (123).	238
FIGURA 67 - ESTEREOGRAMA DO SÍTIO ATIVO DE HSV1 TK, DESTACANDO OS AMINOÁCIDOS ILE 97, ALA 98 E ALA 168, QUE NÃO OCORREM NO SÍTIO ATIVO DE HSV2 TK.	241
FIGURA 68 - ESTEREOGRAMA DO SÍTIO ATIVO DE HSV2 TK, DESTACANDO OS AMINOÁCIDOS LEU 97, THR 98 E SER 168, QUE NÃO OCORREM NO SÍTIO ATIVO DE HSV2 TK.....	242
FIGURA 69 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (126).	245
FIGURA 70 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META E PARA-SUBSTITUÍDOS DA N^2 -FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (132).	249
FIGURA 71 - ATIVIDADE INIBITÓRIA OBSERVADA DOS DERIVADOS META E PARA-SUBSTITUÍDOS DA N^2 -FENILGUANINAS EM FUNÇÃO DA ATIVIDADE PREVISTA PELA EQUAÇÃO (134).	253
FIGURA 72 - CONCLUSÕES GERAIS DO ESTUDO QSAR. OS RESULTADOS INDICARAM QUE A INTERAÇÃO ENTRE O ÁTOMO O ₁₀ E UM ÁTOMO CARREGADO POSITIVAMENTE NO SÍTIO ATIVO É IMPORTANTE PARA A ATIVIDADE INIBITÓRIA. TAMBÉM HOUE INDICAÇÃO DE QUE A DIFERENÇA MAIS IMPORTANTE ENTRE OS SÍTOS ATIVOS DAS ENZIMAS HSV1 TK E HSV2 TK É A OCORRÊNCIA DE EFEITO ESTÉRICO SOBRE OS DERIVADOS META EM HSV2 TK. LIPOFÍLICO+: SUBSTITUÍNTES MAIS LIPOFÍLICOS AUMENTAM A ATIVIDADE; HIDROFÍLICO+: SUBSTITUÍNTES MAIS HIDROFÍLICOS AUMENTAM A ATIVIDADE; RESSONÂNCIA+: SUBSTITUÍNTES CAPAZES DE ATRAIR ELÉTRONS POR MECANISMO DE RESSONÂNCIA AUMENTAM A ATIVIDADE; INDUTIVO+: SUBSTITUÍNTES CAPAZES DE ATRAIR ELÉTRONS POR MECANISMO DE INDUÇÃO AUMENTAM A ATIVIDADE.	255

ÍNDICE DE TABELAS

TABELA 1 - ESTRUTURA QUÍMICA E ATIVIDADE INIBITÓRIA DOS DERIVADOS DA N^2 -FENILGUANINA, CONTRA HSV1 TK E HSV2 TK. IC_{50} É A CONCENTRAÇÃO MOLAR NECESSÁRIA PARA PROVOCAR 50% DE INIBIÇÃO DA ATIVIDADE ENZIMÁTICA (GAMBINO ET AL, 1992).	56
TABELA 2 - ENERGIAS RELATIVAS DOS TAUTÔMEROS DA GUANINA, APRESENTADOS NA FIGURA 11, CALCULADAS PELOS MÉTODOS SEMIEMPÍRICOS MNDO, AM1, PM3 E SCRF-AM1 (SCRF, <i>SELF-CONSISTENT REACTION FIELD</i>). A COLUNA <i>ORDEM DE ENERGIA DOS TAUTÔMEROS</i> EXIBE A ORDEM CRESCENTE DAS ENERGIAS (1, MENOR ENERGIA; 2; 3; 4; 5, MAIOR ENERGIA).	94
TABELA 3 - ENERGIAS RELATIVAS DOS TAUTÔMEROS DA GUANINA, APRESENTADOS NA FIGURA 11, CALCULADAS POR COMBINAÇÃO DE MÉTODOS SEMIEMPÍRICO (AM1) E <i>AB INITIO</i> (HARTREE-FOCK-ROOTHAAN). A COLUNA <i>ORDEM DE ENERGIA DOS TAUTÔMEROS</i> EXIBE A ORDEM CRESCENTE DAS ENERGIAS (1, MENOR ENERGIA; 2; 3; 4; 5, MAIOR ENERGIA).	95
TABELA 4 - ENERGIAS RELATIVAS DOS TAUTÔMEROS DA GUANINA, APRESENTADOS NA FIGURA 11, CALCULADAS POR MÉTODO <i>AB INITIO</i> DE ALTO NÍVEL. A COLUNA <i>ORDEM DE ENERGIA DOS TAUTÔMEROS</i> EXIBE A ORDEM CRESCENTE DAS ENERGIAS (1, MENOR ENERGIA; 2; 3; 4; 5, MAIOR ENERGIA).	96
TABELA 5 - COMPRIMENTOS DE LIGAÇÃO DA GUANINA. COMPAREM-SE OS DIVERSOS VALORES CALCULADOS COM O VALOR EXPERIMENTAL DE REFERÊNCIA.	105
TABELA 6 - ÂNGULOS DE LIGAÇÃO DA GUANINA. COMPAREM-SE OS DIVERSOS VALORES CALCULADOS COM O VALOR EXPERIMENTAL DE REFERÊNCIA.	106
TABELA 7 - TEMPO DE CPU ESTIMADO PARA A EXECUÇÃO DOS CÁLCULOS DE OTIMIZAÇÃO DE GEOMETRIA MOLECULAR NECESSÁRIOS À CONSTRUÇÃO DAS SUPERFÍCIES DE POTENCIAL DE ROTAÇÃO DOS ANÉIS FENILA E GUANINA DAS N^2 -FENILGUANINAS (VER SEÇÃO 3.5, "MAPAS CONFORMACIONAIS DAS N^2 -FENILGUANINAS", PARA MAIORES DETALHES). OS VALORES APRESENTADOS REFEREM-SE AO COMPOSTO 15 (3'-H, 4'-H, TABELA 1, PAG 57), QUE É O MENOR DA SÉRIE.	108
TABELA 8 - VALORES RELATIVOS DOS CALORES DE FORMAÇÃO DOS CONFÔRMEROS ROTACIONAIS DO COMPOSTO 3'-CF ₃ , 4'-H (COMPOSTO 1, TABELA 1, PAG. 57). O VALOR ABSOLUTO DO CALOR DE FORMAÇÃO PARA $\theta_1 = 210$ E $\theta_2 = 330$, QUE CORRESPONDE AO MÍNIMO ABSOLUTO DE ENERGIA, É -70.20 KCAL/MOL.	112
TABELA 9 - VALORES RELATIVOS DOS CALORES DE FORMAÇÃO DOS CONFÔRMEROS ROTACIONAIS DO COMPOSTO 3'-Br, 4'-Br (COMPOSTO 10, TABELA 1, PAG. 57). O VALOR ABSOLUTO DO CALOR DE FORMAÇÃO PARA $\theta_1 = 150$ E $\theta_2 = 300$, QUE CORRESPONDE AO MÍNIMO ABSOLUTO DE ENERGIA, É 96.26 KCAL/MOL.	113
TABELA 10 - VALORES RELATIVOS DOS CALORES DE FORMAÇÃO DOS CONFÔRMEROS ROTACIONAIS DO COMPOSTO 3'-N-Pr, 4'-H (COMPOSTO 11, TABELA 1, PAG. 57). O VALOR ABSOLUTO DO CALOR DE FORMAÇÃO PARA $\theta_1 = 150$ E $\theta_2 = 30$, QUE CORRESPONDE AO MÍNIMO ABSOLUTO DE ENERGIA, É 63.78 KCAL/MOL.	114

TABELA 11 - VALORES RELATIVOS DOS CALORES DE FORMAÇÃO DOS CONFORMEROS ROTACIONAIS DO COMPOSTO 3'-H, 4'-H (COMPOSTO 15, TABELA 1, PAG. 57). O VALOR ABSOLUTO DO CALOR DE FORMAÇÃO PARA $\theta_1 = 210$ E $\theta_2 = 270$, QUE CORRESPONDE AO MÍNIMO ABSOLUTO DE ENERGIA, É 84.25 KCAL/MOL..	115
TABELA 12 - CONFORMAÇÕES DE EQUILÍBRIO GERADAS A PARTIR DOS VALORES INICIAIS DE θ_1 E θ_2 MOSTRADOS NA FIGURA 22.	119
TABELA 13 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO (101).	134
TABELA 14 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO (102).	137
TABELA 15 - PERCENTAGEM DE VARIÂNCIA DA ATIVIDADE INIBITÓRIA DOS DERIVADOS META SOBRE HSV2 TK EXPLICADA POR CADA PARÂMETRO DAS EQUAÇÕES (101), (102) E (103).	139
TABELA 16 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO (103).	140
TABELA 17 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DO CONJUNTO GLOBAL DOS DERIVADOS DA N^2 -FENILGUANINA SOBRE HSV1 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO (104).	144
TABELA 18 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DO CONJUNTO GLOBAL DOS DERIVADOS DA N^2 -FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO (105).	147
TABELA 19 - PARÂMETROS TESTADOS NA ANÁLISE CLÁSSICA DE QSAR (HANSCH ET AL, 1995).	151
TABELA 20 - MATRIZ DE CORRELAÇÃO DOS PARÂMETROS UTILIZADOS NA ANÁLISE CLÁSSICA. VALORES EM TERMOS DE R^2 , QUADRADO DO COEFICIENTE DE CORRELAÇÃO. AS CORRELAÇÕES IGUAIS OU SUPERIORES A 0,5 SÃO MOSTRADAS EM NEGRITO.	153
TABELA 21 - CONTEÚDO DOS ARQUIVOS HSV1_TK_dT_2 E TK_GANC_NOV8G, FORNECIDOS PELO DR. M.R. SANDERSON, COM A ESTRUTURA CRISTALOGRÁFICA DA TIMIDINA-CINASE DO VÍRUS HERPES SIMPLES DO TIPO 1	154
TABELA 22 - DEFINIÇÃO DOS ELEMENTOS DA ESTRUTURA SECUNDÁRIA DE HSV1 TK. DEVIDO AO FATO DO MAPEAMENTO TER SIDO FEITO VISUALMENTE, ALGUMAS PEQUENAS DIFERENÇAS PODEM SER ENCONTRADAS EM RELAÇÃO A OUTROS AUTORES (BROWN ET AL, 1995; WILD ET AL, 1995). OS RESÍDUOS 163 E 169 FAZEM CONEXÃO SIMPLES ENTRE $\beta 3/\alpha 5$ E $\alpha 5/\alpha 6$, RESPECTIVAMENTE. O RESÍDUO 374 NÃO ESTÁ INCORPORADO A $\alpha 15$ E NÃO FAZ PARTE DE QUALQUER ELEMENTO DE ESTRUTURA SECUNDÁRIA.	160
TABELA 23 - VOLTAS PROCURADAS E VOLTAS ESCOLHIDAS PARA OS MONÔMEROS A E B DE HSV1_TK (PRESENTE NO ARQUIVO HSV1_TK_dT_2). NA COLUNA <i>VOLTA ESCOLHIDA</i> , A CONFORMAÇÃO DA SEQUÊNCIA DOS AMINOÁCIDOS EM NEGRITO FOI UTILIZADA PARA MODELAR A SEQUÊNCIA DA VOLTA PROCURADA EM HSV1_TK. A CONFORMAÇÃO DOS AMINOÁCIDOS DAS EXTREMIDADES DA VOLTA ESCOLHIDA FOI USADA COMO REFERÊNCIA PARA FIXAR CADA SEQUÊNCIA NAS EXTREMIDADES DA VOLTA PROCURADA NA MOLÉCULA DE HSV1_TK. RMS (LSQ) (ROOT MEAN SQUARE-LEAST SQUARE FIT)	

CORRESPONDE AO DESVIO QUADRÁTICO MÉDIO DAS DISTÂNCIAS ENTRE OS ÁTOMOS DE CARBONO- α ($C\alpha$) DA SEQUÊNCIA PROCURADA E DA SEQUÊNCIA ENCONTRADA, AJUSTADA POR MÉTODO DE MÍNIMOS QUADRADOS. FIT (DIST) CORRESPONDE À SOMA DAS DIFERENÇAS DAS DISTÂNCIAS ENTRE OS $C\alpha$ DA SEQUÊNCIA PROCURADA E DA SEQUÊNCIA ENCONTRADA.	186
TABELA 24 - VOLTAS PROCURADAS E VOLTAS ESCOLHIDAS PARA OS MONÔMEROS A E B DE HSV1_TK (PRESENTE NO ARQUIVO TK_GANC_NOV8G). VER EXPLICAÇÃO NA TABELA 23, PAG. 190.	188
TABELA 25 - HOMOLOGIA ENTRE OS DIVERSOS ELEMENTOS DA ESTRUTURA SECUNDÁRIA DE HSV1 TK E HSV2 TK. NO. RES. = NÚMERO DE RESÍDUOS EM CADA ELEMENTO DE ESTRUTURA SECUNDÁRIA; HT = HOMOLOGIA TOTAL E HF = HOMOLOGIA FUNCIONAL, DE ACORDO COM A LEGENDA DO DIAGRAMA 5.1, DADAS EM PORCENTAGEM. NOS CASOS EM QUE HT É IGUAL A HF, UTILIZOU-SE HÍFEN NA COLUNA DE HF PARA EFEITO DE SIMPLIFICAÇÃO.	190
TABELA 26 - VOLTAS PROCURADAS E VOLTAS ESCOLHIDAS PARA OS MONÔMEROS A E B DE HSV2_TK, QUE FORAM CONSTRUÍDAS A PARTIR DE HSV1_TK. VER EXPLICAÇÃO NA TABELA 23, PAG. 190.	191
TABELA 27 - VOLTAS PROCURADAS E VOLTAS ESCOLHIDAS PARA OS MONÔMEROS A E B DE HSV2_TK, QUE FORAM CONSTRUÍDAS A PARTIR DE HSV1_TK. VER EXPLICAÇÃO NA TABELA 23, PAG. 190.	191
TABELA 28 - SITUAÇÃO ENERGÉTICA, EM KCAL/MOL, DOS COMPLEXOS DE HSV1_TK-2 MOLÉCULAS DE 2'-DEOXTITIMIDINA-45 MOLÉCULAS DE ÁGUA (10209 ÁTOMOS) E HSV1_TK-2 MOLÉCULAS DE GANCICLOVIR-173 MOLÉCULAS DE ÁGUA (10593 ÁTOMOS) ANTES E DEPOIS DO PROCESSO DE MINIMIZAÇÃO (FASE GASOSA), EXECUTADO PELO PROGRAMA CHARMM VERSÃO 25A1 (BROOKS ET AL, 1983).	194
TABELA 29 - SITUAÇÃO ENERGÉTICA, EM KCAL/MOL, DOS COMPLEXOS DE HSV2_TK-2 MOLÉCULAS DE 2'-DEOXTITIMIDINA-45 MOLÉCULAS DE ÁGUA (10303 ÁTOMOS) E HSV2_TK-2 MOLÉCULAS DE GANCICLOVIR-173 MOLÉCULAS DE ÁGUA (10687 ÁTOMOS) ANTES E DEPOIS DO PROCESSO DE MINIMIZAÇÃO (FASE GASOSA), EXECUTADO PELO PROGRAMA CHARMM VERSÃO 25A1 (BROOKS ET AL, 1983).	195
TABELA 30 - DESVIO QUADRÁTICO MÉDIO ATÔMICO DOS SISTEMAS ENZIMA-LIGANTE-SOLVENTE MINIMIZADOS PELO PROGRAMA CHARMM. O ASTERISCO DEFINE A ENZIMA QUE PARTICIPA DOS COMPLEXOS MOSTRADOS NA SEGUNDA LINHA.	196
TABELA 31 - VARIAÇÃO DA VARIAÇÃO DA ENTALPIA DE INTERAÇÃO HSV1 TK-INIBIDOR ($\Delta\Delta H_{EI}$) DOS DERIVADOS P-H, P-F, P-Cl E P-Br E P-I. NÚMEROS EM NEGRITO FORAM UTILIZADOS PARA DESTACAR A RELAÇÃO LINEAR ENTRE $\log 1/IC_{50}$ E $\Delta\Delta H_{EI}$.	213
TABELA 32 - COMBINAÇÕES DE DUAS VARIÁVEIS, UTILIZADAS PARA REPRESENTAR A VARIÂNCIA DA ATIVIDADE INIBITÓRIA DOS META-DERIVADOS DA N²-FENILGUANINA SOBRE HSV1 TK, QUE RESULTARAM EM COEFICIENTES DE CORRELAÇÃO (R) IGUAIS OU SUPERIORES A 0,9. OS SINAIS DOS COEFICIENTES DAS VARIÁVEIS PRESENTES NAS EQUAÇÕES (112)-(121) SÃO MOSTRADOS NA COLUNA <i>VARIÁVEIS</i>. EM NEGRITO SÃO DESTACADAS AS PRINCIPAIS COMBINAÇÕES DE PARÂMETROS.	228
TABELA 33 - PERCENTAGEM DE VARIÂNCIA DA ATIVIDADE INIBITÓRIA DOS DERIVADOS META-SUBSTITUÍDOS SOBRE HSV1 TK, EXPLICADA POR CADA PARÂMETRO NAS EQUAÇÕES (112)-(114).	232

TABELA 34 - MATRIZ DE CORRELAÇÃO DOS PRINCIPAIS PARÂMETROS APRESENTADOS NA TABELA 32.

SOMENTE OS VALORES RELATIVOS AOS TREZE COMPOSTOS INCLUÍDOS NA ANÁLISE FORAM CONSIDERADOS.

VALORES EM TERMOS DE r^2 233**TABELA 35 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -**

FENILGUANINA SOBRE HSV1 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO

(113). 234

TABELA 36 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -

FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO

(122). 236

TABELA 37 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -

FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO

(123). 237

TABELA 38 - COMBINAÇÕES DE TRÊS VARIÁVEIS, UTILIZADAS PARA REPRESENTAR A VARIÂNCIA DAATIVIDADE INIBITÓRIA DOS META-DERIVADOS DA N^2 -FENILGUANINA SOBRE HSV2 TK, QUE RESULTARAM

EM COEFICIENTES DE CORRELAÇÃO (R) SUPERIORES A 0,97. OS SINAIS DOS COEFICIENTES DAS VARIÁVEIS

PRESENTES NAS EQUAÇÕES (124)-(130) SÃO MOSTRADOS NA COLUNA *VARIÁVEIS*. EM NEGRITO SÃO

DESTACADAS AS PRINCIPAIS COMBINAÇÕES. 239

TABELA 39 - MATRIZ DE CORRELAÇÃO DOS PARÂMETROS APRESENTADOS NAS EQUAÇÕES (124)-(130).

SOMENTE OS VALORES RELATIVOS AOS 13 COMPOSTOS INCLUÍDOS NESSAS EQUAÇÕES FORAM

CONSIDERADOS. VALORES EM TERMOS DE r^2 243**TABELA 40 - PERCENTAGEM DE VARIÂNCIA DA ATIVIDADE INIBITÓRIA DOS DERIVADOS META-SUBSTITUÍDOS**

SOBRE HSV2 TK EXPLICADA POR CADA PARÂMETRO NAS EQUAÇÕES (126), (127) E (129). 243

TABELA 41 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META-SUBSTITUÍDOS DA N^2 -

FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO

(126). 244

TABELA 42 - MATRIZ DE CORRELAÇÃO DOS PARÂMETROS APRESENTADOS NAS EQUAÇÕES (131) E (132).

SOMENTE OS VALORES RELATIVOS AOS 25 COMPOSTOS INCLUÍDOS NESSAS EQUAÇÕES FORAM

CONSIDERADOS. VALORES EM TERMOS DE r^2 247**TABELA 43 - PERCENTAGEM DE VARIÂNCIA DA ATIVIDADE INIBITÓRIA DOS DERIVADOS META E PARA-**

SUBSTITUÍDOS SOBRE HSV1 TK EXPLICADA POR CADA PARÂMETRO NAS EQUAÇÕES (131) E (132). 247

TABELA 44 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META E PARA-SUBSTITUÍDOSDA N^2 -FENILGUANINA SOBRE HSV1 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA

EQUAÇÃO (132). 248

TABELA 45 - COMBINAÇÕES DE QUATRO VARIÁVEIS, UTILIZADAS PARA REPRESENTAR A VARIÂNCIA DAATIVIDADE INIBITÓRIA DOS META E PARA-DERIVADOS DA N^2 -FENILGUANINA SOBRE HSV2 TK, QUE

RESULTARAM EM COEFICIENTES DE CORRELAÇÃO (R) SUPERIORES A 0,9. EM NEGRITO SÃO DESTACADAS AS

PRINCIPAIS COMBINAÇÕES. 250

TABELA 46 - MATRIZ DE CORRELAÇÃO DOS PARÂMETROS APRESENTADOS NAS EQUAÇÕES (133)-(140).

SOMENTE OS VALORES RELATIVOS AOS 25 COMPOSTOS INCLUÍDOS NESSAS EQUAÇÕES FORAM

CONSIDERADOS. 251

TABELA 47 - PERCENTAGEM DE VARIÂNCIA DA ATIVIDADE INIBITÓRIA DOS DERIVADOS META E PARA-SUBSTITUÍDOS SOBRE HSV2 TK EXPLICADA POR CADA PARÂMETRO NAS EQUAÇÕES (134) E (136).....	251
TABELA 48 - ATIVIDADES INIBITÓRIAS OBSERVADA E PREVISTA DOS DERIVADOS META E PARA-SUBSTITUÍDOS DA N^2-FENILGUANINA SOBRE HSV2 TK, E OS RESPECTIVOS RESÍDUOS, CALCULADOS ATRAVÉS DA EQUAÇÃO (134).	252
TABELA 49 - ENERGIAS DOS ORBITAIS MOLECULARES OCUPADO DE ENERGIA MAIS ALTA, ϵ_{HOMO}, E DESOCUPADO DE ENERGIA MAIS BAIXA, ϵ_{LUMO}, DUREZA MOLECULAR, η, VOLUME MOLECULAR DE VAN DER WAALS, V_{vdW}, MOMENTO DIPOLAR, μ, E POLARIZABILIDADE MOLECULAR, α.	256
TABELA 50 - CARGAS ATÔMICAS (Q_N), CALCULADAS A PARTIR DO POTENCIAL ELETROSTÁTICO MOLECULAR, SOBRE O N-ÉSIMO ÁTOMO DA MOLÉCULA DA N^2-FENILGUANINA.	257
TABELA 51 - DENSIDADE ELETROFÍLICA DE FRONTEIRA, NORMALIZADA PELA ENERGIA DO HOMO, DO N-ÉSIMO ÁTOMO DA N^2-FENILGUANINA ($F_N^{(e)}$) . VALORES MULTIPLICADOS POR 100.	260
TABELA 52 - DENSIDADE NUCLEOFÍLICA DE FRONTEIRA, NORMALIZADA PELA ENERGIA DO LUMO, DO N-ÉSIMO ÁTOMO DA N^2-FENILGUANINA ($F_N^{(n)}$) . VALORES MULTIPLICADOS POR 100.	263

ÍNDICE DE QUADROS

QUADRO 1 - CLASSIFICAÇÃO DOS VÍRUS QUANTO À NATUREZA DO GENOMA (MURRAY ET AL, 1992).....	34
QUADRO 2 - VÍRUS HERPES QUE UTILIZAM O HOMEM COMO HOSPEDEIRO E SUAS PRINCIPAIS CONSEQUÊNCIAS CLÍNICAS (COLLIER E OXFORD, 1993).	36
QUADRO 3 - PRINCIPAIS AGENTES ANTI-HERPES (HAYDEN, 1996, MURRAY ET AL, 1992).....	52
QUADRO 4 - SITUAÇÃO FINAL DA REGIÃO DO SÍTIO ATIVO DOS SISTEMAS HSV1 TK-N^2-FENIL-[PARA-SUBSTITUÍDO]-GUANINA APÓS CURTA SIMULAÇÃO DE DINÂMICA MOLECULAR SEGUIDA DE MINIMIZAÇÃO. CARACTERES EM NEGRITO FORAM UTILIZADOS PARA DESTACAR A COMPARAÇÃO ENTRE OS RESULTADOS OBTIDOS E O RESULTADO ESPERADO.	211

ÍNDICE DE DIAGRAMAS

DIAGRAMA 1 - COMPARAÇÃO DAS ESTRUTURAS PRIMÁRIAS DE HSV1 TK, DISPONÍVEIS NO <i>SWISS-PROT PROTEIN SEQUENCE DATABASE</i>, COM A ESTRUTURA PRIMÁRIA DETERMINADA A PARTIR DO PADRÃO DE DIFRAÇÃO DE RAIOS-X DA ENZIMA CRISTALIZADA (BROWN ET AL, 1995). O ALINHAMENTO FOI FEITO ATRAVÉS DO PROGRAMA CAMELEON. AS ABREVIACÕES DAS SEQUÊNCIAS PREVISTAS DE HSV1 TK (KITH_HSV11/C/E/K/S) CORRESPONDEM AO CÓDIGO IDENTIFICADOR (ID) UTILIZADO PELO SWISS-PROT. OS AMINOÁCIDOS APRESENTADOS EM NEGRITO CORRESPONDEM ÀQUELES DESCARTADOS NA ESCOLHA DA SEQUÊNCIA DEFINITIVA DE HSV1 TK. NA LINHA INDICADA POR <i>CONSENSUS</i>, A PRESENÇA DE LETRA MAIÚSCULA INDICA QUE O MESMO AMINOÁCIDO ESTÁ PRESENTE NA MESMA POSIÇÃO, EM TODAS AS SEQUÊNCIAS COMPARADAS; A PRESENÇA DE LETRA MINÚSCULA INDICA QUE O AMINOÁCIDO REPRESENTADO POR ESSA LETRA ESTÁ PRESENTE, NA MESMA POSIÇÃO, EM PELO MENOS 70% DAS
--

SEQÜÊNCIAS ANALISADAS (70% É O VALOR PROPOSTO PELO PROGRAMA CAMELEON, QUE PODE SER MODIFICADO PELO OPERADOR). A PRESENÇA DE CARACTER ESPECIAL (~, !, @, #, \$, ^, % OU & - VER LEGENDA ABAIXO DO DIAGRAMA) INDICA QUE PELO MENOS 70% DOS AMINOÁCIDOS PRESENTES NAQUELA POSIÇÃO APRESENTAM CARACTERÍSTICAS COMUNS, DE ACORDO COM O CARACTER. NA LINHA INDICADA POR <i>IDENTITY</i> , A PRESENÇA DE ESTRELA RESSALTA A COINCIDÊNCIA TOTAL DOS AMINOÁCIDOS COMPARADOS.....	156
DIAGRAMA 2 - ESTRUTURA PRIMÁRIA COMPLETA DA TIMIDINA-CINASE DO VÍRUS HERPES SIMPLES DO TIPO 1. OS AMINOÁCIDOS AUSENTES NA ESTRUTURA CRISTALINA (BROWN ET AL, 1995) SÃO MOSTRADOS EM NEGRITO. ESTES FORAM SELECIONADOS A PARTIR DA SEQÜÊNCIA KITH_HSV1K, QUE É A SEQÜÊNCIA MAIS ATUALIZADA DISPONÍVEL.....	158
DIAGRAMA 3 - ESTRUTURA SECUNDÁRIA DA TIMIDINA-CINASE DO VÍRUS HERPES SIMPLES DO TIPO 1. ELEMENTOS DE ESTRUTURA SECUNDÁRIA: α = α -HÉLICE, β = FITA β , Ω = VOLTA. A NUMERAÇÃO DAS FITAS β PERTENCENTES À FOLHA β MAIOR FOI PRESERVADA, DE ACORDO COM WILD ET AL (1995). DESSA FORMA, NUMEROU-SE UMA FITA β INTERMEDIÁRIA ACRESCENTANDO O ÍNDICE A (β 4A). OS AMINOÁCIDOS AUSENTES NA ESTRUTURA CRISTALINA DISPONÍVEL SÃO APRESENTADOS EM NEGRITO.....	159
DIAGRAMA 4 - ALINHAMENTO DAS SEQÜÊNCIAS DE TIMIDINA-CINASE DE VÍRUS HERPES DISPONÍVEIS NO SEQNET. INFORMAÇÕES SOBRE AS SEQÜÊNCIAS COMPARADAS PODEM SER OBTIDAS NA LEGENDA DO DIAGRAMA 5, PAG. 181.	173
DIAGRAMA 5 - ALINHAMENTO DAS SEQÜÊNCIAS DE AMINOÁCIDO DE TIMIDINA-CINASE DE VÍRUS HERPES DISPONÍVEIS NO SWISS-PROT. SÃO APRESENTADOS APENAS OS ALINHAMENTOS DOS SEGMENTOS QUE CONTÉM OS PRINCIPAIS PADRÕES DE AMINOÁCIDOS (<i>MOTIFS</i>), DESTACADOS EM NEGRITO. O NÚMERO QUE APARECE ANTES DE CADA SEGMENTO DE SEQÜÊNCIA CORRESPONDE À POSIÇÃO DO PRIMEIRO AMINOÁCIDO DO SEGMENTO NA SEQÜÊNCIA ORIGINAL DE CADA ENZIMA. INFORMAÇÕES SOBRE CADA SEQÜÊNCIA SÃO MOSTRADAS NA LEGENDA. INFORMAÇÕES SOBRE OS SÍMBOLOS QUE APARECEM NAS LINHAS <i>CONSENSUS</i> E <i>IDENTITY</i>, PODEM SER OBTIDAS NO DIAGRAMA 1 (PAG. 160).....	177
DIAGRAMA 6 - ALINHAMENTO DAS SEQÜÊNCIAS DE AMINOÁCIDO DA TIMIDINA-CINASE DE VÍRUS HERPES SIMPLES DO TIPO 1 E DA TIMIDINA-CINASE HUMANA. OS PADRÕES DE AMINOÁCIDOS APRESENTADOS EM NEGRITO NO DIAGRAMA 5 TAMBÉM SÃO APRESENTADOS EM NEGRITO NA SEQÜÊNCIA KITH_HSV1 PARA MELHOR COMPARAÇÃO DAS DUAS SEQÜÊNCIAS.	179
DIAGRAMA 7 - COMPARAÇÃO DAS SEQÜÊNCIAS COMPLETAS DA TIMIDINA-CINASE DO VÍRUS HERPES SIMPLES DOS TIPOS 1 E 2. AS SEQÜÊNCIAS AUSENTES NA ESTRUTURA CRISTALINA DE HSV1 TK E AS SEQÜÊNCIAS	181
DIAGRAMA 8 - ALINHAMENTO FEITO INICIALMENTE PELO PROGRAMA CAMELEON. O PROGRAMA INSERIU UM ESPAÇO NA SEQÜÊNCIA DE HSV1 TK NA POSIÇÃO 67.	182
DIAGRAMA 9 - ALINHAMENTO FEITO INICIALMENTE PELO PROGRAMA CAMELEON. O PROGRAMA INSERIU UM ESPAÇO NA SEQÜÊNCIA DE HSV2 TK NA POSIÇÃO 271.....	182

ABREVIACÕES UTILIZADAS

α	- Alfa hélice
α -TIF	- α - <i>Trans-inducing factor</i>
β	- Fita beta (<i>beta strand</i>)
ΔH_f°	- Entalpia-padrão de formação
ϵ_{HOMO}	- Energia do HOMO
ϵ_{LUMO}	- Energia do LUMO
μ	- Constante dielétrica
π	- Parâmetro hidrofóbico de Hansch-Fujita
π_m	- Parâmetro hidrofóbico de Hansch-Fujita de substituintes na posição <i>meta</i>
π_p	- Parâmetro hidrofóbico de Hansch-Fujita de substituintes na posição <i>para</i>
θ_n	- Ângulo torsional n
σ	- Constante eletrônica de Hammett
σ_m	- Constante eletrônica de Hammett de substituintes na posição <i>meta</i>
σ_p	- Constante eletrônica de Hammett de substituintes na posição <i>para</i>
Ω	- Volta (<i>Loop</i>)
Å	- Angstrom
Å ³	- Angstrom cúbico
'	- Minuto
''	- Segundo
1-M-5-FC	- 1-Metil-5-fluorocitosina
1-MC	- 1-Metilcitosina
9-EG	- 9-Etilguanina
a	- Ano
a_n	- n-ésimo coeficiente de equação de regressão linear múltipla
A	- Alanina; Adenosina
AB	- Atividade biológica
Abs	- Valor absoluto
ADK	- Adenil cinase
ADP	- Difosfato de adenosina
Asp	- Ácido Aspártico

Ala	- Alanina
AM1	- <i>Austin Model 1</i>
AMP	- Monofosfato de adenosina
Arg	- Arginina
Asn	- Asparagina
atm	- Atmosfera
ATP	- Trifosfato de adenosina
b	- Intercepto de equação de regressão linear múltipla
B3LYP	- Becke3-Lee-Yang-Parr
C	- Cisteína; Citosina; Concentração molar
C α	- Carbono alfa
Calc.	- Calculado
CHARMM	- <i>Chemistry at Harvard Macromolecular Mechanics</i>
CMV	- Citomegalovírus
Comp.	- Composto
CPU	- <i>Central processor unit</i>
Cys	- Cisteína
CYT	- Cisteína proteinase
d	- Dia
dTMP	- 2'-deoxitimidina
D	- Ácido aspártico
DFR	- Diidrofolato redutase
DNA	- Ácido desoxiribonucleico
F _n ^(e)	- Densidade eletrofílica de fronteira do n-ésimo átomo
F _n ⁽ⁿ⁾	- Densidade nucleofílica de fronteira do n-ésimo átomo
E	- Ácido glutâmico
EBV	- Vírus Epstein-Barr
Elem.	- Elemento
E _s	- Constante estérica de Taft
Exp.	- Experimental
F	- Fenilalanina; Teste de Fischer
F _m	- Constante eletrônica de Swain-Lupton para efeito de campo de substituintes na posição <i>meta</i>

FTT	- <i>Fitness criterion</i>
gB	- Glicoproteína B
G	- Glicina; Guanosina; função gaussiana
G.H ₂ O	- Guanina monohidratada
GAMESS	- <i>General atomic and molecular electronic structure system</i>
Gln	- Glutamina
Glu	- Ácido Glutâmico
Gly	- Glicina
h	- Hora
H	- Histidina
HF	- Homologia funcional; Hartree-Fock
HHV n	- Herpesvírus humano do tipo n
His	- Histidina
HOMO	- <i>Highest occupied molecular orbital</i>
HSV	- Vírus herpes simples
HSV TK	- Timidina-cinase do vírus herpes simples
HSV1	- Vírus herpes simples do tipo 1
HSV2	- Vírus herpes simples do tipo 2
HSV1 TK	- Timidina-cinase do vírus herpes simples do tipo 1
HSV2 TK	- Timidina-cinase do vírus herpes simples do tipo 2
HT	- Homologia total
I	- Isoleucina
IC ₅₀	- Concentração molar de substância capaz de provocar 50% de inibição enzimática
Ile	- Isoleucina
IV	- Intravenoso
k	- Número de variáveis independentes na regressão linear múltipla
kDa	- quiloDalton
K	- Lisina, Kelvin
KITH	- <i>Thymidine kinase</i> (utilizado pelo SeqNet)
L	- Leucina
Leu	- Leucina
Log P	- Logaritmo na base decimal do coeficiente de partição octanol-água
LOO	- <i>Leave-one-out</i>

LUMO	- <i>Lowest unoccupied molecular orbital</i>
Lys	- Lisina
m	- Minuto
M	- Metionina
Mb	- <i>Megabyte</i>
MCP	- <i>Mast cell protease</i>
MDM	- <i>Multal data matrix</i>
Met	- Metionina
MFLOPS	- <i>Millions of floating point operations per second</i>
MHz	- MegaHertz
MIPS	- <i>Millions of informations per second</i>
MM2	- <i>Molecular mechanics 2</i>
MNDO	- <i>Modified neglect of diatomic overlap</i>
MOPAC	- <i>Molecular orbital package</i>
MPn	- Møller-Plesset de n-ésima ordem
MRA	- <i>Multiple Regression Analysis</i>
mRNA	- Ácido ribonucleico mensageiro
n	- Número de compostos na regressão linear múltipla
nf	- Não fornecido
nm	- nanômetro
N	- Asparagina
No. Res.	- Número de resíduos
O	- Oral
Oft	- Oftálmico
OVO	- <i>Ovomucoid</i>
p	- Nível de significância estatística de equação de regressão linear múltipla;
pressão	
P	- Prolina; Coeficiente de partição octanol-água
P_H	- Coeficiente de partição octanol-água do composto não-substituído
P_X	- Coeficiente de partição octanol-água do composto substituído pelo grupo X
PCM	- <i>Polarisable continuum method</i>
PDB	- <i>Protein data bank</i>
Phe	- Fenilalanina
PM3	- <i>Parameterized method 3</i>

PRESS	- <i>Predictive residual sum of squares</i>
Pro	- Prolina
Q	- Glutamina; coeficiente de correlação das previsões
Q_n	- Carga atômica líquida do n-ésimo átomo na molécula
QSAR	- <i>Quantitative structure-activity relationships</i>
r	- Coeficiente de correlação
R	- Arginina
R_p	- Constante eletrônica de Swain-Lupton para efeito de campo de substituintes na posição <i>para</i>
RAM	- <i>Random access memory</i>
Res.	- Resíduo
R_m_{meta}	- Refratividade molar de substituintes na posição <i>meta</i>
R_m_{para}	- Refratividade molar de substituintes na posição <i>para</i>
RMS	- <i>Root mean square</i>
RNA	- Ácido ribonucleico
s	- Desvio-padrão
S	- Serina
SCRF	- <i>Self-consistent reaction field</i>
SDEP	- <i>Standard error of predictions</i>
Ser	- Serina
s_{PRESS}	- Desvio-padrão das previsões
STO	- <i>Slater-type orbital</i>
T	- Treonina; uso tópico; temperatura absoluta
Thr	- Treonina
TK	- Timidina-cinase
TMD	- Timidina monofosfato
Trp	- Triptofano
TTP	- Timidina trifosfato
Tyr	- Tirosina
uma	- Unidade de massa atômica
V	- Valina
Val	- Valina
VHS	- <i>Virion host shut-off</i>
V_{vdw}	- Volume molecular de van der Waals

VZV	- Vírus varicela-zóster
W	- Triptofano
X	- Grupo químico genérico; Resíduo protéico genérico
X-n	- n-ésima variável de equação de regressão linear múltipla
Y	- Tirosina

CAPÍTULO 1

INTRODUÇÃO

Os vírus são importantes agentes infecciosos, são parasitas intracelulares obrigatórios e estão largamente distribuídos entre as espécies vivas. Os vírus já foram descritos como agentes filtráveis, devido ao fato de atravessarem filtros projetados para reter bactérias. Os maiores vírus conhecidos, poxvírus e paramyxovírus, ambos com cerca de 300 nm de diâmetro, têm cerca de 20% do tamanho do bacilo *Escherichia coli*. Os menores exemplares, parvovírus e picornavírus, medem cerca de 25 nm de diâmetro (Murray et al, 1992). Os vírus podem ser classificados de acordo com sua morfologia, tipo de genoma e forma de multiplicação. De acordo com o tipo de genoma, a classificação mais comum, os vírus podem ser divididos em dois grupos: os que apresentam cerne constituído de RNA e os que apresentam cerne de DNA (ver Quadro 1).

Os nomes dos vírus muitas vezes estão associados às suas características, às doenças que provocam, ao tecido afetado, ao local geográfico onde foram inicialmente identificados, etc. Dessa forma, picornavírus (*pico* = pequeno; *rna* = RNA; *virus*) e togavírus (*toga* = manto, referindo-se à membrana que circunda o vírus) descrevem a estrutura do vírus; papovavírus descreve os membros de sua família (*papiloma*, *polioma* e *virus vacuolantes*); retrovírus (*retro* = reverso) refere-se à síntese de DNA a partir do RNA viral; adenovírus (*adenóides*) e reovírus (*respiratório*, *entérico*, *órfão*) referem-se ao local do corpo onde foram isolados pela primeira vez. O reovírus foi designado de vírus órfão devido ao fato de ter sido descoberto antes de ser associado a alguma doença específica. Coxsackievírus recebeu seu nome em homenagem a N.Y. Coxsackie (Murray et al, 1992).

Quadro 1 - Classificação dos vírus quanto à natureza do genoma (Murray et al, 1992)

Genoma	Família	Vírus humanos importantes
RNA	Paramixoviridae	Sarampo, caxumba, sincicial respiratório, parainfluenza, Sendai
	Orthomixoviridae	Influenza A, B e C
	Coronaviridae	Coronavírus
	Arenaviridae	Coriomeningite linfocítica, vírus da febre de Lassa, complexo viral Tacaribe
	Rhabdoviridae	Raiva
	Filoviridae	Vírus de Marburg e Ebola
	Bunyaviridae	Encefalite da Califórnia, febre do flebótomo, febre hemorrágica da Criméia, Congo, etc.
	Retroviridae	Leucemia de células T humanas I e II, imunodeficiência humana
	Reoviridae	Rotavírus, reovírus, febre do carrapato da Califórnia
	Picomaviridae	Rhinovírus, poliovírus, ECHO vírus, Coxsackievírus
	Togaviridae	Rubéola; encefalites eqüinas ocidental, oriental e da Venezuela, Sindbis, Semliki, Forest, etc.
	Flaviviridae	Febre amarela, dengue, encefalite de St. Louis
	Calciviridae	Agente de Norwalk
DNA	Parvoviridae	Parvovírus B19
	Papovaviridae	Papovavírus (JC, BK, SV40), papilomavírus
	Hepadnaviridae	Vírus hepatite B
	Adenoviridae	Adenovírus
	Herpesviridae	Herpes simples tipos 1 e 2, vírus varicela-zoster, vírus Epstein-Barr, citomegalovírus, vírus herpes 6 e 7
	Poxviridae	Vírus da varíola, vírus da vacínia, vírus do molusco contagioso

1.1 HERPESVÍRUS

Os herpesvírus formam grande e importante grupo de agentes infectantes e afetam diversas espécies, tanto homeotermas quanto pecilotermas. Cerca de 100 espécies de herpesvírus já foram pelo menos parcialmente caracterizadas, sendo que sete dessas espécies afetam o homem: vírus herpes simples tipo 1, vírus herpes simples tipo 2, vírus varicela-zóster, citomegalovírus, vírus herpes tipos 6 e 7 e vírus Epstein-Barr (Roizman, 1991; Levy et al, 1994). Essas espécies de vírus são de grande importância em clínica médica. Em seres humanos os vírus herpes causam diversas doenças, variando desde

simples lesões mucocutâneas até infecções fatais, além de estarem relacionados a alguns tipos de cânceres.

1.1.1 Classificação

A família *Herpesviridae* é dividida em três subfamílias: *Alphaherpesvirinae*, *Betaherpesvirinae* e *Gammapherpesvirinae*. A divisão da família *Herpesviridae* em três subfamílias é o resultado do trabalho do Grupo de Estudo de Vírus Herpes do Comitê Internacional de Taxonomia de Vírus (Roizman et al, 1981). A classificação em termos de subfamílias tem o objetivo de estabelecer relações evolucionárias entre as espécies de vírus herpes conhecidas e também de facilitar a previsão de propriedades e a identificação de novas espécies (Roizman, 1991). A subfamília *Alphaherpesvirinae* caracteriza-se pela variedade de hospedeiros capazes de abrigá-la, pelo ciclo reprodutivo relativamente curto, pelo rápido crescimento em cultura, pela eficiente destruição das células infectadas e pela capacidade de estabelecer infecções latentes em gânglios sensoriais. Os membros da subfamília *Betaherpesvirinae* possuem a característica não exclusiva do restrito número de hospedeiros naturais. O ciclo reprodutivo é longo e a infecção progride lentamente em cultura. As células infectadas tornam-se freqüentemente inchadas (citomegalia). O vírus pode ser mantido na forma latente em glândulas secretórias, células linforeticulares, rins e outros tecidos. As infecções causadas por vírus da família *Gammapherpesvirinae* são específicas de linfócitos T ou B. Vírus na forma latente são freqüentemente demonstráveis em tecidos linfóides (Roizman, 1991). O Quadro 2 mostra os sete tipos de vírus herpes que afetam o homem e as principais doenças associadas. Diversas imagens de vírus herpes podem ser obtidas no endereço http://www.tulane.edu/~dmsander/Big_Virology/BVDNAherpes.html.

Quadro 2 - Vírus herpes que utilizam o homem como hospedeiro e suas principais consequências clínicas (Collier e Oxford, 1993).

Família: Herpesviridae			
Subfamília	Vírus	Símbolo	Doenças associadas
Alpha-herpesvirinae	Vírus herpes simples Tipo 1	HSV1 ^(a)	Vesículas ulcerativas nos lábios (herpes labial recorrente), língua e gengiva (gengivostomatite); infecção nos olhos (ceratoconjuntivite) e cérebro (meningite e encefalite).
	Vírus herpes simples Tipo 2	HSV2	Vesículas ulcerativas na região genital e perianal (herpes genital recorrente) e infecção generalizada em recém-nascidos (herpes neonatal).
	Vírus varicela-zóster	VZV	Vesículas ulcerativas no tronco e membros (varicela ou catapora) e infecção no cérebro.
Beta-herpesvirinae	Vírus citomegálico	CMV	Mononucleose infecciosa e infecção nos olhos, rins e cérebro.
	Vírus herpes humano Tipo 6	HHV-6	Erupções cutâneas.
	Vírus herpes humano Tipo 7	HHV-7	Nenhuma até o momento.
Gamma-herpesvirinae	Vírus Epstein-Barr	EBV	Mononucleose infecciosa, linfoma de Burkitt e carcinoma nasofaríngeo.

^a HSV1 = *herpes simplex virus type 1*. Utilizou-se a abreviação em língua inglesa para facilitar a comparação de sequências de aminoácidos adiante. As sequências que serão comparadas no Capítulo 5 são denominadas KITH_HSV11, KITH_HSV23, etc. Essas denominações devem ser mantidas pois correspondem ao código identificador de cada sequência, utilizado para localizá-las no banco de dados de sequências (SwissProt). Dessa forma, utilizando-se a abreviação de *timidina-cinase do vírus herpes simples tipo 1* como HSV1 TK e não como TC VHS1, pretende-se evitar alguma confusão ao longo deste trabalho.

1.1.2 Morfologia

A unidade viral completa é denominada vírion (ver Figura 1). Estes são constituídos por nucleocapsídeo envolvido em membrana lipoprotéica, também chamada de envelope, cuja origem é a membrana nuclear da célula hospedeira. O nucleocapsídeo, por sua vez, é formado pelo capsídeo e pelo núcleo contendo a molécula de DNA. O capsídeo é icosaédrico e constituído de 162 capsômeros tubulares envolvendo o núcleo de DNA (ver Figura 2). O diâmetro do nucleocapsídeo é cerca de 100 nm, enquanto que o vírion mede entre 120 e 200 nm (Collier e Oxford, 1993). O espaço localizado entre o

envelope e o nucleocapsídeo chama-se tegumento. O tegumento é o local onde ficam armazenadas diversas enzimas virais que serão liberadas quando da entrada do vírion na célula hospedeira.

1.1.3 Estrutura do genoma

Todas as viroses herpéticas possuem DNA como material genético. O DNA é linear e em dupla hélice. Seu peso molecular varia de 85×10^6 , para o caso do vírus varicela-zóster, a 160×10^6 , para o vírus citomegálico.

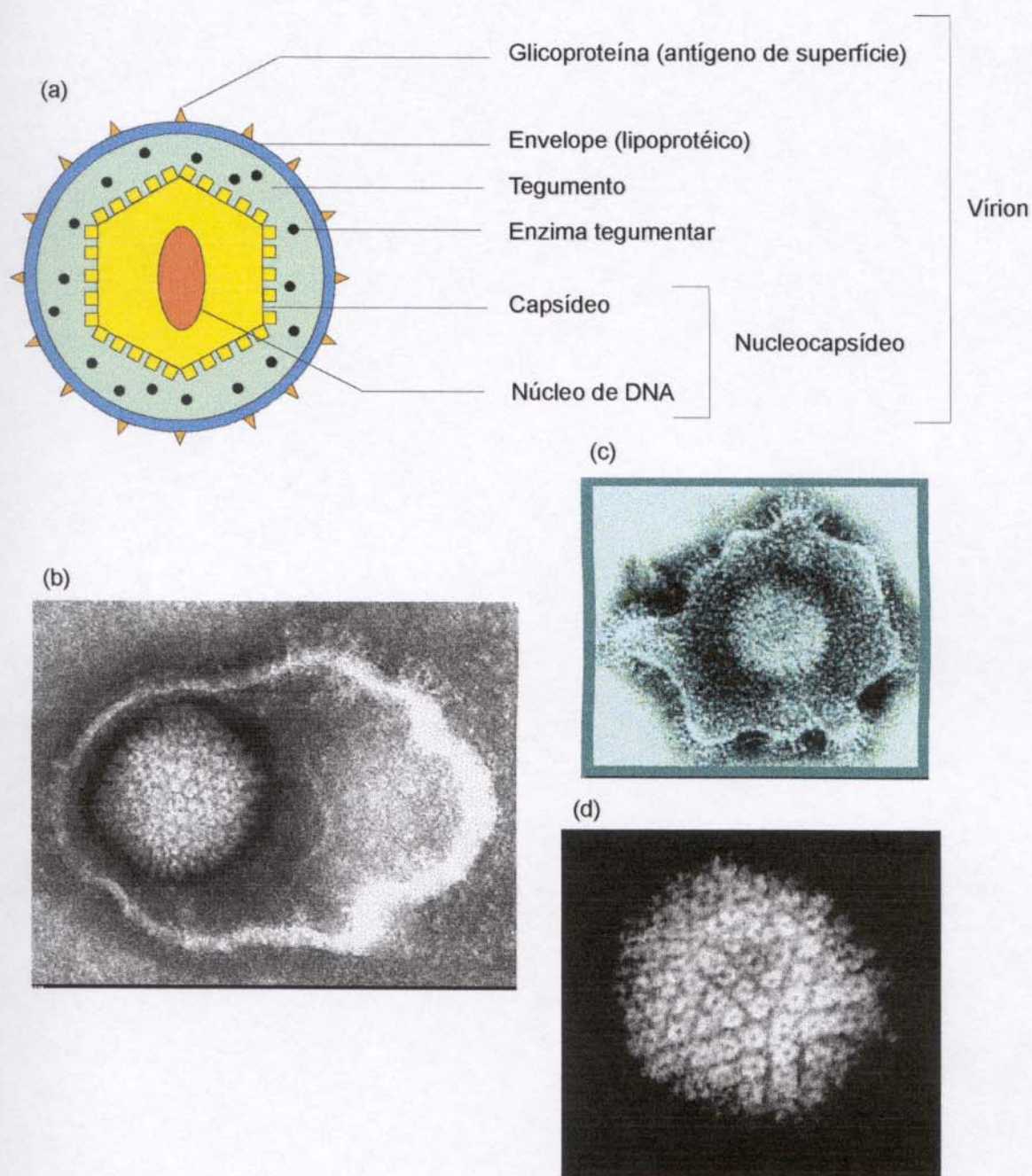


Figura 1 - (a) Representação esquemática da estrutura do vírus herpes simples; (b) Micrografia eletrônica do vírus herpes simples. Os herpesvírus possuem envelope envolvendo o capsídeo icosaédrico de aproximadamente 100 nm de diâmetro, que contém a molécula de DNA. (c) Micrografia eletrônica do vírus herpes simples. Glicoproteínas codificadas pelo HSV são incorporadas no envelope viral e são visíveis como picos localizados na superfície do vírion. Nesta micrografia, a glicoproteína B (gB) é vista como picos agrupados na superfície externa do envelope, tendo aproximadamente 10 nm de comprimento; (d) Micrografia eletrônica do capsídeo do HSV. As micrografias eletrônicas foram obtidas no endereço http://www.tulane.edu/~dmsander/Big_Virology/BVDNAherpes.html.

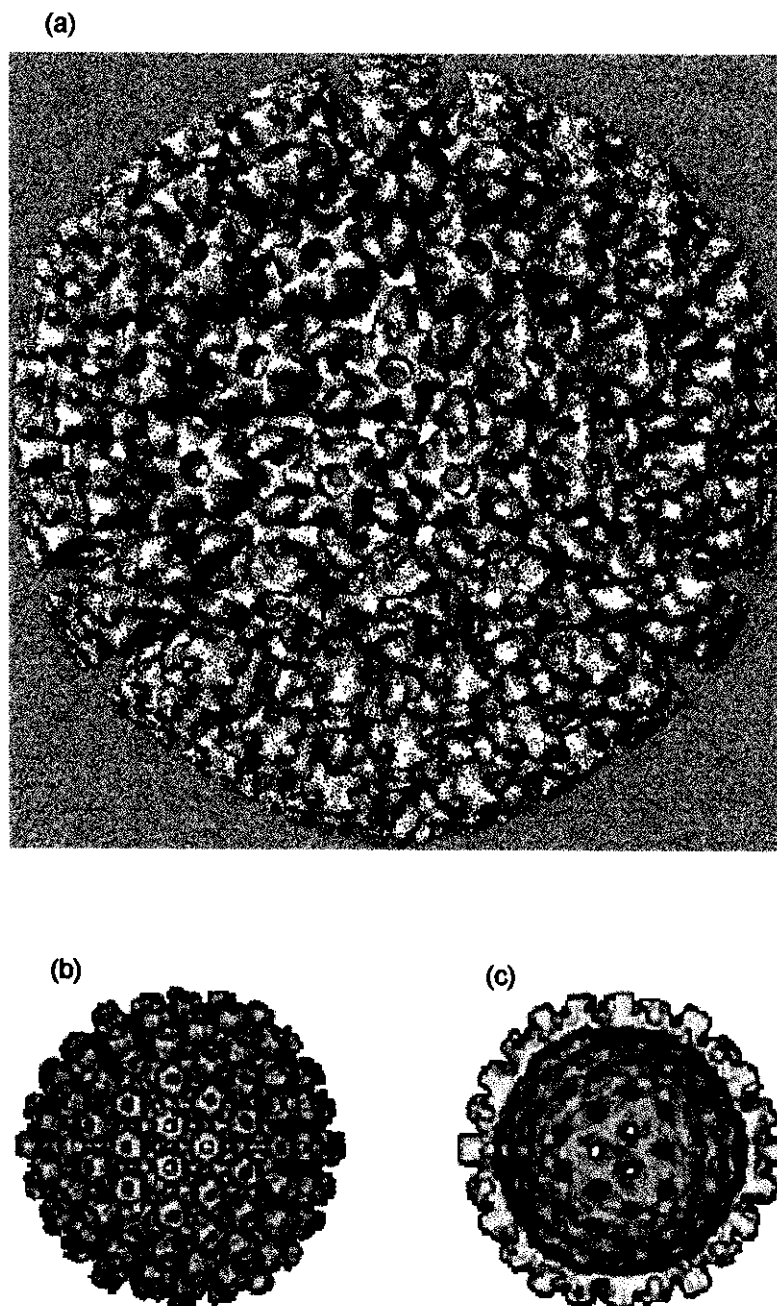


Figura 2 - Imagens computacionais do capsídeo do HSV construídas a partir de informações obtidas através de experimentos de criomicroscopia eletrônica. (a) A unidade estrutural do capsídeo é o capsômero, que possui estrutura protéica. No HSV, o capsídeo é formado por 162 capsômeros, que no conjunto apresentam o formato icosaédrico (imagem obtida no endereço <http://www.bocklabs.wisc.edu/images/herpescapsid.GIF>); (b) Capsídeo do HSV1; (c) Corte do capsídeo de HSV1 (imagens (b) e (c) obtidas em <http://www.med.virginia.edu/ed-programs/biophysics/attrib.html>. A animação da construção do capsídeo pode ser vista em <http://rsb.info.nih.gov/NIH-Image/Java/Rotator/hsv.html>).

1.1.4 Polipeptídios

O genoma do vírus herpes é capaz de produzir cerca de 100 polipeptídios, sendo que muitos desses são enzimas, incluindo a DNA-polimerase que é essencial à replicação viral. HSV1, HSV2 e o vírus varicela-zóster são capazes de produzir a enzima fosforilativa timidina-cinase, que é essencial para a ativação de certas drogas antivirais (Collier e Oxford, 1993). Pelo menos 33 proteínas já foram detectadas em vírions de herpesvirus. O peso molecular dessas proteínas varia desde 15 a 275 kDa. Oito dessas proteínas encontram-se na superfície do vírion e são glicosiladas (ver Figura 1). Essas glicoproteínas superficiais são importantes durante a fase de penetração na célula hospedeira. Muitas outras proteínas são sulfatadas e pelo menos seis são fosforiladas (Levy et al, 1994).

1.1.5 Propriedades biológicas

As viroses herpéticas conhecidas parecem compartilhar quatro significantes propriedades biológicas (Roizman, 1991): (a) todas as viroses herpéticas produzem grande conjunto de enzimas envolvidas no metabolismo dos ácidos nucleicos (timidina-cinase, timidilato-sintetase, ribonucleotídeo-redutase, etc.), síntese de DNA (DNA-polimerase, helicase, etc.), e, possivelmente, processamento de proteínas (proteína-cinase), embora o número exato de enzimas possa variar de uma espécie de herpesvírus para outra; (b) tanto a síntese de DNA quanto a montagem dos capsídeos ocorrem no núcleo. O envelopamento dos capsídeos ocorre quando estes atravessam a membrana nuclear, em seu caminho para fora da célula hospedeira; (c) a produção de prole infecciosa é acompanhada de destruição irreversível da célula infectada; (d) as viroses herpéticas estudadas até o momento são capazes de permanecer em estado latente em seus hospedeiros naturais.

1.2 VÍRUS HERPES SIMPLES

O vírus herpes simples tipo 1 afeta principalmente as partes superiores do corpo, especialmente a boca (herpes labial recorrente e gengivites), os olhos (ceratites e conjuntivites) e o cérebro (meningite e encefalite). A infecção primária por HSV1 ocorre ainda na infância através do beijo ou carinho de adultos infectados. O vírus herpes simples tipo 2 é frequentemente transmitido através do ato sexual e é a causa das ulcerações genitais e perianais (herpes genital primário e recorrente) e da infecção generalizada de recém nascidos, transmitida ao bebê pela mãe infectada durante o trabalho de parto (ver Figura 3). Na verdade, a divisão das regiões do corpo humano onde HSV1 e HSV2 costumam infectar (HSV1 nas partes superiores do corpo e HSV2 na região genital e perianal) não deve ser seguida com rigor. Um indivíduo com herpes labial por HSV1 pode provocar infecção genital por HSV1 no parceiro, após prática de sexo oral. Da mesma forma, um indivíduo infectado na região genital por HSV2 pode provocar lesão bucal por HSV2 no parceiro, pela transmissão através de sexo oral. Importante propriedade de todas as viroses herpéticas é a sua habilidade de causar infecções latentes que podem subsequentemente tornarem-se reativadas (Juel-Jensen & MacCallum, 1972; Collier e Oxford, 1993).

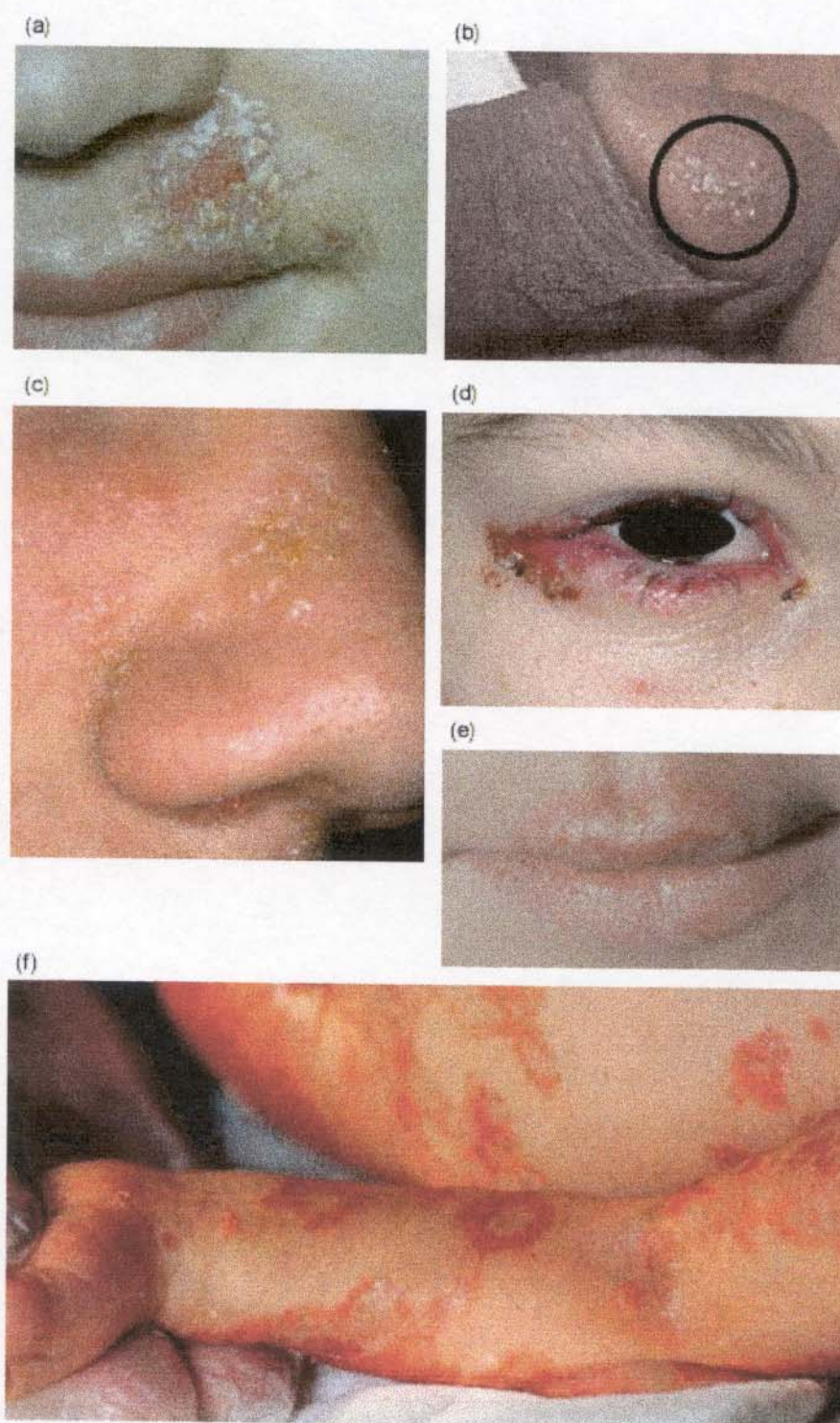


Figura 3 - Algumas das manifestações clínicas do herpes simples. (a) herpes na região entre a boca e o nariz (<http://www.medic.mie-u.ac.ip/derma/bilddb/bilder/cd2/img0020.htm>); (b) herpes genital (<http://med-www.bu.edu/people/sycamore/std/herpes.htm>); (c) herpes nasal (<http://www.medic.mie-u.ac.ip/derma/bilddb/bilder/cd18/img0053.htm>); (d) conjuntivite herpética (<http://www.derma.med.uni-erlangen.de/bilddb/diagnose/englisch/i054900.htm>); (e) herpes labial (<http://www.medic.mie-u.ac.ip/derma/bilddb/bilder/cd18/img0045.htm>); (f) herpes neonatal (<http://tray.dermatology.uiowa.edu/HSVCon01.htm>).

1.2.1 Propriedades biológicas

O ciclo replicativo do vírus herpes simples, de forma similar às outras herpesvíroses, começa com a invasão celular (ver Figura 4). O mecanismo de entrada na célula hospedeira é a fusão do envelope viral com a membrana celular, embora a endocitose também seja possível. As glicoproteínas que revestem o vírion e afloram em sua superfície agem como antígenos de reconhecimento. Admite-se que o sulfato de heparina, um proteoglicano encontrado no exterior de várias células aderentes, seja um dos receptores celulares do vírus herpes simples (Murray et al, 1992). O mecanismo de fusão libera o nucleocapsídeo e as enzimas presentes no tegumento viral diretamente no citoplasma da célula hospedeira. Durante a fusão, duas importantes proteínas presentes no tegumento também são liberadas no citoplasma: uma cinase protéica, VHS (*virion host shut-off*), que interrompe a síntese das proteínas celulares, e uma proteína reguladora transcricional, α -TIF (*α -trans-inducing factor*), que segue para o núcleo, onde coordenará a utilização das enzimas do hospedeiro na transcrição do genoma viral (Boehmer & Lehman, 1997).

Ao chegar à membrana nuclear, o nucleocapsídeo procura um poro e, através deste, libera a molécula de DNA no interior do núcleo celular. Após a penetração no núcleo, o DNA viral imediatamente torna-se circular. A proteína α -TIF induz a transcrição dos genes α e, dessa forma, cinco α -mRNAs são produzidos e transportados para o citoplasma. As α -proteínas virais sintetizadas no citoplasma retornam ao núcleo e auxiliam na segunda etapa de transcrição que resulta na síntese das β -proteínas. Nesse ponto, a cromatina celular é degradada e seus subprodutos são armazenados junto à membrana nuclear. Essa etapa representa uma espécie de limpeza intra-nuclear para que haja espaço para a produção da prole viral.

A seguir ocorre a replicação do DNA viral e a subsequente síntese das γ -proteínas, que são as proteínas estruturais das novas partículas virais. Diversas glicoproteínas produzidas nessa etapa aderem à membrana nuclear. As proteínas do capsídeo são transportadas para o núcleo e, após agrupadas, formam os primeiros capsídeos vazios. As unidades de DNA recém sintetizadas são aspiradas para dentro dos capsídeos e, dessa forma, os nucleocapsídeos-filhos são rapidamente armazenados no núcleo. Finalmente, os nucleocapsídeos atravessam a membrana nuclear em direção ao citoplasma. Nesse processo, partes da membrana nuclear passam a envolver os nucleocapsídeos que recebem assim sua última estrutura, o envelope. A prole de vírions chega ao meio extracelular através do retículo endoplasmático. A saída dos vírions é acompanhada de lise celular.

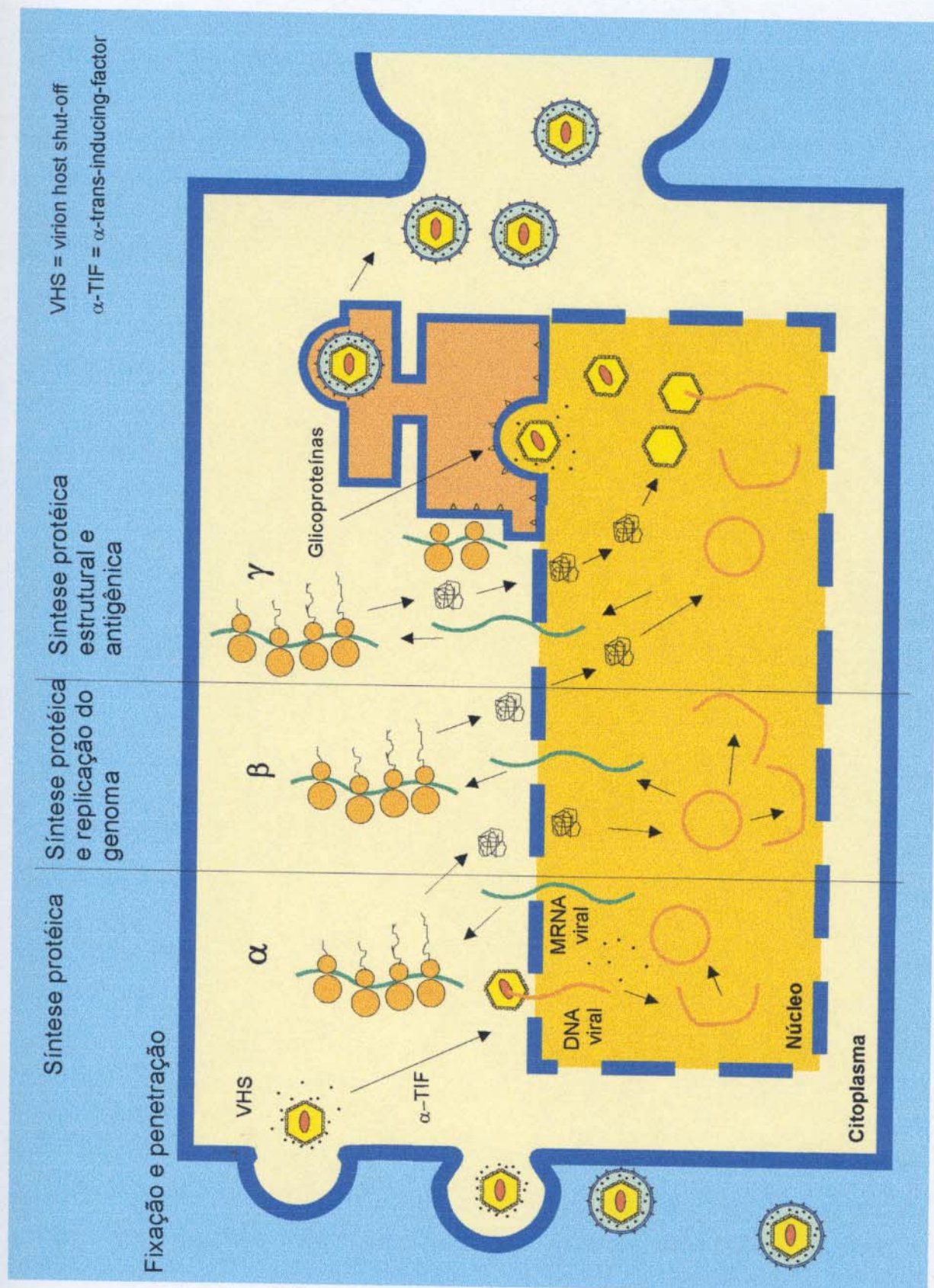


Figura 4 - Replicação viral (ver detalhes no texto).

1.2.2 Patogenia

Cada ciclo infeccioso do vírus herpes simples dura cerca de duas semanas (ver Figura 5). A infecção primária ocorre devido ao contato direto entre o indivíduo sadio e o infectado. O vírus penetra através de soluções de continuidade no epitélio e lesões nas mucosas bucal e genital. A replicação viral ocorre nas células localizadas na base da lesão. Ocorre formação de vesículas e o líquido vesicular contém vírions infecciosos. Poucos dias depois, há formação de úlcera que acaba sarando sem qualquer necessidade de tratamento. Na quase totalidade das vezes a lesão herpética é localizada e desaparece sem deixar vestígios. No entanto, muito cuidado deve ser tomado durante cada recorrência da infecção. Não se deve furar as vesículas devido ao risco de contaminação de outras partes do corpo, especialmente os olhos. As ceratites causadas por herpesvírus em geral têm conseqüências desastrosas, podendo ocasionar úlcera de córnea e, conseqüentemente, perda da visão.

Poucos dias antes do desaparecimento da lesão herpética, os vírions disseminam-se pelos tecidos adjacentes e atingem os neurônios que inervam o tecido afetado. A seguir, migram através de seus axônios até o gânglio mais próximo. No caso de HSV1, o gânglio atingido é o trigêmeo, enquanto que no de HSV2 o gânglio atingido é o lumbosacral. Uma vez no gânglio, estabelece-se uma situação de infecção latente que eventualmente pode ser reativada. Diversos fatores podem ocasionar a reativação da atividade viral. As mais importantes são estresse emocional, trauma, frio, luz solar, irritação gástrica, febre, ciclo menstrual e imunossupressão. Não é conhecido o mecanismo adotado pelo herpesvírus para permanecer em estado de latência. A explicação mais simples é a que considera a possibilidade de o DNA viral tornar-se incorporado ao DNA neuronal (lisogenia). Entretanto, não se sabe explicar como isso poderia ocorrer em células não replicantes como as nervosas. Uma vez reativados, os vírions migram de volta para o tecido

originalmente infectado e o ciclo recomeça. As reativações podem ocorrer em intervalos de semanas ou meses. Esses intervalos vão tornando-se cada vez maiores à medida que o indivíduo envelhece e eventualmente podem cessar as recorrências (Roizman, 1991; Murray et al, 1992; Collier & Orford, 1993).

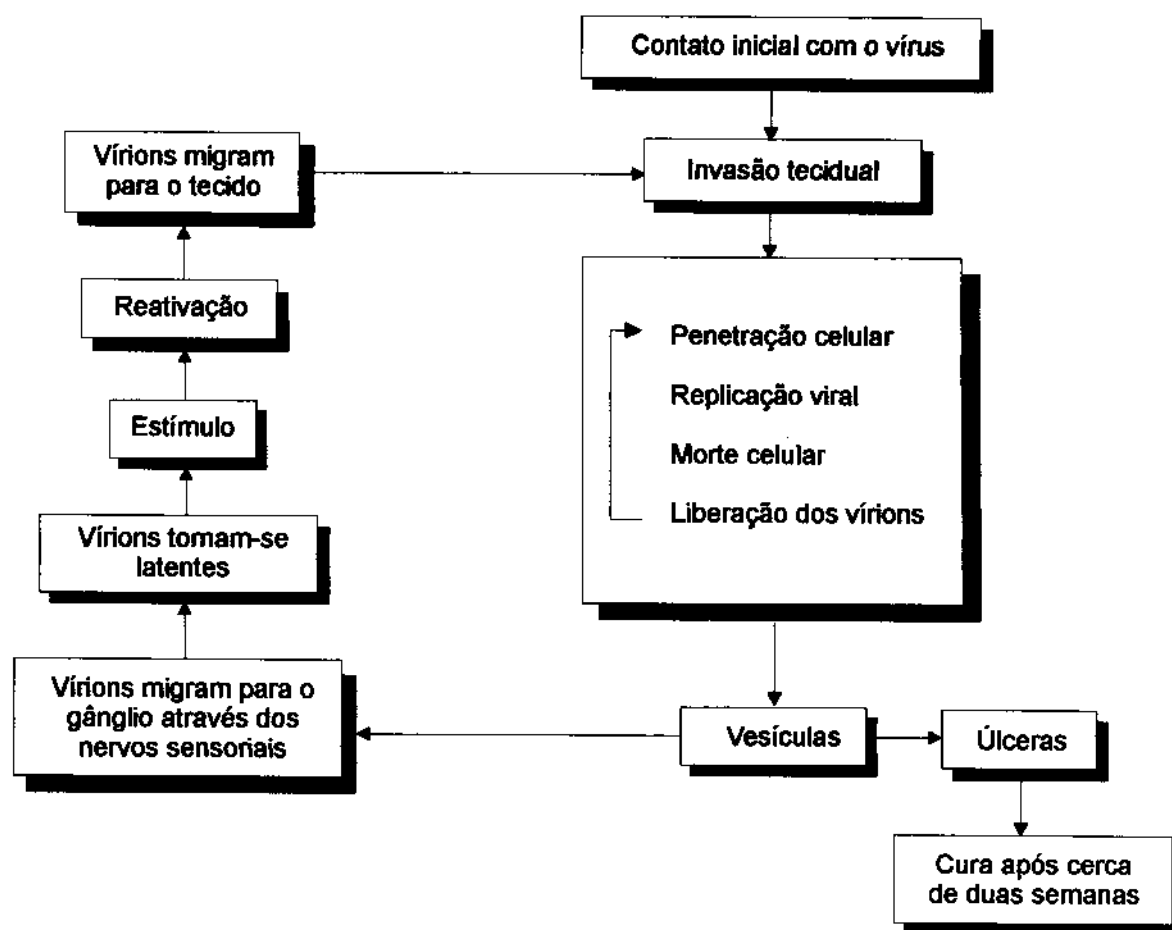


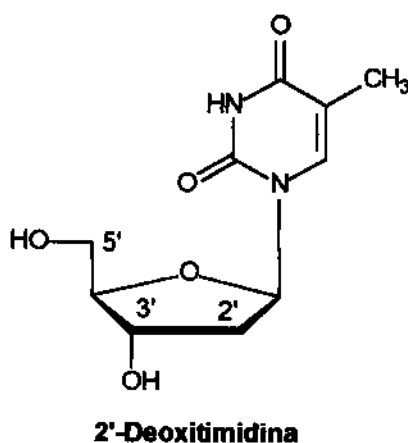
Figura 5 - Esquema do ciclo evolutivo do vírus herpes, ilustrado para o caso de infecções cutâneas.

1.2.3 Combate ao vírus herpes simples

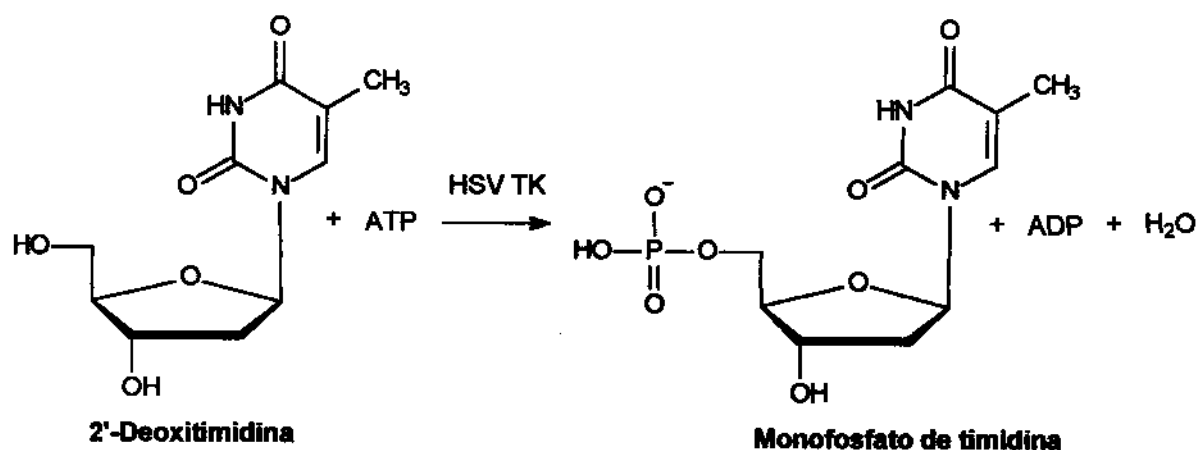
Até o presente momento não há cura para os indivíduos infectados com HSV. Uma vez que o vírus migra para o gânglio, através dos neurônios que enervam a região infectada primariamente, iniciando a fase latente do vírus, o indivíduo sofrerá ataques

reincidentes pelo resto da vida. A periodicidade desses ataques, no entanto, pode ser controlada através do uso contínuo de medicamentos apropriados.

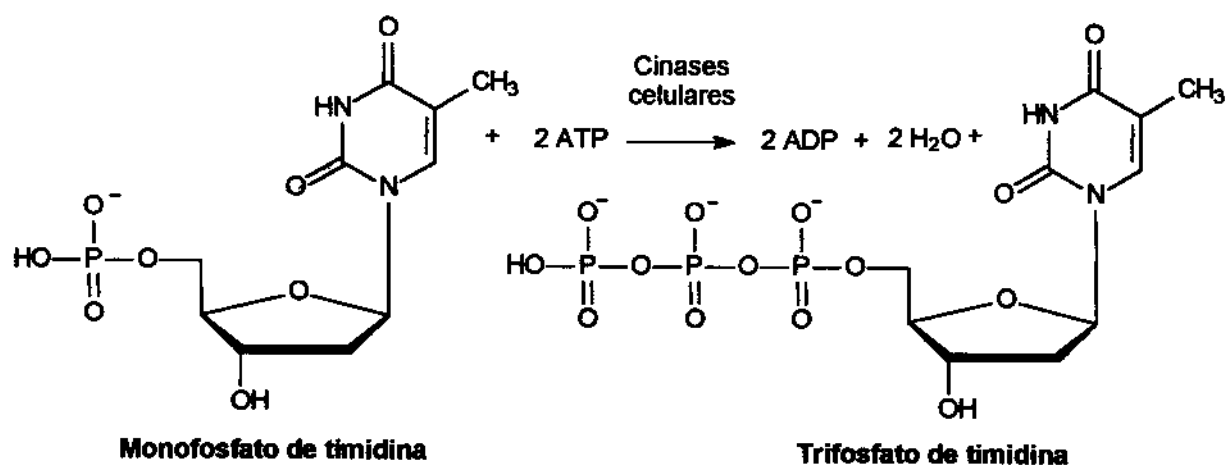
A tática usualmente empregada para combater as infecções por HSV baseia-se na interferência e obstrução de seu ciclo replicativo. HSV1 e HSV2 são capazes de produzir duas enzimas que lhes são essenciais: DNA-polimerase e timidina-cinase (TK). As tentativas mais bem sucedidas no combate ao HSV envolvem o uso de substâncias capazes de inibir simultaneamente TK e DNA-polimerase. Embora o hospedeiro humano também seja capaz de sintetizar essas enzimas, a grande diferença de escala evolutiva entre vírus e homem resultou em grandes diferenças estruturais e catalíticas entre as enzimas codificadas pelo HSV e pelo homem (ver Seção 5.2.8, “Comparação das seqüências de timidina-cinase de HSV1 e humana”, pag.178). Dessa forma, torna-se possível o desenvolvimento de substâncias capazes de inibir a ação das enzimas virais sem que haja considerável inibição das correspondentes enzimas do hospedeiro.



A timidina-cinase é uma enzima chave na rota bioquímica de aproveitamento das bases pirimidínicas (Sanderson et al, 1988). A enzima catalisa a transferência do γ -fosfato do trifosfato de adenosina (ATP) para o grupo 5'-OH da 2'-deoxitimidina (dT), resultando em 5'-monofosfato de timidina (5'-TMP), de acordo com a seguinte equação:



5'-TMP será subseqüentemente fosforilado a 5'-trifosfato de timidina (5'-TTP) por cinases e fosforilases pertencentes ao hospedeiro, de acordo com a equação:



5'-TTP é substrato da DNA-polimerase, e será utilizado como matéria-prima durante a replicação do material genético viral (ver Figura 6).

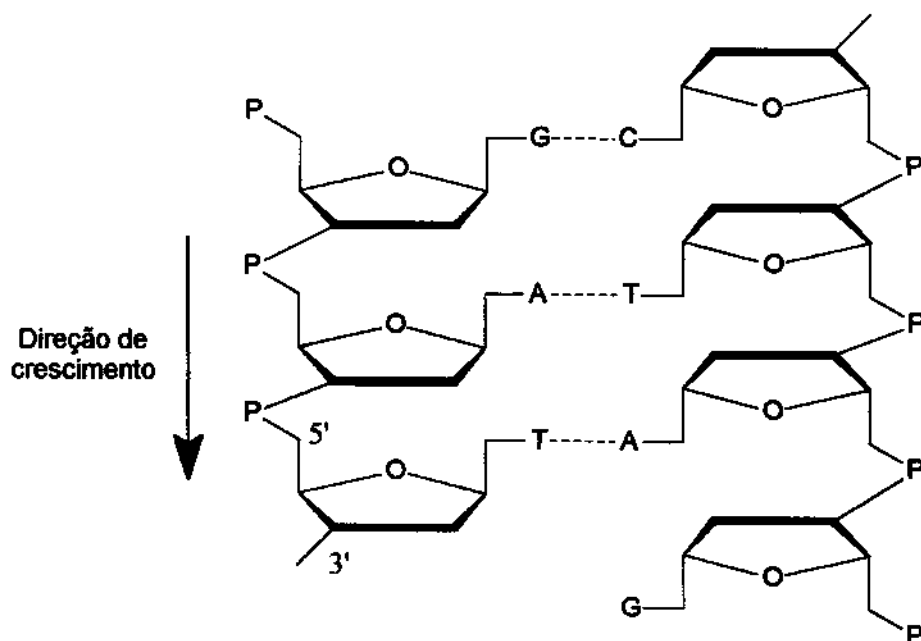
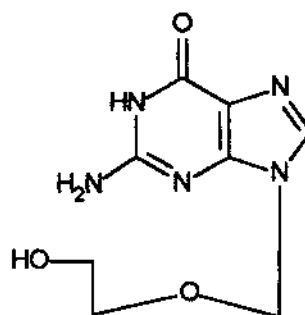


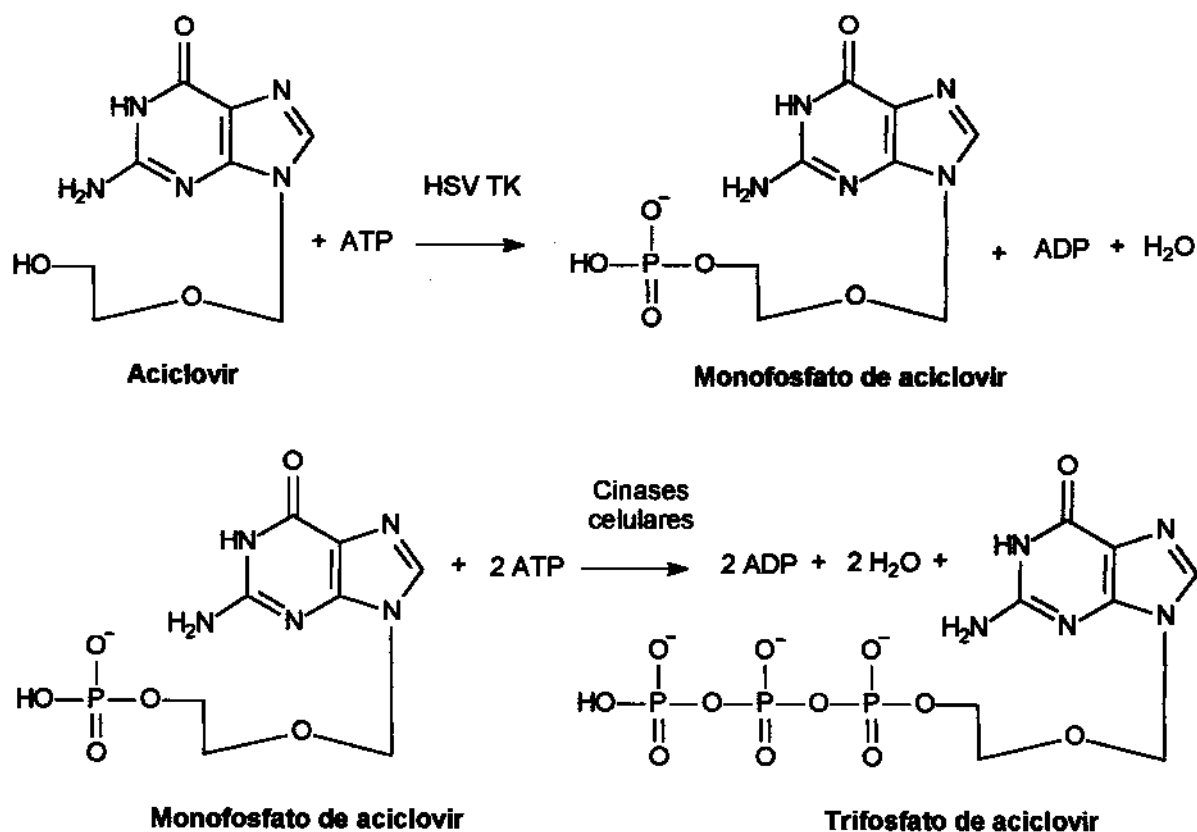
Figura 6 - Replicação da molécula de DNA do HSV. Neste esquema, o crescimento da cadeia dá-se no sentido 5'→ 3'. A DNA-polimerase catalisa o processo de incorporação das bases nitrogenadas (G, C, T, A) na nova cadeia.

O fármaco de escolha no combate às infecções por HSV é o aciclovir, desenvolvido nos laboratórios da Wellcome e comercializado sob o nome de Zovirax[®].



Aciclovir

O aciclovir é inibidor seletivo das enzimas TK e DNA-polimerase, apresentando baixa toxicidade para o homem (Murray et al, 1992). Aciclovir é inibidor competitivo da TK e da DNA-polimerase, sendo fosforilado de forma semelhante à 2'-deoxitimidina, durante o processo de inibição, como pode ser visto nas equações que se seguem:



O trifosfato de aciclovir é substrato da DNA-polimerase viral e, portanto, também será utilizado como matéria-prima durante a replicação do material genético viral. No entanto, devido ao fato de o aciclovir não possuir hidroxila equivalente à 3'-OH da 2'-deoxitimidina, o crescimento da cadeia do DNA replicante será interrompido (Coen, 1996) (ver Figura 7). Por esse motivo, o aciclovir pertence à categoria de agentes *inibidores da síntese de DNA*.

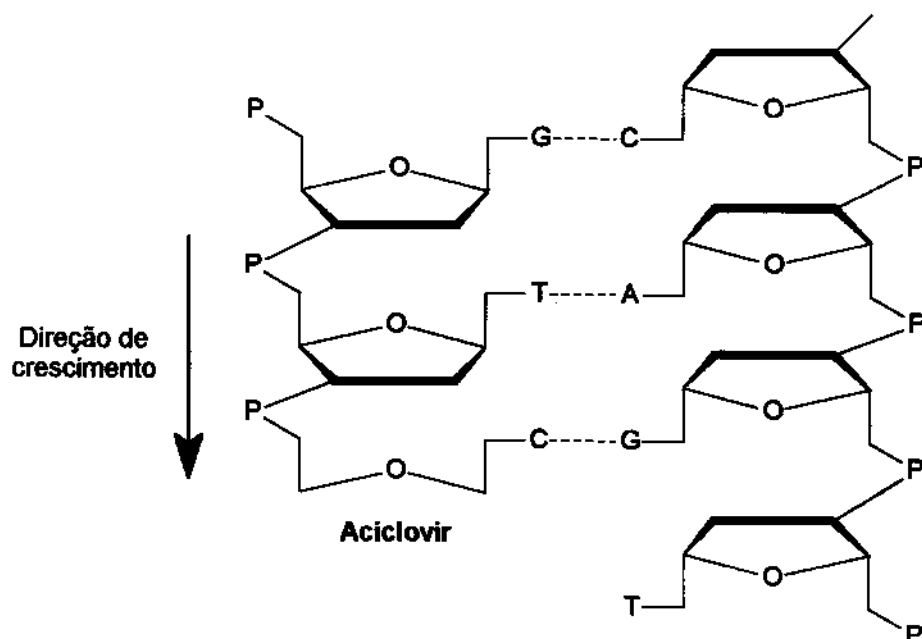


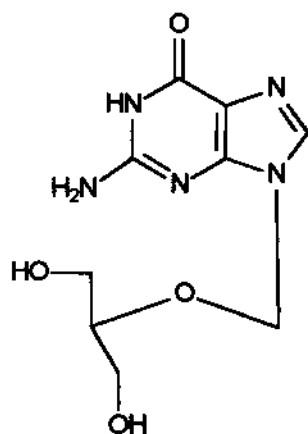
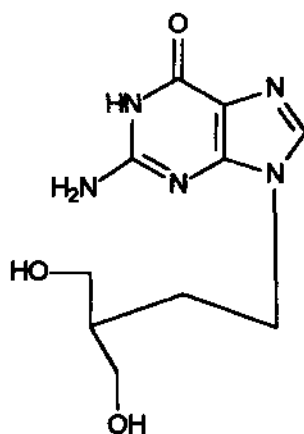
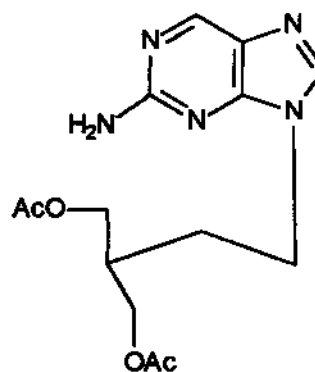
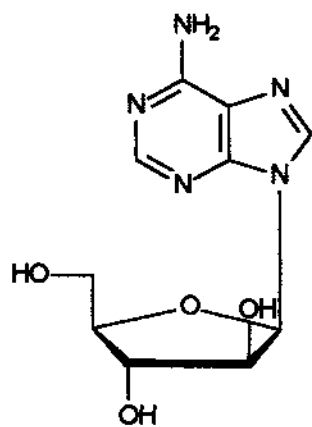
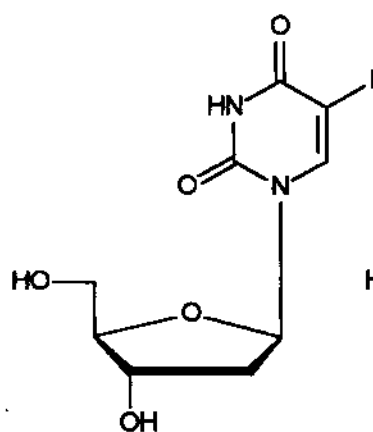
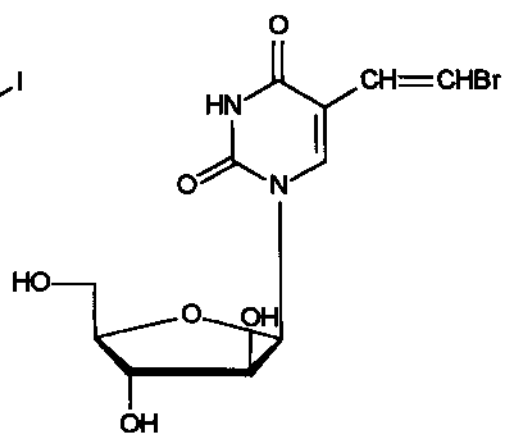
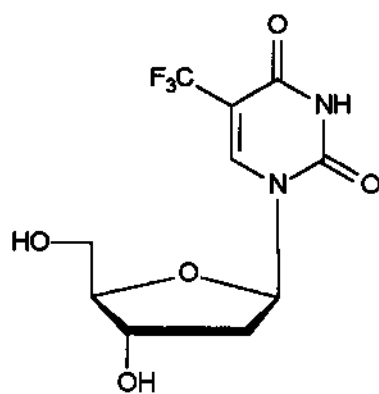
Figura 7 - Replicação da molécula de DNA do HSV. Neste esquema, o crescimento da cadeia dá-se no sentido 5'→ 3'. A inclusão da molécula de aciclovir na cadeia de DNA provoca a interrupção do crescimento da cadeia devido à ausência do grupo 3'-OH.

Quadro 3 - Principais agentes anti-herpes (Hayden, 1996, Murray et al, 1992)

Nome genérico	Nome comercial	Mecanismo de ação	Aplicação ^a
Aciclovir	Zovirax	Inibidor de TK e DNA-polimerase.	IV, O, T, Oft
Famciclovir	Famvir	Inibidor de TK e DNA-polimerase.	O
Foscarnet	Foscavir	Inibidor de DNA-polimerase.	IV, O
Ganciclovir	Cytovene	Inibidor de TK e DNA-polimerase.	IV, O
Idoxuridina	Herpes, Stoxil, Dendrid	Inibidor de DNA-polimerase.	Oft
Penciclovir	–	Inibidor de TK e DNA-polimerase.	–
Sorivudina	–	Inibidor de TK e DNA-polimerase.	–
Trifluridina	Viroptic	Inibidor de DNA-polimerase.	Oft
Valaciclovir	Valtrex	Inibidor de TK e DNA-polimerase.	O
Vidarabina	Vira-A	Inibidor de DNA-polimerase.	IV, Oft

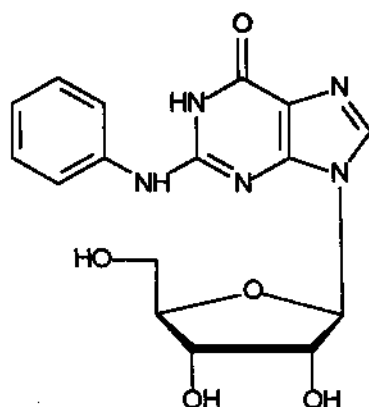
^a IV, intra venoso; O, oral; T, tópico; Oft, oftálmico.

Além do aciclovir, outras substâncias têm sido desenvolvidas como alternativas no combate ao HSV. O Quadro 3 mostra os principais compostos utilizados atualmente no combate ao herpes simples.

**Ganciclovir****Penciclovir****Famciclovir****Vidarabina****Idoxuridina****Sorivudina****Trifluridina**

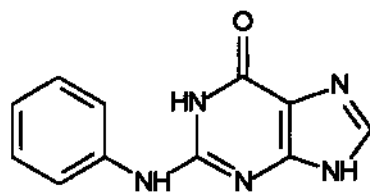
1.2.4 Derivados da N^2 -fenil-(substituído)-guanina

Wright e colaboradores investigaram nova série de inibidores de HSV TK derivados da N^2 -fenilguanossina (Focher et al, 1988). Vários derivados da N^2 -fenilguanossina, substituídos nas posições meta e para do anel fenílico, foram sintetizados e testados em relação à sua capacidade de inibir a fosforilação da 2'-deoxitimidina.



N^2 -Fenilguanossina

As N^2 -fenilguanossinas exercem inibição do tipo competitiva com a 2'-deoxitimidina pelo sítio ativo de TK, sendo, porém, resistentes à fosforilação (Focher et al, 1988). Isso significa dizer que as N^2 -fenilguanossinas não podem atuar como inibidores da DNA-polimerase. Essa observação reduz a potencialidade das N^2 -fenilguanossinas como possíveis agentes anti-herpes para utilização em clínica médica. No entanto, o objetivo do trabalho de Wright e colaboradores não foi propriamente o desenvolvimento de substâncias a serem adotadas em clínica médica, mas sim a execução de estudo que permitisse a ampliação do conhecimento do mecanismo de inibição de TK, no plano molecular. Embora as N^2 -fenilguanossinas tenham se mostrado mais ativas, os autores optaram por executar os estudos de relações estrutura-atividade com derivados da N^2 -fenilguanina, devido às dificuldades de síntese daquelas.

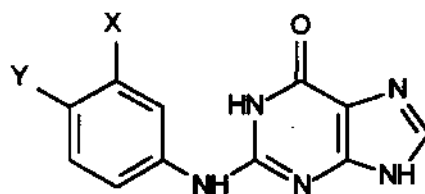


N²-Fenilguanina

Wright e colaboradores realizaram dois trabalhos de relações estrutura-atividade, sendo que um em nível qualitativo (SAR - *Structure-Activity Relationships*) (Hildebrand et al, 1990) e outro em nível quantitativo (QSAR - *Quantitative Structure-Activity Relationships*) (Gambino et al, 1992). Este envolveu 36 derivados da N²-fenilguanina, substituídos nas posições meta e para do anel fenila (ver Tabela 1). O objetivo desses trabalhos foi verificar a influência dos fatores lipofílico, eletrônico e estérico dos substituintes do anel fenila sobre a atividade inibitória dos derivados da N²-fenilguanina. O conhecimento da influência desses fatores podem auxiliar a elaboração de hipóteses sobre o mecanismo de interação dos derivados da N²-fenilguanina com o sítio ativo de TK, bem como sobre a própria configuração do sítio ativo. Resumidamente pode-se dizer que não foram obtidos coeficientes de correlação acima de 0,9, nas equações que correlacionam estrutura química e atividade biológica, e que o fator eletrônico não foi satisfatoriamente parametrizado. Foram também verificadas diferenças acentuadas na influência dos fatores lipofílico e eletrônico sobre a atividade das N²-fenilguaninas sobre HSV1 TK e HSV2 TK. Essas observações abrem espaço para investigação subsequente, em que podem ser utilizadas técnicas diferenciadas para tentar esclarecer pouco mais o mecanismo de ação desses compostos no plano molecular.

Recentemente o grupo do Dr. G.E. Wright realizou os primeiros testes *in vivo* de alguns derivados da N²-fenilguanina (Xu et al, 1995, Gebhardt et al, 1996; Kaufman et al, 1996), sendo que dois dos derivados testados (composto 1, 3'-CF₃, 4'-H, e composto 15, 3'-H, 4'-H) estão presentes na Tabela 1.

Tabela 1 - Estrutura química e atividade inibitória dos derivados da N^2 -fenilguanina, contra HSV1 TK e HSV2 TK. IC_{50} é a concentração molar necessária para provocar 50% de inibição da atividade enzimática (Gambino et al, 1992).



No	X	Y	Log 1/ IC_{50}		No	X	Y	Log 1/ IC_{50}	
			HSV1	HSV2				HSV1	HSV2
1	CF ₃	H	6,82	7,00	19	CF ₃	Cl	4,70	5,30
2	CH ₂ N ₃	H	6,10	5,82	20	CH ₂ OH	H	4,60	4,70
3	H	Br	6,00	6,22	21	CH ₂ NH ₂	H	4,52	4,70
4	CH ₂ Br	H	6,00	5,92	22	H	CF ₃	4,52	4,40
5	I	H	5,96	5,96	23	H	n-Pr	4,52	4,33
6	CH ₂ Cl	H	5,85	5,87	24	-CH=CHCH=CH-		4,40	5,00
7	Cl	H	5,82	6,46	25	Me	n-Bu	4,40	^a
8	Br	H	5,82	6,22	26	H	I	4,40	5,05
9	Et	H	5,60	5,40	27	H	Me	4,30	5,22
10	Br	Br	5,52	5,89	28	F	F	4,30	5,66
11	n-Pr	H	5,48	5,52	29	H	n-Bu	4,30	4,30
12	CN	H	5,30	6,40	30	Cl	Cl	4,00	4,52
13	Et	Me	5,16	5,10	31	H	Cl	4,00	4,30
14	Cl	F	5,10	5,60	32	H	OH	3,85	4,70
15	H	H	5,10	5,80	33	H	i-Pr	3,70	4,00
16	Cl	Me	4,92	5,60	34	CH ₂ OH	Me	3,57	3,70
17	-CH ₂ CH ₂ -CH ₂ -		4,82	5,00	35	H	F	3,52	3,59
18	H	Et	4,70	4,60	36	H	n-C ₅ H ₁₁	3,30	3,60

^a Valor não medido.

CAPÍTULO 2

OBJETIVOS E MÉTODOS

2.1 OBJETIVOS

O objetivo principal deste trabalho é investigar o mecanismo da atividade inibitória das N^2 -fenilguaninas, no plano molecular. Pretende-se ampliar o estudo das relações estrutura-atividade realizado por Gambino e colaboradores (1992), de forma que parâmetros calculados possam ser incluídos. Espera-se que a inclusão de novos parâmetros físico-químicos possa revelar detalhes ainda não conhecidos do mecanismo de interação TK-inibidor. Para levar a cabo essa tarefa, utilizou-se como ponto de partida a pesquisa realizada pelo grupo de Wright e colaboradores (Focher et al, 1988; Hildebrand et al, 1990; Gambino et al, 1992).

Pretende-se também elaborar hipótese sobre o modo de ligação das N^2 -fenilguaninas no sítio ativo de TK. Planeja-se utilizar como referência, o conhecimento dos modos de ligação do substrato natural de HSV1 TK, 2'-deoxitimidina, e do inibidor competitivo ganciclovir. O modo de ligação desses compostos foi elucidado recentemente (Wild et al, 1995; Brown et al, 1995).

Finalmente, pretende-se utilizar o modo de ligação das N^2 -fenilguaninas para verificar a validade dos resultados do estudo das relações estrutura-atividade.

Especificamente, os objetivos deste trabalho consistem em:

- (a) Estudar as características conformacionais e físico-químicas dos derivados da N^2 -fenilguanina;

- (b) Estabelecer relações quantitativas entre a estrutura química e a atividade biológica desses compostos, utilizando-se índices de reatividade e parâmetros moleculares calculados, além de constantes de substituintes clássicas;
- (c) Estudar os detalhes estruturais do sítio ativo de HSV1 TK e os diversos resíduos que participam do processo de interação;
- (d) Desenvolver modelo estrutural para HSV2 TK utilizando-se técnicas de modelagem molecular;
- (e) Propor modelo de interação entre as N^2 -fenilguaninas e o sítio de HSV TK;
- (f) Interpretar os resultados do estudo das relações estrutura-atividade em termos de interação dos derivados da N^2 -fenilguanina e resíduos do sítio ativo de HSV TK.

A seguir são anunciados os métodos gerais utilizados nas principais etapas do presente trabalho. Os detalhes de alguns procedimentos podem ser consultados ao longo dos próximos capítulos.

2.2 SELEÇÃO DOS COMPOSTOS

Tendo em vista a natureza computacional do presente trabalho, fez-se a seleção do grupo de compostos derivados da N^2 -fenilguanina a partir de pesquisa bibliográfica. Alguns critérios foram utilizados nessa seleção: (a) os compostos devem ser derivados de estrutura química comum, em que a diferença entre cada um deles resida apenas no tipo de substituinte químico presente em posições definidas da molécula; (b) os substituintes químicos devem apresentar boa variabilidade de propriedades lipofílica, eletrônica e estérica, para que seja possível detectar a influência dessas propriedades sobre a atividade biológica; (c) o grupo de compostos deve ser suficientemente numeroso para que modelos quantitativos mais complexos possam ser testados; (d) os compostos devem ter atividade

biológica definida e agirem pelo mesmo mecanismo de ação; (e) a atividade biológica deve ser determinada quantitativamente, seus valores devem ter variabilidade adequada, serem uniformemente distribuídos ao longo da série de compostos e, finalmente; (f) devem ser de fonte confiável.

O presente trabalho utiliza as medidas das atividades biológicas da série de 36 derivados da N^2 -fenilguanina, publicadas em três artigos pelo grupo de pesquisa do Dr. George Wright, do *Department of Pharmacology – University of Massachusetts Medical School* (Focher et al, 1988; Hildebrand et al, 1990; Gambino et al, 1992). Até onde pôde-se saber, os requisitos acima foram satisfeitos quando da escolha desses artigos como fontes de dados.

2.3 ANÁLISE CONFORMACIONAL

A análise conformacional realizada no presente trabalho compreende o estudo dos mapas de energia conformacional obtidos pela rotação das ligações sigma que unem os anéis fenila e guanina, dos derivados da N^2 -fenilguanina. Como são duas as ligações sigma que unem esses anéis, os mapas conformacionais serão superfícies de potencial em que a energia potencial conformacional será função dos ângulos torsionais θ_1 e θ_2 , que correspondem àquelas ligações sigma (ver Figura 16, pag.109).

Fez-se a estimativa da energia potencial conformacional dos derivados da N^2 -fenilguanina através do cálculo de seus calores de formação. Na execução do cálculo do calor de formação, utilizou-se o método semiempírico de cálculo de orbital molecular AM1 (Dewar et al, 1985), disponível no programa MOPAC 6.0 (Stewart, 1990a). O calor de formação não mede exclusivamente a energia conformacional, mas o inclui como parte do total. Considerando-se dada molécula capaz de adotar duas conformações diferentes, a diferença entre os calores de formação dessas conformações, calculados pelo método

AM1, corresponde à diferença de energia conformacional que uma possui em relação à outra.

2.3.1 Fundamentos do método AM1

O método AM1 é baseado na aproximação NDDO (*neglect of diatomic differential overlap*), e é considerado o estágio final da série de métodos baseados nessa aproximação (MINDO/3, MNDO, AM1 e PM3). PM3 não é método novo (Stewart, 1989a e 1989b), mas apenas a reparametrização de AM1. A aproximação NDDO despreza apenas o recobrimento (*differential overlap*) dos orbitais atômicos centrados em átomos diferentes. Assim, todas as integrais de dois elétrons e dois centros do tipo $(\mu\nu|\lambda\sigma)$, em que μ e ν estão centradas no mesmo átomo e λ e σ também estão centradas no mesmo átomo (mas não no mesmo átomo onde estão μ e ν) são mantidas. O método AM1, e os demais métodos com aproximação NDDO, dividem o átomo em duas partes: elétrons de valência e cerne (núcleo e elétrons das camadas internas).

A equação de Roothaan-Hall para sistemas de camada fechada é

$$\mathbf{FC} = \mathbf{SCE} \quad , \quad (1)$$

em que $\mathbf{F}$ é a matriz de Fock, $\mathbf{C}$ é a matriz dos autovetores (coeficientes dos orbitais moleculares), $\mathbf{S}$ é a matriz de recobrimento (*overlap*) e $\mathbf{E}$ é a matriz dos autovalores (energias dos orbitais moleculares) (Szabo & Ostlund, 1996). Todos os métodos semiempíricos possuem em comum a aproximação de $\mathbf{S}$ ser substituída por uma matriz identidade, $\mathbf{I}$, em que os elementos da diagonal são iguais a 1 e os elementos fora da diagonal são iguais a zero. Portanto, nos métodos semiempíricos a equação de Roothaan-Hall fica reduzida a

$$\mathbf{FC} = \mathbf{CE} \quad , \quad (2)$$

que é a forma matricial padrão.

É importante ressaltar que embora a matriz de recobrimento seja uma matriz identidade, não significa que todas as integrais de recobrimento sejam iguais a zero, durante o cálculo da matriz de Fock. Em verdade, muitas dessas integrais são diferentes de zero (ver Equação (10), adiante).

Os elementos da diagonal da matriz de Fock, F , são definidos por (Leach, 1996)

$$F_{\mu\mu} = H_{\mu\mu}^{\text{cerne}} + \sum_{v \in A} [P_{vv}(\mu\mu | vv) - \frac{1}{2}(P_{vv}(\mu v | \mu v))] + \sum_{B=A} \sum_{\lambda \in B} \sum_{\sigma \in B} P_{\lambda\sigma}(\mu\mu | \lambda\sigma) \quad , \quad (3)$$

em que $H_{\mu\mu}^{\text{cerne}}$ representa a contribuição da energia resultante do movimento dos elétrons de valência movendo-se no campo do cerne, $P_{\mu\mu}$ são elementos da matriz densidade e a notação $(\mu\nu|\lambda\sigma)$ representa a integral

$$(\mu\nu | \lambda\sigma) = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_{\mu}^*(1) \phi_{\nu}(1) \frac{1}{r_{12}} \phi_{\lambda}^*(2) \phi_{\sigma}(2) \quad , \quad (4)$$

em que $\phi_{\mu}(1)$ é a função de base representada por μ , sobre o átomo 1, e $\phi_{\mu}^*(1)$ é o seu conjugado complexo. O termo $H_{\mu\mu}^{\text{cerne}}$ é definido por

$$H_{\mu\mu}^{\text{cerne}} = U_{\mu\mu} - \sum_{B \neq A} V_{\mu\mu B} \quad , \quad (5)$$

em que $U_{\mu\mu}$ é a energia do orbital ϕ_{μ} no campo do cerne de seu próprio átomo (A), que por sua vez é definido por

$$U_{\mu\mu} = \left(\phi_{\mu} \left| -\frac{1}{2} \nabla^2 - \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|} \right| \phi_{\mu} \right) \quad . \quad (6)$$

O operador $-1/2\nabla^2$ fornece a energia cinética do elétron no orbital ϕ_{μ} e o segundo termo no interior da integral fornece a energia potencial da atração elétron-cerne, sendo que Z_A é a carga do cerne do átomo A, $\mathbf{r}_1$ é o vetor posição do elétron 1 e $\mathbf{R}_A$ é o vetor posição do núcleo do átomo A.

Os demais elementos da matriz de Fock são definidos por

$$F_{\mu\nu} = H_{\mu\nu}^{\text{cerne}} + \frac{3}{2}P_{\mu\nu}(\mu\nu|\mu\nu) - \frac{1}{2}(P_{\mu\nu}(\mu\mu|\nu\nu) + \sum_{B \neq A} \sum_{\lambda \in B} \sum_{\sigma \in B} P_{\lambda\sigma}(\mu\nu|\lambda\sigma)) \quad , \quad \mu \text{ e } \nu \text{ em } A \quad (7)$$

em que

$$H_{\mu\nu}^{\text{cerne}} = - \sum_{B \neq A} V_{\mu\nu B} \quad ; \quad (8)$$

e por

$$F_{\mu\nu} = H_{\mu\nu}^{\text{cerne}} - \frac{1}{2} \sum_{\lambda \in B} \sum_{\sigma \in A} P_{\lambda\sigma}(\mu\sigma|\nu\lambda) \quad , \quad \mu \text{ em } A \text{ e } \nu \text{ em } A \quad (9)$$

em que

$$H_{\mu\nu}^{\text{cerne}} = \frac{1}{2} S_{\mu\nu} (\beta_{\mu} + \beta_{\nu}) \quad (10)$$

O termo $S_{\mu\nu}$ é a integral de recobrimento entre os orbitais ϕ_{μ} e ϕ_{ν} , $(\phi_{\mu}|\phi_{\nu})$, e β_{μ} e β_{ν} são parâmetros ajustáveis. Os termos $V_{\mu\mu B}$ e $V_{\mu\nu B}$ são funções de atração cerne-elétron, de um elétron e dois centros. Esses termos representam a atração existente entre a distribuição eletrônica $\phi_{\mu}\phi_{\mu}$ ou $\phi_{\mu}\phi_{\nu}$, localizada no átomo A, e o cerne do átomo B.

$$V_{\mu\mu B} = -Z_B (\mu_A \mu_A | s_B s_B) \quad (11)$$

$$V_{\mu\nu B} = -Z_B (\mu_A \nu_A | s_B s_B) \quad , \quad (12)$$

em que s_B é um orbital do tipo s localizado no átomo B.

A diferença fundamental entre os métodos MNDO e AM1 está na função de repulsão do cerne (*core repulsion function*), E_{AB} , de acordo com a Equação (13).

$$E_{AB} = E_{\text{MNDO}} + \frac{Z_A Z_B}{R_{AB}} K_{A_i} \left\{ \sum_i \exp[-L_{A_i} (R_{AB} - M_{A_i})^2] + \sum_j K_{B_j} \exp[-L_{B_j} (R_{AB} - M_{B_j})^2] \right\} \quad (13)$$

em que E_{MNDO} é a função de repulsão do cerne do método MNDO e os termos que excedem E_{MNDO} são funções Gaussianas esféricas, cujas dimensões são definidas pelo parâmetro L . Os símbolos K e M representam parâmetros ajustáveis. O termo E_{MNDO} é definido por

$$E_{\text{MNDO}} = Z_A Z_B (s_A s_A | s_B s_B) [1 + \exp(-\alpha_A R_{AB}) + \exp(-\alpha_B R_{AB})] \quad (14)$$

Se o par atômico AB for NH ou OH (par do tipo XH), E_{MNDO} é calculado através de

$$E_{\text{MNDO}} = Z_X Z_H (s_X s_X | s_H s_H) [1 + R_{XH} \exp(-\alpha_X R_{XH}) + R_{XH} \exp(-\alpha_B R_{XH})] \quad (15)$$

2.3.2 Critérios de convergência do cálculo SCF em MOPAC

As iterações do cálculo SCF (*self-consistent field*), em MOPAC, são interrompidas quando dois testes são satisfeitos: (a) quando a diferença na energia eletrônica entre duas iterações consecutivas for menor que um valor de corte ajustável (10^{-4} kcal/mol, ou 10^{-6} kcal/mol se a palavra-chave **PRECISE** for utilizada, como foi o caso deste trabalho), e a diferença entre três iterações consecutivas for menor do que dez vezes aquele valor de corte e (b) quando a diferença entre os elementos da matriz de Fock de duas iterações consecutivas for menor do que um valor pré-estabelecido (Stewart, 1990b).

Os métodos semiempíricos presentes em MOPAC (MINDO/3, MNDO, AM1 e PM3) utilizam seis metodologias para obter a convergência do cálculo SCF: (a) convergência normal (a matriz de Fock inicial gera a matriz densidade inicial, que por sua vez gera nova e melhor matriz de Fock, que gera nova e melhor matriz densidade,...). Em geral esse processo leva à convergência do cálculo SCF. No entanto, as duas principais situações que dificultam a convergência são (i) matriz densidade inicial de má qualidade e (ii) equações que convergem lentamente devido a processo lento de transferência de carga

ou a oscilações persistentes (*long-lived oscillation*). Nesses casos outros métodos de convergência podem ser utilizados: (b) amortecimento de oscilações (*oscillation damping*); (c) interpolação de três pontos da matriz densidade (*three-point interpolation of the density matrix*); (d) deslocamento do nível de energia dos orbitais virtuais (*energy-level shift technique*); (e) método de Pulay; e (f) o convergente de Camp-King (para detalhes desses métodos, ver Stewart, 1990b).

2.3.3 Critérios de otimização de geometria molecular

Em termos ideais, a geometria molecular é considerada otimizada se não for possível baixar o calor de formação através de qualquer distorção da geometria. Embora esta condição raramente seja atingida, MOPAC considera a geometria molecular como otimizada se uma ou mais das seguintes quantidades for *suficientemente* pequena (Stewart, 1990b): (a) a variação prevista da geometria; (b) a variação prevista do calor de formação ou; (c) a atual norma do gradiente, ou seja, o módulo do vetor das derivadas da energia em relação às coordenadas (Stewart, 1990a).

O valor da norma do gradiente pode ser especificada pelo usuário, através da palavra-chave GNORM. O valor padrão de GNORM é 1,0. No presente trabalho, adotou-se o valor de GNORM = 0,1 para a análise conformacional e GNORM = 0,01 para o cálculo das propriedades utilizadas no Capítulo 6.

2.3.4 Cálculo do calor de formação

No método AM1, a energia eletrônica, E_{elet} , e a energia de repulsão nuclear, E_{nuc} , são calculados através de

$$E_{\text{elet}} = \frac{1}{2} \sum_{\mu} \sum_{\nu} P_{\mu\nu} (H_{\mu\nu} + F_{\mu\nu}) \quad (16)$$

$$E_{\text{nucl}} = \sum_{A < B} \sum E_{AB} \quad (17)$$

Nas Equações (16) e (17), $P_{\mu\nu}$ e $F_{\mu\nu}$ são os elementos da matriz densidade (**P**) e da matriz de Fock (**F**), obtidas durante o processo de SCF, e $H_{\mu\nu}$ corresponde aos elementos da matriz de integrais de um elétron (**H**). O termo E_{AB} é função de repulsão nuclear de cada par de átomos A e B (Dewar et al, 1985). A soma de E_{elctr} e E_{nucl} representa a energia liberada quando os átomos ionizados e seus elétrons de valência combinam-se para formar a molécula. O calor de formação molecular no estado padrão, ΔH_f° , é obtido acrescentando-se a essa soma a energia necessária para ionizar os elétrons de valência dos átomos da molécula, $E_{\text{el}}(A)$, calculada através de parâmetros semiempíricos, e a entalpia padrão de atomização, $\Delta H_f^\circ(A)$ (Stewart, 1990a).

$$\Delta H_f^\circ = E_{\text{elctr}} + E_{\text{nucl}} + \sum_A E_{\text{el}}(A) + \sum_A \Delta H_f^\circ(A) \quad (18)$$

É importante ressaltar que o calor de formação liberado nos resultados do método AM1 corresponde ao calor de formação experimental, ou seja, referente à temperatura de 25°C, e não ao calor de formação da geometria de equilíbrio a 0 K !

2.4 CÁLCULO DAS PROPRIEDADES FÍSICO-QUÍMICAS

O estudo de relações quantitativas entre estrutura química e atividade biológica requer a utilização de propriedades físico-químicas para a construção de modelos quantitativos que possam explicar a atividade biológica de dada série de compostos. Muitas das propriedades utilizadas nesse tipo de estudo são constantes de substituintes, ou seja, propriedades físico-químicas que se referem aos substituintes da série de compostos, e não propriamente aos compostos. No presente estudo, utilizaram-se não somente constantes de substituintes, tais como a constante lipofílica de Hansch-Fujita (Fujita et al, 1964) e a constante eletrônica de Hammett (Hammett, 1937), mas também propriedades

moleculares. Neste trabalho, utilizaram-se as seguintes propriedades moleculares: energia do orbital molecular ocupado de mais alta energia (ϵ_{HOMO}), energia do orbital molecular desocupado de mais baixa energia (ϵ_{LUMO}), dureza molecular (η), volume molecular de van der Waals (V_{vdw}), momento dipolar (μ) e polarizabilidade molecular (α). Além disso, utilizaram-se propriedades que estão relacionadas a átomos específicos da molécula: carga atômica líquida do n-ésimo átomo (Q_n) e índices de reatividade atômicos relacionados aos orbitais de fronteira. Os índices de reatividade utilizados foram a densidade eletrofílica de fronteira do n-ésimo átomo ($F_n^{(e)}$) e a densidade nucleofílica de fronteira ($F_n^{(n)}$) (Chin e Song, 1981).

2.4.1 Energias do HOMO (ϵ_{HOMO}) e LUMO (ϵ_{LUMO})

O cálculo de ϵ_{HOMO} e ϵ_{LUMO} é feito durante o processo SCF do método AM1. Os valores de ϵ_{HOMO} e ϵ_{LUMO} aparecem na matriz de autovalores, E , e, naturalmente, referem-se à geometria molecular otimizada.

2.4.2 Dureza molecular (η)

A dureza molecular é uma medida da intensidade da atração entre os núcleos atômicos e a nuvem eletrônica molecular. A definição qualitativa da dureza molecular está intimamente relacionada à da polarizabilidade, uma vez que a diminuição da diferença entre as energias do HOMO e LUMO geralmente resulta no aumento da polarizabilidade da molécula. A dureza molecular é calculada segundo a Equação (19) (Parr & Pearson, 1983).

$$\eta = (\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}) / 2 \quad (19)$$

2.4.3 Volume de van der Waals (VvdW)

O volume molecular de van der Waals foi calculado pelo algoritmo de Higo e Gō (1989), implementado computacionalmente (programa SURF) pelo autor (Gaudio e Takahata, 1992). O método baseia-se na construção de um paralelogramo capaz de conter a molécula, cujos átomos são representados por esferas interpenetrantes, centradas nas coordenadas atômicas e possuidoras de raio de van der Waals. O paralelogramo é dividido em cubos de 2 Å de aresta, que são denominados cubos do nível 1. A seguir os cubos do nível 1 são subdivididos sucessivamente em cubos de nível 2 (1 Å de aresta), 3 (0,5 Å de aresta), 4 (0,25 Å de aresta) e 5 (0,125 Å de aresta). Se o operador desejar, a subdivisão pode chegar até o nível 8 (0,015625 Å de aresta). Quanto mais elevado o nível, melhor a qualidade do cálculo. Quando se atinge o nível desejado, começa-se a contar o volume dos cubos que estão localizados no interior da molécula e a desprezar aqueles que se localizam no exterior. Os cubos localizados na superfície das esferas atômicas têm apenas 0,482 de seu volume (e não 0,5, devido à curvatura da superfície da esfera) computado para compor o volume molecular. O V_{vdw} dos derivados da N^2 -fenilguanina foram calculados utilizando-se o nível 7.

Alguns métodos alternativos de cálculo do volume molecular estão disponíveis (Hermann, 1972; Richmond, 1984; Karfunkel & Eyraud, 1989; Pascual-Ahuir & Silla, 1990) e seus resultados apresentam, em geral, boa concordância.

2.4.4 Cargas atômicas eletrostáticas (Q_n)

As cargas atômicas derivadas do potencial eletrostático molecular, também chamadas de cargas atômicas eletrostáticas, são o conjunto de cargas atômicas capazes de reproduzir o potencial eletrostático molecular, obtido a partir da função de onda molecular (Scrocco & Tomasi, 1973; Williams & Yan, 1987).

O potencial eletrostático molecular no ponto $\mathbf{r}$, $V(\mathbf{r})$, é definido como o trabalho necessário para mover uma carga positiva (+1) do infinito até o ponto $\mathbf{r}$, através do campo eletrostático molecular. A energia de interação eletrostática entre a molécula e uma carga q , colocada no ponto $\mathbf{r}$, é $qV(\mathbf{r})$. O potencial eletrostático molecular é resultante da contribuição nuclear, $V_{\text{nuc}}(\mathbf{r})$, e eletrônica, $V_{\text{elec}}(\mathbf{r})$.

$$V(\mathbf{r}) = V_{\text{nuc}}(\mathbf{r}) + V_{\text{elec}}(\mathbf{r}) \quad (20)$$

No caso da molécula que possui M núcleos, a contribuição nuclear é dada por

$$V_{\text{nuc}}(\mathbf{r}) = \sum_{A=1}^M \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \quad , \quad (21)$$

enquanto que a contribuição eletrônica é dada por

$$V_{\text{elec}}(\mathbf{r}) = - \int \frac{d\mathbf{r}' \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad , \quad (22)$$

em que $\mathbf{r}'$ é o vetor posição do elemento de carga eletrônica responsável pelo potencial eletrostático na posição $\mathbf{r}$ e $\rho(\mathbf{r}')$ é a função que define a densidade eletrônica na posição $\mathbf{r}'$.

A representação discretizada $V_{\text{elec}}(\mathbf{r})$ é dada por

$$V_{\text{elec}}(\mathbf{r}) = - \sum_{\mu} \sum_{\nu} P_{\mu\nu} \int \frac{\phi_{\mu}(\mathbf{r}') \phi_{\nu}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (23)$$

O potencial eletrostático molecular é uma função contínua (contribuição eletrônica) e de forma analítica difícil de representar. A implementação computacional do cálculo do potencial eletrostático molecular implica na discretização de sua representação. O objetivo é calcular o valor de um conjunto de cargas, localizadas nas coordenadas atômicas, capazes de reproduzir o potencial eletrostático mecânico-quântico numa série de pontos (*grid*) ao redor da molécula (Leach, 1996). Cox e Williams (1981) propuseram solução para este problema.

De acordo com Cox e Williams, o potencial eletrostático molecular deve ser calculado sobre cada um de uma série de pontos da grade ao redor da molécula, a partir da

função de onda molecular. Utiliza-se procedimento de ajuste de mínimos quadrados (*least-squares fit*) para determinar o melhor conjunto de cargas que reproduzem o potencial eletrostático naqueles pontos, sujeito à condição de que a soma total dessas cargas seja igual à carga líquida da molécula. O processo de ajuste tem como objetivo minimizar a soma dos quadrados das diferenças no potencial eletrostático. Dessa forma, se o potencial eletrostático, calculado a partir da função de onda, num ponto é V_i^0 e o valor do potencial eletrostático, calculado a partir das cargas atômicas, nesse ponto é V_i^{calc} , então o objetivo é minimizar a função (Leach, 1996)

$$R = \sum_{i=1}^{N_{\text{pontos}}} w_i (V_i^0 - V_i^{calc})^2 \quad , \quad (24)$$

em que N_{pontos} é o número de pontos na grade e w_i é um fator-pêso que atribui a importância relativa dos pontos da grade.

A N-ésima carga atômica depende do valor das demais, porque a soma deve ser igual a Z, a carga molecular. O valor da N-ésima carga é dada por

$$q_N = Z - \sum_{j=1}^{N-1} q_j \quad . \quad (25)$$

O potencial eletrostático devido às cargas q_j no ponto i é dado pela lei de Coulomb

$$V_i^{calc} = \sum_{j=1}^{N-1} \frac{q_j}{4\pi\epsilon_0 r_{ij}} + \frac{Z - \sum_{j=1}^{N-1} q_j}{4\pi\epsilon_0 r_{iN}} \quad , \quad (26)$$

em que r_{ij} é a distância entre a carga j e o ponto i . A minimização da função R implica em fazer

$$\frac{\partial R}{\partial q_k} = -2 \sum_{i=1}^{N_{\text{pontos}}} w_i (V_i^0 - V_i^{calc}) \left(\frac{\partial V_i^{calc}}{\partial q_k} \right) = 0 \quad . \quad (27)$$

Esta equação pode ser escrita da seguinte forma:

$$\sum_{i=1}^{N_{\text{pontos}}} w_i \left(V_i^0 - \frac{Z}{r_{iN}} \right) \left(\frac{1}{r_{ik}} - \frac{1}{r_{iN}} \right) = \sum_{j=1}^{N-1} \left[\sum_{i=1}^{N_{\text{pontos}}} w_i \left(\frac{1}{r_{ik}} - \frac{1}{r_{iN}} \right) \left(\frac{1}{r_{ij}} - \frac{1}{r_{iN}} \right) \right] \frac{q_j}{4\pi\epsilon_0} \quad (28)$$

A Equação (28) pode ser rearranjada na forma matricial $\mathbf{A}\mathbf{Q} = \mathbf{a}$. A matriz das cargas é determinada através de $\mathbf{Q} = \mathbf{A}^{-1}\mathbf{a}$ (Cox & Williams, 1981; Williams & Yan, 1987; Leach, 1996).

As cargas atômicas utilizadas neste trabalho foram derivadas a partir do potencial eletrostático molecular gerado pela função de onda AM1. A superioridade das cargas geradas a partir do potencial eletrostático molecular sobre as cargas de Mulliken e de Löwdin têm sido frequentemente citada na literatura (Ferenkzy et al, 1990; Besler et al, 1990; Orozco e Luque, 1990).

2.4.5 Momento dipolar (μ)

O momento dipolar molecular corresponde à soma das contribuições dos momentos de dipolo nuclear e eletrônico.

$$\mu = \mu_{\text{nuc}} + \mu_{\text{eletr}} \quad (29)$$

A contribuição dos núcleos atômicos para o momento dipolar total pode ser calculada de forma semelhante ao cálculo do momento dipolar de um conjunto de cargas discretas. Considerando-se dada molécula que possui M átomos, sendo que o A -ésimo átomo possui carga Z_A e posição $\mathbf{R}_A$, tem-se

$$\mu_{\text{nuc}} = \sum_{A=1}^M Z_A \mathbf{R}_A \quad (30)$$

A contribuição eletrônica para o momento dipolar total deve ser calculada levando-se em consideração a distribuição contínua da nuvem eletrônica molecular. Neste caso, deve-se aplicar o operador momento dipolar, $\mathbf{r}$, na função de onda molecular e integrar

sobre todo o espaço ($d\tau = dx dy dz$). No sistema molecular com N elétrons e cuja função de onda total é Ψ_0 , tem-se

$$\mu_{\text{elétr}} = \int d\tau \Psi_0 \left(\sum_{i=1}^N -r_i \right) \Psi_0 \quad (31)$$

Na Equação (31), o operador momento dipolar, r , foi substituído pela soma das contribuições de cada elétron para o momento dipolar eletrônico. Pode-se rescrever a Equação (31) em termos da matriz densidade, P , obtida no cálculo SCF do método AM1.

$$\mu_{\text{elétr}} = \sum_{\mu=1}^K \sum_{\nu=1}^K P_{\mu\nu} \int d\tau \phi_{\mu} (-r) \phi_{\nu} \quad (32)$$

Devido ao fato de o operador momento dipolar possuir componentes nas direções x , y e z , a Equação (32) é a resultante das contribuições nessas três direções. A componente na direção x da contribuição eletrônica do momento dipolar pode ser escrita como

$$\mu_x = \sum_{\mu=1}^K \sum_{\nu=1}^K P_{\mu\nu} \int d\tau \phi_{\mu} (-x) \phi_{\nu} \quad (33)$$

2.4.6 Polarizabilidade molecular (α)

Em MOPAC, dois métodos são utilizados no cálculo da polarizabilidade molecular: expansão em série do momento dipolar e expansão em série da energia total do sistema molecular (Kurtz et al, 1990).

Se dada molécula encontra-se sob a influência de um campo elétrico externo, F , seu momento dipolar, μ_i , pode ser expandido em série, de acordo com a Equação (34).

$$\mu_i = \mu_i^0 + \alpha_{ij} F_j + (1/2) \beta_{ijk} F_j F_k + (1/6) \gamma_{ijkl} F_j F_k F_l + \dots \quad (34)$$

em que μ_i^0 é o momento dipolar permanente e α_{ij} , β_{ijk} e γ_{ijkl} são elementos dos tensores da polarizabilidade linear e das hiperpolarizabilidades de primeira e segunda ordens da molécula, respectivamente.

De forma semelhante ao momento dipolar, a energia da molécula sujeita ao campo elétrico externo também pode ser expandida em série.

$$E_{(F)} = E_{(0)} - \mu_i F_i - (1/2!) \alpha_{ij} F_i F_j - (1/3!) \beta_{ijk} F_i F_j F_k - (1/4!) \gamma_{ijkl} F_i F_j F_k F_l - \dots, \quad (35)$$

em que $E_{(0)}$ é a energia do sistema na ausência do campo elétrico e F_i , F_j , etc. são os componentes do campo aplicado.

Alinhando-se o campo elétrico externo a um dos eixos cartesianos, eixo x por exemplo ($F_x, 0, 0$), os valores do momento dipolar perturbado (μ_x) e das polarizabilidades ao longo desse eixo (α_{xx} , β_{xxx} , γ_{xxxx}) podem ser obtidos. Nesse caso, as eqs. (34) e (35) tornam-se, respectivamente,

$$\mu_x = \mu_x^0 + \alpha_{xx} F_x + (1/2) \beta_{xxx} F_x^2 + (1/6) \gamma_{xxxx} F_x^3 + \dots, \quad (36)$$

e

$$E_{(F)} = E_{(0)} - \mu_x F_x - (1/2) \alpha_{xx} F_x^2 - (1/3!) \beta_{xxx} F_x^3 - (1/4!) \gamma_{xxxx} F_x^4 - \dots, \quad (37)$$

Truncando-se a Equação (36) após F_x^3 e estimando-se μ_x para quatro intensidades de campo elétrico diferentes ($\pm F_x$ e $\pm 2F_x$), produz-se quatro equações a quatro incógnitas.

Resolvendo-se o sistema para α_{xx} , obtém-se

$$\alpha_{xx} F_x^2 = (2/3) [\mu_{x(F_x)} - \mu_{x(-F_x)}] - (1/12) [\mu_{x(2F_x)} - \mu_{x(-2F_x)}] \quad (38)$$

Truncando-se a Equação (37) após F_x^4 e procedendo-se de maneira idêntica à Equação (36), obtém-se

$$\alpha_{xx} F_x^2 = (5/2) E_{(0)} - (4/3) [E_{(F_x)} + E_{(-F_x)}] + (1/12) [E_{(2F_x)} + E_{(-2F_x)}], \quad (39)$$

Alinhando-se o campo elétrico externo aos demais eixos cartesianos (0, F_y , 0) e (0, 0, F_z), obtém-se os valores de α_{yy} e α_{zz} . A polarizabilidade média é calculada através da média aritmética das componentes da polarizabilidades nos eixos x, y e z.

$$\alpha = 1/3(\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) \quad (40)$$

Métodos alternativos de cálculo da polarizabilidade molecular estão disponíveis na literatura (Dewar, et al, 1974; Abdul-Ahad et al, 1980; Dewar & Stewart, 1984; Lewis, 1989)

2.4.7 Índices de fronteira

A teoria dos orbitais de fronteira foi elaborada por Fukui e colaboradores (Fukui et al, 1952, 1954 e 1957) para explicar a reatividade preferencial de alguns átomos sobre outros, numa molécula orgânica particular, especialmente compostos aromáticos. De acordo com a teoria dos orbitais de fronteira, a reatividade preferencial de um átomo sobre outro frente a uma reação química, como por exemplo substituição eletrofílica, nucleofílica ou radicalar, está relacionada à densidade dos orbitais de fronteira, HOMO, LUMO ou ambos, respectivamente, sobre os átomos considerados. Assim, desprezando-se possíveis impedimentos estéricos, o átomo com maior densidade eletrônica relativa ao HOMO tem preferência na substituição eletrofílica.

Muitos índices de reatividade surgiram em função da teoria dos orbitais de fronteira. Alguns exemplos são as superdeslocalizabilidades eletrofílica, nucleofílica e radicalar, as densidades eletrofílica, nucleofílica e radicalar de fronteira, densidade de elétrons π , polarizabilidade átomo-átomo, etc. (Ching & Song, 1981). No presente trabalho, utilizou-se apenas as densidades eletrofílica e nucleofílica de fronteira.

A densidade eletrofílica de fronteira do n-ésimo átomo, $F_n^{(e)}$ e a densidade nucleofílica de fronteira, $F_n^{(n)}$, são calculadas de acordo com as seguintes equações:

$$F_n^{(e)} = \frac{2 \sum (c_n^{\text{HOMO}})^2}{-\epsilon_{\text{HOMO}}} \quad , \quad (41)$$

$$F_n^{(n)} = \frac{2 \sum (c_n^{\text{LUMO}})^2}{\epsilon_{\text{LUMO}}} \quad , \quad (42)$$

em que c_n^{HOMO} e c_n^{LUMO} são os coeficientes do HOMO e LUMO do n-ésimo átomo, respectivamente. Os índices de fronteira utilizados neste trabalho, $F_n^{(e)}$ e $F_n^{(n)}$, foram calculados a partir dos coeficientes da função de onda AM1.

Um ponto discutível é a utilização de parâmetros calculados a partir da energia dos orbitais virtuais (LUMO) provenientes de cálculos semiempíricos. São bem conhecidas as deficiências dos métodos semiempíricos em relação ao cálculo dos coeficientes e energias dos orbitais virtuais. A utilização de parâmetros calculados a partir de propriedades de orbitais virtuais é bastante questionável caso não sejam tomadas as devidas precauções. Em primeiro lugar, não é particularmente necessário que o valor absoluto de dada propriedade seja conhecida para que ela possa ser utilizada no método de Hansch-Fujita. Na verdade, é suficiente o conhecimento do valor relativo daquela propriedade para os diversos compostos. Em geral, as moléculas utilizadas nas análises de QSAR são bastantes semelhantes, pois as diferenças entre cada molécula é resumida à presença de substituintes variados em posições definidas da molécula-mãe. Nesta circunstância, torna-se aceitável a utilização de propriedades derivadas de orbitais virtuais, visto que o que se deseja é a tendência geral dos valores das propriedades do grupo de moléculas, em termos quantitativos.

2.5 RELAÇÕES ENTRE ESTRUTURA QUÍMICA E ATIVIDADE BIOLÓGICA

Utilizou-se o método de Hansch-Fujita (Hansch e Fujita, 1964; Kubinyi, 1993) no estudo de relações quantitativas entre estrutura química e atividade biológica. Esse método utiliza regressão linear múltipla como instrumento matemático. As equações de regressão

apresentadas neste trabalho foram geradas através do programa não publicado denominado MRA (*Multiple Regression Analysis*), desenvolvido pelo autor. As variáveis apresentadas em cada equação de regressão foram selecionadas por busca sistemática, em que todas as combinações envolvendo dado número de variáveis foram analisadas. A escolha das melhores equações foi baseada nos valores do coeficiente de correlação e do desvio-padrão. Entretanto, o limite de confiabilidade de 95% do coeficiente de cada variável e a distribuição dos resíduos também foram observados.

2.5.1 Regressão linear múltipla

Dado um conjunto de variáveis independentes (ortogonais) $X_1, X_2, \dots, X_i, \dots, X_k$, e uma variável dependente (de X_i) Y , promover a regressão linear de Y em função das variáveis X_i significa criar um modelo do tipo

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k, \quad (43)$$

em que b_0, b_1 , etc., são constantes de ajuste.

Aplicar modelo de dependência linear sobre um conjunto de dados não necessariamente significa que a dependência linear exista. Significa apenas que se acredita que tal dependência exista. Ou seja, assume-se que a variável Y depende da combinação linear das variáveis X_i e constrói-se o modelo. Faz-se, posteriormente, testes de validação que possam caracterizar a significância estatística do modelo linear.

O processo de regressão linear múltipla envolve os seguintes vetores e matriz:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{1,1} & X_{1,2} & \dots & X_{1,k} \\ 1 & X_{2,1} & X_{2,2} & \dots & X_{2,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n,1} & X_{n,2} & \dots & X_{n,k} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad 1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

em que n é o número de objetos ou observações (compostos, no presente trabalho) e k é o número de variáveis independentes (constantes de substituintes e parâmetros físico-químicos, no presente trabalho).

$Y \rightarrow$ vetor da variável dependente ($n \times 1$);

$X \rightarrow$ matriz das variáveis independentes ($n \times k+1$);

$\beta \rightarrow$ vetor dos parâmetros a serem estimados ($k+1 \times 1$);

$\varepsilon \rightarrow$ vetor dos erros ($n \times 1$);

$1 \rightarrow$ vetor de 1 ($n \times 1$)

Em termos exatos, executar a regressão linear múltipla significa determinar o vetor de parâmetros β e o vetor dos erros ε , tal que

$$Y = X\beta + \varepsilon \quad (44)$$

Em geral, a determinação de β e ε não é possível. Para isso seria necessário examinar todos os valores possíveis de X e Y . Na prática, utilizam-se os valores disponíveis de X e Y , ou seja, X e Y , para obter a estimativa de β . O vetor b é o valor estimado de β .

$$b = E(\beta) \quad (45)$$

O cálculo de b é feito através da minimização da soma dos quadrados dos desvios dos valores previstos de Y , $\hat{Y}$, e os valores observados de Y , Y (Draper & Smith, 1981). Sem entrar nos detalhes da diferenciação da soma de ε_i^2 , o cálculo de b é obtido através da operação

$$b = (X'X)^{-1} X'Y \quad (46)$$

em que o apóstrofo indica transposição matricial e “ -1 ” indica a inversão matricial. O vetor b também é conhecido como vetor dos coeficientes da regressão linear.

O valor estimado de Y_i , $\hat{Y}_i$, é dado por

$$\hat{Y}_i = b_0 + b_1 X_{i,1} + b_2 X_{i,2} + \dots + b_k X_{i,k} \quad (47)$$

A precisão da regressão estimada é feita através da análise da variância da regressão. A construção da tabela da análise de variância (Draper & Smith, 1981) permite o cálculo de três propriedades capazes de medir o grau de ajuste do modelo linear ao conjunto de dados em questão: o coeficiente de correlação (r), o desvio-padrão (s) e o teste de Fischer, razão das variâncias ou teste-F (F). O cálculo destas propriedades é baseado na identidade

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y}_i)^2 \quad (48)$$

que, pode ser lida como

SS dos valores observados em relação média = SS dos valores previstos em relação à média + SS dos resíduos,

em que SS significa soma dos quadrados (*sum of squares*) e $\bar{Y}$ é a média dos valores de Y_i . Quanto menor for *SS dos resíduos*, ou seja, quanto mais próximo *SS dos valores previstos em relação à média* for de *SS dos valores observados em relação média*, melhor o ajuste do modelo linear ao conjunto de dados. Baseado nesse argumento, define-se o quadrado do coeficiente de correlação pela relação

$$r^2 = \frac{\text{SS dos valores previstos em relação à média}}{\text{SS dos valores observados em relação à média}} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \leq 1. \quad (49)$$

A definição acima implica em que modelos lineares bem ajustados possuem r^2 próximos à unidade.

O quadrado do desvio-padrão da regressão é definido por

$$s^2 = \frac{\text{SS dos resíduos}}{n - k - 1 \text{ graus de liberdade}} = \frac{\sum (Y_i - \hat{Y}_i)^2}{n - k - 1} \quad (50)$$

que implica em quanto menor o valor de s , melhor a regressão.

O teste de Fischer é definido por

$$F = \frac{\frac{\text{SS dos valores previstos em relação à média}}{(k-1) \text{ graus de liberdade}}}{\frac{\text{SS dos resíduos}}{(n-k-1) \text{ graus de liberdade}}} = \frac{\frac{\sum (\hat{Y}_i - \bar{Y})^2}{k-1}}{\frac{\sum (Y_i - \hat{Y}_i)^2}{n-k-1}}, \quad (51)$$

que implica em

$$F = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{(k-1) s^2} \quad (52)$$

Como o termo $\sum (\hat{Y}_i - \bar{Y})^2$ deve ser maior possível e s^2 , menor possível, um valor de F grande (ordem de dezenas ou centenas, até milhares) revela a boa qualidade de ajuste da regressão. O valor de F, mostrado nas equações de regressão como $F_{(k,n-k-1)}$, deve ser comparado com o valor mínimo aceitável para o teste, disponível em tabelas que podem ser encontradas na maioria dos livros de estatística (Draper & Smith, 1981). Muitas tabelas desse tipo podem ser encontradas, cada uma apresentando valores de referência para o teste-F que correspondem a dado nível de confiabilidade (p) que se espera que a regressão possua. O valor geralmente tomado como referência corresponde ao nível de confiabilidade superior a 95% ($p < 1 - 0,95 = 0,05$). Em geral aconselha-se que uma regressão de qualidade razoável deva apresentar teste-F no mínimo cinco vezes maior do que o valor mínimo aceitável. Como não é costume apresentar esse valor mínimo nos resultados da regressão, deve-se apresentar o nível de confiabilidade atual da regressão, que corresponde ao valor obtido para o teste-F ($p < 0,0001$ é considerado ótimo resultado).

A análise de variância indica a capacidade do modelo de fazer previsões em relação aos objetos incluídos na análise. Para medir a capacidade do modelo de fazer previsões em relação a objetos não incluídos na análise, deve-se lançar mão da validação cruzada (*cross validation*). A forma mais popular de validação cruzada é a do tipo LOO (*leave one out*), em que se exclui um dos objetos, faz-se a regressão com $n-1$ objetos e utiliza-se a regressão obtida para prever o objeto deixado de fora do modelo. Repetindo-se o processo

para todos os objetos, pode-se derivar o parâmetro PRESS (*predictive residual sum of squares*).

$$\text{PRESS} = \sum (Y_i - \hat{Y}_i)^2 \quad (53)$$

Nota-se que PRESS é a soma dos resíduos entre o valor observado e o valor previsto do objeto deixado de fora no processo de validação cruzada. Por isso, PRESS deverá ter menor valor possível para que a regressão possa ser considerada como tendo boa capacidade preditiva. O parâmetro PRESS pode ser utilizado para obter os principais índices que medem a capacidade de previsão da regressão linear: o quadrado do coeficiente de correlação das previsões (Q^2), o desvio-padrão das previsões (s_{PRESS}) e o erro padrão das previsões (SDEP).

$$Q^2 = 1 - \frac{\text{PRESS}}{\sum (\hat{Y}_i - \bar{Y})^2} \quad (54)$$

Viu-se anteriormente que o termo $\sum (\hat{Y}_i - \bar{Y})^2$ corresponde à *SS dos valores previstos em relação à média*, e que deveria ter maior valor possível, tendendo ao valor da *SS dos valores observados em relação média*, $\sum (Y_i - \bar{Y})^2$. Portanto, a relação $\text{PRESS} / \sum (\hat{Y}_i - \bar{Y})^2$ deve ter menor valor possível. Consequentemente, Q^2 deverá ser próximo da unidade para que a regressão tenha boa capacidade preditiva.

O desvio-padrão das previsões (s_{PRESS}) e o erro padrão das previsões (SDEP) são calculados de acordo com as seguintes equações:

$$s_{\text{PRESS}} = \sqrt{\frac{\text{PRESS}}{n - k - 1}} \quad (55)$$

$$\text{SDEP} = \sqrt{\frac{\text{PRESS}}{n}} \quad (56)$$

Da mesma forma que PRESS, os parâmetros s_{PRESS} e SDEP deverão apresentar menores valores possíveis para que a regressão tenha boa capacidade preditiva.

2.5.2 Seleção de variáveis por busca sistemática

Fez-se a seleção de variáveis por busca sistemática. Para cada grupo de compostos incluídos na análise (meta-substituídos, para-substituídos, etc.) foram analisadas todas as regressões de $\log(1/IC_{50})$ em função de uma variável independente, todas as regressão de $\log(1/IC_{50})$ em função de duas variáveis independentes, e assim por diante, até o máximo de cinco variáveis independentes. A enorme quantidade de modelos a serem testados faz com que seja necessário optar-se por alguma propriedade da regressão que indique sua capacidade de ajuste e utilizá-la como parâmetro de seleção. Além do mais, é necessário utilizar metodologia eficiente de cálculo dessa propriedade para viabilizar a verificação de todos os modelos no menor espaço de tempo possível. A propriedade escolhida foi o coeficiente de correlação. A metodologia de cálculo rápido que viabilizou a busca sistemática é mostrada a seguir (Sachs, 1984; Kubinyi, 1994).

O coeficiente de correlação da variável dependente Y em função da i-ésima variável independente X_i , $r_{y,i}$, é dado por

$$r_{y,i} = \frac{S_{yi}}{\sqrt{S_{yy} \cdot S_{ii}}} \quad (57)$$

em que S_{yi} , S_{yy} e S_{ii} são definidos de acordo com a Equações (58)-(60)

$$S_{yi} = \sum Y \cdot X_i - \sum Y \cdot \sum X_i / n \quad (58)$$

$$S_{yy} = \sum Y^2 - (\sum Y)^2 / n \quad (59)$$

$$S_{ii} = \sum X_i^2 - (\sum X_i)^2 / n \quad (60)$$

No caso dos modelos com duas variáveis independentes, X_i e X_j , o coeficiente de correlação de Y em função de X_i e X_j , $r_{y,ij}$, pode ser rapidamente calculado por

$$r_{y,ij}^2 = \frac{r_{y,i}^2 + r_{y,j}^2 - 2 \cdot r_{y,i} \cdot r_{y,j} \cdot r_{ij}}{1 - r_{ij}^2}, \quad (61)$$

em que r_{ij} , o coeficiente de correlação entre X_i e X_j , é dado por

$$r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii} \cdot S_{jj}}} \quad (62)$$

O termo S_{ij} é dado por

$$S_{ij} = \sum X_i \cdot X_j - \sum X_i \cdot \sum X_j / n \quad (63)$$

O coeficiente de correlação de Y em função de X_i , X_j e X_k , $r_{y.ijk}$, pode ser calculado de acordo com

$$r_{y.ijk}^2 = 1 - (1 - r_{yi}^2) \cdot (1 - r_{yj}^2) \cdot (1 - r_{yk}^2) \quad (64)$$

em que $r_{yj.i}$ e $r_{yk.ij}$ são dados por

$$r_{yj.i} = \frac{r_{y.j} - r_{y.i} \cdot r_{ij}}{\sqrt{(1 - r_{yi}^2) \cdot (1 - r_{ij}^2)}} \quad (65)$$

e

$$r_{yk.ij}^2 = \frac{(r_{y.k} - r_{yji} \cdot r_{jk.i})^2}{(1 - r_{yj.i}^2) \cdot (1 - r_{jk.i}^2)} \quad (66)$$

O termo $r_{jk.i}$ é dado por

$$r_{jk.i} = \frac{r_{j.k} - r_{ij} \cdot r_{ik}}{\sqrt{(1 - r_{ij}^2) \cdot (1 - r_{ik}^2)}} \quad (67)$$

De forma semelhante, o coeficiente de correlação entre Y e X_i , X_j , X_k e X_l pode ser calculado pelas seguintes equações

$$r_{y.ijkl}^2 = 1 - (1 - r_{yi}^2) \cdot (1 - r_{yj}^2) \cdot (1 - r_{yk}^2) \cdot (1 - r_{yl}^2) \quad (68)$$

$$r_{yl.ijk}^2 = \frac{(r_{y.l} - r_{yjk.ij} \cdot r_{kl.ij})^2}{(1 - r_{yjk.ij}^2) \cdot (1 - r_{kl.ij}^2)} \quad (69)$$

$$r_{yl.ij} = \frac{r_{y.l} - r_{yji} \cdot r_{jl.i}}{\sqrt{(1 - r_{yj.i}^2) \cdot (1 - r_{jl.i}^2)}} \quad (70)$$

$$r_{yk \cdot ij} = \frac{r_{yk \cdot i} - r_{yj \cdot i} \cdot r_{jk \cdot i}}{\sqrt{(1 - r_{yj \cdot i}^2) \cdot (1 - r_{jk \cdot i}^2)}} \quad (71)$$

$$r_{kl \cdot ij} = \frac{r_{kl \cdot i} - r_{jk \cdot i} \cdot r_{jl \cdot i}}{\sqrt{(1 - r_{jk \cdot i}^2) \cdot (1 - r_{jl \cdot i}^2)}} \quad (72)$$

Finalmente, o coeficiente de correlação entre Y e X_i , X_j , X_k , X_l e X_m pode ser calculado pelas equações

$$r_{y \cdot ijk}^2 = 1 - (1 - r_{y \cdot i}^2) \cdot (1 - r_{y \cdot j}^2) \cdot (1 - r_{y \cdot k}^2) \cdot (1 - r_{y \cdot l}^2) \cdot (1 - r_{y \cdot m}^2) \quad (73)$$

$$r_{ym \cdot ijk}^2 = \frac{(r_{ym \cdot ijk} - r_{yl \cdot ijk} \cdot r_{lm \cdot ijk})^2}{(1 - r_{yl \cdot ijk}^2) \cdot (1 - r_{lm \cdot ijk}^2)} \quad (74)$$

$$r_{ym \cdot ijk} = \frac{r_{ym \cdot ij} - r_{yk \cdot ij} \cdot r_{km \cdot ij}}{\sqrt{(1 - r_{yk \cdot ij}^2) \cdot (1 - r_{km \cdot ij}^2)}} \quad (75)$$

$$r_{yl \cdot ijk} = \frac{r_{yl \cdot ij} - r_{yk \cdot ij} \cdot r_{kl \cdot ij}}{\sqrt{(1 - r_{yk \cdot ij}^2) \cdot (1 - r_{kl \cdot ij}^2)}} \quad (76)$$

$$r_{lm \cdot ijk} = \frac{r_{lm \cdot ij} - r_{kl \cdot ij} \cdot r_{km \cdot ij}}{\sqrt{(1 - r_{kl \cdot ij}^2) \cdot (1 - r_{km \cdot ij}^2)}} \quad (77)$$

Recentemente, nota-se o interesse pela utilização dos algoritmos genéticos no processo de seleção de variáveis (ver as referências Kubinyi, 1994a e 1994b, bem como as referências citadas nesses artigos).

2.6 ALINHAMENTO DAS SEQUÊNCIAS DE AMINOÁCIDOS

O alinhamento dos grupos de seqüências de aminoácidos, apresentados no Capítulo 5, foram feitos através do método MDM (*Mutal Data Matrix*) (Dayhoff, 1978; Schwartz e Dayhoff, 1978), implementado no programa Cameleon (1994). O alinhamento das seqüências não foi feito de forma completamente automatizada. Fez-se manualmente o alinhamento prévio dos padrões de aminoácidos (*motifs*) que reconhecidamente são conservados em cada seqüência. As partes previamente alinhadas foram mantidas juntas

durante o processo de alinhamento executado pelo algoritmo MDM. Esse procedimento oferece maior segurança no alinhamento de seqüências longas, tais como aquelas analisadas neste trabalho.

2.7 MODELAGEM MOLECULAR

Boa parte dos trabalhos executados no Capítulo 5 envolveram a manipulação da estrutura cristalográfica de HSV1 TK, que foi gentilmente cedida por Mark R. Sanderson e John Champness, do Kings College, Londres. Como regra geral, os modelos de estruturas tridimensionais de macromoléculas (arquivos PDB – *protein data bank*), que são gerados a partir da interpretação do padrão de difração de raios-X produzido pelo cristal da macromolécula, não se encontram em condições de serem submetidos a cálculos moleculares (a discussão pormenorizada sobre esse assunto pode ser vista na Seção 5.4.1, “Preparação da molécula protéica para cálculos moleculares”, pag. 183). O fato é que geralmente o modelo macromolecular tridimensional precisa ser preparado (*clean up*) para ser devidamente reconhecido por programas de cálculo molecular, via de regra mecânica molecular. Fez-se a preparação da estrutura cristalográfica de HSV1 TK através do módulo *Protein Design* do programa Quanta-CHARMM. Utilizou-se a versão 22.3 do campo de força CHARMM (Brooks et al, 1983; MacKerell, Jr. et al, 1995). A preparação dos ligantes associados à estrutura cristalográfica de HSV1 TK, 2'-deoxitimidina e ganciclovir, foi feita através do módulo *Molecular Editor* do programa Quanta-CHARMM.

Fez-se construção do modelo tridimensional de HSV2 TK, a partir da estrutura cristalográfica de HSV1 TK, que foi utilizada como molde, também através do módulo *Protein Design* do programa Quanta-CHARMM. A construção do modelo de HSV2 TK envolveu a realização de mutações sucessivas na seqüência de aminoácidos do modelo

tridimensional de HSV1 TK, até que esta se transformasse na sequência de aminoácidos de HSV2 TK, seguido de minimizações de energia. A Seção 5.4.4 ("Modelagem das estruturas terciária e quaternária de HSV2 TK", pag. 188) apresenta os detalhes desse procedimento.

2.7.1 Campo de força CHARMM

A função energia total do programa CHARMM (versão 22.3) é dada por (CHARMM, 1994; MacKerell, Jr. et al, 1995)

$$E = E_{\text{lig}} + E_{\text{ang}} + E_{\text{diedr}} + E_{\text{UB}} + E_{\text{impr}} + E_{\text{nl}} \quad (78)$$

em que E_{lig} é a energia relativa ao estiramento das ligações químicas, E_{ang} refere-se às deformações angulares, E_{diedr} à rotação dos ângulos diedros, E_{UB} é a função de Urey-Bradley para interações 1-3, E_{impr} refere-se à deformação dos diedros impróprios (diedro no qual os átomos não estão ligados na ordem 1-2-3-4, como por exemplo no NH_3) e E_{nl} às interações entre átomos não ligados. A definição de cada termo é mostrada a seguir:

$$E_{\text{lig}} = \sum K_b (b - b_0)^2 \quad (79)$$

em que K_b é a constante de estiramento, b é o comprimento da ligação e b_0 é o valor de referência do comprimento da ligação.

$$E_{\text{ang}} = \sum K_\theta (\theta - \theta_0)^2 \quad (80)$$

em que K_θ é a constante de deformação angular, θ é o ângulo da ligação e θ_0 é o valor de referência do ângulo da ligação.

$$E_{\text{UB}} = \sum K_{\text{UB}} (S - S_0)^2 \quad (81)$$

em que K_{UB} é a constante de Urey-Bradley, S é a distância 1-3 de Urey-Bradley e S_0 é o valor de referência da distância Urey-Bradley.

$$E_{\text{diedr}} = \sum K_\chi (1 + \cos(n\chi - \delta)) \quad (82)$$

em que K_x é a constante de deformação torsional, n é um número inteiro, χ é o ângulo diedro e δ é a constante de fase do ângulo.

$$E_{\text{impr}} = K_{\text{impr}} (\varphi - \varphi_0)^2 \quad (83)$$

em que K_{impr} é a constante de deformação do diedro impróprio, φ é o valor do diedro impróprio e φ_0 é o valor de referência do diedro impróprio.

$$E_{\text{el}} = \sum_{i,j} \epsilon \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (84)$$

em que o termo entre colchetes corresponde às interações atrativa de van der Waals e repulsiva nuclear (potencial de Lennard-Jones) e o último termo do somatório corresponde às interações eletrostáticas (termo coulômbico).

2.7.2 Simulação da dinâmica molecular

Promover a simulação da dinâmica (clássica) de um sistema molecular significa tratar o sistema como se fora composto por partículas clássicas, cujas interações são governadas pela Equação (78), e gerar sucessivas configurações para o sistema através da integração das leis do movimento de Newton. O resultado da simulação é a trajetória das partículas do sistema, que mostra como a posição e a velocidade das partículas variam com o tempo. A trajetória é obtida resolvendo-se as equações diferenciais embutidas na segunda lei de Newton ($F = ma$):

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad , \quad (85)$$

que descreve o movimento da partícula de massa m_i , cuja posição é x_i , ao longo da coordenada x , sendo F_{x_i} a força nessa direção que atua sobre a partícula.

Utilizou-se o algoritmo de Verlet (1967) para integrar as equações de movimento do sistema. O algoritmo de Verlet utiliza a posição e a aceleração no tempo t , e a posição

no passo de tempo anterior, $r(t - \delta t)$, para calcular a nova posição no tempo $t + \delta t$, $r(t + \delta t)$.

Todos os algoritmos de integração das equações de movimento utilizados em dinâmica molecular assumem que as relações entre posição, velocidade, aceleração, etc. podem ser aproximadas por expansões em séries de Taylor:

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \frac{1}{6} \delta t^3 b(t) + \frac{1}{24} \delta t^4 c(t) + \dots \quad (86)$$

$$v(t + \delta t) = v(t) + \delta t a(t) + \frac{1}{2} \delta t^2 b(t) + \frac{1}{6} \delta t^3 c(t) + \dots \quad (87)$$

$$a(t + \delta t) = a(t) + \delta t b(t) + \frac{1}{2} \delta t^2 c(t) + \dots \quad (88)$$

O algoritmo de Verlet utiliza as seguintes relações para produzir a trajetória durante a simulação:

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \dots \quad (89)$$

$$r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) - \dots \quad (90)$$

A soma dessas duas equações resulta em

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad , \quad (91)$$

que não inclui a velocidade explicitamente. A velocidade é calculada fazendo-se a diferença entre as Equações (89) e (90) e dividindo-se o resultado por $2\delta t$. Isso resulta em

$$v(t) = [r(t + \delta t) - r(t - \delta t)] / 2\delta t \quad (92)$$

2.7.3 Cargas atômicas

Nas simulações de dinâmica molecular realizadas neste trabalho, utilizou-se o método de Gasteiger e Marsili (1980) para calcular as cargas atômicas do sistema. Esse método calcula as cargas atômicas utilizando apenas informações sobre os átomos

presentes no sistema e o modo como eles estão conectados, que o torna bastante rápido. O método utiliza o princípio da equiparação parcial da eletronegatividade do orbital.

Gasteiger e Marsili assumiram dependência polinomial entre a eletronegatividade $\chi_{\mu A}$ do orbital ϕ_{μ} no átomo A e a carga Q_A do átomo A:

$$\chi_{\mu A} = a_{\mu} + b_{\mu A} Q_A + c_{\mu A} Q_A^2 \quad , \quad (93)$$

em que os valores dos coeficientes a, b e c foram derivados para os elementos químicos mais comuns em seus estados de valência usuais. Segundo o método, elétrons fluem dos elementos menos eletronegativos para os mais eletronegativos. Este fluxo resulta em carga positiva sobre o elemento menos eletronegativo e carga negativa no mais eletronegativo, tendendo a equiparar as eletronegatividades, uma vez que a eletronegatividade do orbital depende de sua ocupação. A equiparação total, no entanto, não conduz a resultados quimicamente razoáveis (Leach, 1996). Gasteiger e Marsili adotaram processo iterativo de equiparação das eletronegatividades, em que menos e menos carga é transferida entre dois átomos em cada iteração. A carga transferida do átomo A para o átomo B (supondo B mais eletronegativo que A) na k-ésima iteração é dada por:

$$Q^{(k)} = \frac{\chi_B^{(k)} - \chi_A^{(k)}}{\chi_A^+} \alpha^k \quad , \quad (94)$$

em que $Q^{(k)}$ é a carga transferida, $\chi_A^{(k)}$ e $\chi_B^{(k)}$ são as eletronegatividades dos átomos A e B, χ_A^+ é a eletronegatividade do cátion do elemento menos eletronegativo e α é um fator de amortecimento, elevado à potência k. No início do processo, cargas-padrão são atribuídas a cada átomo. A cada iteração as eletronegatividades são calculadas pela Equação (93) e a carga a ser transferida, pela Equação (94). A carga total de dado átomo, ao final de cada iteração, é obtida somando-se a carga transferida de todas as ligações para o átomo ao valor da carga obtida na iteração anterior. O fator de amortecimento, α , reduz a

influência dos átomos mais eletronegativos. Utilizando-se $\alpha = 1/2$, a convergência é obtida rapidamente, geralmente em quatro ou cinco passos (Leach, 1996).

CAPÍTULO 3

ANÁLISE CONFORMACIONAL

3.1 INTRODUÇÃO

A presente análise conformacional tem como objetivos principais o estudo da estrutura geométrica e o estudo dos fatores que influem na estrutura geométrica dos derivados da N^2 -fenilguanina. Decidiu-se por realizar este estudo em três etapas. A primeira etapa refere-se ao estudo do fenômeno da tautomeria existente na molécula da guanina e seu impacto sobre as aproximações adotadas nos cálculos moleculares. A segunda etapa consiste na padronização do método de cálculo molecular a ser utilizado na construção dos mapas conformacionais. A terceira e última etapa consiste no estudo dos mapas de energia conformacional obtidos através da rotação das ligações sigma que unem os anéis fenila e guanina, nos derivados da N^2 -fenilguanina. O estudo desses mapas de energia conformacional é especialmente importante nos casos em que o receptor farmacológico e/ou o modo de ligação dos fármacos em estudo não são conhecidos.

3.2 TAUTOMERIA NA GUANINA

3.2.1 Tautomeria

O fenômeno da tautomeria é bem conhecido na química. Aos estudantes de química são geralmente apresentadas suas duas formas clássicas: as tautomerias aldo-enólica e ceto-enólica (ver Figura 8).

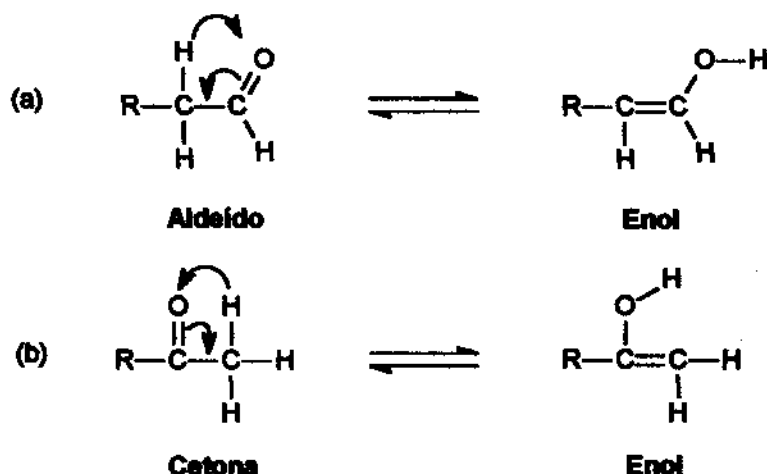


Figura 8 - (a) Tautomeria aldo-enólica; (b) tautomeria ceto-enólica.

A tautomeria é o equilíbrio dinâmico entre duas ou mais espécies químicas que diferem entre si pela posição de próton e/ou de elétrons de ligação, com subsequente rearranjo da configuração molecular. A tautomeria é geralmente classificada em dois tipos: tautomeria de próton (Figura 9a), em que ocorre migração de próton, e tautomeria de valência (Figura 9b), em que ocorre reorganização das ligações moleculares, sem que ocorra migração atômica (Orchin, 1980). Cada forma tautomérica, ou tautômero, possui energia interna própria, que define a população relativa dos tautômeros, de acordo com a distribuição de Boltzmann. Quanto menor a energia interna de um tautômero em relação a outro, maior será a população relativa daquela forma tautomérica.

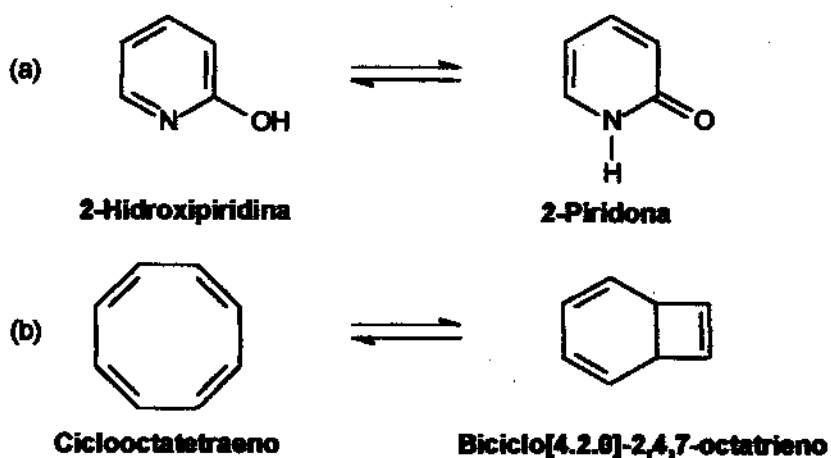


Figura 9 - (a) Tautomeria de próton; (b) tautomeria de valência.

3.2.2 Tautomeria nos processos bioquímicos

O mecanismo de muitas reações químicas pode ser explicado com base no tautomerismo. Livros-texto básicos de química orgânica citam muitos exemplos dessas reações (Morrison & Boyd, 1992; Streitwieser et al, 1992; Solomons, 1996). Menos comuns aos químicos, no entanto, são as implicações do tautomerismo em bioquímica. Existem modelos que consideram o tautomerismo de ligantes químicos como sendo importante na ativação de receptores histaminérgicos H_2 (Topiol et al, 1984; Weinstein et al, 1986) e receptores de serotonina (Osman et al, 1985 e 1987). O tautomerismo do ligante pode influenciar a interação enzima-ligante em outros sistemas bioquímicos (Topiol, 1987). De forma semelhante, o tautomerismo do sítio ativo é importante no mecanismo de ação de enzimas como proteases da serina (Blow & Steitz, 1970) e DNase (Suck & Oefner, 1986). Esses exemplos mostram que a tautomeria pode ser muito importante em alguns fenômenos bioquímicos. Por conseguinte, deve-se considerar a tautomeria sempre que estiver presente em moléculas bioquímicas.

3.2.3 Tautomeria na guanina

As bases nitrogenadas dos ácidos nucleicos, adenina, citosina, guanina, timina e uracila, também apresentam tautomeria. A Figura 10 mostra os tautômeros mais estáveis das bases nitrogenadas dos ácidos nucleicos. A tautomeria é muito importante no processo de replicação do DNA. Alguns tipos de mutações genéticas estão associadas ao emparelhamento das bases nitrogenadas da cadeia replicante aos tautômeros de alta energia de seus pares correspondentes (Topal & Fresco, 1976; Stryer, 1988). Em seu mais famoso trabalho, Watson e Crick (1953) reconheceram a importância do equilíbrio entre as formas tautoméricas e iônicas no reconhecimento das bases do DNA, que é modulado pela intrincada rede de interações via pontes de hidrogênio.

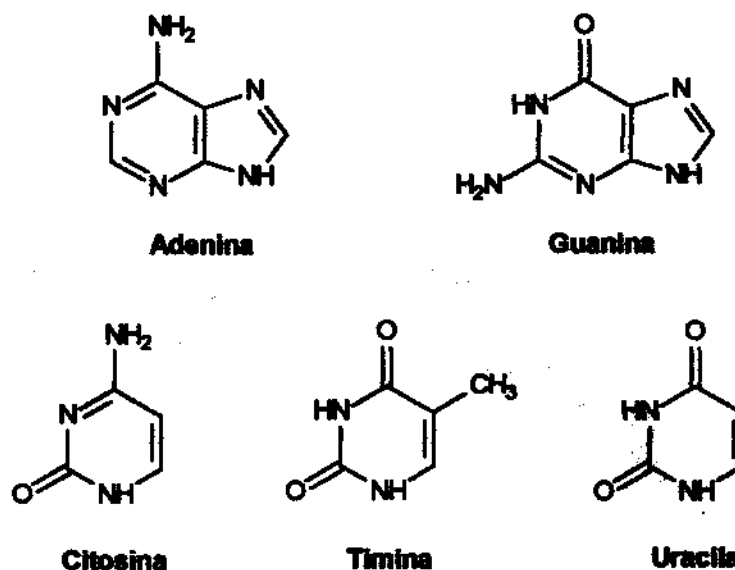


Figura 10 - Tautômeros mais estáveis das bases nitrogenadas dos ácidos nucleicos.

Existe farta literatura sobre os diversos aspectos do tautomerismo das bases dos ácidos nucleicos. Os estudos comparativos mais relevantes da energia dos tautômeros da guanina podem ser vistos nas seguintes referências: Sigulá & Buda, 1983; Sabio et al, 1990; Katritzky & Karelson, 1991; Fabian, 1991; Colominas et al, 1996. Outros estudos comparativos das energias dos tautômeros da guanina podem ser encontrados nas referências: Gould & Hillier, 1989; Leszczyński, 1990; Kwiatkowski & Leszczyński, 1990; Gould et al, 1995. Vários outros estudos sobre os diversos aspectos do tautomerismo da adenina e guanina também podem ser vistos nas referências 5a-b, 6a-j e 7a-p, citadas no artigo Colominas et al, 1996.

Os tautômeros da guanina frequentemente citados como sendo os mais estáveis (Sigulá & Buda, 1983; Sabio et al, 1990; Katritzky & Karelson, 1991; Fabian, 1991; Colominas et al, 1996), são mostrados na Figura 11.

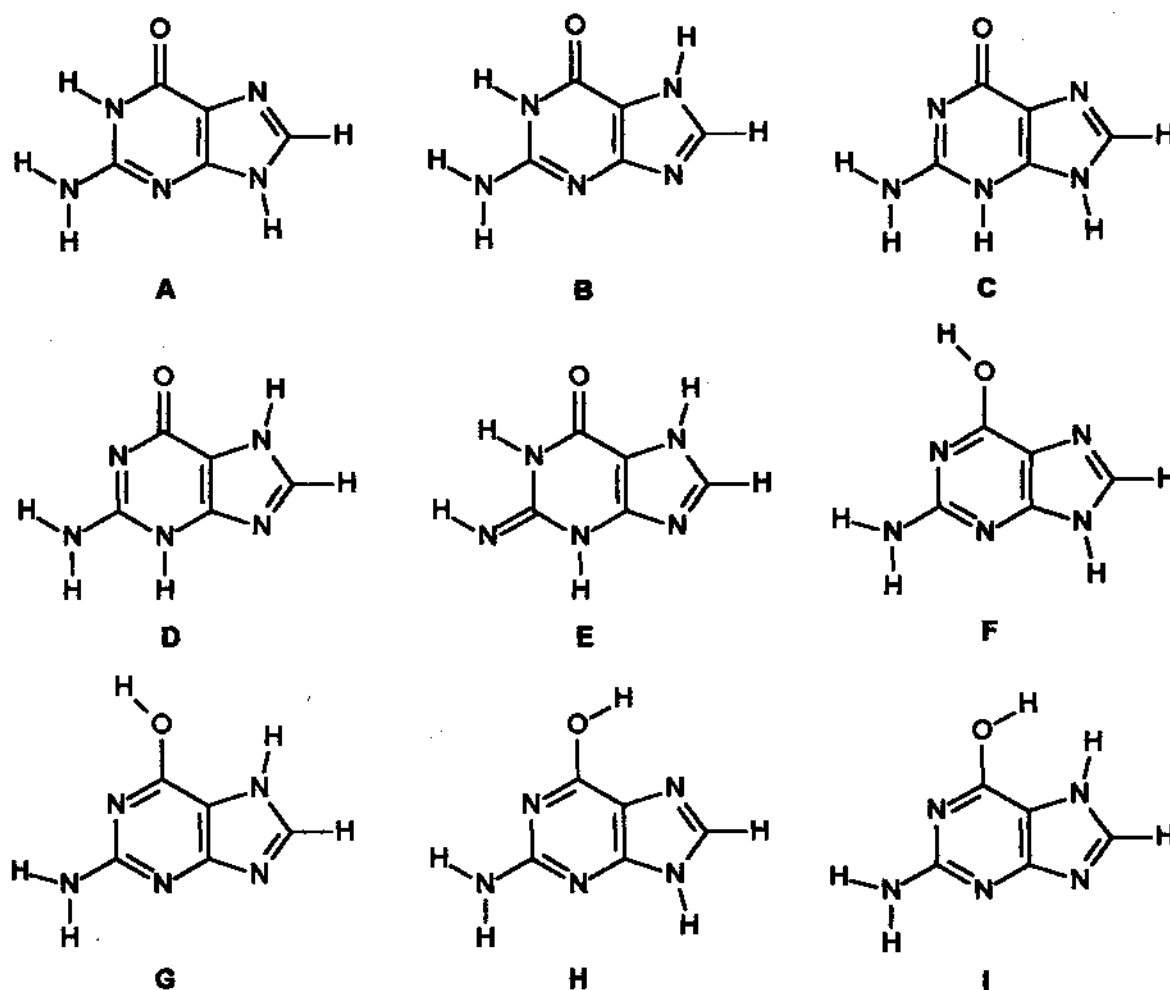


Figura 11 - Tautômeros da guanina mais frequentemente citados como sendo os mais estáveis.

Sigula e Buda (1983), Katritzky e Karelson (1991) e Fabian (1991) utilizaram os métodos semiempíricos MNDO, AM1 e PM3 no cálculo da energia de vários tautômeros da guanina (ver Tabela 2). Houve boa concordância nos resultados apresentados pelos métodos AM1 e PM3, considerados mais evoluídos. Estes métodos indicaram que os tautômeros A e B são os mais estáveis e, portanto, mais abundantes. O método SCRF-AM1 (SCRF, *self-consistent reaction field*) simula a presença do tautômero em meio aquoso, colocando-o num meio dielétrico polarizável ($\mu = 80$). Da mesma forma que os métodos AM1 e PM3 em fase gasosa ($\mu = 1$), o método SCRF-AM1 indicou serem os tautômeros A e B os de maior estabilidade.

Tabela 2 - Energias relativas dos tautômeros da guanina, apresentados na Figura 11, calculadas pelos métodos semiempíricos MNDO, AM1, PM3 e SCRF-AM1 (SCRF, *self-consistent reaction field*). A coluna *Ordem de energia dos tautômeros* exibe a ordem crescente das energias (1, menor energia; 2; 3; 4; 5, maior energia).

Método	Ordem de energia dos tautômeros, kcal.mol ⁻¹					Referência
	1	2	3	4	5	
MNDO	F -7,2	G -2,9	A 0,0	B 2,9	D 7,6	Sigala & Buda, 1983.
MNDO	F -7,2	G -2,9	A 0,0	B 2,6	D 7,6	Fabian, 1991.
AM1	A 0,0	B 1,5	F 3,3	G 6,5	E 7,2	
PM3	A 0,0	B 0,9	F 1,0	D 2,7	E 3,1	
AM1	A 0,0	B 1,5	F 3,2	G 6,5	E 7,0	Katritzky & Karelson, 1991.
SCRF-AM1 ($\mu = 80$)	A 0,0	B 5,3	F 6,3	C 7,3	G 8,9	

Sabio e colaboradores (1990) utilizaram combinação de métodos semiempírico e *ab initio* no cálculo da energia de 36 tautômeros da guanina. As energias relativas dos tautômeros mais estáveis são mostrados na Tabela 3. Pode-se notar que, à exceção do método STO-3G//AM1, os resultados apresentados na Tabela 3 mostram boa concordância com os resultados em que foram empregados apenas métodos semiempíricos. Os resultados mostraram ser os tautômeros A e B os mais estáveis. O método STO-3G//AM1 apresentou resultados que se assemelham aos produzidos pelo método MNDO.

Tabela 3 - Energias relativas dos tautômeros da guanina, apresentados na Figura 11, calculadas por combinação de métodos semiempírico (AM1) e *ab initio* (Hartree-Fock-Roothaan). A coluna Ordem de energia dos tautômeros exibe a ordem crescente das energias (1, menor energia; 2; 3; 4; 5, maior energia).

Método ^a	Ordem de energia dos tautômeros, kcal.mol ⁻¹					Referência
	1	2	3	4	5	
AM1//AM1	A	B	F	H	G	Sabo et al, 1990.
	0,0	1,5	3,2	4,5	6,4	
STO-3G//AM1	F	H	G	A	I	
	-9,6	-8,4	-4,2	0,0	3,1	
3-21G//AM1	A	B	F	H	E	
	0,0	2,7	3,3	3,8	4,4	
3-21G//3-21G	A	B	F	H	E	
	0,0	2,5	4,9	5,7	7,0	

^a A notação X/Y significa que as geometrias moleculares dos tautômeros de cada linha foram otimizadas pelo método Y e suas energias foram calculadas pelo método X.

Finalmente, Colominas e colaboradores (1996) executaram cálculos *ab initio* de alto nível na determinação das energias de diversos tautômeros da guanina (ver Tabela 4). Foram utilizados três níveis de cálculo *ab initio*: Hartree-Fock; teoria de perturbação de Møller-Plesset, com correções de segunda e terceira ordens para a correlação eletrônica; e teoria do funcional de densidade, utilizando funcional de Becke3-Lee-Yang-Parr (para mais detalhes e referências específicas de cada metodologia, ver a seção **Methods** da referência Colominas et al, 1996). Os cálculos em nível Hartree-Fock indicaram ser o tautômero A o mais estável. Os cálculos que incluíram a correlação eletrônica, através da teoria de perturbação, apresentaram resultados conflitantes. O cálculo em nível MP2 indicou ser o tautômero B o mais estável, enquanto que o cálculo MP4 apontou para o tautômero A. Os cálculos baseados na teoria do funcional de densidade indicaram o tautômero B como sendo o mais estável, seguido de perto por A. Em todos os resultados

apresentados na Tabela 4, a diferença de energia entre os dois tautômeros mais estáveis, calculada pelos diversos métodos, foi inferior a 1 kcal/mol.

Tabela 4 - Energias relativas dos tautômeros da guanina, apresentados na Figura 11, calculadas por método *ab initio* de alto nível. A coluna **Ordem de energia dos tautômeros** exibe a ordem crescente das energias (1, menor energia; 2; 3; 4; 5, maior energia).

Método	Ordem de energia dos tautômeros, kcal.mol ⁻¹					Referência
	1	2	3	4	5	
HF/6-31G(d)//	A	B	F	H	G	Colominas et al, 1996.
HF/6-31G(d)	0,0	0,7	1,5	2,7	6,1	
HF/6-311++G(d,p)//	A	F	B	H	G	
HF/6-31G(d)	0,0	0,2	0,4	1,2	4,4	
MP2/6-311++G(d,p)//	B	F	A	H	G	
HF/6-31G(d)	-0,5	-0,1	0,0	0,5	3,2	
MP4/6-311++G(d,p)//	A	B	F	H	G	
MP2/6-31G(d)	0,0	0,2	1,1	1,8	4,7	
B3LYP(6-31G(d))//	B	A	F	H	G	
MP2/6-31G(d)	-0,1	0,0	1,9	2,8	5,3	
B3LYP(6-311++G(d,p))//	B	A	F	H	G	
MP2/6-31G(d)	-0,5	0,0	1,1	1,9	4,0	

* A notação X/Y significa que as geometrias moleculares dos tautômeros de cada linha foram otimizadas pelo método Y e suas energias foram calculadas pelo método X. HF, Hartree-Fock; MPn, correção de n-ésima ordem da correlação eletrônica de Møller-Plesset; B3LYP, funcional de densidade de Becke3-Lee-Yang-Parr (para mais detalhes, ver a seção **Methods** da referência Colominas et al, 1996).

Os resultados apresentados na Tabela 2, Tabela 3 e Tabela 4 indicam de forma consistente que os tautômeros A e B, apresentados na Figura 11, são os de mais baixa energia e, conseqüentemente, aqueles que devem existir em maior abundância. A maioria desses resultados aponta especificamente para o tautômero A.

3.2.4 Forma tautomérica preferencial da guanina

A natureza computacional do presente trabalho faz com que seja necessária a escolha de estrutura molecular específica para a guanina, ou seja, deve-se optar por uma de suas formas tautoméricas. Isto é necessário para que se possa executar os cálculos

moleculares. Os resultados discutidos na seção anterior indicam o tautômero A como sendo o mais estável. Isto não é surpresa. Há muito estamos familiarizados com a estrutura da guanina. Os livros-texto de bioquímica sempre indicaram o tautômero A como sendo a fórmula estrutural da guanina. A molécula do trifosfato de guanosina, que é incorporada ao DNA replicante, pressupõe o tautômero A da guanina, pois a furanose encontra-se preferencialmente ligada ao átomo N(9).

Portanto, decidiu-se utilizar no presente trabalho a fórmula estrutural do tautômero A da guanina nos cálculos moleculares futuros. Não serão feitas outras considerações sobre a tautomeria da guanina nos resultados que serão obtidos. Deve-se ter em mente, no entanto, que a tautomeria estará sempre presente.

3.3 GEOMETRIA EXPERIMENTAL DA GUANINA

3.3.1 Comprimentos de ligação

A geometria molecular de alguns derivados da guanina foi determinada experimentalmente através de cristalografia de raios-X. O'Brien (1967) determinou a estrutura da 9-etilguanina em complexo de 1:1 com 1-metilcitosina, e em complexo de 1:1 com 1-metil-5-fluorocitosina, com 0,006 Å de resolução. Thewalt e colaboradores determinaram a geometria molecular da guanosina (Thewalt et al, 1970) e da guanina monohidratada (Thewalt et al, 1971), com 0,004 Å de resolução. Destro e colaboradores (1974) determinaram a estrutura da 9-etilguanina, com 0,004 Å de resolução. Os valores médios dos comprimentos de ligação da molécula da guanina, nos compostos citados, são mostrados na Figura 12.

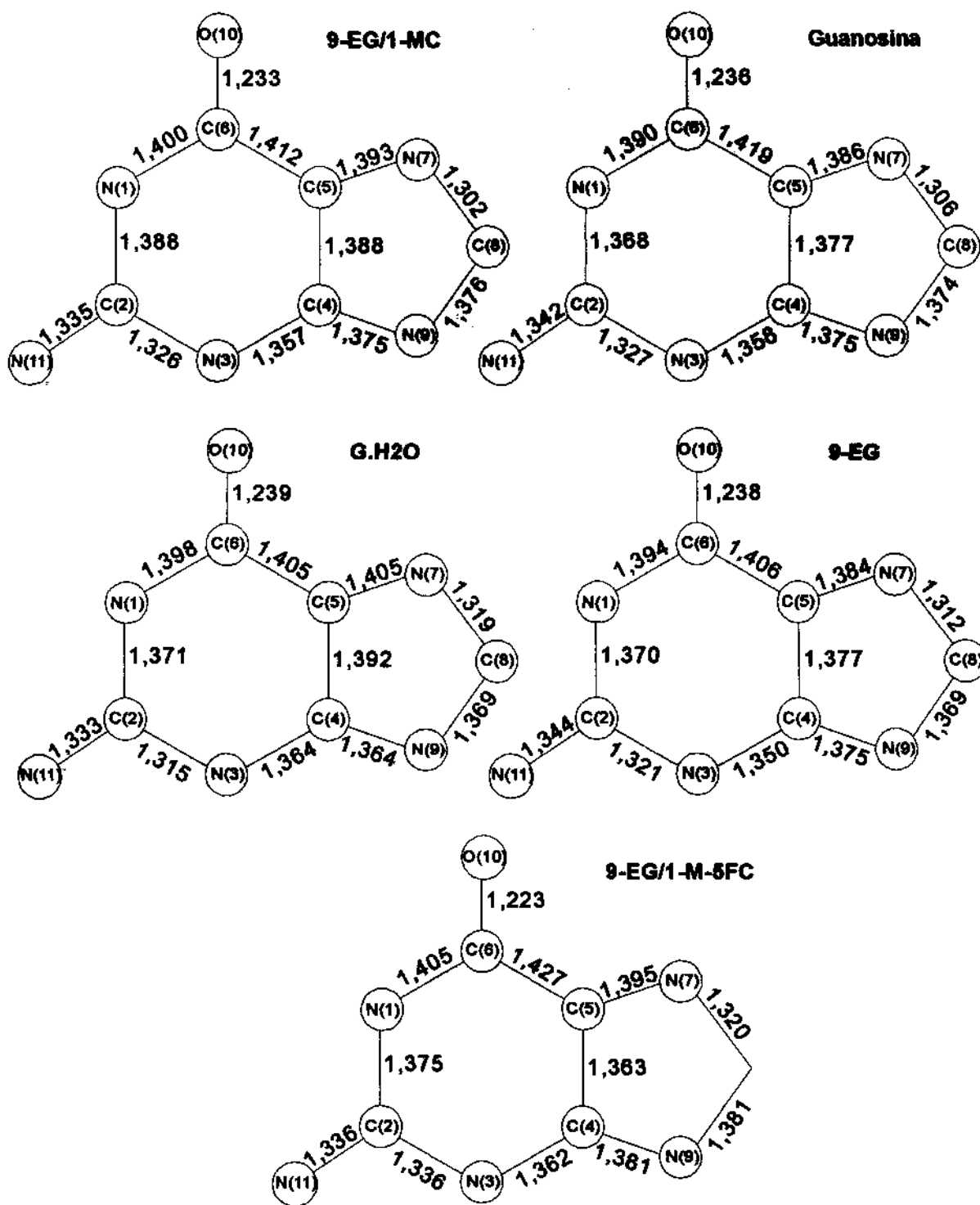


Figura 12 - Comprimentos de ligação (em Å) dos derivados da guanina, obtidos através de cristalografia de raios-X. 9-EG/1-MC e 9-EG/1-M-5-FC, 9-etilguanina em complexo de 1:1 com 1-metilcitosina, e com 1-metil-5-fluorocitosina (O'Brien, 1967); guanosina (Thewalt et al, 1970); G.H₂O, guanina monohidratada (Thewalt et al, 1971); 9-EG, 9-etilguanina (Destro et al, 1974).

A Figura 13 mostra a média aritmética dos comprimentos de ligação apresentados na Figura 12. Os números que aparecem entre parênteses correspondem aos milésimos de

unidade do desvio-padrão da média. Dessa forma, o valor 1,374 (8), que corresponde à média dos comprimentos da ligação N(1)–C(2), deve ser lido como 1,374 ($\pm 0,008$), em que 0,008 é o desvio-padrão da média. Os comprimentos de ligação da guanina apresentam desvio-padrão médio de 0,007 Å.

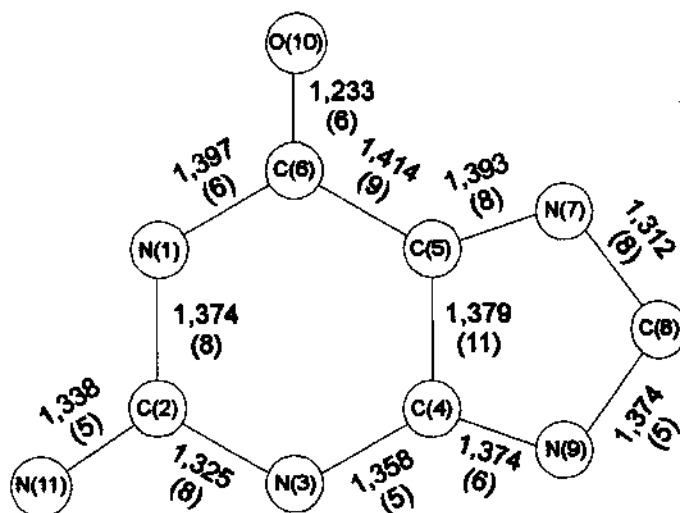


Figura 13 - Média aritmética dos comprimentos de ligação (em Å) dos derivados da guanina, apresentados na Figura 12. Os números entre parênteses correspondem aos milésimos de unidade do desvio-padrão da média.

3.3.2 Ângulos de ligação

Os ângulos de ligação dos derivados da guanina, determinados experimentalmente, são mostrados na Figura 14.

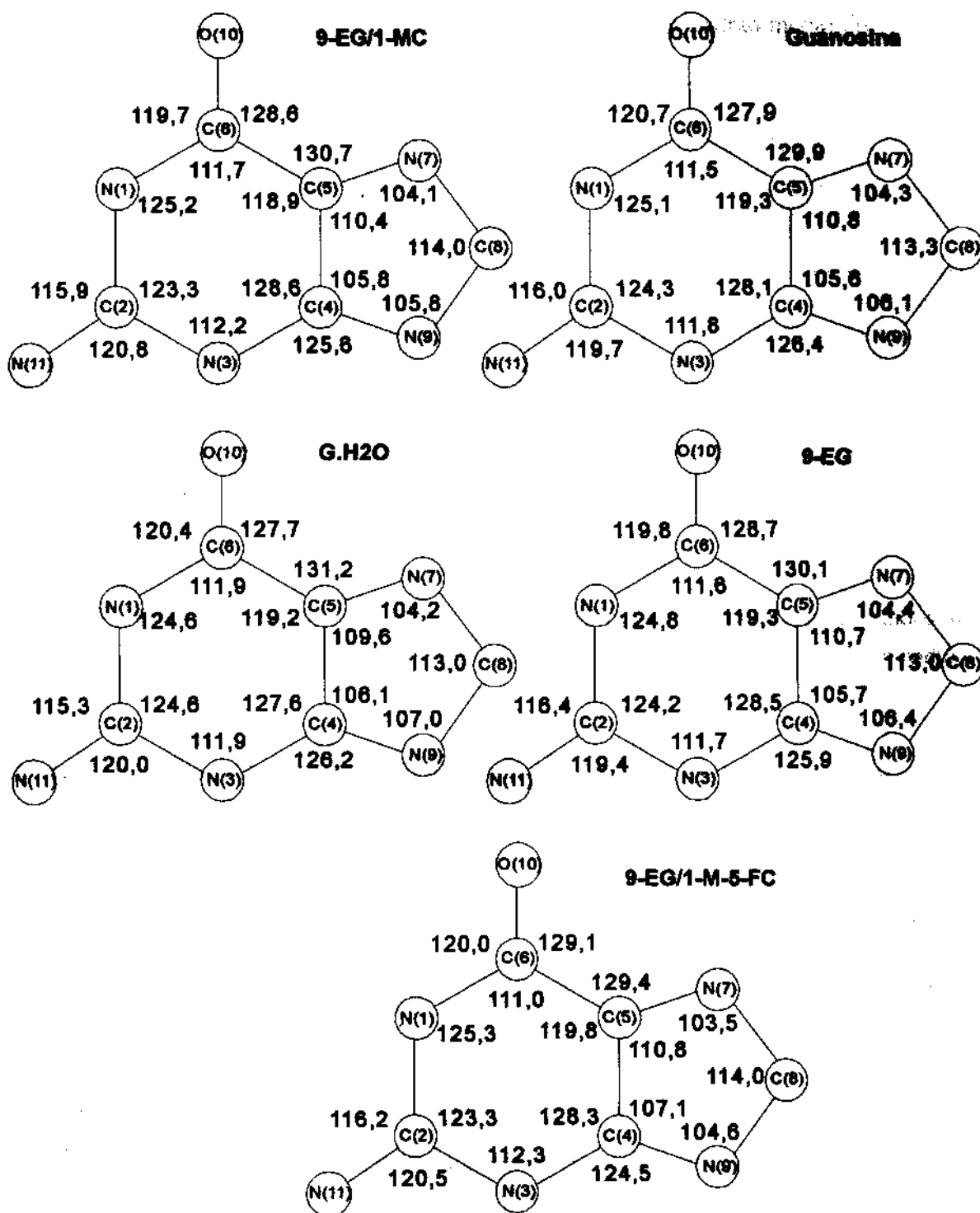


Figura 14 - Ângulos de ligação, em graus, dos derivados da guanina, obtidos através de cristalografia de raios-X. 9-EG/1-MC e 9-EG/1-M-5-FC, 9-etilguanina em complexo de 1:1 com 1-metilcitosina, e com 1-metil-5-fluorocitosina (O'Brien, 1967); guanosina (Thewalt et al, 1970); G.H₂O, guanina monohidratada (Thewalt et al, 1971); 9-EG, 9-etilguanina (Destro et al, 1974).

A média aritmética dos valores dos ângulos de ligação, apresentados na Figura 14, são mostrados na Figura 15. De forma semelhante aos comprimentos de ligação, os

ângulos de ligação apresentam pequena variabilidade. O desvio-padrão médio dos ângulos de ligação é $0,5^\circ$.

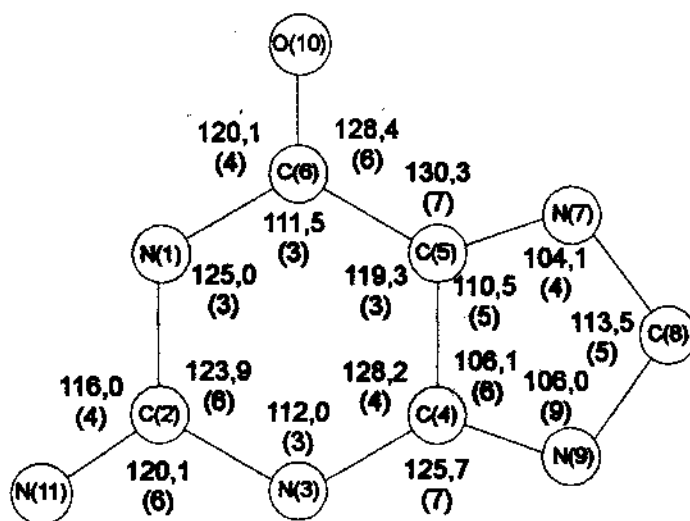


Figura 15 - Média aritmética dos ângulos de ligação dos derivados da guanina, apresentados na Figura 14. Os números entre parênteses correspondem aos décimos de unidade do desvio-padrão da média.

3.4 PADRONIZAÇÃO DA METODOLOGIA DE CÁLCULO MOLECULAR

Escolher a metodologia de cálculo molecular não é tarefa fácil, embora existam basicamente apenas dois métodos gerais de cálculo molecular (Leach, 1996): métodos baseados na mecânica molecular e métodos baseados na teoria do orbital molecular. Estes, por sua vez, podem ser classificados em métodos semiempíricos e em métodos *ab initio*. Há vantagens e desvantagens na utilização de cada um desses métodos e a escolha apropriada deve ser baseada na relação custo/benefício do cálculo. No presente contexto, o custo refere-se ao esforço computacional requerido para a execução do cálculo, que em geral é medido pelo tempo de computação. Cálculos baseados na mecânica molecular são processados mais rapidamente, porém há necessidade da correta parametrização do sistema molecular, o que nem sempre está disponível. Além disso, a mecânica molecular não inclui os elétrons explicitamente. Isto impede o estudo de fenômenos que dependam da densidade e/ou do fluxo de elétrons nas espécies químicas consideradas. Os métodos

baseados na teoria do orbital molecular consideram os elétrons explicitamente, mas são computacionalmente dispendiosos. Para o cálculo de espécies químicas estruturalmente simples, é possível a utilização de método *ab initio* em que a representação dos orbitais atômicos pode ser feita através de extensos conjuntos de funções gaussianas. No entanto, a complexidade do cálculo *ab initio* aumenta exponencialmente com o número dessas funções. Isto torna os métodos *ab initio* restritos às espécies químicas mais simples. As espécies químicas de interesse biológico não são, em geral, estruturalmente simples. A execução de cálculo molecular *ab initio* nessas espécies fica restrita à utilização dos conjuntos de bases mais simples (STO-3G, 3-21G, 4-31G, etc.), que limita a representação do sistema molecular e produz resultados distantes dos valores observados experimentalmente. Restam, pois, os métodos semiempíricos de cálculo de orbital molecular. Estes utilizam parâmetros experimentais que possibilitam grande economia de computação em relação aos métodos *ab initio*. Além disso, a correlação eletrônica está incluída nos métodos semiempíricos, em alguma extensão. Os métodos semiempíricos mais populares, AM1 (Dewar et al, 1985) e PM3 (Stewart, 1989a e 1989b), foram parametrizados para reproduzirem a geometria de equilíbrio e o calor de formação. No entanto, muitas outras propriedades moleculares podem ser ao menos estimadas por esses métodos (Stewart, 1990b). A última geração de método semiempírico, SAM1 – *Semi-Ab-initio Model 1* – (Dewar et al, 1993), parece que ainda não é usado de forma disseminada (talvez pelo fato de que o programa não seja distribuído gratuitamente, como foram AM1 e PM3 durante algum tempo). No momento da escolha do método a ser utilizado no cálculo molecular, todos esses fatores devem ser considerados.

Para auxiliar na escolha do método a ser utilizado no estudo das características conformacionais das N^2 -fenilguaninas, decidiu-se fazer um experimento computacional para tentar reproduzir a geometria molecular da guanina, cujos valores experimentais

foram estudados na Seção 3.3 (“Geometria experimental da guanina”, pag.97). Infelizmente não foi possível utilizar nenhum dos derivados da N^2 -fenilguanina como referência, pelo fato de não haver medidas experimentais da geometria molecular disponíveis. Como as moléculas dos derivados da N^2 -fenilguanina são basicamente formadas pelos anéis guanina e fenila, a reprodução correta da geometria da guanina é passo importante na reprodução da geometria das N^2 -fenilguaninas, uma vez que o anel guanina é a parte mais complexa dessas moléculas. Utilizou-se como sistema de referência os valores médios das medidas experimentais apresentadas na Figura 13 e Figura 15. Poder-se-ia, por outro lado, optar por algum dos cinco sistemas apresentados na Figura 12 e Figura 14 para ser utilizado como referência. Porém, devido à variação natural de alguns comprimentos e ângulos de ligação nas espécies estudadas e também ao fato de não se dispor da geometria molecular experimental da guanina pura, optou-se por utilizar a média dos comprimentos e ângulos de ligação como sistema de referência.

Foram testados o método de mecânica molecular MM2 (Allinger, 1977), os métodos semiempíricos AM1 e PM3, implementados no programa MOPAC (Stewart, 1990a e 1990b), e o método de cálculo de orbital molecular *ab initio* em nível Hartree-Fock, utilizando-se o conjunto de bases STO-3G, implementado no programa GAMESS (Schmidt et al, 1993). Fez-se, quando possível, duas modalidades de cálculo para cada um desses métodos: cálculo em fase gasosa, $\epsilon = 1$ (ϵ , constante dielétrica), e em fase aquosa, $\epsilon = 78,4$, utilizando-se o método PCM (*polarisable continuum method*), de Miertus e colaboradores (1981).

Nos cálculos apresentados nas seções seguintes, o modelo molecular da guanina foi construído através do programa HyperChem (1995). O programa HyperChem possui ambiente de edição molecular que torna possível a construção de moléculas através da escolha apropriada dos átomos e a subsequente formação das ligações químicas com o

auxílio do *mouse*. Os modelos moleculares gráficos construídos dessa maneira são planos e há necessidade de torná-los tridimensionais. Para isso, aplicou-se o recurso *Model Builder*, que utiliza comprimentos e ângulos de ligação-padrão e transfere estes valores para o modelo. Para efeito de refinamento, aplicou-se o campo de força MM+, interno ao programa HyperChem, projetado para moléculas orgânicas e que é similar ao campo de força MM2. As coordenadas atômicas da guanina foram exportadas como arquivo do tipo ZMT (coordenadas internas), que foi submetido aos métodos AM1 e PM3. O arquivo ZMT foi submetido ao programa Babel (Walters & Stahl, 1996), que o converteu para os formatos de entrada dos programas MM2 e GAMESS.

3.4.1 Comprimentos de ligação

A Tabela 5 mostra os comprimentos de ligação da guanina, calculados pelas diversas metodologias citadas anteriormente. Estes valores podem ser comparados aos valores médios dos comprimentos de ligação da guanina, apresentados na Figura 13. Em suas duas últimas linhas, a Tabela 5 mostra, respectivamente, a média dos desvios absolutos e os correspondentes desvios relativos. De acordo com os resultados apresentados, os comprimentos de ligação da guanina foram melhor reproduzidos pelo método MM2. Os demais métodos apresentaram equilíbrio na capacidade de reproduzirem o valor experimental de referência, com ligeira vantagem para o método HF – STO-3G em fase aquosa.

3.4.2 Ângulos de ligação

Os valores calculados dos ângulos de ligação, bem como os valores experimentais de referência, são mostrados na Tabela 6. Em suas duas últimas linhas, a Tabela 6 mostra, respectivamente, a média dos desvios absolutos e os correspondentes desvios relativos. Observa-se que o método HF-STO-3G em fase aquosa apresentou os menores desvios em

relação ao valor de referência. Os métodos MM2 e PM3 apresentaram as piores performances. O provável motivo de o método MM2 ter apresentado resultado ruim para os ângulos de ligação é o fato de nem todas as constantes de força referentes às deformações angulares da guanina estarem disponíveis.

Tabela 5 - Comprimentos de ligação da guanina. Comparem-se os diversos valores calculados com o valor experimental de referência.

Ligação	Comprimento de ligação/Método de cálculo, Å							Exp. ^a
	$\epsilon = 1$				$\epsilon = 78,4$			
	MM2	AM1	PM3	STO-3G	AM1	PM3	STO-3G	
N(1)-C(2)	1,35	1,41	1,42	1,40	1,42	1,42	1,40	1,37
N(1)-C(6)	1,36	1,42	1,46	1,46	1,40	1,42	1,45	1,40
N(1)-H(12)	1,01	1,00	1,00	1,02	1,00	1,00	1,03	–
C(2)-N(3)	1,31	1,36	1,35	1,32	1,38	1,38	1,33	1,33
C(2)-N(11)	1,37	1,39	1,39	1,39	1,37	1,35	1,39	1,34
N(3)-C(4)	1,41	1,38	1,40	1,41	1,37	1,39	1,42	1,36
C(4)-C(5)	1,39	1,44	1,41	1,37	1,45	1,41	1,37	1,38
C(4)-C(9)	1,37	1,40	1,39	1,38	1,39	1,40	1,38	1,37
C(5)-C(6)	1,47	1,45	1,45	1,48	1,44	1,43	1,47	1,41
C(5)-N(7)	1,37	1,40	1,40	1,41	1,40	1,41	1,42	1,39
C(6)-O(10)	1,22	1,24	1,22	1,22	1,26	1,25	1,22	1,23
N(7)-C(8)	1,31	1,35	1,34	1,31	1,35	1,35	1,31	1,31
C(8)-N(9)	1,37	1,42	1,41	1,40	1,41	1,40	1,40	1,37
C(8)-H(13)	1,10	1,10	1,09	1,08	1,10	1,10	1,09	–
N(9)-H(14)	1,05	0,99	0,99	1,02	0,99	0,99	1,02	–
N(11)-H(15)	1,05	0,99	0,99	1,01	1,00	0,99	1,02	–
N(11)-H(16)	1,05	0,99	0,99	1,01	1,00	0,99	1,01	–
Média desvio absoluto ^b	0,022	0,033	0,033	0,029	0,032	0,029	0,028	
Desvio relativo ^c	0,00	0,50	0,50	0,32	0,45	0,32	0,27	

^a Média dos comprimentos de ligação (ver Figura 13) dos derivados da guanina apresentados na Figura 12; ^b Apenas os comprimentos de ligação com valores de referência experimental foram incluídos; ^c $(MDA_x - MDA_{\text{menor valor}})/MDA_{\text{menor valor}}$, MDA = média dos desvios absolutos.

Tabela 6 - Ângulos de ligação da guanina. Comparem-se os diversos valores calculados com o valor experimental de referência.

Ângulo	Ângulo de ligação/Método de cálculo, graus							Exp. ^a
	ε = 1				ε = 78,4			
	MM2	AM1	PM3	STO-3G	AM1	PM3	STO-3G	
C(2)-N(1)-C(6)	128,1	122,8	122,3	125,2	122,4	121,9	125,2	125,0
C(2)-N(1)-H(12)	113,8	120,3	118,7	117,9	119,6	118,5	118,7	—
C(6)-N(1)-H(12)	118,1	116,8	119,0	116,8	118,0	119,6	116,1	—
N(1)-C(2)-N(3)	120,8	124,5	124,2	126,1	123,6	123,4	125,6	123,9
N(1)-C(2)-N(11)	115,9	117,7	119,8	115,3	118,0	120,1	116,0	116,0
N(3)-C(2)-N(11)	123,2	117,7	116,1	118,6	118,5	116,5	118,5	120,1
C(2)-N(3)-C(4)	117,5	114,0	114,6	110,1	114,6	114,2	110,7	112,0
N(3)-C(4)-C(5)	122,7	125,7	125,7	129,8	125,1	125,3	128,6	128,2
N(3)-C(4)-N(9)	133,1	128,7	127,7	124,6	129,3	128,0	125,2	125,7
C(5)-C(4)-N(9)	104,2	105,6	106,6	105,5	105,6	106,7	106,0	106,1
C(4)-C(5)-C(6)	118,8	119,1	120,4	120,1	118,7	120,2	120,5	119,3
C(4)-C(5)-N(7)	111,3	110,3	108,6	111,4	109,8	108,4	111,0	110,5
C(6)-C(5)-N(7)	129,8	130,6	131,0	128,5	131,5	131,5	128,5	130,3
N(1)-C(6)-C(5)	112,0	113,9	112,8	106,6	115,6	114,9	109,3	111,5
N(1)-C(6)-O(10)	121,7	117,3	115,2	120,0	118,3	116,5	119,3	120,1
C(5)-C(6)-O(10)	126,3	128,8	132,0	131,3	126,1	128,6	131,4	128,4
C(5)-N(7)-C(8)	104,9	104,9	108,0	103,9	105,1	107,3	103,8	104,1
N(7)-C(8)-N(9)	111,9	113,1	109,4	112,6	113,1	109,9	112,8	113,5
N(7)-C(8)-H(13)	126,1	125,6	125,1	126,0	126,0	124,8	125,4	—
N(9)-C(8)-H(13)	121,9	121,3	125,4	121,4	120,9	125,2	121,7	—
C(4)-N(9)-C(8)	107,6	106,0	107,3	106,5	106,3	107,6	106,4	106,0
C(4)-N(9)-H(14)	126,1	126,5	126,5	126,1	127,1	126,6	128,3	—
C(8)-N(9)-H(14)	126,3	127,5	126,2	127,4	126,6	125,8	125,4	—
C(2)-N(11)-H(16)	119,1	119,2	120,1	118,6	120,3	120,7	118,9	—
C(2)-N(11)-H(15)	120,1	121,0	120,7	122,0	121,2	121,0	122,5	—
H(15)-N(11)-H(16)	120,8	119,8	119,3	119,5	118,5	118,3	118,6	—
Média desvio absoluto ^b	2,34	1,32	2,42	1,22	1,69	2,30	0,99	
Desvio relativo ^c	1,36	0,33	1,44	0,23	0,71	1,32	0,00	

^a Média dos ângulos de ligação (ver Figura 15) dos derivados da guanina apresentados na Figura 14; ^b Apenas os ângulos de ligação com valores de referência experimental foram incluídos; ^c $(MDA_x - MDA_{\text{menor valor}})/MDA_{\text{menor valor}}$, MDA = média dos desvios absolutos.

3.4.3 Método de escolha

Antes da escolha do método de cálculo, é preciso que se diga que na construção dos mapas conformacionais será necessária a execução de 1728 cálculos de otimização de geometria molecular (ver a Seção 3.5, “Mapas conformacionais das N^2 -fenilguaninas”, para os detalhes). Os compostos que serão submetidos aos cálculos têm, em média, 19 átomos pesados (C, N e O) e 10 átomos de hidrogênio.

A ordem decrescente de qualidade dos resultados apresentados na Tabela 5 e Tabela 6 é HF-STO-3G ($\epsilon = 78,4$), 0,27 > HF-STO-3G ($\epsilon = 1$), 0,55 > AM1 ($\epsilon = 1$), 0,83 > AM1 ($\epsilon = 78,4$), 1,16 > MM2, 1,36 > PM3 ($\epsilon = 78,4$), 1,64 > PM3 ($\epsilon = 1$), 1,94¹. Baseando-se apenas nos resultados apresentados na Tabela 5 e Tabela 6, a escolha natural seria a utilização do método HF-STO-3G ($\epsilon = 78,4$). Mas, como foi dito anteriormente, a escolha apropriada da metodologia deve ser baseada na relação custo/benefício do cálculo e a Tabela 7 mostra a dimensão do custo de cada metodologia. Analisando-se a Tabela 7 percebe-se que a escolha do método HF-STO-3G ($\epsilon = 1$ ou $\epsilon = 78,4$) fica inviabilizada, especialmente se o objetivo do cálculo é conhecer a forma geral da superfície de potencial da rotação dos anéis fenila e guanina. A próxima opção é o método AM1. Este parece ser adequado à presente necessidade, pois a qualidade de reprodução da geometria da guanina foi apenas ligeiramente inferior ao resultado apresentado pelo método HF-STO-3G ($\epsilon = 1$). Utilizando-se o método AM1 o tempo de CPU estimado para a execução dos cálculos necessários à construção da superfície de potencial é razoável, ou seja, apenas alguns dias. Portanto, decidiu-se pela utilização do método AM1.

¹ Na determinação da ordem de qualidade dos métodos de cálculo molecular, os desvios relativos de cada método, que aparecem na última linha das Tabelas 3.4 e 3.5, foram somados e o resultado da soma é mostrado juntamente com a ordem de qualidade. A melhor qualidade foi associada à menor soma.

Tabela 7 - Tempo de CPU estimado para a execução dos cálculos de otimização de geometria molecular necessários à construção das superfícies de potencial de rotação dos anéis fenila e guanina das N^2 -fenilguaninas (ver Seção 3.5, "Mapas conformacionais das N^2 -fenilguaninas", para maiores detalhes). Os valores apresentados referem-se ao composto 15 (3'-H, 4'-H, Tabela 1, pag 56), que é o menor da série.

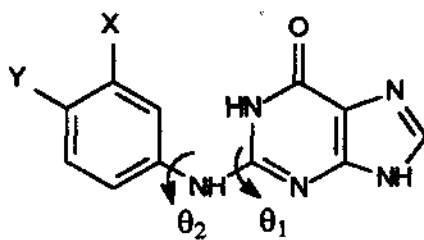
Tempo de CPU ^a	Método				
	$\epsilon = 1$			$\epsilon = 78,4$	
	MM2	AM1/PM3	STO-3G	AM1/PM3	STO-3G
Tempo médio de cada passo de otimização de geometria ^b	0,34"	4,5"	1,1'	1,4'	1,3h
Tempo total estimado de uma otimização de geometria com 50 passos	17"	3,8'	54,2'	1,2h	2,6d
Tempo estimado de 1728 otimizações de geometria de 50 passos cada	8,2h	4,5d	65d	85d	12,3a

^a Medido em microcomputador com processador Pentium II Intel 233 MHz com 128 Mb de memória RAM, performance equivalente a 141 MFLOPS (490 MIPS) e ambiente Windows 95; ^b a, intervalo de tempo em anos; d, dias; h, horas; ', minutos e; ", segundos.

3.5 MAPAS CONFORMACIONAIS DAS N^2 -FENILGUANINAS

3.5.1 Introdução

Os derivados da N^2 -fenilguanina possuem dois conjuntos de anéis, fenila e guanina, que apresentam apreciável rigidez às forças de torção. No entanto, a existência de duas ligações simples do tipo sigma entre esses anéis permite o movimento rotacional dos mesmos (ângulos θ_1 e θ_2 , Figura 16), gerando muitas possibilidades conformacionais. Portanto, pode-se esperar grande variedade de conformeros rotacionais nas moléculas dos derivados da N^2 -fenilguanina.



N²-Fenil-(substituído)-guanina

Figura 16 - Estrutura molecular dos derivados da N²-fenilguanina. X e Y são substituintes diversos (ver Tabela 1, pag. 56). Os ângulos torsionais θ_1 e θ_2 fornecem graus de liberdade rotacionais capazes de gerar grande número de confómeros.

Devido ao fato de trabalhar-se com grupo numeroso de derivados da N²-fenilguanina, optou-se por analisar o mapa conformacional de apenas alguns desses derivados. Os critérios para a escolha dos derivados a serem analisados foram:

- (a) derivado com maior barreira rotacional (composto 1 – 3'-CF₃, 4'-H; Tabela 1, pag. 56);
- (b) derivado com menor barreira rotacional (composto 15 – 3'-H, 4'-H);
- (c) derivado com substituintes nas posições meta e para (composto 10 – 3'-Br, 4'-Br) e
- (d) derivado com substituinte ramificado (composto 11 – 3'-n-Prop, 4'-H).

A escolha destes critérios foi arbitrária e procurou ser representativa. Optou-se pela escolha de apenas quatro compostos devido à grande quantidade de cálculo necessário à construção dos mapas conformacionais. Tendo a estrutura dos derivados da N²-fenilguanina dois graus de liberdade rotacional entre os anéis guanina e fenila, e fazendo-se otimizações de geometria molecular a cada 30°, isto representa 144 cálculos de otimização por molécula. Esses cálculos são feitos em triplicata, partindo-se de estruturas ligeiramente modificadas. Assim, tomando-se quatro derivados como amostra, foram necessários 1728 cálculos de otimização de geometria.

3.5.2 Detalhes metodológicos

Fez-se a análise conformacional dos derivados da N^2 -fenilguanina com base na construção de mapas de energia conformacional. Estes são gráficos tridimensionais, ou projeções bidimensionais, que mostram a dependência da energia em função dos ângulos de rotação interna, neste caso θ_1 e θ_2 (Figura 16). Para construírem-se os mapas, variou-se sistematicamente os ângulos torsionais θ_1 e θ_2 e, subsequente, otimizou-se a geometria molecular em fase gasosa. No cálculo da energia molecular, utilizou-se o método semiempírico AM1. Fez-se a construção das moléculas através do programa HyperChem. Fez-se a pré-otimização através do campo de força MM+. A seguir, fez-se com que os ângulos θ_1 e θ_2 adotassem o valor de 0° e o modelo foi exportado no formato do arquivo de entrada do programa AM1 (formato ZMT). O arquivo foi editado de forma que os ângulos θ_1 e θ_2 variem sistematicamente em intervalos de 30° . Assim, iniciando-se o processo com $\theta_1 = 0^\circ$, θ_2 foi sistematicamente variado de 0° a 360° . Fazendo $\theta_1 = 30^\circ$, θ_2 foi novamente variado de 0° a 360° . O processo foi executado até varrer todo o espaço conformacional dos ângulos torsionais θ_1 e θ_2 . Em cada etapa de otimização de geometria molecular, com exceção da primeira ($\theta_1 = 0^\circ$ e $\theta_2 = 0^\circ$), o método AM1 utiliza a geometria do passo anterior como ponto de partida. Dessa forma, a geometria molecular não sofreu alterações maiores do que 30° , em θ_1 ou θ_2 , de um passo de otimização para outro. Os cálculos foram executados em fase gasosa ($\epsilon = 1$). Os ângulos θ_1 e θ_2 foram mantidos congelados durante cada passo de otimização da geometria molecular. Os comprimentos e ângulos de ligação e os demais ângulos torsionais foram deixados livres para que pudessem adotar a conformação local de energia mínima. O processo foi repetido por mais duas vezes, como forma de confirmação. Na primeira repetição, fez-se com que os ângulos θ_1 e θ_2 adotassem valores dois graus a menos (-2° , 28° , 58° , ...) e dois graus a mais (2° ,

32°, 62°, ...) do que no primeiro cálculo, respectivamente. Na segunda repetição, fez-se com que os ângulos θ_1 e θ_2 adotassem valores dois graus a mais (2°, 32°, 62°, ...) e dois graus a menos (-2°, 28°, 58°, ...) do que no primeiro cálculo, respectivamente. Esse procedimento foi adotado para evitar que valores incorretos da energia molecular fossem incorporados aos resultados. Valores incorretos da energia podem decorrer da impossibilidade de atingir a convergência no processo de minimização da energia do cálculo molecular (Stewart, 1990b). Apenas os resultados obtidos no primeiro cálculo foram utilizados na construção dos mapas conformacionais. Os valores das energias obtidas nas duas repetições foram usados apenas como referência. Nos casos em que houve diferenças significativas entre os resultados do primeiro cálculo e os resultados das duas repetições, fez-se nova repetição do cálculo discordante, fosse ele o primeiro cálculo ou uma das repetições. Nos casos em que a discordância persistiu, novo cálculo foi feito, desta vez utilizando-se geometria molecular inicial ligeiramente diferente das três já testadas. O processo foi repetido até que houvesse consistência entre pelo menos três cálculos.

O parâmetro de energia obtido nos cálculos AM1 e utilizado como medida da energia conformacional foi o calor de formação-padrão, ΔH_f° ($T = 298\text{ K}$, $p = 1\text{ atm}$). Embora o calor de formação não meça especificamente a energia conformacional, ele a contém. A diferença entre os calores de formação das conformações A e B de um certo composto pode ser atribuída à diferença de energia conformacional de A em relação à B.

3.5.3 Resultados

A Tabela 8 - Tabela 11 mostram os valores relativos do calor de formação-padrão para cada conformação definida pelos ângulos torsionais θ_1 e θ_2 dos quatro derivados do

grupo de teste. Os gráficos de contorno e as superfícies de potencial correspondentes aos valores da Tabela 8 - Tabela 11 são mostrados na Figura 17 - Figura 20.

Tabela 8 - Valores relativos dos calores de formação dos confórmeros rotacionais do composto 3'-CF₃, 4'-H (composto 1, Tabela 1, pag. 56). O valor absoluto do calor de formação para $\theta_1 = 210$ e $\theta_2 = 330$, que corresponde ao mínimo absoluto de energia, é -70.20 kcal/mol

θ_1	θ_2												
	0	30	60	90	120	150	180	210	240	270	300	330	360
0	1,05	1,17	2,17	2,37	2,29	1,18	1,54	1,18	2,51	2,54	2,33	1,18	1,05
30	0,96	1,92	2,52	3,59	2,71	1,50	2,58	2,26	2,51	2,61	2,63	1,53	0,96
60	5,08	2,04	2,59	3,77	3,29	1,45	1,42	2,34	2,89	4,01	3,01	1,08	5,08
90	4,18	4,23	5,04	6,18	6,15	5,04	4,34	4,46	5,33	6,32	6,22	5,09	4,18
120	0,80	0,61	0,97	0,83	0,51	1,03	2,57	0,97	1,10	0,79	0,17	0,83	0,80
150	2,06	4,28	9,55	0,63	0,51	1,30	2,86	5,75	11,1	0,42	0,09	0,71	2,06
180	1,59	6,50	0,77	1,81	1,08	0,86	3,02	9,17	1,05	1,64	0,88	0,25	1,59
210	2,36	0,84	0,09	0,43	0,72	0,56	4,33	1,27	0,50	0,62	0,59	0,00	2,36
240	2,10	0,70	0,13	0,66	1,03	0,89	2,58	0,99	0,45	0,81	0,94	0,44	2,10
270	4,18	5,09	6,22	6,32	5,32	4,46	4,34	5,04	6,18	6,17	5,04	4,21	4,18
300	3,30	3,13	4,35	4,31	3,73	3,59	3,54	3,07	4,50	4,27	3,63	3,49	3,30
330	2,14	1,53	2,65	2,68	2,52	2,27	1,09	2,72	2,71	2,79	2,52	2,20	2,14
360	1,05	1,17	2,17	2,37	2,29	1,18	1,54	4,34	2,51	2,54	2,33	1,18	1,05

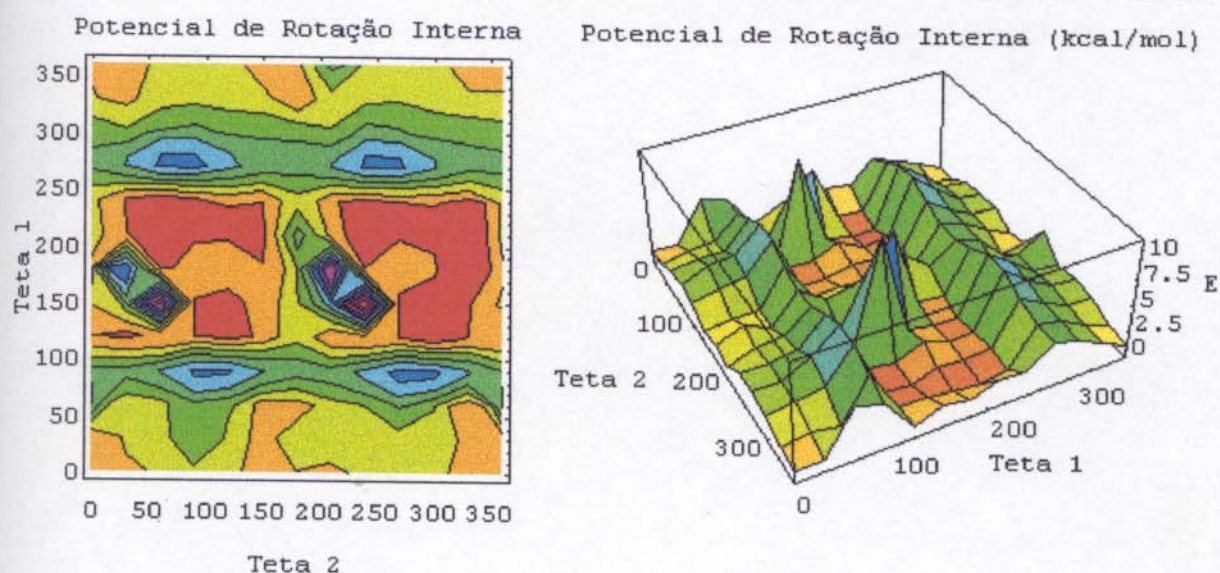


Figura 17 - Mapas conformacionais bidimensional e tridimensional correspondentes aos valores relativos dos calores de formação apresentados (E) na Tabela 8, referentes ao composto 3'-CF₃, 4'-H (composto 1, Tabela 1, pag. 56). O gráfico tridimensional sofreu rotação de aproximadamente 60° no sentido horário em relação ao gráfico bidimensional para que seu relevo fique mais evidente.

Tabela 9 - Valores relativos dos calores de formação dos confórmeros rotacionais do composto 3'-Br, 4'-Br (composto 10, Tabela 1, pag. 56). O valor absoluto do calor de formação para $\theta_1 = 150$ e $\theta_2 = 300$, que corresponde ao mínimo absoluto de energia, é 96.26 kcal/mol

θ_1	θ_2												
	0	30	60	90	120	150	180	210	240	270	300	330	360
0	1,19	3,72	2,39	2,64	2,66	1,19	1,46	1,19	2,46	2,65	2,39	2,18	1,19
30	2,13	2,10	2,52	2,77	2,83	1,81	2,69	2,13	2,54	2,72	2,77	1,60	2,13
60	3,24	3,28	3,57	4,33	4,49	3,19	3,56	3,37	3,60	4,30	4,44	3,20	3,24
90	4,08	4,23	5,16	6,32	6,26	5,07	4,20	4,16	5,10	6,24	6,23	5,12	4,08
120	1,18	0,96	2,94	4,05	2,96	2,44	1,19	1,25	3,32	4,35	3,19	2,21	1,18
150	2,26	0,05	0,66	0,48	0,16	0,78	2,33	5,28	10,8	0,45	0,00	0,49	2,26
180	1,46	0,23	0,77	1,95	0,86	0,51	2,56	8,74	0,86	1,72	0,77	0,24	1,46
210	1,82	0,49	0,01	0,46	0,69	0,29	3,92	0,78	0,17	0,70	0,66	0,00	1,82
240	2,03	0,45	0,07	0,74	1,16	0,80	2,21	0,58	0,21	0,78	1,09	0,61	2,03
270	4,09	5,14	6,23	6,24	5,10	4,17	4,30	5,07	6,26	6,20	4,93	4,01	4,09
300	1,18	0,97	2,94	4,05	2,96	2,31	1,19	1,25	3,30	4,02	3,03	2,21	1,18
330	0,89	1,99	2,77	2,79	2,53	2,12	0,96	2,73	2,82	2,78	2,49	2,09	0,89
360	1,19	3,72	2,39	2,64	2,66	1,19	1,46	1,19	2,46	2,65	2,39	2,18	1,19

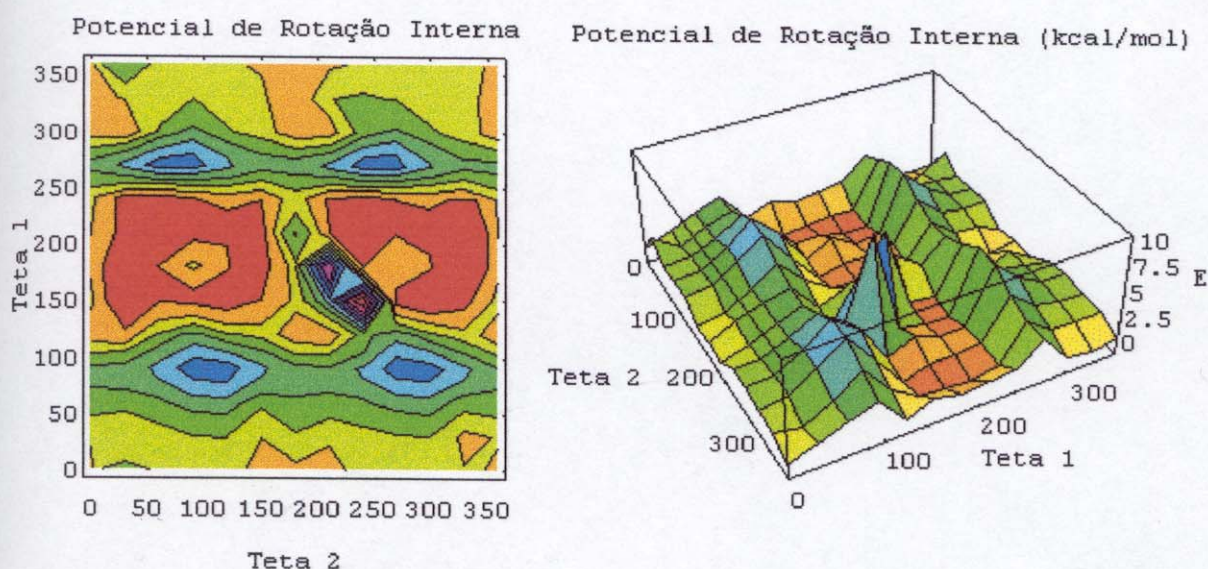


Figura 18 - Mapas conformacionais bidimensional e tridimensional correspondentes aos valores relativos dos calores de formação (E) apresentados na Tabela 9, referentes ao composto 3'-Br, 4'-Br (composto 10, Tabela 1, pag. 56). O gráfico tridimensional sofreu rotação de aproximadamente 60° no sentido horário em relação ao gráfico bidimensional para que seu relevo fique mais evidente.

Tabela 10 - Valores relativos dos calores de formação dos conformêros rotacionais do composto 3'-n-Pr, 4'-H (composto 11, Tabela 1, pag. 56). O valor absoluto do calor de formação para $\theta_1 = 150$ e $\theta_2 = 30$, que corresponde ao mínimo absoluto de energia, é 63.78 kcal/mol

θ_1	θ_2												
	0	30	60	90	120	150	180	210	240	270	300	330	360
0	2,18	4,95	3,33	3,49	3,21	2,16	2,57	2,14	3,23	3,50	3,34	2,24	2,18
30	3,16	3,34	3,48	3,54	3,48	2,48	3,69	3,33	3,48	3,53	3,50	2,50	3,16
60	4,27	4,55	4,51	5,03	5,06	3,95	4,50	4,55	4,51	4,99	5,07	4,07	4,27
90	5,07	5,17	5,86	6,86	6,74	5,72	5,08	5,18	5,91	6,84	6,74	5,84	5,07
120	4,26	4,01	5,08	1,00	0,58	1,18	2,67	0,82	1,23	1,14	0,65	1,31	4,26
150	2,44	0,00	0,54	0,38	0,21	1,08	2,76	5,63	11,1	0,45	0,27	1,08	2,44
180	2,01	0,42	0,76	1,50	0,66	0,40	2,54	8,91	0,68	1,51	0,76	0,40	2,01
210	2,58	1,08	0,27	0,45	0,50	0,05	3,73	1,08	0,22	0,40	0,56	0,01	2,58
240	4,91	6,16	7,41	1,20	1,31	0,82	2,66	1,18	0,59	1,02	1,27	0,82	4,91
270	5,09	5,83	6,74	6,82	5,90	5,17	5,07	5,72	6,74	6,88	5,88	5,19	5,09
300	2,29	2,17	4,08	5,03	3,76	3,05	2,17	2,36	4,38	5,06	3,88	5,16	2,29
330	2,13	3,25	3,49	3,61	3,46	2,96	2,15	3,96	3,47	3,64	3,48	3,35	2,13
360	2,18	4,95	3,33	3,49	3,21	2,16	2,57	2,14	3,23	3,50	3,34	2,24	2,18

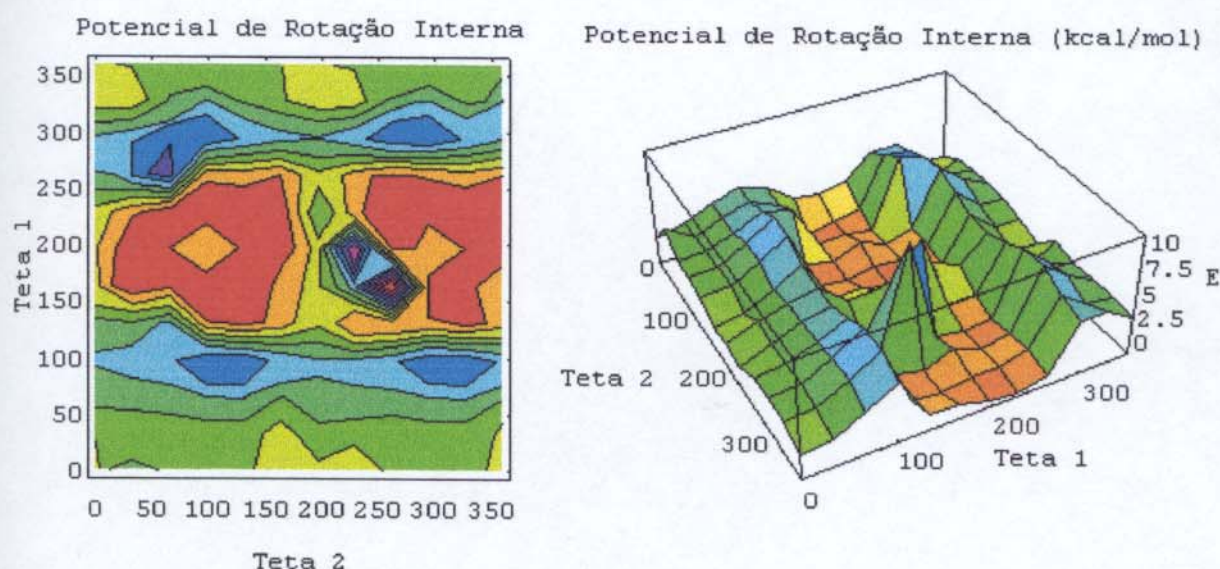


Figura 19 - Mapas conformacionais bidimensional e tridimensional correspondentes aos valores relativos dos calores de formação (E) apresentados na Tabela 10, referentes ao composto 3'-n-Prop, 4'-H (composto 11, Tabela 1, pag. 56). O gráfico tridimensional sofreu rotação de aproximadamente 60° no sentido horário em relação ao gráfico bidimensional para que seu relevo fique mais evidente.

Tabela 11 - Valores relativos dos calores de formação dos conformêros rotacionais do composto 3'-H, 4'-H (composto 15, Tabela 1, pag. 56). O valor absoluto do calor de formação para $\theta_1 = 210$ e $\theta_2 = 270$, que corresponde ao mínimo absoluto de energia, é 84.25 kcal/mol

θ_1	θ_2												
	0	30	60	90	120	150	180	210	240	270	300	330	360
0	4,07	3,04	1,66	2,58	3,17	2,74	4,07	2,75	1,71	2,67	3,19	3,03	4,07
30	4,01	2,55	2,03	3,00	3,92	3,69	4,00	2,20	2,16	3,10	4,06	3,85	4,01
60	4,29	4,52	5,00	4,26	3,65	4,46	4,29	4,50	4,98	4,17	3,75	4,50	4,29
90	1,59	1,45	1,56	2,00	2,45	2,48	1,49	1,34	1,60	2,00	2,47	2,49	1,59
120	0,63	1,97	2,12	0,63	1,24	0,67	0,69	2,16	1,78	0,66	1,16	0,68	0,63
150	2,35	5,08	1,16	0,01	0,65	0,23	2,70	5,00	0,68	0,11	0,52	0,11	2,35
180	6,33	2,31	0,26	1,40	1,30	2,09	6,31	1,82	0,30	1,42	1,30	2,10	6,33
210	2,32	0,11	0,52	0,09	0,58	4,91	2,33	0,10	0,54	0,00	1,16	5,08	2,32
240	0,62	0,69	1,15	0,66	1,79	2,15	0,69	0,67	1,20	0,65	1,79	2,10	0,62
270	5,26	4,95	4,89	5,81	6,67	6,41	5,18	4,86	4,94	5,92	6,69	6,60	5,26
300	4,27	4,49	3,75	4,14	4,98	4,52	4,29	4,45	3,65	4,27	5,00	4,52	4,27
330	4,01	4,14	2,34	2,78	3,43	3,25	4,02	3,69	2,22	2,88	3,43	3,18	4,01
360	4,07	3,04	1,66	2,58	3,17	2,74	4,07	2,75	1,71	2,67	3,19	3,03	4,07

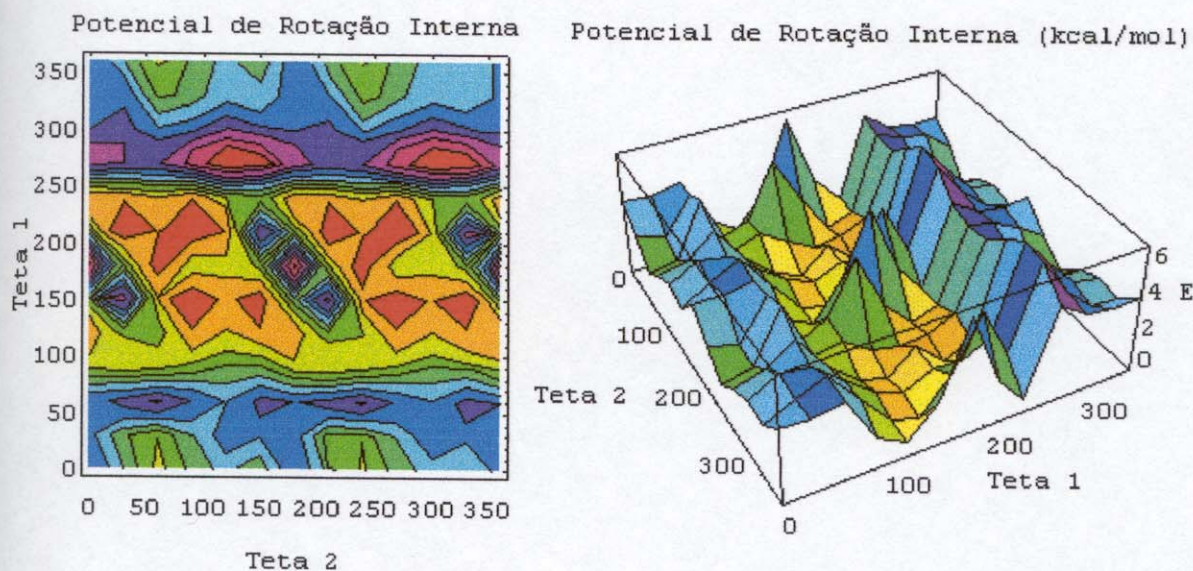


Figura 20 - Mapas conformacionais bidimensional e tridimensional correspondentes aos valores relativos dos calores de formação (E) apresentados na Tabela 11, referentes ao composto 3'-H, 4'-H (composto 15, Tabela 1, pag. 56). O gráfico tridimensional sofreu rotação de aproximadamente 60° no sentido horário em relação ao gráfico bidimensional para que seu relevo fique mais evidente.

A análise das superfícies de potencial mostra que as conformações de baixa energia das N^2 -fenilguaninas estão localizadas na região definida por $100^\circ \leq \theta_1 \leq 250^\circ$ e $90^\circ \leq \theta_2 \leq 140^\circ$ e $250^\circ \leq \theta_2 \leq 320^\circ$. A sobreposição dos mapas conformacionais sugere a existência de quatro regiões de baixa energia, que foram denominadas A, B, C e D. A Figura 21 mostra essas regiões de maneira esquemática. Devido ao fato de as quatro moléculas tomadas como exemplo para a análise conformacional serem bem variadas do ponto de vista dos substituintes presentes no anel fenila, acredita-se que os mapas conformacionais dos demais derivados da N^2 -fenilguanina não sejam apreciavelmente diferentes dos mapas apresentados na Figura 17 - Figura 20.

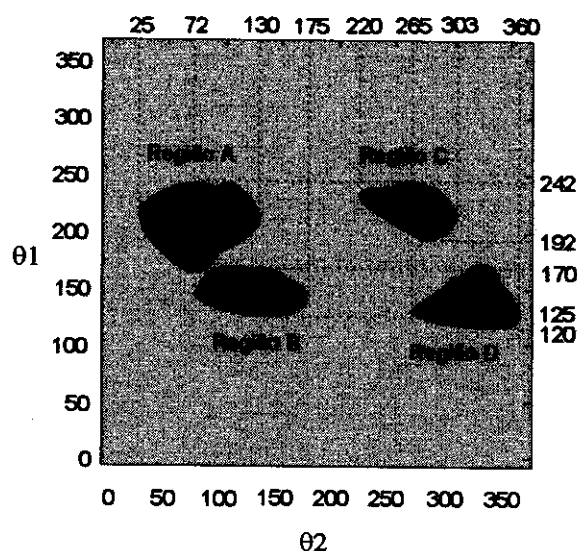


Figura 21 - Mapa conformacional das N^2 -fenilguaninas. As áreas destacadas correspondem às regiões de baixa energia conformacional.

Os resultados da presente análise conformacional serão a base da escolha da geometria molecular a ser adotada no cálculo de diversas propriedades físico-químicas. Essas propriedades serão utilizadas no estudo quantitativo das relações entre estrutura química e atividade biológica, utilizando-se parâmetros calculados. No Capítulo 6 serão feitas as devidas considerações sobre a escolha da geometria molecular a ser utilizada nos cálculos das propriedades físico-químicas.

3.6 GEOMETRIA MOLECULAR DAS $\underline{N}^2$ -FENILGUANINAS

Nos cálculos executados na Seção 3.5, os ângulos torsionais θ_1 e θ_2 foram mantidos fixos enquanto o restante da molécula procurou adaptar-se aos seus valores. Enquanto isto é aceitável na situação do cálculo das barreiras de potencial, torna-se inadequado caso haja interesse em conhecer a geometria de equilíbrio gerada a partir de dado conjunto de valores iniciais dos ângulos θ_1 e θ_2 . Neste caso, há necessidade de deixar todos os parâmetros geométricos livres durante o processo de otimização.

3.6.1 Principais conformações de equilíbrio das $\underline{N}^2$ -fenilguaninas

Investigou-se a geometria molecular de equilíbrio das $\underline{N}^2$ -fenilguaninas a partir do conhecimento das superfícies de potencial determinadas na Seção 3.5. Fez-se a otimização da geometria molecular dos quatro derivados da $\underline{N}^2$ -fenilguanina já utilizados (compostos 1, 3'-CF₃, 4'-H; 10, 3'-Br, 4'-Br; 11, 3'-n-Prop, 4'-H e; 15, 3'-H, 4'-H), partindo-se de valores de θ_1 e θ_2 localizados nas regiões A, B, C e D dos respectivos mapas conformacionais. Decidiu-se pelos valores de θ_1 e θ_2 localizados nos limites e no centro das quatro regiões destacadas na Figura 21. Os valores iniciais de θ_1 e θ_2 são mostrados na Figura 22.

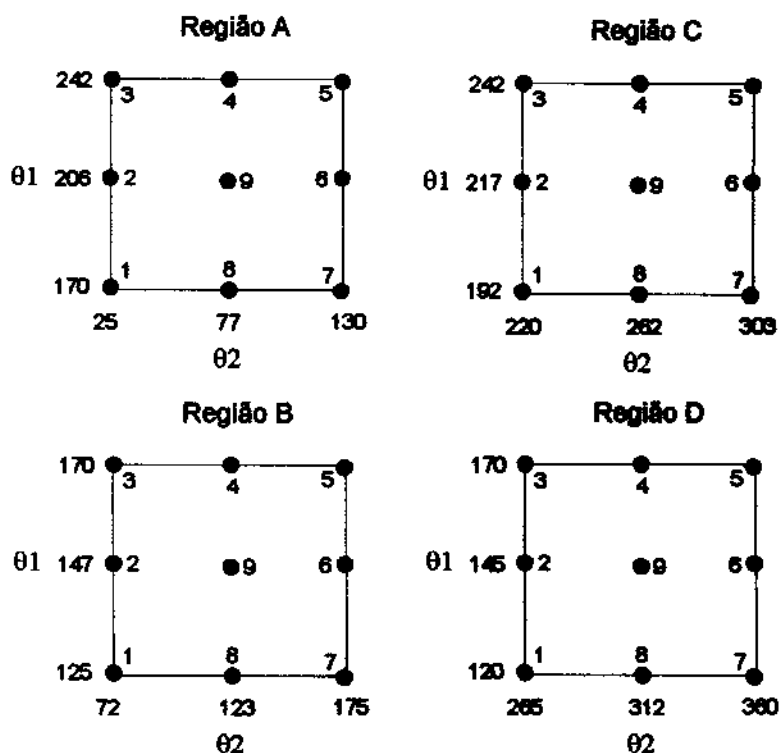


Figura 22 - Valores iniciais de θ_1 e θ_2 utilizados na localização das principais conformações de equilíbrio das N^2 -fenilguaninas. Cada ponto significa um conjunto particular de valores de θ_1 e θ_2 , que definem a geometria molecular inicial a ser submetida à otimização.

Observando-se a Figura 22, nota-se que para cada região de estabilidade conformacional foram feitas 9 otimizações de geometria molecular. Isto significa que, para cada um dos quatro compostos selecionados, fez-se 36 otimizações de geometria molecular com liberdade conformacional total. Os cálculos foram feitos através do método AM1. A geometria inicial foi preparada através do programa HyperChem e pré-otimizada pelo campo de força MM+.

A Tabela 12 mostra os valores de θ_1 e θ_2 , e também a entalpia de formação, dos quatro derivados da N^2 -fenilguanina em suas principais conformações de equilíbrio. Apesar da quantidade de conjuntos de valores iniciais de θ_1 e θ_2 , apenas quatro conformações de equilíbrio (C1, C2, C3 e C4, ver Tabela 12) foram atingidas para cada composto. Essas conformações, embora diferentes entre si, são praticamente as mesmas para cada um dos compostos estudados, como pode ser evidenciado pela semelhança dos

valores de θ_1 e θ_2 de cada conformação. Neste sentido, a conformação C1 dos compostos 1, 10, 11 e 15 é definida, respectivamente, por $\theta_1 = 184$ e $\theta_2 = 58$; $\theta_1 = 190$ e $\theta_2 = 55$; $\theta_1 = 182$ e $\theta_2 = 60$; $\theta_1 = 182$ e $\theta_2 = 60$.

Tabela 12 - Conformações de equilíbrio geradas a partir dos valores iniciais de θ_1 e θ_2 mostrados na Figura 22.

Composto	Parâmetro	Conformação de equilíbrio			
		C1	C2	C3	C4
1, 3'-CF ₃ , 4'-H	θ_1^a	184°	174°	186°	176°
	θ_2^a	58°	127°	233°	303°
	ΔH_f^b	-69.37	-69.72	-69.72	-69.63
	Num.conv. ^c	6	10	10	9
10, 3'-Br, 4'-Br	θ_1	190°	170°	190°	170°
	θ_2	55°	130°	230°	305°
	ΔH_f^b	96.96	96.84	96.84	96.96
	Num.conv.	9	9	7	11
11, 3'-n-Prop,4'-H	θ_1	182°	178°	181°	178°
	θ_2	60°	125°	235°	300°
	ΔH_f^b	65.47	65.44	65.45	65.49
	Num.conv.	9	9	6	12
15, 3'-H, 4'-H	θ_1	182°	178°	182°	178°
	θ_2	60°	125°	235°	301°
	ΔH_f^b	85.95	85.95	85.77	85.93
	Num.conv.	8	10	6	12
Número total de convergências		32	38	29	44

^a ângulos torsionais da conformação de equilíbrio; ^b Entalpia de formação-padrão da conformação de equilíbrio, em kcal.mol⁻¹; ^c Número de vezes que os cálculos convergiram para uma conformação em particular.

A Figura 23 mostra a localização das conformações de baixa energia C1-C4 nos mapas conformacionais dos quatro derivados da N²-fenilguanina. Mostra também os valores de θ_1 e θ_2 utilizados como pontos de partida para a determinação das conformações de baixa energia.

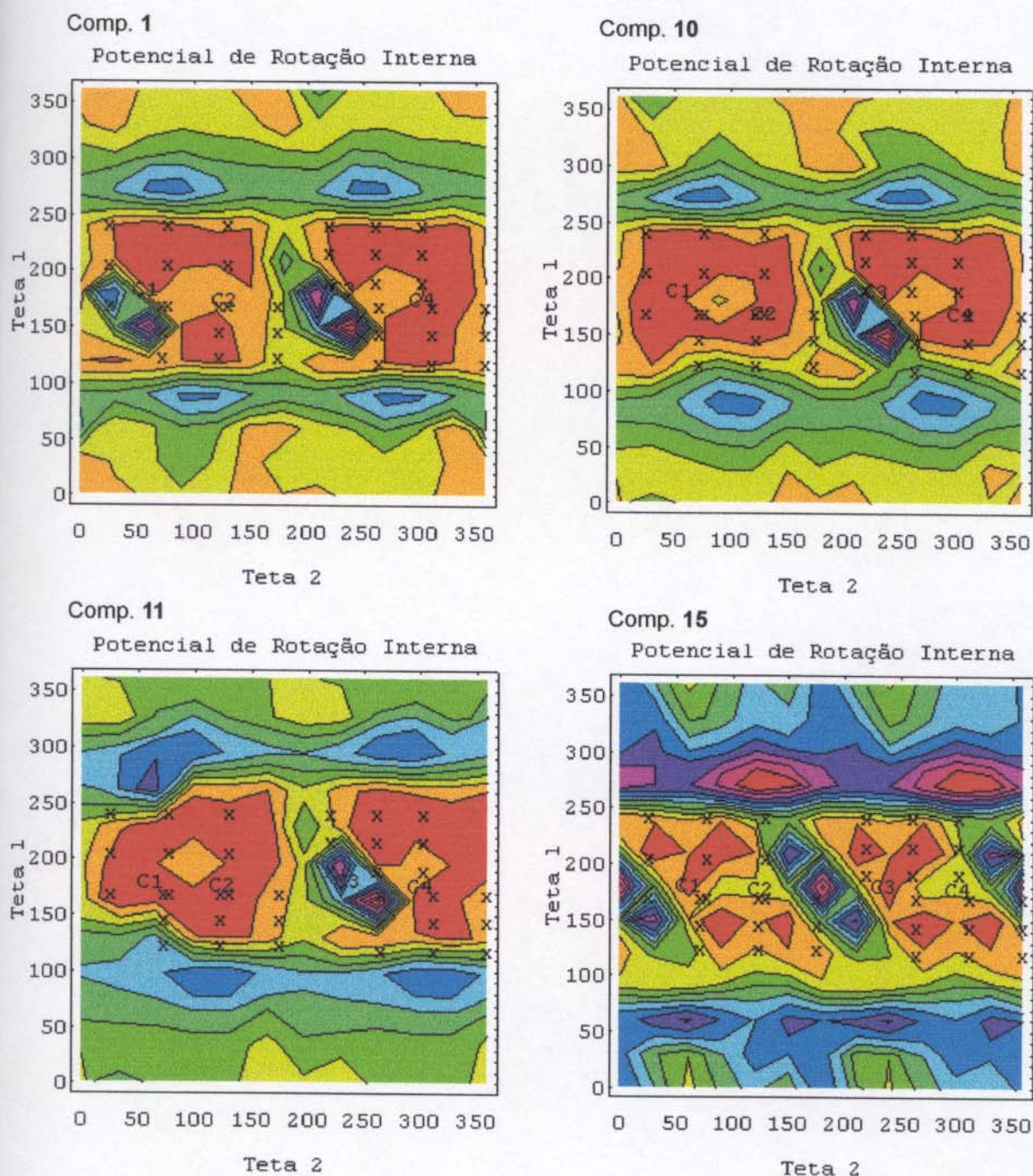


Figura 23 - Mapas conformacionais dos compostos 1 (3'-CF₃, 4'-H), 10 (3'-Br, 4'-Br), 11 (3'-n-Prop, 4'-H) e 15 (3'-H, 4'-H), mostrando a posição das principais conformações de equilíbrio, C1, C2, C3 e C4 (ver também a Tabela 12). O símbolo "x" indica as coordenadas de partida (θ_1 e θ_2 iniciais) das diversas otimizações da geometria molecular.

É interessante ressaltar que no mapa conformacional referente ao composto 11 (3'-n-Prop, 4'-H) a conformação C3 aparece numa região que aparentemente é de máximo de energia conformacional, o que seria inconsistente. Na verdade, isto é o resultado do fato

dos mapas conformacionais terem sido construídos com intervalos de valores de θ_1 e θ_2 igual a 30° . Utilizando-se intervalos de 30° , muitos detalhes acabam se perdendo e algumas regiões de baixa (ou de alta) energia conformacional podem não ser adequadamente mapeadas. Em situação ideal, esse intervalo deveria ser bem menor, talvez 5° , para que o relevo do mapa fosse melhor delineado. No entanto, a utilização de intervalo igual a 5° elevaria o número de otimizações de geometria de 1728 para 62208, o que demandaria tempo estimado de 162 dias de cálculo, de acordo com a Tabela 7 (página 108).

As Figura 24 - Figura 27 mostram as conformações C1-C4 do composto 1 (3'-CF₃, 4'-H), como exemplo representativo das principais conformações de equilíbrio das N²-fenilguaninas.

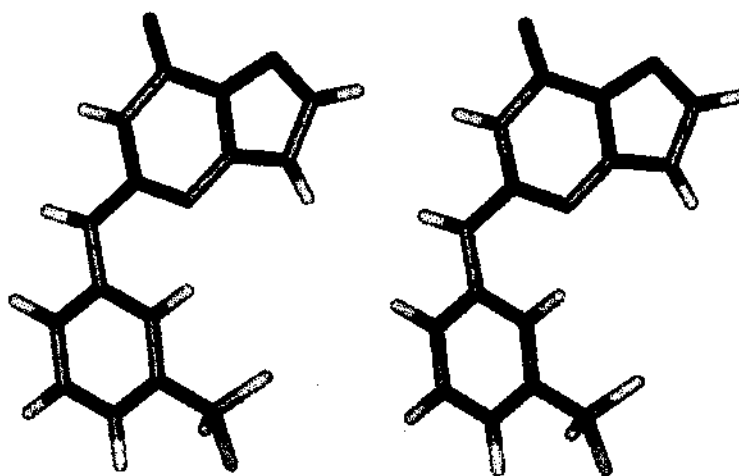


Figura 24 - Estereograma² da conformação C1 do composto 1 (3'-CF₃, 4'-H, Tabela 1, pag. 56). $\theta_1 = 184^\circ$ e $\theta_2 = 58^\circ$.

² O estereogramas apresentados neste trabalho foram construídos para serem observados a olho nu, embora a utilização de óculos estereoscópicos facilitem a observação. Os estereogramas foram testados por pelo menos três pessoas para verificação da facilidade de obtenção da visão estereoscópica, sendo que não houve grande dificuldade observação, exceto no caso das pessoas que nunca haviam observado esse tipo de imagem.

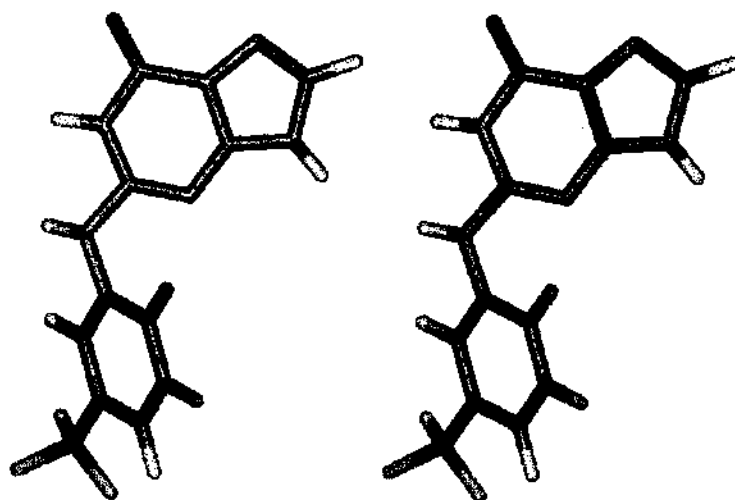


Figura 25 - Estereograma da conformação C2 do composto 1 (3'-CF₃, 4'-H, Tabela 1, pag. 56). $\theta_1 = 174^\circ$ e $\theta_2 = 127^\circ$.

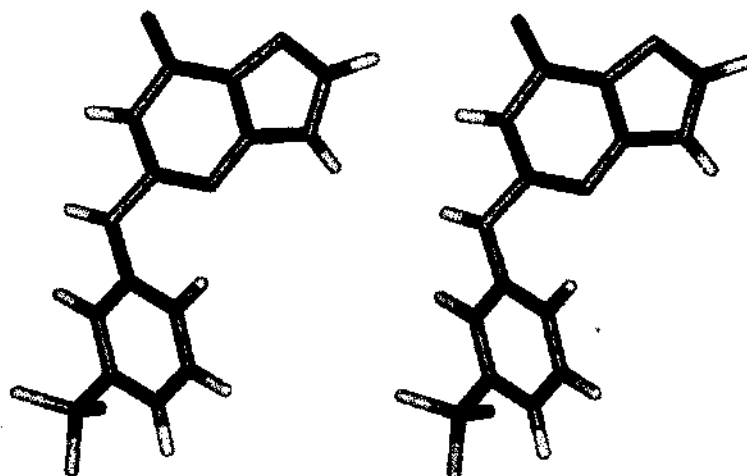


Figura 26 - Estereograma da conformação C3 do composto 1 (3'-CF₃, 4'-H, Tabela 1, pag. 56). $\theta_1 = 186^\circ$ e $\theta_2 = 233^\circ$.

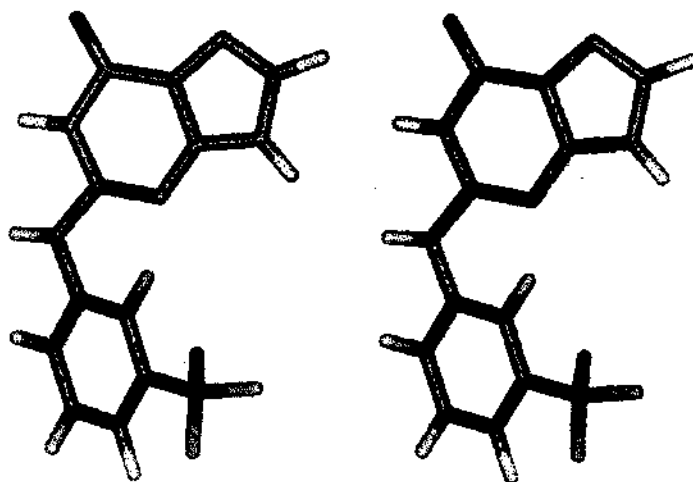


Figura 27 - Estereograma da conformação C4 do composto 1 (3'-CF₃, 4'-H, Tabela 1, pag. 56). $\theta_1 = 176^\circ$ e $\theta_2 = 303^\circ$.

3.7 CONSIDERAÇÕES FINAIS

O objetivo da análise conformacional dos compostos bioativos quase sempre é descobrir a geometria que o composto adota quando em processo de interação com seu receptor farmacológico, ou seja, a geometria bioativa. Exceto no caso das moléculas com estrutura rígida, não é possível concluir com segurança o que quer que seja a respeito da geometria molecular bioativa apenas analisando a molécula do fármaco. É necessário levar em consideração algum conhecimento sobre a estrutura do receptor ou quaisquer outras informações que relacionem a geometria molecular à atividade biológica.

Os resultados da presente análise conformacional são o produto apenas da análise da molécula do inibidor. Não foi feita qualquer consideração a respeito do receptor, ou de informações que possam indicar esta ou aquela conformação como sendo preferencial durante a interação enzima-inibidor. Portanto, não é possível avançar mais nas conclusões a respeito da geometria bioativa. É possível que a geometria bioativa seja parecida com

alguma das conformações apresentadas na Seção 3.6, mas não é possível afirmar com segurança.

O problema da determinação da geometria bioativa voltará a ser atacado no Capítulo 5, considerando a geometria do receptor das N^2 -fenilguaninas na timidina-cinase.

3.8 CONCLUSÕES DA ANÁLISE CONFORMACIONAL

- (a) A análise dos mapas conformacionais dos derivados da N^2 -fenilguanina indicou a existência de quatro regiões de estabilidade conformacional, A, B, C e D, localizadas nos seguintes intervalos dos ângulos torsionais θ_1 e θ_2 : $100^\circ \leq \theta_1 \leq 250^\circ$ e $90^\circ \leq \theta_2 \leq 140^\circ$ e $250^\circ \leq \theta_2 \leq 320^\circ$;
- (b) Apenas quatro conformações estáveis das N^2 -fenilguaninas, C1, C2, C3 e C4, foram identificadas na análise conformacional. Os ângulos torsionais θ_1 e θ_2 que caracterizam essas conformações são: C1, $\theta_1 = 184^\circ$ e $\theta_2 = 58^\circ$; C2, $\theta_1 = 174^\circ$ e $\theta_2 = 127^\circ$; C3, $\theta_1 = 186^\circ$ e $\theta_2 = 233^\circ$ e; C4, $\theta_1 = 176^\circ$ e $\theta_2 = 303^\circ$.

CAPÍTULO 4

QSAR COM PARÂMETROS CLÁSSICOS

4.1 INTRODUÇÃO

O método clássico de análise das relações quantitativas entre estrutura química e atividade biológica (QSAR), ou método de Hansch-Fujita (Hansch & Fujita, 1964), baseia-se na utilização de parâmetros físico-químicos linearmente relacionados com a energia livre para a construção de modelos multidimensionais do tipo

$$AB = a_1 X-1 + a_2 X-2 + \dots + a_n X-n + a_{n+1} \quad , \quad (95)$$

em que AB é a atividade biológica, medida em escala apropriada (Martin, 1978; Kubinyi, 1993), X-n são os parâmetros físico-químicos que descrevem as propriedades de grupos substituintes e a_n são coeficientes obtidos por análise de regressão linear múltipla. Modelos desse tipo correlacionam a atividade biológica com as propriedades medidas pelos parâmetros físico-químicos. Esses parâmetros, chamados clássicos, são propriedades moleculares aditivas e constitutivas (Martin, 1978) e são normalmente classificados como sendo dos tipos lipofílico, eletrônico e estérico. Os parâmetros mais utilizados em QSAR clássico são o logaritmo do coeficiente de partição octanol/água ($\log P$), a constante lipofílica de Hansch-Fujita (Fujita et al, 1964), π^2 , a constante eletrônica de Hammett (Hammett, 1937), σ e a constante estérica de Taft (Taft, 1952), Es. A validade do método de Hansch pode ser comprovada a partir das dezenas de milhares de equações publicadas desde que o método foi formulado, em 1964 (Hansch & Leo, 1995). Ótimos trabalhos de

³ $\pi = \log P_X - \log P_H$, em que P_X é o coeficiente de partição octanol/água da molécula-mãe substituída pelo grupo químico genérico X e P_H é o coeficiente de partição da molécula-mãe não substituída.

revisão sobre o método de Hansch, também chamado de abordagem extratermodinâmica, podem ser encontrados nas referências: Tute, 1971; Redl, 1974; Martin, 1978 e 1981; Fujita, 1990; Kubinyi, 1993; Hansch, 1995.

Recentemente publicou-se em periódico de circulação nacional a revisão dos métodos clássicos de estudo das relações entre estrutura química e atividade biológica (Gaudio, 1996). Nessa revisão, relatou-se o histórico do desenvolvimento desses métodos e fez-se a análise em pormenor dos métodos de Free-Wilson, Fujita-Ban, Hansch-Fujita e Kubinyi, arrolando exemplos ilustrativos da aplicação de cada um deles.

Até o presente momento, têm-se conhecimento de apenas dois trabalhos QSAR sobre as N^2 -fenilguaninas como inibidores de HSV TK (Gambino et al, 1992; Hadjipavlou-Litina, 1995). Mostram-se, a seguir, as principais conclusões desses trabalhos.

4.1.1 Estudo QSAR de Gambino e colaboradores

Gambino e colaboradores (1992) sintetizaram e determinaram a atividade de 36 derivados da N^2 -fenilguanina sobre HSV1 TK e HSV2 TK. O isolamento e purificação das enzimas utilizadas nos testes biológicos são descritos por Focher e colaboradores (1988). A atividade foi medida em termos de IC_{50} , a concentração molar de cada derivado capaz de proporcionar 50% de inibição da atividade enzimática, em presença de concentração fixa de 2'-deoxitimidina e tendo o ATP como agente doador de grupo fosfato (Hildebrand et al, 1990). Os valores de IC_{50} dos derivados da N^2 -fenilguanina mostraram-se proporcionais e muito próximos aos valores de K_i , a constante de inibição da reação. Portanto, a atividade representada por IC_{50} pode ser considerada como sendo a própria medida da afinidade do inibidor pela enzima (Gambino et al, 1992). Nenhum dos compostos testados provocou inibição da timidina-cinase da célula hospedeira, dentro dos

limites de concentração dos reagentes utilizados (Hildebrand et al, 1990). Os parâmetros físico-químicos testados nessa análise foram: (a) a constante lipofílica de Hansch-Fujita dos substituintes nas posições meta e para do anel fenila, π_m e π_p ; (b) a constante eletrônica de Hammett dos substituintes nas posições meta, σ_m (efeito indutivo), e para, σ_p (efeito de ressonância); (c) as constantes eletrônicas de Swain-Lupton (Swain & Lupton, 1968) dos substituintes nas posições meta (efeito indutivo), F_m , e para (ressonância), R_p , e; (d) a refratividade molar (Martin, 1978) dos substituintes nas posições meta, RM_m , e para, RM_p , para detectar efeitos polares e possíveis impedimentos estéricos. Cabe ressaltar que foram detectadas algumas incorreções e também conjuntos de parâmetros desatualizados na matriz de parâmetros físico-químicos utilizados por Gambino e colaboradores (ver detalhes na Seção 4.3). Como resultado da análise, foram elaborados apenas dois modelos lineares, baseados no método de Hansch. A Equação (96) mostra o modelo proposto para explicar a atividade das N^2 -fenilguaninas sobre HSV1 TK.

$$\begin{aligned} \log(1/C) = & 0,56 (\pm 0,16) \pi_m - 0,45 (\pm 0,14) \pi_p + 0,73 (\pm 0,49) \sigma_m + 2,45 (\pm 0,65) R_p + \\ & + 5,10 (\pm 0,16) \end{aligned} \quad (96)$$

(n = 38; r = 0,797; s = n^f; F = 13,48; p = n^f)

Na equação (96), n é o número de compostos analisados, r é o coeficiente de correlação, s é o desvio-padrão, F é o teste de Fischer ou razão da variância, p é o nível global de confiabilidade da equação, e os números apresentados entre parênteses, após cada coeficiente, correspondem ao desvio-padrão⁴ dos coeficientes da equação (ver Seção 2.5.1, “Regressão linear múltipla”, pag.75).

⁴ não fornecido.

⁵ A apresentação do desvio-padrão dos coeficientes nas equações de regressão são desaconselhados por Kubinyi (Kubinyi, 1993). É aconselhável a apresentação do limite de confiabilidade de 95% dos coeficientes da equação (Draper & Smith, 1981).

De acordo com a Equação (96), a atividade das N^2 -fenilguaninas sobre HSV1 TK é basicamente controlada pelo parâmetro lipofílico, sendo que substituintes mais lipofílicos na posição meta e mais hidrofílicos na posição para do anel fenila intensificam a atividade inibitória. Substituintes na posição para capazes de extrair elétrons via mecanismo de ressonância também parecem ser importantes para a atividade desses compostos. Nota-se que a significância estatística da Equação (96) não é boa. Seu coeficiente de correlação é baixo (0,797), embora o teste de Fischer seja cerca de cinco vezes maior do que o valor mínimo aceitável, que é de 2,6 (Draper & Smith, 1981). Além disso, convertendo-se os números entre parênteses para o limite de confiabilidade de 95%, verifica-se que o coeficiente de σ_m é nulo, pois seu limite de confiabilidade é maior do que o próprio coeficiente, incluindo, pois, o valor zero. $(0,72 (\pm 0,98) \sigma_m)$.

Gambino e colaboradores eliminaram os compostos menos ajustados (também chamados *outliers*) ao modelo representado pela Equação (96), compostos 3 (4'-Br) e 30 (3',4'-Cl), e a exclusão desses compostos produziu a Equação (97).

$$\log(1/C) = 0,61 (\pm 0,12) \pi_m - 0,44 (\pm 0,11) \pi_p + 1,05 (\pm 0,38) \sigma_m + 2,53 (\pm 0,49) R_p + \\ + 5,03 (\pm 0,12) \quad (97)$$

$$(n = 34; r = 0,887; s = nf; F = 26,21; p = nf)$$

Embora a Equação (97) apresente melhor ajuste do que a Equação (96), não há diferenças estruturais importantes entre essas equações. Apenas o coeficiente do parâmetro σ_m aumentou significativamente. Devido ao fato de os compostos eliminados terem substituintes constituídos de halogêneos e apenas σ_m ter variado, Gambino e colaboradores concluíram que algum efeito eletrônico não detectado pelas constantes testadas deve ser importante na interação enzima-inibidor.

A atividade inibitória dos 35 derivados da N^2 -fenilguanina sobre HSV2 TK foi modelada pela Equação (98).

$$\log(1/C) = 0,40 (\pm 0,10) \pi_m - 0,53 (\pm 0,15) \pi_p + 1,72 (\pm 0,49) \sigma_m + 1,49 (\pm 0,64) R_p +$$

$$+ 5,27 (\pm 0,16) \quad (98)$$

$$(n = 35; r = 0,805; s = \text{nf}; F = 13,89; p = \text{nf})$$

Embora os parâmetros que aparecem na Equação (98) sejam os mesmos da Equação (96), os coeficientes de σ_m e R_p sofreram alterações significativas. O coeficiente de σ_m aumentou de 0,73 para 1,72, enquanto que o coeficiente de R_p diminuiu de 2,45 para 1,49. Esses números indicam que fatores eletrônicos distintos são responsáveis pela atividade inibitória das N^2 -fenilguaninas sobre HSV1 TK e HSV2 TK.

A Equação (99) foi produzida após a eliminação dos compostos 3 e 30, mais uma vez os menos ajustados ao modelo.

$$\log(1/C) = 0,45 (\pm 0,12) \pi_m - 0,52 (\pm 0,11) \pi_p + 2,05 (\pm 0,38) \sigma_m + 1,57 (\pm 0,49) R_p +$$

$$+ 5,21 (\pm 0,12) \quad (99)$$

$$(n = 33; r = 0,893; s = \text{nf}; F = 27,61; p = \text{nf})$$

Pode-se notar que o coeficiente de σ_m sofreu a alteração mais significativa. Esta observação indica que há diferenças no modo como as N^2 -fenilguaninas interagem com o sítio ativo de HSV1 TK e HSV2 TK. Isto provavelmente se deve às possíveis diferenças nos sítios ativos dessas enzimas e não no modo de ligação das N^2 -fenilguaninas, uma vez que há indicações de que esses compostos agem pelo mesmo mecanismo em ambas enzimas (Focher et al, 1988).

4.1.2 Estudo QSAR de Hadjipavlou-Litina

Hadjipavlou-Litina (1995) apresentou estudo QSAR clássico complementar ao de Gambino e colaboradores. Nesse trabalho, foram utilizados apenas os valores da atividade inibitória das N^2 -fenilguaninas sobre HSV1 TK, produzidos por Gambino. A melhor (e única) equação apresentada por Hadjipavlou-Litina foi a Equação (100).

$$\log(1/C_{50}) = 1,73 (\pm 0,596) \pi_{3,4} - 2,54 (\pm 0,82) \log (\beta \cdot 10 \pi_{3,4} + 1) + 1,97 (\pm 0,47) B_{1,3} - 1,72 (\pm 0,64) F + 4,07 (\pm 0,80) \quad (100)$$

$$(n = 30; r = 0,907; s = 0,401; F = 22,30; p = \text{nf}; \pi_0 = 0,141; \log \beta = 0,198)$$

Na Equação (100), os números apresentados entre parênteses, após cada coeficiente, são os limites de confiabilidade de 95% dos coeficientes da equação; $\pi_{3,4}$ é a soma dos valores da constante lipofílica de Hansch-Fujita (Fujita et al, 1964) dos substituintes presentes nas posições meta e para do anel fenila das N^2 -fenilguaninas; π_0 é o valor ótimo de $\pi_{3,4}$; $B_{1,3}$ é a constante estérica de Verloop (Kubinyi, 1993) dos substituintes na posição meta ($B_{1,3} = B_1$ na posição 3', que corresponde à posição meta; B_1 é a largura mínima do substituinte, medida a partir do eixo de substituição e em direção ortogonal a este); F é a constante eletrônica de Swain-Lupton para efeito indutivo (Swain & Lupton, 1968) e; β é o parâmetro de ajuste bilinear de Kubinyi (Kubinyi, 1977). Também foram testados por Hadjipavlou-Litina os seguintes parâmetros: RM_m e RM_p , σ_p e R_p . A qualidade dos valores das constantes dos substituintes de sua matriz de parâmetros físico-químicos é equivalente à de Gambino (ver Seção 4.3).

De acordo com a Equação (100), a atividade inibitória dos derivados da N^2 -fenilguanina sobre HSV1 TK obedece ao modelo bilinear⁶ assimétrico (Kubinyi, 1977), ou seja, a atividade inibitória possui dependência bilinear (não-linear) em relação à lipofilicidade. As declividades da curva bilinear são 1,73 e -0,81 ($1,73 - 2,54 = -0,81$). Destaca-se no modelo apresentado por Hadjipavlou-Litina a ocorrência de efeito estérico na posição meta. De acordo com o modelo, quanto maior a largura mínima do substituinte

⁶ O modelo bilinear pode ser representado pela função $\log 1/C = a \log P - b \log (\beta P + 1) + c$, em que C é a concentração molar do fármaco, P é o coeficiente de partição octanol/água e a , b , c e β são constantes de ajuste. A curva bilinear é semelhante à parábola, sendo que as partes ascendente e descendente são retilíneas.

ligado à posição meta, maior a atividade inibitória. Há que se destacar o fato de a Equação (100) incluir apenas 30 compostos. Seis compostos (2, 3'-CH₂N₃; 3, 4'-Br; 10, 3',4'-Br; 17, 3',4'-CH₂CH₂CH₂; 21, 3'-CH₂NH₂ e; 24, 3',4'-CH=CHCH=CH) não se ajustaram ao modelo bilinear.

4.2 ESTRATÉGIA DA ANÁLISE

Fez-se a análise clássica das relações estrutura-atividade separando-se os compostos em grupos de acordo com a posição do substituinte no anel fenila. Dessa forma, analisou-se os compostos meta, para, meta e para em conjunto, di-substituídos, e, finalmente, todo o conjunto de compostos. Para a seleção de variáveis, utilizou-se o método de busca sistemática, em que, para dado grupo de compostos analisados, verificou-se todas as possíveis regressões com uma variável, todas as possíveis regressões com duas variáveis, etc. (ver Seção 2.5.2, "Seleção de variáveis por busca sistemática", pag.80). A seleção das melhores equações foi feita com base no coeficiente de correlação, desvio-padrão e teste de Fischer, simultaneamente (Draper & Smith, 1981). Como método complementar de seleção de modelos, utilizou-se a validação cruzada (*cross-validation*) do tipo LOO (*leave-one-out*) para identificar os modelos com maior capacidade preditiva. Foram calculados o quadrado do coeficiente de correlação das previsões (Q^2), o desvio-padrão das previsões (s_{PRESS}) e o erro padrão das previsões (SDEP) (ver Seção 2.5.1, "Regressão linear múltipla", pag.75). Para as melhores equações de regressão serão apresentadas tabelas contendo a atividade observada, prevista e os respectivos resíduos, bem como gráficos da atividade observada em função da atividade prevista.

Pode-se utilizar a seguinte formulação alternativa para o modelo bilinear $\log 1/C = a\pi - b \log (\beta 10^\pi + 1) + c$, em que π é o parâmetro lipofílico. O valor de P_o , π_o (P e π ótimos) são dados por $a/(\beta.(b-a))$, ($b > a$).

No processo de seleção de variáveis e cálculo das equações de regressão, utilizou-se o programa MLA (*Multiple Linear Regression*), construído pelo autor. MLA foi codificado em linguagem de programação Basic e possui cerca de cinco mil linhas de código fonte.

4.3 PARÂMETROS CLÁSSICOS UTILIZADOS

A matriz de parâmetros físico-químicos de Gambino e colaboradores foi revista, o que implicou em diversas correções, e ampliada para a execução da presente análise. As principais correções foram (entre parênteses são apresentados os valores originais incorretos): (a) π_m - composto 2: 1,12⁷ (0,72 - valor incorreto), composto 6: 0,17 (0,64); (b) π_p - composto 36: 2,67 (2,39); (c) σ_m - pequenas correções no composto 6: 0,11 (0,10), composto 11: -0,06 (-0,07) e composto 21: -0,03 (0,00) e (d) R_p . Os valores de R_p foram todos modificados devido à reparametrização dessa constante feita por Hansch e colaboradores (1991). Todos esses valores fazem parte da mais recente atualização de parâmetros disponível e, portanto, mais confiáveis do que os parâmetros utilizados por Gambino e colaboradores (Hansch, 1973). Além disso, incluíram-se na presente análise os parâmetros estéricos de Verloop, L , B_1 e B_5 , que medem, respectivamente, o comprimento, a largura mínima e a largura máxima do substituinte, em relação ao eixo de substituição (Kubinyi, 1993). Finalmente, incluíram-se os quadrados dos parâmetros π_m e π_p como tentativa para detectar possível não-linearidade entre a atividade inibitória das N^2 -fenilguaninas e a lipofilicidade.

⁷ O valor de $\pi_m = 1,12$ para o substituinte CH_2N_3 foi calculado e gentilmente fornecido pelo Dr. Corwin Hansch, Pomona College, Claremont, California.

Os valores dos parâmetros clássicos utilizados na análise de QSAR, bem como a matriz de correlação dos parâmetros, são mostrados, respectivamente, na Tabela 19 (pag. 151) e Tabela 20 (pag. 153).

4.4 ANÁLISE DOS COMPOSTOS META-SUBSTITUÍDOS

Nas equações que se seguem, os números apresentados entre parênteses, após cada coeficiente, correspondem ao limite de confiabilidade de 95% dos coeficientes da equação (Draper & Smith, 1981).

4.4.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

4.4.2 Atividade inibitória sobre HSV2 TK

O único modelo significativo e com duas variáveis capaz de representar a atividade dos derivados meta-substituídos sobre HSV2 TK é mostrado na Equação (101).

$$\log(1/C) = 0,334 (\pm 0,238) \pi_m + 2,284 (\pm 0,920) \sigma_m + 5,307 (\pm 0,286) \quad (101)$$

$$(n = 13; r = 0,897; s = 0,315; F_{(2,10)} = 20,676; p < 0,00028)$$

$$(Q^2 = 0,690; s_{PRESS} = 0,398; SDEP = 0,363)$$

A Equação (101) não possui equivalente na análise de Gambino e colaboradores. Essa equação indica que a atividade inibitória dos compostos meta-substituídos contra HSV2 TK depende da lipofilicidade e da capacidade do substituinte ligado à posição meta de extrair elétrons do anel fenila, sendo que substituintes mais lipofílicos e com maior capacidade de extrair elétrons do anel fenila por efeito indutivo intensificam a atividade. A Equação (101) está em acordo com o modelo de Gambino, apresentado pela Equação (98) (pag. 129).

Na construção da Equação (101), o parâmetro eletrônico, σ_m , é o mais importante e é capaz de explicar cerca de 62% da variância total da atividade inibitória. Ao acrescentar-se o parâmetro lipofílico, π_m , este contribui com aumento de apenas 18% na explicação da variância. Na Equação (101), o parâmetro π_m é pouco significativo, na medida em que seu coeficiente possui limite de confiabilidade muito estreito. Portanto, a Equação (101) indica que a atividade dos derivados meta das N^2 -fenilguaninas sobre HSV2 TK é basicamente controlada pelo parâmetro eletrônico.

A Tabela 13 mostra os valores da atividade inibitória previstos pela Equação (101), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (101) é mostrado na Figura 28.

Tabela 13 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (101).

Comp.	Substituinte		log 1/C ₅₀		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	6,57	0,43
2	CH ₂ N ₃	H	5,82	5,68	0,14
4	CH ₂ Br	H	5,92	5,84	0,08
5	I	H	5,96	6,47	-0,51
6	CH ₂ Cl	H	5,87	5,61	0,26
7	Cl	H	6,46	6,38	0,08
8	Br	H	6,22	6,48	-0,26
9	Et	H	5,40	5,49	-0,09
11	n-Pr	H	5,52	5,69	-0,17
12	CN	H	6,40	6,38	0,02
15	H	H	5,80	5,31	0,49
20	CH ₂ OH	H	4,70	4,96	-0,26
21	CH ₂ NH ₂	H	4,70	4,89	-0,19

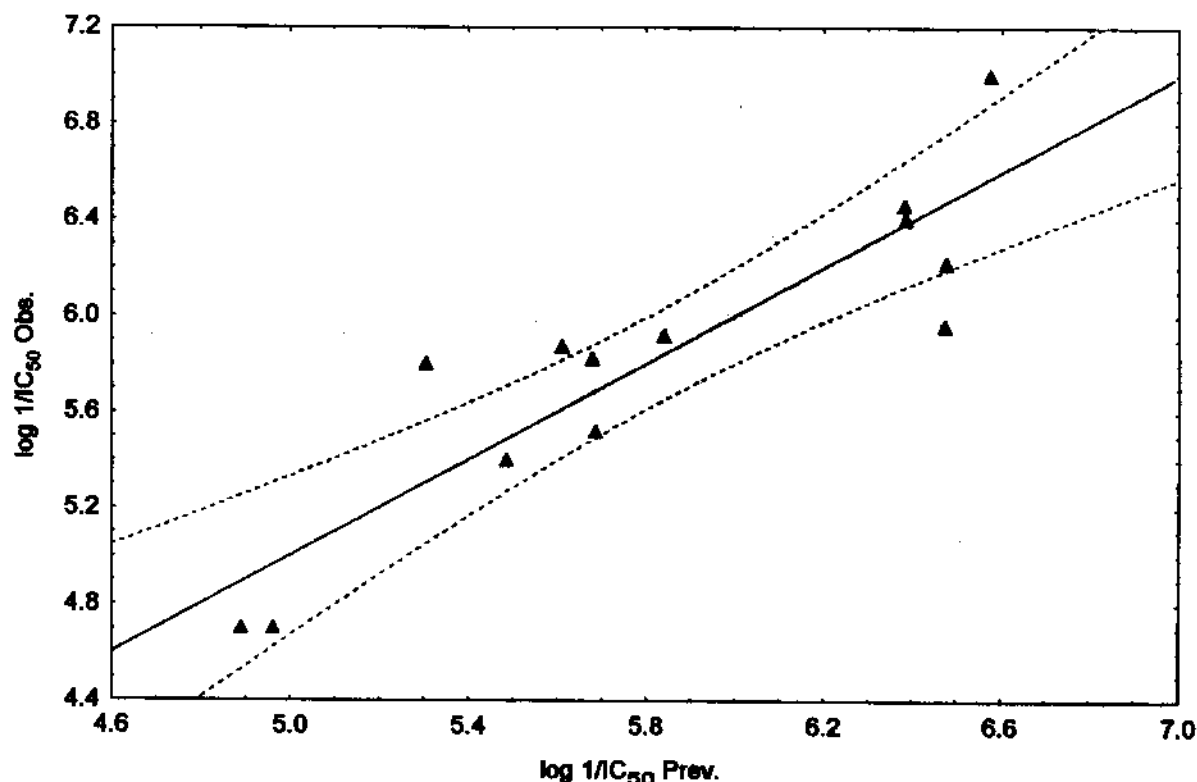


Figura 28 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (101). As linhas pontilhadas são hipérboles que demarcam a região correspondente a 95% de confiança dos valores previstos.

Pode-se verificar como se comporta a Equação (101) ao reconstruí-la excluindo-se os dois compostos menos ajustados. São eles os compostos 5 (3'-I, 4'-H), cujo valor calculado de $\log(1/IC_{50})$ foi superestimado em 1,6 desvios-padrão, e 15 (3'-H, 4'-H), cujo valor calculado da atividade inibitória foi subestimado em 1,6 desvios-padrão. A exclusão desses dois compostos deu origem à Equação (102):

$$\log(1/C) = 0,409 (\pm 0,178) \pi_m + 2,616 (\pm 0,699) \sigma_m + 5,220 (\pm 0,208) \quad (102)$$

$$(n = 11; r = 0,982; s = 0,219; F_{(2,8)} = 49,178; p < 0,00003)$$

$$(Q^2 = 0,873 \quad s_{PRESS} = 0,284 \quad SDEP = 0,254)$$

A Equação (102) possui ajuste superior à Equação (101). Na Equação (102), o parâmetro eletrônico isoladamente passa a ser capaz de explicar cerca de 66% da variância

total da atividade biológica. A inclusão do parâmetro lipofílico contribui com aumento de 26% em termos de explicação da variância. No entanto, ao observar-se o comportamento dos valores dos limites de confiabilidade de 95% dos coeficientes das variáveis presentes, vê-se que tanto o coeficiente π_m quanto o de σ_m tornaram-se mais confiáveis, simultaneamente. Na Equação (101), os limites de confiabilidade de π_m e σ_m representam, respectivamente, 72% e 41% do valor de seus coeficientes (quanto menor for o valor apresentado entre parênteses nas equações de regressão, maior será a confiabilidade do respectivo coeficiente). Na Equação (102), esses valores são, respectivamente, 42% e 26%, ou seja, os coeficientes das variáveis presentes na Equação (102) são mais confiáveis do que os da Equação (101). Esse fato é uma indicação de que tanto π_m quanto σ_m são responsáveis pelo desajuste dos compostos 5 e 15 à Equação (101).

A Tabela 14 mostra os valores da atividade inibitória previstos pela Equação (102), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (102) é mostrado na Figura 29.

Tabela 14 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (102).

Comp.	Substituinte		log 1/C ₅₀		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	6,71	0,29
2	CH ₂ N ₃	H	5,82	5,68	0,14
4	CH ₂ Br	H	5,92	5,86	0,06
6	CH ₂ Cl	H	5,87	5,58	0,29
7	Cl	H	6,46	6,48	-0,02
8	Br	H	6,22	6,59	-0,37
9	Et	H	5,40	5,45	-0,05
11	n-Pr	H	5,52	5,70	-0,18
12	CN	H	6,40	6,45	-0,05
20	CH ₂ OH	H	4,70	4,80	-0,10
21	CH ₂ NH ₂	H	4,70	4,72	-0,02

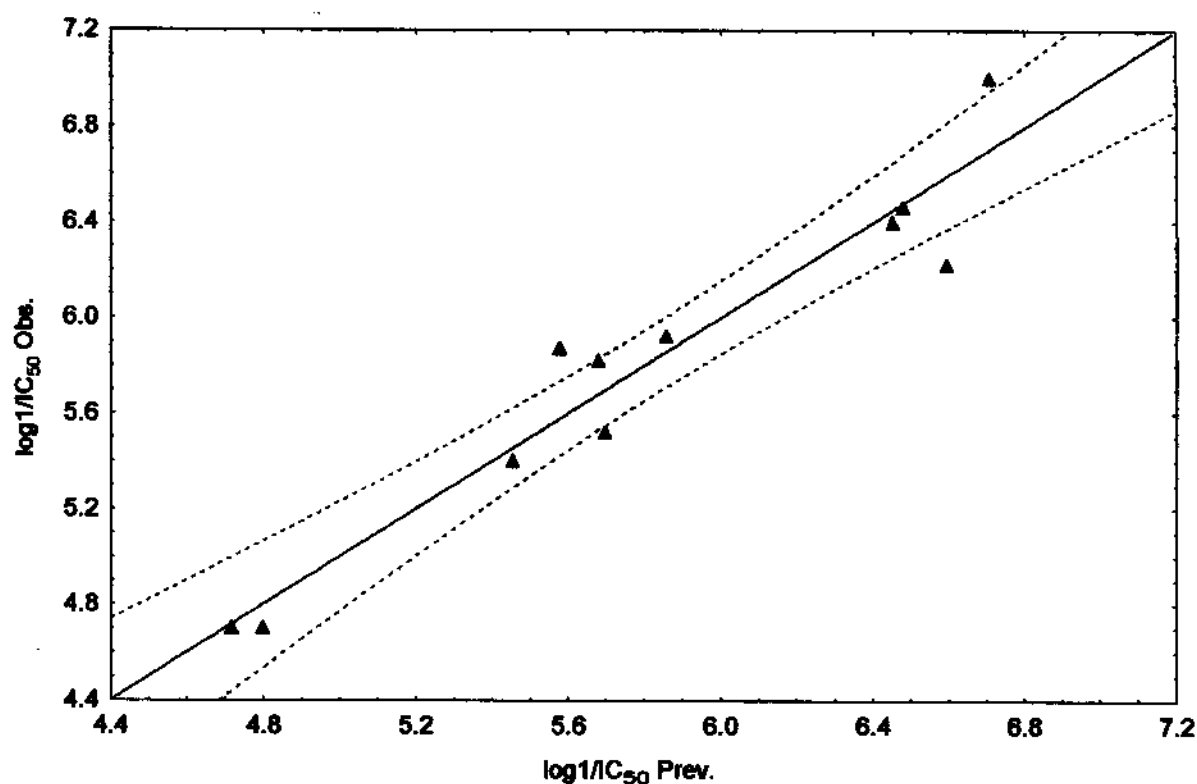


Figura 29 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (102).

Embora o número de derivados meta-substituídos seja de apenas treze, o que limita em dois o número de variáveis que podem ser analisadas com segurança, é importante verificar a ocorrência de modelos com três variáveis. Sob este aspecto, deve-se considerar a Equação (103).

$$\log(1/C) = 0,475 (\pm 0,232) \pi_m + 1,946 (\pm 0,805) \sigma_m - 0,546 (\pm 0,488) RM_m + 5,809 (\pm 0,498) \quad (103)$$

$$(n = 13; r = 0,941; s = 0,254; F_{(3,9)} = 23,372; p < 0,00014)$$

$$(Q^2 = 0,795; s_{PRESS} = 0,341; SDEP = 0,295)$$

É interessante a presença da refratividade molar na Equação (103). A refratividade molar é uma espécie de camaleão entre os parâmetros físico-químicos, apesar de ser amplamente utilizada em estudos QSAR (Martin, 1978; Kubinyi, 1993). Os valores da RM costumam ter correlação elevada com a lipofilicidade e o volume molar. Este fato costuma gerar dificuldades na interpretação do significado da presença da RM em modelos de QSAR. Kubinyi (1993) defende a seguinte tese: se a RM aparece com coeficiente positivo, isto pode representar a interação do substituinte com regiões polares do sítio ativo; coeficiente de RM negativo significa possível impedimento estérico, especialmente se a correlação com a lipofilicidade for pequena. Na Equação (103), a correlação entre π_m e RM_m , considerando-se somente os treze compostos, é de apenas cerca de 26%. Isto pode ser indicação de impedimento estérico na região meta do sítio ativo de HSV2 TK.

Apesar de Equação (103) possuir boas características gerais, nota-se que o coeficiente de RM_m possui limite de confiabilidade muito estreito. O limite de confiabilidade de 0,488 representa cerca de 90% do valor do próprio coeficiente de RM_m . No entanto, devido às boas características dessa equação, ela será mantida para futura comparação com os resultados da análise incluindo parâmetros calculados (ver Seção 6.2.2, “Atividade inibitória sobre HSV2 TK.”, pag. 235).

A Tabela 15 mostra a percentagem de explicação da variância de cada parâmetro, à medida em que foram introduzidos na Equação (103).

Tabela 15 - Percentagem de variância da atividade inibitória dos derivados meta sobre HSV2 TK explicada por cada parâmetro das Equações (101), (102) e (103).

Equação	Percentagem de variância explicada		
	σ_m	π_m	RM_m
101	62%	+18%	—
102	66%	+26%	—
103	62%	+18%	+9%

Na Tabela 15, e nas demais tabelas semelhantes a esta que serão mostradas ao longo do presente trabalho, cada linha mostra os parâmetros físico-químicos em ordem decrescente de capacidade de explicação da variância da atividade inibitória. Como exemplo de como deve ser interpretada este tipo de tabela, na linha referente à Equação (103), o parâmetro σ_m é o mais importante e capaz de explicar, sozinho, 62% da variância da atividade inibitória. Nenhum outro parâmetro testado é capaz de explicar, sozinho, maior percentual de variância. Ao incluir-se o parâmetro π_m , este, na presença de σ_m , acrescenta 18% à explicação da variância da atividade da Equação (103). Finalmente, quando se inclui o parâmetro RM_m , este, na presença de σ_m e π_m , acrescenta 9% à capacidade de explicação da variância da atividade da Equação (103). Isto significa que σ_m , π_m e RM_m são capazes de explicar cerca de 89% da variância dos valores da atividade dos compostos meta-substituídos, o que equivale ao quadrado do coeficiente de correlação da Equação (103), multiplicado por 100 ($0,941^2 \times 100 \cong 89\%$).

A Tabela 16 mostra os valores da atividade inibitória previstos pela Equação (103), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (103) é mostrado na Figura 30.

Tabela 16 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (103).

Comp.	Substituinte		log 1/C ₅₀		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	6,79	0,21
2	CH ₂ N ₃	H	5,82	5,54	0,28
4	CH ₂ Br	H	5,92	5,69	0,23
5	I	H	5,96	6,26	-0,30
6	CH ₂ Cl	H	5,87	5,53	0,34
7	Cl	H	6,46	6,54	-0,08
8	Br	H	6,22	6,49	-0,27
9	Et	H	5,40	5,60	-0,20
11	n-Pr	H	5,52	5,61	-0,09
12	CN	H	6,40	6,28	0,12
15	H	H	5,80	5,75	0,05
20	CH ₂ OH	H	4,70	4,93	-0,23
21	CH ₂ NH ₂	H	4,70	4,76	-0,06

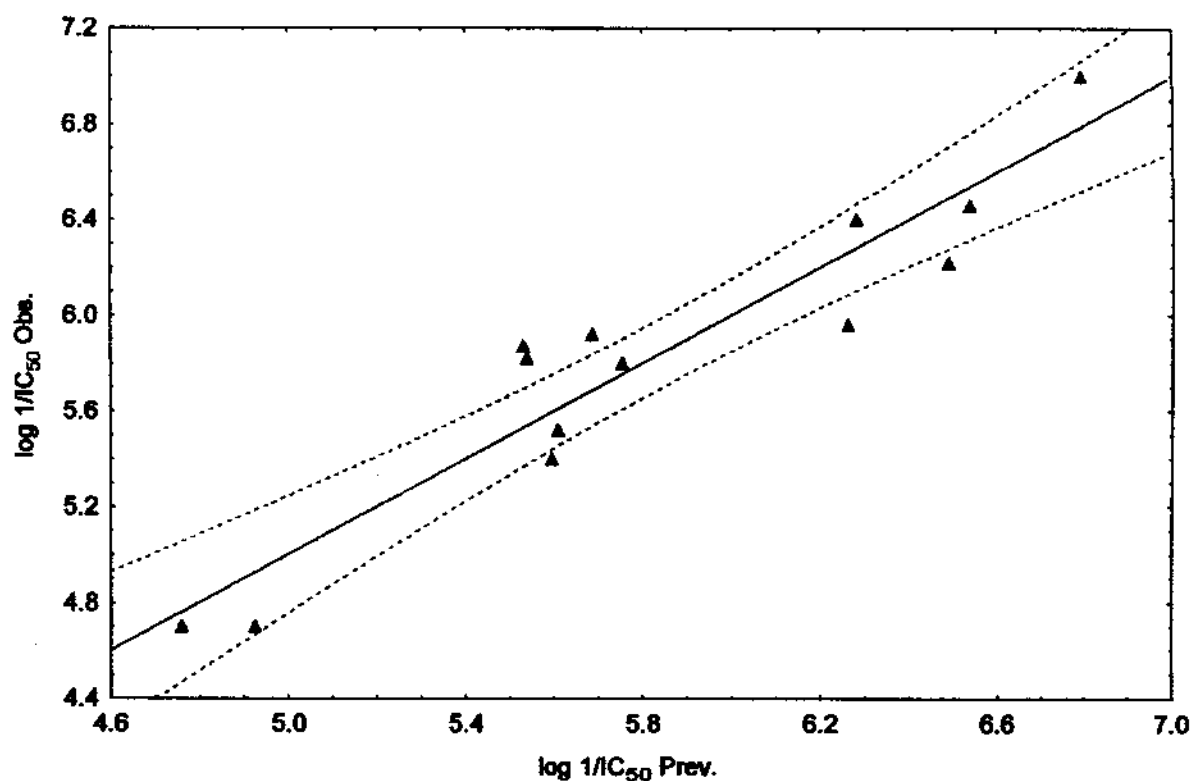


Figura 30 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (103).

4.5 ANÁLISE DOS COMPOSTOS PARA-SUBSTITUÍDOS

4.5.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

4.5.2 Atividade inibitória sobre HSV2 TK

Não houve modelos significativos.

4.6 ANÁLISE DOS COMPOSTOS META E PARA-SUBSTITUÍDOS

4.6.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

4.6.2 Atividade inibitória sobre HSV2 TK

Não houve modelos significativos.

4.7 ANÁLISE DOS COMPOSTOS DI-SUBSTITUÍDOS

4.7.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

4.7.2 Atividade inibitória sobre HSV2 TK

Não houve modelos significativos.

4.8 ANÁLISE GLOBAL

4.8.1 Atividade inibitória sobre HSV1 TK

O melhor modelo capaz de representar a atividade de toda a série de N^2 -fenilguaninas, com exceção dos compostos **3** (4'-Br) e **30** (3,4'-Cl₂), sobre HSV1 TK foi o modelo de Gambino, apresentado na Equação (97). A Equação (104) mostra o modelo de Gambino recalculado com base na matriz de parâmetros físico-químicos apresentada na Tabela 19.

$$\begin{aligned} \log(1/C) = & 0,594 (\pm 0,229) \pi_m - 0,414 (\pm 0,200) \pi_p + 1,048 (\pm 0,746) \sigma_m + \\ & + 2,381 (\pm 0,871) R_p + 5,077 (\pm 0,244) \end{aligned} \quad (104)$$

(n = 34; r = 0,897; s = 0,399; F_(4,29) = 29,767; p < 0,00001)

(Q² = 0,741; S_{PRESS} = 0,459; SDEP = 0,430)

Na Equação (104), a ordem de entrada de cada variável em relação à sua capacidade de explicar a variância total da atividade inibitória sobre HSV1 TK é a seguinte: π_m 36%, R_p +21%, π_p +5% e σ_m +3%. Individualmente, a capacidade de explicação da variância dos parâmetros presentes na Equação (104) é a seguinte: π_m 36%, R_p 27%, π_p 25% e σ_m 20%. Estes números indicam que a atividade dos derivados das N^2 -fenilguaninas sobre HSV1 TK é basicamente controlada pelos parâmetros lipofílico e eletrônico nas posições meta e para. Substituintes mais lipofílicos na posição meta e mais hidrofílicos na posição para do anel fenila intensificam a atividade inibitória. Substituintes na posição para capazes de extrair elétrons via mecanismo de ressonância intensificam a atividade inibitória. Substituintes na posição meta capazes de extrair elétrons por efeito indutivo intensificam a atividade inibitória.

Na Equação (104), os valores dos coeficientes de π_m e π_p têm significado importante em termos do ambiente de interação enzima-inibidor. A lipofilicidade é geralmente considerada como sendo o resultado da variação da energia livre do processo de dessolvatação de um composto em seu movimento da fase aquosa para a biofase (Hansch & Leo, 1995). Coeficientes de log P e π próximos da unidade geralmente são interpretados como sendo consequência do processo de dessolvatação completa do composto. Isto pode significar que a interação deste com a biofase dá-se através da penetração completa do fármaco na estrutura onde se localiza o seu receptor. Coeficientes de log P e π próximos de 0,5, por outro lado, são interpretados como sendo consequência do processo de dessolvatação parcial do composto, que pode ser decorrente do fato de o sítio ativo estar localizado na superfície do receptor. No entanto, essa interpretação não é infalível. Pode acontecer de o sítio de ligação do fármaco, bem como o caminho que dá acesso ao sítio, serem espaçosos suficientes para impedir que ocorra dessolvatação apreciável.

O sinal dos coeficientes de π_m e π_p na Equação (104) indica que as regiões do sítio ativo de HSV1 TK próximas às posições meta e para do anel fenila são, respectivamente, lipofílicas e hidrofílicas. No entanto, a presença de σ_m na Equação (104) indica que substituintes polares em meta também intensifica a atividade inibitória das N^2 -fenilguaninas. Isso pode significar que a região do sítio ativo próxima à posição meta tenha caráter duplo ou anfifílico (lipofílico por parte de π_m e polar por parte de σ_m), sendo que predomina a natureza lipofílica (ver também a Seção 5.5.6, "Natureza das regiões do sítio ativo próximas às posições meta e para", pag.218).

A Tabela 17 mostra os valores calculados da atividade inibitória das N^2 -fenilguaninas sobre HSV1 TK, bem como os respectivos resíduos, produzidos pela

Equação (104). O gráfico da atividade observada em função da atividade calculada pela Equação (104) é mostrado na Figura 31.

Tabela 17 - Atividades inibitórias observada e prevista do conjunto global dos derivados da N²-fenilguanina sobre HSV1 TK, e os respectivos resíduos, calculados através da Equação (104).

Comp.	Substituinte		log 1/C ₅₀		Resíduo
	X	Y	Observado	Previsto	
1	CF ₃	H	6,82	6,06	0,77
2	CH ₂ N ₃	H	6,10	5,74	0,36
4	CH ₂ Br	H	6,00	5,67	0,33
5	I	H	5,96	6,11	-0,15
6	CH ₂ Cl	H	5,85	5,29	0,56
7	Cl	H	5,82	5,89	-0,07
8	Br	H	5,82	6,00	-0,18
9	Et	H	5,60	5,61	-0,01
10	Br	Br	5,52	5,12	0,40
11	n-Pr	H	5,48	5,93	-0,45
12	CN	H	5,30	5,33	-0,03
13	Et	Me	5,16	4,95	0,21
14	Cl	F	5,10	4,90	0,20
15	H	H	5,10	5,08	0,02
16	Cl	Me	4,92	5,23	-0,31
17	-CH ₂ CH ₂ CH ₂ -		4,82	4,77	0,05
18	H	Et	4,70	4,30	0,40
19	CF ₃	Cl	4,70	5,30	-0,60
20	CH ₂ OH	H	4,60	4,47	0,13
21	CH ₂ NH ₂	H	4,52	4,43	0,09
22	H	CF ₃	4,52	5,09	-0,57
23	H	n-Pr	4,52	4,10	0,42
24	-CH=CHCH=CH-		4,40	5,17	-0,77
25	Me	n-Bu	4,40	4,10	0,30
26	H	I	4,40	4,04	0,36
27	H	Me	4,30	4,42	-0,12
28	F	F	4,30	4,53	-0,23
29	H	n-Bu	4,30	3,84	0,46
31	H	Cl	4,00	4,33	-0,33
32	H	OH	3,85	3,69	0,16
33	H	i-Pr	3,70	3,99	-0,29
34	CH ₂ OH	Me	3,57	3,81	-0,24
35	H	F	3,52	4,09	-0,57
36	H	n-C ₆ H ₁₁	3,30	3,64	-0,34

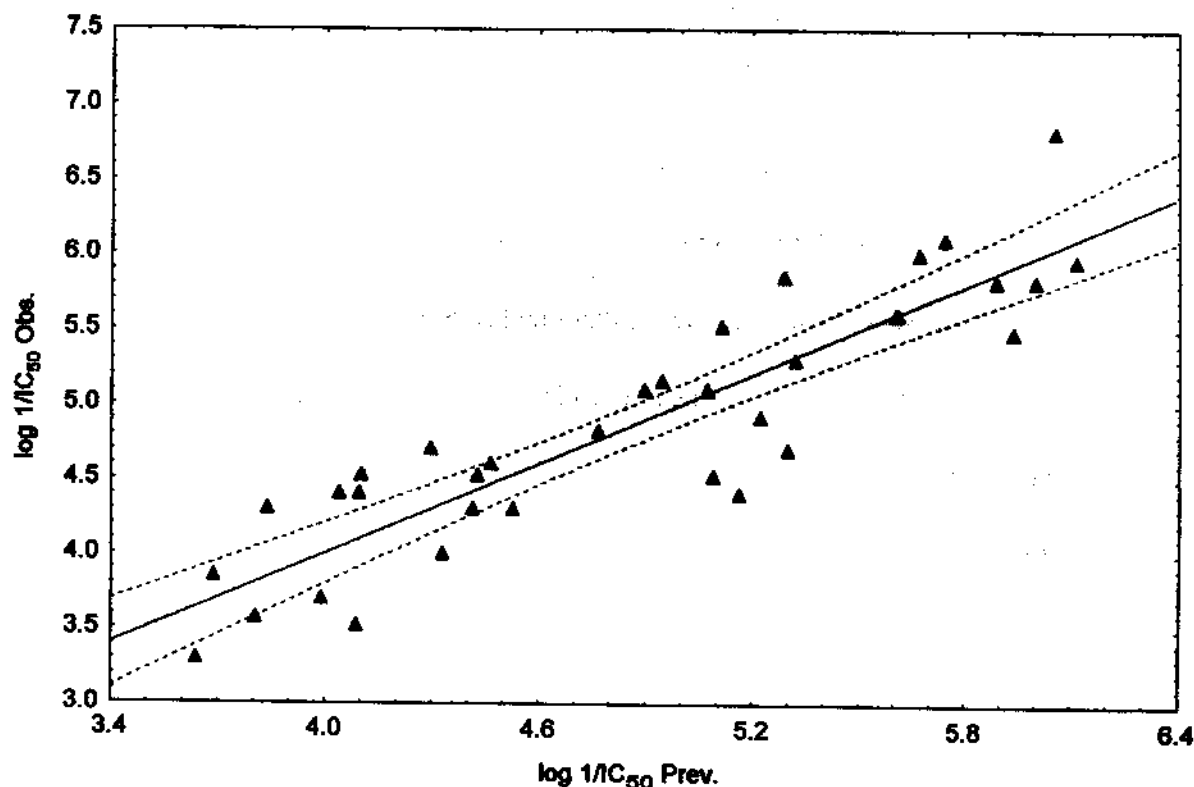


Figura 31 - Atividade inibitória observada do conjunto global dos derivados da N^2 -fenilguanina sobre HSV1 TK em função da atividade prevista pela Equação (104).

4.8.2 Atividade inibitória sobre HSV2 TK

O melhor modelo capaz de explicar a atividade das N^2 -fenilguaninas sobre HSV2 TK, com exceção dos compostos 3 (3'-H, 4'-Br), 30 (3'-Cl, 4'-Cl) e 35 (3'-H, 4'-F) é representado pela Equação (105).

$$\log(1/C) = 0,410 (\pm 0,211) \pi_m - 0,535 (\pm 0,197) \pi_p + 1,928 (\pm 0,687) \sigma_m + \\ + 1,123 (\pm 0,828) R_p + 5,269 (\pm 0,224) \quad (105)$$

$$(n = 32; r = 0,913; s = 0,362; F_{(4,27)} = 33,822; p < 0,00001)$$

$$(Q^2 = 0,773; S_{PRESS} = 0,423; SDEP = 0,395)$$

A Equação (105) é equivalente à Equação (98), proposta por Gambino e colaboradores (pag. 129), e bastante semelhante à Equação (104), o que reforça a

indicação de que a interação das N^2 -fenilguaninas com HSV1 TK e HSV2 TK dá-se pelo mesmo mecanismo de ação.

Na construção da Equação (105), a ordem de entrada de cada variável em relação à explicação da variância total da atividade inibitória sobre HSV2 TK é a seguinte: σ_m 43%, π_p +24%, π_m +11% e R_p +5%. Individualmente, a capacidade de explicação da variância dos parâmetros presentes na Equação (105) é a seguinte: σ_m 43%, π_p 41%, π_m 28%, R_p 6%.. Estes números indicam que a atividade dos derivados das N^2 -fenilguaninas sobre HSV2 é controlada pelos parâmetros eletrônico e lipofílico, sendo que os substituintes capazes de extrair elétrons por efeito indutivo na posição meta e substituintes hidrofílicos na posição para e lipofílicos na posição meta do anel fenila intensificam a atividade.

Podem-se comparar as Equações (104) e (105) sob os aspectos qualitativo e quantitativo. Qualitativamente essas equações são muito semelhantes, considerando-se que os parâmetros que nelas aparecem são os mesmos e que seus coeficientes têm aproximadamente a mesma ordem de grandeza. Esta observação indica que o processo de interação das N^2 -fenilguaninas com HSV1 TK e HSV2 TK é muito semelhante, ou seja, acontece em sítios ativos equivalentes e através do mesmo modo de ligação. Do ponto de vista quantitativo, percebem-se pequenas diferenças nas importâncias relativas dos parâmetros que explicam a atividade inibitória sobre HSV1 TK e HSV2 TK. Os parâmetros relativos à posição meta, π_m e σ_m , apresentam inversão da capacidade de explicação da variância nas Equações (104) e (105) (36% e 20% em HSV1 TK e 28% e 43% em HSV2 TK, respectivamente). Seguindo o mesmo raciocínio que para HSV1 TK (ver Seção 4.8.1, "Atividade inibitória sobre HSV1 TK", pag.142), o sítio ativo de HSV2 TK, na região próxima à posição meta, deve apresentar natureza anfifílica, sendo que predomina o caráter polar. Isto pode significar a existência de pequenas diferenças nas estruturas dos sítios ativos das duas enzimas, principalmente na região próxima à posição

meta. No entanto, a natureza do ambiente de interação do anel fenila, ou seja, anfifílica em meta e hidrofílica em para, parece ser idêntica nas duas enzimas.

A Tabela 18 mostra os valores calculados da atividade inibitória das N^2 -fenilguaninas sobre HSV1 TK, bem como os respectivos resíduos, produzidos pela Equação (105). O gráfico da atividade observada em função da atividade calculada pela Equação (105) é mostrado na Figura 32.

Tabela 18 - Atividades inibitórias observada e prevista do conjunto global dos derivados da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (105).

Comp.	Substituinte		log 1/C ₅₀		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	6,46	0,54
2	CH ₂ N ₃	H	5,82	5,73	0,09
4	CH ₂ Br	H	5,82	5,82	0,10
5	I	H	5,96	6,40	-0,44
6	CH ₂ Cl	H	5,87	5,55	0,32
7	Cl	H	6,46	6,27	0,19
8	Br	H	6,22	6,37	-0,15
9	Et	H	5,40	5,55	-0,15
10	Br	Br	5,89	5,67	0,22
11	n-Pr	H	5,52	5,79	-0,27
12	CN	H	6,40	6,11	0,29
13	Et	Me	5,10	5,05	0,05
14	Cl	F	5,60	5,76	-0,16
15	H	H	5,80	5,27	0,53
16	Cl	Me	5,60	5,77	-0,17
17	-CH ₂ CH ₂ CH ₂ -		5,00	4,62	0,38
18	H	Et	4,60	4,55	0,05
19	CF ₃	Cl	5,30	5,87	-0,57
20	CH ₂ OH	H	4,70	4,85	-0,15
21	CH ₂ NH ₂	H	4,70	4,78	-0,08
22	H	CF ₃	4,40	4,98	-0,58
23	H	n-Pr	4,33	4,28	0,05
24	-CH=CHCH=CH-		5,00	5,23	-0,23
26	H	I	5,05	4,40	0,65
27	H	Me	5,22	4,77	0,45
28	F	F	5,66	5,47	0,19
29	H	n-Bu	4,30	3,96	0,34
31	H	Cl	4,30	4,68	-0,38
32	H	OH	4,70	4,84	-0,14
33	H	i-Pr	4,00	4,24	-0,24
34	CH ₂ OH	Me	3,70	4,34	-0,64
36	H	n-C ₅ H ₁₁	3,80	3,68	-0,08

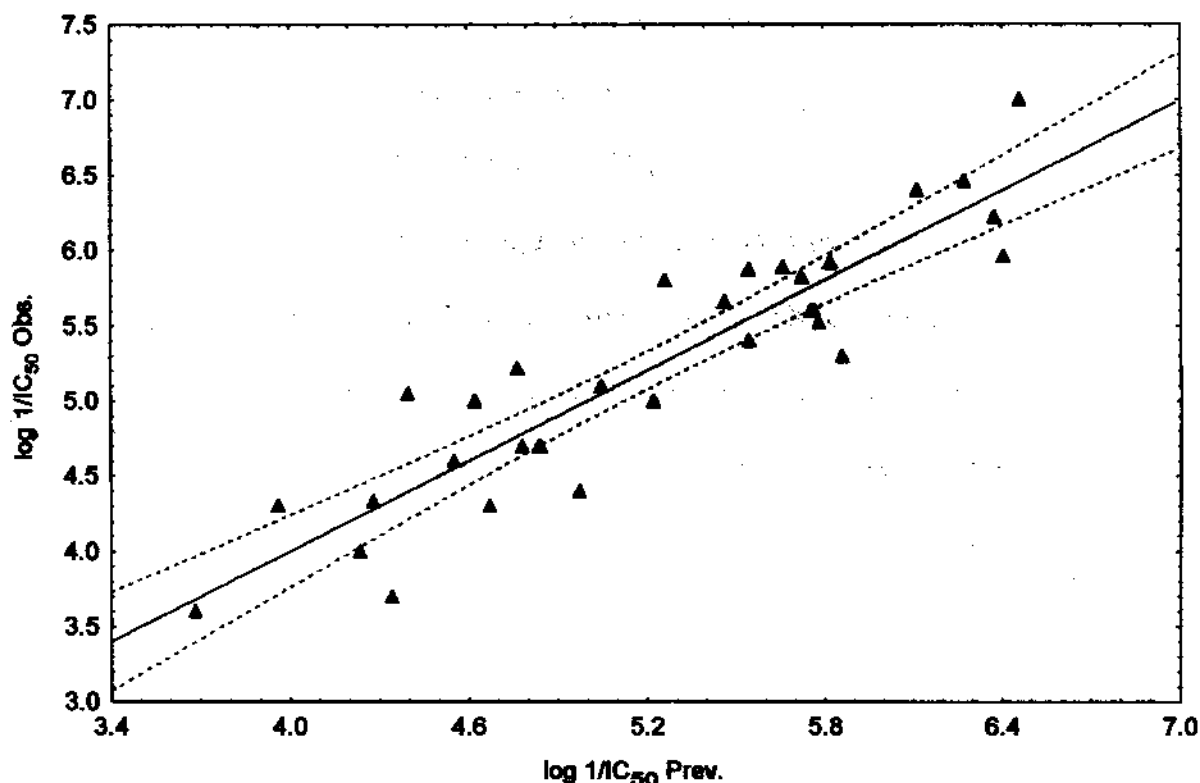


Figura 32 - Atividade inibitória observada do conjunto global dos derivados da N^2 -fenilguanina sobre HSV2 TK em função da atividade prevista pela Equação (105).

4.9 CONCLUSÕES DO ESTUDO QSAR COM PARÂMETROS CLÁSSICOS

- (a) A atividade dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK é governada principalmente pelo parâmetro eletrônico, sendo que substituintes capazes de extrair elétrons do anel fenila intensificam a atividade. A lipofilicidade exerce influência menor na atividade desses compostos sobre HSV2, sendo que substituintes meta mais lipofílicos intensificam a atividade inibitória;
- (b) A atividade inibitória global das N^2 -fenilguaninas sobre HSV1 TK e HSV2 TK depende dos parâmetros π_m , π_p , σ_m e R_p ;
- (c) Qualitativamente, os parâmetros π_m , π_p , σ_m e R_p exercem influência semelhante sobre a atividade das N^2 -fenilguaninas sobre HSV1 TK e HSV2 TK (os coeficientes desses

parâmetros nas Equações (104) e (105) têm os mesmos sinais: $+\pi_m$, $-\pi_p$, $+\sigma_m$ e $+R_p$).

Isto pode ser interpretado como uma indicação da semelhança do mecanismo de ação das N^2 -fenilguaninas sobre HSV1 TK e HSV2 TK;

- (d) Quantitativamente, ocorrem diferenças na influência dos parâmetros π_m , π_p , σ_m e R_p sobre a atividade das N^2 -fenilguaninas sobre HSV1 TK e HSV2 TK: a atividade inibitória global sobre HSV1 TK é influenciada pelos parâmetros lipofílico em meta ($+\pi_m$) e hidrofílico em para ($+\pi_p$), pelo efeito de ressonância em para ($+R_p$) e pelo efeito indutivo em meta ($+\sigma_m$) de forma aproximadamente equivalente, enquanto que a atividade inibitória global sobre HSV2 TK é principalmente influenciada pelo efeito indutivo em meta ($+\sigma_m$) e pela lipofilicidade em meta ($+\pi_m$) e para ($-\pi_p$). Isto pode ser interpretado como sendo reflexo de pequenas diferenças nos sítios ativos dessas enzimas;
- (e) A região do sítio ativo de HSV1 TK, próxima à posição meta do anel fenila das N^2 -fenilguaninas, possui natureza anfifílica (caráter duplo), predominando o caráter apolar ($+\pi_m$ e $+\sigma_m$), enquanto que a região próxima à posição para possui natureza hidrofílica ($-\pi_p$);
- (f) A região do sítio ativo de HSV2 TK, próxima à posição meta do anel fenila das N^2 -fenilguaninas, também possui natureza anfifílica, predominando o caráter polar ($+\sigma_m$ e $+\pi_m$), enquanto que a região próxima à posição para possui natureza hidrofílica ($-\pi_p$);

A Figura 33 mostra, de forma resumida, as conclusões acerca da natureza do sítio ativo das N^2 -fenilguaninas de HSV1 TK e HSV2 TK.

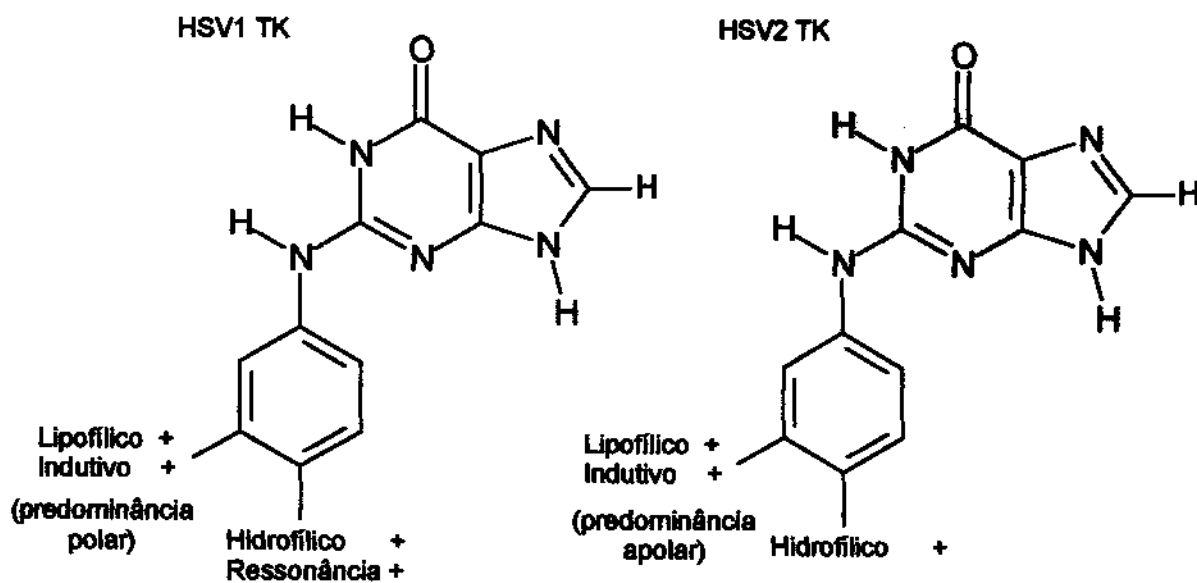
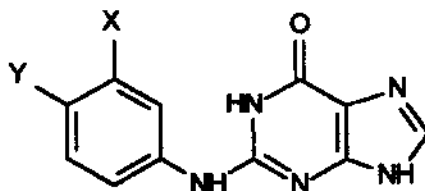


Figura 33 - Conclusões gerais do estudo QSAR com parâmetros clássicos. Lipofílico+, substituintes mais lipofílicos aumentam a atividade; Hidrofílico+, substituintes mais hidrofílicos aumentam a atividade; Ressonância+, substituintes capazes de atrair elétrons por mecanismo de ressonância aumentam a atividade; Indutivo+, Substituintes capazes de atrair elétrons por mecanismo de indução aumentam a atividade.

4.10 LISTAGEM DOS PARÂMETROS CLÁSSICOS

A Tabela 19 mostra os valores das constantes dos substituintes utilizadas no presente capítulo. A matriz de correlação dessas constantes é mostrada na Tabela 20.

Tabela 19 - Parâmetros testados na análise clássica de QSAR (Hansch et al, 1995)



No.	X	Y	Log 1/IC ₅₀ ^a		π_m^b	π_p	RM _m ^c	RM _p	σ_m^d	σ_p
			HSV1	HSV2						
1	CF ₃	H	6,82	7,00	0,88	0,00	0,50	0,10	0,43	0,00
2	CH ₂ N ₃	H	6,10	5,82	^a 1,12	0,00	1,47	0,10	0,00	0,00
3	H	Br	6,00	6,22	0,00	0,86	0,10	0,89	0,00	0,23
4	CH ₂ Br	H	6,00	5,92	0,79	0,00	1,34	0,10	0,12	0,00
5	I	H	5,96	5,96	1,12	0,00	1,39	0,10	0,35	0,00
6	CH ₂ Cl	H	5,85	5,87	0,17	0,00	1,05	0,10	0,11	0,00
7	Cl	H	5,82	6,46	0,71	0,00	0,60	0,10	0,37	0,00
8	Br	H	5,82	6,22	0,86	0,00	0,89	0,10	0,39	0,00
9	Et	H	5,60	5,40	1,02	0,00	1,03	0,10	-0,07	0,00
10	Br	Br	5,52	5,89	0,86	0,86	0,89	0,89	0,39	0,23
11	n-Pr	H	5,48	5,52	1,55	0,00	1,50	0,10	-0,06	0,00
12	CN	H	5,30	6,40	-0,57	0,00	0,63	0,10	0,56	0,00
13	Et	Me	5,16	5,10	1,02	0,56	1,03	0,56	-0,07	-0,17
14	Cl	F	5,10	5,60	0,71	0,14	0,60	0,09	0,37	0,06
15	H	H	5,10	5,80	0,00	0,00	0,10	0,10	0,00	0,00
16	Cl	Me	4,92	5,60	0,71	0,56	0,60	0,56	0,37	-0,17
17	-CH ₂ CH ₂ CH ₂ -		4,82	5,00	0,60	0,60	0,70	0,70	-0,26	-0,26
18	H	Et	4,70	4,60	0,00	1,02	0,10	1,03	0,00	-0,15
19	CF ₃	Cl	4,70	5,30	0,88	0,71	0,50	0,60	0,43	0,23
20	CH ₂ OH	H	4,60	4,70	-1,03	0,00	0,72	0,10	0,00	0,00
21	CH ₂ NH ₂	H	4,52	4,70	-1,04	0,00	0,91	0,10	-0,03	0,00
22	H	CF ₃	4,52	4,40	0,00	0,88	0,10	0,50	0,00	0,54
23	H	n-Pr	4,52	4,33	0,00	1,55	0,10	1,50	0,00	-0,13
24	-CH=CHCH=CH-		4,40	5,00	0,66	0,66	0,88	0,88	0,04	0,04
25	Me	n-Bu	4,40	^f	0,56	2,13	0,56	1,96	-0,07	-0,16
26	H	I	4,40	5,05	0,00	1,12	0,10	1,39	0,00	0,18
27	H	Me	4,30	5,22	0,00	0,56	0,10	0,56	0,00	-0,17
28	F	F	4,30	5,66	0,14	0,14	0,09	0,09	0,34	0,06
29	H	n-Bu	4,30	4,30	0,00	2,13	0,10	1,96	0,00	-0,16
30	Cl	Cl	4,00	4,52	0,71	0,71	0,60	0,60	0,37	0,23
31	H	Cl	4,00	4,30	0,00	0,71	0,10	0,60	0,00	0,23
32	H	OH	3,85	4,70	0,00	-0,67	0,10	0,28	0,00	-0,37
33	H	i-Pr	3,70	4,00	0,00	1,53	0,10	1,50	0,00	-0,15
34	CH ₂ OH	Me	3,57	3,70	-1,03	0,56	0,72	0,56	0,00	-0,17
35	H	F	3,52	3,59	0,00	0,14	0,10	0,09	0,00	0,06
36	H	n-C ₅ H ₁₁	3,30	3,60	0,00	^g 2,67	0,10	2,42	0,00	-0,15

^a IC₅₀: concentração molar capaz de provocar 50% de inibição da atividade inibitória da enzima; ^b

π : constante lipofílica de Hansch-Fujita (Fujita et al, 1964); ^c RM: refratividade molar (Martin, 1978);

^d σ : constante eletrônica de Hammett (Hammett, 1937); ^e Valor calculado; ^f Valor não medido; ^g

Obtido da referência (Kubinyi, 1993); Subscritos m e p: posições meta e para do anel fenila.

Tabela 19 - Parâmetros testados na análise clássica de QSAR (Hansch et al, 1995) (continuação).

No.	X	Y	F _m ^a	R _p ^b	L _m ^c	L _p	B _{1m} ^d	B _{1p}	B _{5m} ^e	B _{5p}
1	CF ₃	H	0,38	0,00	3,30	2,06	1,99	1,00	2,61	1,00
2	CH ₂ N ₃	H	^f 0,15	0,00	^g 4,62	2,06	^h 1,50	1,00	^g 4,80	1,00
3	H	Br	0,00	-0,22	2,06	3,82	1,00	1,85	1,00	1,95
4	CH ₂ Br	H	0,14	0,00	4,09	2,06	1,52	1,00	3,75	1,00
5	I	H	0,42	0,00	4,23	2,06	2,15	1,00	2,15	1,00
6	CH ₂ Cl	H	0,13	0,00	3,89	2,06	1,52	1,00	3,46	1,00
7	Cl	H	0,42	0,00	3,52	2,06	1,80	1,00	1,80	1,00
8	Br	H	0,45	0,00	3,82	2,06	1,95	1,00	1,95	1,00
9	Et	H	0,00	0,00	4,11	2,06	1,52	1,00	3,17	1,00
10	Br	Br	0,45	-0,22	3,82	3,82	1,95	1,95	1,95	1,95
11	n-Pr	H	0,01	0,00	4,92	2,06	1,52	1,00	3,49	1,00
12	CN	H	0,51	0,00	4,23	2,06	1,60	1,00	1,60	1,00
13	Et	Me	0,00	-0,18	4,11	2,87	1,52	1,52	3,17	2,04
14	Cl	F	0,42	-0,39	3,52	2,65	1,80	1,35	1,80	1,35
15	H	H	0,00	0,00	2,06	2,06	1,00	1,00	1,00	1,00
16	Cl	Me	0,42	-0,18	3,52	2,87	1,80	1,52	1,80	2,04
17	-CH ₂ CH ₂ CH ₂ -		-0,20	-0,06	^h 4,11	^h 4,11	^h 1,52	^h 1,52	^h 3,17	^h 3,17
18	H	Et	0,00	-0,15	2,06	4,11	1,00	1,52	1,00	3,17
19	CF ₃	Cl	0,38	-0,19	3,30	3,52	1,99	1,80	2,61	1,80
20	CH ₂ OH	H	0,03	0,00	3,97	2,06	1,52	1,00	2,70	1,00
21	CH ₂ NH ₂	H	0,04	0,00	4,02	2,06	1,52	1,00	3,05	1,00
22	H	CF ₃	0,00	0,16	2,06	3,30	1,00	1,99	1,00	2,61
23	H	n-Pr	0,00	-0,14	2,06	4,92	1,00	1,52	1,00	3,49
24	-CH=CHCH=CH-		0,07	-0,03	ⁱ 4,29	ⁱ 4,29	ⁱ 1,60	ⁱ 1,60	ⁱ 3,09	ⁱ 3,09
25	Me	n-Bu	0,01	-0,15	2,87	6,17	1,52	1,52	2,04	4,54
26	H	I	0,00	-0,24	2,06	4,23	1,00	2,15	1,00	2,15
27	H	Me	0,00	-0,18	2,06	2,87	1,00	1,52	1,00	2,04
28	F	F	0,45	-0,39	2,65	2,65	1,35	1,35	1,35	1,35
29	H	n-Bu	0,00	-0,15	2,06	6,17	1,00	1,52	1,00	4,54
30	Cl	Cl	0,42	-0,19	3,52	3,52	1,80	1,80	1,80	1,80
31	H	Cl	0,00	-0,19	2,06	3,52	1,00	1,80	1,00	1,80
32	H	OH	0,00	-0,70	2,06	2,74	1,00	1,35	1,00	1,93
33	H	i-Pr	0,00	-0,19	2,06	4,11	1,00	1,90	1,00	3,17
34	CH ₂ OH	Me	0,03	-0,18	3,97	2,87	1,52	1,52	2,70	2,04
35	H	F	0,00	-0,39	2,06	2,65	1,00	1,35	1,00	1,35
36	H	n-C ₆ H ₁₁	0,00	-0,14	2,06	6,97	1,00	1,52	1,00	4,94

^a F, parâmetro eletrônico de Swain-Lupton para efeito indutivo; ^b R: constante eletrônica de Swain-Lupton para efeito de ressonância (Swain & Lupton, 1968); ^c L, comprimento do substituinte ao longo do eixo de substituição; ^d B₁, menor largura do substituinte, medida a partir do eixo de L e ortogonalmente a este; ^e B₅, maior largura do substituinte, medida a partir do eixo de L e ortogonalmente a este. As constantes L, B₁ e B₅ são denominadas STERIMOL e foram propostas por Verloop (ver detalhes em Kubinyi, 1993); ^f Valor estimado; ^g Valor referente ao substituinte NNN; ^h Valor referente ao substituinte -CH₂CH₃; ⁱ Valor referente ao substituinte -CH=CH₂. Subscritos m e p: posições meta e para do anel fenila.

Tabela 20 - Matriz de correlação dos parâmetros utilizados na análise clássica. Valores em termos de r^2 , quadrado do coeficiente de correlação. As correlações iguais ou superiores a 0,5 são mostradas em negrito.

[illegible]

CAPÍTULO 5

MODELAGEM MOLECULAR E CONFORMAÇÃO BIOATIVA

5.1 INTRODUÇÃO

Todo o trabalho de modelagem molecular que se segue foi realizado no *Physical and Theoretical Chemistry Laboratory (PTCL), University of Oxford*, sob a supervisão do Prof. Dr. William Graham Richards. Além disso, sua execução somente foi possível após a obtenção dos arquivos HSV1_TK_dT_2 e tk_ganc_nov8g, em formato PDB (*Protein Data Bank*). O primeiro arquivo contém a estrutura cristalográfica de HSV1 TK, determinada através de difração de raios-X com 2,8 Å de resolução, em complexo com seu substrato natural, 2'-deoxitimidina, e um grupo sulfato (Brown et al, 1995). O segundo arquivo contém a estrutura cristalográfica de HSV1 TK, determinada com 2,2 Å de resolução, em complexo com ganciclovir, seu inibidor competitivo, e um grupo sulfato. Esses arquivos foram gentilmente cedidos pelo Dr. Mark R. Sanderson e Dr. John Champness, do *Division of Biomedical Sciences, The Randall Institute, King's College, Londres*. Até o momento da cessão, a estrutura cristalográfica de HSV1 TK não estava disponível ao público nos bancos de dados cristalográficos conhecidos. A Tabela 21 mostra o conteúdo dos arquivos HSV1_TK_dT_2 e tk_ganc_nov8g.

Tabela 21 - Conteúdo dos arquivos HSV1_TK_dT_2 e tk_ganc_nov8g, fornecidos pelo Dr. M.R. Sanderson, com a estrutura cristalográfica da timidina-cinase do vírus herpes simples do tipo 1

HSV1_TK_dT_2			tk_ganc_nov8g		
Segmento	Quantidade	No. de átomos	Segmento	Quantidade	No. de átomos
HSV1 TK	2	4679	HSV1 TK	2	5735
2'-deoxitimidina	2	34	ganciclovir	2	48
Sulfato	2	10	Sulfato	2	10
Água	45	45	Água	173	519
Total	—	4768	Total	—	6312

O arquivo HSV1_TK_dT_2 foi obtido no início do mês de setembro/96, enquanto que o segundo arquivo só foi obtido no final de novembro/96. Devido ao fato de dispor-se apenas de doze meses para a realização desta parte do trabalho, sua maior parte foi feita a partir da estrutura contida no primeiro arquivo.

Wild et al (1995) também determinaram a estrutura tridimensional de HSV1 TK, com 3,0 Å de resolução, em complexo com 2'-deoxitimidina e ATP. Infelizmente não foi possível obter as coordenadas cartesianas dessa determinação.

5.2 ANÁLISE ESTRUTURAL DE HSV1 TK

5.2.1 Estrutura primária de HSV1 TK

A estrutura de HSV1 TK, presente no arquivo HSV1_TK_dT_2, apresenta duas subunidades, monômeros, que serão chamadas de HSV1_TK_A, que apresenta 304 aminoácidos, e HSV1_TK_B, que apresenta 308 aminoácidos. Sabe-se que a estrutura completa de cada uma dessas subunidades possui 376 aminoácidos, cujo peso molecular é aproximadamente 41000 uma (Irmiere, 1989). A estrutura cristalográfica que se dispõe apresenta, portanto, regiões que não foram adequadamente mapeadas pela difração de raios-X. Para definir a composição de aminoácidos das seqüências ausentes na estrutura cristalográfica, as seqüências das subunidades A e B serão comparadas com as seqüências previstas a partir do gene responsável pela codificação de HSV1 TK.

O Diagrama 1 mostra o alinhamento das seqüências de aminoácidos de HSV1 TK, disponíveis no *Swiss-Prot Protein Sequence Databank*⁸, com as seqüências de aminoácidos dos segmentos HSV1_TK_A e HSV1_TK_B, do arquivo HSV1_TK_dT_2. Utilizou-se


```

131      140      150      160      170      180      190
.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+
KITH_HSV11 PYAVTDAVLAPHIGGEAGSSHAPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
KITH_HSV1C PYAVTDAVLAPHIGGEAGSSHAPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
KITH_HSV1E PYAVTDAVLAPHIGGEAGSSHAPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
KITH_HSV1K PYAVTDAVLAPHIGGEAGSSHAPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
KITH_HSV1S PYAVTDAVLAPHIGGEAGSSHAPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
HSV1_TK_A PYAVTDAVLAPHIGGEAGS-APPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
HSV1_TK_B PYAVTDAVLAPHIGGEAGS-APPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP

Consensus: PYAVTDAVLAPHIGGEAGShAPPALTTLIFDRHPAIALLCYPAARYLMGSMTPOAVLAFVALIP
Identity: *****

196      205      215      225      235      245      255
.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+
KITH_HSV11 PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
KITH_HSV1C PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
KITH_HSV1E PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
KITH_HSV1K PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
KITH_HSV1S PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
HSV1_TK_A PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
HSV1_TK_B PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG

Consensus: PTLPGTNIVLGALPEDRHIDRLAKRQRPGERLDLAMLAAIRRVYGLLANTVRYLQGGGSWREDWG
Identity: *****

261      270      280      290      300      310      320
.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+
KITH_HSV11 QLSGAAVPPQGAEPQSNAGPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRPMHVF
KITH_HSV1C QLSGTAVPPQGAEPQSNAGPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRPMHVF
KITH_HSV1E QLSGTAVPPQGAEPQSNAGPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRPMHVF
KITH_HSV1K QLSGTAVPPQGAEPQSNAGPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRPMHVF
KITH_HSV1S QLSGTAVPPQGAEPQSNAGPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRPMHVF
HSV1_TK_A Q-----PRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRSMHVF
HSV1_TK_B QL-----AGPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRSMHVF

Consensus: Qlsq avppqgaepqsnagPRPHIGDTLFTLFRAPELLAPNGDLYNVFANALDVLAKRLRPMHVF
Identity: *

326      335      345      355      365      375      385
.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+
KITH_HSV11 ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN
KITH_HSV1C ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN
KITH_HSV1E ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN
KITH_HSV1K ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN
KITH_HSV1S ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN
HSV1_TK_A ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGE--
HSV1_TK_B ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGE--

Consensus: ILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN
Identity: *****

```

Legenda:

```

~ - Ácido           @ - Amida           $ - Básico           % - Hidroxila
! - Hidrofóbico    # - Aromático        ^ - Prolina         & - Enxofre

```

Nota: O cálculo do percentual de homologia funcional (HF%) do alinhamento das diversas seqüências ao longo deste trabalho será feito considerando-se a classificação dos resíduos de acordo com os símbolos acima.

KITH_HSV11: Vírus herpes simples (Tipo 1, Linhagem 17) (McGeoch et al, 1988)
KITH_HSV1C: Vírus herpes simples (Tipo 1, Linhagem CL101) (Wagner et al, 1981)
KITH_HSV1E: Vírus herpes simples (Tipo 1, Linhagem HFEM) (Gompels e Minson, 1986)
KITH_HSV1K: Vírus herpes simples (Tipo 1, Linhagem KOS) (Irmiere, 1989)
KITH_HSV1S: Vírus herpes simples (Tipo 1, Linhagem SC16) (Darby et al, 1986)
HSV1_TK_A/B: Vírus herpes simples (Tipo 1, Linhagem SY211) (Brown et al, 1995)

A análise das seqüências de aminoácidos apresentadas no Diagrama 1 permite-nos estabelecer a seqüência completa de HSV1 TK, que foi arbitrariamente denominada

Tabela 22 - Definição dos elementos da estrutura secundária de HSV1 TK. Devido ao fato do mapeamento ter sido feito visualmente, algumas pequenas diferenças podem ser encontradas em relação a outros autores (Brown et al, 1995; Wild et al, 1995). Os resíduos 163 e 169 fazem conexão simples entre $\beta 3/\alpha 5$ e $\alpha 5/\alpha 6$, respectivamente. O resíduo 374 não está incorporado a $\alpha 15$ e não faz parte de qualquer elemento de estrutura secundária.

Elementos	Resíduos	Elementos	Resíduos
$\alpha 1$	61 - 72	$\beta 1$	46 - 56
$\alpha 2$	84 - 89	$\beta 2$	77 - 83
$\alpha 3$	95 - 109	$\beta 3$	156 - 162
$\alpha 4$	113 - 142	$\beta 4$	200 - 208
$\alpha 5$	164 - 168	$\beta 4a$	252 - 254
$\alpha 6$	170 - 179	$\beta 5$	321 - 329
$\alpha 7$	183 - 193	$\beta 6$	347 - 352
$\alpha 8$	209 - 219		
$\alpha 9$	229 - 251	$\Omega 1$	57 - 60
$\alpha 10$	255 - 262	$\Omega 2$	73 - 76
$\alpha 11$	283 - 286	$\Omega 3$	90 - 94
$\alpha 12$	287 - 293	$\Omega 4$	110 - 112
$\alpha 13$	305 - 320	$\Omega 5$	143 - 155
$\alpha 14$	332 - 346	$\Omega 6$	180 - 182
$\alpha 15$	356 - 373	$\Omega 7$	194 - 199
		$\Omega 8$	220 - 228
		$\Omega 9$	278 - 282
		$\Omega 10$	294 - 304
		$\Omega 11$	330 - 331
		$\Omega 12$	353 - 355

5.2.3 Estruturas terciária e quaternária de HSV1 TK

HSV1 TK possui estrutura quaternária dimérica, sendo que cada monômero, HSV1_TK_A e HSV1_TK_B, possui 376 aminoácidos. A Figura 34 mostra o esquema da estrutura terciária do monômero de HSV1 TK, que foi baseado na Figura 1 do artigo de Wild e colaboradores (1995). A região de interface do dímero é formada por α -hélices e voltas (*loops*). As α -hélices $\alpha 4$ e $\alpha 7$ estão localizadas na região central da interface.

Outros elementos que também estão presentes nessa região são as voltas $\Omega 3$ e $\Omega 10$ e a parte final de $\alpha 13$ e $\alpha 15$. A Figura 34 também mostra a presença de folha β (β -sheet) formada por cinco fitas β . Enzimas da classe adenil-cinase (ADK) também apresentam folha β similar a essa (Schulz et al, 1974). Essa folha β está situada próxima à superfície da estrutura protéica, na região diametralmente oposta à interface do dímero.

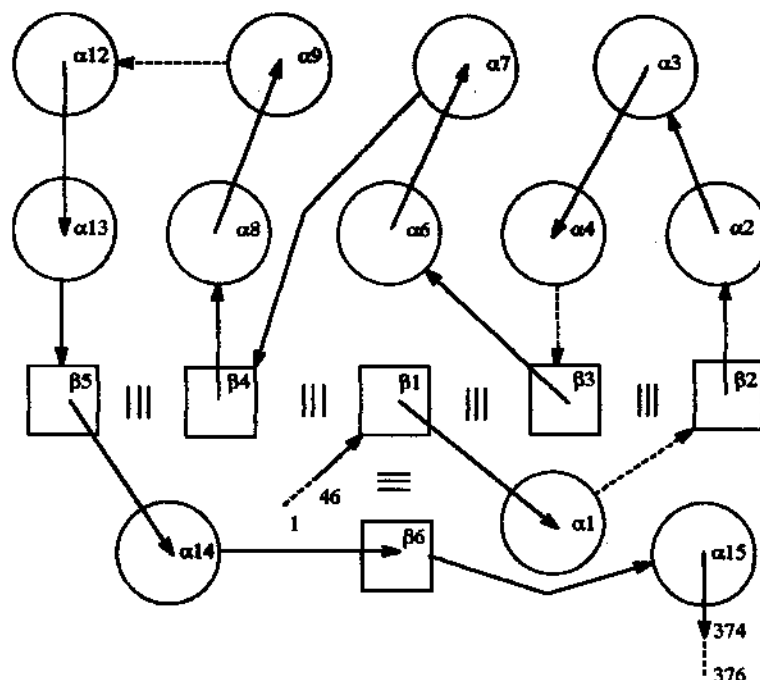


Figura 34 - Esquema da estrutura terciária do monômero de HSV1 TK, presente no arquivo HSV1_TK_dT_2. Círculos representam α -hélices e quadrados representam fitas β . Linhas tracejadas indicam sequências ausentes. A orientação espacial dos elementos presentes é indicada pela disposição das setas. Nota-se que todas as fitas β que formam a folha β maior estão projetadas para fora do plano do papel. A posição de cada elemento ao longo da sequência de aminoácidos pode ser observada no Diagrama 3 (pag. 159). Os elementos de estrutura secundária pequenos ou incompletos não foram incluídos neste esquema (Wild et al, 1995).

A Figura 35 mostra a representação gráfica da estrutura cristalográfica de HSV1 TK, presente no arquivo HSV1_TK_dT_2, em sua forma original e sem modificações.



Figura 35 - Representação gráfica da estrutura cristalográfica de HSV1 TK, presente no arquivo HSV1_TK_dT_2. A representação mostra a estrutura na forma original e sem modificações. A enzima apresenta-se na forma de dímero e com simetria C_2 (o eixo de simetria passa pelo centro do dímero, na direção vertical). Pode-se notar a presença do substrato 2'-deoxitimidina e do grupo sulfato nos sítios ativos de cada monômero. As bolas vermelhas isoladas representam os átomos de oxigênio das moléculas de água (45 átomos de oxigênio). As α -hélices são mostradas em vermelho, as fitas β em azul e as voltas e demais segmentos em cinza.

A Figura 36 mostra a representação gráfica da estrutura cristalográfica de HSV1 TK, presente no arquivo tk_gang_nov8g, em sua forma original e sem modificações.



Figura 36 - Representação gráfica da estrutura cristalográfica de HSV1 TK, presente no arquivo tk_gang_nov8g, cedido pelo Dr. M.R. Sanderson. A representação mostra a estrutura na forma original e sem modificações. As diferenças em relação à Figura 35 são a presença de ganciclovir no lugar de 2'-deoxitimidina, e a maior quantidade de moléculas de água (173).

A Figura 37 mostra a estrutura tridimensional do dímero de HSV1 TK.

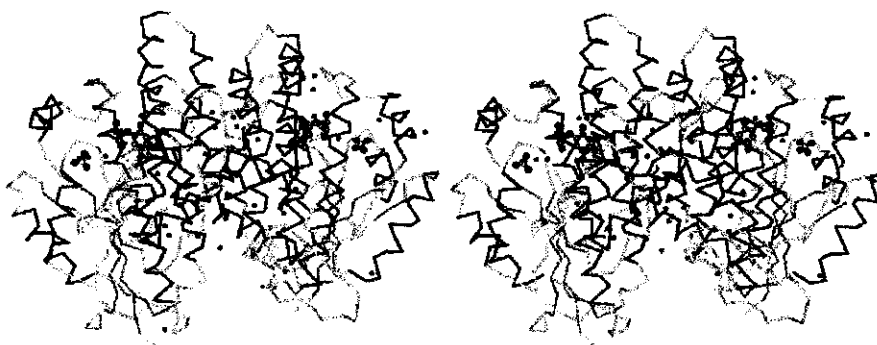


Figura 37 - Estereograma do dímero de HSV1 TK, construído a partir do arquivo HSV1_TK_dT_2.

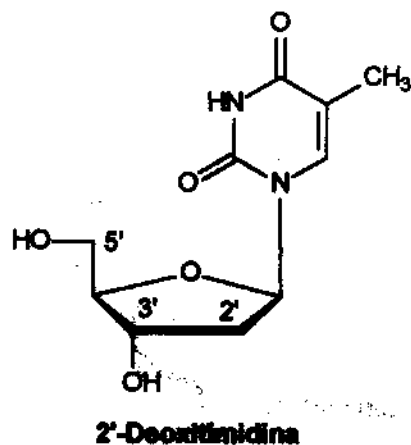
A Figura 38 mostra a estrutura tridimensional do monômero de HSV1 TK.



Figura 38 - Estereograma do monômero de HSV1 TK, construído a partir do arquivo HSV1_TK_dT_2.

5.2.4 Sítio de ligação da 2'-deoxitimidina

O sítio de ligação da 2'-deoxitimidina está localizado entre as α -hélices α_3 , α_4 e α_6 e a volta Ω_7 , que une α_8 e α_9 . O anel pirimidínico de 2'-deoxitimidina é mantido em sua posição através da formação de ponte de hidrogênio dupla com Gln 125 (ver Figura 39). Essa ponte de hidrogênio é similar à encontrada no emparelhamento adenina-timina, na dupla hélice da molécula de DNA. Lateralmente, o anel pirimidínico é equilibrado pelas interações com Tyr 172 e Met 128. A presença de Tyr 172 é muito importante na manutenção da posição correta do anel pirimidínico. A análise mutacional demonstrou que Tyr 172 pode ser substituída apenas por Phe (Munir et al, 1992), o que é consistente com aquela observação. O grupo SCH_3 de Met 128 aponta diretamente para o centro do anel pirimidínico, permitindo estabilidade extra para a manutenção do mesmo em sua posição.



O anel 2'-deoxiribose da 2'-deoxitimidina é mantido em sua posição basicamente pela formação de duas pontes de hidrogênio. A hidroxila ligada a C(3') forma ponte de hidrogênio com Tyr 101 e a hidroxila ligada a C(5') forma ponte de hidrogênio com Arg 163. Esta interação é responsável pelo posicionamento correto de HO-C(5') para fosforilação (Winter & Herdewijn, 1996).

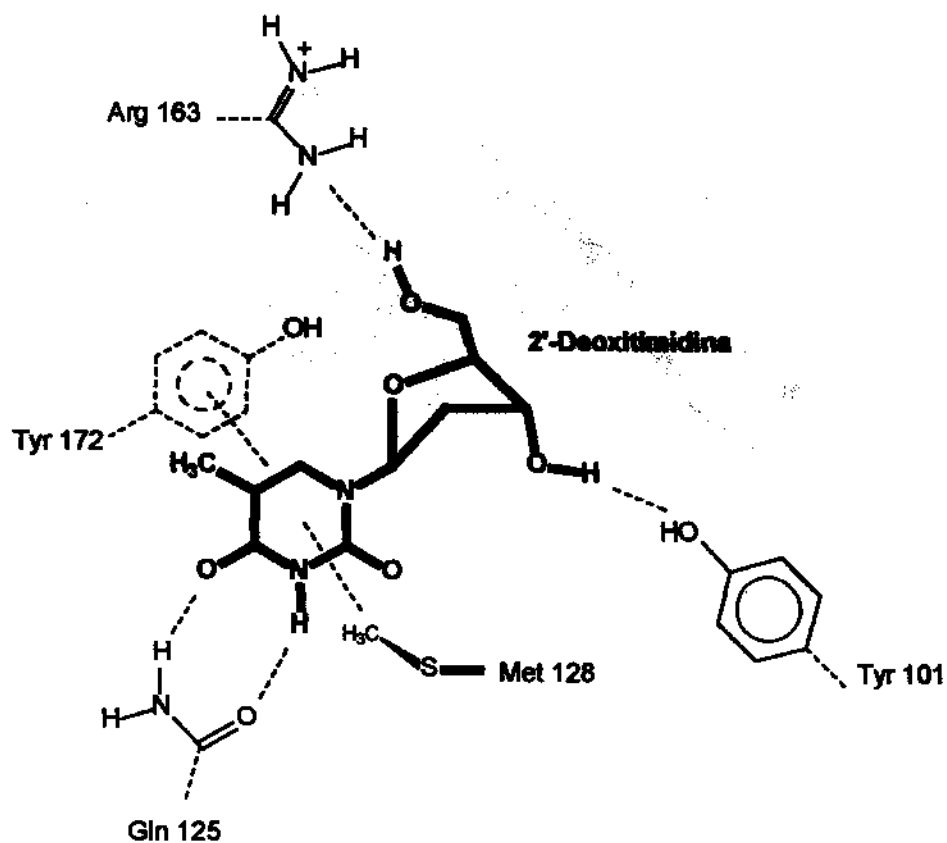


Figura 39 - Representação do sítio de ligação de 2'-deoxitimidina em HSV1 TK. Alguns átomos e ligações da molécula de 2'-deoxitimidina foram omitidos para maior clareza. O fragmento Tyr 172 encontra-se atrás do plano do papel, enquanto que Met 128 está acima do plano do papel. As linhas retas tracejadas, entre 2'-deoxitimidina e os diversos fragmentos, representam interações intermoleculares.

A Figura 40 mostra a representação gráfica do modo de ligação de 2'-deoxitimidina, referente ao esquema apresentado na Figura 39.

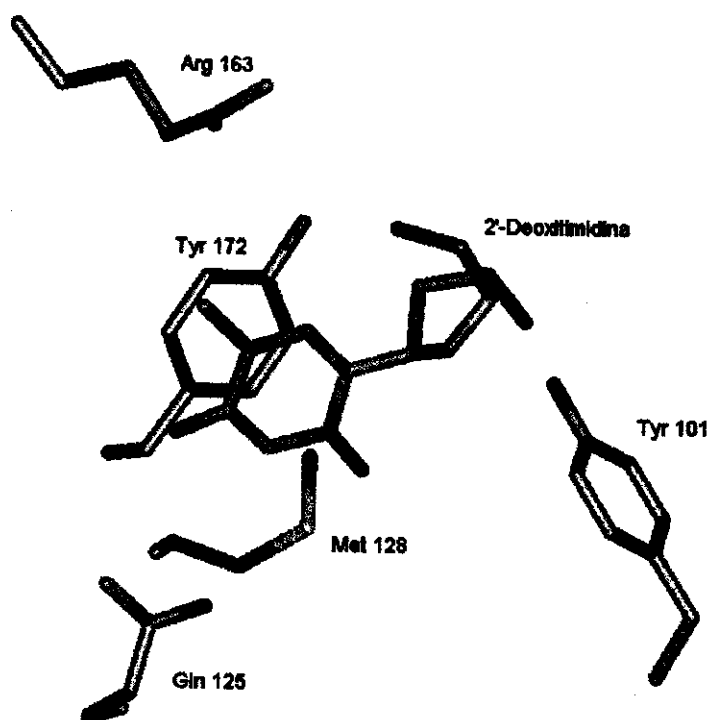


Figura 40 - Representação gráfica do modo de ligação de 2'-deoxitimidina. Ver detalhes na legenda da Figura 39.

A Figura 41 mostra o estereograma correspondente ao modo de ligação de 2'-deoxitimidina.

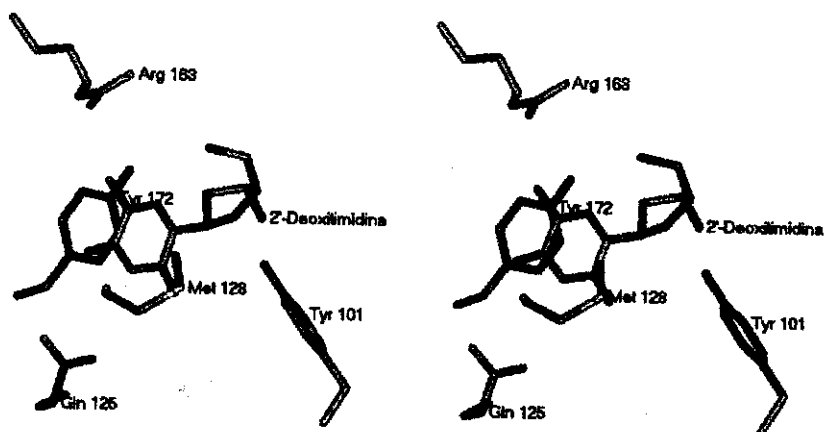
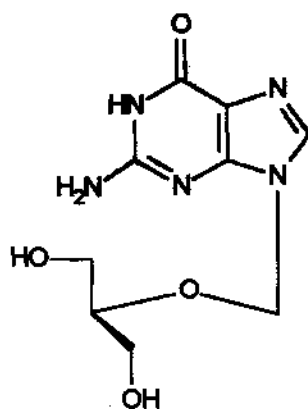


Figura 41 - Estereograma do modo de ligação de 2'-deoxitimidina no sítio ativo de HSV1 TK.

Como foi dito anteriormente (ver Seção 5.1, “Introdução”, pag. 154), dispõe-se da estrutura cristalográfica de HSV1 TK em complexo com ganciclovir. Ganciclovir é inibidor competitivo de 2'-deoxitimidina e, portanto, ocupa o mesmo sítio de ligação deste.



Ganciclovir

No entanto, ocorre diferença importante na orientação espacial da cadeia lateral de Gln 125. Ganciclovir obriga essa cadeia lateral a posicionar-se de forma invertida em relação à orientação adotada quando em presença de 2'-deoxitimidina (Figura 42). Devido ao fato das N^2 -fenilguaninas guardarem maior similaridade com ganciclovir do que com 2'-deoxitimidina, a configuração de ganciclovir no sítio ativo de HSV1 TK será importante ponto de partida para a determinação da conformação bioativa das N^2 -fenilguaninas (ver Seção 5.5, “Modo de ligação das N^2 -fenilguaninas”, pag. 196).

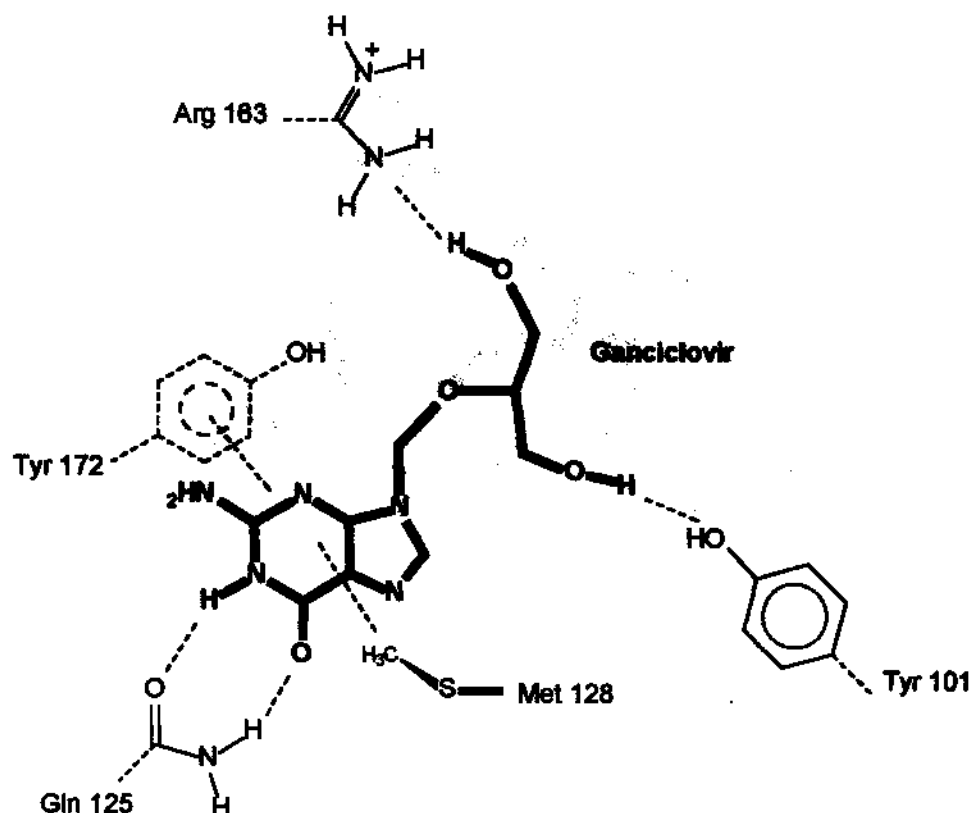


Figura 42 - Representação do sítio de ligação de ganciclovir em HSV1 TK. Observa-se a orientação espacial invertida do grupo amida de Gln 125, em relação à Figura 39. Ver explicação dos detalhes na Figura 39.

A Figura 43 mostra a representação gráfica do modo de ligação de 2'-deoxitimidina, referente ao esquema apresentado na Figura 42.

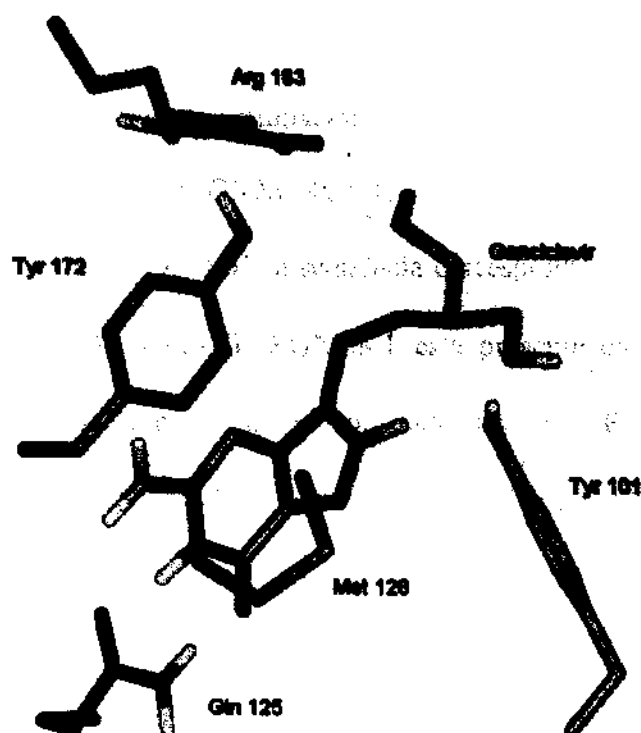


Figura 43 - Representação gráfica do modo de ligação de ganciclovir. Ver detalhes na legenda da Figura 39.

A Figura 42 mostra o estereograma correspondente ao modo de ligação de ganciclovir.

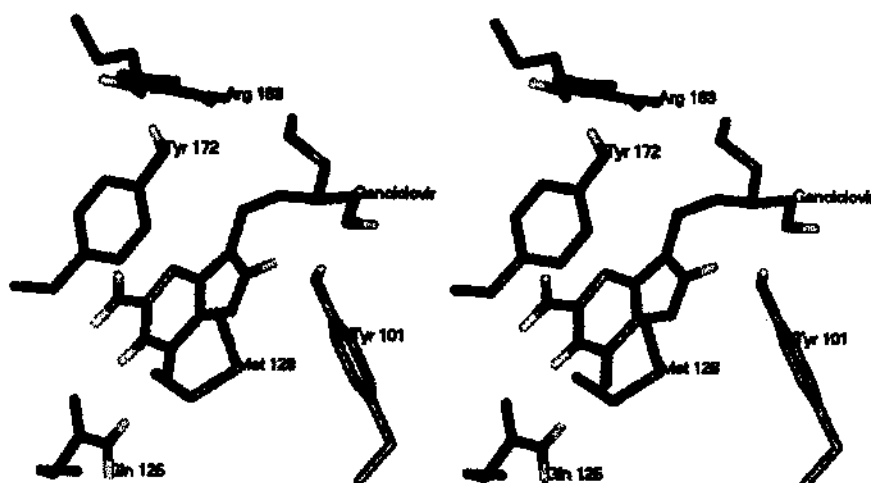
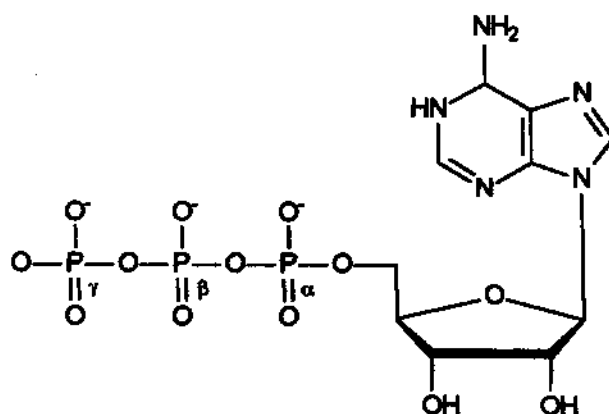


Figura 44 - Estereograma do modo de ligação de ganciclovir no sítio ativo de HSV1 TK.

5.2.5 Sítio do ATP

A principal sequência de aminoácidos onde se localiza o sítio de ligação do ATP, em HSV1 TK, é a sequência GXXGXGKT, em que X é um resíduo de aminoácido qualquer (Black & Hruby, 1990). Essa sequência corresponde à região final de $\beta 1$, $\Omega 1$ e à região inicial de $\alpha 1$. O padrão GXXGXGKT está presente em todas as sequências de timidina-cinase de vírus herpes simples depositadas no Swiss-Prot (ver nota de rodapé 8, pag. 156), embora com algumas poucas mutações (ver Diagrama 5, pag. 177). Esse padrão, que também está presente em adenil-cinase (Dreusicke et al, 1988; Müller e Schulz, 1992), forma a cavidade que acomoda o β -fosfato do ATP. Por ser rico em glicina, que não têm cadeia lateral, essa volta oferece bastante espaço para a entrada do ATP em seu sítio de ligação. A interação do β -fosfato com $\Omega 1$ posiciona o γ -fosfato para sua posterior transferência para 2'-deoxitimidina. Outros elementos da estrutura secundária de HSV1 TK que também compõem o sítio do ATP são $\Omega 8$, que força o β -fosfato contra $\Omega 1$, e $\alpha 8$, que parece manter alguma interação com o anel purínico. Na estrutura cristalográfica que se dispõe, há um grupo sulfato, na posição correspondente ao β -fosfato do ATP, em íntima interação com resíduos de $\Omega 1$.



Trifosfato de adenosina (ATP)

O ATP está presente no ambiente celular na forma de complexo com íon Mg^{2+} , que se coloca entre os grupos fosfato β e γ . A existência do complexo ATP- Mg^{2+} é importante para atenuar a grande concentração de carga resultante do encadeamento de grupos fosfato na molécula do ATP. Portanto, não é propriamente o ATP que age como doador de grupo fosfato à 2'-deoxitimidina, mas sim o complexo ATP- Mg^{2+} (Voet e Voet, 1995).

5.2.6 Sítio do Mg^{2+}

A presença de Mg^{2+} é essencial para a transferência de grupo fosfato do ATP para o nucleosídeo. Em muitas cinases, como por exemplo creatina-cinase e hexocinase, Mg^{2+} não chega a interagir propriamente com a enzima. Nesses casos, Mg^{2+} é coordenado por dois átomos de oxigênio, um de β -fosfato e outro de γ -fosfato do ATP, e pelos átomos de oxigênio de quatro moléculas de água (York, 1992; Mathews & van Holde, 1996). Os sítios ativos de creatina-cinase e hexocinase são suficientemente espaçosos para acomodar o complexo ATP- Mg^{2+} -4.H₂O. No caso de timidina-cinase e algumas outras cinases, como por exemplo adenil-cinase, admite-se que Mg^{2+} é coordenado por resíduo Asp ou Glu. De acordo com Brown e colaboradores (1995), o resíduo responsável pela coordenação de Mg^{2+} em HSV1 TK é Asp-162. Vários argumentos são utilizados para sustentar essa hipótese. (a) a sequência F¹⁶¹-D-R-H, que contém Asp 162, é altamente conservada em diversas TKs (Folkers et al, 1987 e Balasubramanian et al, 1990); (b) este é o único resíduo Asp que é conservado em todas as TKs de vírus herpes (Robertson e Whalley, 1988) (ver Diagrama 5, pag. 177). (c) a substituição de Asp 162 por Asn, em HSV1 TK, resulta em perda da atividade catalítica (Fetzer et al, 1993). (d) Asp 162 está localizado na extremidade carboxi de $\beta 3$, em posição espacialmente similar ao Asp 93 de adenil-cinase (Yan e Tsai, 1991; Abele & Schulz, 1995), que também coordena Mg^{2+} . Não restam dúvidas de que esses são argumentos bastante fortes em favor de Asp 162 como elemento

coordenador de Mg^{2+} em HSV1 TK. No entanto, após atenta observação da configuração tridimensional da região em torno de Asp 162, notou-se a presença do resíduo Glu 83, em posição vizinha ao resíduo Asp 162, e muito bem localizado em relação à posição de ligação de β -fosfato e γ -fosfato, que trazem consigo Mg^{2+} em complexo.

Decidiu-se averiguar a possibilidade de Glu 83 participar da coordenação de Mg^{2+} . Para isto, fez-se o alinhamento de todas as seqüências conhecidas de timidina-cinase de vírus herpes disponíveis e verificar se Glu 83 tem resíduo equivalente nessas seqüências. A conservação de um resíduo ao longo de seqüências diferentes, é forte indicação de alguma propriedade funcional importante. O Diagrama 4 mostra parte do alinhamento das seqüências de timidina-cinase de vírus herpes disponíveis no Swiss-Prot (ver nota de rodapé 8, pag. 156). Para nossa satisfação, verificou-se que Glu 83 (cujo símbolo é E) de HSV1 TK tem resíduo equivalente em todas as seqüências de timidina-cinase de vírus herpes conhecidas.

Diagrama 4 - Alinhamento das seqüências de timidina-cinase de vírus herpes disponíveis no SeqNet. Informações sobre as seqüências comparadas podem ser obtidas na legenda do Diagrama 5, pag. 177.

```

+.....+.....+.....+.....+
KITH_HSV1  80 YVPMPMTYWRVLGASETIANIYTTQHR
KITH_HSV23 81 YVPMPMTYWRVLGASETLTNIYNTQHR
KITH_HSVBQ 42 FFPMPMAYWRTMFVTDALSGILAASAR
KITH_HSVB6 42 FFPMPMAYWRTMFGTDALSGILAASAR
KITH_HSVBH 33 YLPMPMSYWSGAGEDDLVARVYTAQHR
KITH_HSVEM 267 TIPMPMREWTEVY-TNVLSQLV--KIN
KITH_HSVB5 41 FFPMPMAYWRTMFPTDALSGILAASAR
KITH_HSVB3 57 YFPMPMAYWRTLFADVISGIYDTQNR
KITH_HSVB4 57 YFPMPMAYWRTLFETDVISGIYDTQNR
KITH_HSVF  53 YFPMPMLYWRSLFETDVVGGIYAVQDR
KITH_HSVSA 240 AFHMPFIAYWTDVF-SNSLEEVY--KLT
KITH_HSVMD 51 KVFMPMKYWRYYF-TDLVTTVN--DTC
KITH_HSVMR 43 MLOMPMAYWRSTFASDAITEIYDTQHR
KITH_HSVTF 42 KVMMPMKYWRCS-TDLVVAAN--ETP
KITH_HSVTU 42 KVMMPMKYWRCS-TDLVVAAN--ETP

```

```

Consensus:      EPm yWr  %d!!  !
Identity:       **  *

```

Legenda:

- - Ácido	@ - Amida	\$ - Básico	% - Hidroxila
! - Hidrofóbico	# - Aromático	^ - Prolina	& - Enxofre

De acordo com o Diagrama 4, Glu 83 de HSV1 TK faz parte do padrão EPMXYWR, em que X é um resíduo de aminoácido qualquer, altamente conservado ao longo das seqüências presentes no diagrama.

Também verificou-se a possibilidade do Mg^{2+} , em complexo com ATP, interagir com os resíduos Glu 83 e Asp 162 de HSV1 TK. Ou seja, verificou-se se o posicionamento de β -fosfato na região de $\Omega 1$ (ver Seção 5.2.5, "Sítio do ATP", pag. 171) permite que Mg^{2+} possa interagir com Glu 83, de acordo com o esquema apresentado na Figura 45.

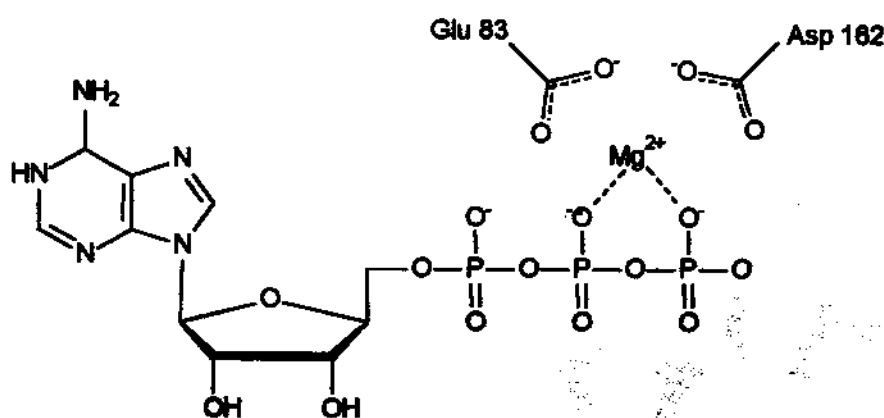


Figura 45 - Sítio de ligação do Mg^{2+} em HSV TK. De acordo com Brown e colaboradores (Brown et al, 1995), o resíduo responsável pela coordenação de Mg^{2+} em HSV1 TK é Asp 162. No entanto, há indicações de que Glu 83 também participa na coordenação desse íon.

Preparou-se o modelo do complexo $ATP-Mg^{2+}$ no módulo *Molecular Editor* do programa Quanta-CHARMM (1996). A posição do complexo $ATP-Mg^{2+}$, no sítio de HSV1 TK, foi ajustada utilizando-se o grupo sulfato presente na estrutura cristalina como referência. O grupo sulfato está localizado na posição correspondente a β -fosfato. A conformação utilizada para o complexo $ATP-Mg^{2+}$ foi baseada na estrutura do ATP no sítio de ligação de HSV1 TK determinada por cristalografia de raios-X, apresentada por Wild e colaboradores (1995). A conformação e a posição do complexo $ATP-Mg^{2+}$ foram ajustadas manualmente com o auxílio de óculos 3-D, no terminal gráfico. As conformações das cadeias laterais de Glu 83 e Asp 162 também foram manualmente ajustadas para serem aproximadas de Mg^{2+} .

A Figura 46 mostra o resultado do posicionamento do complexo ATP-Mg²⁺, em relação a Glu 83, Asp 162 e 2'-deoxitimidina. Pode-se dizer que a posição de Mg²⁺ é plenamente compatível com sua coordenação pelos resíduos Glu 83 e Asp 162, embora não seja possível, no momento, especular sobre qual resíduo é realmente o responsável pela coordenação, ou se ocorre coordenação simultânea por ambos resíduos .

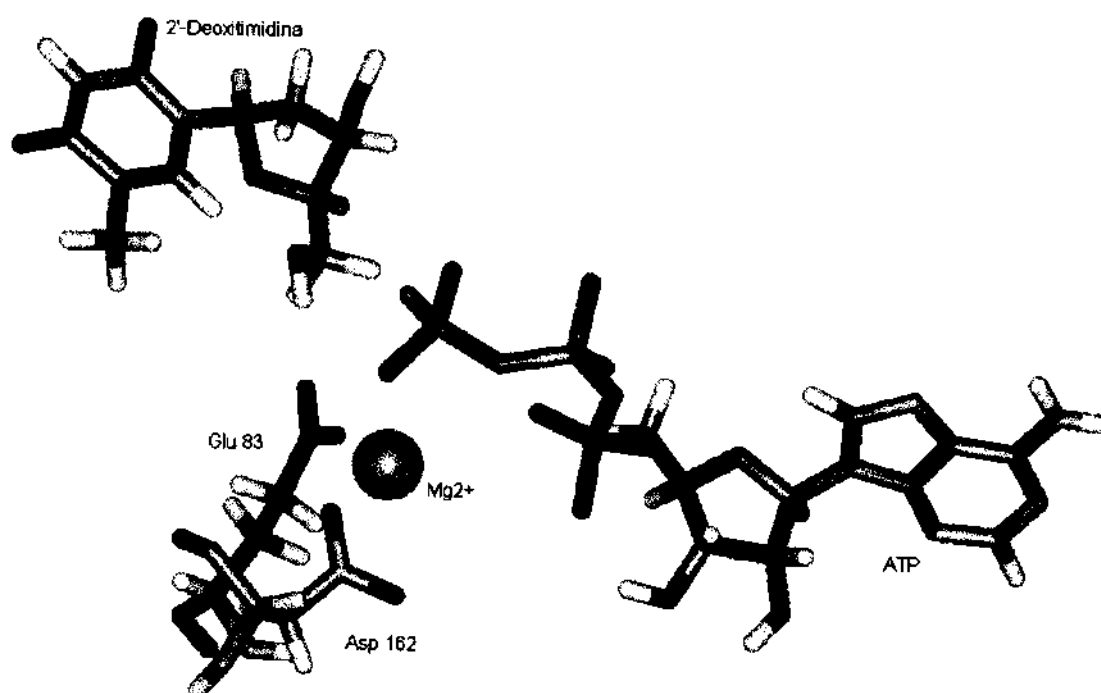


Figura 46 - Complexo ATP-Mg²⁺ na posição de transferência do grupo fosfato para a 2'-deoxitimidina. A posição espacial de Glu 83 é compatível para que possa participar da coordenação de Mg²⁺.

A Figura 47 mostra o estereograma do complexo ATP-Mg²⁺.

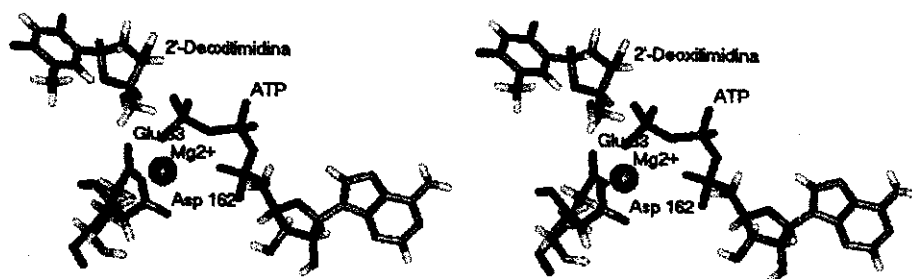


Figura 47 - Estereograma complexo ATP- Mg^{2+} na posição de transferência do grupo fosfato para a 2'-deoxitimidina.

5.2.7 Comparação das seqüências de outras timidinas cinases de vírus herpes

É interessante a comparação das seqüências de timidina-cinase conhecidas para conhecer os principais padrões de aminoácidos (*motifs*) que aparecem ao longo dessas seqüências. O SeqNet contém, até o presente momento, quatorze seqüências de timidina-cinase de vírus herpes, não incluindo as cinco seqüências do vírus herpes simples do tipo 1 que aparecem no Diagrama 1 (pag. 156). Em geral, as seqüências de TK dos diferentes tipos de vírus herpes não apresentam grande variação no comprimento da seqüência. Como pode ser visto na legenda do Diagrama 5, KITH_HSV23 apresenta 375 aminoácidos (AA), HSVBQ apresenta 357, HSVEB 352, HSVF 343, HSVMD 352, HSVMR 361 e HSVTU 350. Duas das quatorze seqüências apresentam número de aminoácidos muito acima do valor médio das seqüências mencionadas acima, que é de 356 aminoácidos. KITH_HSVSA apresenta 527 aminoácidos e HSVBM apresenta 686. Outras duas seqüências apresentam número de aminoácidos um pouco abaixo daquele valor médio. A seqüência KITH_HSVTF apresenta 310 aminoácidos e HSVBH 306. As duas seqüências

maiores dificultaram bastante o alinhamento das quinze seqüências. No entanto, o resultado do trabalho de alinhamento foi bastante satisfatório, pois os principais padrões de aminoácidos funcionalmente importantes, discutidos nas seções anteriores, mostraram-se bastante conservados ao longo das diversas seqüências de timidina-cinase dos diversos tipos de vírus herpes. O Diagrama 5 mostra o alinhamento das seqüências de aminoácidos da timidina-cinase disponíveis no Swiss-Prot (ver nota de rodapé 8, pag. 156).

Diagrama 5 - Alinhamento das seqüências de aminoácido de timidina-cinase de vírus herpes disponíveis no Swiss-Prot. São apresentados apenas os alinhamentos dos segmentos que contém os principais padrões de aminoácidos (*motifs*), destacados em negrito. O número que aparece antes de cada segmento de seqüência corresponde à posição do primeiro aminoácido do segmento na seqüência original de cada enzima. Informações sobre cada seqüência são mostradas na legenda.

	+.... ..		+.....+.....+.....
KITH_HSV1	52 VYIDGPHG MGKT -TTT-QLL ...	80	YV PEPM TYWR VLGASETIANIYTTQHR
KITH_HSV23	52 VYIDGPHG VGKT TTTSA-QLM ...	81	YV PEPM TYWQ VLGASETLTNIYNTQHR
KITH_HSVBQ	13 IYLDGAHGL GKT -TTG-RAL ...	42	FF PEPM AYWR TMEVTDALSGILAASAR
KITH_HSVB6	13 IYLDGAHGL GKT -TTG-RAL ...	42	FF PEPM AYWR TMEVTDALSGILAASAR
KITH_HSVBH	7 VYVDGPHGL GKT -TAA-SRL ...	33	YL PEPM SYWS GAGEDDLVARVYTAQHR
KITH_HSVBM	239 VFFEGCMAA GKT TTLLN-FAR ...	267	TI PEPM RYWR TEVY--TNVLSQIV--KIN
KITH_HSVB5	11 IYLDGAHGL GKT -TTG-RAL ...	41	FF PEPM AYWR TMEPTDALSGILAASAR
KITH_HSVBEB	28 IYLDGVYGI GKS -TTG-RVM ...	57	YF PEPM AYWR TLFEADVISGIYDTQNR
KITH_HSVE4	28 IYLDGVYGI GKS -TTG-RVM ...	57	YF PEPM AYWR TLFETDVISGIYDTQNR
KITH_HSVF	23 IYIDGAYGI GKS -LTA-KYL ...	53	YF PEPM LYWR SLFETDVVGGIYAVQDR
KITH_HSVSA	212 IFLEGSI GVGKT TLK-SMN ...	240	AF HEPI AYWR TDVDF-SNSLEEVY--KLT
KITH_HSVMD	22 VYLDGSMGI GKT SMLNEIPT ...	51	KV PEPM KYWR YF-TDLVTTVN--DTC
KITH_HSVMR	13 VYLDGPHG VGKT TTAE-ALV ...	43	ML QPEM AYWR STFASDAITEIYDTQHR
KITH_HSVTF	13 VYLDGPF GI GKT SILNAMPD ...	42	KVY PEM KYWR CQS-TDLVVAAN--ETP
KITH_HSVTU	13 VYLDGPF GI GKT SILNAMPD ...	42	KVY PEM KYWR CQS-TDLVVAAN--ETP
Consensus:	!yldG glGkt		EPm yWr #d!! !
Identity:	* **		** *
	+.....+.....		+.....+.....
KITH_HSV1	116 ---DAAVVMTSA QITMGM PYAV ...	156	----ALTLI DRHP LAALLCYPAARYLM
KITH_HSV23	117 ---EAAVMTSA QITMST PYAA ...	157	----ALTLV DRHP IASLLCYPAARYLM
KITH_HSVBQ	90 ---DAAGLVAYY QARFA PYLI ...	128	----VTLV DRHP VAACLCYPARYCL
KITH_HSVB6	92 ---GPGCVLP- QVR -- GPVLN ...	127	----VTLV DRHP VAACLCYPARYCL
KITH_HSVBH	69 ---EAAGVVLGA QLTMST PYVA ...	109	----DLILI DRHP TASLLCYPLARYLT
KITH_HSVBM	300 ---TTAELVSC- QLKFATP --L ...	335	VGPLDKWVIV- DRHQ LSALVVFPLVMRR
KITH_HSVB5	79 ---DAAGLVAYY QARFA PYLV ...	128	----VVTLV DRHP LAACLCYPARYCL
KITH_HSVBEB	93 ---DAASITAHY QSRFTT PYLI ...	132	----DLTV DRHP VASAVCFPAARYLI
KITH_HSVE4	93 ---DAALITAHY QSRFTT PYLI ...	132	----DLTLV DRHP VASTVCFPAARYLL
KITH_HSVF	89 ---DAAYITAHY QARFA PYLL ...	128	----DVTLLI DRHP LASLVCFPLARYFV
KITH_HSVSA	273 ---NSAKLLAC- QLKFASP --L ...	305	LLSSDMWVME- DRHP LSATVVFPPYMHFQN
KITH_HSVMD	84 LFQSSMIVTAL- QSKFADP --Y ...	119	TRGNPSLILIL DRHP ISATVCFPIARHLT
KITH_HSVMR	79 ---EAGAEMTSL QLHMG TPYAL ...	119	LPQRDFVLVV DRH AVASMVCPYLAREFM
KITH_HSVTF	75 RFQSDMIMASI- QARFADP --Y ...	110	ICDTPAII MLDRHP VAAILCFPITRYLL
KITH_HSVTU	75 RFQSDMIMASI- QARFADP --Y ...	110	ICDTPAII MLDRHP VAAILCFPITRYLL
Consensus:	! ! QIS P !		! !! DRHp!a !c#PIary
Identity:	* *		*** *

```

Consensus:      g  %  !!  !!  p  pG n!!!  l  e  !  Rl  R  r  ge  !d
Identity:              *

```

~ - Ácido @ - Amida \$ - Básico % - Hidroxila
 ! - Hidrofóbico # - Aromático ^ - Prolina & - Enxofre

KITH_HSV23: Vírus herpes simples Tipo 2 (Linhagem 333) (375 AA; PM 40370) (Kit et al, 1983; Swain e Galloway, 1983)
KITH_HSVBQ: Vírus herpes bovino Tipo 1.2 (Linhagem Q3932) (357 AA; PM 36759) (Smith et al, 1990)
KITH_HSVB6: Vírus herpes bovino Tipo 1 (Linhagem 6660) (357 AA; PM 36759) (Mittal e Field, 1989)
KITH_HSVBH: Vírus herpes bovino Tipo 2 (Linhagem BHM-1) (306 AA; PM 32929) (Sheppard e May, 1989)
KITH_HSVBM: Vírus herpes bovino Tipo 3 (Linhagem WC11) (686 AA; PM 77073) (Hsu et al, 1990)
KITH_HSVB5: Vírus herpes bovino Tipo 5 (Linhagem N569) (350 AA; PM 36262) (Smith et al, 1991)
KITH_HSVEB: Vírus herpes equino Tipo 1 (Linhagem AB4P) (EHV-1), e vírus herpes equino Tipo 1 (HVS25A isolado) (EHV-1) (352 AA; PM 38750) (Telford et al, 1992; Robertson e Whalley, 1988)
KITH_HSVE4: Vírus herpes equino Tipo 4 (Linhagem 1942) (EHV-4) (Vírus herpes equino Tipo 1 Subtipo 2) (352 AA; PM 38784) (Nicolson et al, 1990)
KITH_HSVF: Vírus herpes felino (343 AA; PM 38922) (Nunberg et al, 1989)
KITH_HSVSA: Vírus herpes saimiri (*Saimiri herpesvirus*) (Linhagem 11) (527 AA; PM 59806) (Honest et al, 1989; Albrecht et al, 1992)
KITH_HSVMD: Vírus herpes de Marek (*Marek's disease herpesvirus*) (Linhagem RB-1B) (MDHV) (352 AA; PM 40365) (Scott et al, 1989)
KITH_HSVMR: Vírus herpes símio (*Marmoset herpesvirus*) (361 AA; PM 39458) (Otsuka e Kit, 1984)
KITH_HSVTF: Vírus herpes de peru (Linhagem FC126) (310 AA; PM 35512) (Martin et al, 1989)
KITH_HSVTU: Vírus herpes de peru (350 AA; PM 39968) (Scott et al, 1989)

Também é interessante comparar-se a sequência de aminoácidos da timidina-cinase de vírus herpes simples com a sequência da timidina-cinase humana. A especificidade dos inibidores enzimáticos depende da similaridade existente entre a enzima-alvo e a correspondente enzima do hospedeiro. Quanto menor a similaridade entre essas enzimas, maiores as chances de conseguir-se desenvolver inibidor específico, ou seja, inibidor que atue eficazmente contra a enzima-alvo sem interferir apreciavelmente na atividade da enzima do hospedeiro.

A timidina-cinase humana possui sequência de 234 aminoácidos e peso molecular 25501 uma. O código identificador dessa enzima, no Swiss-Prot (ver nota de rodapé 8, pag. 156), é KITH_HUMAN. O alinhamento das sequências de KITH_HSV1 e KITH_HUMAN é mostrado no Diagrama 6.

Diagrama 6 - Alinhamento das seqüências de aminoácido da timidina-cinase de vírus herpes simples do tipo 1 e da timidina-cinase humana. Os padrões de aminoácidos apresentados em negrito no Diagrama 5 também são apresentados em negrito na seqüência KITH_HSV1 para melhor comparação das duas seqüências.

	1	10	20	30	40	50	60
KITH_HSV1	MASYPCHQHASFDAQARSRGHSNRRRTALRPRRQOEATEVRLEQKMPILLRVITDGPHEGAKTTT						
KITH_HUMAN	MS---CINLPTVL-PGSPSKTR-----GQIQVILQPMFSGK--S						
Consensus:	M	C @ @!	! S\$ \$			!	I GP GK %
Identity:	*	*	*				* ** **
	66	75	85	95	105	115	125
KITH_HSV1	TQLLVALLGSRRDIVYV ELMTI TVRLVGASETIANIYTTOHRLDQGEISAGDAAVVM TSAQTITMGM						
KITH_HUMAN	TELMRRVRRFQIAQYKCLVIKY----AKDTRYSSSFCTHDRNTMEALPA---CLLRDVVAQEALGV						
Consensus:	T L !	! Y	Y !	# T R	! A	!! AQ	G
Identity:	**	*	*	* *	*	**	*
	131	140	150	160	170	180	190
KITH_HSV1	P-YAVTDVAVLAPHIGGEAGSSSHAPPPALTLI EDEEPIA LALCYPAARYLMGSMTPQAVLAFLVALI						
KITH_HUMAN	AVIGIDEGQFFPDI--MEFCEAMANAGKTVI-----VAAL----DGTFQRKPFG--AILNLVPLA						
Consensus:	!! ~!	P I	! TII	!AAL	! #	AIL V LI	
Identity:		**	**	***		** *	
	195	205	215	225	235	245	255
KITH_HSV1	PPTLPGTNIIVLGALPEDRHIDRLAKR RPGER LDLAMLAAIRRVYGLLAN TVRYLQCGGSWREDW						
KITH_HUMAN	ESVVKLTAVCMCECFREAAAYTKRLGTEK-----EVEVIGGADK-Y---HSVCR-L-C--YFKKAS						
Consensus:	I IT !	E	RI!	~! IIII \$ Y	%V L C	# \$	
Identity:	*	*	**	*	*	**	
	260	270	280	290	300	310	320
KITH_HSV1	GQLSGTAVPPQGAEPQSNA GPRPHIGDTLFTLLFRAPELLAPNGDLNVFAWALDVLA KRLRSMHV						
KITH_HUMAN	GQ-----PAGPDNKENCVPVGP KPGEAV---AARKLFAPQQ-----						
Consensus:	GQ	P G ~ N	\$ G~ I	A L AP@			
Identity:	**	** *	*	* * **			
	325	335	345	355	365	375	385
KITH_HSV1	FILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMGEAN						
KITH_HUMAN	-ILQ-----C-----SPAN-----						
Consensus:	IL	C		@P!			
Identity:	**	*		*			

Legenda :

~ - Ácido @ - Amida \$ - Básico % - Hidroxila
 ! - Hidrofóbico # - Aromático ^ - Prolina & - Enxofre

KITH_HUMAN: Homo sapiens (humano) (234 AA; PM 25501) (Bradshaw e Deininger, 1984; Flemington et al, 1987)

O alinhamento de KITH_HSV1 e KITH_HUMAN provavelmente deve conter algumas incorreções. Isso é esperado devido à baixa homologia entre as duas seqüências. De acordo com o alinhamento acima, o percentual de homologia entre as duas seqüências é 21% (49/234), considerando-se coincidência total, e 42 % (98/234), considerando-se homologia funcional, de acordo com os símbolos apresentados na legenda do Diagrama 1 (pag. 156). A baixa homologia entre as timidina-cinases viral e humana representa fator positivo no processo de desenvolvimento de inibidores contra HSV1 TK. A baixa homologia entre as seqüências certamente significa baixa homologia entre as estruturas tridimensionais. Isto significa que um composto capaz de inibir HSV1 TK muito pouco provavelmente será capaz de inibir apreciavelmente a TK humana.

5.3 ANÁLISE ESTRUTURAL DE HSV2 TK

5.3.1 Estrutura primária de HSV2 TK

Enquanto dispõe-se de cinco previsões da seqüência de aminoácidos de HSV1 TK e também de sua estrutura tridimensional, determinada experimentalmente, dispõe-se apenas da previsão da seqüência de HSV2 TK (Kit et al, 1983; Swain e Galloway, 1983). O Diagrama 7 mostra o alinhamento da seqüência completa de HSV1 TK com a previsão da seqüência de HSV2 TK, cuja identificação (ID) no Swiss-Prot (ver nota de rodapé 8, pag. 156) é KITH_HSV23. É interessante notar que HSV2 TK possui um aminoácido a menos que HSV1 TK em sua seqüência. No entanto, será adotada a mesma numeração para os aminoácidos homólogos de HSV1 TK e HSV2 TK, apesar dos espaços (*gaps*) existentes no alinhamento apresentado no Diagrama 7.

Diagrama 7 - Comparação das seqüências completas da timidina-cinase do vírus herpes simples dos tipos 1 e 2. As seqüências ausentes na estrutura cristalina de HSV1 TK e as seqüências equivalentes e essas em HSV2 TK são mostradas em negrito.

	1	10	20	30	40	50	60
KITH_HSV1	+.....+.....+.....+.....+.....+.....+.....+.....+						
KITH_HSV23	MAS TFCHEASAFDQAARSGESENRTALRPPQOAEIVLEKMPPTLLRVYIDGPHGMGKT TTT						
Consensus:	MAS QHA AF QAAR G % R A RP \$ Q A\$E!R!~ PTLLRVYIDGPHG GKTTT						
Identity:	*** ** * **** * * * * * ***** *						
	66	75	85	95	105	115	125
KITH_HSV1	+.....+.....+.....+.....+.....+.....+.....+.....+						
KITH_HSV23	TQLLV-ALGSRD IVVPEPMTYWVRVLGASETIANIYTQHRLDQEISAGDAAVVM TSAQITMG						
Consensus:	% L ALG-RD IVVPEPMTYW VLGASET! NIY TQHRLD GEISAG-AAVVM TSAQITM						
Identity:	* *** * ***** ***** * ***** ***** *****						
	130	140	150	160	170	180	190
KITH_HSV1	+.....+.....+.....+.....+.....+.....+.....+.....+.....+						
KITH_HSV23	MPYAVTD AVLAPHIGGEAGS SH APPPALTIFDRHPIALLCYPAARYLMGSMTFQAVLAFVALI						
Consensus:	FYA!TDAVLAPHIGGEA! APPPALTL!FDRHPIA LLCYPAARYLMGSMTFQAVLAFVAL						
Identity:	*** ***** ***** ***** ***** ***** *****						
	195	205	215	225	235	245	255
KITH_HSV1	:.....+.....+.....+.....+.....+.....+.....+.....+.....+						
KITH_HSV23	PPTLPGTNIVLGALPEDRHIDLAKRQRPERGLDLAMLAAIRRVYGLLANTVRYLQCGGSWREDW						
Consensus:	PPT!PGTN!VLG!LPE HIDRLA\$RQRPERGLDLAML AIRRVY LLANTVRYLQ GG WREDW						
Identity:	*** **** * **** * ***** ***** ***** ***** *						
	260	270	280	290	300	310	320
KITH_HSV1	+.....+.....+.....+.....+.....+.....+.....+.....+.....+						
KITH_HSV23	GQL SGTAVTFQGAEPQ SAGPRPHIGDTLFTLFRAPELLAPNGDLYNF FAWALDVLA KRLRS MHV						
Consensus:	G L&G A! -P AG P\$ I DTLE LFR!PELLAPNGDLY !FAW!LDVLA RL MH!						
Identity:	* *						
	325	335	345	355	365	375	385
KITH_HSV1	:.....+.....+.....+.....+.....+.....+.....+.....+.....+						
KITH_HSV23	FILDYDQSPAGCRDALLQLTSGMVQTHVTTPGSIPTICDLARTFAREMG EAM						
Consensus:	FILDYDQSP!GCRDALL LT GM! T\$VTT GSI I DLARTFARE G !						
Identity:	* ***** ***** * * * * * * * * * * *						

Legenda:

~ - Ácido @ - Amida \$ - Básico % - Hidroxila
 ! - Hidrofóbico # - Aromático ^ - Prolina & - Enxofre

KITH_HSV23: Vírus herpes simples (Tipo 2; Linhagem 333) (375 AA; PM 40370) (Kit et al, 1983; Swain e Galloway, 1983)

HSV1 TK e HSV2 TK apresentam 74% (276/375) de homologia, considerando-se coincidência total da sequência de aminoácidos. Incluindo-se a coincidência funcional, de acordo com a legenda do Diagrama 1 (pag. 156), a homologia chega a 82% (306/375).

Duas porções das seqüências apresentadas no Diagrama 7 merecem consideração especial. Ao executar o alinhamento das duas seqüências de aminoácidos, o programa Cameleon (1994) inseriu um espaço (*gap*) na posição 67 de HSV1 TK de forma a obter homologia máxima no alinhamento (Diagrama 8).

Diagrama 8 - Alinhamento feito inicialmente pelo programa Cameleon. O programa inseriu um espaço na seqüência de HSV1 TK na posição 67.

```

        66      75      85      95      105      115      125
KITH_HSV1  T-QLLVALGSRDIDIVYVPEPMTYWRVLGASETIANIYTTQHRLDQGEISAGDAAVVMTSAQITMG
KITH_HSV23 SAQLMEALGPRDNIVYVPEPMTYWRVLGASETLTNIYNTQHRLDGEISAGEAAVVM TSAQITMS

```

Porém, a análise da seqüência de nucleotídios do gene de HSV2 TK (Swain e Galloway, 1983) mostrou que a posição correta para a inserção daquele espaço na seqüência de HSV1 TK é 71, como indicado no Diagrama 7. A Figura 2 do artigo de Swain e Galloway (1983), que exibe a seqüência de nucleotídios do gene responsável pela codificação de HSV1 TK, mostra a posição do códon GAG (G = guanosina e A = adenosina), responsável pela codificação de Glu 70 da seqüência de HSV2 TK (Kit et al, 1983), e mostra que esse códon não está presente no gene de HSV1 TK.

A segunda porção que merece destaque corresponde à posição 269 do Diagrama 7. O programa Cameleon inseriu mais um espaço, desta vez na posição 271 da seqüência de HSV2 TK (Diagrama 9).

Diagrama 9 - Alinhamento feito inicialmente pelo programa Cameleon. O programa inseriu um espaço na seqüência de HSV2 TK na posição 271.

```

        260      270      280      290      300      310      320
KITH_HSV1  +.....+.....+.....+.....+.....+.....+.....
KITH_HSV23 GQLSGTAVVPFGAEFQSHAGPRPHIGDTLFTLFRAPPELLAPNGDLYNVEFAWALDVLAKRLRSMHV
          GRLTGVAATAIR-PDPEDGAGSLPRIEDTLFALFRVPELLAPNGDLYHIFAWVLDVLADRLLEPMHL

```

Em princípio, isso significaria que, quando da construção de HSV2 TK a partir da estrutura de HSV1 TK (ver Seção 5.4.4, "Modelagem das estruturas terciária e quaternária de HSV2 TK", pag. 188), dever-se-ia eliminar a Gly 271 da seqüência de HSV1 TK. Porém, a comparação das seqüências de nucleotídios apresentadas em Swain e Galloway (1983) e Kit et al (1983) revela diferença importante. Na posição 1118 do gene de HSV1

TK (Figura 2 de Swain e Galloway, 1983) aparece o códon CCC (C = Citosina), responsável pela codificação da prolina. Esse códon ocorre na posição correspondente à Pro 269 da sequência de HSV1 TK. O códon 1118-CCC não está presente na sequência do gene de HSV1 TK apresentado por Kit et al (1983, Figura 2) e nem na sequência atualizada desse gene que pode ser encontrado no GenBank⁹ (ID = HEHSV1TK). Dessa forma, modificou-se a posição do espaço inserido na sequência de HSV2 TK pelo programa Cameleon para a posição correspondente à Pro 269 de HSV1 TK, como pode ser conferido no Diagrama 7 (pag. 181).

As intervenções do operador nos resultados apresentados por programas do tipo *pacote fechado*, como é costume falar, são muito importantes. Não se deve tomar como verdade os resultados apresentados por programas desse tipo, principalmente porque o operador freqüentemente não tem conhecimento (embora devesse ter) dos métodos de cálculo e operações por eles utilizados. No entanto, programas como Cameleon e muitos outros são ferramentas muito úteis quando usados com a devida cautela.

5.4 MODELAGEM DE HSV1 TK E HSV2 TK

5.4.1 Preparação da molécula protéica para cálculos moleculares

De forma geral, não se consegue fazer cálculos de mecânica e dinâmica molecular utilizando-se diretamente as estruturas protéicas obtidas de bancos de dados cristalográficos. Antes é preciso preparar a molécula (*cleanup*) para tal. A preparação da molécula é necessária devido a diversos fatores: a estrutura protéica não possui átomos de hidrogênio, algumas regiões podem estar mal definidas, voltas podem estar ausentes, pode haver presença de aminoácidos que não estejam pré-definidos, átomos não definidos

⁹ O GenBank pode ser acessado através do SeqNet no endereço da Internet <http://salpha2.dl.ac.uk/srs/srsc>.

(especialmente de ligantes), ligações não definidas, etc. Portanto, há necessidade de tratamento completo antes de qualquer procedimento que envolva cálculo molecular. Isso requer o uso de programas computacionais específicos. Os principais programas capazes de executar essa tarefa que estão disponíveis no *Physical and Theoretical Chemistry Laboratory* são Sybyl (1991) e Quanta-CHARMM (1996). No presente trabalho, preferiu-se utilizar o programa Quanta na preparação das moléculas de HSV1 TK e HSV2 TK devido à maior facilidade de uso e, principalmente, por dispor-se de versão bastante atualizada.

Decidiu-se fazer a preparação das moléculas de HSV1 TK, presentes nos arquivos HSV1_TK_dT_2 e tk_ganc_nov8g, separadamente. As enzimas, cujas estruturas estão definidas nesses dois arquivos, foram cristalizadas com ligantes diferentes ocupando seus respectivos sítios ativos; 2'-deoxitimidina em HSV1_TK_dT_2 e ganciclovir em tk_ganc_nov8g. A presença de moléculas muito diferentes como 2'-deoxitimidina e ganciclovir no sítio ativo de HSV1 TK é capaz de provocar alterações conformacionais importantes nas cadeias laterais de alguns aminoácidos. Neste caso, a alteração mais importante refere-se à cadeia lateral de Gln 125, que faz ponte de hidrogênio dupla com os anéis pirimidínico de 2'-deoxitimidina (ver Figura 39, pag. 166) e com o anel purínico de ganciclovir (ver Figura 42, pag. 169). As cadeias laterais de Gln 125, nessas duas estruturas, apresentam-se espacialmente invertidas.

Para efeito de diferenciação das moléculas de HSV TK obtidas a partir dos arquivos HSV1_TK_dT_2 e tk_ganc_nov8g, adotou-se a seguinte terminologia: as moléculas de HSV TK obtidas a partir da manipulação do arquivo HSV1_TK_dT_2, serão chamadas HSV1_TK (obtida diretamente por *cleanup*) e HSV2_TK (obtida através de mutações na sequência de aminoácidos de HSV1 TK); e as moléculas de HSV TK obtidas a partir de tk_ganc_nov8g serão chamadas hsv1_tk (obtida diretamente por *cleanup*) e

hsv2_tk (obtida através de mutações na sequência de aminoácidos de hsv1 tk). Acredita-se que essa diferenciação irá poupar alguma dor de cabeça no futuro.

5.4.2 Reconstrução das seqüências ausentes de HSV1_TK_dT_2

A reconstrução das seqüências ausentes na estrutura de HSV1_TK pode ser dividida em duas partes: reconstrução de voltas e reconstrução de seqüências N e C terminais. A reconstrução das voltas é tarefa relativamente simples. Em geral as seqüências ausentes pertencentes às voltas não são muito extensas e podem ser comparadas com as voltas similares encontradas em outras enzimas. Conhecendo-se a seqüência de aminoácidos que constitui a volta desejada, executa-se busca sistemática em banco de dados cristalográfico de forma a verificar possíveis similaridades com as diversas voltas depositadas no banco.

Para executar a reconstrução das voltas ausentes, utilizou-se o programa Quanta (1996). Possuindo banco de dados cristalográfico interno, Quanta compara a seqüência de aminoácidos da volta desejada com as diversas voltas existentes em seu banco de dados. Terminada a comparação, Quanta apresenta as voltas selecionadas, mencionando índices que refletem o grau de similaridade existente entre a volta procurada e a encontrada, e também o grau de ajuste das extremidades das voltas encontradas às extremidades onde estas deverão ser encaixadas na molécula protéica. Na tarefa de escolha da conformação da volta para determinada seqüência, não é boa idéia tomar decisão em favor desta ou daquela conformação baseado apenas em índices. É preciso observar cuidadosamente na tela do terminal gráfico, se possível utilizando óculos 3D, para verificar possíveis conflitos de cadeias laterais da volta escolhida com o restante da estrutura protéica. Esse tipo de conflito não é considerado nos índices apresentados. No PTCL dispõe-se de estação gráfica Power Challenge Silicon Graphics acoplada a sistema que permite a observação de

imagens tridimensionais, composto de óculos 3D e sensor infravermelho. A utilização dos óculos 3D permitiu o exame bastante realista da estrutura protéica e facilitou a tomada de decisão na escolha das voltas de HSV1_TK, que são mostradas na Tabela 23.

Tabela 23 - Voltas procuradas e voltas escolhidas para os monômeros A e B de HSV1_TK (presente no arquivo HSV1_TK_dT_2). Na coluna *Volta escolhida*, a conformação da sequência dos aminoácidos em negrito foi utilizada para modelar a sequência da volta procurada em HSV1_TK. A conformação dos aminoácidos das extremidades da volta escolhida foi usada como referência para fixar cada sequência nas extremidades da volta procurada na molécula de HSV1_TK. RMS (lsq) (*root mean square-least square fit*) corresponde ao desvio quadrático médio das distâncias entre os átomos de carbono- α ($C\alpha$) da sequência procurada e da sequência encontrada, ajustada por método de mínimos quadrados. FIT (dist) corresponde à soma das diferenças das distâncias entre os $C\alpha$ da sequência procurada e da sequência encontrada.

Monôm.	Volta procurada	Volta escolhida		RMS (lsq)	FIT (dist)
		Enzima: Início	Sequência		
A	Leu-72-Asp-76	3CYT: Gly 35	GLFGRKTGQ	2,318	1,085
B	Gly-73-Asp-76	2OVO: Pro 22	PLCGSDNK	1,564	0,546
A	Ser-150-His-151	3CYT: Met 66	MEYL	0,919	0,008
B	Ser-150-His-151	3CYT: Asp 63	DTLM	0,828	0,008
A	Leu-262-Gly-279	5CHA: Gln 110	QTVSAVCLPSAS DDFAAGTTCVTT	2,920	1,365
B	Ser-263-Asn-277	5CHA: Gly 197	GAWTLVGIVSWG SSTCSTSTP	3,575	1,399

A reconstrução das sequências N ou C terminais são mais complicadas do que a reconstrução das voltas. No caso destas, além da sequência de aminoácidos em geral ser curta, têm-se dois pontos de referência (as duas extremidades onde a volta será conectada) para basear a busca pela conformação adequada. No caso das sequências terminais, dispõe-se apenas de um ponto (o resíduo terminal). No caso específico de HSV1_TK, a sequência C terminal não apresenta problema sério por tratar-se de apenas dois aminoácidos (Ala - Asn). No entanto, a sequência N terminal apresenta séria dificuldade de ajuste por tratar-se de sequência de aminoácidos muito longa (45 aminoácidos) (ver Diagrama 3, pag.159). Devido às dificuldades de ajuste das sequências terminais à estrutura original de HSV1 TK

e também à sua pouca importância funcional, decidimos por não incluí-las neste trabalho. Em verdade, experimentos mostram que as primeiras três ou quatro dezenas de aminoácidos N-terminais de HSV1 TK não têm influência detectável sobre a atividade catalítica da enzima (Wild et al, 1995).

Após a inserção das voltas e inclusão dos átomos de hidrogênio, HSV1_TK passou a ter 10012 átomos, sendo 4994 átomos não-hidrogenóides e 5018 átomos de hidrogênio.

5.4.3 Reconstrução das seqüências ausentes de tk_ganc_nov8g

Repetimos o procedimento acima para hsv1_tk (arquivo tk_ganc_nov8g). Da mesma forma, optou-se por não completar as extremidades amino e carboxi.

Enquanto cada monômero de HSV1_TK apresentou três voltas ausentes, os monômeros de hsv1_tk apresentaram apenas duas voltas ausentes cada. As seqüências ausentes nessa estrutura são: A: Leu 72 – Asp 77, A: Leu 262 – Pro 280, B: Gly 73 – Asp 77 e B: Ser 263 – Asn 277. É importante lembrar que o arquivo tk_ganc_nov8g foi obtido cerca de três meses após a obtenção de HSV1_TK_dT_2. Muitas simulações com a estrutura de HSV1_TK já haviam sido feitas nessa ocasião. Por esse motivo, não se utilizaram informações sobre a conformação da volta presente no segundo arquivo para a construção da volta correspondente que estava ausente no primeiro. Porém, a superposição dos dois segmentos mostrou que a diferença conformacional entre as voltas selecionadas para o segmento A/B: Ser 150 – His 151 de HSV1_TK (ver Tabela 23, pag. 186) e os mesmos segmentos encontrados na estrutura cristalográfica de hsv1_tk é insignificante. A Tabela 24 mostra as voltas procuradas e as escolhidas para os segmentos ausentes em hsv1_tk.

Tabela 24 - Voltas procuradas e voltas escolhidas para os monômeros A e B de hsv1_tk (presente no arquivo tk_ganc_nov8g). Ver explicação na Tabela 23, pag. 186.

Monôm.	Volta procurada	Volta escolhida		RMS (lsq)	FIT (dist)
		Enzima: Início	Seqüência		
A	Leu-72-Asp-77	3CYT: Pro 72	PKKYIPGTKM	2,136	0,741
B	Gly-73-Asp-77	1LYZ: Ala 10	AAMKRHGLD	1,571	0,699
A	Leu-262-Pro-280	4DFR: Lys 76	KSVDEAIAACGDV PEIMVIGGGRVY	3,095	1,032
B	Ser-263-Asn-277	5CHA: Gly 13	GEEAVPGSWPW QVSLQDKTGF	2,584	0,563

A comparação da Tabela 23 com a Tabela 24 mostra que voltas diferentes foram selecionadas para os mesmos segmentos ausentes em HSV1_TK e hsv1_tk. Na prática, o efeito disso é praticamente nenhum, pois mesmo sendo diferentes, o resultado foi visualmente similar. Além disso, essas voltas não apresentam importância funcional, pois não fazem parte do sítio de ligação de 2'-deoxitimidina, ATP ou do íon magnésio. E, mais importante, essas estruturas serão submetidas a simulações de dinâmica molecular à temperatura de 310 K. Como, em geral, voltas maiores são mais flexíveis, espera-se alguma mudança na conformação dessas voltas. Portanto, admite-se que pequenas diferenças conformacionais nessas voltas não representam problema sério. Lembra-se, finalmente, que, na reconstrução das voltas, apenas a conformação da volta escolhida é utilizada e não a volta propriamente dita. O que o programa Quanta faz é tentar ajustar a seqüência de aminoácidos da volta original à conformação da volta escolhida.

5.4.4 Modelagem das estruturas terciária e quaternária de HSV2 TK

A modelagem da estrutura tridimensional de HSV2 TK é necessária devido ao fato da mesma não ter sido determinada experimentalmente até o momento. Para isso, partiu-se da estrutura cristalográfica original de HSV1_TK, sem as voltas que foram adicionadas na

Seção 5.4.2 (“Reconstrução das seqüências ausentes de HSV1_TK_dT_2”, pag. 185), e fez-se mutações sucessivas em sua seqüência de aminoácidos, ou seja, substituiu-se diversos aminoácidos da seqüência de HSV1_TK de forma a convertê-la na seqüência de HSV2_TK. Para isso, utilizou-se o programa Quanta (1996) que possui ferramenta de análise, manipulação e edição de estruturas poliméricas biológicas, denominada *Protein Design Module*. Dessa forma, conseguiu-se converter a seqüência de HSV1_TK na seqüência de HSV2_TK, utilizando a estrutura tridimensional de HSV1_TK como base. Devido à grande homologia existente entre as duas enzimas (74%), espera-se que exista grande similaridade nas correspondentes estruturas tridimensionais (ver Seção 5.3.1, “Estrutura primária de HSV2 TK”, pag.180). A Tabela 25 mostra os percentuais de homologia entre os diversos elementos da estrutura secundária de HSV1 TK e HSV2 TK. Está claro que, neste ponto do trabalho, não há condições de afirmar se os elementos da estrutura secundária de HSV2 TK possuem ou não a mesma extensão, em termos de número de aminoácidos, que os correspondentes elementos de HSV1 TK. Em princípio considerou-se que isso é verdadeiro.

Tabela 25 - Homologia entre os diversos elementos da estrutura secundária de HSV1 TK e HSV2 TK. No. Res. = número de resíduos em cada elemento de estrutura secundária; HT = homologia total e HF = homologia funcional, de acordo com a legenda do Diagrama 5.1, dadas em porcentagem. Nos casos em que HT é igual a HF, utilizou-se hífen na coluna de HF para efeito de simplificação.

Elem.	No. Res.	HT %	HF %	Elem.	No. Res.	HT %	HF %
$\alpha 1$	13	62	69	$\beta 1$	11	91	-
$\alpha 2$	6	83	-	$\beta 2$	7	86	-
$\alpha 3$	15	73	80	$\beta 3$	7	86	100
$\alpha 4$	30	87	93	$\beta 4$	9	78	100
$\alpha 5$	5	80	-	$\beta 4a$	3	67	-
$\alpha 6$	10	100	-	$\beta 5$	9	89	100
$\alpha 7$	11	100	-	$\beta 6$	6	50	83
$\alpha 8$	11	64	82				
$\alpha 9$	23	87	-	$\Omega 1$	4	75	-
$\alpha 10$	8	88	-	$\Omega 2$	4	75	-
$\alpha 11$	4	50	75	$\Omega 3$	5	100	-
$\alpha 12$	7	86	-	$\Omega 4$	3	100	-
$\alpha 13$	16	69	81	$\Omega 5$	13	69	77
$\alpha 14$	15	80	87	$\Omega 6$	3	100	-
$\alpha 15$	18	78	-	$\Omega 7$	6	67	83
				$\Omega 8$	9	100	-
				$\Omega 9$	20	30	45
				$\Omega 10$	11	91	100
				$\Omega 11$	2	100	-
				$\Omega 12$	3	67	-

Completada a primeira parte da modelagem de HSV2_TK, fez-se a modelagem das voltas ausentes da estrutura. Houve necessidade de refazer esta parte porque há pouca homologia (cerca de 30%) entre HSV1 TK e HSV2 TK, nas seqüências ausentes da estrutura cristalina que se dispõe. A Tabela 26 mostra as voltas selecionadas para HSV2_TK.

Tabela 26 - Voltas procuradas e voltas escolhidas para os monômeros A e B de HSV2_TK, que foram construídas a partir de HSV1_TK. Ver explicação na Tabela 23, pag. 186.

Monôm.	Volta procurada	Volta escolhida		RMS (lsq)	FIT (dist)
		Enzima: Início	Seqüência		
A	Leu-72-Asp-76	3CYT: Pro 72	PKKYIPGTK	1,641	1,138
B	Gly-73-Asp-76	1MCP: Glu 201	EATHKTST	1,674	0,418
A	Pro-150-Gln-151	1HMQ: Tyr 67	YAGY	0,983	0,038
B	Pro-150-Gln-151	3CYT: Asp 63	DTLM	0,828	0,008
A	Leu-262-Gly-279	5CHA: Gln 312	QKLKIAKVFKN KYNSTINNDIG	4,941	1,357
B	Thr-263-Gly-277	5CHA: Gly 250	GEEAVPGSWP WQVSLQDKTGL	3,821	1,333

Após a inserção das voltas e inclusão dos átomos de hidrogênio, HSV2_TK passou a ter 10106 átomos, sendo 5030 átomos não-hidrogenóides e 5076 átomos de hidrogênio.

Repetiu-se o procedimento acima para construir o dímero de hsv2_tk a partir da estrutura de hsv1_tk. As voltas utilizadas para construir esse modelo são mostradas na Tabela 27.

Tabela 27 - Voltas procuradas e voltas escolhidas para os monômeros A e B de hsv2_tk, que foram construídas a partir de hsv1_tk. Ver explicação na Tabela 23, pag. 186.

Monôm.	Volta procurada	Volta escolhida		RMS (lsq)	FIT (dist)
		Enzima: Início	Seqüência		
A	Leu-72-Asn-77	3CYT: Gly 24	GKGHKVGP NL	2,731	0,534
B	Gly-73-Asn-77	3CYT: Gly 57	GIVWNNDTL	1,690	0,510
A	Leu-262-Ser-280	1MCP: Ala 86	AEDLAVYYCQND HSYPLTFGAGTK	2,806	1,295
B	Thr-263-Gly-277	5CHA: Gly 434	GAWTLVGIVSWG SSTCSTST	3,893	1,401

5.4.5 Relaxamento da estrutura cristalina de HSV1 TK e do modelo de HSV2 TK

A minimização da energia é necessária para aliviar o acúmulo de energia potencial de determinadas regiões da molécula protéica. Se isso não for feito, esse acúmulo de energia potencial pode gerar forças muito grandes durante a execução da dinâmica

molecular, que por conseguinte podem modificar de forma indevida a conformação geral da estrutura enzimática.

Os sistemas cujas energias potenciais serão minimizadas compreendem:

- (a) HSV1_TK + 2 2'-deoxitimidina + 45 H₂O
- (b) HSV2_TK + 2 2'-deoxitimidina + 45 H₂O
- (c) hsv1_tk + 2 ganciclovir + 173 H₂O
- (d) hsv2_tk + 2 ganciclovir + 173 H₂O

Decidiu-se excluir o íon sulfato do sistema a ser estudado. A origem do íon sulfato é o sulfato de amônio utilizado durante o processo de cristalização (Brown et al, 1995). Na ausência de ATP, o íon sulfato ocupou o sítio de ligação de β -fosfato do ATP. Como há interesse apenas no sítio de ligação de 2'-deoxitimidina, e o sítio de ligação de β -fosfato fica relativamente longe daquele (há cerca de 6 Å entre os dois átomos mais próximos de 2'-deoxitimidina e do íon sulfato), não há motivo razoável para incluir o íon sulfato nesta etapa do trabalho.

Existem diversas maneiras de executar a minimização de energia e também muitos algoritmos podem ser utilizados. Parece ser bastante coerente não submeter todo o sistema enzima-solvente-ligante ao processo de minimização num único estágio. Ao invés disso, optou-se por minimizar a energia potencial em diversas etapas, o que significa manter fixas algumas partes da estrutura e fazer a minimização das partes restantes. Especificamente, adotou-se a seguinte estratégia para o processo de minimização em fase gasosa:

- (a) Aplicação de potencial harmônico forte (constante de força igual a 200 kcal.mol⁻¹.Å⁻²) aos átomos localizados na região mais externa do sistema, numa faixa de 3 Å ao redor da enzima (para prevenir o colapso da mesma, devido à ausência de forças externas

- que contrabalançam as interações atrativas dos átomos da superfície com os átomos localizados no interior do sistema);
- (b) 1000 passos de minimização *steepest descents* somente para as ligações X-H, onde X é um átomo não-hidrogenóide;
 - (c) 1000 passos de minimização *steepest descents* seguido de 1000 passos de minimização *conjugate gradient* somente para as cadeias laterais dos aminoácidos e o solvente;
 - (d) 1000 passos de minimização *steepest descents* seguido de 1000 passos de minimização *conjugate gradient* somente para o ligante e o solvente;
 - (e) 1000 passos de minimização *steepest descents* para toda a enzima e o solvente e;
 - (f) 1000 passos de minimização *steepest descents* seguido de 1000 passos de minimização *conjugate gradient* para todo o sistema.

Utilizou-se o programa CHARMM (Brooks et al, 1983) versão 22 para executar o processo de minimização de energia. Todo o processo de minimização descrito acima demandou cerca de 37 horas de tempo de CPU, em estação Indigo II Silicon Graphics, para cada sistema enzima-solvente-ligante.

Optou-se por não utilizar o sistema sob condições periódicas de fronteira, devido à imensa quantidade de moléculas de solvente que seriam introduzidas no mesmo. Apenas para ter-se idéia do problema, pode-se considerar a seguinte estimativa. Aproximando-se a forma irregular de HSV1 TK por um cubo de 80 Å de aresta. A caixa de solvente capaz de acomodar confortavelmente a enzima cúbica em seu interior deverá possuir algo como 120 Å de aresta. Uma estimativa grosseira indica que existem $3,34 \times 10^{-2}$ moléculas de água/Å³. Portanto, o número de moléculas de água necessárias para preencher o espaço vazio no interior da caixa de 120 Å, ao redor da caixa de 80 Å, é cerca de 40 mil, ou seja, cerca de dez vezes mais átomos que o conteúdo do próprio sistema enzima-inibidor. Preferiu-se, portanto, aplicar potencial harmônico forte nos átomos mais externos da

enzima e executar a minimização de energia apenas aos átomos localizados em seu interior. A aplicação do potencial harmônico previne o colapso da enzima, pela ausência de forças atrativas externas (minimização de energia em fase gasosa).

A situação energética dos complexos HSV1_TK/HSV2_TK-2'-deoxitimidina-água e hsv1_tk/hsv2_tk-ganciclovir-água, antes e depois do processo de minimização, pode ser vista na Tabela 28 e Tabela 29.

Tabela 28 - Situação energética, em kcal/mol, dos complexos de HSV1_TK-2 moléculas de 2'-deoxitimidina-45 moléculas de água (10209 átomos) e hsv1_tk-2 moléculas de ganciclovir-173 moléculas de água (10593 átomos) antes e depois do processo de minimização (fase gasosa), executado pelo programa CHARMM versão 25a1 (Brooks et al, 1983).

Contribuição	Energia, kcal/mol			
	HSV1_TK-2'-deoxitimidina-água		hsv1_tk-ganciclovir-água	
	Inicial	Final	Inicial	Final
Estiramento de ligação	1809	201	984	221
Deformação angular	2693	1659	2790	1751
Deformação torsional	1969	1876	1875	1823
Deformação fora do plano	187	151	166	140
Lennard-Jones	4766271	-5444	$1,2172 \cdot 10^{10}$	-5471
Eletrostática	-21645	-31568	-20084	-33791
Outras	0	0	-396	0
Total	4751284	-33125	$1,2172 \cdot 10^{10}$	-35327

Tabela 29 - Situação energética, em kcal/mol, dos complexos de HSV2_TK-2 moléculas de 2'-deoxitimidina-45 moléculas de água (10303 átomos) e hsv2_tk-2 moléculas de ganciclovir-173 moléculas de água (10687 átomos) antes e depois do processo de minimização (fase gasosa), executado pelo programa CHARMM versão 25a1 (Brooks et al, 1983).

Contribuição	Energia, kcal/mol			
	HSV2_TK-2'-deoxitimidina-água		hsv2_tk-ganciclovir-água	
	Inicial	Final	Inicial	Final
Estiramento de ligação	1929	202	1119	219
Deformação angular	2865	1676	2935	1729
Deformação torsional	1912	1934	1851	1855
Deformação fora do plano	234	148	265	145
Lennard-Jones	$3,0892 \cdot 10^{12}$	-5456	$5,6782 \cdot 10^{11}$	-5503
Eletrostática	-21091	-32365	-20932	-34580
Outras	21091	0	20932	0
Total	$3,0892 \cdot 10^{12}$	-33861	$5,6782 \cdot 10^{11}$	-36135

A energia inicial dos complexos envolvendo HSV2_TK e hsv2_tk é muito maior do que a energia dos correspondentes complexos envolvendo HSV1_TK e hsv1_tk. Isso é razoável porque as moléculas de HSV2 TK foram construídas a partir das moléculas de HSV1 TK. Diversas mutações foram feitas ao longo da sequência da molécula de HSV1 TK, e os aminoácidos introduzidos nessas mutações muitas vezes não encontraram ambiente adequado para suas cadeias laterais. Isso resultou em grande proximidade de algumas dessas cadeias laterais e, conseqüentemente, no aumento da energia potencial (interação não-ligante). No entanto, as energias finais dos complexos são bastantes próximas, sendo que os complexos envolvendo HSV2 TK possuem energia potencial ligeiramente superior aos correspondentes complexos envolvendo HSV1 TK. Isso também é razoável porque nossos modelos da molécula HSV2 TK possuem 94 átomos a mais do que os modelos da molécula de HSV1 TK.

Vale lembrar que os valores absolutos das energias apresentadas na Tabela 28 e Tabela 29 não possuem qualquer significado. Apenas são relevantes as diferenças de energia entre as estruturas.

A verificação da variação das posições atômicas antes e depois do processo de minimização é procedimento usual para avaliar se a estrutura protéica teve sua geometria incorretamente alterada durante o processo. Para isso, calculou-se o desvio quadrático médio das posições atômicas antes e depois da minimização de cada conjunto enzima-ligante-solvente. Os resultados dessa análise são mostrados na Tabela 30.

Tabela 30 - Desvio quadrático médio atômico dos sistemas enzima-ligante-solvente minimizados pelo programa CHARMM. O asterisco define a enzima que participa dos complexos mostrados na segunda linha.

Subsistema	Desvio quadrático médio (RMS) , Å			
	Enzima*-2'-deoxitimidina-45 H ₂ O		Enzima*-Ganciclovir-173 H ₂ O	
	* HSV1_TK	* HSV2_TK	* hsv1_tk	* hsv2_tk
Carbono α (CA)	0,924	0,977	0,861	0,897
Esqueleto protéico	1,037	1,078	0,955	0,980
Enzima completa	1,409	1,600	1,470	1,554
Ligante	1,260	1,452	1,356	0,737
Solvente	2,622	2,588	2,951	3,155

Os resultados dos desvios quadráticos médios apresentados na Tabela 30 podem ser considerados satisfatórios. O desvio quadrático de apenas cerca de 1 Å para o esqueleto protéico e de cerca de 1,5 Å para toda a estrutura enzimática é aceitável. Isso indica que não houve alteração brusca da molécula da enzima, mas apenas o ajuste esperado de suas coordenadas atômicas.

5.5 MODO DE LIGAÇÃO DAS N²-FENILGUANINAS

No presente trabalho, denominou-se *configuração bioativa* de um ligante no sítio ativo da molécula-alvo (receptor) à orientação espacial da molécula do ligante em relação

aos grupos químicos do receptor. Ignorou-se a utilização desse termo para designar a conectividade atômica numa molécula particular. A *conformação bioativa* da molécula ligante no sítio ativo da molécula-alvo diz respeito aos valores específicos de suas coordenadas internas, principalmente ângulos torsionais. Ao conjunto configuração e conformação bioativas chamou-se *modo de ligação*.

A atividade dos compostos bioativos é geralmente dependente de seu modo de ligação. Isto se deve ao caráter complementar da interação fármaco-receptor. A interação fármaco-receptor requer que grupos químicos da molécula do fármaco interajam com grupos químicos complementares do receptor (cadeias laterais de aminoácidos, no caso de enzimas). Estes possuem distribuição espacial definida e, em geral, não podem sofrer grandes alterações conformacionais devido à proximidade de outros grupos vizinhos. Disso resulta que o fármaco deverá adaptar-se mais ao receptor do que o receptor ao fármaco. No caso de moléculas de fármaco rígidas, como por exemplo anéis aromáticos condensados, não há muito o que considerar. Na interação de compostos de estrutura molecular rígida com o mesmo receptor, a plasticidade do sítio ativo deverá ser o fator conformacional mais importante a ser considerado. Entenda-se por plasticidade do receptor a capacidade de seus diversos grupos químicos poderem sofrer alterações conformacionais de forma a acomodar a molécula do fármaco. No caso da interação do receptor com moléculas flexíveis, a combinação de fatores entálpicos e entrópicos determinará o modo de ligação. O conhecimento do modo de ligação do fármaco é fundamental para qualquer procedimento de cálculo molecular, seja ele *ab initio*, semiempírico, de mecânica ou dinâmica molecular. Mas como descobrir a configuração e conformação bioativas ?

O método mais confiável que se dispõe atualmente para determinar o modo de ligação é a determinação das coordenadas dos átomos do fármaco através da análise do padrão de difração de raios-X do cristal contendo a enzima em complexo com o fármaco.

Se a determinação for feita com os devidos cuidados, as coordenadas revelam o modo de ligação do fármaco e também mostram as interações intermoleculares que ocorrem entre grupos químicos do fármaco e do receptor.

Freqüentemente dispõe-se da estrutura cristalográfica da molécula que contém o receptor em complexo com algum substrato, no caso de enzimas, ou em complexo com composto que guarde alguma similaridade com o fármaco em estudo. Neste caso, deve-se utilizar o modo de ligação da molécula complexada ao receptor como ponto de partida para a dedução da conformação bioativa do fármaco em estudo. No presente trabalho, o caso é precisamente este. Dispõe-se da estrutura cristalina de HSV1 TK em complexo com 2'-deoxitimidina e também em complexo com ganciclovir. Pretende-se utilizar as informações contidas nesses complexos, e também as informações reveladas no estudo QSAR do Capítulo 4, para tentar chegar à conformação bioativa das N^2 -fenilguaninas. Também serão utilizados cálculos de mecânica molecular para verificar a estabilidade dos prováveis modos de ligação.

5.5.1 Possibilidades

A similaridade dos modos de ligação de 2'-deoxitimidina e ganciclovir (ver Figura 39, pag. 166 e Figura 42, pag. 169) sugere que as N^2 -fenilguaninas também devam adotar modo de ligação parecido. Fatores como a formação da ponte de hidrogênio dupla com Gln 125 e as interações com Met 128 e Tyr 172 também devem ser importantes para o modo de ligação das N^2 -fenilguaninas. Devido ao fato de as N^2 -fenilguaninas guardarem maior similaridade com ganciclovir do que com 2'-deoxitimidina, deve-se considerar em primeiro lugar o modo de ligação de ganciclovir como molde para as N^2 -fenilguaninas (Gaudio et al, 1998).

A Figura 48a mostra que a configuração bioativa do anel purínico de ganciclovir é basicamente determinada pela formação de ponte de hidrogênio dupla com Gln 125 (ver também Figura 42, pag. 169). É possível ajustar a posição do anel purínico das N^2 -fenilguaninas de forma equivalente (Figura 48b).

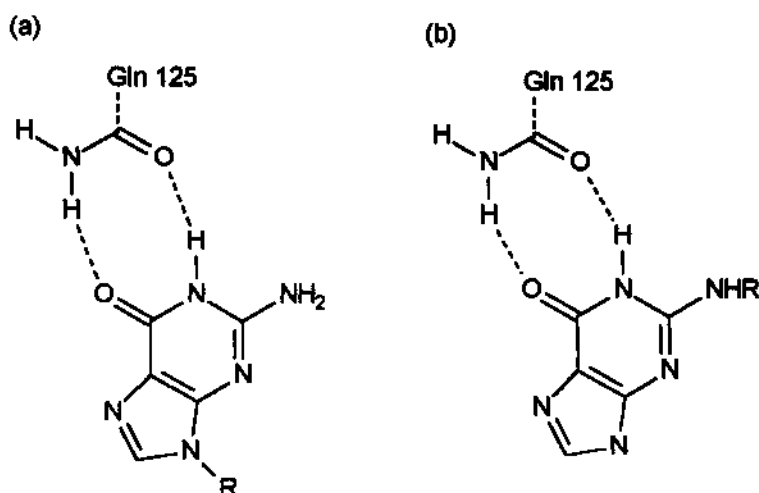


Figura 48 - Primeira possível configuração bioativa do anel purínico das N^2 -fenilguaninas. (a) A configuração do anel purínico de ganciclovir foi utilizado como molde (Brown et al, 1995). (b) A ponte de hidrogênio dupla com Gln 125 é usada como referência para o estabelecimento da configuração bioativa do anel purínico das N^2 -fenilguaninas.

Considerando-se a configuração bioativa das N^2 -fenilguaninas da Figura 48b, há duas possíveis conformações para o anel fenila: uma na qual o ângulo torsional θ_1 é próximo de 0° (Figura 49a) e outra em que θ_1 é próximo de 180° (Figura 49b).

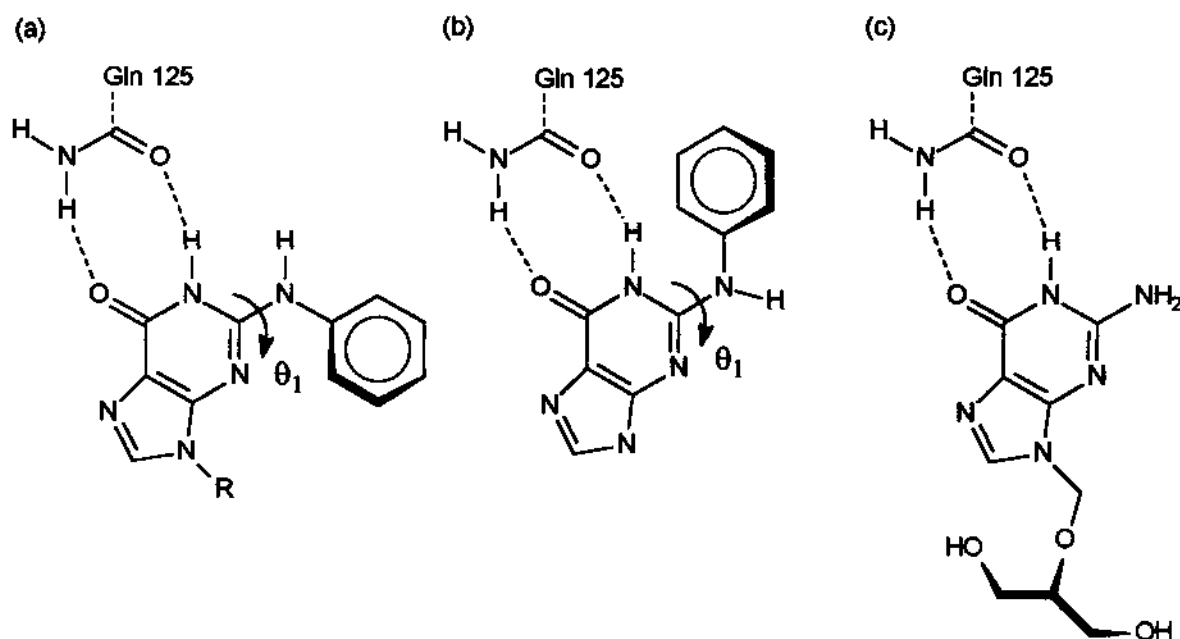


Figura 49 - Possíveis conformações do anel fenila de $\underline{N}^2$ -fenilguanina, considerando-se a configuração apresentada na Figura 48b. Na conformação (a) o ângulo torsional θ_1 é cerca de 0° e na conformação (b) θ_1 é aproximadamente 180° . Devido às interações intramoleculares, o anel fenila não se encontra no mesmo plano do anel purínico. (c) O modo de ligação de ganciclovir também é apresentado para efeito de comparação.

Não há espaço no sítio ativo de HSV1 TK para acomodar as conformações das $\underline{N}^2$ -fenilguaninas em que o ângulo torsional θ_1 seja apreciavelmente diferente de 0° ou 180° . Nas conformações da Figura 49a e Figura 49b, o anel fenila não se encontra no mesmo plano do anel purínico devido às interações intramoleculares. Neste ponto da discussão não há condições de especificar valores para os ângulos torsionais θ_1 e θ_2 , que determinam a orientação relativa dos anéis fenila e purínico. Mais tarde far-se-ão as devidas considerações sobre a possível orientação do anel fenila substituído na posição *meta* (ver Seção 5.5.7, “Ângulos torsionais θ_1 e θ_2 na conformação bioativa”, pag. 221).

A segunda possível configuração do anel purínico das $\underline{N}^2$ -fenilguaninas é similar à adotada pelo anel pirimidínico de 2'-deoxitimidina (Figura 50a). A formação de ponte de hidrogênio dupla com Gln 125 também é a principal âncora da molécula de 2'-

deoxitimidina. No entanto, a ponte de hidrogênio dupla que 2'-deoxitimidina forma com Gln 125 tem orientação invertida em relação à ponte de hidrogênio dupla de ganciclovir. O grupo CONH₂ de Gln 125 sofre rotação de 180° para tornar possível a formação da ponte de hidrogênio em cada caso (ver também a Figura 39, pag. 166 e Figura 42, pag. 169). Nesta segunda configuração (Figura 50b), o anel de purínico das N²-fenilguaninas é ajustado ao sítio ativo de HSV1 TK de forma a reproduzir a ponte de hidrogênio dupla de 2'-deoxitimidina.

Também poder-se-ia considerar duas possíveis conformações para o anel fenila de N²-fenilguanina, tendo em vista a configuração da Figura 50b, ou seja $\theta_1 \cong 0^\circ$ e $\theta_1 \cong 180^\circ$. No entanto, a análise da estrutura cristalográfica mostra claramente que não há espaço no sítio ativo de HSV1 TK para a conformação em que θ_1 é 0°. Nessa conformação há sério impedimento estérico que somente seria evitado às custas das interações com Gln 125, Met 128 e Tyr 172, que se considera a base para o estabelecimento do modo de ligação das N²-fenilguaninas. Portanto essa conformação não será considerada viável. A Figura 51a mostra a única conformação possível do anel fenila ($\theta_1 \cong 180^\circ$) considerando-se o modo de ligação de 2'-deoxitimidina como molde (Figura 51b).

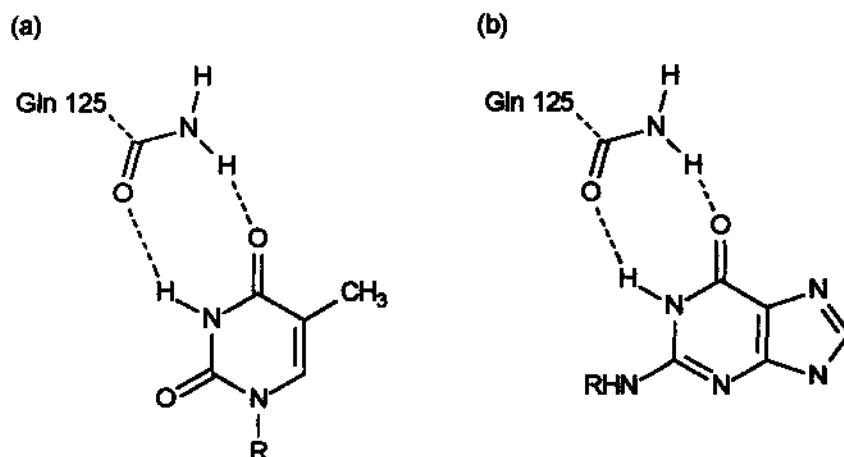


Figura 50 - Segunda possível configuração bioativa do anel purínico das N^2 -fenilguaninas. (a) A configuração do anel pirimidínico de 2'-deoxitimidina foi utilizada como molde (Brown et al, 1995); (b) a ponte de hidrogênio dupla com Gln 125 é usada como referência para o estabelecimento da configuração bioativa do anel purínico das N^2 -fenilguaninas.

À primeira vista não parece tarefa fácil optar-se por algum desses modos de ligação, em detrimento dos outros. Mas a determinação do modo de ligação correto é importante para a sequência deste trabalho. Para isso, fez-se a verificação da estabilidade relativa de cada um desses modelos, qualitativa e quantitativamente, e utilizaram-se as informações do estudo QSAR do Capítulo 4 para auxiliar na decisão final.

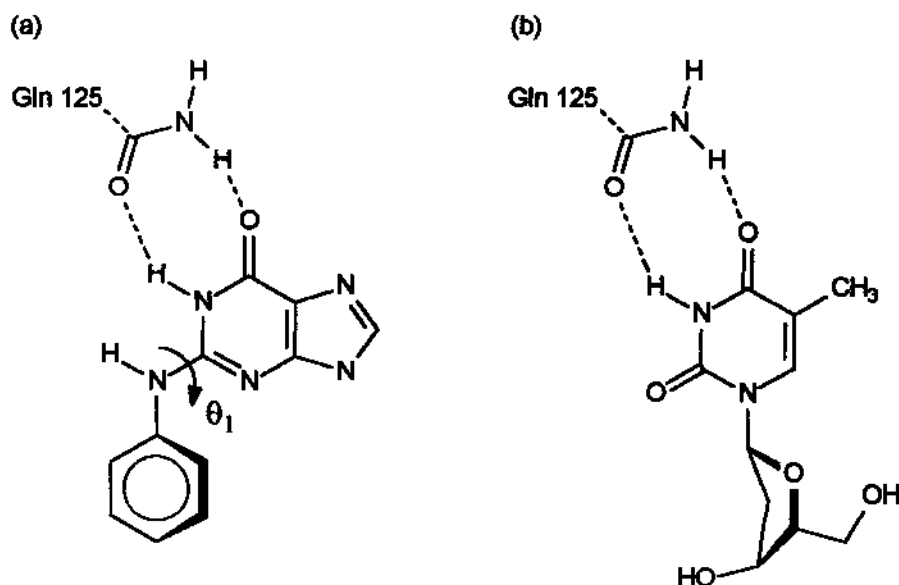


Figura 51 - Única conformação possível do anel fenila das *N*²-fenilguaninas, considerando-se a configuração apresentada na Figura 50b. (a) Nessa conformação, o ângulo torsional θ_1 é cerca de 180°. A conformação em que θ_1 é cerca de 0° não é permitida devido ao impedimento estérico entre o anel fenila e resíduos de aminoácidos do receptor. Devido às interações intramoleculares, o anel fenila não se encontra no mesmo plano do anel purínico. (b) O modo de ligação de 2'-deoxitimidina é apresentado para efeito de comparação.

5.5.2 Experimento computacional

Planejou-se experimento computacional para definir o modo de ligação das *N*²-fenilguaninas. O experimento tem dois objetivos. O primeiro objetivo é checar alguns parâmetros qualitativos, de forma a avaliar o aspecto visual da interação HSV1 TK-*N*²-fenilguaninas. O segundo objetivo é estimar a entalpia de interação HSV1 TK-*N*²-fenilguaninas, para alguns compostos da série, e relacionar esses valores às respectivas atividades inibitórias ($\log 1/IC_{50}$). Os modos de ligação a serem testados são aqueles discutidos na Seção 5.5.1 (“Possibilidades”, pag. 198).

Os parâmetros qualitativos a serem checados no experimento computacional são:

(a) estabilidade da ponte de hidrogênio dupla com Gln 125;

- (b) grau de paralelismo entre anel benzênico de Tyr 172 e o anel purínico nas N^2 -fenilguaninas;
- (c) direção para onde aponta o grupo SCH_3 de Met 128;
- (d) ambiente vizinho às posições meta e para do anel fenílico (hidro e lipofilicidade);

Espera-se que, no modo de ligação correto, o comportamento dos três primeiros parâmetros esteja de acordo com a Figura 39 (pag. 166) e Figura 42 (pag. 169). A natureza do ambiente vizinho às posições meta e para do anel fenílico requer alguma discussão.

No estudo QSAR clássico apresentado no Capítulo 4, apresentou-se a Equação (104) (pag. 142) como resultado da análise de 34 N^2 -fenilguaninas como inibidores de HSV1 TK.

$$\begin{aligned} \log(1/C) = & 0,594 (\pm 0,229) \pi_m - 0,414 (\pm 0,200) \pi_p + 1,048 (\pm 0,746) \sigma_m + 2,381 (\pm 0,871) R_p + \\ & + 5,077 (\pm 0,244) \end{aligned} \quad (106)$$

(n = 34; r = 0,897; s = 0,399; $F_{(4,29)} = 29,767$; p < 0,00001)

($Q^2 = 0,741$; $s_{PRESS} = 0,459$; SDEP = 0,430)

De acordo com a Equação (106), a atividade inibitória das N^2 -fenilguaninas sobre HSV1 TK é intensificada pela presença de substituinte lipofílico na posição meta e de substituinte hidrofílico na posição para do anel fenila. Ao mesmo tempo, a atividade inibitória é intensificada pela presença de substituinte na posição meta capaz de retirar elétrons por efeito indutivo e pela presença de substituinte na posição para capaz de retirar elétrons por efeito de ressonância. Como, em geral, substituintes que exercem efeito indutivo positivo são polares e substituintes lipofílicos são apolares, pode-se suspeitar que os resíduos de aminoácido que constituem a vizinhança da posição meta geram um ambiente anfifílico. Isto significa que se pode esperar nessa região a presença de aminoácidos polares e apolares. A Equação 5.1 também sugere que o ambiente vizinho à posição para do sítio ativo de HSV1 TK seja hidrofílico.

O objetivo do cálculo da energia de interação enzima-inibidor é comparar as entalpias de interação de uma série de derivados das N^2 -fenilguaninas com os respectivos valores de $\log 1/IC_{50}$. Nessa comparação serão considerados, para cada composto da série, os três modos de ligação discutidos na Seção 5.5.1 ("Possibilidades", pag. 198).

O experimento baseia-se no seguinte argumento. Considere o processo de interação enzima-inibidor representado pela equação



em que as linhas tracejadas indicam interações não-ligantes. A variação da energia livre de interação desse processo (ΔG_{inter}), como de qualquer outro processo químico, é representada pela Equação (107)

$$\Delta G_{\text{inter}} = \Delta H_{\text{inter}} - T \Delta S_{\text{inter}} , \quad (107)$$

em que ΔH_{inter} é a variação da entalpia de interação, T é a temperatura absoluta e ΔS_{inter} é a variação da entropia de interação. Numa série de inibidores em que ΔS_{inter} seja aproximadamente constante, implicaria em $\Delta \Delta S_{\text{inter}} \cong 0$, ao longo da série. Como consequência, a Equação (108) seria válida para a série considerada.

$$\Delta \Delta G_{\text{inter}} \cong \Delta \Delta H_{\text{inter}} \quad (108)$$

Pode-se considerar ΔH_{inter} como sendo o resultado da soma representada por

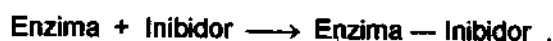
$$\Delta H_{\text{inter}} = \Delta H_{\text{EI}} + \Delta H_{\text{conf}} + \Delta H_{\text{solv}} , \quad (109)$$

em que ΔH_{EI} corresponde à variação de entalpia devido às interações não-ligantes (van der Waals, ponte de hidrogênio, etc.) entre enzima e inibidor, ΔH_{conf} é a variação de entalpia devido às alterações conformacionais e ΔH_{solv} é a variação de entalpia devido ao processo de solvatação e dessolvatação de enzima e inibidor. Se na série hipotética de compostos

ΔH_{conf} e ΔH_{solv} também fossem aproximadamente constantes, ter-se-ia $\Delta\Delta H_{\text{conf}} \cong \Delta\Delta H_{\text{solv}} \cong 0$, o que implicaria em

$$\Delta\Delta H_{\text{inter}} \cong \Delta\Delta H_{\text{EI}} \quad . \quad (110)$$

Pode-se representar o processo, cujo ΔH é medido por ΔH_{EI} , pela equação



em que não estão computadas as alterações conformacionais e nem a influência do solvente, uma vez que se considerou $\Delta\Delta H_{\text{conf}} \cong \Delta\Delta H_{\text{solv}} \cong 0$.

Portanto, para uma série de compostos que satisfaça todas as condições acima, a seguinte aproximação poderia ser válida:

$$\Delta\Delta G_{\text{inter}} \cong \Delta\Delta H_{\text{EI}} \quad . \quad (111)$$

É muito difícil encontrar uma série de compostos na qual todas as aproximações acima sejam válidas. Considerando-se a série de 36 derivados da N^2 -fenil-substituído-guanina apresentada na Tabela 1 (pag. 56), a subsérie mais próxima de satisfazer essas exigências é o conjunto de derivados para-substituídos: H, F, Cl, Br e I. O fato de serem para-substituídos e monoatômicos minimiza o problema de alterações conformacionais do anel fenila que certamente ocorreriam no caso dos derivados meta-substituídos ou poliatômicos. Isso implica na redução do valor de $\Delta\Delta H_{\text{conf}}$. O fato de serem monoatômicos, e portanto não muito volumosos, implica na redução do valor de $\Delta\Delta H_{\text{solv}}$. Para ter-se idéia da variação de volume nessa série, o derivado p-I é apenas cerca de 16% mais volumoso (toda a molécula, e não apenas o substituinte) do que o derivado não substituído (p-H). Como a maior contribuição de ΔS_{inter} é devido ao processo de solvatação e dessolvatação, espera-se que ΔS_{inter} também possa ser mantido em níveis aceitáveis para os nossos

propósitos. Outra vantagem em escolher-se a série 4'-H, 4'-F, 4'-Cl, 4'-Br e 4'-I reside nos valores de $\log 1/IC_{50}$ apresentados por esses derivados (ver Tabela 1, pag. 56): 4'-Br (6,00), 4'-H (5,10), 4'-I (4,40), 4'-Cl (4,00) e 4'-F (3,52). Pode-se notar que os valores de $\log 1/IC_{50}$ são regularmente distribuídos entre 3,5 e 6,0, que deve propiciar boas condições para a comparação com os valores de $\Delta\Delta H_{EI}$.

Os valores de $\log 1/IC_{50}$ apresentados na Tabela 1 (pag. 56) são linearmente relacionados com a variação da energia livre de interação enzima-inibidor. Valores elevados de $\log 1/IC_{50}$ (compostos muito ativos) correspondem a valores muito negativos de ΔG_{int} . Dessa forma, espera-se que os valores de $\Delta\Delta H_{EI}$ obtidos utilizando-se o modo de ligação correto sejam linearmente relacionados com os valores de $\log 1/IC_{50}$, pelo menos em termos aproximados. Ou seja, espera-se correlação linear entre $\log 1/IC_{50}$ e $\Delta\Delta H_{EI}$, como representado esquematicamente como a seguir.

$\log (1/IC_{50})_1$	$\log (1/IC_{50})_2$	$\log (1/IC_{50})_3$	$\log (1/IC_{50})_4$	$\log (1/IC_{50})_5$
$\Delta\Delta H_{EI1}$	$\Delta\Delta H_{EI2}$	$\Delta\Delta H_{EI3}$	$\Delta\Delta H_{EI4}$	$\Delta\Delta H_{EI5}$

Espera-se que a valores elevados de $\log 1/IC_{50}$ correspondam valores baixos de $\Delta\Delta H_{EI}$ e vice-versa.

5.5.3 Detalhes do experimento

A seguir são enunciados os detalhes do experimento computacional utilizado para a determinação do modo de ligação das N^2 -fenilguaninas, em que o programa Quanta-CHARMM foi utilizado.

- preparação dos derivados 4'-Br, 4'-H, 4'-I, 4'-Cl e 4'-F da N^2 -fenilguanina;
- eliminação das moléculas de solvente cristalizado de HSV1_TK e hsv1_tk;
- ajuste manual da posição da molécula do inibidor ao sítio ativo da enzima, sendo que os modos de ligação representados na Figura 49a-b foram ajustados à molécula

hsv1_tk e o modo de ligação representado na Figura 51a foi ajustado à molécula HSV1_TK;

- (d) aplicação de potencial harmônico fraco (constante de força igual a $10 \text{ kcal.mol}^{-1}.\text{\AA}^{-2}$) aos resíduos de aminoácidos completos, em que ao menos um de seus átomos estivessem num raio de $15 \text{ \AA}$ da molécula do inibidor;
- (e) liberação total dos resíduos de aminoácido completos em que ao menos um de seus átomos estivessem num raio de $8 \text{ \AA}$ da molécula do inibidor;
- (f) imobilização da molécula do inibidor e dos resíduos Gln 125, Met 128 e Tyr 172;
- (g) 1000 passos de minimização *steepest descents*;
- (h) liberação da molécula do inibidor e dos resíduos Gln 125, Met 128 e Tyr 172;
- (i) 1000 passos de minimização *conjugate gradient*;
- (j) aquecimento de 0 a 310 K durante 4 ps, com incrementos de 3,875 K a cada 0,05 ps;
- (j) equilíbrio em 310 K durante 4 ps, com correção de temperatura (velocidades atômicas) a cada 0,05 ps;
- (k) simulação livre em 310 K durante 8 ps;
- (l) resfriamento de 310 a 0 K durante 4 ps com decréscimos de 3,875 K a cada 0,05 ps;
- (m) 1000 passos de minimização *steepest descents*;
- (n) 1000 passos de minimização *conjugate gradient*;
- (o) avaliação dos parâmetros qualitativos mencionados na seção anterior;
- (p) cálculo da energia potencial do sistema enzima-inibidor, e da energia potencial do inibidor e da enzima, separadamente, utilizando-se as estruturas finais minimizadas;
- (q) cálculo de ΔH_{EI} e $\Delta \Delta H_{EI}$.

A ordem dessas operações é importante. Como exemplo, a imobilização dos resíduos Gln 125, Met 128 e Tyr 172, determinada no item (f), revoga a liberação dos

resíduos exigida pelo item (e). Cada sistema demandou cerca de 30 horas de CPU em estação Silicon Graphics Indigo 2.

A sequência de operações acima é o resultado da experiência obtida após algumas tentativas frustradas de executar a simulação da dinâmica molecular de maneira *bem comportada*. Enfrentou-se diversos problemas nessas tentativas, sendo que o principal problema foi o excessivo deslocamento da molécula do inibidor da posição inicial de interação com os resíduos Gln 125, Met 128 e Tyr 172. Naturalmente que algum deslocamento é esperado, pois trata-se de processo dinâmico. Além disso, submeteu-se à dinâmica molecular modos de ligação diferentes para as mesmas moléculas. É sempre esperado que, no modo de ligação incorreto, o ligante sintá-se desconfortável no sítio ativo e sofra algum tipo de modificação no modo de ligação. Porém, quando a modificação da situação inicial é muito intensa, é preciso suspeitar da forma como a simulação está sendo conduzida. Nas tentativas anteriores, começou-se a desconfiar da forma como as condições iniciais estavam sendo estabelecidas.

O primeiro fator que pode ter interferido negativamente nos primeiros resultados foi a maneira como foi feita a seleção dos átomos incluídos na simulação. Inicialmente optou-se por incluir os átomos da enzima localizados num raio de 12 Å da molécula do inibidor, sem a preocupação de incluir resíduos completos. Isso representou a inclusão parcial de diversos resíduos de aminoácidos no sistema. Pareceu ser mais prudente diminuir o valor da distância de corte (8 Å, por exemplo) e incluir apenas resíduos completos, ou seja, se dentro do raio de 8 Å da molécula do inibidor alguns resíduos tiverem parte de seus átomos selecionados, então incluir-se-iam todos os átomos daqueles resíduos. A inclusão na simulação de resíduos completos parece ser mais coerente do que a inclusão de partes de resíduos. Também criou-se zona de transição entre a liberdade total e a imobilidade total, aplicando constante harmônica de baixo valor aos resíduos em que ao

menos um de seus átomos estivessem num raio de 15 Å da molécula do inibidor. Os resíduos não incluídos por esse processo seletivo também estão presentes no sistema, mas só que mantidos fixos em suas posições iniciais.

Em algumas tentativas executou-se o processo inicial de minimização de energia mantendo todo o sistema livre, qual seja, molécula do inibidor e os resíduos de aminoácidos que compõe o sítio ativo. Concluiu-se que melhor seria manter a molécula do inibidor e os resíduos Gln 125, Met 128 e Tyr 172 fixos nos primeiros mil passos de minimização e só liberá-los na segunda parte da minimização inicial. Esse procedimento manteve a posição inicial do inibidor por algum tempo, enquanto as cadeias laterais de aminoácidos próximas iriam se adaptando à presença da nova molécula.

O último fator importante foi a presença de solvente cristalizado. A presença de determinada molécula no sítio ativo de HSV TK, seja substrato ou inibidor, induz moléculas de solvente a localizarem-se em diferentes posições. Portanto, pode haver erro importante ao substituir-se determinada molécula ligada ao sítio ativo de HSV TK por outra, sem fazer correções na posição do solvente cristalizado nas vizinhanças. Na estrutura cristalográfica contida no arquivo HSV1_TK_dT_2 há uma molécula de água bem próxima à ponte de hidrogênio dupla. No entanto não é tarefa simples determinar a posição correta de moléculas de solvente ao redor do inibidor em curtos períodos de simulação computacional. É preciso ainda considerar que pode haver moléculas de solvente nas redondezas do sítio ativo que, por terem maior grau de movimento, não foram detectadas na difração de raios-X. Dessa forma, pareceu ser mais seguro não incluir as moléculas de solvente cristalizado nesse experimento. Além disso, uma das aproximações utilizadas no experimento computacional é de $\Delta\Delta H_{\text{solv}} \equiv 0$ (ver Seção 5.5.2, “Experimento computacional”, pag. 203)

5.5.4 Resultados

Os resultados obtidos na parte qualitativa do experimento computacional são mostrados no Quadro 4.

Quadro 4 - Situação final da região do sítio ativo dos sistemas HSV1 TK- N^2 -fenil-[para-substituído]-guanina após curta simulação de dinâmica molecular seguida de minimização. Caracteres em negrito foram utilizados para destacar a comparação entre os resultados obtidos e o resultado esperado.

Modo de ligação	Sítio ^a	Gln 125 ^b	Met 128 ^c	Tyr 172 ^d	Meta ^e	Para ^e
Resultado esperado	A	2/2	Sim	Sim	Anfi Anfi	Hidro
	B	2/2	Sim	Sim	Anfi Anfi	Hidro
Tipo ganciclovir ^f ($\theta_1 = 0^\circ$)	A	1/2	Sim	Não	Anfi Lipo	Lipo
	B	1/2	Não	Sim	Lipo Anfi	Lipo
Tipo ganciclovir ^g ($\theta_1 = 180^\circ$)	A	0/2	Sim	Não	Lipo Lipo	Lipo
	B	0/2	Não	Sim	Lipo Lipo	Lipo
Tipo 2'-deoxitimidina ^h ($\theta_1 = 180^\circ$)	A	1/2	Não	Sim	Anfi Anfi	Hidro
	B	1/2	Sim	Sim	Anfi Hidro	Hidro

^a A e B indicam em qual monômero, do dímero de HSV TK, a análise está sendo feita. ^b Ponte de hidrogênio dupla entre Gln 125 e a molécula do inibidor. 2/2 = duas pontes de hidrogênio ocorreram; 1/2 = embora duas pontes de hidrogênio fossem possíveis, apenas uma ocorreu; 0/2 = embora duas pontes de hidrogênio fossem possíveis, nenhuma ocorreu. ^c Direção para a qual aponta o grupo SCH₃ de Met 128. Sim = apontando para o anel guanina da molécula do inibidor; Não = apontando para outra direção. ^d Posição espacial do anel benzênico de Tyr 172. Sim = paralelo ao anel guanina do inibidor; Não = não paralelo ao anel guanina. ^e Ambiente do sítio ativo nas vizinhanças das posições meta e para. Na coluna meta, há duas designações para o ambiente por causa das duas posições meta do anel fenila. Hidro = hidrofílico; Lipo = lipofílico; Anfi = anfifílico. ^f Ver Figura 49a, pag. 200. ^g Ver Figura 49b, pag. 200. ^h Ver Figura 51a, pag. 203.

Esses resultados são apresentados em termos de média sobre o comportamento dos derivados 4'-H, 4'-F, 4'-Cl, 4'-Br e 4'-I- N^2 -fenilguanina. Como exemplo, na análise da formação da ponte de hidrogênio dupla entre Gln 125 e a molécula do inibidor (modo de

ligação do tipo ganciclovir e $\theta_1 = 0^\circ$) no sítio ativo A de HSV1 TK, o resultado apresentado é 1/2. Isso significa que apenas metade da ponte de hidrogênio dupla foi formada no sítio A na maioria das cinco moléculas analisadas (4'-H, 4'-F, 4'-Cl, 4'-Br e 4'-I).

Os resultados apresentados no Quadro 4 apontam para o modo de ligação de 2'-deoxitimidina como sendo o modo de ligação preferencial das N^2 -fenilguaninas. Nesse modo de ligação, houve relativa estabilidade da ponte de hidrogênio dupla com Gln 125 e satisfatório grau de paralelismo entre o anel guanina e Tyr 172. Além disso, o modo de ligação do tipo 2'-deoxitimidina é o único que permite ambiente hidrofílico próximo à região para do anel fenila, e também permite ambiente anfifílico próximo a uma das posições meta, de acordo com o previsto (ver Seção 5.5.2, “Experimento computacional”, pag. 203). Considerando-se a similaridade entre as N^2 -fenilguaninas e a molécula de ganciclovir, poder-se-ia esperar que os resultados apontassem para o modo de ligação similar ao de ganciclovir.

A Tabela 31 mostra os resultados da parte quantitativa do experimento computacional.

Tabela 31 - Variação da variação da entalpia de interação HSV1 TK-inibidor ($\Delta\Delta H_{EI}$) dos derivados p-H, p-F, p-Cl e p-Br e p-I. Números em negrito foram utilizados para destacar a relação linear entre $\log 1/IC_{50}$ e $\Delta\Delta H_{EI}$.

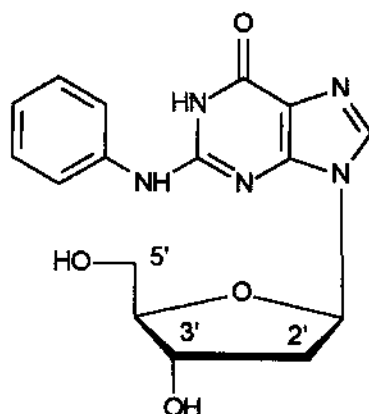
Derivado	$\log 1/IC_{50}$	$\Delta\Delta H_{EI}$, kcal.mol ⁻¹		
		Tipo ganciclovir ($\phi_1 = 0^\circ$)	Tipo ganciclovir ($\phi_1 = 180^\circ$)	Tipo 2'- deoxitimidina ($\phi_1 = 180^\circ$)
p-Br	6,00	26,0	0,0	0,0
p-H	5,10	4,7	28,2	9,7
p-I	4,40	0,0	21,3	21,3
p-Cl	4,00	27,1	25,3	24,7
p-F	3,52	12,4	17,9	18,0

A análise dos resultados apresentados na Tabela 31 mais uma vez aponta na direção do modo de ligação de 2'-deoxitimidina como sendo o modo de ligação preferencial das N²-fenilguaninas no sítio ativo de HSV1 TK. O coeficiente de correlação entre os valores de $\Delta\Delta H_{EI}$, para cada um dos três modos de ligação, e $\log 1/IC_{50}$ são: 0,185 para o modo de ligação do tipo ganciclovir com $\theta_1 = 0^\circ$, 0,585 para o modo de ligação do tipo ganciclovir com $\theta_1 = 180^\circ$ e 0,879 para o modo de ligação do tipo 2'-deoxitimidina com $\theta_1 = 180^\circ$. Este último coeficiente de correlação torna-se 0,995 se o derivado 4'-F não for considerado. O desvio da relação linear do valor de $\Delta\Delta H_{EI}$ do derivado 4'-F pode estar relacionado ao comportamento anômalo da entropia de solvatação de muitos derivados fluorinados (Cuthbertson e Richards, 1986).

Neste ponto do trabalho, há, portanto, duas previsões computacionais de que o modo de ligação do tipo 2'-deoxitimidina, em que $\theta_1 = 180^\circ$, seja o modo de ligação dos derivados da N²-fenilguanina. Mas haverá alguma indicação experimental de que isso seja verdadeiro?

5.5.5 Indicação experimental do modo de ligação das N^2 -fenilguaninas

Focher e colaboradores (1988) verificaram que os derivados da N^2 -fenil-(substituído)-2'-deoxiguanosina não sofrem fosforilação durante o processo de interação com o sítio ativo de HSV1 TK, ao contrário de aciclovir e ganciclovir, que são fosforilados. O motivo da não fosforilação dos derivados da N^2 -fenilguanina pode ser explicado a partir da análise do modo de ligação das N^2 -fenilguaninas.



N^2 -Fenil-2'-deoxiguanosina

O modo de ligação de ganciclovir com HSV1 TK já foi apresentado na Figura 42 (pag. 169). Esse modo de ligação permite que uma das hidroxilas da molécula de ganciclovir fique em posição espacialmente favorável para ser fosforilada (Figura 52a). Se os derivados da N^2 -fenilguanina utilizassem o modo de ligação de ganciclovir, não haveria razão para que não houvesse fosforilação durante a interação com HSV1 TK. A superposição do anel guanina de N^2 -fenil-2'-deoxi-guanosina com o anel guanina de ganciclovir (Figura 52b), na estrutura cristalográfica de hsv1_tk, mostra que os átomos de oxigênio passíveis de fosforilação das moléculas desses inibidores ficam separados por apenas 0,5 Å. Esse pequeno deslocamento provavelmente não seria suficiente para impedir a fosforilação.

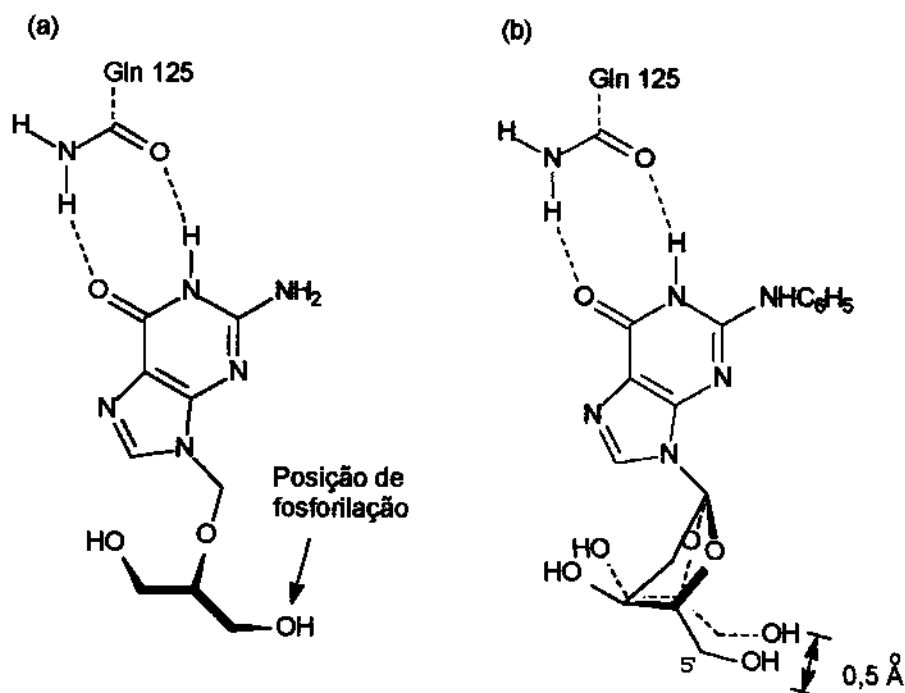


Figura 52 - Comparação da posição dos átomos passíveis de fosforilação de ganciclovir e N^2 -fenil-2'-deoxi-guanosina. (a) Modo de ligação de ganciclovir no sítio ativo de HSV1 TK e posição espacial do átomo de oxigênio que é fosforilado; (b) Superposição das moléculas de ganciclovir e N^2 -fenil-2'-deoxi-guanosina. A distância entre os átomos de oxigênio das hidroxilas passíveis de fosforilação das duas moléculas é cerca de 0,5 Å. Se as N^2 -fenilguanossinas adotassem o modo de ligação de ganciclovir, o átomo O5' ficaria em posição espacial favorável para ser fosforilado.

A Figura 53 mostra o modelo gráfico da sobreposição da molécula de ganciclovir, em sua conformação bioativa, e N^2 -fenil-2'-deoxiguanosina, que corresponde à representação esquemática da Figura 52b.

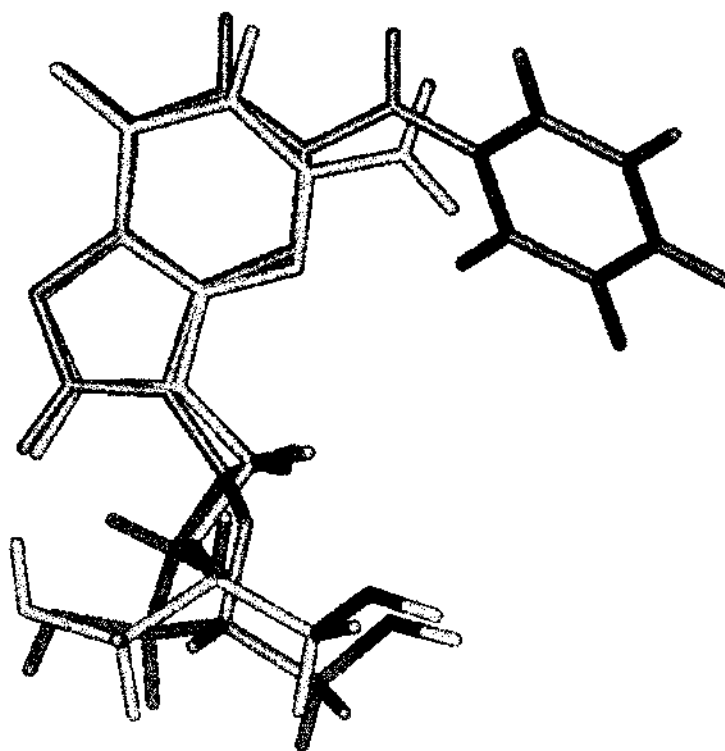


Figura 53 - Sobreposição da molécula de ganciclovir (amarelo), em sua conformação bioativa, e N^2 -fenil-2'-deoxiguanosina (azul). A distância entre os átomos de oxigênio das hidroxilas destacadas é cerca de 0,5 Å.

Se, por outro lado, for considerado que o modo de ligação dos derivados de N^2 -fenilguanossina com o sítio ativo de HSV1 TK ocorre de forma similar ao modo de interação de 2'-deoxitimidina (Figura 54a), então esse modo de ligação não permitirá o correto posicionamento do átomo O5' do anel ribose para fosforilação. A sobreposição do anel timidina de 2'-deoxitimidina com o anel guanina de N^2 -fenil-2'-deoxi-guanosina, mostra que a menor distância que os átomos O5' das duas moléculas podem ser posicionados é cerca de 2,8 Å (Figura 54b). Ao computar essa distância, o aspecto energético da conformação do anel ribose da N^2 -fenilguanossina não foi considerado, ou seja, manipulou-se os ângulos torsionais ϕ_1 e ϕ_2 de forma a obter a máxima aproximação entre os átomos O5' das duas moléculas sem preocupação com o fator energia de cada conformação. Isso significa que a distância entre os átomos O5' pode ser ainda maior. Tal

distância da posição ideal de fosforilação pode impedir a interação entre O5' e γ -fosfato de ATP com subsequente inibição do processo de fosforilação.

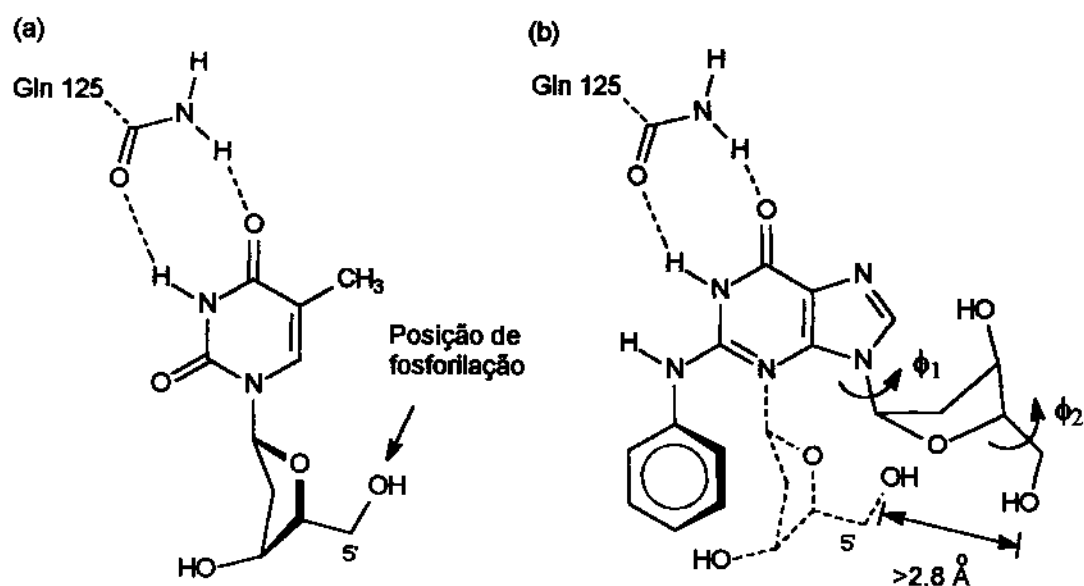


Figura 54 - Comparação da posição dos átomos O5' de 2'-deoxitimidina e N²-fenil-2'-deoxiguanosina. (a) Modo de ligação de 2'-deoxitimidina ao sítio ativo de HSV1 TK e posição espacial do átomo O5'. (b) Superposição das moléculas de 2'-deoxitimidina e N²-fenil-2'-deoxiguanosina. A menor distância entre os átomos O5' das duas moléculas que pode ser obtida por manipulação dos ângulos torsionais ϕ_1 e ϕ_2 é cerca de 2.8 Å.

A Figura 55 mostra o modelo gráfico da sobreposição da molécula de 2'-deoxitimidina, em sua conformação bioativa, e N²-fenil-2'-deoxiguanosina, que corresponde à representação esquemática da Figura 54b.

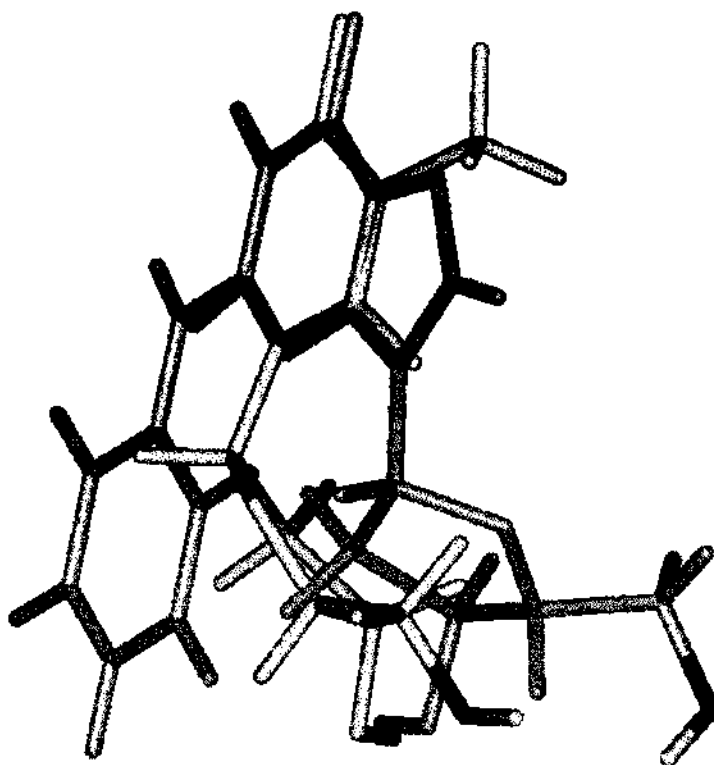


Figura 55 - Sobreposição da molécula de 2'-deoxitimidina (amarelo), em sua conformação bioativa, e N^2 -fenil-2'-deoxiguanosina (azul). A distância entre os átomos de oxigênio das hidroxilas destacadas é superior a 2,8 Å.

5.5.6 Natureza das regiões do sítio ativo próximas às posições meta e para

Os resultados do estudo QSAR do Capítulo 4 indicaram que as regiões dos sítios ativos de HSV1 TK e HSV2 TK próximas à posição meta do anel fenila das N^2 -fenilguaninas apresentam caráter lipofílico/polar diferentes (Seção 4.9, “Conclusões do estudo QSAR com parâmetros clássicos”, pag. 148). De acordo com a Seção 4.9, a região do sítio ativo de HSV1 TK próxima à posição meta apresenta natureza anfifílica, predominando o caráter apolar, enquanto que em HSV2 TK a região próxima à posição meta apresenta natureza anfifílica, predominando o caráter polar. Sugeriu-se que essas observações são reflexo de possíveis diferenças na natureza dos aminoácidos que constituem a região próxima à posição meta do anel fenila das N^2 -fenilguaninas. Pode-se verificar se essas informações são consistentes com o modo de ligação das N^2 -

fenilguaninas através da comparação dos aminoácidos que estão presentes naquela região em HSV1 TK e HSV2 TK.

A Figura 56 mostra a os aminoácidos que constituem a vizinhança da posição meta das N^2 -fenilguaninas em HSV1 TK. São mostrados os resíduos Tyr 101 (polar e não ionizável), Met 231 (apolar), Ile 97 (apolar) e Glu 225 (polar e ionizável, embora somente a parte hidrocarbônica de sua cadeia lateral esteja próxima à posição meta). Também é mostrado o resíduo Ala 98, pouco mais afastado da posição meta.

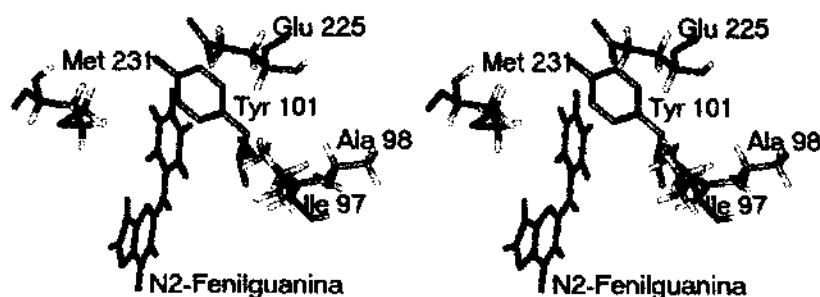


Figura 56 – Região do sítio ativo de HSV1 TK nas vizinhanças da posição meta do anel fenila das N^2 -fenilguaninas.

A Figura 57 mostra os aminoácidos que constituem a vizinhança da posição meta das N^2 -fenilguaninas em HSV2 TK. Nota-se grande semelhança na constituição dos resíduos em relação à HSV1 TK, exceto pela presença de Leu 97 (apolar) no lugar de Ile 97 e, mais afastado, Thr 98 (polar e neutro) no lugar de Ala 98.

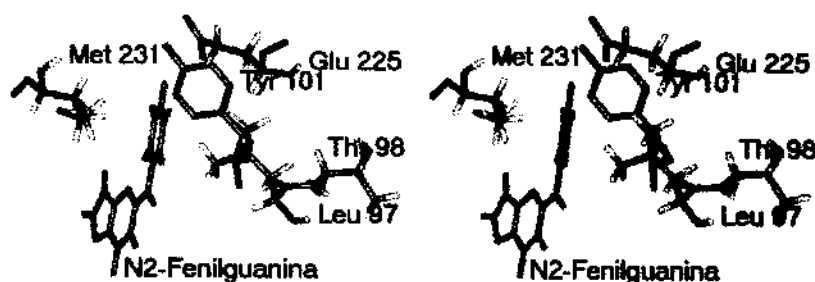


Figura 57 - Região do sítio ativo de HSV2 TK nas vizinhanças da posição meta do anel fenila das N^2 -fenilguaninas.

Observando-se a Figura 56 e a Figura 57, percebe-se que a substituição de Ile 97 (apolar) por Leu 97 (apolar) não é capaz de justificar a predominância do caráter polar do sítio ativo de HSV1 TK (ver Seção 4.8.1, “Atividade inibitória sobre HSV1 TK”, pag. 142) nem tampouco a predominância do caráter polar do sítio ativo de HSV2 TK previsto na Seção 4.8.2 (“Atividade inibitória sobre HSV2 TK”, pag. 145). Uma substituição do tipo Ala 98 (ápolar) por Thr 98 (polar e neutro) certamente o seria. Porém, parece que essa substituição ocorre muito longe da região meta, pelo menos de acordo com os modelos apresentados na Figura 56 e na Figura 57 (na Figura 56, a distância entre o átomo de carbono- α de Leu 98 e o átomo de carbono da posição meta, mais próximo de Leu 98, é cerca de 9 Å). No caso de alguma alteração conformacional conseguir aproximar mais o resíduo 98 da posição meta, ou no caso de o anel fenila dos derivados da N^2 -fenilguanina conseguir aproximar-se mais do resíduo 98, ou ambos aproximarem-se simultaneamente, ter-se-ia boa justificativa para explicar a diferença na polaridade dos sítios de HSV1 TK e HSV2 TK. No momento, a diferença de polaridade permanece sem explicação razoável.

5.5.7 Ângulos torsionais θ_1 e θ_2 na conformação bioativa

O Quadro 4 (pag. 211) indicou que apenas o modo de ligação do tipo 2'-deoxitimidina permitiu ambiente hidrofílico próximo à posição para do anel fenila e ambiente anfifílico próximo a uma das posições meta. No experimento computacional, o valor inicial de θ_1 foi de 180° . Após as simulações, seu valor médio foi de $180^\circ (\pm 10^\circ)^{10}$, computado em dez observações (sítios A e B de HSV1 TK com os cinco derivados para-substituídos analisados). O valor médio do ângulo torsional θ_2 , que permite a proximidade entre a posição meta e a região anfifílica, é de $218^\circ (\pm 11^\circ)$. Esses valores de θ_1 e θ_2 , que definem a conformação bioativa das N^2 -fenilguaninas, são bastante próximos dos valores de θ_1 e θ_2 da conformação C3 (186° e 233° , respectivamente), discutida no Capítulo 3 (ver 3.6.1, "Principais conformações de equilíbrio das N^2 -fenilguaninas", pag. 117).

Os resíduos de aminoácido encontrados com mais frequência nas proximidades da posição meta, que formam ambiente anfifílico, foram Tyr 101 (polar e não ionizável), Met 231 (apolar), Ile 97 (apolar) e a parte hidrocarbônica da cadeia lateral de Glu 225 (polar e ionizável). O substituinte ligado na posição para do anel fenila fica próximo à região altamente polar, constituída pelos resíduos His 58 (polar e ionizável), Gln 221 (polar e neutro), Arg 222 (polar e ionizável) e Glu 225 (polar e ionizável).

A Figura 58 mostra a conformação bioativa das N^2 -fenilguaninas, de acordo com as conclusões da presente análise.

¹⁰ Desvio-padrão da média.

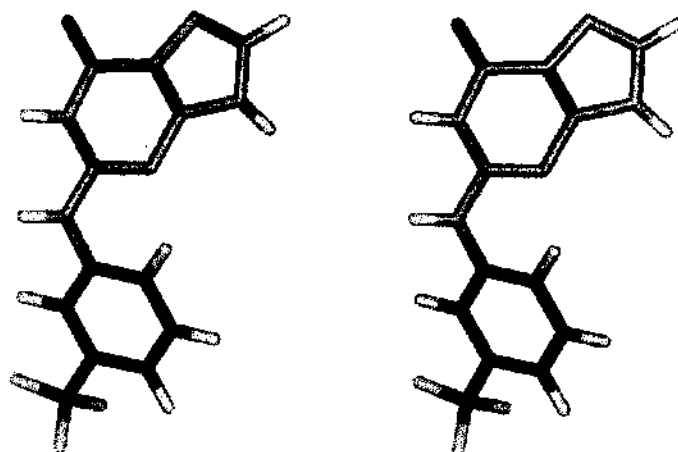


Figura 58 - Estereograma da conformação bioativa das N^2 -fenilguaninas, de acordo com as conclusões do presente estudo, exemplificada pelo derivado 3'-CF₃ (composto 1, Tabela 1, pag. 56). Esta conformação é semelhante à conformação C3, apresentada na Seção 3.6.1 ("Principais conformações de equilíbrio das N2-fenilguaninas", pag.117)

5.6 CONCLUSÕES DA ANÁLISE CONFORMACIONAL

- (a) Os derivados da N^2 -fenilguanina interagem com o sítio ativo de HSV1 TK adotando o modo de ligação de 2'-deoxitimidina;
- (b) Na conformação bioativa, os valores dos ângulos torsionais θ_1 e θ_2 , valem aproximadamente 180° e 220°, respectivamente, que corresponde à conformação C3 prevista pela análise conformacional;
- (c) o substituinte ligado na posição meta do anel fenila fica preferencialmente orientado na direção dos resíduos Tyr 101 (polar e não ionizável), Met 231 (apolar), Ile 97 (apolar) e a parte hidrocarbônica da cadeia lateral da cadeia lateral de Glu 225 (polar e ionizável), que compõe o ambiente anfifílico previsto pelo estudo QSAR do Capítulo 4;
- (d) o substituinte ligado na posição para do anel fenila fica próximo à região altamente polar, constituída pelos resíduos His 58 (polar e ionizável), Gln 221 (polar e neutro),

Arg 222 (polar e ionizável) e Glu 225 (polar e ionizável), em acordo com a previsão do estudo QSAR do Capítulo 4;

- (e) a observação experimental acerca da interação não fosforilativa entre os derivados da N^2 -fenilguanina e HSV1 TK fornece suporte adicional para a hipótese do modo de ligação proposto;
- (f) as conclusões acima reforçam a hipótese de que o anel guanina das N^2 -fenilguaninas interagem com os resíduos Gln 125, Met 128 e Tyr 172 de forma semelhante à interação desses resíduos com 2'-deoxitimidina.

A representação gráfica do modo de ligação das N^2 -fenilguaninas é mostrada na Figura 59.

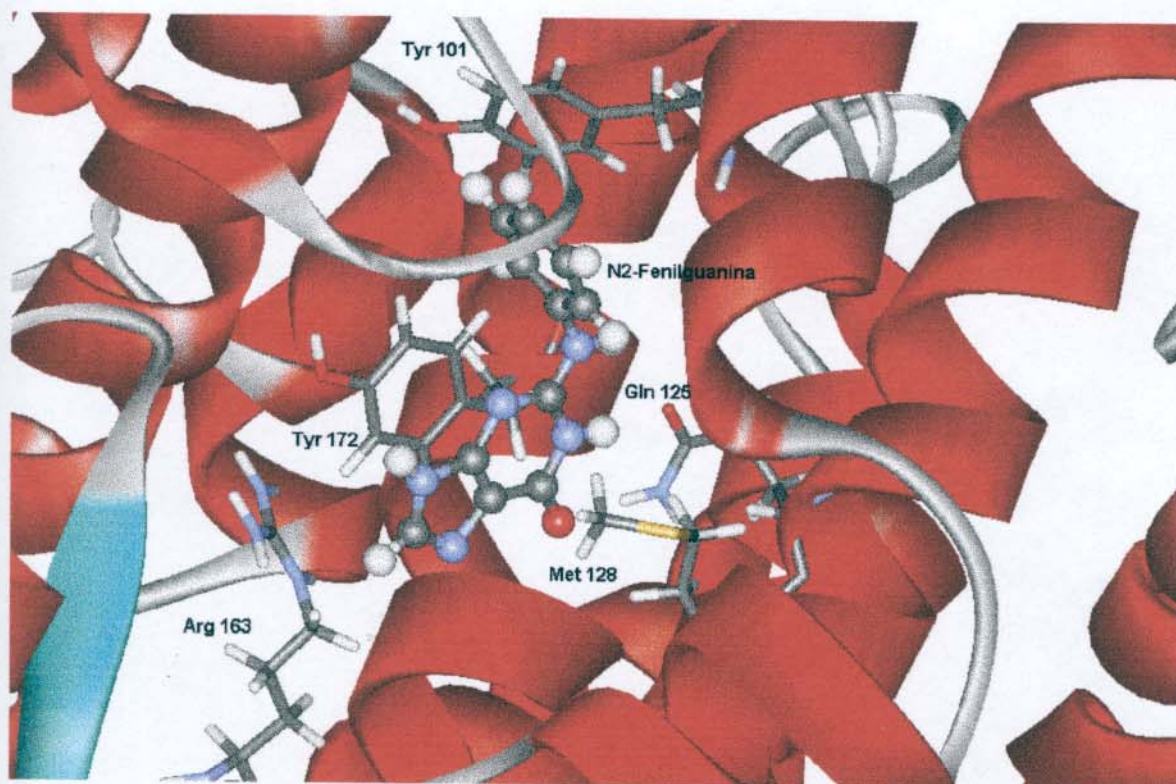


Figura 59 - Representação gráfica do modo de ligação das N^2 -fenilguaninas. Os substituintes ligados à posição meta do anel fenila orientam-se em direção ao resíduo Tyr 101.

A Figura 60 mostra o estereograma do modo de ligação das N^2 -fenilguaninas, proposto com base nos resultados obtidos neste capítulo.



Figura 60 - Estereograma do modo de ligação proposto para as N^2 -fenilguaninas.

CAPÍTULO 6

QSAR COM PARÂMETROS CLÁSSICOS E CALCULADOS

6.1 INTRODUÇÃO

No método clássico de Hansch-Fujita, os parâmetros físico-químicos utilizados são constantes de substituintes linearmente relacionadas com a energia livre (Kubinyi, 1993). No entanto, tornou-se prática comum a inclusão de parâmetros não linearmente relacionados com a energia livre, tais como parâmetros que descrevem propriedades localizadas da molécula, como por exemplo cargas e índices de reatividade atômicos, e propriedades da molécula como um todo, como volume molecular e energia de orbitais moleculares (Karelson et al, 1996). A extensão do método de Hansch-Fujita no sentido de acomodar tais variáveis tem resultado não raras vezes em ferramenta útil para esclarecer alguns tipos de interação fármaco-receptor, em nível molecular, que não são possíveis com o uso de outros parâmetros clássicos (Gaudio et al, 1994; Pires et al, 1997; Karelson et al, 1996)

Na presente análise, construiu-se a matriz de parâmetros físico-químicos a partir das constantes de substituintes clássicas da Tabela 19 (pag.151) e adicionando-se diversos parâmetros calculados. O resultado foi uma matriz de parâmetros físico-químicos mistos. No desenvolvimento das equações que serão apresentadas adiante, fez-se a seleção das variáveis utilizando-se toda a matriz mista.

6.1.1 Geometria Molecular

No cálculo das propriedades atômicas e moleculares, utilizou-se a geometria bioativa proposta na Seção 5.5.7 (“Ângulos torsionais θ_1 e θ_2 na conformação bioativa”,

pag.221). Como foi discutido na Seção 5.5.7, a geometria bioativa proposta corresponde à conformação C3 (ver Seção 3.6.1 “Principais conformações de equilíbrio das N^2 -fenilguaninas”, pag. 117), em que os ângulos torsionais θ_1 e θ_2 valem aproximadamente 180° e 230° (ver Figura 26, pag. 122).

6.1.2 Parâmetros incluídos na análise

Os parâmetros incluídos na presente análise foram:

- (a) Carga atômica (Q_n) do n -ésimo átomo. Foram calculadas as cargas dos átomos que estão presentes em todas as moléculas dos derivados da N^2 -fenilguanina: N(1)–H(22), H(25) e H(26) (ver Figura 61). As cargas atômicas foram calculadas com base no ajuste (*fitting*) do potencial eletrostático molecular, utilizando-se o método AM1;
- (b) Energias do HOMO (ϵ_{HOMO}) e LUMO (ϵ_{LUMO}). As energias do orbital molecular ocupado de energia mais alta e do orbital desocupado de energia mais baixa foram calculadas pelo método AM1;
- (c) Dureza molecular (η). Definida pela equação $(\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}})/2$;
- (d) Momento dipolar (μ). Calculado a partir das cargas eletrostáticas do item (a);
- (e) Polarizabilidade molecular (α). Calculada a partir da expansão do momento dipolar (Kurtz et al, 1990), através do método AM1;
- (f) Índices de fronteira. Foram calculados os seguintes índices de reatividade: densidade eletrofílica de fronteira ($F_n^{(e)}$) e densidade nucleofílica de fronteira ($F_n^{(n)}$) do n -ésimo átomo da molécula (Chin & Song, 1981). Estes índices foram normalizados pelas energias do HOMO e LUMO (ver Seção 2.4.7, “Índices de fronteira”, pag. 73) para maiores detalhes. De forma semelhante às cargas atômicas, foram calculados os índices de fronteira dos átomos que estão presentes em todas as moléculas dos derivados da

N^2 -fenilguanina. Esses índices foram calculados a partir da função de onda produzida através do método AM1;

(g) Volume molecular de van der Waals (V_{vdw}). O volume de van der Waals foi calculado através do programa SURF, criado pelo autor, com cerca de 1000 linhas de código fonte em Fortran 77 (Gaudio & Takahata, 1992).

A numeração dos átomos da N^2 -fenilguanina, que serviu de base para definir os índices de reatividade e as cargas atômicas, é mostrada na Figura 61. Eventualmente utilizar-se-á numeração especial para se referir aos átomos de carbono do anel fenila, como por exemplo 3'-F para se referir ao substituinte fluoreto ligado na posição meta, ou 4'-Me para se referir ao grupo metila ligado na posição para.

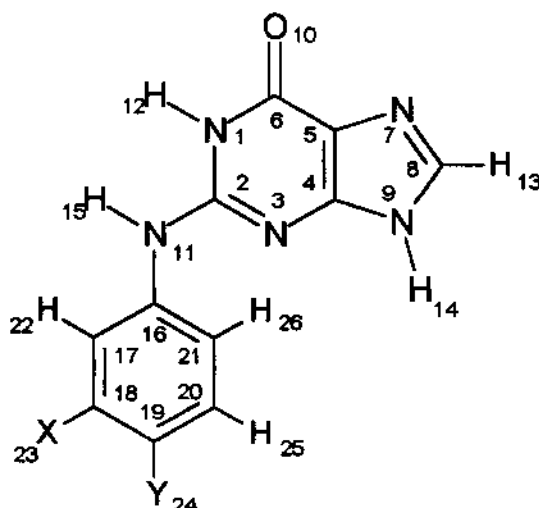


Figura 61 - Numeração dos átomos na molécula da N^2 -fenilguanina.

Os valores das diversas propriedades calculadas no presente trabalho são mostradas na Tabela 49 -Tabela 52, no final deste capítulo.

6.1.3 Correlação entre os parâmetros calculados

Não será apresentada a matriz de correlação entre todos os parâmetros utilizados na presente análise devido às dimensões da mesma (118×118). No entanto, à medida em que

as equações de regressão forem sendo apresentadas, mostrar-se-ão as matrizes de correlação contendo os parâmetros incluídos nessas equações.

6.2 ANÁLISE DOS COMPOSTOS META-SUBSTITUÍDOS

6.2.1 Atividade inibitória sobre HSV1 TK

Devido ao pequeno número de derivados meta-substituídos (treze, incluindo-se o derivado 3'-H, 4'-H), apenas as combinações de dois parâmetros podem ser testadas com segurança, na construção de equações de regressão linear múltipla (Topliss & Costello, 1972; Kubinyi, 1993). Foram muitas as combinações de dois parâmetros que resultaram em equações de regressão estatisticamente significativas para representar a atividade dos derivados meta-substituídos sobre HSV1 TK. A Tabela 32 mostra as combinações de dois parâmetros que resultaram em coeficientes de correlação iguais ou superiores a 0,93 (Equações (112)-(121)).

Tabela 32 - Combinações de duas variáveis, utilizadas para representar a variância da atividade inibitória dos meta-derivados da N^2 -fenilguanina sobre HSV1 TK, que resultaram em coeficientes de correlação (r) iguais ou superiores a 0,9. Os sinais dos coeficientes das variáveis presentes nas Equações (112)-(121) são mostrados na coluna *Variáveis*. Em negrito são destacadas as principais combinações de parâmetros.

Variáveis	r	s	F	Equação
$+ \pi_m + F_1^{(e)}$	0,949	0,215	45,751	112
$+ \pi_m + F_2^{(e)}$	0,953	0,209	49,057	113
$+ \pi_m - F_3^{(e)}$	0,943	0,229	39,859	114
$+ \pi_m + F_4^{(e)}$	0,942	0,229	39,713	115
$+ \pi_m + F_7^{(e)}$	0,942	0,231	39,174	116
$+ \pi_m - F_9^{(e)}$	0,933	0,247	33,438	117
$+ \pi_m + F_{10}^{(e)}$	0,940	0,235	37,705	118
$+ \pi_m - F_{11}^{(e)}$	0,939	0,236	37,305	119
$+ \pi_m - F_{22}^{(e)}$	0,932	0,248	33,204	120
$+ \pi_m - F_{26}^{(e)}$	0,941	0,231	38,914	121

Todas as combinações de variáveis apresentadas na Tabela 32 têm em comum o parâmetro π_m , que se revelou essencial à explicação da variância dos valores da atividade inibitória dos derivados meta-substituídos sobre HSV1 TK. O motivo do surgimento da grande quantidade de combinações de parâmetros estatisticamente significativas (o número de combinações de dois parâmetros com $r \geq 0.9$ foi de 28) é sem dúvida a forte correlação existente entre os índices de fronteira atômicos. Esta característica é inconveniente e a interpretação dos resultados envolvendo os índices de fronteira deve ser feita com cautela.

As três melhores equações da Tabela 32 são apresentadas explicitamente a seguir.

$$\log(1/IC_{50}) = 0,498 (\pm 0,164) \pi_m + 8,754 (\pm 3,335) F_1^{(e)} + 1,829 (\pm 1,359) \quad (112)$$

$$(n = 13; r = 0,949; s = 0,215; F_{(2,10)} = 45,751; p < 0,00001)$$

$$(Q^2 = 0,788; S_{PRESS} = 0,316; SDEP = 0,288)$$

$$\log(1/IC_{50}) = 0,486 (\pm 0,160) \pi_m + 8,201 (\pm 3,000) F_2^{(e)} - 6,180 (\pm 4,230) \quad (113)$$

$$(n = 13; r = 0,953; s = 0,209; F_{(2,10)} = 49,057; p < 0,00001)$$

$$(Q^2 = 0,805; S_{PRESS} = 0,303; SDEP = 0,276)$$

$$\log(1/IC_{50}) = 0,515 (\pm 0,174) \pi_m - 4,531 (\pm 1,872) F_3^{(e)} + 19,205 (\pm 5,717) \quad (114)$$

$$(n = 13; r = 0,943; s = 0,229; F_{(2,10)} = 39,859; p < 0,00002)$$

$$(Q^2 = 0,767; S_{PRESS} = 0,331; SDEP = 0,302)$$

Nota-se que as Equações (112)–(114) não somente apresentam boa qualidade de ajuste (r e F elevados e s e p baixos), como também boa capacidade preditiva (Q^2 elevado e S_{PRESS} e $SDEP$ baixos).

As Equações (112)–(114) indicam que a atividade inibitória dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV1 TK depende basicamente dos parâmetros lipofílico e eletrônico. A influência da lipofilicidade sobre a atividade dos compostos meta-substituídos não é surpresa. As Equações (101) (pag. 133) e (102) (pag. 135)

acusaram essa influência, pelo menos no caso de HSV2 TK. Mesmo assim, a correlação obtida na Equação (101) não foi boa, provavelmente por causa da parametrização deficiente do efeito eletrônico do substituinte. Os parâmetros clássicos (σ , R e F) não foram capazes de detectar a influência do efeito eletrônico na atividade desses compostos sobre HSV1 TK (ver Seção 4.4.1, “Atividade inibitória sobre HSV1 TK”, pag. 133). É possível que a atividade dos derivados meta-substituídos das N^2 -fenilguaninas dependa de efeitos eletrônicos em posições específicas da molécula, como indicam as Equações (112)-(121) (Tabela 32, pag. 228). Tais efeitos não são facilmente detectados por constantes de substituintes.

A presença de grupos polares na molécula de inibidores enzimáticos é sempre importante, por causa da possibilidade da interação com aminoácidos polares do sítio ativo. Sob este aspecto, a Equação (118) parece ser a mais interessante, pois indica que a atividade dos derivados meta depende da densidade eletrofílica de fronteira sobre o átomo de oxigênio da carbonila do anel guanina. A densidade eletrofílica de fronteira é calculada a partir dos coeficientes do HOMO da função de onda molecular. Os valores de $F_n^{(e)}$ para determinado átomo indicam a possibilidade deste átomo interagir com agente eletrófilo, ou seja, alguma espécie carregada positivamente. Pelo fato de os valores de $F_n^{(e)}$ serem positivos (ver Tabela 51, pag. 260), o sinal positivo do coeficiente de $F_n^{(e)}$ na Equação (118) indica que a interação do átomo O(10) com alguma espécie carregada positivamente é importante e capaz de intensificar a atividade dos derivados meta sobre HSV1 TK. A seleção de $F_{10}^{(e)}$ como parâmetro relevante na explicação da atividade sobre HSV1 TK era esperada, pois o experimento computacional realizado na determinação do modo de ligação das N^2 -fenilguaninas (ver Seção 5.5, “Modo de ligação das N^2 -fenilguaninas”, pag. 196) baseou-se na hipótese da formação de ponte de hidrogênio dupla entre os átomos O(10) e H(12) do ligante e Gln 125 (ver Figura 39, pag. 166). No entanto, também era

esperada a seleção do parâmetro $F_{12}^{(e)}$, ou $F_{12}^{(n)}$, para assinalar a existência da ponte de hidrogênio dupla. Isto não aconteceu. Os valores de $F_{12}^{(e)}$ (ver Tabela 51, pag. 260) e $F_{12}^{(n)}$ (ver Tabela 52, pag. 263) são muito pequenos. Isto é o resultado da baixa densidade eletrônica de HOMO e LUMO sobre o átomo H(12) nas N^2 -fenilguaninas (ver Figura 62).

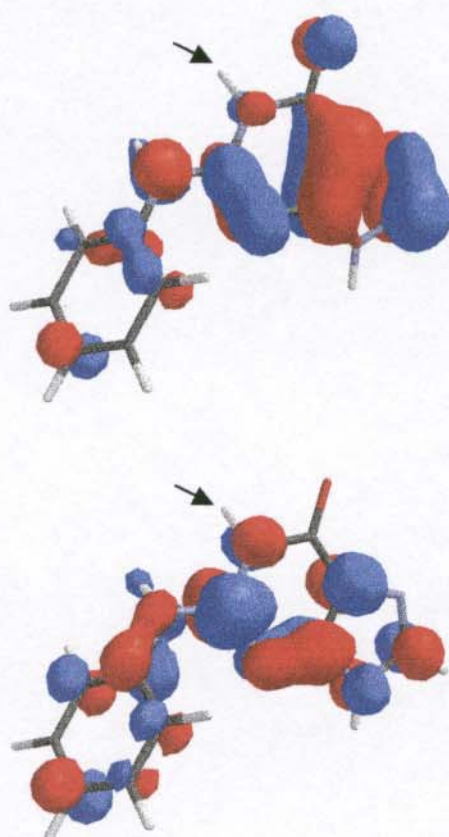


Figura 62 - Densidade eletrônica do HOMO (modelo de cima) e LUMO do composto **15** (3'-H, 4'-H). Esses orbitais não apresentam densidade eletrônica sobre o átomo H(12) (seta).

O resultado da parte qualitativa do experimento computacional apresentado no Quadro 4 (pag 211) indicou que apenas a ponte de hidrogênio entre O(10) e Gln 125 foi conservada após a simulação de dinâmica molecular, na maioria dos casos estudados. A Figura 63 mostra um típico exemplo de situação pós-simulação de dinâmica molecular, do composto **15** (3'-H, 4'-H) no sítio ativo B de HSV1 TK, que foi iniciada no modo de ligação do tipo 2'-deoxitimidina.

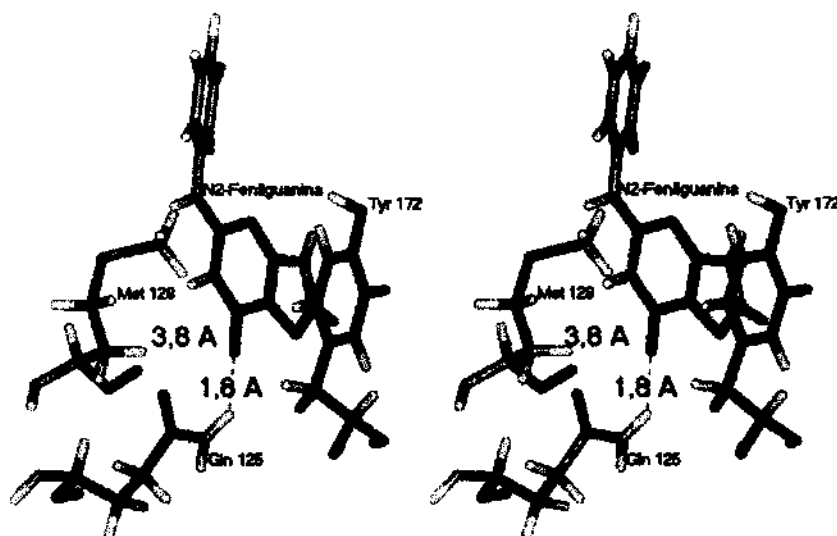


Figura 63 - Estereograma da situação final típica do modo de ligação das N^2 -fenilguaninas após a simulação de dinâmica molecular executada na Seção 5.5.3 ("Detalhes do experimento", pag. 207). Neste exemplo, apenas uma ponte de hidrogênio foi mantida entre o composto 1 e Gln 125 (linha tracejada, distância de 1,8 Å). A outra ponte não pôde ser formada devido à longa distância entre o segundo par de átomos (3,8 Å).

Na Figura 63, apenas a ponte de hidrogênio entre O(10) e Gln 125 foi conservada após a simulação. A acomodação natural da molécula do ligante ao ambiente do sítio ativo fez com que uma das pontes de hidrogênio não fosse mantida. Esta observação está em acordo com o fato de nenhum dos índices atômicos de H(12) ter sido selecionado para compor as Equações (112)-(121).

A ordem de entrada dos parâmetros físico-químicos na construção das Equações (112)-(114) é mostrada na Tabela 33.

Tabela 33 - Percentagem de variância da atividade inibitória dos derivados meta-substituídos sobre HSV1 TK, explicada por cada parâmetro nas Equações (112)-(114).

Equação	Percentagem de variância explicada			
	π_m	$F_1^{(e)}$	$F_2^{(e)}$	$F_3^{(e)}$
112	56%	+34%	—	—
113	56%	—	+34%	—
114	56%	—	—	+33%

Os resultados da Tabela 33 mostram que a atividade dos derivados meta-substituídos sobre HSV1 TK é principalmente dominada pelo parâmetro lipofílico, que é capaz de explicar cerca de 56% da variância da atividade inibitória.

A Tabela 34 mostra a matriz de correlação dos parâmetros presentes nas Equações (112)-(121).

Tabela 34 - Matriz de correlação dos principais parâmetros apresentados na Tabela 32. Somente os valores relativos aos treze compostos incluídos na análise foram considerados. Valores em termos de r^2 .

	π_m	$F_1^{(e)}$	$F_2^{(e)}$	$F_3^{(e)}$	$F_4^{(e)}$	$F_7^{(e)}$	$F_9^{(e)}$	$F_{10}^{(e)}$	$F_{11}^{(e)}$	$F_{26}^{(e)}$
π_m	1	,016	,024	,008	,016	,015	,012	,027	,012	,018
$F_1^{(e)}$		1	,989	,974	,981	,934	,975	,969	,988	,962
$F_2^{(e)}$			1	,943	,977	,947	,950	,974	,971	,951
$F_3^{(e)}$				1	,922	,841	,970	,905	,958	,935
$F_4^{(e)}$					1	,976	,961	,988	,991	,956
$F_7^{(e)}$						1	,882	,954	,943	,926
$F_9^{(e)}$							1	,963	,986	,935
$F_{10}^{(e)}$								1	,979	,935
$F_{11}^{(e)}$									1	,965
$F_{26}^{(e)}$										1

A Tabela 35 mostra os valores da atividade inibitória previstos pela Equação (113), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (113) é mostrado na Figura 64.

Tabela 35 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV1 TK, e os respectivos resíduos, calculados através da Equação (113).

Comp.	Substituinte		log $1/C_{50}$		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	6,82	6,46	0,36
2	CH ₂ N ₃	H	6,10	6,22	-0,12
4	CH ₂ Br	H	6,00	5,93	0,07
5	I	H	5,96	6,11	-0,15
6	CH ₂ Cl	H	5,85	5,62	0,23
7	Cl	H	5,82	5,84	-0,02
8	Br	H	5,82	5,94	-0,12
9	Et	H	5,60	5,41	0,19
11	n-Pr	H	5,48	5,69	-0,21
12	CN	H	5,30	5,61	-0,31
15	H	H	5,10	5,08	0,02
20	CH ₂ OH	H	4,60	4,66	-0,06
21	CH ₂ NH ₂	H	4,52	4,41	0,11

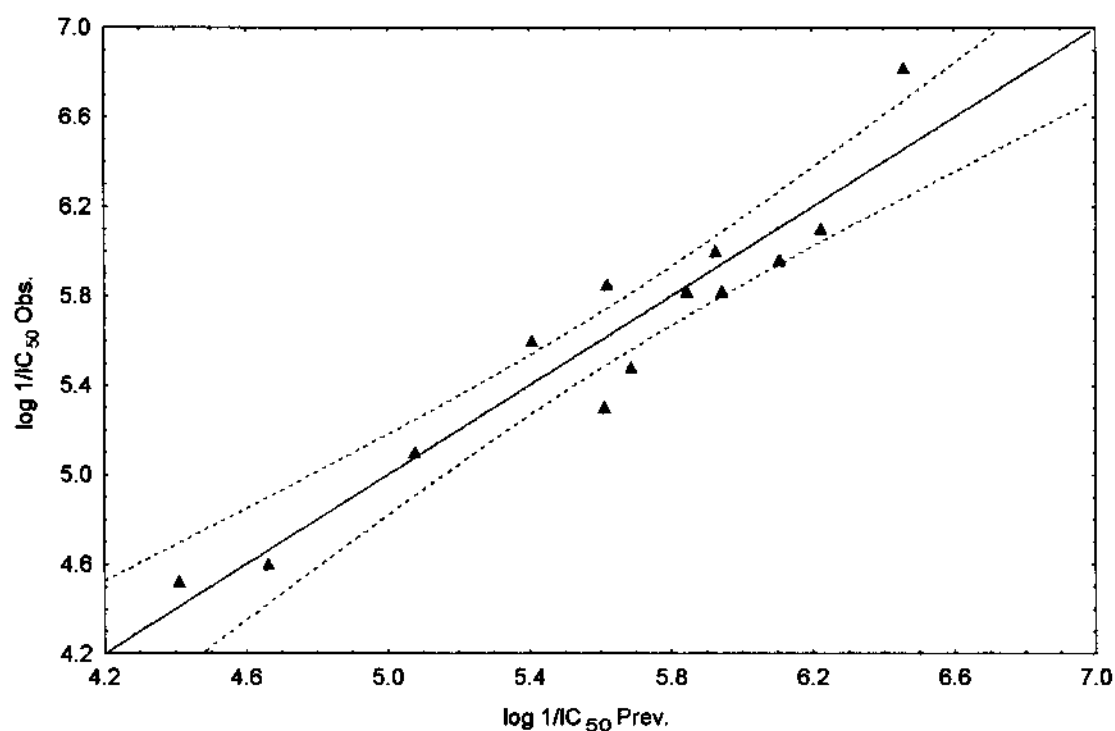


Figura 64 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (113).

6.2.2 Atividade inibitória sobre HSV2 TK.

Considerando-se modelos com apenas uma variável, obteve-se a Equação (122).

$$\log(1/IC_{50}) = - 0,168 (\pm 0,054) F_{16}^{(n)} + 1,611 (\pm 1,373) \quad (122)$$

$$(n = 13; r = 0,899; s = 0,298; F_{(1,11)} = 46,535; p < 0,00003)$$

$$(Q^2 = 0,737; s_{PRESS} = 0,349; SDEP = 0,334)$$

Isoladamente, $F_{16}^{(n)}$ é capaz de explicar cerca de 81% da variância da atividade dos compostos meta-substituídos sobre HSV2 TK. A interpretação da presença dessa variável na Equação (122) não é trivial (pode ser apenas coincidência). É difícil imaginar de que maneira o átomo C(16), isoladamente, poderia determinar a atividade dos compostos meta-substituídos sobre HSV2 TK. Por enquanto, o que se pode dizer apenas é que há indicação da influência de fator eletrônico na atividade desses compostos sobre HSV2 TK, fato que está em acordo com os resultados apresentados na Seção 4.4.2 (“Atividade inibitória sobre HSV2 TK”, pag. 133).

A Tabela 36 mostra os valores da atividade inibitória previstos pela Equação (122), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (122) é mostrado na Figura 65.

Tabela 36 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (122).

Comp.	Substituinte		Log $1/IC_{50}$		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	7,05	-0,05
2	CH ₂ N ₃	H	5,82	5,50	0,32
4	CH ₂ Br	H	5,92	5,79	0,13
5	I	H	5,98	6,22	-0,26
6	CH ₂ Cl	H	5,87	5,85	0,02
7	Cl	H	6,46	6,21	0,25
8	Br	H	6,22	6,23	-0,01
9	Et	H	5,40	5,29	0,11
11	n-Pr	H	5,52	5,28	0,24
12	CN	H	6,40	6,52	-0,12
15	H	H	5,80	5,43	0,37
20	CH ₂ OH	H	4,70	5,28	-0,58
21	CH ₂ NH ₂	H	4,70	5,10	-0,40

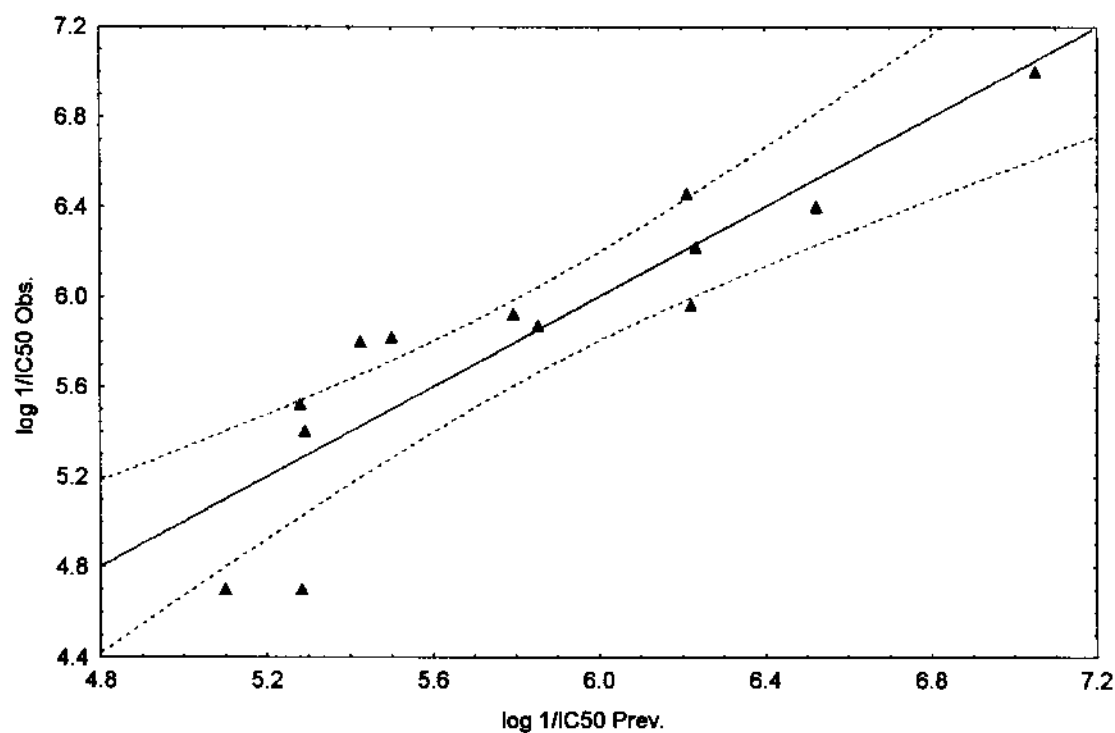


Figura 65 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (122).

Verificou-se a viabilidade de modelos com duas variáveis para representar a atividade dos compostos meta-substituídos sobre HSV2 TK. Apenas uma combinação de duas variáveis resultou em equação razoável. (Equação (123)).

$$\log(1/IC_{50}) = 0,224 (\pm 0,183) \pi_m - 0,157 (\pm 0,045) F_{16}^{(n)} + 1,805 (\pm 1,116) \quad (123)$$

$$(n = 13; r = 0,943; s = 0,237; F = 40,530; p < 0,00002)$$

$$(Q^2 = 0,824; s_{PRESS} = 0,300; SDEP = 0,274)$$

A Tabela 37 mostra os valores da atividade inibitória previstos pela Equação (123), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (123) é mostrado na Figura 66.

Tabela 37 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N²-fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (123).

Comp.	Substituinte		log 1/IC ₅₀		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	7,07	-0,07
2	CH ₂ N ₃	H	5,82	5,68	0,14
4	CH ₂ Br	H	5,92	5,88	0,04
5	I	H	5,96	6,35	-0,39
6	CH ₂ Cl	H	5,87	5,79	0,08
7	Cl	H	6,46	6,25	0,21
8	Br	H	6,22	6,30	-0,08
9	Et	H	5,40	5,46	-0,06
11	n-Pr	H	5,52	5,57	-0,05
12	CN	H	6,40	6,25	0,15
15	H	H	5,80	5,36	0,44
20	CH ₂ OH	H	4,70	4,99	-0,29
21	CH ₂ NH ₂	H	4,70	4,82	-0,12

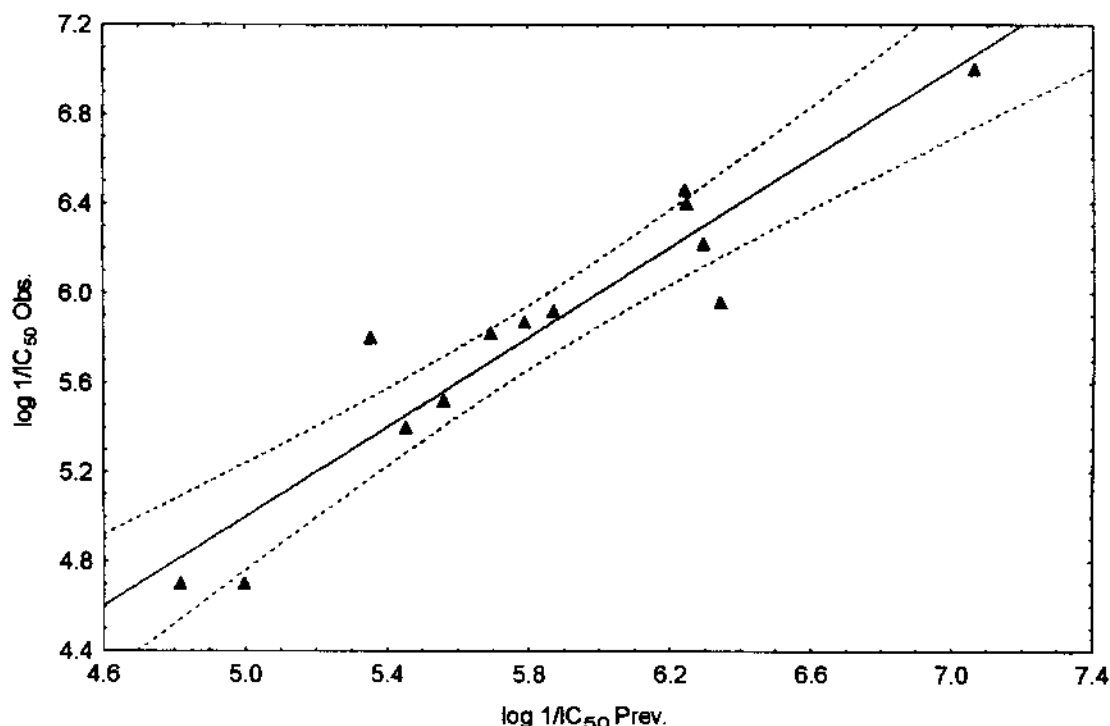


Figura 66 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (123).

A Equação (123), embora com elevado coeficiente de correlação e boa capacidade preditiva, têm um problema sério. O intervalo de confiabilidade de π_m é muito estreito (representa cerca de 82% do valor do próprio coeficiente de π_m), o que torna esta variável pouco significativa. Além disso, a observação do gráfico apresentado na Figura 66 mostra distribuição irregular dos valores de $\log 1/IC_{50}$, com aglomeração excessiva de pontos na região central do gráfico, dois pontos isolados no canto inferior esquerdo e apenas um ponto no canto superior direito. Considerou-se, portanto, a viabilidade de modelos com três variáveis, a despeito do risco de incorrer em correlação por coincidência.

A Tabela 38 mostra as combinações de três variáveis que produziram equações com coeficientes de correlação superiores a 0,97.

Tabela 38 - Combinações de três variáveis, utilizadas para representar a variância da atividade inibitória dos meta-derivados da N^2 -fenilguanina sobre HSV2 TK, que resultaram em coeficientes de correlação (r) superiores a 0,97. Os sinais dos coeficientes das variáveis presentes nas Equações (124)-(130) são mostrados na coluna *Variáveis*. Em negrito são destacadas as principais combinações.

Variáveis	R	s	F	Equação
+ π_m - RM_m + $F_1^{(e)}$	0,970	0,183	48,086	124
+ π_m - RM_m - $F_3^{(e)}$	0,974	0,172	54,737	125
+ π_m - B_{1m} - $F_{16}^{(n)}$	0,988	0,117	120,946	126
+ π_m - V_{vdW} - $F_3^{(e)}$	0,980	0,150	72,213	127
+ π_m - V_{vdW} - $F_{16}^{(n)}$	0,971	0,179	50,041	128
+ π_m - α - $F_{13}^{(n)}$	0,981	0,146	76,802	129
+ π_m - α - $F_{19}^{(n)}$	0,973	0,173	54,055	130

As combinações de variáveis apresentadas em negrito na Tabela 38 são mostradas explicitamente a seguir.

$$\log(1/IC_{50}) = 0,292 (\pm 0,096) \pi_m - 0,922 (\pm 0,370) B_{1m} - 0,205 (\pm 0,030) F_{16}^{(n)} + 2,067 (\pm 0,570) \quad (126)$$

$$(n = 13; r = 0,988; s = 0,117; F_{(3,9)} = 120,946; p < 0,000001)$$

$$(Q^2 = 0,952; s_{PRESS} = 0,166; SDEP = 0,143)$$

$$\log(1/IC_{50}) = 0,400 (\pm 0,121) \pi_m - 0,022 (\pm 0,008) V_{vdW} - 5,587 (\pm 1,278) F_3^{(e)} + 27,332 (\pm 3,916) \quad (127)$$

$$(n = 13; r = 0,980; s = 0,150; F_{(3,9)} = 72,213; p < 0,000001)$$

$$(Q^2 = 0,923; s_{PRESS} = 0,209; SDEP = 0,181)$$

$$\log(1/IC_{50}) = 0,545 (\pm 0,121) \pi_m - 0,276 (\pm 0,075) \alpha - 663,623 (\pm 124,871) F_{13}^{(n)} + 11,299 (\pm 1,739) \quad (129)$$

$$(n = 13; r = 0,981; s = 0,146; F_{(3,9)} = 76,802; p < 0,000001)$$

$$(Q^2 = 0,935; s_{PRESS} = 0,192; SDEP = 0,167)$$

Além do fato das Equações (126), (127) e (129) apresentarem elevada qualidade estatística e capacidade preditiva, o intervalo de confiabilidade de π_m aumentou sensivelmente em relação às Equações (123) (pag. 237), tornando π_m mais significativo. Além disso, houve sensível diminuição de s , s_{PRESS} e $SDEP$ e aumento significativo dos valores de F e Q^2 . Esses fatores são indicativos da real importância das Equações (126), (127) e (129).

Da mesma forma que para HSV1 TK, os parâmetros lipofílico e eletrônico são importantes para a atividade dos derivados meta-substituídos sobre HSV2 TK. No entanto, as Equações (124)-(130) indicam que a atividade dos derivados meta sobre HSV2 TK também sofre influência de fator estérico. Este já havia sido detectado anteriormente na Equação (103) (pag. 138), em que aparece o parâmetro RM_m . Observa-se que na Equação (126) o limite de confiabilidade do parâmetro B_{1m} representa apenas cerca de 40% do valor de seu coeficiente, que é razoável. Na Equação (103), o limite de confiabilidade do coeficiente de RM_m representa cerca de 90%. Os coeficientes de B_{1m} , V_{vdw} , α e RM_m são negativos, indicando a influência negativa do tamanho dos substituintes ligados na posição meta na atividade inibitória sobre HSV2 TK.

A comparação das Equações (124)-(130) com as Equações (112)-(121) (pag. 228) sugere a existência de importante diferença entre os sítios ativos de HSV1 TK e HSV2 TK, nas proximidades da posição meta do anel fenila. De acordo com essas equações, a atividade dos derivados meta-substituídos sobre HSV2 TK é influenciada por fator estérico, que está ausente em HSV1 TK (não há modelos de três variáveis estatisticamente significativos capazes de explicar a variância da atividade dos derivados meta-substituídos sobre HSV1 TK). O fator estérico pode ser resultante da presença de aminoácidos diferentes nos sítios ativos de HSV1 TK e HSV2 TK, em posições equivalentes próximas à posição meta do anel fenila. Para investigar essa possibilidade, fez-se a comparação dos

sítios ativos de HSV1 TK (Figura 67) e HSV2 TK (Figura 68). Utilizou-se, na comparação, o modelo de HSV2 TK, construído a partir de HSV1 TK, de acordo com a Seção 5.4.4 (“Modelagem das estruturas terciária e quaternária de HSV2 TK”, pag. 188). Nessa comparação foram considerados todos os resíduos afastados do ligante por distância menor ou igual a 10 Å. Resíduos localizados além desta distância não podem interagir de forma significativa com a molécula do ligante.

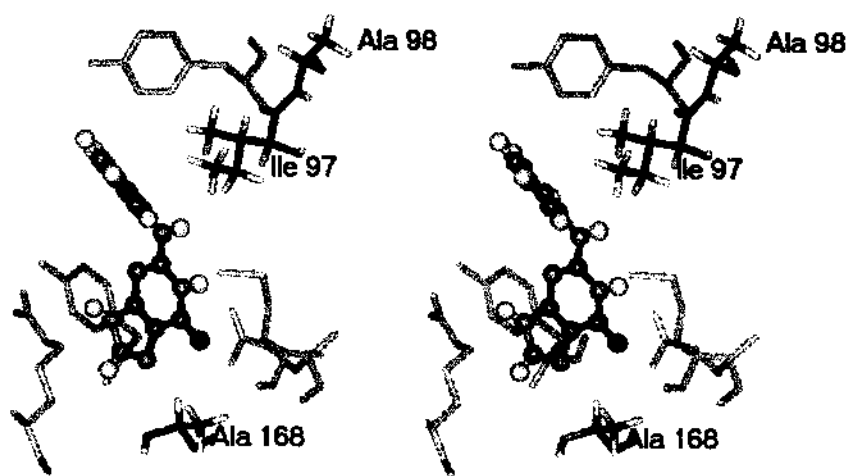


Figura 67 - Estereograma do sítio ativo de HSV1 TK, destacando os aminoácidos Ile 97, Ala 98 e Ala 168, que não ocorrem no sítio ativo de HSV2 TK.

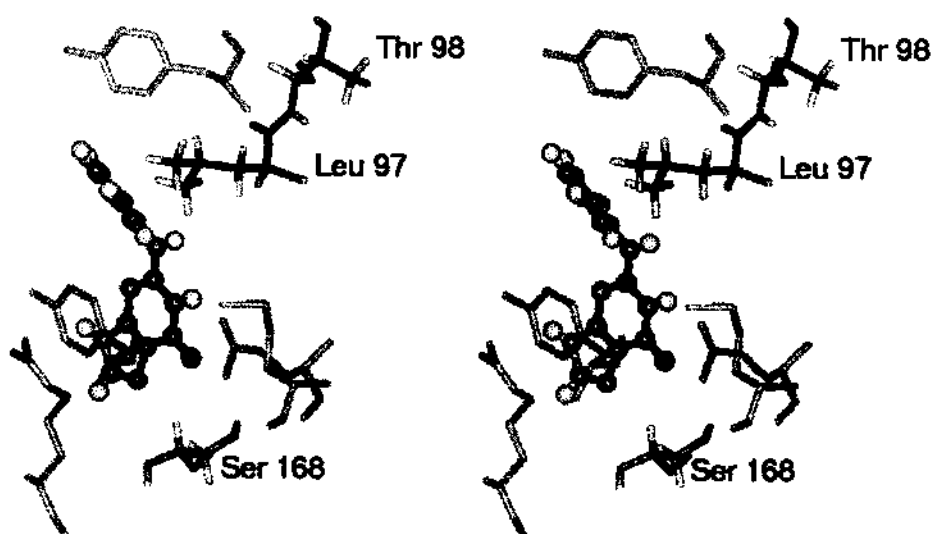


Figura 68 - Estereograma do sítio ativo de HSV2 TK, destacando os aminoácidos Leu 97, Thr 98 e Ser 168, que não ocorrem no sítio ativo de HSV2 TK.

A Figura 67 e a Figura 68 mostram que há apenas três aminoácidos diferentes nas proximidades do sítio ativo de HSV1 TK e HSV2 TK; Ile 97, Ala 98 e Ala 168 em HSV1 TK, que são substituídos por Leu 97, Thr 98 e Ser 168 em HSV2 TK. No entanto, apenas o resíduo 97 parece em condições de influenciar de maneira significativa a interação com o ligante meta-substituído. O motivo é que o resíduo 98 possui cadeia lateral projetada para longe do ligante, enquanto que o resíduo 168 está bastante afastado do anel fenila.

A presença de Leu no lugar de Ile em HSV2 TK é capaz de explicar a fonte do efeito estérico observado nas Equações (124)- (130) (pag. 239). Pelo fato da cadeia lateral de Ile possuir carbono terciário próximo ao esqueleto protéico, apenas um grupo metila fica próximo ao anel fenila (Figura 67). Por outro lado, o arranjo estrutural de Leu apresenta carbono terciário no final da cadeia lateral. Neste caso, há dois grupos metila próximos ao anel fenila, na região da posição meta (Figura 68). A presença de dois grupos metila nessa posição pode ser capaz de causar algum grau de impedimento estérico no caso de ligantes volumosos ligados na posição meta. Pode não haver espaço suficiente para acomodar de maneira confortável ligantes com substituintes grandes na posição meta.

Neste caso, toda a molécula do ligante pode ser afastada da posição ideal de interação, diminuindo a potência inibitória.

A Tabela 39 mostra a matriz de correlação das variáveis presentes nas Equações (124)-(130).

Tabela 39 - Matriz de correlação dos parâmetros apresentados nas Equações (124)-(130). Somente os valores relativos aos 13 compostos incluídos nessas equações foram considerados. Valores em termos de r^2 .

	π_m	RM_m	B_{1m}	V_{vdW}	α	$F_1^{(e)}$	$F_3^{(e)}$	$F_{13}^{(n)}$	$F_{16}^{(n)}$	$F_{19}^{(n)}$
π_m	1	,259	,122	,071	,136	,016	,008	,001	,045	,009
RM_m		1	,063	,652	,662	,015	,019	,020	,070	,002
B_{1m}			1	,027	,000	,301	,355	,080	,450	,125
V_{vdW}				1	,514	,040	,034	,034	,115	,034
α					1	,000	,002	,038	,072	,050
$F_1^{(e)}$						1	,974	,783	,824	,843
$F_3^{(e)}$							1	,743	,864	,780
$F_{13}^{(n)}$								1	,625	,951
$F_{16}^{(n)}$									1	,573
$F_{19}^{(n)}$										1

A ordem de entrada dos parâmetros na construção das Equações (126), (127) e (129) é mostrada na Tabela 40.

Tabela 40 - Percentagem de variância da atividade inibitória dos derivados meta-substituídos sobre HSV2 TK explicada por cada parâmetro nas Equações (126), (127) e (129).

Equação	Percentagem de variância explicada						
	$F_{16}^{(n)}$	$F_3^{(e)}$	$F_{13}^{(n)}$	π_m	B_{1m}	V_{vdW}	α
126	81%	—	—	+8%	+9%	—	—
127	—	64%	—	+16%	—	+16%	—
129	—	—	44%	+24%	—	—	+28%

Viu-se anteriormente na Tabela 33 (pag. 232) que a atividade dos compostos meta substituídos sobre HSV1 TK é mais intensamente influenciada pelo parâmetro lipofílico. De acordo com a Tabela 40, a atividade dos compostos meta-substituídos sobre HSV2 TK

é mais influenciada pelo parâmetro eletrônico. Esta conclusão está em acordo com as conclusões da análise clássica.

A Tabela 41 mostra os valores da atividade inibitória previstos pela Equação (126), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (126) é mostrado na Figura 69.

Tabela 41 - Atividades inibitórias observada e prevista dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (126).

Comp.	Substituinte		log 1/ IC_{50}		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	7,11	-0,11
2	CH ₂ N ₃	H	5,82	5,75	0,07
4	CH ₂ Br	H	5,92	5,99	-0,07
5	I	H	5,96	6,02	-0,06
6	CH ₂ Cl	H	5,87	5,88	-0,01
7	Cl	H	6,46	6,22	0,24
8	Br	H	6,22	6,15	0,07
9	Et	H	5,40	5,44	-0,04
11	n-Pr	H	5,52	5,58	-0,06
12	CN	H	6,40	6,40	0,00
15	H	H	5,80	5,79	0,01
20	CH ₂ OH	H	4,70	4,83	-0,13
21	CH ₂ NH ₂	H	4,70	4,61	0,09

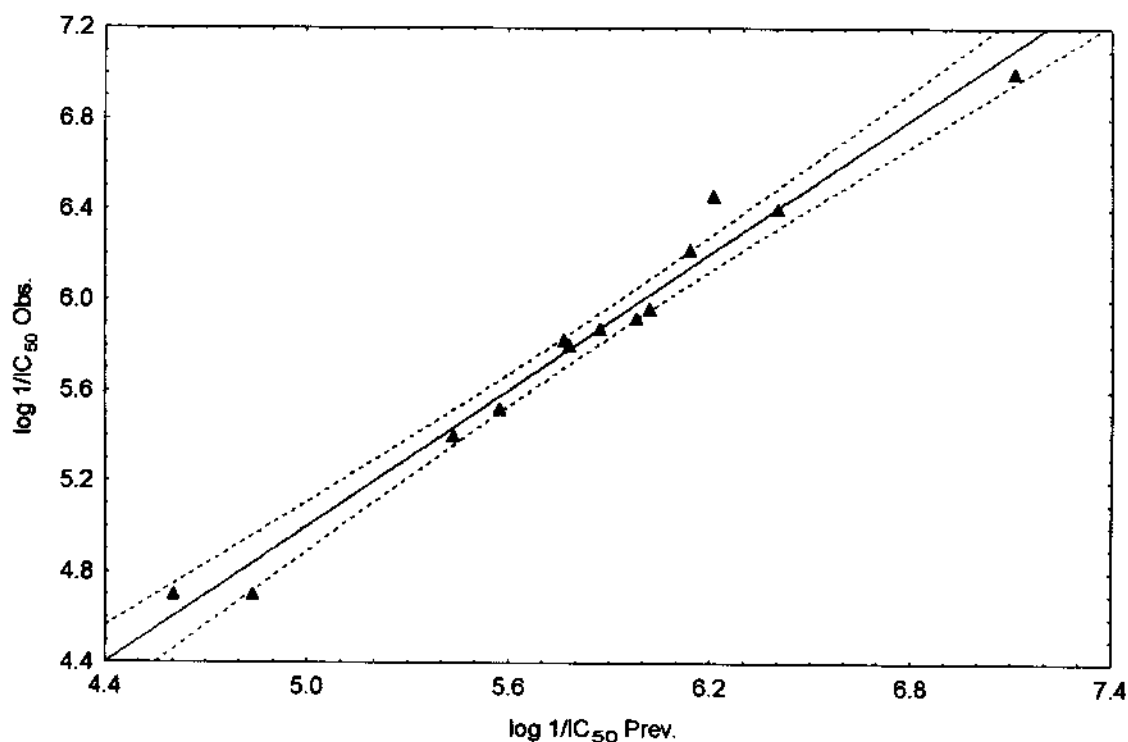


Figura 69 - Atividade inibitória observada dos derivados meta-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (126).

O gráfico da Figura 69 mostra a sensível melhora da linearidade do modelo representado pela Equação (126) quando comparado à Equação (123) (pag. 237), cujo gráfico é mostrado na Figura 66 (pag. 238).

6.3 ANÁLISE DOS COMPOSTOS PARA-SUBSTITUÍDOS

6.3.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

6.3.2 Atividade inibitória sobre HSV2 TK.

Não houve modelos significativos.

6.4 ANÁLISE DOS COMPOSTOS DI-SUBSTITUÍDOS

6.4.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

6.4.2 Atividade inibitória sobre HSV2 TK

Não houve modelos significativos.

6.5 ANÁLISE DOS COMPOSTOS META E PARA-SUBSTITUÍDOS

6.5.1 Atividade inibitória sobre HSV1 TK

Apenas dois modelos de quatro variáveis foram selecionados para representar a atividade dos derivados meta e para-substituídos sobre HSV1 TK (Equações (131) e (132)).

$$\begin{aligned} \log(1/IC_{50}) = & 0,573 (\pm 0,325) \pi_m - 0,416 (\pm 0,269) \pi_p + 2,236 (\pm 1,211) \sigma_m - \\ & - 2,788 (\pm 1,316) Q_{19} + 4,672 (\pm 0,296) \end{aligned} \quad (131)$$

(n = 25; r = 0,892; s = 0,469; F_(4,20) = 19,535; p < 0,000001)
(Q² = 0,690; s_{PRESS} = 0,577; SDEP = 0,527)

$$\begin{aligned} \log(1/IC_{50}) = & 0,499 (\pm 0,328) \pi_m - 0,373 (\pm 0,271) \pi_p + 2,450 (\pm 1,277) F_m - \\ & - 2,765 (\pm 1,293) Q_{19} + 4,600 (\pm 0,310) \end{aligned} \quad (132)$$

(n = 25; r = 0,896; s = 0,461; F = 20,359; p < 0,000001)
(Q² = 0,711; s_{PRESS} = 0,558; SDEP = 0,510)

De acordo com as Equações (131) e (132), a atividade dos compostos meta e para-substituídos depende da presença de substituintes lipofílicos em meta e hidrofílicos em para. Também depende da presença de substituintes em meta capazes de exercer efeito

indutivo positivo. Estas conclusões estão em acordo com os resultados do Capítulo 4. Parece haver consistência na presença de Q_{19} nas Equações (131) e (132). O átomo C(19) corresponde à posição para do anel fenila. A carga do átomo C(19) é negativa em 92% dos compostos meta-substituídos e 69% dos para-substituídos (ver Tabela 50, pag. 257). A variável Q_{19} possui coeficiente negativo nas Equações (131) e (132), o que indica que a atividade inibitória é intensificada à medida que a carga Q_{19} for mais negativa. O efeito indutivo positivo do substituinte em meta pode provocar o aumento da carga negativa em para, através de efeito orto-diretor.

A Tabela 42 mostra a matriz de correlação dos parâmetros das Equações (131) e (132).

Tabela 42 - Matriz de correlação dos parâmetros apresentados nas Equações (131) e (132). Somente os valores relativos aos 25 compostos incluídos nessas equações foram considerados. Valores em termos de r^2 .

	π_m	π_p	σ_m	F_m	Q_{19}
π_m	1	,052	,040	,099	,004
π_p		1	,102	,160	,014
σ_m			1	,941	,028
F_m				1	,023
Q_{19}					1

A ordem de entrada dos parâmetros na construção das Equações (131) e (132) é mostrada na Tabela 43.

Tabela 43 - Percentagem de variância da atividade inibitória dos derivados meta e para-substituídos sobre HSV1 TK explicada por cada parâmetro nas Equações (131) e (132).

Equação	Percentagem de variância explicada					
	π_m	F_m	σ_m	Q_{19}	π_m	π_p
131	33%	–	+19%	+17%	–	+11%
132	–	38%	–	+22%	+12%	+8%

A Tabela 44 mostra os valores da atividade inibitória previstos pela Equação (132), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (132) é mostrado na Figura 70.

Tabela 44 - Atividades inibitórias observada e prevista dos derivados meta e para-substituídos da N^2 -fenilguanina sobre HSV1 TK, e os respectivos resíduos, calculados através da Equação (132).

Comp.	Substituinte		log 1/ IC_{50}		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	6,82	6,09	0,73
2	CH ₂ N ₃	H	6,10	5,77	0,33
3	H	Br	6,00	5,08	0,92
4	CH ₂ Br	H	6,00	5,91	0,09
5	I	H	5,96	6,05	-0,09
6	CH ₂ Cl	H	5,85	5,47	0,38
7	Cl	H	5,82	6,09	-0,27
8	Br	H	5,82	6,00	-0,18
9	Et	H	5,60	5,60	0,00
11	n-Pr	H	5,48	5,89	-0,41
12	CN	H	5,30	5,62	-0,32
15	H	H	5,10	5,00	0,10
18	H	Et	4,70	4,25	0,45
20	CH ₂ OH	H	4,60	4,55	0,05
21	CH ₂ NH ₂	H	4,52	4,33	0,19
22	H	CF ₃	4,52	5,25	-0,73
23	H	n-Pr	4,52	4,06	0,46
26	H	I	4,40	4,78	-0,38
27	H	Me	4,30	4,40	-0,10
29	H	n-Bu	4,30	3,87	0,43
31	H	Cl	4,00	4,80	-0,80
32	H	OH	3,85	3,87	-0,02
33	H	i-Pr	3,70	3,97	-0,27
35	H	F	3,52	3,81	-0,29
36	H	n-C ₅ H ₁₁	3,30	3,58	-0,28

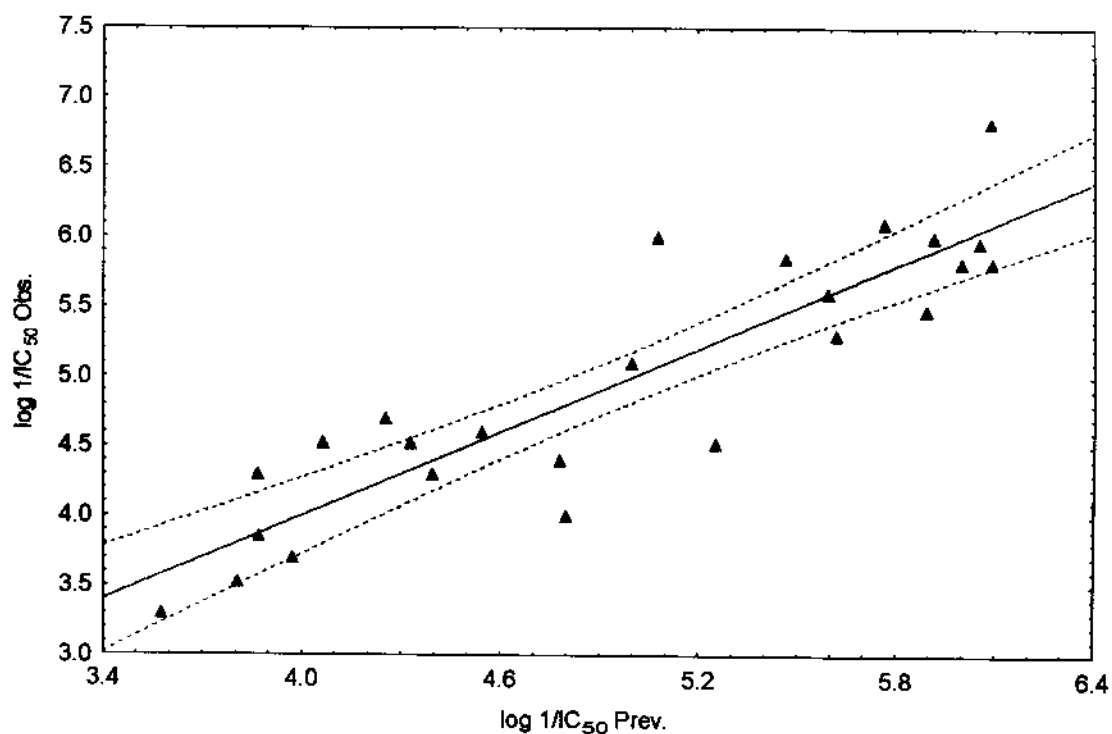


Figura 70 - Atividade inibitória observada dos derivados meta e para-substituídos da N^2 -fenilguaninas em função da atividade prevista pela Equação (132).

6.5.2 Atividade inibitória sobre HSV2 TK

A Tabela 45 mostra as combinações de quatro parâmetros utilizadas para representar a variância da atividade dos derivados meta e para sobre HSV2 TK e que resultaram em coeficientes de correlação iguais ou superiores a 0.9.

Tabela 45 - Combinações de quatro variáveis, utilizadas para representar a variância da atividade inibitória dos meta e para-derivados da N^2 -fenilguanina sobre HSV2 TK, que resultaram em coeficientes de correlação (r) superiores a 0,9. Em negrito são destacadas as principais combinações.

Variáveis	r	s	F	Equação
$+ \pi_m - \pi_p + \sigma_m + Q_{20}$	0,904	0,444	22,271	133
$+ \pi_m - \pi_p + \sigma_m - Q_{25}$	0,929	0,383	31,611	134
$+ \pi_m - \pi_p + F_m + Q_{20}$	0,901	0,450	21,553	135
$+ \pi_m - \pi_p + F_m - Q_{25}$	0,928	0,387	31,004	136
$+ \pi_m - RM_p + \sigma_m - Q_{25}$	0,921	0,404	27,933	137
$+ \pi_m - RM_p + F_m - Q_{25}$	0,919	0,408	27,251	138
$+ \pi_m + \sigma_m - Lp - Q_{25}$	0,927	0,389	30,536	139
$+ \pi_m + \sigma_m - B_{1p} - Q_{25}$	0,912	0,426	24,605	140

As equações destacadas em negrito na Tabela 45 são apresentadas explicitamente a seguir.

$$\log(1/IC_{50}) = 0,465 (\pm 0,265) \pi_m - 0,505 (\pm 0,222) \pi_p + 3,166 (\pm 0,997) \sigma_m - 25,753 (\pm 10,096) Q_{25} + 7,620 (\pm 1,017) \quad (134)$$

$$(n = 25; r = 0,929; s = 0,383; F_{(4,20)} = 31,611; p < 0,000001)$$

$$(Q^2 = 0,792; s_{PRESS} = 0,473; SDEP = 0,432)$$

$$\log(1/IC_{50}) = 0,366 (\pm 0,273) \pi_m - 0,452 (\pm 0,228) \pi_p + 3,407 (\pm 1,087) F_m - 26,535 (\pm 10,233) Q_{25} + 7,603 (\pm 1,025) \quad (136)$$

$$(n = 25; r = 0,928; s = 0,387; F_{(4,20)} = 31,004; p < 0,000001)$$

$$(Q^2 = 0,795; s_{PRESS} = 0,470; SDEP = 0,429)$$

A matriz de correlação dos parâmetros presentes nas Equações (133)-(140) são mostrados na Tabela 46.

Tabela 46 - Matriz de correlação dos parâmetros apresentados nas Equações (133)-(140). Somente os valores relativos aos 25 compostos incluídos nessas equações foram considerados.

	π_m	π_p	RM_p	σ_m	F_m	L_p	B_{1p}	Q_{20}	Q_{25}
π_m	1	,052	,060	,040	,099	,067	,091	,019	,027
π_p		1	,928	,102	,160	,912	,418	,084	,051
RM_p			1	,117	,184	,961	,441	,045	,048
σ_m				1	,941	,131	,179	,074	,068
F_m					1	,207	,282	,097	,091
L_p						1	,434	,039	,032
B_{1p}							1	,277	,189
Q_{20}								1	,870
Q_{25}									1

A estrutura das Equações (134) e (136) é muito parecida com a das Equações (131) e (132) (pag. 246), evidenciando a semelhança no mecanismo de ação dos derivados meta e para-substituídos sobre HSV1 e HSV2 TKs. A principal diferença entre esses dois grupos de equações está na ordem de explicação da variância da atividade inibitória, como mostram a Tabela 43 (pag. 247) e a Tabela 47. O estudo de QSAR do Capítulo 4 revelou que a atividade sobre HSV1 TK é principalmente governada pelo parâmetro lipofílico, enquanto que a atividade sobre HSV2 TK é principalmente governada pelo parâmetro eletrônico. Os resultados apresentados na Tabela 43 (pag. 247) e Tabela 47 confirmam essa tendência, exceto pela Equação (134), que mostra o inverso.

Tabela 47 - Percentagem de variância da atividade inibitória dos derivados meta e para-substituídos sobre HSV2 TK explicada por cada parâmetro nas Equações (134) e (136).

Equação	Percentagem de variância explicada					
	σ_m	F_m	Q_{25}	π_p	Q_{25}	π_m
134	45%	—	—	+15%	+17%	+9%
136	—	51%	+16%	+14%	—	+5%

A Tabela 48 mostra os valores da atividade inibitória previstos pela Equação (134), bem como os respectivos resíduos. O gráfico mostrando a atividade observada em função da atividade prevista pela Equação (134) é mostrado na Figura 71.

Tabela 48 - Atividades inibitórias observada e prevista dos derivados meta e para-substituídos da N^2 -fenilguanina sobre HSV2 TK, e os respectivos resíduos, calculados através da Equação (134).

Comp.	Substituinte		log 1/IC ₅₀		
	X	Y	Observado	Previsto	Resíduo
1	CF ₃	H	7,00	6,78	0,22
2	CH ₂ N ₃	H	5,82	5,52	0,30
3	H	Br	6,22	5,81	0,41
4	CH ₂ Br	H	5,92	6,01	-0,09
5	I	H	5,96	6,14	-0,18
6	CH ₂ Cl	H	5,87	5,45	0,42
7	Cl	H	6,46	6,36	0,10
8	Br	H	6,22	6,13	0,09
9	Et	H	5,40	5,43	-0,03
11	n-Pr	H	5,52	5,79	-0,27
12	CN	H	6,40	6,68	-0,28
15	H	H	5,80	5,22	0,58
18	H	Et	4,60	4,45	0,15
20	CH ₂ OH	H	4,70	4,60	0,10
21	CH ₂ NH ₂	H	4,70	4,62	0,08
22	H	CF ₃	4,40	4,54	-0,14
23	H	n-Pr	4,33	4,22	0,11
26	H	I	5,05	5,47	-0,42
27	H	Me	5,22	4,83	0,39
29	H	n-Bu	4,30	3,99	0,31
31	H	Cl	4,30	5,32	-1,02
32	H	OH	4,70	4,81	-0,11
33	H	i-Pr	4,00	4,10	-0,10
35	H	F	3,59	4,18	-0,59
36	H	n-C ₅ H ₁₁	3,60	3,63	-0,03

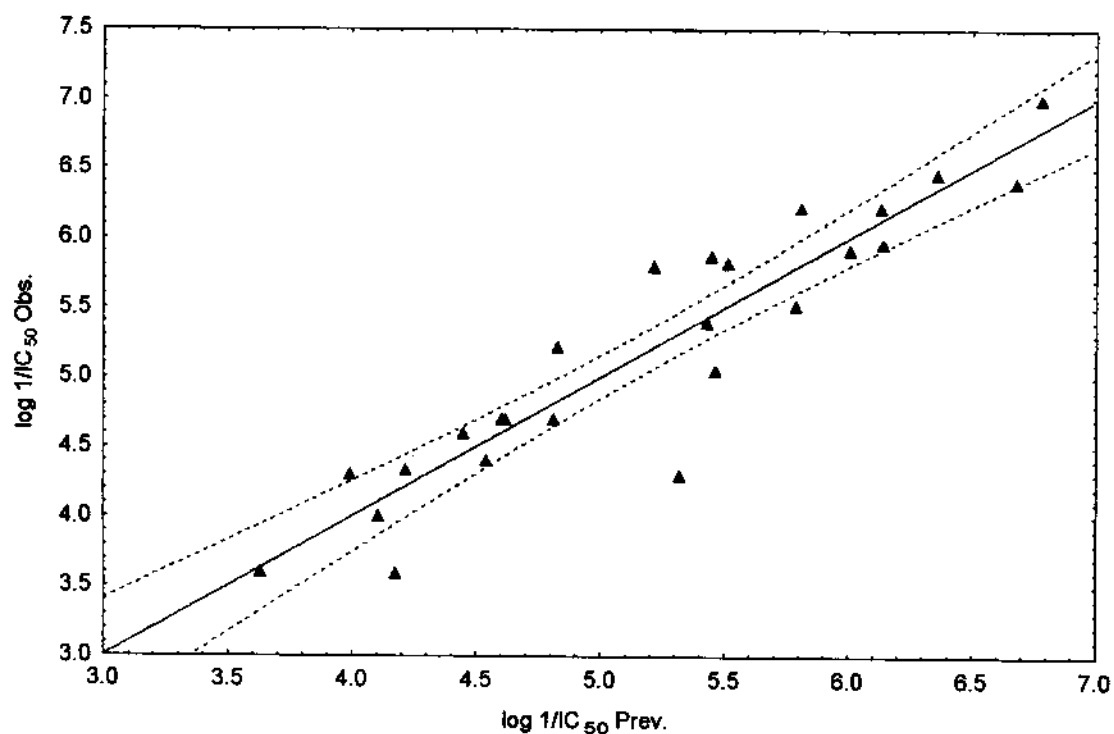


Figura 71 - Atividade inibitória observada dos derivados meta e para-substituídos da N²-fenilguaninas em função da atividade prevista pela Equação (134).

6.6 ANÁLISE GLOBAL

6.6.1 Atividade inibitória sobre HSV1 TK

Não houve modelos significativos.

6.6.2 Atividade inibitória sobre HSV2 TK

Não houve modelos significativos.

6.7 CONCLUSÕES DO ESTUDO QSAR COM PARÂMETROS CLÁSSICOS E CALCULADOS

- (a) A atividade inibitória dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV1 TK depende dos parâmetros lipofílico e eletrônico, sendo que, isoladamente, o parâmetro lipofílico é o mais importante;
- (b) A atividade dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK depende dos parâmetros lipofílico, eletrônico e estérico, sendo que, isoladamente, o parâmetro eletrônico é mais importante;
- (c) Há indicações de que o resíduo Leu 97 é o responsável pelo efeito estérico observado na atividade dos derivados meta-substituídos sobre HSV2 TK. Na posição 97, HSV1 TK apresenta o aminoácido Ile;
- (d) A densidade eletrofílica de fronteira do átomo O(10) pode ser importante para a atividade inibitória sobre HSV1 TK e HSV2 TK. No entanto, apenas no caso dos derivados meta-substituídos sobre HSV1 TK houve modelo que incluiu explicitamente essa variável.

A Figura 72 mostra, de forma resumida, as conclusões acerca da natureza do sítio ativo das N^2 -fenilguaninas de HSV1 TK e HSV2 TK.

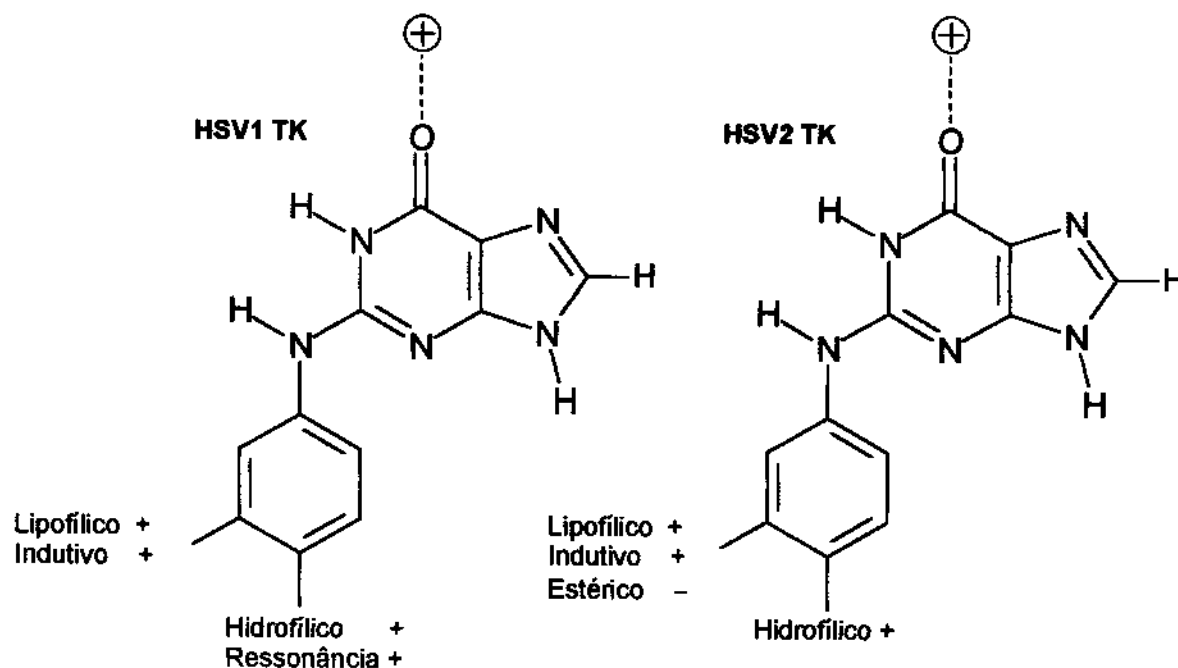


Figura 72 - Conclusões gerais do estudo QSAR. Os resultados indicaram que a interação entre o átomo O₁₀ e um átomo carregado positivamente no sítio ativo é importante para a atividade inibitória. Também houve indicação de que a diferença mais importante entre os sítios ativos das enzimas HSV1 TK e HSV2 TK é a ocorrência de efeito estérico sobre os derivados meta em HSV2 TK. Lipofílico+: substituintes mais lipofílicos aumentam a atividade; Hidrofílico+: substituintes mais hidrofílicos aumentam a atividade; Ressonância+: substituintes capazes de atrair elétrons por mecanismo de ressonância aumentam a atividade; Indutivo+: Substituintes capazes de atrair elétrons por mecanismo de indução aumentam a atividade.

6.8 LISTAGEM DOS PARÂMETROS CALCULADOS

Mostram-se na Tabela 49 - Tabela 52 os parâmetros calculados que, juntamente com os parâmetros clássicos apresentados na Tabela 19 (pag. 151), foram testados e/ou utilizados na construção das equações de regressão apresentadas ao longo do presente capítulo.

Tabela 49 - Energias dos orbitais moleculares ocupado de energia mais alta, ϵ_{HOMO} , e desocupado de energia mais baixa, ϵ_{LUMO} , dureza molecular, η , volume molecular de van der Waals, V_{vdW} , momento dipolar, μ , e polarizabilidade molecular, α .

No.	ϵ_{HOMO}^a	ϵ_{LUMO}^a	η^a	V_{vdW}^b	μ^c	α^b
1	-8,91	-0,78	4,07	210,50	4,03	22,99
2	-8,79	-0,65	4,07	221,30	4,98	27,07
3	-8,81	-0,68	4,06	204,60	4,32	23,07
4	-8,75	-0,56	4,09	223,10	5,48	24,14
5	-8,79	-0,61	4,09	214,90	4,99	23,43
6	-8,74	-0,55	4,10	216,80	5,55	23,62
7	-8,78	-0,60	4,09	200,00	4,96	22,44
8	-8,80	-0,62	4,09	206,60	4,95	22,89
9	-8,65	-0,44	4,10	217,40	6,29	23,96
10	-8,90	-0,83	4,04	225,50	3,39	24,70
11	-8,65	-0,44	4,10	232,90	6,32	25,28
12	-8,89	-0,76	4,07	200,10	4,08	23,60
13	-8,61	-0,42	4,09	231,20	6,64	25,55
14	-8,86	-0,77	4,05	201,20	3,39	22,90
15	-8,68	-0,47	4,10	185,00	6,05	21,21
16	-8,74	-0,58	4,08	214,00	5,41	24,08
17	-8,60	-0,42	4,09	221,30	6,71	25,34
18	-8,64	-0,45	4,09	216,30	6,38	24,17
19	-8,99	-0,98	4,00	226,30	3,01	24,38
20	-8,74	-0,54	4,10	225,50	4,65	23,19
21	-8,67	-0,46	4,11	212,50	5,61	23,61
22	-9,02	-0,90	4,06	208,50	2,85	23,06
23	-8,64	-0,45	4,09	231,80	6,42	25,48
24	-8,66	-0,68	3,99	223,40	6,05	28,26
25	-8,62	-0,43	4,09	266,40	6,67	28,14
26	-8,82	-0,68	4,07	214,20	4,39	23,60
27	-8,63	-0,45	4,09	199,00	6,39	22,90
28	-8,90	-0,81	4,04	189,20	3,07	22,00
29	-8,64	-0,45	4,09	248,20	6,42	26,74
30	-8,87	-0,76	4,05	212,90	3,63	23,78
31	-8,77	-0,63	4,07	198,10	4,52	22,58
32	-8,64	-0,46	4,09	190,20	5,66	22,20
33	-8,64	-0,45	4,10	231,60	6,42	25,35
34	-8,70	-0,52	4,09	222,00	4,88	24,75
35	-8,77	-0,63	4,07	186,20	4,33	21,66
36	-8,64	-0,45	4,09	268,50	6,44	27,98

^a Em elétron-Volt; ^b em Å³; ^c em Debye.

Tabela 50 - Cargas atômicas (Q_n), calculadas a partir do potencial eletrostático molecular, sobre o n -ésimo átomo da molécula da N^2 -fenilguanina.

No.	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8
1	-0,7190	0,8816	-0,6863	0,4163	-0,1597	0,7204	-0,4230	0,2023
2	-0,6862	0,8461	-0,6781	0,4160	-0,1497	0,6993	-0,4256	0,2044
3	-0,7068	0,8613	-0,6651	0,3984	-0,1587	0,7155	-0,4215	0,2068
4	-0,7180	0,8742	-0,6682	0,3961	-0,1554	0,7191	-0,4196	0,1893
5	-0,6911	0,8425	-0,6493	0,3797	-0,1448	0,7066	-0,4267	0,2107
6	-0,6961	0,8529	-0,6553	0,3911	-0,1452	0,7013	-0,4306	0,2267
7	-0,6932	0,8508	-0,6564	0,3856	-0,1453	0,7040	-0,4266	0,2090
8	-0,7051	0,8665	-0,6690	0,4052	-0,1526	0,7084	-0,4257	0,2015
9	-0,6852	0,8397	-0,6505	0,3807	-0,1414	0,6984	-0,4261	0,1982
10	-0,7191	0,8768	-0,6761	0,4100	-0,1557	0,7200	-0,4259	0,2089
11	-0,6861	0,8456	-0,6580	0,3860	-0,1323	0,6855	-0,4234	0,1873
12	-0,7069	0,8680	-0,6713	0,4019	-0,1573	0,7151	-0,4210	0,2019
13	-0,6923	0,8447	-0,6623	0,4003	-0,1520	0,7036	-0,4213	0,1898
14	-0,6858	0,8356	-0,6464	0,3781	-0,1364	0,6976	-0,4292	0,2189
15	-0,7012	0,8612	-0,6651	0,4001	-0,1576	0,7102	-0,4239	0,1982
16	-0,6966	0,8548	-0,6651	0,4018	-0,1558	0,7059	-0,4215	0,2071
17	-0,6922	0,8483	-0,6637	0,3947	-0,1500	0,7044	-0,4244	0,1947
18	-0,6960	0,8525	-0,6688	0,4041	-0,1580	0,7074	-0,4237	0,2075
19	-0,7042	0,8624	-0,6715	0,4004	-0,1456	0,7078	-0,4244	0,2084
20	-0,6945	0,8609	-0,6688	0,3943	-0,1440	0,6965	-0,4241	0,2003
21	-0,6962	0,8697	-0,6683	0,3857	-0,1321	0,6879	-0,4229	0,1851
22	-0,7165	0,8673	-0,6619	0,3960	-0,1513	0,7208	-0,4270	0,2221
23	-0,6954	0,8539	-0,6668	0,4018	-0,1540	0,7055	-0,4263	0,2108
24	-0,7010	0,8523	-0,6599	0,3971	-0,1551	0,7122	-0,4272	0,2187
25	-0,6897	0,8473	-0,6603	0,3886	-0,1401	0,6957	-0,4263	0,1986
26	-0,7005	0,8538	-0,6580	0,3932	-0,1500	0,7108	-0,4284	0,2254
27	-0,7017	0,8578	-0,6699	0,4066	-0,1534	0,7079	-0,4223	0,1880
28	-0,7039	0,8395	-0,6504	0,3929	-0,1537	0,7163	-0,4263	0,2234
29	-0,7043	0,8636	-0,6700	0,4000	-0,1523	0,7085	-0,4246	0,1993
30	-0,7040	0,8635	-0,6729	0,4077	-0,1542	0,7098	-0,4282	0,2245
31	-0,6993	0,8600	-0,6684	0,3969	-0,1469	0,7030	-0,4244	0,2022
32	-0,6911	0,8479	-0,6653	0,3959	-0,1487	0,6995	-0,4272	0,2128
33	-0,6956	0,8531	-0,6644	0,3981	-0,1515	0,7048	-0,4269	0,2102
34	-0,7034	0,8578	-0,6629	0,3953	-0,1587	0,7124	-0,4206	0,2021
35	-0,6961	0,8418	-0,6543	0,3884	-0,1459	0,7065	-0,4222	0,1952
36	-0,7120	0,8687	-0,6714	0,4024	-0,1659	0,7224	-0,4186	0,1893

Tabela 50 - Cargas atômicas (continuação).

No.	Q ₉	Q ₁₀	Q ₁₁	Q ₁₂	Q ₁₃	Q ₁₄	Q ₁₅	Q ₁₆
1	-0,3484	-0,4906	-0,8003	0,3576	0,1221	0,3084	0,3718	0,3740
2	-0,3569	-0,4903	-0,7818	0,3502	0,1211	0,3102	0,3656	0,3767
3	-0,3460	-0,4933	-0,7954	0,3553	0,1180	0,3082	0,3680	0,4729
4	-0,3322	-0,4966	-0,7772	0,3559	0,1230	0,3035	0,3625	0,3736
5	-0,3406	-0,4932	-0,7586	0,3487	0,1153	0,3045	0,3648	0,2781
6	-0,3580	-0,4926	-0,7772	0,3526	0,1096	0,3114	0,3605	0,3712
7	-0,3434	-0,4913	-0,7788	0,3507	0,1172	0,3064	0,3682	0,3451
8	-0,3470	-0,4907	-0,7747	0,3508	0,1209	0,3066	0,3713	0,2917
9	-0,3361	-0,4951	-0,7611	0,3480	0,1185	0,3029	0,3578	0,3529
10	-0,3508	-0,4896	-0,8042	0,3570	0,1211	0,3088	0,3781	0,4200
11	-0,3348	-0,4921	-0,7582	0,3499	0,1221	0,3033	0,3557	0,3461
12	-0,3416	-0,4914	-0,7999	0,3552	0,1205	0,3054	0,3702	0,3710
13	-0,3405	-0,4965	-0,7563	0,3494	0,1213	0,3039	0,3550	0,3416
14	-0,3474	-0,4880	-0,7636	0,3493	0,1147	0,3083	0,3677	0,3063
15	-0,3416	-0,4950	-0,7791	0,3502	0,1198	0,3039	0,3645	0,3695
16	-0,3513	-0,4922	-0,7771	0,3519	0,1170	0,3088	0,3666	0,3410
17	-0,3397	-0,4965	-0,7569	0,3474	0,1199	0,3042	0,3583	0,3415
18	-0,3478	-0,4961	-0,7699	0,3500	0,1142	0,3059	0,3603	0,3604
19	-0,3484	-0,4861	-0,8035	0,3563	0,1213	0,3096	0,3748	0,4151
20	-0,3432	-0,4912	-0,7927	0,3512	0,1183	0,3064	0,3666	0,4148
21	-0,3318	-0,4925	-0,7997	0,3522	0,1224	0,3026	0,3645	0,4164
22	-0,3503	-0,4893	-0,8046	0,3611	0,1162	0,3107	0,3730	0,4803
23	-0,3549	-0,4957	-0,7715	0,3497	0,1146	0,3088	0,3613	0,3567
24	-0,3545	-0,4955	-0,7734	0,3522	0,1122	0,3098	0,3632	0,3801
25	-0,3398	-0,4952	-0,7596	0,3498	0,1172	0,3043	0,3536	0,3751
26	-0,3534	-0,4928	-0,7945	0,3536	0,1107	0,3087	0,3696	0,4730
27	-0,3417	-0,4962	-0,7667	0,3509	0,1231	0,3047	0,3593	0,3521
28	-0,3524	-0,4904	-0,7455	0,3531	0,1130	0,3094	0,3617	0,3503
29	-0,3431	-0,4964	-0,7689	0,3514	0,1179	0,3047	0,3593	0,3591
30	-0,3606	-0,4883	-0,7980	0,3549	0,1147	0,3124	0,3739	0,3916
31	-0,3447	-0,4908	-0,7915	0,3535	0,1199	0,3078	0,3645	0,4340
32	-0,3558	-0,4941	-0,7544	0,3492	0,1140	0,3093	0,3553	0,2734
33	-0,3540	-0,4957	-0,7637	0,3499	0,1148	0,3093	0,3588	0,3271
34	-0,3414	-0,4962	-0,7730	0,3530	0,1165	0,3049	0,3587	0,3412
35	-0,3370	-0,4930	-0,7604	0,3517	0,1217	0,3045	0,3610	0,3276
36	-0,3354	-0,5002	-0,7691	0,3536	0,1214	0,3024	0,3576	0,3530

Tabela 50 - Cargas atômicas (continuação).

No.	Q ₁₇	Q ₁₈	Q ₁₉	Q ₂₀	Q ₂₁	Q ₂₂	Q ₂₅	Q ₂₆
1	-0,1121	-0,2823	-0,0422	-0,0350	-0,2191	0,1250	0,1015	0,1582
2	-0,1462	-0,1612	-0,0865	-0,0348	-0,2344	0,1334	0,1019	0,1627
3	-0,3058	0,1110	-0,2897	0,2010	-0,3875	0,1406	0,0534	0,1844
4	-0,2298	0,1025	-0,2082	0,0192	-0,2810	0,1343	0,0916	0,1654
5	-0,0395	-0,1710	0,0491	-0,1323	-0,1842	0,0941	0,1206	0,1582
6	-0,2281	0,0790	-0,1680	-0,0201	-0,2608	0,1447	0,1008	0,1638
7	-0,1052	-0,1079	-0,0384	-0,0604	-0,2398	0,1077	0,1073	0,1622
8	-0,0230	-0,2049	0,0485	-0,1143	-0,1904	0,0858	0,1213	0,1516
9	-0,1970	-0,0131	-0,1767	-0,0073	-0,2705	0,1223	0,0948	0,1617
10	-0,1502	-0,0066	-0,1768	0,1390	-0,3271	0,1072	0,0685	0,1774
11	-0,1964	0,0146	-0,1794	-0,0007	-0,2789	0,1262	0,0916	0,1661
12	-0,0862	-0,3150	-0,0206	-0,0330	-0,2242	0,1115	0,0952	0,1582
13	-0,1896	-0,0635	-0,0025	-0,0643	-0,2648	0,1293	0,1112	0,1642
14	-0,0631	-0,2336	0,3495	-0,1701	-0,2224	0,1100	0,1426	0,1749
15	-0,1871	-0,0889	-0,1445	-0,0121	-0,2721	0,1149	0,0932	0,1634
16	-0,0921	-0,1584	0,0906	-0,0822	-0,2559	0,1067	0,1098	0,1707
17	-0,1854	-0,0380	-0,0580	-0,0259	-0,2787	0,1214	0,0979	0,1670
18	-0,1651	-0,1614	-0,0122	-0,0377	-0,2799	0,1163	0,1031	0,1671
19	-0,1627	-0,1349	-0,1299	0,0954	-0,2817	0,1342	0,0776	0,1739
20	-0,2413	-0,0455	-0,1400	-0,0156	-0,2925	0,1217	0,0985	0,1709
21	-0,2105	-0,1915	-0,0542	-0,0202	-0,2972	0,1165	0,0941	0,1705
22	-0,2420	0,0224	-0,3552	0,0803	-0,3087	0,1262	0,1024	0,1641
23	-0,1611	-0,1590	-0,0153	-0,0393	-0,2778	0,1144	0,1018	0,1683
24	-0,2457	0,0290	0,0037	-0,0489	-0,2802	0,1400	0,1080	0,1709
25	-0,2237	0,0076	-0,0961	0,0025	-0,3137	0,1325	0,0980	0,1722
26	-0,3062	0,0712	-0,2170	0,1613	-0,3900	0,1431	0,0616	0,1878
27	-0,1577	-0,1591	-0,0014	-0,0333	-0,2845	0,1125	0,0974	0,1687
28	-0,2807	0,1839	0,1523	-0,1156	-0,2672	0,1670	0,1425	0,1777
29	-0,1691	-0,1399	-0,0221	-0,0253	-0,2906	0,1146	0,0992	0,1700
30	-0,1360	-0,0428	-0,0885	0,0632	-0,2927	0,1132	0,0846	0,1741
31	-0,2457	0,0094	-0,1680	0,1079	-0,3458	0,1330	0,0753	0,1825
32	-0,0903	-0,2994	0,3541	-0,2075	-0,2039	0,1204	0,1221	0,1710
33	-0,1278	-0,2107	0,0204	-0,0656	-0,2525	0,1112	0,1065	0,1625
34	-0,1722	-0,1992	0,0761	-0,0938	-0,2351	0,1313	0,1150	0,1619
35	-0,1458	-0,2325	0,2683	-0,1370	-0,2412	0,1251	0,1310	0,1727
36	-0,1598	-0,1639	0,0103	-0,0416	-0,2845	0,1150	0,1027	0,1696

Tabela 51 - Densidade eletrofílica de fronteira, normalizada pela energia do HOMO, do n-ésimo átomo da N²-fenilguanina ($F_n^{(e)}$). Valores multiplicados por 100.

No.	$F_1^{(e)}$	$F_2^{(e)}$	$F_3^{(e)}$	$F_4^{(e)}$	$F_5^{(e)}$	$F_6^{(e)}$	$F_7^{(e)}$	$F_8^{(e)}$
1	0,48381	1,48881	2,90545	2,86435	4,97162	0,14399	1,01991	4,16459
2	0,43057	1,44604	3,03910	2,77689	5,02192	0,14753	1,00867	4,14917
3	0,39684	1,36389	2,89309	2,59934	4,66692	0,13680	0,94387	3,86388
4	0,41959	1,42927	3,01480	2,69405	4,89808	0,14543	0,98397	4,04112
5	0,42423	1,43174	3,02356	2,73951	4,96368	0,14573	0,99517	4,09825
6	0,41499	1,42864	3,03941	2,70323	4,93667	0,14657	0,98913	4,06648
7	0,41929	1,42404	3,01935	2,71866	4,93504	0,14514	0,98953	4,07322
8	0,42279	1,42741	3,00876	2,72520	4,97639	0,14534	0,98406	4,08758
9	0,35459	1,35246	3,14708	2,57177	4,91117	0,14636	0,95988	3,98550
10	0,45176	1,42324	2,81328	2,71787	4,71686	0,13702	0,97013	3,94957
11	0,35657	1,35514	3,14058	2,57381	4,90258	0,14615	0,95961	3,98259
12	0,46764	1,47142	2,93013	2,82243	4,95513	0,14399	1,00992	4,13255
13	0,29684	1,25248	3,09428	2,38645	4,66578	0,13904	0,90165	3,75595
14	0,41105	1,38882	2,88793	2,67264	4,74919	0,13834	0,96589	3,94569
15	0,36893	1,37242	3,13847	2,61400	4,94460	0,14696	0,96996	4,02689
16	0,36640	1,33978	2,99130	2,55852	4,73408	0,13938	0,94141	3,88036
17	0,28290	1,22561	3,07502	2,33575	4,59646	0,13720	0,88700	3,69348
18	0,32097	1,29319	3,10882	2,46679	4,75499	0,14146	0,92590	3,84613
19	0,49191	1,47560	2,79265	2,85247	4,83321	0,13890	1,00692	4,08145
20	0,38382	1,38306	3,08685	2,64498	4,92152	0,14529	0,97480	4,02751
21	0,35493	1,35309	3,13955	2,58096	4,91236	0,14626	0,96253	3,99143
22	0,59407	1,56135	2,66957	2,91313	4,75370	0,13839	1,01303	4,07153
23	0,32030	1,29291	3,10841	2,46662	4,75344	0,14139	0,92570	3,84508
24	0,23102	1,04382	2,55174	1,96967	3,76693	0,11049	0,73655	3,04857
25	0,30598	1,26488	3,08238	2,39969	4,66618	0,13914	0,90554	3,76430
26	0,41601	1,38868	2,89061	2,63841	4,70136	0,13790	0,95514	3,90283
27	0,31337	1,28051	3,10288	2,44529	4,72921	0,14063	0,91884	3,82085
28	0,42838	1,39524	2,81828	2,66841	4,67397	0,13563	0,95700	3,89947
29	0,31943	1,28985	3,10854	2,46168	4,74837	0,14137	0,92495	3,84037
30	0,42263	1,39302	2,84823	2,66066	4,69000	0,13676	0,95837	3,90748
31	0,37437	1,34840	2,95881	2,58067	4,71829	0,13868	0,94553	3,88261
32	0,27183	1,20283	3,02626	2,30750	4,53259	0,13450	0,87386	3,64178
33	0,32093	1,29373	3,12069	2,47150	4,77215	0,14198	0,92913	3,85848
34	0,33561	1,30648	3,05120	2,50286	4,73746	0,14001	0,93187	3,85447
35	0,36072	1,32791	2,95709	2,54558	4,67978	0,13732	0,93483	3,84311
36	0,32011	1,29327	3,10933	2,46872	4,75034	0,14143	0,92715	3,84785

Tabela 51 - Densidade eletrofílica de fronteira (continuação).

No.	F ₉ ^(e)	F ₁₀ ^(e)	F ₁₁ ^(e)	F ₁₂ ^(e)	F ₁₃ ^(e)	F ₁₄ ^(e)	F ₁₅ ^(e)	F ₁₆ ^(e)
1	0,06209	1,72118	1,48039	0,00211	0,00014	0,00008	0,08927	0,18952
2	0,07245	1,68799	1,65868	0,00201	0,00009	0,00009	0,08432	0,24458
3	0,06975	1,56301	1,81555	0,00282	0,00008	0,00010	0,09433	0,42747
4	0,07512	1,64726	1,75776	0,00254	0,00016	0,00010	0,09584	0,34178
5	0,07257	1,66677	1,69376	0,00217	0,00003	0,00009	0,08775	0,27636
6	0,07598	1,65346	1,76005	0,00237	0,00014	0,00010	0,09261	0,32564
7	0,07317	1,65440	1,71158	0,00232	0,00014	0,00009	0,08725	0,28733
8	0,06961	1,67357	1,68887	0,00225	0,00010	0,00008	0,08701	0,27695
9	0,08608	1,59069	1,96732	0,00236	0,00012	0,00010	0,08505	0,42394
10	0,06133	1,62503	1,61283	0,00264	0,00017	0,00009	0,09481	0,31733
11	0,08585	1,58928	1,96249	0,00241	0,00015	0,00010	0,08587	0,42611
12	0,06457	1,70258	1,53764	0,00220	0,00014	0,00008	0,08967	0,21880
13	0,08897	1,46916	2,15166	0,00247	0,00015	0,00012	0,07933	0,66235
14	0,06667	1,59733	1,69853	0,00236	0,00012	0,00010	0,08637	0,37854
15	0,08363	1,61143	1,92014	0,00232	0,00011	0,00010	0,08606	0,38661
16	0,07687	1,55285	1,88777	0,00249	0,00015	0,00011	0,08580	0,46610
17	0,08976	1,43794	2,18348	0,00245	0,00014	0,00012	0,07582	0,72528
18	0,08725	1,51622	2,08702	0,00252	0,00011	0,00010	0,08316	0,57537
19	0,05623	1,68980	1,44875	0,00220	0,00011	0,00009	0,09086	0,22489
20	0,07963	1,61912	1,83916	0,00237	0,00013	0,00009	0,08556	0,35497
21	0,08528	1,59013	1,94714	0,00233	0,00013	0,00010	0,08278	0,42157
22	0,05064	1,75635	1,30500	0,00289	0,00022	0,00008	0,11249	0,19424
23	0,08729	1,51480	2,08681	0,00246	0,00015	0,00011	0,08267	0,57571
24	0,07231	1,17486	2,11810	0,00284	0,00019	0,00015	0,07819	1,07908
25	0,08823	1,48183	2,12210	0,00250	0,00009	0,00012	0,08220	0,64264
26	0,06882	1,58966	1,76956	0,00292	0,00024	0,00010	0,09683	0,38112
27	0,08750	1,50184	2,10549	0,00240	0,00012	0,00012	0,08135	0,60312
28	0,06262	1,59223	1,65864	0,00261	0,00007	0,00010	0,09130	0,38061
29	0,08756	1,51336	2,09232	0,00242	0,00015	0,00011	0,08224	0,57968
30	0,06531	1,59222	1,69334	0,00265	0,00012	0,00009	0,09212	0,37617
31	0,07409	1,55617	1,86883	0,00260	0,00014	0,00011	0,08843	0,45931
32	0,08705	1,41003	2,16950	0,00207	0,00009	0,00013	0,06781	0,85911
33	0,08792	1,52266	2,08795	0,00246	0,00016	0,00011	0,08194	0,56372
34	0,08182	1,52476	1,99218	0,00244	0,00012	0,00010	0,08153	0,54609
35	0,07477	1,53368	1,91737	0,00265	0,00016	0,00011	0,08837	0,53121
36	0,08787	1,51405	2,08728	0,00244	0,00015	0,00011	0,08240	0,57484

Tabela 51 - Densidade eletrofílica de fronteira (continuação).

No.	F ₁₇ ^(e)	F ₁₈ ^(e)	F ₁₉ ^(e)	F ₂₀ ^(e)	F ₂₁ ^(e)	F ₂₂ ^(e)	F ₂₅ ^(e)	F ₂₆ ^(e)
1	0,24895	0,05434	0,28183	0,01054	0,24312	0,00009	0,00053	0,00394
2	0,26988	0,06676	0,33550	0,01622	0,27685	0,00015	0,00058	0,00450
3	0,35406	0,12988	0,64210	0,04381	0,43273	0,00018	0,00072	0,00484
4	0,27507	0,11315	0,48553	0,01838	0,39978	0,00009	0,00059	0,00489
5	0,26189	0,08935	0,39725	0,01530	0,32996	0,00024	0,00072	0,00490
6	0,28062	0,10084	0,45367	0,01924	0,37015	0,00011	0,00064	0,00498
7	0,23189	0,10615	0,44837	0,01091	0,39133	0,00013	0,00053	0,00470
8	0,24564	0,09633	0,41299	0,01321	0,35083	0,00018	0,00063	0,00489
9	0,29061	0,14071	0,58092	0,02472	0,46296	0,00021	0,00082	0,00565
10	0,28608	0,11647	0,52061	0,02559	0,37464	0,00013	0,00053	0,00415
11	0,28599	0,14414	0,58650	0,02410	0,47011	0,00030	0,00084	0,00584
12	0,23634	0,07160	0,32736	0,01050	0,28523	0,00011	0,00050	0,00409
13	0,34562	0,25891	0,85812	0,07341	0,55898	0,00027	0,00099	0,00595
14	0,25850	0,17204	0,53554	0,04016	0,37889	0,00014	0,00054	0,00403
15	0,31701	0,10830	0,51442	0,02769	0,40114	0,00029	0,00088	0,00586
16	0,30047	0,19951	0,67238	0,04233	0,45543	0,00018	0,00068	0,00493
17	0,26707	0,29682	0,91273	0,07293	0,70455	0,00039	0,00096	0,00582
18	0,38793	0,18296	0,72634	0,06984	0,46253	0,00035	0,00112	0,00580
19	0,27820	0,07579	0,33151	0,02431	0,23861	0,00009	0,00048	0,00349
20	0,27151	0,12214	0,49406	0,01954	0,40429	0,00023	0,00068	0,00536
21	0,28219	0,14900	0,55657	0,02731	0,44130	0,00031	0,00074	0,00579
22	0,26359	0,03865	0,38812	0,00432	0,33206	0,00005	0,00045	0,00374
23	0,38577	0,18384	0,72523	0,07047	0,46253	0,00036	0,00114	0,00587
24	1,60097	0,25424	0,45588	0,50420	0,03633	0,00028	0,00104	0,00468
25	0,32645	0,25653	0,85845	0,06332	0,57520	0,00029	0,00109	0,00608
26	0,34216	0,10721	0,59868	0,03063	0,42754	0,00016	0,00068	0,00489
27	0,39599	0,19641	0,75183	0,07914	0,46414	0,00038	0,00102	0,00590
28	0,22787	0,19244	0,59754	0,03231	0,45084	0,00007	0,00041	0,00376
29	0,38876	0,18516	0,73095	0,07113	0,46411	0,00036	0,00115	0,00587
30	0,27238	0,16037	0,58593	0,03241	0,42084	0,00012	0,00050	0,00415
31	0,34516	0,15045	0,61967	0,05585	0,41429	0,00024	0,00080	0,00491
32	0,35996	0,35890	0,78660	0,19002	0,49495	0,00048	0,00103	0,00528
33	0,38378	0,17800	0,71358	0,06697	0,45894	0,00037	0,00102	0,00605
34	0,34244	0,20870	0,67481	0,06837	0,43579	0,00033	0,00089	0,00559
35	0,36047	0,18587	0,64612	0,07951	0,41784	0,00027	0,00085	0,00482
36	0,38595	0,18355	0,72473	0,07049	0,46099	0,00036	0,00114	0,00584

Tabela 52 - Densidade nucleofílica de fronteira, normalizada pela energia do LUMO, do n-ésimo átomo da N²-fenilguanina ($F_n^{(n)}$). Valores multiplicados por 100.

No.	$F_1^{(n)}$	$F_2^{(n)}$	$F_3^{(n)}$	$F_4^{(n)}$	$F_5^{(n)}$	$F_6^{(n)}$	$F_7^{(n)}$	$F_8^{(n)}$
1	-12,7771	-61,1315	-12,6171	-19,6174	-18,9092	-0,60582	-0,06000	-9,5259
2	-9,3422	-45,0519	-8,6328	-15,0289	-13,7443	-0,53762	-0,04367	-6,9548
3	-16,3660	-77,9551	-15,7126	-26,8391	-24,9874	-0,63617	-0,06736	-12,5927
4	-22,2080	-105,7473	-19,4219	-40,4389	-35,1171	-0,68685	-0,07393	-17,8210
5	-19,9472	-95,3054	-17,6736	-35,3759	-31,1156	-0,68519	-0,07067	-15,7282
6	-23,0362	-109,8789	-19,9227	-42,3987	-36,5074	-0,70877	-0,07576	-18,4964
7	-20,3765	-97,3547	-18,0598	-36,3280	-31,8854	-0,68249	-0,07082	-16,1191
8	-19,7174	-94,1618	-17,5105	-34,8480	-30,6704	-0,67563	-0,06705	-15,4717
9	-30,2125	-144,7641	-23,8667	-60,6586	-49,0353	-0,80741	-0,07653	-24,9330
10	-10,6086	-50,8341	-11,2915	-15,0695	-15,3328	-0,57003	-0,05704	-7,6856
11	-30,2487	-144,9888	-23,9942	-60,6333	-49,1256	-0,80317	-0,07474	-25,0164
12	-10,6667	-51,3251	-10,5070	-16,2383	-15,6138	-0,54577	-0,05079	-7,8833
13	-30,7345	-147,9859	-24,2517	-61,7249	-49,7124	-0,88341	-0,07569	-25,2958
14	-11,9950	-57,7176	-12,0441	-18,0336	-17,5113	-0,67179	-0,06189	-8,7658
15	-28,2761	-135,4482	-22,8007	-55,6381	-45,5914	-0,78028	-0,07519	-23,1839
16	-20,3365	-97,3473	-18,0163	-36,1449	-31,6808	-0,72538	-0,06961	-16,0373
17	-31,1201	-149,9021	-24,4089	-62,6282	-50,2821	-0,89407	-0,07690	-25,5742
18	-28,8557	-138,5237	-23,1657	-56,8799	-46,4323	-0,83854	-0,07413	-23,6034
19	-5,9291	-29,0088	-6,8559	-7,1239	-7,9128	-0,47890	-0,04177	-3,9516
20	-24,4597	-117,0025	-20,2168	-46,8318	-39,1748	-0,69837	-0,07071	-19,8630
21	-29,1802	-140,0242	-23,0486	-58,5947	-47,3833	-0,78156	-0,07533	-24,0572
22	-11,0587	-51,6645	-12,2095	-15,6363	-16,4416	-0,41855	-0,05678	-8,3089
23	-28,7876	-138,4048	-23,1698	-56,7268	-46,3483	-0,83891	-0,07468	-23,5802
24	-7,1410	-34,6875	-7,5427	-10,3100	-10,1764	-0,51331	-0,03681	-5,0564
25	-30,5778	-146,7147	-24,1853	-61,1213	-49,3976	-0,86538	-0,07656	-25,1337
26	-16,6995	-79,3107	-16,0483	-27,4685	-25,6051	-0,60768	-0,06664	-12,9308
27	-28,6206	-137,5661	-23,0005	-56,3145	-45,9808	-0,84333	-0,07552	-23,3700
28	-11,4497	-54,7422	-11,8130	-16,8511	-16,7452	-0,60698	-0,05951	-8,3727
29	-28,8078	-138,3699	-23,1425	-56,7480	-46,3356	-0,83361	-0,07421	-23,5707
30	-12,9689	-61,9552	-13,0736	-19,7710	-19,2229	-0,62342	-0,06245	-9,6507
31	-18,1312	-86,7853	-16,6664	-31,0175	-27,9249	-0,70254	-0,07047	-14,0796
32	-27,6277	-133,3241	-21,6437	-55,1424	-44,3968	-0,85663	-0,07244	-22,5306
33	-29,3536	-140,8766	-23,3543	-58,3225	-47,3224	-0,85467	-0,07583	-24,0496
34	-24,2307	-116,3797	-20,1023	-46,1207	-38,5840	-0,77262	-0,07326	-19,5463
35	-18,1834	-86,9955	-16,5511	-31,2824	-28,0149	-0,69594	-0,06857	-14,1321
36	-28,7559	-138,2897	-23,1527	-56,6657	-46,3036	-0,83479	-0,07504	-23,5858

Tabela 52 - Densidade nucleofílica de fronteira (continuação).

No.	$F_9^{(n)}$	$F_{10}^{(n)}$	$F_{11}^{(n)}$	$F_{12}^{(n)}$	$F_{13}^{(n)}$	$F_{14}^{(n)}$	$F_{15}^{(n)}$	$F_{66}^{(n)}$
1	-11,9309	-0,23964	-2,42773	-0,18807	-0,00252	-0,04231	-1,23689	-32,3328
2	-8,8617	-0,20975	-2,34253	-0,17461	-0,00197	-0,03367	-0,84457	-23,1230
3	-16,0008	-0,27315	-2,88300	-0,21561	-0,00122	-0,03624	-1,73568	-32,4000
4	-23,1597	-0,34366	-4,96096	-0,27417	-0,00124	-0,04412	-2,89555	-24,8666
5	-20,3432	-0,32383	-4,55760	-0,26384	-0,00087	-0,04349	-2,44596	-27,4148
6	-24,1384	-0,36119	-5,39546	-0,28241	-0,00114	-0,04564	-3,01059	-25,2171
7	-20,9038	-0,32833	-4,57726	-0,26199	-0,00106	-0,03325	-2,52466	-27,3622
8	-20,0538	-0,31887	-4,41796	-0,25714	-0,00115	-0,03936	-2,40113	-27,4778
9	-33,2408	-0,47733	-8,79256	-0,34763	-0,00052	-0,03352	-4,25622	-21,8841
10	-9,4947	-0,21310	-1,77377	-0,15090	-0,00254	-0,03534	-0,82663	-37,9473
11	-33,2916	-0,47452	-8,74848	-0,34440	-0,00056	-0,03641	-4,29020	-21,8079
12	-9,8670	-0,21046	-2,14860	-0,16983	-0,00296	-0,03475	-0,99394	-29,2017
13	-33,7647	-0,50235	-9,12623	-0,37091	-0,00070	-0,03796	-4,21009	-25,3238
14	-11,0166	-0,25793	-2,38250	-0,19632	-0,00227	-0,03889	-1,09487	-37,5974
15	-30,7519	-0,44222	-7,90442	-0,33201	-0,00056	-0,03349	-3,89798	-22,6804
16	-20,7751	-0,33822	-4,61063	-0,27014	-0,00134	-0,04198	-2,45673	-29,9770
17	-34,1909	-0,51019	-9,37621	-0,37262	-0,00068	-0,03481	-4,23293	-25,9243
18	-31,3481	-0,46666	-8,16130	-0,35039	-0,00069	-0,03729	-3,91073	-25,2816
19	-4,7310	-0,16621	-1,39200	-0,10079	-0,00294	-0,03356	-0,30865	-37,2399
20	-26,1572	-0,38258	-6,53871	-0,29501	-0,00068	-0,02255	-3,34395	-21,8352
21	-32,1040	-0,46443	-8,72584	-0,33945	-0,00070	-0,03167	-4,14039	-20,7423
22	-10,0866	-0,16089	-1,28259	-0,10497	-0,00268	-0,02956	-0,79114	-32,8486
23	-31,2895	-0,46330	-8,14696	-0,34791	-0,00073	-0,03683	-3,89777	-25,5417
24	-6,3670	-0,18794	-1,40182	-0,11806	-0,00138	-0,02529	-0,58012	-37,0366
25	-33,4860	-0,49487	-8,84918	-0,36409	-0,00058	-0,03905	-4,24632	-24,7487
26	-16,4144	-0,26558	-2,81067	-0,20906	-0,00229	-0,03344	-1,79621	-31,1295
27	-31,0365	-0,46321	-8,13004	-0,34969	-0,00086	-0,03737	-3,83912	-26,1389
28	-10,4300	-0,23274	-2,02368	-0,17080	-0,00175	-0,03724	-1,00513	-36,4411
29	-31,2895	-0,46262	-8,14188	-0,33335	-0,00081	-0,03965	-3,88589	-25,6530
30	-12,1001	-0,24632	-2,26950	-0,18704	-0,00214	-0,03761	-1,21094	-35,7498
31	-18,1168	-0,31160	-3,77865	-0,25369	-0,00147	-0,03758	-2,07598	-31,6042
32	-30,0808	-0,46938	-8,81807	-0,34820	-0,00089	-0,04474	-3,68466	-25,1330
33	-32,0329	-0,48196	-8,41189	-0,35233	-0,00071	-0,03697	-3,97426	-25,0488
34	-25,7680	-0,40066	-6,64349	-0,32003	-0,00071	-0,02728	-3,23283	-24,7486
35	-18,2070	-0,31089	-3,92254	-0,25259	-0,00174	-0,03550	-2,11218	-30,4673
36	-31,2801	-0,46045	-8,14353	-0,34750	-0,00067	-0,03684	-3,88875	-25,6761

Tabela 52 - Densidade nucleofílica de fronteira (continuação).

No.	$F_{17}^{(n)}$	$F_{18}^{(n)}$	$F_{19}^{(n)}$	$F_{20}^{(n)}$	$F_{21}^{(n)}$	$F_{22}^{(n)}$	$F_{25}^{(n)}$	$F_{26}^{(n)}$
1	-5,3634	-9,9038	-34,9780	-3,5094	-18,9513	-0,00107	-0,00040	-0,00159
2	-0,1213	-18,9771	-33,1579	-0,1904	-28,8326	-0,00168	-0,01231	-0,00639
3	-12,7360	-4,5640	-28,8288	-9,3024	-8,3752	-0,00082	-0,00019	-0,00075
4	-5,7581	-6,4118	-26,4874	-2,7189	-15,2054	-0,00298	0,00000	-0,00354
5	-10,3271	-3,8598	-26,2913	-5,1693	-11,2041	-0,00239	-0,00004	-0,00282
6	-8,3239	-4,4813	-25,1517	-3,7867	-12,5010	-0,00353	-0,00003	-0,00425
7	-12,2674	-2,7489	-24,3666	-6,2969	-9,0975	-0,00237	-0,00007	-0,00291
8	-10,5058	-3,7821	-25,9825	-5,3285	-10,8801	-0,00182	-0,00031	-0,00375
9	-12,1088	-1,8009	-19,7444	-4,5839	-8,3568	-0,00358	-0,00010	-0,00688
10	-8,1233	-10,6980	-37,7069	-8,1681	-13,6129	-0,00047	-0,00024	-0,00042
11	-12,0398	-1,8037	-19,6936	-4,5656	-8,3558	-0,00295	-0,00023	-0,00692
12	-0,1109	-22,8606	-40,1197	-0,2458	-34,7034	-0,00095	-0,00011	-0,00253
13	-13,1175	-2,5120	-24,1235	-5,4592	-9,5991	-0,00390	-0,00013	-0,00776
14	-13,5486	-6,0009	-38,8091	-9,2344	-10,8878	-0,00117	-0,00021	-0,00107
15	-12,7938	-1,5809	-20,2087	-5,1885	-7,9918	-0,00275	-0,00033	-0,00651
16	-13,5950	-3,3808	-28,2259	-7,1720	-9,2068	-0,00204	-0,00022	-0,00296
17	-11,9167	-2,3975	-24,5786	-5,5179	-11,5243	-0,00353	-0,00030	-0,00842
18	-13,9743	-1,9802	-23,4613	-6,0094	-8,5706	-0,00348	-0,00033	-0,00661
19	-2,3484	-19,4597	-46,3131	-3,6705	-23,6551	-0,00031	-0,00015	-0,00053
20	-11,1358	-1,8431	-19,7538	-4,6077	-8,1840	-0,00282	-0,00029	-0,00527
21	-11,7099	-1,3777	-18,8465	-4,2492	-8,2194	-0,00319	-0,00014	-0,00792
22	-9,8034	-6,5445	-26,3747	-10,4136	-6,7509	-0,00010	-0,00028	-0,00001
23	-14,0397	-2,0203	-23,6961	-6,1057	-8,6443	-0,00340	-0,00044	-0,00706
24	-46,1985	-1,4812	-4,9652	-32,3908	-2,7865	-0,00107	-0,00001	-0,00009
25	-12,7917	-2,4541	-23,3778	-5,4275	-9,2616	-0,00332	-0,00042	-0,00913
26	-12,0085	-4,5293	-26,9546	-8,9687	-8,0693	-0,00080	-0,00018	-0,00064
27	-14,4848	-2,0511	-24,4476	-6,3341	-8,6337	-0,00328	-0,00014	-0,00664
28	-14,4705	-4,5081	-34,7728	-10,6161	-8,4546	-0,00107	-0,00027	-0,00067
29	-14,0986	-2,0369	-23,7974	-6,1243	-8,6596	-0,00333	-0,00048	-0,00703
30	-11,8900	-6,3104	-34,1323	-8,8886	-10,4410	-0,00107	-0,00016	-0,00085
31	-14,1350	-3,5592	-29,3504	-8,7064	-8,6447	-0,00147	-0,00020	-0,00174
32	-17,6036	-0,8159	-26,5889	-6,4116	-7,5349	-0,00406	-0,00017	-0,00871
33	-13,8195	-1,9433	-23,1531	-5,8527	-8,6912	-0,00299	-0,00031	-0,00708
34	-11,3647	-2,6068	-24,3499	-5,0008	-9,9038	-0,00193	-0,00019	-0,00587
35	-14,8837	-2,8603	-30,3849	-8,1908	-8,3611	-0,00146	-0,00021	-0,00194
36	-14,1035	-2,0352	-23,8138	-6,1440	-8,6601	-0,00353	-0,00042	-0,00721

CONCLUSÕES

ANÁLISE CONFORMACIONAL

- (a) A análise dos mapas conformacionais dos derivados da $\underline{\text{N}}^2$ -fenilguanina indicou a existência de quatro regiões de estabilidade conformacional, A, B, C e D, localizadas nos seguintes intervalos dos ângulos torsionais θ_1 e θ_2 : $100^\circ \leq \theta_1 \leq 250^\circ$ e $90^\circ \leq \theta_2 \leq 140^\circ$ e $250^\circ \leq \theta_2 \leq 320^\circ$;
- (b) Apenas quatro conformações estáveis das $\underline{\text{N}}^2$ -fenilguaninas, C1, C2, C3 e C4, foram identificadas na análise conformacional. Os ângulos torsionais θ_1 e θ_2 que caracterizam essas conformações são: C1, $\theta_1 = 184^\circ$ e $\theta_2 = 58^\circ$; C2, $\theta_1 = 174^\circ$ e $\theta_2 = 127^\circ$; C3, $\theta_1 = 186^\circ$ e $\theta_2 = 233^\circ$ e; C4, $\theta_1 = 176^\circ$ e $\theta_2 = 303^\circ$.

QSAR COM PARÂMETROS CLÁSSICOS

- (a) A atividade dos derivados meta-substituídos da $\underline{\text{N}}^2$ -fenilguanina sobre HSV2 TK é governada principalmente pelo parâmetro eletrônico, sendo que substituintes capazes de extrair elétrons do anel fenila intensificam a atividade. A lipofilicidade exerce influência menor na atividade desses compostos sobre HSV2, sendo que substituintes meta mais lipofílicos intensificam a atividade inibitória;
- (b) A atividade inibitória global das $\underline{\text{N}}^2$ -fenilguaninas sobre HSV1 TK e HSV2 TK depende dos parâmetros π_m , π_p , σ_m e R_p ;
- (c) Qualitativamente, os parâmetros π_m , π_p , σ_m e R_p exercem influência semelhante sobre a atividade das $\underline{\text{N}}^2$ -fenilguaninas sobre HSV1 TK e HSV2 TK. Isto pode ser interpretado como uma indicação das semelhança do mecanismo de ação das $\underline{\text{N}}^2$ -fenilguaninas sobre HSV1 TK e HSV2 TK;

- (d) Quantitativamente, ocorrem diferenças na influência dos parâmetros π_m , π_p , σ_m e R_p sobre a atividade das N^2 -fenilguaninas sobre HSV1 TK e HSV2 TK: a atividade inibitória global sobre HSV1 TK é influenciada pelos parâmetros lipofílico em meta ($+\pi_m$) e hidrofílico em para ($+\pi_p$), pelo efeito de ressonância em para ($+R_p$) e pelo efeito indutivo em meta ($+\sigma_m$) de forma aproximadamente equivalente, enquanto que a atividade inibitória global sobre HSV2 TK é principalmente influenciada pelo efeito indutivo em meta ($+\sigma_m$) e pela lipofilicidade em meta ($+\pi_m$) e para ($-\pi_p$). Isto pode ser interpretado como sendo reflexo de pequenas diferenças nos sítios ativos dessas enzimas;
- (e) A região do sítio ativo de HSV1 TK, próxima à posição meta do anel fenila das N^2 -fenilguaninas, possui natureza anfifílica (caráter duplo), predominando o caráter apolar ($+\pi_m$ e $+\sigma_m$), enquanto que a região próxima à posição para possui natureza hidrofílica ($-\pi_p$);
- (f) A região do sítio ativo de HSV2 TK, próxima à posição meta do anel fenila das N^2 -fenilguaninas, também possui natureza anfifílica, predominando o caráter polar ($+\sigma_m$ e $+\pi_m$), enquanto que a região próxima à posição para possui natureza hidrofílica ($-\pi_p$);

MODELAGEM MOLECULAR E CONFORMAÇÃO BIOATIVA

- (a) Os derivados da N^2 -fenilguanina interagem com o sítio ativo de HSV1 TK adotando o modo de ligação de 2'-deoxitimidina;
- (b) Na conformação bioativa, os valores dos ângulos torsionais θ_1 e θ_2 , valem aproximadamente 180° e 220° , respectivamente, que corresponde à conformação C3 prevista pela análise conformacional;
- (c) o substituinte ligado na posição meta do anel fenila fica preferencialmente orientado na direção dos resíduos Tyr 101 (polar e não ionizável), Met 231 (apolar), Ile 97 (apolar)

- e a parte hidrocarbônica da cadeia lateral de Glu 225 (polar e ionizável), que compõe o ambiente anfifílico previsto pelo estudo QSAR do Capítulo 4;
- (d) o substituinte ligado na posição para do anel fenila fica próximo à região altamente polar, constituída pelos resíduos His 58 (polar e ionizável), Gln 221 (polar e neutro), Arg 222 (polar e ionizável) e Glu 225 (polar e ionizável), em acordo com a previsão do estudo QSAR do Capítulo 4;
 - (e) a observação experimental acerca da interação não fosforilativa entre os derivados da N^2 -fenilguanina e HSV1 TK fornece suporte adicional para a hipótese do modo de ligação proposto;
 - (f) as conclusões acima reforçam a hipótese de que o anel guanina das N^2 -fenilguaninas interagem com os resíduos Gln 125, Met 128 e Tyr 172 de forma semelhante à interação desses resíduos com 2'-deoxitimidina.

QSAR COM PARÂMETROS CLÁSSICOS E CALCULADOS

- (a) A atividade inibitória dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV1 TK depende dos parâmetros lipofílico e eletrônico, sendo que, isoladamente, o parâmetro lipofílico é o mais importante;
- (b) A atividade dos derivados meta-substituídos da N^2 -fenilguanina sobre HSV2 TK depende dos parâmetros lipofílico, eletrônico e estérico, sendo que, isoladamente, o parâmetro eletrônico é mais importante;
- (c) Há indicações de que o resíduo Leu 97 é o responsável pelo efeito estérico observado na atividade dos derivados meta-substituídos sobre HSV2 TK. Na posição 97, HSV1 TK apresenta o aminoácido Ile;
- (d) A densidade eletrofílica de fronteira do átomo O(10) pode ser importante para a atividade inibitória sobre HSV1 TK e HSV2 TK. No entanto, apenas no caso dos

derivados meta-substituídos sobre HSV1 TK houve modelo que incluiu explicitamente essa variável.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDUL-AHAD, P.G.; BLAIR, T.; WEBB, G.A. Quantitative structure-activity study of some enzyme-inhibitory quinazolines. *Int. J. Quantum Chem.*, v.17, p.821-831, 1980.
- ABELE, U.; SCHULZ, G.E. High-resolution structures of adenylate kinase from yeast ligated with inhibitor Ap5A, showing the pathway of phosphoryl transfer. *Prot. Sci.*, v.4, p. 1262-1271, 1995.
- ALBRECHT, J.-C.; NICHOLAS, J.; BILLER, D.; CAMERON, K.R.; BIESINGER, B.; NEWMAN, C.; WITTMANN, S.; CRAXTON, M.A.; COLEMAN, H.; FLECKENSTEIN, B.; HONESS, R.W. Primary structure of the herpesvirus saimiri genome. *J. Virol.*, v.66, p.5047-5058, 1992.
- ALLINGER, N.L. Conformational analysis 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J. Am. Chem. Soc.*, v.99, p.8127-8134, 1977.
- BALASUBRAMANIAM, N.K.; VEERISSETTY, V.; GENTRY, G.A. Herpesviral deoxythymidine kinases contain a site analogous to the phosphoryl-binding arginine-rich region of porcine adenylate kinase - comparison of secondary structure predictions and conservation. *J. Gen. Virol.*, v.71, p.2979-2987, 1990.
- BESLER, B.H.; MERZ, K.M.; KOLLMAN, P.A. Atomic charges derived from semiempirical methods. *J. Comp. Chem.*, v.11, p.431-439, 1990.
- BLACK, M.E.; HRUBY, D.E. Identification of the ATP-binding domain of vaccinia virus thymidine kinase. *J. Biol. Chem.*, v.265, p.17584-17592, 1990.
- BLOW, D.M.; STEITZ, T.A. X-ray diffraction studies of enzymes. *Ann. Rev. Biochem.*, v.39, p.63-100, 1970.

- BOEHMER, P.E.; LEHMAN, I.R. Herpes simplex virus DNA replication. *Ann. Ver. Biochem.*, v.66, p.347-384, **1997**.
- BRADSHAW, H.D. JR.; DEININGER, P.L. Human thymidine kinase gene - molecular-cloning and nucleotide-sequence of a cdna expressible in mammalian-cells. *Mol. Cell Biol.*, v.4, p.2316-2320, **1984**.
- BRANDEN, C.; TOOZE, J. *Introduction to Protein Structure*. New York, Garland Publishing, **1991**, pp.302.
- BROOKS, B.R.; BRUCCOLERI, R.E.; OLAFSON, B.D.; STATES, D.J.; SWAMINATHAN, S.; KARPLUS, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculation. *J. Comp. Chem.*, v.4, p.187-217, **1983**.
- BROWN, D.G.; VISSE, R.; SANDHU, G.; DAVIES, A.; RIZKALLAH, P.J.; MELITZ, C.; SUMMERS, W.C.; SANDERSON, M.R. Crystal structures of the thymidine kinase from herpes simplex virus type-I in complex with deoxythymidine and ganciclovir. *Nature Struct. Biol.* v.2, p.876-881, **1995**.
- CAMELEON v.3.1, Oxford Molecular Ltd., Magdalen Centre, Oxford Science Park, Sandford-on-Thames, Oxford, OX4 4GA, England, **1994**.
- CHARMM User Guide, Release 22.3. San Diego, Molecular Simulations, **1994**.
- CHIN, C.-A.; SONG, P.-S. Reactivity Indices for Biomolecules, In: *Graduate Studies*: Texas Tech University, No. 24, Texas, Texas Tech Press, **1981**.
- COEN, D.M. Nucleosides and foscarnet – Mechanisms. In: *Antiviral drug resistance*. RICHMAN, D.D., ed., Guildford, John Wiley & Sons, **1996**, p.81-102.
- COLOMINAS, C.; LUQUE, F.J.; OROZCO, M. Tautomerism and protonation of guanine and cytosine. Implications in the formation of hydrogen-bonded complexes. *J. Am. Chem. Soc.*, v.118, p.6811-6821, **1996**.

- COLLIER, L.; OXFORD, J. *Human Virology: A Text for Students of Medicine, Dentistry and Microbiology*. Oxford, Oxford University Press, 1993, p.185-212.
- COX, S.R.; WILLIAMS, D.E. Representation of the molecular electrostatic potential by a new atomic charge model. *J. Comp. Chem.*, v.2, p.304-323, 1981.
- CUTHBERTSON, A.F.; RICHARDS, W.G., The effect of fluorination on the binding-energies of drugs. *J. Mol. Struct. (Theochem)*, v.134, p.411-414, 1986.
- DARBY, G.; LARDER, B.A.; INGLIS, M.M. Evidence that the active-center of the herpes-simplex virus thymidine kinase involves an interaction between 3 distinct regions of the polypeptide. *J. Gen. Virol.* v.67: p.753-758, 1986.
- DAYHOFF, M.O. (Ed.) *Atlas of Protein Sequences and Structure*, NBRF, Washington, D.C. 1978, 5 (Suppl.3), 435.
- DESTRO, R.; KISTENMACHER, T.J.; MARSH, R.E. The crystal structure of 9-ethylguanine. *Acta Cryst.*, v.B30, p.79-85, 1974.
- DEWAR, M.J.S.; HADDON, R.C.; SUCK, S.H. MINDO/3 calculations of molecular electric polarizabilities. *J. Chem. Soc., Chem. Comm.*, p.611-612, 1974.
- DEWAR, M.J.S.; STEWART, J.J.P. A new procedure for calculating molecular polarizabilities. Applications using MNDO. *Chem. Phys. Lett.*, v.111, p.416-420, 1984.
- DEWAR, M.J.S.; ZOEBISCH, E.G.; HEALY, E.F.; STEWART, J.J.P. AM1 - a new general-purpose quantum-mechanical molecular-model *J. Am. Chem. Soc.*, v.107, p.3902-3909, 1985.
- DEWAR, M.J.S.; JIE, C.; YU, J. SAM1; The first of a new series of general purpose quantum mechanical molecular models. *Tetrahedron*, v.49, p.5003-5038, 1993.
- DRAPER, N.R.; SMITH, H. *Applied Regression Analysis*. 2.ed., New York, John Wiley & Sons, 1981, p.1-136.

- DREUSICKE, D.; KARPLUS, P.A.; SCHULZ, G.E. Refined structure of porcine cytosolic adenylate kinase at 2,1 Å resolution. *J. Mol. Biol.*, v.199, p.359-371, 1988.
- FABIAN, W.M.F. Tautomeric equilibria of heterocyclic molecules. A test of the semiempirical AM1 and MNDO-PM3 methods. *J. Comp. Chem.*, v.12, p.17-35, 1991.
- FETZER, J.; FOLKERS, G.; MÜLLER, I.; KEIL, G.M. Site-directed mutagenesis in the active-site of the herpes-simplex virus type-1 thymidine kinase gene. *Virus Genes*, v.7, p.205-209, 1993.
- FERENCZY, G.G.; REYNOLDS, C.A.; RICHARDS, W.G. Semiempirical AM1 electrostatic potentials and AM1 electrostatic potential derived charges - a comparison with *ab initio* values. *J. Comp. Chem.*, v.11, p.159-169, 1990.
- FLEMINGTON, E.; BRADSHAW, H.D. JR.; TRAINA-DORGE, V.; SLAGEL, V.; DEININGER, P.L. Sequence, structure and promoter characterization of the human thymidine kinase gene. *Gene*, v.52, p.267-277, 1987.
- FOCHER, F.; HILDEBRAND, C.; FREESE, S.; CIARROCHI, G.; NOONAN, T.; SANGALLI, S.; BROWN, N.; SPADARI, S.; WRIGHT, G. N²-phenyldeoxyguanosine: a novel selective inhibitor of herpes simplex virus thymidine kinase. *J. Med. Chem.*, v.31, p.1496-1500, 1988.
- FOLKERS, G.; TRUMPP, S.; EGER, K.; ROTH, H.J.; SCHWOEBEL, W. Sequence homologies in the C-terminal region and structural relationships of the herpes-simplex virus thymidine kinases to other thymidine kinases. *Med. Sci. Res.-Biochem.*, v.15, p.633-634, 1987.
- FUJITA, T. The extratermodynamic approach to drug design. In: *Comprehensive medicinal chemistry*. The rational design, mechanistic study, and therapeutic

- application of chemical compounds. HANSCH, C.; SAMMES, P.G.; TAYLOR, J.B., eds., New York, Pergamon Press, **1990**, p.497-560.
- FUJITA, T.; IWASA, J.; HANSCH, C. A new substituent constant, π , derived from partition coefficients. *J. Am. Chem. Soc.*, v.86, p.5175-5180, **1964**.
- FUKUI, K.; YONEZAWA, T.; SHINGU, H. A molecular orbital theory of reactivity in aromatic hydrocarbons. *J. Chem. Phys.*, v.20, p.722-725, **1952**.
- FUKUI, K.; YONEZAWA, T.; NAGATA, C.; SHINGU, H. Molecular orbital theory of orientation in aromatic, heteroaromatic, and other conjugated molecules. *J. Chem. Phys.*, v.22, p.1433-1442, **1954**.
- FUKUI, K.; YONEZAWA, T.; NAGATA, C. MO-theoretical approach to the mechanism of charge transfer in the process of aromatic substitutions. *J. Chem. Phys.*, v.27, p.1247-1259, **1957**.
- GAMBINO, J.; FOCHER, F.; HILDEBRAND, C.; MAJA, G.; NOONAN, T.; SPADARI, S.; WRIGHT, G. Quantitative structure-activity relationships of N²-phenylguanines as inhibitors of herpes simplex virus thymidine kinases. *J. Med. Chem.*, v.35, p.2979-2983, **1992**.
- GASTEIGER, J.; MARSILI, M. Iterative partial equalization of orbital electronegativity - rapid access to atomic charges. *Tetrahedron*, v.36, p.3219-3288, **1980**.
- GAUDIO, A.C. Modelos de estudo quantitativo das relações entre estrutura química e atividade biológica. *Química Nova*, v.19, p.278-289, **1996**.
- GAUDIO, A.C.; TAKAHATA, Y. Calculation of molecular surface area with numerical factors. *Comput. Chem.*, v.16, p.277-284, **1992**.
- GAUDIO, A.C.; KOROLKOVAS, A.; TAKAHATA, Y. Quantitative structure-activity relationships for 1,4-dihydropyridine calcium channel antagonists (nifedipine

- analogues): a quantum chemical/classical approach. *J. Pharm. Sci.*, v.83, p.1110-1115, 1994.
- GAUDIO, A.C.; TAKAHATA, Y.; RICHARDS, W.G. Prediction of the binding mode of N^2 -phenylguanine derivative inhibitors to herpes simplex virus type 1 thymidine kinase. *J. Comp.-Aid. Mol. Des.*, v.12, p.15-25, 1998.
- GEBHARDT, B.M.; WRIGHT, G.E.; XU, H.; FOCHER, F.; SPADARI, S.; KAUFMAN, H.E. 9-(4-Hydroxybutyl)- N^2 -phenylguanine (HBPG), a thymidine kinase inhibitor, supresses herpes virus reactivation in mice. *Antivir. Res.*, v.30, p.87-94, 1996.
- GOMPELS, U.; MINSON, A. The properties and sequence of glycoprotein-H of herpes-simplex virus type-1. *Virology*, v.153, p.230-247, 1986.
- GOULD, I.R.; HILLIER, I.H. Accurate calculations of the relative energies of the tautomers of cytosine and guanine. *Chem. Phys. Lett.*, v.161, p.185-187, 1989.
- GOULD, I.R.; BURTON, N.A.; HALL, R.J.; HILLIER, I.H. Tautomerism in uracil, cytosine and guanine: a comparison of electron correlation predicted by *ab initio* and density functional theory methods. *J. Mol. Struct. (Theochem)*, v.331, p.147-154, 1995.
- HADJIPAVLOU-LITINA, D. QSAR of N^2 -phenylguanidines as inhibitors of herpes simplex virus-1 thymidine kinase. *Pharmazie*, v.50, p.796-798, 1995.
- HAMMETT, L.P. The effect of structure upon the reactions of organic compounds. Benzene derivatives. *J. Am. Chem. Soc.*, v.59, p.96-103, 1937.
- HANSCH, C. *Structure-Activity Relationships*. v.1, CAVALLITO, C., ed., London, Pergamon Press, 1973.
- HANSCH, C.; FUJITA, T. ρ - σ - π analysis. A method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.*, v.86, p.1616-1626, 1964.

- HANSCH, C.; LEO, A. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*. Washington DC, American Chemical Society, 1995, pp.557.
- HANSCH, C.; LEO, A.; TAFT, R. A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.*, v.91, p.165-195, 1991.
- HANSCH, C.; LEO, A.; HOEKMAN, D. *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*. Washington, DC , American Chemical Society, 1995.
- HAYDEN, F.G. Antimicrobial Agents: Antiviral Agents. In HARDMAN, J.G.; LIMBIRD, L.E.; MOLINOFF, P.B.; RUDDON, R.W.; GILMAN, A.G., eds., *Goodman & Gilman's The Pharmacological Basis of Therapeutics*. 9.ed., New York, McGraw-Hill, 1996, cap.50.
- HERMANN, R.B. Theory of hydrophobic bonding. II. The correlation of hydrocarbon solubility in water with solvent cavity surface area. *J. Phys. Chem.*, v.76, p. 2754-2759, 1972.
- HIGO, J.; GÖ, N. Algorithm for rapid calculation of excluded volume of large molecules. *J. Comp. Chem.*, v.10, p.376-379, 1989.
- HILDEBRAND, C.; SANDOLI, D.; FOCHER, F.; GAMBINO, J.; CIARROCHI, G.; SPADARI, S.; WRIGHT, G. Structure-activity relationships of N²-substituted guanines as inhibitors of HSV1 and HSV2 thymidine kinases. *J. Med. Chem.*, v.33, p.203-206, 1990.
- HONESS, R.W.; CRAXTON, M.A.; WILLIAMS, L.; GOMPELS, U.A. A comparative-analysis of the sequence of the thymidine kinase gene of a gamma-herpesvirus, herpesvirus saimiri. *J. Gen. Virol.*, v.70, p.3003-3013, 1989.
- HSU, D.; SHIH, L.M.; ZEE, Y.C. Nucleotide-sequence of a 3.5 kilobase fragment of malignant catarrhal fever virus strain-wc11. *Arch. Virol.*, v.113, p.53-60, 1990.

- HYPERCHEM 4.0. HyperCube, Inc., 419 Phillip St, Waterloo, Ontario, Canada, N2L3X2, 1995.
- IRMIERE, A.F.; MANOS, M.M.; JACOBSON, J.G.; GIBBS, J.S.; COEN, D.M. Effect of an amber mutation in the herpes-simplex virus thymidine kinase gene on polypeptide-synthesis and stability. *Virology*, v.168, p.210-220, 1989.
- JUEL-JENSEN, B.E.; MACCALLUM, F.O. *Herpes simplex, varicella and zoster: Clinical manifestations and treatment*. Philadelphia, J.B. Lippincott Co., 1972.
- KARELSON, M.; LOBANOV, V.S.; KATRITZKY, A.R. Quantum-chemical descriptors in QSAR/QSPR studies. *Chem. Rev.*, v.96, p.1027-1043, 1996.
- KARFUNKEL, H.R.; EYRAUD, V. An algorithm for the representation and computation of supermolecular surfaces and volumes. *J. Comp. Chem.*, v.10, p. 628-634, 1989.
- KATRITZKY, A.R.; KARELSON, M. AM1 calculations of reaction field effects on the tautomeric equilibria of nucleic acid pyrimidine and purine bases and their 1-methyl analogues. *J. Am. Chem. Soc.*, v.113, p.1561-1566, 1991.
- KAUFMAN, H.E.; VARNELL, E.D.; WRIGHT, G.E.; XU, H.; GEBHARDT, B.M.; THOMPSON, H.W. Effect of 9-(4-hydroxybutyl)-N²-phenylguanine (HBPG), a thymidine kinase inhibitor, on clinical recurrences of ocular herpetic keratitis in squirrel monkeys. *Antivir. Res.*, v.33, p. 65-72, 1996.
- KIT, S.; KIT, M.; QAVI, H.; TRKULA, D.; OTSUKA, H. Nucleotide-sequence of the herpes-simplex virus type-2 (HSV-2) thymidine kinase gene and predicted amino-acid-sequence of thymidine kinase polypeptide and its comparison with the HSV-1 thymidine kinase gene. *Biochim. Biophys. Acta.*, v.741, p.158-170, 1983.
- KUBINYI, H. Quantitative structure-activity relationships. 7. The bilinear model, a new model for nonlinear dependence of biological activity on hydrophobic character. *J. Med. Chem.*, v.20, p.625-629, 1977.

- KUBINYI, H. QSAR: Hansch analysis and related approaches. In: *Methods and principles in medicinal chemistry*. MANNHOLD, R.; KROGSGAARD-LARSEN, P.; TIMMERMAN, H., eds., v.1, Weinheim, VCH Publishers, 1993, pp.240.
- KUBINYI, H. Variable selection in QSAR studies. I. An evolutionary algorithm. *Quant. Struct.-Act. Relat.*, v.13, p.285-294, 1994a.
- KUBINYI, H. Variable selection in QSAR studies. II. A highly efficient combination of systematic search and Evolution. *Quant. Struct.-Act. Relat.*, v.13, p.393-401, 1994b.
- KURTZ, H.A.; STEWART, J.J.P.; DIETER, K.M. Calculation of the nonlinear optical properties of molecules. *J. Comp. Chem.*, v.11, p.82-87, 1990.
- KWIATKOWSKI, J.S.; LESZCZYŃSKI, J. An *ab initio* quantum-mechanical study of tautomerism of purine, adenine and guanine. *J. Mol. Struct. (Theochem)*, v.208, p.35-44, 1990.
- LEACH, A.R. *Molecular modeling*. Principles and applications. Singapore, Longman Singapore Publishers, 1996, pp.595.
- LESZCZYŃSKI, J. Electron correlation effects in the *ab initio* study on tautomerism of guanine. *Chem. Phys. Lett.*, v.174, p.347-354, 1990.
- LEVY, J.A.; FRAENKEL-CONRAT, H.; OWENS, R.A. *Virology*. 3.ed., New Jersey, Prentice-Hall, 1994, p.192-198.
- LEWIS, D.F.V. The calculation of molar polarizabilities by the CNDO/2 method: Correlation with the hydrophobic parameter, log P. *J. Comp. Chem.*, v.10, p.145-151, 1989.
- MACKERELL, JR., A.D.; WIÓRKIEWICZ-KUCZERA, J.; KARPLUS, M. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.*, v.117, p.11946-11975, 1995.

- MARTIN, S.L.; APARISIO, D.I.; BANDYOPADHYAY, P.K. Genetic and biochemical-characterization of the thymidine kinase gene from herpesvirus of turkeys. *J. Virol.*, v.63, p.2847-2852, 1989.
- MARTIN, Y. *Quantitative Drug Design: A Critical Introduction*. New York, Marcel Dekker, 1978.
- MARTIN, Y. A practioner's perspective of the role of quantitative structure-activity analysis in medicinal chemistry. *J. Med. Chem.*, v.24, p.229-237, 1981.
- MATHEMATICA, Enhanced Version 2.2, Wolfram Research, Inc, Champaign, IL, 1993.
- MATHEWS, C.K.; VAN HOLDE, K.E. *Biochemistry*, 2.ed., California, Benjamin/Cummings, 1996, p.77-80.
- MCGEOCH, D.J.; DALRYMPLE, M.A.; DAVISION, A.J.; DOLAN, A.; FRAME, M.C.; MCNAB, D.; PERRY, L.J.; SCOTT, J.E.; TAYLOR, P. The complete DNA-sequence of the long unique region in the genome of herpes-simplex virus type-1. *J. Gen. Virol.*, v.69, p.1531-1574, 1988.
- MIERTUS, S.; SCROCCO, E.; TOMMASI, J. Electrostatic interaction of a solute with a continuum. A direct utilization of *ab initio* molecular potentials for the provision of solvent effects. *Chem. Phys.*, v.55, p.117-129, 1981.
- MITTAL, S.K.; FIELD, H.J. Analysis of the bovine herpesvirus type-1 thymidine kinase (TK) gene from wild-type virus and TK-deficient mutants. *J. Gen. Virol.*, v.70, p.901-918, 1989.
- MORRISON, R.T.; BOYD, R.N. *Organic Chemistry*. 6.ed., New Jersey, Prentice Hall, 1992, pp.1325.
- MÜLLER, C.W.; SCHULZ, G.E. Structure of the complex between adenylate kinase from escherichia-coli and the inhibitor Ap5A refined at 1.9 a resolution - a model for a catalytic transition-state. *J. Mol. Biol.*, v.224, p.159-177, 1992.

- MUNIR, K.M.; FRENCH, D.C.; DUBE, D.K.; LOEB, L.A. Permissible amino-acid substitutions within the putative nucleoside-binding site of herpes-simplex virus type-1 established by random sequence mutagenesis. *J. Biol. Chem.*, v.267, p.6584-6589, 1992.
- MURRAY, P.R.; DREW, W.L.; KOBAYASHI, G.S.; THOMPSON, J.H. *Microbiologia Médica*. Rio de Janeiro, Guanabara Koogan, 1992.
- NICOLSON, L.; CULLINANE, A.A.; ONIONS, D.E. The nucleotide-sequence of the equine herpesvirus-4 thymidine kinase gene. *J. Gen. Virol.*, v.71, p.1801-1805, 1990.
- NUNBERG, J.H.; WRIGHT, D.K.; COLE, G.E.; PETROVSKIS, E.A.; POST, L.E.; COMPTON, T.; GILBERT, J.H. Identification of the thymidine kinase gene of feline herpesvirus - use of degenerate oligonucleotides in the polymerase chain-reaction to isolate herpesvirus gene homologs. *J. Virol.*, v.63, p.3240-3249, 1989.
- O'BRIEN, E.J. Crystal structures of two complexes containing guanine and cytosine derivatives. *Acta Cryst.*, v.23, p.92-106, 1967.
- ORCHIN, M.; KAPLAN, F.; MACOMBER, R.S.; WILSON, R.M.; ZIMMER, H. *The vocabulary of organic chemistry*. New York, John Wiley & Sons, 1980, p.120-121.
- OROZCO, M.; LUQUE, F.J. On the use of AM1 and MNDO wave-functions to compute accurate electrostatic charges. *J. Comp. Chem.*, v.11, p.909-923, 1990.
- OSMAN, R.; WEINSTEIN, H.; TOPIOL, S.; RUBENSTEIN, L. A molecular theory of recognition and activation at a 5-HT receptor based on a quantum chemical approach to structure activity relationships. *Clin. Physiol.*, v.3, p.80-88, 1985.
- OSMAN, R.; TOPIOL, S.; RUBENSTEIN, L.; WEINSTEIN, H. A molecular-model for activation of a 5-hydroxytryptamine receptor. *Mol. Pharmacol.*, v.32, p.699-705, 1987.

- OTSUKA, H.; KIT, S. Nucleotide-sequence of the marmoset herpesvirus thymidine kinase gene and predicted amino-acid-sequence of thymidine kinase polypeptide. *Virology*, v.135, p.316-330, 1984.
- PARR, R.G.; PEARSON, R.G. Absolute hardness: Companion parameter to absolute electronegativity. *J. Am. Chem. Soc.*, v.105, p.7512-7516, 1983.
- PASCUAL-AHUIR, J.L.; SILLA, E. GEPOL: an improved description of molecular surfaces. I. Building the spherical surface set. *J. Comp. Chem.*, v.11, p. 1047-1060, 1990.
- PIRES, J.M.; FLORIANO, W.B.; GAUDIO, A.C. Extension of the frontier reactivity indices to groups of atoms and application to quantitative structure-activity relationship studies. *J. Mol. Struct. (Theochem)*, v.389, p.159-167, 1997.
- QUANTA-CHARMM, v.4.0, Molecular Simulations Inc., San Diego, CA, 1996.
- REDL, G.; CRAMER III, R.D.; BERKOFF, C.E. Quantitative drug design. *Chem. Soc. Rev.*, v.3, p.273-293, 1974.
- RICHMOND, T. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.*, v.178, p. 63-89, 1984.
- ROBERTSON, G.R.; WHALLEY, J.M. Evolution of the herpes thymidine kinase - identification and comparison of the equine herpesvirus-1 thymidine kinase gene reveals similarity to a cell-encoded thymidylate kinase. *Nucleic Acids Res.*, v.16, p.11303-11317, 1988.
- ROIZMAN, B. Herpesviridae: A Brief Introduction. In: FIELDS, B.N.; KNIPE, D.M.; CHANOCK, R.M.; HIRSCH, M.S.; MELNICK, J.L.; MONATH, T.P.; ROIZMAN, B., eds., *Fundamental Virology*. 2.ed., New York, Raven Press, 1991, p.841-847.

- ROIZMAN, B.; CARMICHAEL, L.E.; DEINHARDT, F. Herpesviridae - definition, provisional nomenclature, and taxonomy. *Intervirology*, v.16, p.201-217, 1981.
- SABIO, M.; TOPIOL, S.; LUMMA, JR., W.C. An investigation of tautomerism of adenine and guanine. *J. Phys. Chem.*, v.94, p.1366-1372, 1990.
- SACHS, L. *Applied statistics: A handbook of techniques*. 2.ed., New York, Springer Verlag, 1984.
- SANDERSON, M.R.; FREEMONT, P.S.; MURTHY, H.M.K.; KRANE, J.F.; SUMMERS, W.C.; STEITZ, T.A. Purification and crystallization of thymidine kinase from herpes-simplex virus type-1. *J. Mol. Biol.*, v.202, p.917-919, 1988.
- SCHMIDT, M.W.; BALDRIDGE, K.K.; BOATZ, J.A.; ELBERT, S.T.; GORDON, M.S.; JENSEN, J.H.; KOSEKI, S.; MATSUNAGA, N.; NGUYEN, K.A.; SU, S.J.; WINDUS, T.L.; DUPUIS, M.; MONTGOMERY, J.A. General atomic and molecular electronic structure system. *J. Comp. Chem.*, v.14, p.1347-1363, 1993.
- SCHULZ, G.E.; ELZINGA, M.; MARX, F.; SCHIRMER, R.H. Three-dimensional structure of adenil kinase. *Nature*, v.250, p.120-123, 1974.
- SCHWARTZ, R.M.; DAYHOFF, M.O. Improved scoring matrix for identifying evolutionary relatedness among proteins. *Biophys. J.*, v.21, p.A198-A198, 1978.
- SCOTT, S.D.; ROSS, N.L.J.; BINNS, M.M. Nucleotide and predicted amino-acid sequences of the mareks-disease virus and turkey herpesvirus thymidine kinase genes - comparison with thymidine kinase genes of other herpesviruses. *J. Gen. Virol.*, v.70, p.3055-3065, 1989.
- SCROCCO, E.; TOMASI, J. The electrostatic molecular potential as a tool for the interpretation of molecular properties. *Topics in Curr. Chem.*, v.42, p.95-170, 1973.
- SHEPPARD, M.; MAY, J.T. Location and characterization of the bovine herpesvirus type-2 thymidine kinase gene. *J. Gen. Virol.*, v.70, p.3067-3071, 1989.

- SIGULA, A.; BUDA, A. MNDO study of the the tautomers of nucleic acids. Part II. Adenine and guanine. *J. Mol. Struct. (Theochem)*, v.92, p.267-277, **1983**.
- SMITH, G.A.; YOUNG, P.L.; MATTICK, J.S. The location and nucleotide-sequence of the thymidine kinase gene of bovine herpesvirus type-1.2. *J. Gen. Virol.*, v.71, p.2417-2424, **1990**.
- SMITH, G.A.; YOUNG, P.L.; MATTICK, J.S. Nucleotide and amino-acid-sequence analysis of the thymidine kinase gene of a bovine encephalitis herpesvirus. *Arch. Virol.*, v.119, p.199-210, **1991**.
- SOLOMONS, T.W.G. *Organic Chemistry*. 6.ed., New York, John Wiley & Sons, **1996**, pp.1218.
- STEWART, J.J.P. Optimization of parameters for semi-empirical methods. I. Method. *J. Comp. Chem.*, v.10, p.209-220, **1989a**.
- STEWART, J.J.P. Optimization of parameters for semi-empirical methods. II. Applications. *J. Comp. Chem.*, v.10, p.221-264, **1989b**.
- STEWART, J.J.P. MOPAC: A semiempirical molecular orbital program. *J. Comp.-Aid. Mol. Des.*, v.4, p.1-105, **1990a**.
- STEWART, J.J.P MOPAC manual. 6.ed., Colorado Springs, US Air Force Academy, **1990b**, p.6-6-6-8.
- STREITWIESER, A.; HEATHCOCK, C.H.; KOSOWER, E.M. *Introduction to Organic Chemistry*. 4.ed., New York, Macmillan, **1992**, pp.1218.
- STRYER, L. *Biochemistry*. 3.ed., New York, W.H. Freeman & Co., **1988**, p.675-677.
- SUCK, D.; OEFNER, C. Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature*, v.321, p.620-625, **1986**.
- SWAIN, M.A.; GALLOWAY, D.A. Nucleotide-sequence of the herpes-simplex virus type-2 thymidine kinase gene. *J. Virol.*, v.46, p.1045-1050, **1983**.

- SWAIN, C.G.; LUPTON JR, E.C. Field and resonance components of substituent effects. *J. Am. Chem. Soc.*, v.90, p.4328-4337, 1968.
- SYBYL, Trypos Associates, Inc., 1991.
- SZABO, A.; OSTLUND, N.S. *Modern quantum chemistry: Introduction to advanced electronic structure theory*. New York, Dover, 1996, pp.466.
- TAFT, R.W. Polar and steric substituent constants for aliphatic and o-benzoate groups from rates of esterification and hydrolysis of esters. *J. Am. Chem. Soc.*, v.74, p.3120-3128, 1952.
- TELFORD, E.A.R.; WATSON, M.S.; MCBRIDE, K.; DAVISON, A.J. The DNA-sequence of equine herpesvirus-1. *Virology*, v.189, p.304-316, 1992.
- THEWALT, U.; BUGG, C.E.; MARSH, R.E. The crystal structure of guanosine dihydrate and inosine dihydrate. *Acta Cryst.*, v.B26, p.1089-1101, 1970.
- THEWALT, U.; BUGG, C.E.; MARSH, R.E. The crystal structure of guanine monohydrate. *Acta Cryst.*, v.B27, p.2358-2363, 1971.
- TOPAL, M.D.; FRESCO, J.R. Complementary base pairing and the origin of substitution mutations. *Nature*, v.263, p.285-289, 1976.
- TOPIOL, S. The deletion model for the origin of receptors. *Trends Biochem. Sci.*, v.12, p.419-421, 1987.
- TOPIOL, S.; WEINSTEIN, H.; OSMAN, R. A theoretical investigation of histamine tautomerism. *J. Med. Chem.*, v.27, p.1531-1534, 1984.
- TOPLISS, J.G.; COSTELLO, R.J. Chance correlations in structure-activity studies using multiple regression analysis. *J. Med. Chem.*, v.15, p.1066-1068, 1972.
- TUTE, M.S. Principles and practice of Hansch analysis: a guide to structure-activity correlation for the medicinal chemist. *Adv. Drug Res.*, v.6, p.1-77, 1971.

- VERLET, L. Computer 'experiments' on classical fluids. I. Thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, v.159, p.69-95, 1967.
- VOET, D.; VOET, J.G. *Biochemistry*. 2.ed., New York, John Wiley & Sons Inc., 1995.
- WAGNER, M.J.; SHARP, J.A.; SUMMERS, W.C. Nucleotide-sequence of the thymidine kinase gene of the herpes-simplex virus type-1. *Proc. Natl. Acad. Sci. U.S.A.-Biol.*, v.78, p.1441-1445, 1981.
- WALTERS, P.; STAHL, M. BABEL 1.6. University of Arizona, 1996.
- WATSON, J.D.; CRICK, F.H.C. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature*, v.171, p.737-738, 1953.
- WEINSTEIN, H.; MAZUREK, A.P.; OSMAN, R.; TOPIOL, S. Theoretical studies on the activation mechanism of the histamine H₂-receptor: The proton transfer between histamine and a receptor model. *Mol. Pharmacol.*, v.29, p.28-33, 1986.
- WILD, K.; BOHNER, T.; AUBRY, A.; FOLKERS, G.; SCHULZ, G.E. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett.* v.368, p.289-292, 1995.
- WILLIAMS, D.E.; YAN, J.M. Point-charge models for molecules derived from least-squares fitting of the electrical potential. *Adv. Atom. Mol. Phys.*, v.23, p.87-131, 1987.
- WINTER, H.D.; HERDEWIJN, P. Understanding the binding of 5-substituted 2'-deoxythymidine substrates to thymidine kinase of herpes simplex virus type-1. *J. Med. Chem.*, v.39, p.4727-4737, 1996.
- XU, H.; MAGA, G.; FOCHER, F.; SMITH, E.R.; SPADARI, S.; GAMBINO, J. WRIGHT, G.E. Synthesis, properties and pharmacokinetic studies of N²-phenylguanine derivatives as inhibitors of herpes simplex virus thymidine kinases. *J. Med. Chem.*, v.38, p.49-57, 1995.

YAN, H.G.; TSAI, M.-D. Mechanism of adenylate kinase. Demonstration of a functional-relationship between aspartate-93 and Mg^{2+} by site-directed mutagenesis and proton, P-31, and magnesium-25 nmr. *Biochemistry*, v.30, p.5539-5546, 1991.

YORK, J.L. Enzymes: Classification, Kinetics and Control. In: *Textbook of Biochemistry with Clinical Correlations*, 3.ed., Devlin, T.M., ed., New York, John Wiley & Sons, 1992.