



**Universidade Estadual de Goiás  
Unidade Universitária de Ciências Exatas e Tecnológicas**

Lilian Tatiane Ferreira de Melo Camargo

**Estudo Químico-Quântico da Relação Estrutura Atividade da  
Indolo [2,1b] Quinazolina e seus derivados análogos contra o  
câncer de mama**

Orientador: Professor Dr. Ademir João Camargo

Mestrado  
Ciências Moleculares  
Anápolis – 2008

Lilian Tatiane Ferreira de Melo Camargo

Estudo Químico-Quântico da Relação  
Estrutura Atividade da Indolo [2,1b]  
Quinazolina e seus derivados análogos  
contra o câncer de mama. [Anápolis] 2008  
XIV, 68p. 29,7 cm (UnUCET/UEG, MSc.,  
Ciências Moleculares, 2008) Dissertação  
– Universidade Estadual de  
Goiás, UnCET  
1. Câncer de mama  
2. Mecânica Quântica  
3. DFT  
4. PCA  
I. UnUCET/UEG II. Título (série)

Estudo Químico-Quântico da Relação Estrutura Atividade da Indolo  
[2,1b] Quinazolina e seus derivados análogos contra o câncer de  
mama

Lilian Tatiane Ferreira de Melo Camargo

Dissertação apresentada ao corpo docente do Programa de Mestrado em  
Ciências Moleculares da Universidade Estadual de Goiás como parte dos  
requisitos necessários para a obtenção do título de Mestre em Ciências  
Moleculares.

Aprovado por:

\_\_\_\_\_  
Dr. (Orientador)

\_\_\_\_\_  
Dr.

\_\_\_\_\_  
Dra.

ANÁPOLIS  
Março/2008

A Deus pela oportunidade a mim concedida  
a minha família e, especialmente, ao meu  
esposo Eduardo Francisco Camargo, pela  
compreensão e apoio durante a realização  
deste trabalho

## AGRADECIMENTO

Agradeço primeiramente a Deus, pois acredito que a realização deste trabalho é devida a Sua permissão, porque Dele, por Ele e para Ele são todas as coisas.

Agradeço também à Unidade Universitária de Ciências Exatas e Tecnológicas, pelo apoio institucional e às facilidades a mim concebidas.

Meus sinceros agradecimentos ao professor Dr. Ademir João Camargo pela compreensão e dedicação na orientação dada.

A todos os professores que fazem parte deste programa de pós-graduação, pelo incentivo.

Aos Colegas, José Maria Maia, Guilhermina, Valter Henrique, Sergio, Vítor, Wellington e Eduardo, pela paciência e apoio.

A minha família pelo incentivo.

A meu esposo Eduardo Francisco Camargo, pelo amor, carinho e compreensão.

A todos que, de alguma forma, contribuíram para realização deste trabalho.

## **CURRICULUM VITAE**

### **Dados pessoais:**

**Nome:** Lilian Tatiane Ferreira de Melo

### **Formação acadêmica:**

2004 - Graduação em Química Licenciatura. Universidade Estadual de Goiás, UEG, Brasil.

### **Formação Complementar:**

2001 - Princípios de Toxicologia Aplicada. (Carga horária: 20h).  
Universidade Estadual de Goiás, UEG, Brasil.

2003 - Utilização de Plantas no Cerrado. (Carga horária: 40h).  
Universidade Estadual de Goiás, UEG, Brasil.

2003 - Biotecnologia Enzimática. (Carga horária: 40h). Universidade  
Estadual de Goiás, UEG, Brasil.

2006 Extensão universitária em Comunicação em sala de aula. (Carga  
horária:300h).

Unidade da Fundação Brasileira de educação, FUBRAE, Brasil.

2006 Extensões universitárias em Aceleração da aprendizagem no  
ensino fundamental. (Carga horária: 320h). Unidade da Fundação  
Brasileira de educação, FUBRAE, Brasil.

2006 Extensão universitária em O desafio da redação profissional.  
Unidade da Fundação Brasileira de educação, FUBRAE, Brasil.

2006 - Extensão universitária em Avaliação na alfabetização. (Carga  
horária:280h).

Unidade da Fundação Brasileira de educação, FUBRAE, Brasil.

### **Participação em eventos:**

I Encontro de Química. 2001. (Participações em eventos/Encontro).

Encontro de Química da Universidade Estadual de Goiás. 2001.  
(Participações em eventos/Encontro).

II Encontro de Química. 2002. (Participações em eventos/Encontro). III  
Encontro de Química. 2003. (Participações em eventos/Encontro).

Simpósio Brasileiro de Química Teórica. Estudo Teórico químico quântico  
da relação estrutura atividade da indolo[2,1b] quinazoline contra o  
câncer de mama. 2007. (Participações em eventos/Congresso).

Semana da química. Qsar e a Produção de novos Farmacos. 2008.  
(Participações em eventos/Encontro).

Simpósio de Estrutura Eletônica e Dinâmica Molecular. Estudo Químico  
Quântico Teórico da relação estrutura atividade da Piperina e seus  
derivados análogos. 2008. (Participações em eventos/Simpósio).

#### **Atuação Profissional:**

**2003** - Vínculo: Monitora, Enquadramento Funcional: Monitora de  
Química Orgânica I, Carga horária: 4

**2005** - Vínculo: Servidor Público, Enquadramento Funcional: Professora,  
Carga horária: 20

**2005/2008**: Colégio Estadual Senador Onofre Quinam- Disciplina  
ministrada Química. Carga horária: 14

**2008**: Universidade Estadual de Goiás (UEG) - Disciplina ministradas  
Físico - Química Experimental I e II. Carga horária: 16

**2008**: Faculdade do Instituto Brasil ( FIBRA) - Disciplina ministradas  
Físico - Química Teórica e Experimental. Carga horária: 03

#### **Produção Bibliográfica:**

1. MELO, L. T. F. ; Camargo.A.J ; Gilberto.L.B Aquino . Estudo Químico  
quântico da relação estrutura atividade da Piperina e seus derivados  
análogos na inibição de enzimas do citocromo P450.. 2008.  
(Apresentação de Trabalho/Simpósio).
2. MELO, L. T. F. ; Camargo.A.J . Estudo Teórico químico quântico da  
relação estrutura e atividade da indolo[2,01b] quinazoline e seus  
análogos contra o câncer de mama.. 2007. (Apresentação de  
Trabalho/Simpósio).

## RESUMO

As atividades biológicas da Indolo[2,1b]quinazolina e de trinta e três derivados análogos sintetizados foram testadas *in vitro* contra o câncer de mama humano. Os resultados dos testes mostraram que vinte e dois destes derivados apresentam atividade contra as células cancerosas, enquanto que os demais mostraram-se inativos. Com o objetivo de relacionar os parâmetros físico-químicos destes derivados com a atividade anticancerígena, cálculos de mecânica quântica molecular foram realizados usando a Teoria do Funcional da Densidade (DFT) com o funcional de troca e correlação híbrido B3LYP e com o conjunto de base 6-31G\* implementados no pacote de programa Gaussian 03. Os descritores obtidos nos cálculos teóricos foram relacionados com a atividade anticancerígena usando métodos da estatística multivariada implementados no programa quimiométrico Unscrambler 9.1. A Análise de Componentes Principais (PCA) mostrou que cinco descritores eletrônicos são suficientes para separar os compostos em duas classes: ativos e inativos. A primeira componente foi a responsável pela separação, a qual descreveu 50% da variância total dos dados. Com base nos descritores selecionados é possível sugerir novos compostos para síntese cujas atividades biológicas sejam superiores.



## ABSTRACT

The biological activities of the Indolo[2,1b]quinalozine and thirty and three analogue compounds were synthesized and tested *in vitro* against the human breast cancer cells. The assays show that twenty and two of these compounds exhibited cytotoxic activity while other compounds are not cytotoxic. With the aim to correlate the physical chemical parameters of those compounds with the anticancer activity, quantum molecular calculations were carried out using the Density Functional Theory (DFT) with the exchange correlation functional B3LYP and the basis set 6-31G\* implemented in Gaussian 03 suite of program. The descriptors obtained from the theoretical calculations were related to the anticancer activity using the multivariate statistical methods implemented in the Unscrambler 9.1 chemometrics program. The Principal Component Analysis (PCA) showed that five electronic descriptors are able to separate the compounds into two classes: active and inactive. The first component alone was responsible for the separation and describes 50% of the total variance of the original data. From the selected descriptors it is possible rationally to design new compounds for synthesis that can be more active against breast cancer cells.

<b>CAPÍTULO I</b> .....	<b>1</b>
1.1 <i>Introdução</i> .....	1
1.2 <i>DERIVADOS ANÁLOGOS DA INDOLO [ 2,1b] QUINAZOLINA</i> .....	2
1.2 <i>CÂNCER</i> .....	5
1.2 <i>CÂNCER DE MAMA</i> .....	6
<b>CAPÍTULO II</b> .....	<b>8</b>
<b>METODOLOGIA</b> .....	<b>8</b>
1.1 <i>EQUIPAMENTOS E SOFWARES</i> .....	9
1.2 <i>PROCEDIMENTO COMPUTACIONAL</i> .....	9
1.2.1 <i>Escolha da conformação mais adequada para levantamento das propriedades físico-químico e geométrica</i> .....	9
1.2.2 <i>Otimizações das moléculas</i> .....	10
1.2.3 <i>Descritores geométricos e eletrônicos calculados usando o pacote programa gaussian 03 com nível de teoria b3lyp/6-31G*</i> .....	10
1.3 <i>ANÁLISE ESTATÍSTICA DOS DADOS</i> .....	10
<b>CAPÍTULO III</b> .....	<b>12</b>
<b>FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>12</b>
1.1 <i>INTRODUÇÃO</i> .....	12
1.2 <i>APROXIMAÇÃO DE BORH-OPPENHEIMER</i> .....	13
1.3 <i>MÉTODO DA APROXIMAÇÃO DE HARTREE – FOCK (HF)</i> .....	14
1.4 <i>AS EQUAÇÕES DE HARTREE - FOCK - Roothan</i> .....	16
1.5 <i>TEORIA DO FUNCIONAL DA DENSIDADE</i> .....	16
1.5.1 <i>O modelo de Thomas-Fermi (TF)</i> .....	17
1.5.2 <i>Os teoremas de Hohenberg e kohn</i> .....	20
1.5.3 <i>O Método de Kohn - Sham</i> .....	22
1.5.4 <i>Aproximação LDA ( Local - Density Approximation)</i> .....	25
1.5.5 <i>Aproximação LSDA ( Local – Spin - Density Approximation</i> .....	27
1.5.6 <i>GGA ( Generalized - Gradiente Approximation )</i> .....	28

III.5.7 Aproximação híbrida .....	29
III.6 METODO DE ESTATÍSTICA MULTIVARIADA DE ANÁLISE .....	31
III.6.1 Matriz de média e matriz de covariância e correlação .....	32
III.6.2 escalamento dos dados .....	34
III.6.3 Pesos de Fisher .....	35
III.6.4 Análise de Componentes Principais .....	36
<b>CAPÍTULO IV.....</b>	<b>38</b>
RESULTADOS E DISCUSSÃO .....	38
IV.1 ANÁLISE CONFORMACIONAL .....	38
IV.2 ANÁLISE DAS PROPRIEDADES CALCULADAS .....	38
IV.2.1 Energia dos orbitais moleculares de fronteira .....	39
IV.2.2 Gap ( HOMO – LUMO ) .....	39
IV.2.3 Dureza( .....	39
IV.2.4 Eletronegatividade (.....	41
IV.2.5 Cargas atômicas parciais ( C ) .....	41
IV.2.6 Ordens de ligação ( b ) .....	45
IV.2.7 Ângulos ( a ) .....	50
IV.2.8 Ângulos diedrais ( d ) .....	53
IV.2.9 Volume molar ( V ) .....	54
IV.2.10.Coefficiente de Partição ( LogP ) .....	54
IV.2.11 Momento de dipolo ( ) .....	54
IV.2.9 Volume molar ( V ) .....	54
IV.2.10.Coefficiente de Partição ( LogP ) .....	54
IV.2.11 Momento de dipolo ( ) .....	54
IV.2.12 polarizabilidade (.....	57
IV.3. ANÁLISE ESTATÍSTICA DOS DADOS .....	57
IV.3.1Análise da atividade contra o câncer de mama .....	57
<b>CAPÍTULO V.....</b>	<b>65</b>

<i>CONCLUSÃO</i> .....	65
<i>REFERÊNCIAS BIBLIOGRÁFICAS</i> .....	66

## ÍNDICE DE FIGURA

Figura I.1 Estrutura do Tryptanthrine e Bactracyclin.....	2
Figura I.2 Indolo [2,1b] quinazoline e seus derivados análogos .....	4
Figura I.3 Derivados análogos da indolo [2,1b] quinazolina.....	5
Figura IV.1 Estrutura base das indolos [2,1b] quinazoline e seus derivados análogos .....	38
Figura IV.2 Representação gráfica dos scores obtidos usando o programa Unscrambler 9.1 .....	59
Figura IV.3 Representação gráfica dos loading obtidos usando o programa Unscrambler 9.1 .....	60
Figura IV.4 Representação gráfica dos loading score obtidos usando o programa Unscrambler 9.1 .....	62
Figura IV.4 Representação da formação de carbocátion devido a adição de um grupo retirador de elétrons, obtidos pelo programa Unscrambler 9.1 .....	64

## ÍNDICE DE TABELA

Tabela I.1 Valor de GI <sub>50</sub> da indolo[2,1b] quinazoline e seus derivados análogos contra o câncer de mama. Fonte[6].....	8
Tabela IV.1 valores de energia dos orbitais moleculares de fronteira, Gap( $\epsilon$ ), dureza( $\eta$ ) e eletronegatividade ( $\chi$ ) obtidas a nível de teoria B3LYP/6-31G*. todos valores são dados em eV.....	40
Tabela IV.2 Carga atômica ( $q$ ), derivadas do potencial eletrostático com método CHELPG, a nível de teoria B3LYP/6-31G*. Todos valores são dados em eV.....	42
Tabela IV.3 Descrição de ordem de ligação obtidas a nível de teoria B3LYP/6-31G* .....	46
Tabela IV.4 Ângulos obtidas a nível de teoria B3LYP/6-31G*. Todos valores são dados em graus ( $^{\circ}$ ).....	51
Tabela IV.5 Ângulos diedrais obtidas a nível de teoria B3LYP/6-31G*. Todos valores são dados em graus ( $^{\circ}$ ).....	51
Tabela IV.6 Momento de dipolo ( $\mu$ ); obtido a nivel teoria B3LYP/6-31G*. O Volume( $V$ ), coeficiente de partição ( $\log P$ ) e Polarizabilidade ( $\alpha$ ) foram calculados pela rotina do pacote Hyperchem 7.5.Todos valores são dados em graus ( $^{\circ}$ ).....	56
Tabela IV.7 Peso de Fisher.....	58
Tabela IV.8 Porcentagem de variância explicada pelas Componentes Principais obtidos pelo programa Unscrambler 9.1 .....	60
Tabela IV.9 loading das variáveis para as 2 primeiras componentes obtidos pelo programa Unscrambler 9.1 .....	61

## **CAPÍTULO I**

### **CONSIDERAÇÕES BÁSICAS**

#### **1.1 INTRODUÇÃO**

Segundo dados publicados pela University's Center for the Study of Drugs Development [1], contando com equipes multidisciplinares de pesquisa e desenvolvimento, um investimento de aproximadamente 1,2 bilhões de dólares é aplicado no desenvolvimento de novos fármacos [1], com estimativa de 10 a 15 anos para a sua introdução no mercado. Esta busca por substâncias com atividade biológica vem desde tempos remotos, sendo que diversos métodos já foram desenvolvidos nesta tentativa, desde métodos místicos até o mais atual, que é fundamentado em parâmetros energéticos e estruturais [2].

Fatores estéricos, eletrônicos e hidrofóbicos das moléculas estão diretamente relacionados com a ação terapêutica do fármaco nos sistemas biológicos. Assim, mudar um átomo de posição ou trocar um substituinte em uma molécula pode ser fundamental para explicar a atividade biológica de uma substância, possibilitando a modelagem de novas drogas que sejam potencialmente mais ativas [3,4].

Com o progresso dos *hardwares* e o desenvolvimento de eficientes algoritmos computacionais, o método de Relação entre a Estrutura e a Atividade (SAR -Structure-Activity Relationships) que associe as análises de química quântica com as técnicas de quimiometria, tem sido utilizados vastamente para um estudo racional das propriedades físico-químicas e estruturais das moléculas, possibilitando a modelagem de novas drogas [5].

O estudo SAR visa fazer uma análise computacional, em relação à estrutura química dos compostos e sua atividade biológica, diminuindo os

gastos e o tempo de pesquisa na produção de novas drogas. Neste trabalho, foi analisado usando a técnica SAR, a atividade da indolo[2,1b]quinazolina e seus derivados análogos contra o câncer de mama.

## 1.2 DERIVADOS ANÁLOGOS DA INDOLO[2,1b]QUINAZOLINA

Do extrato da folha de uma planta medicinal da Índia (*Wrightia tinctoria*), foi isolada a indolo[2,1b]quinazolina (Tryptanthrine). O que tornou este composto atrativo para síntese e estudos biológico foi a similaridade de sua estrutura química com um excelente agente citostático, o batracyclin. (Ver Figura I.1) [6].

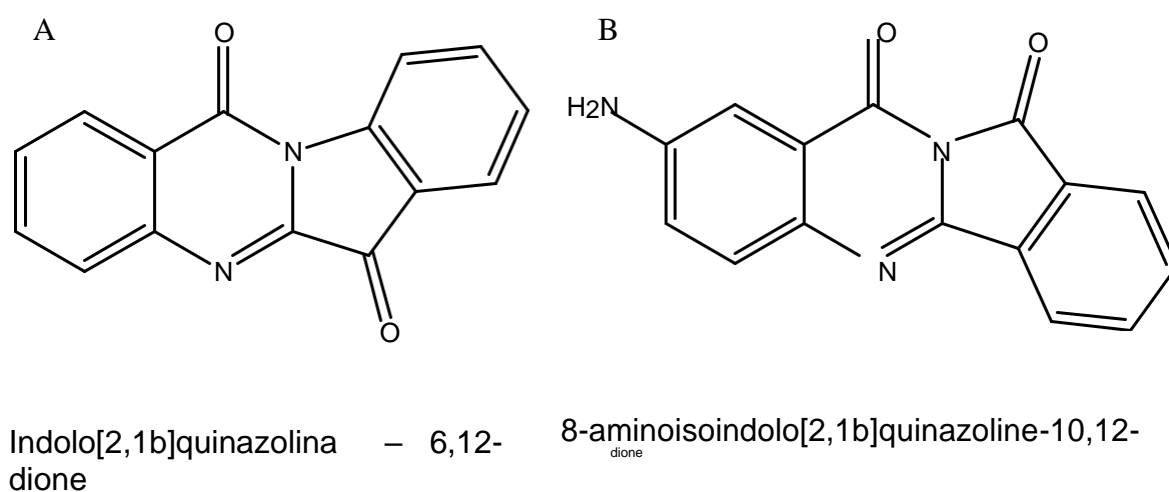
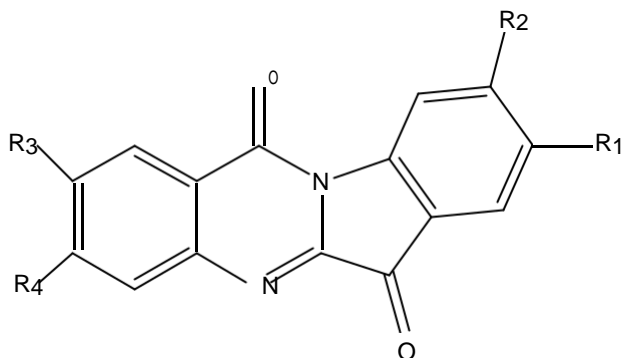


Figura I.1 Estrutura do Tryptanthrine e Bactracyclin

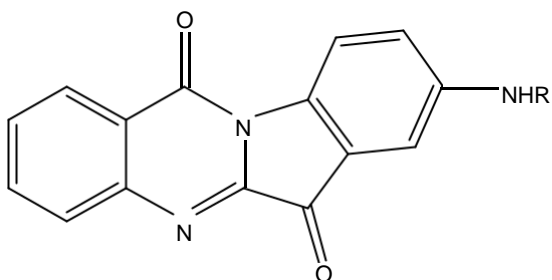
A partir da indolo[2,1b]quinazoline, Figura I.1, foram sintetizados alguns análogos; os quais foram testados *in vitro* e *in vivo* contra 12 tipos de câncer humano, sendo que alguns compostos análogos mostraram-se ativos contra

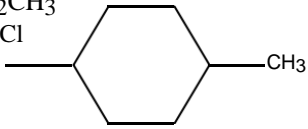


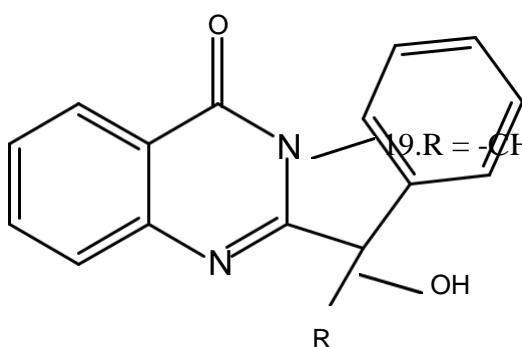
determinados tipo de câncer. A Figura I.2 e a Figura I.3 mostram a estrutura química da indolo[2,1b]quinazolina e seus derivados análogos [6].



1.  $R_1 = R_2 = R_3 = R_4 = H$
2.  $R_3 = Br; R_1 = R_2 = R_4 = H$
3.  $R_3 = NO_2; R_1 = R_2 = R_3 = R_4 = H$
4.  $R_3 = R_4 = OMe; R_1 = R_2 = H$
5.  $R_3 = NH_2; R_1 = R_2 = R_4 = H$
6.  $R_1 = Br; R_2 = R_4 = H$
7.  $R_1 = NO_2; R_2 = R_3 = R_4 = H$
8.  $R_1 = NH_2; R_2 = R_3 = R_4 = H$
9.  $R_1 = R_2 = -OCH_2-; R_3 = R_4 = H$
10.  $R_1 = OMe; R_2 = R_3 = R_4 = H$
11.  $R_1 = Cl; R_2 = R_3 = R_4 = H$

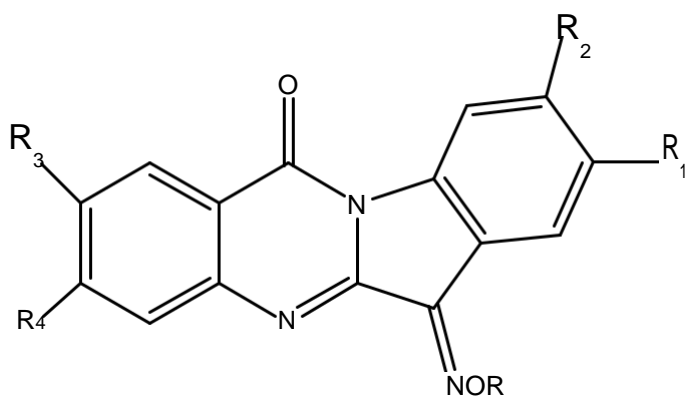


12.  $R = C(O)CH_3$
13.  $R = C(O)CH_2CH_2CH_3$
14.  $R = C(O)CH_2CH_3$
15.  $R = C(O)CH_2Cl$
16.  $R = C(O)CH_2$  
17.  $R = SO_2CH_3$



18.  $R = -CH_3$
19.  $R = -CH_2CH_2CH_2CH_3$
20.  $R = -OCH_2CH_2O-$   
Ph

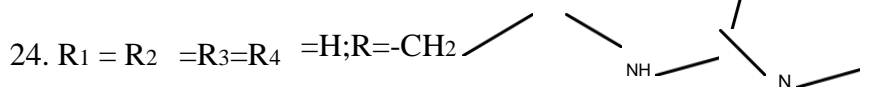
Figura I.2 Indolo [2,1b] quinazoline e seus derivados análogos



21.  $R_1=R_2=R_3=R_4=R=H$

22.  $R_1 = R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2OH$

23.  $R_1 = R_2 = R_3 = R_4 = H$ ;  $R = -CH_2COOCH_3$

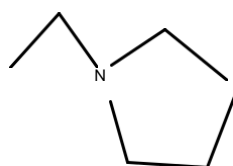


25.  $R_1 = R_2 = R_3 = R_4 = H$ ;  $R = -C(O)CH=CH-Ph$

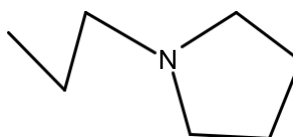
26.  $R_1 = R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2CH_2-NMe_2$

27.  $R_1 = R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2CH_2-NMe_2$

28.  $R_1=R_2=R_3=R_4=H$ ;  $R=-CH_2$



29.  $R_1=R_2=R_3=R_4=H$ ;  $R=-CH_2$



30.  $R_1 = NO_2$ ;  $R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2-NMe_2$

31.  $R_1 = NO_2$ ;  $R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2CH_2-NMe_2$

32.  $R_1 = NO_2$ ;  $R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2OH$

33.  $R_1 = Br$ ;  $R_2 = R_3 = R_4 = H$ ;  $R = -CH_2CH_2-NMe_2$

34.  $R_1 = Br$ ;  $R_2 = R_3 = R_4 = H$ ;  $R = -CH_2$

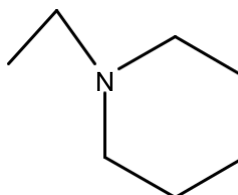


Figura I.3 Derivados análogos da indolo [2,1b] quinazolina

### 1.3 CÂNCER

O câncer é um termo genérico utilizado para designar mais de 100 doenças que podem invadir qualquer parte do corpo. Outros termos empregados para referir a estas doenças são: tumor e neoplasia. Uma característica do câncer é o crescimento descontrolado de células anormais capazes de invadir tecidos e órgãos vizinhos ou distantes. A migração do tumor para outras regiões do corpo é denominada de metástase [7].

A explicação bioquímica para ocorrência da neoplasia e sua proliferação é a alteração de um gene especial chamado protooncogene que, quando ativado, transforma-se em oncogene, responsável pela transformação das células. Estas células transformadas (células tumorais) multiplicam-se de forma descontrolada, levando à formação de novos vasos sanguíneos para sua nutrição. A manutenção e o acúmulo de massa dessas células formam os tumores malignos que podem adquirir a capacidade de migrar para outra parte do corpo. À medida que as células cancerosas vão substituindo as normais, os tecidos e órgãos invadidos vão perdendo suas funções [8].

Estatisticamente, em pesquisa realizada pela Organização Mundial de Saúde, o câncer foi responsável por 7,6 milhões (13%) de todas as mortes ocorridas em todo mundo, no ano de 2005. Sendo que 70% dos óbitos por câncer ocorreram em países de economia baixa ou média, devido às limitações no diagnóstico e tratamento da doença [7]. No Brasil, o número de morte por câncer em 2005 foi de 190 mil, sendo que 113 mil destas pessoas tinham menos de 70 anos de idade. A estimativa para 2020, em nível mundial, é 10 milhões de morte por câncer ao ano e aproximadamente de 30 milhões de portadores da doença [7].

As causas do câncer são muito variadas, podendo ser interna ou externa ao organismo. As causas externas estão relacionadas com o meio ambiente, social e cultural, tais como: tabagismo, hábitos alimentares, agrotóxicos, medicamentos, radiação solar, fatores ocupacionais, hábitos sexuais e alcoolismo. As causas internas, em geral são genéticas e pré-determinadas [7].

O processo de formação do tumor é lento e pode ser dividido em três estágios: estágio de iniciação que corresponde à alteração do gene da célula

por um agente cancerígeno; nesta etapa ainda não é possível detectar o tumor. O segundo estágio é o de promoção, em que as células alteradas são transformadas em células malignas de forma lenta. O último estágio é o de progressão, em que as células começam a multiplicar descontroladamente [8].

Os tratamentos para o câncer provocam muitos efeitos colaterais e não mostram 100% de eficácia. Os três mais tradicionais são: cirurgia para a retirada do tumor, radioterapia, que é a utilização de radiação de alta energia para bloquear o crescimento das células cancerígenas e evitar a metástase; e a quimioterapia, que é o uso de drogas potentes cujo objetivo principal é destruir as células malignas e conservar as células normais. [8]. Este tratamento provoca muitos efeitos colaterais como: anemia, diminuição da resistência a infecções, náuseas, vômitos, diarreia e queda de pêlos e cabelos [8].

#### *1.4 CÂNCER DE MAMA*

No mundo, milhares de mulheres são acometidos anualmente pelo câncer de mama, representando o tipo de neoplasia com maior número de mortalidade entre pessoas do sexo feminino. No ano de 2005, o câncer de mama foi responsável por 502 mil mortes no mundo [7].

O câncer de mama não tem causas totalmente esclarecidas, acredita-se que os principais fatores responsáveis pelo aumento de sua incidência sejam: a mudança no estilo de vida e reprodutivo das mulheres nos últimos anos e o aprimoramento do diagnóstico [9]. Ao citar mudanças no estilo de vida referem-se ao aumento do consumo de álcool por mulheres, maior índice de obesas, maior sedentarismo e maus hábitos alimentares. Quanto às mudanças reprodutivas, referem-se ao vasto uso de anticoncepcionais dos últimos anos, menarca precoces, reposição hormonal, redução do número de gestação e do tempo de amamentação [9,10].

O tratamento do câncer de mama, em geral, é feito com cirurgia que pode ser: quadrantectomia (uma parte da mama é retirada) ou mastectomia (retirada total da mama), quimioterapia e radioterapia. Empregam-se também alguns fármacos para o controle da doença, os mais comuns são: mostardas nitrogenadas, ciclofosfamida, ifosfamida, tiotepa, metotrexato, fluoracil, vimblastina, taxol, docetaxel, etoposida, teniposida, caproato de

hidroxiprogesterona, dietilestilbestrol, propionato de testosterona e tamoxifeno {2-[4-(1,2-difenil-but-1-enil)-fenóxi]-etil}-dimetil-amina que é o mais amplamente empregado, pertence ao grupo dos trifeniletilênicos e por isto apresenta várias atividades estrogênicas e antiestrogênicas[9].

Atualmente não existe cura para o câncer de mama, o tratamento disponível para o seu controle não tem excelente eficácia e, em geral, apresenta efeitos colaterais indesejáveis. Portanto deseja-se investigar teoricamente, usando os métodos da mecânica quântica molecular e da estatística multivariada, a relação entre as propriedades eletrônicas e estruturais da indolo[2,1-b]quinazoline e seus derivados análogos, com a sua atividade no combate ao câncer de mama[8,9].

## CAPÍTULO II

### METODOLOGIA

Pesquisadores da Índia sintetizaram a indolo[2,1b]quinazolina e seus derivados análogos e fizeram os testes *in vitro* contra várias linhas de câncer humano [6]. Compostos que apresentaram baixo índice de GI<sub>50</sub> (micro mol da substância necessária para inibir o crescimento de 50% das células cancerígenas) foram testados também *in vivo* usando o modelo Hollow Fibre Assay (HFA) em ratos albinos[6], que consiste em um método rápido para a triagem de agentes antitumorais. Neste método, as células cancerígenas são colocadas em capilares, que são implantadas por via subcutânea ou intraperitoneal em ratos anestesiados, de forma que um único animal é utilizado para fazer testes de diferentes linhagens de células cancerosas. Após a introdução do capilar, os animais são tratados com os compostos em estudo, que são capazes de difundir através dos capilares. No final do ensaio o material é removido do animal [11]. Os valores de GI<sub>50</sub> de cada substância estão registrados na Tabela I.1.

Tabela I.1 Valor de GI<sub>50</sub> da indolo[2,1b]quinazoline e seus derivados análogos contra o câncer de mama. Fonte[6]

Compostos	GI <sub>50</sub> referente	Compostos	GI <sub>50</sub> referente
1	30	18	>100
2	>100	19	>100
3	>100	20	>100
4	>100	21	15,0
5	40,0	22	0,01
6	0,7	23	0,01
7	1,0	24	8,0
8	5,0	25	7,5
9	>100	26	7,5
10	>100	27	0,4
11	>100	28	6,0
12	0,4	29	5,0
13	10,0	30	4,0
14	20,0	31	4,0
15	0,06	32	0,4
16	20,0	33	0,06
17	>100	34	0,3

Os pesquisadores consideraram ativas as moléculas cujo valor de GI<sub>50</sub> fosse de no máximo 20 . Este critério foi adotada para as análises realizadas no presente trabalho.

## *II.1 EQUIPAMENTOS E SOFWARES*

Para realização deste trabalho foi usado um *cluster* de 10 microcomputadores e os seguintes *Softwares*: Gaussian 03[17], HyperChem Release 7.5 [12], Unscrambler 9.1 e Microsoft Office Exel 2007.

## *II.2 PROCEDIMENTO COMPUTACIONAL*

As moléculas em estudo foram construída usando o programa HyperChem 7.5 e submetidas à análise conformacional no mesmo programa. A realização desta análise é fundamental, pois consiste na determinação da geometria cuja estrutura molecular será utilizada para calcular as propriedades físico-químicas e geométricas.

As análises conformacionais foram aplicadas apenas em moléculas que possuíam eixo de rotação livre.

### *II.2.1 Escolha da conformação mais adequada para levantamento das propriedades físico-químico e geométricas*

Nem sempre a conformação de menor energia é a mais adequada para o estudo da atividade biológica, pois ao planejar um fármaco, procura-se a conformação que melhor ajuste ao receptor; sendo assim, a escolha da conformação a ser utilizada na pesquisa foi feita da seguinte forma:

1. Selecionou-se a conformação mais estável da molécula mais ativa
2. A conformação dos demais compostos, cuja estrutura geométrica mais se assemelha à estrutura da conformação mais estável da molécula mais ativa, foi escolhida para o estudo.

### *II.2.2 Otimização das moléculas*

As estruturas escolhidas na etapa anterior foram otimizadas pelo programa Gaussian 03, fazendo uso da Teoria do Funcional da Densidade,



utilizando o funcional de correlação e troca híbrido B3LYP [13,14], com a combinação de funções de base 6-31G\*[15,16]

*II 2.3 Descritores geométricos e eletrônicos calculados usando o pacote programa Gaussian 03 [17], com o nível de teoria b3lyp/6-31G\**

Foram calculados os orbitais: HOMO (mais alto orbital ocupado), HOMO-1 (um orbital abaixo do mais alto orbital ocupado), LUMO (mais baixo orbital desocupado), LUMO+1 (um orbital acima do mais baixo orbital desocupado), momento de dipolo ( $\mu$ ), ângulos interatômicos ( $\alpha$ ), ângulos diedrais ( $\delta$ ), eletronegatividade ( $\chi$ ), energia do Gap (HOMO-LUMO) representado por  $E_{\text{gap}}$ , dureza ( $\eta$ ), Ordem de ligação natural ( $b$ ) (NBO) e cargas ( $C$ ) derivadas do potencial eletrostático (CHELPG).

Usando o pacote de programa Hyperchem 7.5 com o método semiempírico PM3, foram calculados as seguintes variáveis: coeficiente de partição (LogP), volume molar ( $V$ ) e polarizabilidade ( $\alpha$ ). Incluindo propriedades estruturais e geométricas, 84 descritores foram obtidos para a análise estatística.

### *II.3 ANÁLISE ESTATÍSTICA DOS DADOS*

Foi feita uma relação entre os parâmetros eletrônicos e geométricos calculados para cada composto e sua atividade contra o câncer de mama. Para isto foram utilizados dois programas: The Unscrambler 9.1 e Microsoft Office Excel 2007, seguindo as seguintes etapas:

Construiu-se uma tabela de dados em que as linhas representam as amostras e as colunas representam as variáveis (parâmetros) calculadas;

Fez-se o cálculo dos Pesos de Fisher para ter uma visão preliminar de quais variáveis são mais significativas na separação entre compostos ativos e inativos;

Gerou-se uma tabela de correlação entre as variáveis;

Fez-se a Análise de Componentes Principais (PCA) dos dados.

## CAPÍTULO III

### FUNDAMENTAÇÃO TEÓRICA

#### III.1 INTRODUÇÃO

Com o avanço da mecânica quântica, o conhecimento em nível molecular foi aprimorado, uma vez que as propriedades dos sistemas químicos podem ser obtidas por meio da resolução da equação de Schrödinger. Contudo, as ferramentas matemáticas, até o presente momento, não são suficientes para resolução das equações quânticas, sendo preciso recorrer a métodos de aproximação [18]. Os principais métodos de aproximação utilizados são: *semiempírico*, que empregam parâmetros experimentais em suas equações e *ab initio*, que utilizam apenas constantes universais [19].

A equação de Schrödinger não relativística independente do tempo para sistemas multieletrônicos pode ser descrita como:

$$\hat{H}\psi = E\psi \quad (III.1)$$

O hamiltoniano ( $\hat{H}$ ) é um operador diferencial que fornece a energia total do sistema, dado pela Equação III.2:

$$\hat{H} = T_n + T_e + V_{nn} + V_{ee} + V_{ne} \quad (III.2)$$

Na Equação III.1 o termo  $E$  representa a energia total do sistema,  $T_n$  representa o operador da energia cinética dos núcleos,  $T_e$  é o operador da energia cinética dos elétrons,  $V_{nn}$  representa o operador da energia potencial devido à interação núcleo – núcleo, o termo  $V_{ee}$  é o operador da energia

potencial devido à interação elétron-elétron e  $\nabla^2$  é o operador da energia potencial devido a interação núcleo - elétron. A função de onda  $\Psi$  depende das coordenadas de posição e de *spin* de todos  $N$  elétrons e  $M$  núcleos.

A forma explícita das energias cinéticas e potenciais dados na Equação III.1 podem ser escritas em unidades atômicas, como:

$$\begin{aligned}
 & -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{a=1}^M \frac{Z_a}{r_{ia}} + \sum_{i < j}^N \frac{1}{r_{ij}} + \sum_{a < b}^M \frac{Z_a Z_b}{r_{ab}} \\
 & \Psi = \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \sigma_1, \dots, \sigma_N, \mathbf{R}_1, \dots, \mathbf{R}_M)
 \end{aligned}$$

onde os  $N$  elétrons representados possuem coordenadas  $\mathbf{r}_i$  e as cargas dos núcleos são representadas pelas coordenadas  $\mathbf{R}_a$ . O operador Laplaciano representado por  $\nabla^2$  em coordenadas cartesianas é dado pela Equação III.3.

$$\nabla^2 = \nabla_x^2 + \nabla_y^2 + \nabla_z^2 \tag{III.3}$$

### III.2 APROXIMAÇÃO DE BORH – OPPENHEIMER

A equação diferencial (Equação III.1), só pode ser resolvida pelo método da separação de variáveis para sistemas monoelétrônicos, isto é, para o átomo de hidrogênio e cátions hidrogenóides. No caso de sistemas multieletrônicos não existe resolução analítica [19].

Uma simplificação proposta por Borh–Oppenheimer é empregada em quase todos os métodos para resolução da equação de Schrödinger. Esta aproximação leva em consideração o fato de o núcleo ser mais pesado que o elétron. Portanto, pode-se supor que o núcleo movimenta devagar quando comparado com o movimento dos elétrons, assim como uma boa aproximação, pode-se considerar que os elétrons movimentam-se em um campo fixo formado pelos núcleos [19].

Considerando a aproximação proposta por Borh–Oppenheimer o termo que se refere à energia cinética entre os núcleos pode ser desprezado e o termo que se refere à repulsão núcleo–núcleo se torna constante [19]. Simplificando o hamiltoniano multieletrônico dado pela Equação III.2, no hamiltoniano eletrônico (

$$H_{el} = -\sum_i \frac{\hbar^2 \nabla_i^2}{2m_e} - \sum_a \frac{\hbar^2 \nabla_a^2}{2M_a} + \sum_{i,j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i,a} \frac{-Ze^2}{|\mathbf{r}_i - \mathbf{r}_a|} + \sum_{a,b} \frac{Ze^2}{|\mathbf{r}_a - \mathbf{r}_b|} \quad (III.4)$$

Em que  $H_{el}$  refere-se ao hamiltoniano eletrônico. Os índices  $i$  e  $j$  refere-se aos elétrons e  $a$  refere-se ao núcleo.

### III.3 MÉTODO DA APROXIMAÇÃO DE HARTREE – FOCK (HF)

Apesar da aproximação de Born–Oppenheimer ter simplificado o operador hamiltoniano, o termo de repulsão entre os elétrons não permite a resolução analítica da Equação III.1 para sistema multieletrônicos. Para resolução desta problemática, Hartree propôs a seguinte aproximação: descrever a função de onda de sistemas multieletrônicos como sendo o produto de  $n$  funções de onda monoelétrônicas ortogonais, onde cada elétron interage com o potencial médio formado pelos núcleos fixos e os demais elétrons [20]. Matematicamente pode ser representado como:

$$\Psi(\mathbf{r}), \quad (III.5)$$

em que  $\psi_i$  é a função de onda monoelétrica e  $\mathbf{r}$  representa coordenada espacial e de *spin*.

Ao fazer esta aproximação Hartree não leva em consideração o princípio da exclusão de Pauli [21], pelo qual as funções de onda multieletrônicas precisam ser anti-simétricas. Fock e Slater propuseram uma solução para o problema de anti-simetrização, utilizando um determinante normalizado [22], descrito como:

$$\dots) = \dots$$

Escrevendo em uma notação mais simplificada fica :

$$\dots) = \det \dots, \quad (III.6)$$

onde  $\psi_i$  é o  $i$ -ésimo spin orbital monoelétrônico ortogonal com coordenadas espaciais e de *spin*. O uso do determinante de Slater (Equação III.6) nos assegura a anti-simetrização em relação à permutação dos elétrons.

Sabe-se que a energia do estado fundamental é descrita como [20]:

$$E = \langle \Psi | \hat{H} | \Psi \rangle. \quad (III.7)$$

Na Equação III. 7 foi utilizado a notação braket de Dirac, e o hamiltoniano  $\hat{H}$  é dado pela Equação III.4. Substituindo as equações (III.4) e (III.6) na equação (III.7), chega ao seguinte valor da energia para sistemas multieletrônicos:

$$\dots \quad (III.8)$$

O primeiro termo da Equação III.8 representa a energia cinética dos elétrons e a energia de atração núcleo – elétron. O segundo termo representa a energia de Coulomb e o último termo refere-se à energia de troca.

A energia de Coulomb pode ser interpretada como a repulsão entre as nuvens eletrônicas. Já a energia de troca não permite uma interpretação clássica; este termo surge com a resolução do problema de anti- simetria das



funções de onda de Hartree, mediante a utilização do determinante de Slater [20].

O procedimento para encontrar a energia de Hartree-Fock é iterativo, em que uma suposição inicial do *spin* orbital origina o potencial de Hartree-Fock, (o potencial sentido por um elétron devido à presença dos demais elétrons e do núcleo), formando o operador de Fock e, conseqüentemente, suas equações, dando origem a um novo conjunto de *spin* orbital e a energia. Através deste novo *spin* orbital, melhora o potencial de HF e repete o procedimento até que a energia convirja [18].

#### III.4 AS EQUAÇÕES DE HARTREE – FOCK - ROOTHAN

A aproximação de Hartree-Fock para átomos isolados é viável, porém sua aplicação para moléculas é inconcebível [23], sendo necessária a expansão da parte espacial dos *spins* orbitais usando um conjunto de funções de base que, em geral, são do tipo Slater ou Gaussianas [21].

(III.9)

$$i = 1, 2, 3 \dots n$$

As funções de base são representadas por  $\phi_i$  e o coeficiente de expansão por  $c_i$ . A resolução desta equação é iterativa em termos dos coeficientes ajustáveis.

#### III.5 - TEORIA DO FUNCIONAL DA DENSIDADE

O método proposto por Hartree-Fock é tão bem aceito que às vezes se esquece que as aproximações que se encontram por trás de tal esquema não sejam boas. Por outro lado, a interação de configuração (IC) [19], a princípio, leva à solução da função de onda exata, com a qual se pode calcular a maioria das propriedades de interesse. Uma vez que o número de configurações explode com o número de elétrons considerados, só podem estudar com

confiabilidade sistemas que possuem pequeno número de elétrons. Além do mais, a interpretação dos resultados é dificultada devido à complexidade da solução resultante.

Um novo modelo baseado na densidade eletrônica do sistema,  $\rho(\mathbf{r})$  foi proposto por Thomas em 1927 [24] e Fermi em 1928 [25]. Esta proposta é considerável, uma vez que se substitui a complexa função de onda de  $N$  elétrons e a equação de *Schrödinger* pela densidade eletrônica muito mais simples. Nesta aproximação os sistemas reais assumem propriedades eletrônicas semelhantes àsquelas dos *gases de fermions*, *i.e.*, um clássico exemplo de partículas não interagentes. Este modelo não teve bom êxito na descrição de sistemas reais, contudo foi a base da teoria do funcional da densidade que considera a densidade eletrônica como variável fundamental.

### III.5.1 O modelo de Thomas-Fermi (TF)

O interesse é encontrar as propriedades eletrônicas do sistema descrito pelo seguinte Hamiltoniano [18]:

$$H = \sum_i \left( -\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Z}{r_i} \right) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V(\mathbf{r}_i) \quad (III.10)$$

onde

$$V(\mathbf{r}_i) = \sum_a \frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|} \quad (III.11)$$

é o potencial externo que representa a interação entre os elétrons  $i$  e o núcleo  $a$ . O termo que se refere à repulsão eletrônica, Thomas(1927) e Fermi(1928) [24,25] consideraram apenas a interação colombiana representado pela Equação III.12.

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (III.12)$$



O termo referente à energia cinética na Equação III.10 foi obtido por Thomas e Fermi a partir da suposição de que os elétrons estão distribuídos uniformemente em um espaço de fase, na razão de dois elétrons para cada elemento de volume  $h^3$ :

$$T[\rho] = \frac{3}{10} \int \rho(r) \sqrt{2m} \, dr \quad (III.13)$$

Esta equação mostra que a energia cinética é determinada explicitamente como um funcional da densidade  $\rho$ . Esta idéia é fundamental para a teoria do funcional da densidade: *local density approximation* (LDA), a qual será detalhada mais a frente. O termo local se refere ao fato de que o potencial depende da densidade no ponto  $r$ , apenas. A energia eletrônica total pode ser escrita como [18]:

$$E[\rho] = \int \rho(r) \left[ \frac{1}{2} \nabla^2 + V(r) \right] dr + E_{TF}[\rho] \quad (III.14)$$

em que  $E_{TF}[\rho] = \int \rho(r) \epsilon_{TF}(r) dr$ . É importante destacar que a energia proposta por Thomas e Fermi  $E_{TF}$  é um funcional do potencial externo  $V(r)$  e da densidade eletrônica  $\rho(r)$  [18]. Observe que os termos de troca ou correlação são completamente negligenciados na Equação III.14.

Minimizando o funcional da energia (Equação III.14) sujeita a restrição  $N[\rho] = N$ , temos:

$$\delta \left( E[\rho] - \mu \left( N[\rho] - N \right) \right) = 0 \quad (III.15)$$

Em que  $\mu$  é o multiplicador de Lagrange, interpretado como sendo o potencial químico na teoria do funcional da densidade [26]. Utilizando a Equação III.15, a solução do problema variacional pode ser expressa como:

$$\epsilon(r) = \epsilon_0 \left( 1 + \frac{4\pi e^2 \rho(r)}{k} \right) \quad (III.16)$$

que expressa a relação básica entre a densidade eletrônica  $\rho(r)$  e a energia potencial  $\epsilon(r)$ . Usando o valor de  $C_F = \frac{3}{4} \frac{h^2}{2m} \left( \frac{3}{\pi} \right)^{1/3}$  na Equação III.16 e considerando  $\epsilon(r)$  como potencial eletrostático ou interno produzido pela densidade eletrônica como:

$$\epsilon(r) = \int \frac{\rho(r')}{|r-r'|} d^3r' \quad (III.17)$$

a densidade eletrônica pode ser escrita como:

$$\rho(r) = \sum_i |\psi_i(r)|^2, \text{ com } \psi_i(r) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}} \quad (III.18)$$

Uma maneira de determinar a forma de  $\epsilon(r)$  é relacionar esta quantidade com a densidade eletrônica  $\rho(r)$  (Equação III.18), através da equação de Poisson[20]

$$\nabla^2 \epsilon(r) = -4\pi \rho(r) \quad (III.19)$$

Substituindo a Equação III.18 na Equação III.19, obtemos a Equação de Thomas-Fermi:

$$\nabla^2 \epsilon(r) = -4\pi \sum_i |\psi_i(r)|^2 \quad (III.20)$$

A Equação III.20 demonstra ser possível determinar a densidade eletrônica a partir do potencial  $\epsilon(r)$ , sem empregar a função de onda. Esta é a principal realização desta teoria. Pode-se enumerar alguns pontos fracos do método de Thomas-Fermi, como: falta de estrutura de camadas (a variação periódica observada quando se muda o número atômico não pode ser reproduzida), falta de uma justificativa coerente para a utilização da densidade eletrônica como variável fundamental e a não precisão dos resultados dos cálculos [27].

### III.5.2 Os teoremas de Hohenberg e Kohn(1964)

A partir da forma do operador do Hamiltoniano total

$$\dots \dots \dots \tag{III.21}$$

fica evidente que conhecendo os valores de:  $N$  e a função de onda e a energia do sistema são determinados, em outras palavras pode se dizer que o número total de elétrons e a distribuição nuclear determinam o hamiltoniano e conseqüentemente as propriedades de um sistema descrito por este operador. O primeiro teorema de Hohenberg e Kohn (1964), [28], mostraram que pode se usar a densidade eletrônica como variável básica ao invés de  $N$  e  $u(r)$ , *i.e.*, o potencial externo é univocamente determinado, a menos de uma constante física aditiva irrelevante, pela densidade eletrônica  $\rho$ .

Esta afirmação deixa claro que se  $\rho(r)$  determina  $u(r)$ , e  $N$ , conseqüentemente  $\rho(r)$  determina o Hamiltoniano  $\hat{H}$ ; daí, implicitamente, todas propriedades deriváveis do  $\hat{H}$ , *i.e.*,

$$\rho(r) \xrightarrow{HK} N \rightarrow \hat{H} \rightarrow \Psi \rightarrow \text{todos os observáveis!!!}$$

$$u(r)$$

A prova é feita através de um absurdo, partindo do princípio de que dois potenciais externos diferentes,  $u$  e  $u'$ , levam a mesma densidade de carga [28].

*Prova.* Seja  $u$  o potencial externo de um sistema com  $N$  partículas, com  $N$  [estado fundamental], associado a um estado fundamental de densidade  $\rho(r)$ , Hamiltoniano  $\hat{H}$ , com função de onda e energia do estado fundamental,  $\Psi$  e  $E$ , respectivamente. Esquemáticamente dado por um sistema A:

E um segundo sistema B, com  $N$  partículas caracterizado por:

Em que  $\psi \neq C$ , conseqüentemente  $\psi \neq 0$ . Pelo princípio variacional de Rayleigh-Ritz,

$$\langle \psi | H | \psi \rangle < \langle \psi_0 | H | \psi_0 \rangle = E_0 \quad (III.22)$$

Semelhantemente:

$$\langle \psi_0 | H | \psi_0 \rangle < \langle \psi | H | \psi \rangle \quad (III.23)$$

Somando as equações III.22 e III.23, obtém-se:

$$\langle \psi | H | \psi \rangle + \langle \psi_0 | H | \psi_0 \rangle < \langle \psi | H | \psi \rangle + \langle \psi_0 | H | \psi_0 \rangle \quad (III.24)$$

Se considerarmos que  $\psi_0$  é o mesmo que afirmar que  $\psi_0 = \psi$ . Concluindo assim que qualquer potencial  $V(\mathbf{r})$ , exceto  $\psi \neq C$ , leva a  $\psi_0 = \psi$ . A energia do estado fundamental do sistema pode ser escrita como [27]:

$$E_0[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V[\rho(\mathbf{r})] + E_{ee}[\rho(\mathbf{r})] \quad (III.25)$$

Em que  $T[\rho(\mathbf{r})]$  é o funcional de Hohenberg-Kohn.  $T[\rho(\mathbf{r})]$  não depende do potencial externo  $V(\mathbf{r})$ , depende somente da densidade eletrônica  $\rho$ . Definido para qualquer sistema físico com  $\rho$  correspondendo à algum  $\psi$  (tal densidade é chamada de  $\rho$  - representativo).  $T[\rho(\mathbf{r})]$  representa a energia cinética e  $V[\rho(\mathbf{r})]$  representa os efeitos de interação elétron-elétron, incluindo a repulsão clássica  $J[\rho]$  e os termos não clássicos [20, 28].

O segundo teorema de Hohenberg-Kohn (HK) [28] provê o princípio

variacional para a teoria do funcional da densidade, *i.e.*, para uma densidade

*tentativa*  $(\mathbf{r})$  tal que  $\geq 0$  e  $N[\cdot]$ , tem-se que  $[\mathbf{r}]$ , onde  $[\mathbf{r}]$  é o funcional da energia definido por III.25.

*Prova.* Seja  $\psi$  uma função de onda tentativa. O primeiro teorema assegura que  $\psi$  determina seu próprio  $E$ . Então, pelo princípio variacional de Rayleigh-Ritz,

$$\int \psi^* \left[ -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi d\mathbf{r} = E \int \psi^* \psi d\mathbf{r} \quad (\text{III.26})$$

Só havendo igualdade quando  $\psi$  é a verdadeira função de onda. Uma vez conhecida a forma explícita do funcional pode – se empregar este formalismo a qualquer sistema, independentemente do número de elétrons  $N$ . O que pode ser claramente percebido se aplicar explicitamente o princípio variacional:

$$\delta E = 0 \quad (\text{III.27})$$

Isto equivale a:

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) - E \right] \psi = 0 \quad (\text{III.28})$$

Hohenberg e Kohn (HK) mostraram que todos os termos na expressão da energia total podem ser avaliados como uma soma de funcionais da densidade de carga, embora a forma matemática destes funcionais não seja conhecida, ou seja, o primeiro teorema de HK é um teorema de existência [28].

### III.5.3 O Método de Kohn-Sham

Como visto, os teoremas de Hohenberg e Kohn não provêm um procedimento para a construção do funcional  $E[\rho]$ . Foi em 1965 que Kohn e Sham (KS) resolveram o problema prático para obter  $E[\rho]$  a partir de  $\rho$ . Os autores propuseram a introdução dos orbitais monoelétrônicos (os quais levam, a uma representação mais apropriada para a descrição da contribuição da energia cinética) e a inclusão dos efeitos de correlação eletrônica explicitamente. A priori Kohn e Sham consideraram um sistema de referência

*não interagente* de  $N$  partículas, descrito pelo seguinte Hamiltoniano [20, 21, 26] ,

$$- \tag{III.29}$$

o termo *não interagente*, significa que não existe interação do tipo elétron-elétron. A principal afirmação de KS é: *para qualquer sistema de partículas interagentes, existe um potencial monoelétrônico local (potencial fictício) tal que a exata densidade do estado fundamental do sistema interagente satisfaz* , onde representa a densidade de carga do sistema não interagente [29]. O determinante da função de onda do estado fundamental do sistema não interagente é:

$$\dots ) = \det \tag{III.30}$$

A função de onda monoelétrica  $\varphi_i$ , assumindo ser não degenerada, define por meio da equação:

$$\tag{III.31}$$

Um procedimento semelhante ao aplicado nas equações monoelétrônicas de Hartree-Fock são empregados para obter os orbitais monoelétrônicos de Kohn e Sham[26,29]:

$$\varphi - \quad i = 1,2,\dots,N. \tag{III.32}$$

Uma vez assumido a existência do potencial "*fictício*" , o qual gera por meio das equações III.31 e III.32, o termo de energia cinética [ é representado por:

$$\dots - \dots - \dots \quad (III.33)$$

Os componentes do funcional  $E_{xc}$  podem ser escritos da seguinte maneira:

$$\dots \quad (III.34)$$

A quantidade  $E_{xc}$  é a energia de troca e correlação (*exchange-correlation*). Este termo contém a diferença entre a energia do estado fundamental e a energia cinética do referido sistema não interagente ( $E_{xc}^0$ ) e também a parte não clássica do funcional  $E_{xc}$ . Substituindo a Equação III.34 na Equação III.28, temos

$$\dots - \dots - \dots$$

em que:

$$\dots \quad (III.34)$$

Kohn e Sham fez  $\phi_i$  na Equação III.32, para obter as equações monoeletrônica

$$\phi_i - \dots \quad i = 1, 2, \dots, N \quad (III.35)$$

$$\dots - \dots - \dots \quad (III.36)$$



Estas são as equações de Kohn-Sham [29]. Como o operador de KS, depende da densidade eletrônica, estas equações devem ser resolvidas iterativamente, de modo similar a resolução das equações de Hartree-Fock. A energia total é dada por:

$$E_{KS} = \sum_i \langle \psi_i | \hat{h} | \psi_i \rangle + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + \int \rho(\mathbf{r})v_{ext}(\mathbf{r})d\mathbf{r} + E_{xc}[\rho]$$

onde:

$$\hat{h} = -\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r})$$

A densidade eletrônica é dada por:

.

#### III.5.4 Aproximação LDA (Local-Density Approximation)

As equações de KS não determinam o termo de correlação e troca, sendo necessário fazer algumas aproximações. Em 1965 Kohn e Sham propuseram a *aproximação local da densidade* (LDA – Local-Density Approximation)[26]. Como varia lentamente com a posição, então a energia pode ser dada por:

(III.37)

A integral é sobre todo espaço e  $E_{xc}$  é a energia de troca mais a energia correlação de cada elétron em um gás homogêneo de elétrons cuja densidade é  $\rho$ . Aplicando a derivada no funcional  $E_{xc}$ , obtém-se:

$$\frac{\delta E_{xc}}{\delta \rho} = \frac{\delta E_{xc}}{\delta \rho} \quad (III.38)$$

As equações III.37 e III.38 são empregadas na resolução das equações propostas por KS. O funcional pode ser dividido em duas partes: de troca e correlação:

(III.39)

em que

--

Vosko, Wilk e Nusair (VWN) calcularam a parte de correlação, que é dada por uma complexa expressão (como pode ser visto no apêndice E do livro de *Parr e Yang*). Assim,  $E_{xc} = E_x + E_c$ . Usando a relação  $E_c = E_{xc} - E_x$  e derivando o funcional (veja apêndice A do livro de *Parr & Yang*) [26], tem-se:

(III.40)

—

(III.41)

O procedimento para encontrar o funcional  $E_{xc}$ , ou qualquer outro funcional é iterativo: faz uma suposição da densidade eletrônica molecular 26

inicial usando a densidade eletrônica de átomos, a partir desta densidade calcula o potencial (que inclui o potencial e aplica este potencial na equação de Kohn – Sham(KS) dada pela Equação III.42 [18, 20, 21].

$$\begin{aligned} & - \quad \text{—————} \quad \text{—————} \\ & \hspace{20em} \text{(III.42)} \end{aligned}$$

Ao resolver as equações de KS, os orbitais iniciais são determinados. Sendo que estes orbitais são, geralmente, expandidos em termos de um conjunto de funções de base gerando um conjunto de equações parecidas com as equações de Hartree-Fock-Roothaan, com a substituição dos elementos da matriz de Fock pelos elementos da matriz de KS [21]. As equações de Kohn-Sham, com os orbitais expandidos em termos das funções do conjunto de base, podem ser escritas como:

As funções de base mais utilizadas nos cálculos a nível de DFT são as funções Gaussianas contraídas, mas alguns programas usam STO's ou outro tipo de funções de base. As equações de KS também podem ser resolvidas numericamente.

### III.5.5 Aproximação LSDA (Local-Spin-Density Approximation)

Sistemas cujo número de elétrons é ímpar (camada aberta) ou nas geometrias moleculares onde os átomos estão quase dissociados, aplica-se o método LSDA. O qual considera diferentes partes espaciais nos orbitais de KS para elétrons com *spins* opostos. Assim o funcional no método LSDA aplica separadamente as densidades eletrônicas: e :

, . As equações de KS são tratadas separadamente[21], com os potenciais de troca e correlação dados por

\_\_\_\_\_ e \_\_\_\_\_ Em sistemas em que todos os elétrons estão emparelhados ou nas geometrias moleculares próximas do equilíbrio, a aproximação LSDA se reduz à aproximação LDA. Apesar de o método LSDA considerar a densidade eletrônica uniforme em moléculas, comparando com os cálculos de HF, produz bons resultados das geometrias moleculares, frequências vibracionais e momentos dipolares; mesmo para compostos contendo metais de transição [18, 21].

### III.5.6 GGA - Generalized-Gradiente Approximation

Para sistemas com variação da densidade uniforme, aplica-se os métodos LDA (cuja a integração da Equação III.37 para o funcional é função somente de  $\rho$ ), ou LSDA (cuja integração na expressão é função de  $\rho$  e  $\nabla\rho$ ).

Para os funcionais que vão além do LSDA deve apresentar correções referente à variação da densidade  $\rho$  com a posição [26]. Para tal efeito inclui-se o gradiente das densidades  $\rho_\uparrow$  e  $\rho_\downarrow$  no integrando:

$$\int d^3r \left[ \frac{1}{2} \nabla \rho \cdot \nabla \rho + f(\rho, \nabla \rho) \right] \quad (III.43)$$

onde  $f$  é uma função da densidade de spin e do gradiente da densidade de spin. Desta forma consegue-se obter uma aproximação não local. O funcional de correlação e troca, pode ser tratado separadamente como sugere a equação III.44:

$$\quad (III.44)$$

Para obter os funcionais da energia de troca e correlação são feitas algumas suposições teóricas referentes ao comportamento da densidade em situações variadas e geralmente são adicionados alguns parâmetros empíricos nas equações [21]. Alguns funcionais de troca, com a densidade corrigida pelo gradiente, comumente usados são: os funcionais de Perdew e Wang de 1986 (não contêm parâmetros experimentais) e são denotados por PW86 ou PWx86,

o funcional de Becke de 1988 denotados por B88, Bx88, Becke88, ou B e PWx91. A forma explícita do funcional de troca B88[13] é:

$$\frac{1}{2} \left( \frac{1}{2} + \frac{1}{2} \right) \quad (III.45)$$

onde  $\alpha = \frac{1}{2}$ ,  $\beta = \ln$ ,  $\gamma$ , é um parâmetro semi-empírico cujo valor é 0,0042 u.a. foi ajustado usando a energia de troca conhecida de cálculos HF para vários átomos e

$$\frac{1}{2} \left( \frac{1}{2} + \frac{1}{2} \right) \quad (III.46)$$

Os funcionais de correlação corrigidos pelo gradiente mais utilizados são de Lee-Yang-Parr (LYP) [14], o funcional de correlação de Perdew (P86 ou P86), o funcional de correlação de Perdew-Wang 1991 (PW91 ou PWc91) e os funcionais de Becke designados como Bc95 ou B96. [18, 21]. A princípio qualquer funcional de correlação pode ser combinado com qualquer funcional de troca resultando, por exemplo, no BLYP/6-31G\*, notação que descreve um cálculo a nível DFT feito com o funcional de troca de Becke 1988 combinado com o funcional de correlação de Lee-Yang-Parr com os orbitais de KS expandidos com o conjunto de base 6-31G\*.

### III.5.7 Aproximação híbrida

O funcional de troca da equação de KS pode ser descrito da mesma forma que é descrito o termo de troca da equação de HF, diferindo apenas pelo fato dos orbitais de HF serem substituídos pelos orbitais de KS. Para sistemas de camadas fechadas, o termo de troca pode ser escrito como:[21]

$$\frac{1}{2} \left( \frac{1}{2} + \frac{1}{2} \right) \quad (III.47)$$

Nesta expressão a soma é feita sobre os elétrons  $e$ , portanto o valor  $\frac{1}{4}$  substitui o valor  $\frac{1}{2}$  na equação. O funcional híbrido mistura as fórmulas III.47 para o  $E_X$  com o funcional  $E_X$  corrigido pelo gradiente e as fórmulas para a energia de correlação  $E_C$ . Nesta dissertação, foi empregado, principalmente, o funcional híbrido B3LYP[13] (ou Becke3LYP); uma das aproximações mais populares (o número 3 indica que este funcional possui três parâmetros). Este funcional é definido por:

$$(III.48)$$

onde o termo  $E_X$  é o funcional de troca desenvolvido por Becker,  $E_C$  é dado pela Equação III.8 e o termo  $E_{XC}$  denota o funcional de correlação desenvolvido por Vosko-Wilk-Nusair no procedimento LSDA e  $E_{XC}$  denota o funcional de correlação descrito por Lee, Yang e Parr. Os valores das constantes obtidos pelo ajuste com os valores experimentais das energias de atomizações moleculares são:  $a_0=0,20$ ,  $a_X = 0,72$  e  $a_C=0,81$ .

O funcional híbrido B3PW91 substitui, na Equação III.48, o funcional pelo funcional  $E_X$ , mas sem mudar as constantes de ajuste experimental. Outro funcional híbrido bastante utilizado é o B3PW91. Este funcional é dado por  $E_X$ , onde  $c_X$  é um parâmetro experimental e  $E_X$  e  $E_C$  são funcionais que contêm três e seis parâmetros, respectivamente [18,20,21].

Cálculos utilizando os funcionais corrigidos pelo gradiente (GGA) e os funcionais híbridos, têm produzidos bons resultados tanto para otimizações moleculares, como para a obtenção de propriedades como: momentos dipolares, frequências vibracionais e energias de atomizações. Estes resultados são obtidos com custo computacional aproximadamente equivalente ao de HF[20].

### III.6 MÉTODOS DE ESTATÍSTICA MULTIVARIADOS DE ANÁLISE

Os métodos de análise multivariada constituem em uma ferramenta aplicável na análise de conjunto de dados complexos.

Três conceitos fundamentais devem ser entendidos para as análises de dados, são eles: fatos, dados e informações. Fato refere-se a tudo que tem existência real (pessoas, caroço de feijão, etc.). Os dados são medidas, notícias, utilizados no desenvolvimento e raciocínio, cálculos e discussões. A análise dos fatos e dados leva a algum conhecimento, isto é informação [30].

Os métodos multivariados podem ser classificados com base em duas considerações: natureza da informação e propósito da análise para estruturação dos dados. Podem-se definir seis tipos de variáveis[30]:

1. Variáveis binárias: Só podem assumir dois valores (sim/não, 1/0)
2. Classes ordenadas: Podem assumir i valores alfa numéricos (magro e obeso, grupo 1, grupo e grupo 3).
3. Classes desordenadas: Variável que não leva em consideração a hierarquia entre as classes (colunas cromatográficas)
4. Variáveis discretas: Como por exemplo: 1,2,3, etc.
5. Variáveis contínuas: Variáveis que assumem um determinado valor em um intervalo definido (20-40).
6. Variáveis de categoria: é uma variável não utilizada em métodos não paramétrico em poucos exemplo de utilização por métodos multivariados.

A escolha do método é limitada pelo tipo de variável disponível em um determinado conjunto de variáveis que se pretende estudar

Os métodos estatísticos ainda podem ser relacionados com os seus propósitos. O objetivo a ser alcançado é quem vai determinar o modelo a ser escolhido. As principais metas a serem alcançadas são [30]:

1. Predição: neste caso faz-se uma relação entre variáveis independente e outra variável ou conjunto de variáveis dependentes, para obter resultados satisfatórios.
2. Tomada de decisão: neste caso prevalecem os aspectos qualitativos entre as variáveis, leva em consideração a relação existente entre elas.

3. Agrupamento (não supervisionado): formam-se grupos que possuam similaridades, definindo, desta forma, grupos homogêneos.
4. Discriminação (supervisionado): São feitas observações, as quais são classificadas como pertencente a uma classe pré definida, isto após cálculos probabilísticos.

Os métodos de estatística multivariada podem ser subdivididos em dois grandes grupos: método de Análises de fatores onde novas variáveis ou fatores são calculados a partir da combinação linear das variáveis originais como, por exemplo, a Análise de Componentes Principais e método de classificação nos quais colunas ou linhas das matrizes são analisados e aquelas que possuem similaridade são agrupadas em classe. O método de classificação pode ser dividido em dois grupos: métodos de partição (as classes são predefinidas) e método hierárquico (inicia pelos dois elementos mais similares, produzindo uma série de classes, formando um dendograma), como por exemplo, HCA (Análise Hierárquica de Cluster) [31].

### *III.6.1 Vetores de média e matriz de covariância e correlação*

Variáveis que representam um dado populacional ou amostral são freqüentemente descritos pelos seus valores médios e suas variâncias. A média de uma amostra é dada pela Equação III.49 e a estimativa de sua variância pela equação III.50. Em que n corresponde ao número de medidas realizadas sobre a variável X[ 32].

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (III.49)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (III.50)$$

O vetor que representa as médias e a matriz de covariância são utilizadas de forma similar para descrever as populações e amostras multivariadas. Considere por exemplo que existam n variáveis e os valores destas variáveis para o iésimo indivíduo da



amostra são: então a média da variável na amostra é [32]:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (III.51)$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (III.52)$$

A covariância é a propensão que duas variáveis têm de variar de forma mais ou menos similar das respectivas médias [32]. A equação que descreve a covariância é:

$$s_{jk} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1} \quad (III.53)$$

em que  $j$  e  $k$  são variáveis e  $n$  o número de amostras.

Um problema previsto nos cálculos de covariância é o fato de que cada variável pode possuir escala diferente e assim tornar complicado a análise do grau de associação estatística entre estas variáveis [2]. Esta problemática pode ser resolvida com a obtenção de um coeficiente de correlação que, para duas variáveis, é dado por:

$$r_{jk} = \frac{s_{jk}}{s_j s_k} \quad (III.54)$$

onde  $r_{jk}$  é dado pela Equação III.53 e  $s_j$  e  $s_k$  são os desvios padrão das respectivas variáveis cuja a equação que o representa é dado a seguir:

$$s_j = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (III.55)$$

Por definição, pode-se afirmar que o coeficiente de correlação tem seus valores variando no intervalo  $[-1,1]$ ; desta forma, é possível comparar duas variáveis, pois suas grandezas são adimensionais [32].

Se o coeficiente de correlação  $= 0$ , indica que as variáveis são estatisticamente independentes, caso  $> 0$ , indica que quando uma das variáveis é maior que sua média a outra também é, e da mesma forma se uma das variáveis for menor que sua média a outra também é. Já para  $< 0$  indica que quando uma das variáveis é maior que sua média a outra é menor [32].

### III.6.2 Escalamento dos dados

Para averiguar a importância deste procedimento, partiremos do pressuposto que variáveis de diferentes escalas estão sendo analisadas para uma mesma amostra, consideremos a ordem de ligação dada em angstrom e o calor de formação dado em calorias. Os valores das ordens de ligação são bem menores que os valores do calor de formação; contudo, isto não significa que a variável ordem de ligação tenha pouca importância em detrimento ao calor de formação, então para que ambas as variáveis possuam a mesma importância dentro da classe de variáveis é necessário fazer o escalamento [2].

Os dois tipos de escalamento mais empregados em análises multivariada são:

#### a) Escalamento por intervalo

Este escalamento procede-se da seguinte forma: na origem o valor mínimo de cada amostra e divide pelo intervalo de dados como mostra a Equação III.56.

---

$$\text{(III.56)}$$

Onde:  $0$ , este método é sensível a pontos atípicos.

### b) Autoescalamento

Este método exclui qualquer influência de unidades, contudo não é muito sensível a pontos atípicos (*outliers*). O autoescalamento pode ser feito de duas formas: auto escalamento por unidade de variância em que os dados são centrados na média e divididos pelo desvio padrão ( Equação III.57). Ou outro tipo de uso comum, que é dado pela Equação III.58.

$$\text{-----} \tag{III.57}$$

$$\text{-----} \tag{III.58}$$

No presente trabalho foi utilizado a equação III.57

### III.6.3 Pesos de Fisher

Com a finalidade de reduzir a matriz de dados, pode-se aplicar o Peso de Fisher que é uma técnica capaz de discriminar as variáveis mais significativas das variáveis menos significativa para as análises [33]. O peso de Fisher é dado pela Equação III.59, em que estão sendo consideradas duas classes: A e B.

$$\text{-----} \tag{III.59}$$

Em que  $i$  representa as amostras,  $\bar{x}$  a média e  $S^2$  a variância.

Analisando a Equação III.59, observa-se que quanto maior a diferença entre as médias das classes A e B e quanto menor as variâncias dentro das classes, maior é o valor do Peso Fisher. Como variâncias menores são melhores para discriminação das classes, quanto maior o Peso Fisher maior a capacidade da variável na separação.

### III.6.4 Análise de Componentes Principais (PCA)

A técnica de análise de componentes principais foi introduzida na química por Malinowski por volta de 1960 [31], dez anos mais tarde um grande número de trabalhos começou a ser publicados com suas aplicações. A PCA consiste em um método não supervisionado que resulta em um conjunto de vetores, denominados componentes principais (PC). Estas componentes são a combinação linear de  $m$  variáveis originais, como descritos pelas Equações III.60, III.61 e III.62.

$$PC_1 = a_{1,1}X_1 + a_{1,2}X_2 + a_{1,3}X_3 + \dots + a_{1,m}X_m \quad (III.60)$$

$$PC_2 = a_{2,1}X_1 + a_{2,2}X_2 + a_{2,3}X_3 + \dots + a_{2,m}X_m \quad (III.61)$$

$$\dots \dots \dots$$
$$PC_m = a_{m,1}X_1 + a_{m,2}X_2 + a_{m,3}X_3 + \dots + a_{m,m}X_m \quad (III.62)$$

As componentes principais conservam a maior parte da variância das variáveis originais entre si [34], sendo ortogonais, isto é, não existe correlação entre as mesmas, o que é importante, pois garante que cada componente descreva uma “característica” diferente dos dados originais [35].

A PCA é um método de decompor uma matriz de dados originais  $X$  de posto  $r$  em uma soma de  $r$  matrizes de posto 1, em que o posto é o número de vetores linearmente independentes. As matrizes de posto 1 são o produto dos vetores scores ( $S_h$ ) pelos transposto dos loading ( $I_h$ ) como mostra a Equação III.63 [36]:

$$(III.63)$$

Uma das vantagens de utilizar a técnica de PCA é diminuir a dimensão dos dados originais. Um conjunto formado por  $m$  variáveis e  $k$  amostra, representa  $k$  pontos em  $m$  dimensões, se  $m$  for maior que três não é possível a

visualização gráfica. As primeiras PCs ,isto é, 2 ou 3, na maioria das vezes, são capazes de explicar a maior parte da variância dos dados, o espaço m dimensional se reduz para aproximadamente 3 dimensões, sendo possível a análise gráfica do mesmo[2].

## CAPITULO IV

### RESULTADOS E DISCUSSÃO

#### IV. 1. ANÁLISE CONFORMACIONAL

A indolo [2,1b] quinazolina e seus derivados análogos apresentam conformações planares, isso se deve aos quatro anéis rígidos que constituem a estrutura base destas moléculas, como pode ser visto na Figura IV.1.

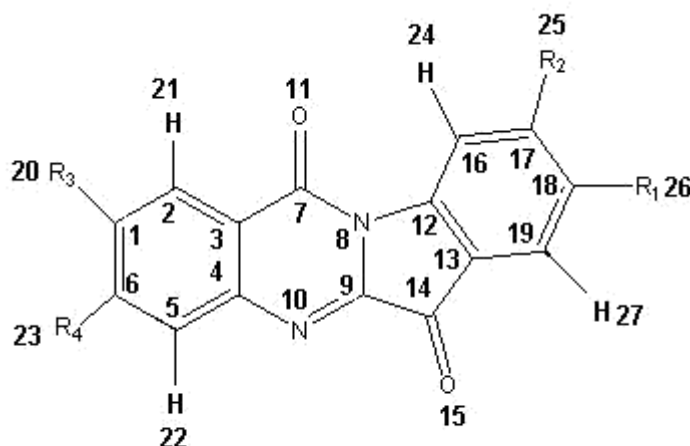


Figura IV.1 Estrutura base das indolos [2,1b] quinazoline e seus derivados análogos.

#### IV. 2. ANÁLISE DAS PROPRIEDADES CALCULADAS

Para estabelecer a relação da estrutura da molécula e sua atividade contra o câncer de mama, foram calculadas treze propriedades, incluindo diferentes descritores eletrônicos e estruturais, os cálculos e discussão serão apresentados nas próximas linhas.

#### *IV.2.1 Energias dos orbitais moleculares de fronteira*

As energias dos orbitais moleculares de fronteira são muito populares na descrição SAR, uma vez que determinam as bandas eletrônicas de muitos sólidos e são responsáveis pela transferência de cargas na formação de complexos.

A energia do HOMO (mais alto orbital ocupado) e do LUMO (o mais baixo orbital desocupado) corresponde à energia necessária para remover e adicionar elétrons respectivamente. Assim a energia do HOMO é proporcional a energia de ionização e à energia do LUMO a afinidade eletrônica [5]. Além das energias dos orbitais de fronteira foi calculado também a energia do HOMO-1 (um orbital molecular, abaixo do mais alto orbital molecular ocupado) e a energia do LUMO+1 (um orbital de fronteira acima do mais baixo orbital molecular desocupado).

Analisando a Tabela IV.1, pode se observar que as energias do orbital mais alto ocupado (HOMO) das 34 moléculas possuem pequena variância, isto pode ser percebido por meio do baixo valor de desvio-padrão encontrado [5].

#### *IV. 2.2 Gap ( HOMO – LUMO )*

A energia do Gap ( diferença de energia do mais alto orbital ocupado e do mais baixo orbital desocupado) é um importante índice de estabilidade da molécula. Altos valores de Gap indicam que a molécula é altamente estável e tem baixa reatividade. Baixos valores do Gap indicam instabilidade na molécula e, portanto, alta reatividade [5]. Como mostra a Tabela IV.1, as moléculas que apresentaram maior Gap foram as 18,19 e 20, as quais não apresentam atividade contra o câncer de mama, um fator que justifica esta inatividade pode estar relacionado com sua baixa reatividade, quando comparada com as demais moléculas do grupo.

#### *IV. 2.3 Dureza (*

Segundo a teoria do orbital molecular, a dureza corresponde à facilidade com que os elétrons são rearranjados após uma perturbação; Uma molécula é dita dura se possuir certa resistência ao rearranjo de elétrons após

uma perturbação, a molécula é dito mole, se tiver facilidade de rearranjo molecular após uma perturbação [2]. A dureza é dada por:

$$= \frac{1}{2} (\text{Potencial de Ionização} - \text{Afinidade eletrônica})^{\frac{1}{2}} (-\text{HOMO} + \text{LUMO})$$

Tabela IV.1 Valores de energias dos orbitais moleculares de fronteira, Gap( , dureza ( e eletronegatividade( obtidas a nível de teoria B3LYP/6-31G\* . Todos valores são dados em ( eV).

Indolos	HOMO	HOMO-1	LUMO	LUMO-1		
1	-0,24	-0,26	-0,11	-0,06	7	7
2	-0,24	-0,26	-0,11	-0,07	7	8
3	-0,26	-0,28	-0,13	-0,09	7	9
4	-0,23	-0,24	-0,10	-0,05		
5	-0,21	-0,24	-0,10	-0,05		
6	-0,24	-0,26	-0,11	-0,06		8
7	-0,26	-0,28	-0,12	-0,10	7	9
8	-0,21	-0,25	-0,10	-0,05		
9	-0,22	-0,25	-0,10	-0,05		
10	-0,22	-0,25	-0,10	-0,05		
11	-0,24	-0,26	-0,11	-0,06		8
12	-0,23	-0,26	-0,11	-0,06		7
13	-0,23	-0,26	-0,11	-0,06		7
14	-0,23	-0,26	-0,11	-0,06		7
15	-0,23	-0,25	-0,11	-0,06		7
16	-0,22	-0,23	-0,10	-0,05		
17	-0,23	-0,26	-0,11	-0,06		7
18	-0,23	-0,25	-0,05	-0,04	7	9
19	-0,22	-0,25	-0,05	-0,04	7	9
20	-0,21	-0,23	-0,06	-0,05		8
21	-0,22	-0,25	-0,09	-0,05		7
22	-0,22	-0,25	-0,09	-0,05		7
23	-0,22	-0,25	-0,09	-0,05		7
24	-0,22	-0,24	-0,09	-0,05		7
25	-0,23	-0,24	-0,10	-0,07		7
26	-0,21	-0,22	-0,08	-0,05		7
27	-0,20	-0,22	-0,09	-0,05		
28	-0,21	-0,22	-0,08	-0,05		
29	-0,21	-0,22	-0,08	-0,05		7
30	-0,21	-0,24	-0,10	-0,09		
31	-0,21	-0,23	-0,10	-0,09		
32	-0,24	-0,27	-0,11	-0,09		7
33	-0,22	-0,22	-0,09	-0,05		
34	-0,22	-0,22	-0,09	-0,05		
Média	-0,23	-0,25	-0,10	-0,06		
Variância	1,69E-04	3,00E-04	3,00E-04	2,00E-04		
D.padrão	0,01	0,02	0,02	0,02		



As moléculas 18,19 e 20 que são inativas, são as mais resistentes a um rearranjo eletrônico, uma vez que apresentaram os maiores valores de dureza, como pode ser visto na Tabela IV.1.

#### *IV. 2.4 Eletronegatividade ( )*

É a tendência que um átomo tem de atrair elétrons em uma ligação química, sua equação matemática é dada por [2,37]:

$$= \frac{1}{2} (\text{potencial de ionização} + \text{afinidade eletrônica}) = \frac{1}{2}(-\text{HOMO} - \text{LUMO})$$

Uma vez que as moléculas apresentam um pequeno desvio padrão, pode se afirmar que há uma grande semelhança entre os valores de eletro negatividade das 34 moléculas estudadas, vide Tabela IV.1

#### *IV. 2.5 Cargas atômicas parciais (C)*

Neste trabalho, as cargas obtidas foram derivadas do potencial eletrostático utilizando CHELPG (cargas derivadas do potencial eletrostático usando o modelo de grade). Este método vem sendo utilizado com maior frequência em relação ao tradicional modelo proposto por Mullikan, uma vez que as cargas derivadas do potencial eletrostático superam alguns problemas apresentados pelo método de Mullikan, como dividir igualmente uma população de elétrons em uma ligação entre dois átomos [38].

As cargas ChelpG foram propostas por Breneman e por Wiberg [38]. Esta análise utiliza o método de grade, onde se constrói um modelo em que a molécula é colocada em um cubo e vários pontos são adicionados a sua estrutura. Após o procedimento de seleção de pontos, o potencial eletrostático de cada amostra é calculado analiticamente e as cargas que melhor representam este potencial eletrostático são obtidas[38].

As cargas derivadas do potencial eletrostático são ajustadas para reproduzir o potencial eletrostático ao redor da molécula. Como a interação inicial entre o substrato e sítio ativo se dá inicialmente, através de uma interação eletrostática as cargas obtidas a partir deste potencial eletrostático constituem um importante descritor no estudo da relação estrutura atividade [38].

Na Tabela IV.2 estão as cargas atômicas parciais dos 27 átomos da estrutura básica das 34 moléculas em estudo.

Tabela IV.2 Cargas atômicas (c), derivadas do potencial eletrostático com método CHELPG, obtidas a nível de teoria B3LYP/6-31G\*.

Indolo	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>	C <sub>9</sub>
1	-0,13	-0,03	-0,27	0,52	-0,24	-0,04	0,63	-0,37	0,42
2	-0,35	0,27	-0,56	0,78	-0,42	0,16	0,76	-0,52	0,55
3	-0,02	-0,03	-0,29	0,56	-0,29	-0,05	0,67	-0,43	0,51
4	0,20	-0,10	-0,33	0,54	-0,38	0,23	0,75	-0,46	0,49
5	0,38	-0,21	-0,28	0,50	-0,22	-0,20	0,72	-0,45	0,46
6	-0,17	0,02	-0,51	0,73	-0,34	-0,05	0,83	-0,63	0,59
7	-0,11	-0,02	-0,30	0,54	-0,26	-0,04	0,66	-0,39	0,47
8	-0,13	-0,02	-0,28	0,53	-0,25	-0,03	0,63	-0,35	0,42
9	-0,12	-0,03	-0,27	0,52	-0,24	-0,04	0,63	-0,29	0,37
10	-0,11	-0,04	-0,27	0,53	-0,25	-0,05	0,64	-0,36	0,43
11	-0,12	-0,02	-0,31	0,55	-0,26	-0,04	0,68	-0,40	0,47
12	-0,12	-0,02	-0,27	0,52	-0,24	-0,03	0,64	-0,35	0,43
13	-0,13	-0,01	-0,30	0,54	-0,27	-0,02	0,65	-0,36	0,43
14	-0,13	-0,03	-0,28	0,54	-0,16	-0,02	0,63	-0,34	0,41
15	-0,15	-0,01	-0,28	0,54	-0,27	-0,01	0,61	-0,31	0,41
16	-0,14	-0,02	-0,28	0,53	-0,26	-0,02	0,64	-0,37	0,45
17	-0,13	-0,01	-0,30	0,53	-0,25	-0,02	0,65	-0,36	0,42
18	-0,14	-0,02	-0,27	0,49	-0,25	-0,03	0,64	-0,44	0,56
19	-0,13	-0,02	-0,28	0,46	-0,23	-0,04	0,67	-0,50	0,69
20	-0,14	-0,02	-0,27	0,48	-0,26	-0,04	0,62	-0,37	0,53
21	-0,18	0,01	-0,34	0,61	-0,32	0,02	0,66	-0,37	0,44
22	-0,15	-0,01	-0,29	0,53	-0,28	-0,01	0,65	-0,38	0,48
23	-0,12	-0,03	-0,31	0,54	-0,25	-0,05	0,68	-0,45	0,53
24	-0,13	-0,01	-0,31	0,55	-0,26	-0,04	0,66	-0,38	0,46
25	-0,13	-0,03	-0,28	0,52	-0,27	-0,03	0,64	-0,36	0,46
26	-0,14	-0,02	-0,30	0,55	-0,26	-0,03	0,65	-0,40	0,50
27	-0,15	-0,01	-0,30	0,55	-0,27	-0,03	0,66	-0,41	0,49
28	-0,13	-0,01	-0,30	0,53	-0,26	-0,04	0,66	-0,40	0,50
29	-0,15	-0,02	-0,31	0,56	-0,28	-0,02	0,70	-0,44	0,51
30	-0,12	-0,05	-0,27	0,51	-0,24	-0,04	0,66	-0,41	0,49
31	-0,14	-0,01	-0,29	0,53	-0,25	-0,03	0,66	-0,39	0,49
32	-0,14	-0,01	-0,29	0,52	-0,27	-0,01	0,63	-0,37	0,49
33	-0,19	0,01	-0,46	0,72	-0,36	-0,03	0,75	-0,47	0,55
34	-0,19	0,02	-0,50	0,74	-0,37	-0,04	0,80	-0,53	0,60

Indolo	C <sub>10</sub>	C <sub>11</sub>	C <sub>12</sub>	C <sub>13</sub>	C <sub>14</sub>	C <sub>15</sub>	C <sub>16</sub>	C <sub>17</sub>	C <sub>18</sub>
1	-0,62	-0,51	0,23	-0,16	0,43	-0,42	-0,17	-0,06	-0,11
2	-0,72	-0,50	0,32	-0,17	0,40	-0,40	-0,24	-0,04	-0,18
3	-0,65	-0,49	0,24	-0,13	0,37	-0,40	-0,18	-0,02	-0,14
4	-0,63	-0,55	0,25	-0,14	0,38	-0,41	-0,19	-0,04	-0,14
5	-0,63	-0,54	0,26	-0,14	0,39	-0,42	-0,19	-0,04	-0,15
6	-0,63	-0,51	0,21	-0,09	0,37	-0,39	-0,14	-0,08	0,01
7	-0,63	-0,52	0,17	-0,15	0,45	-0,43	-0,14	-0,21	0,36
8	-0,61	-0,52	0,17	-0,20	0,49	-0,44	-0,30	0,26	0,24
9	-0,63	-0,52	0,15	-0,09	0,42	-0,41	-0,13	-0,21	0,35
10	-0,62	-0,50	0,60	-0,12	0,44	-0,42	-0,13	-0,16	0,16
11	-0,63	-0,52	0,23	-0,14	0,39	-0,40	-0,19	-0,02	0,03
12	-0,62	-0,51	0,17	-0,13	0,43	-0,42	-0,12	-0,21	0,36
13	-0,62	-0,51	0,18	-0,14	0,43	-0,41	-0,14	-0,18	0,31
14	-0,62	-0,51	0,17	-0,14	0,44	-0,42	-0,11	-0,21	0,33
15	-0,62	-0,50	0,15	-0,11	0,43	-0,42	-0,13	-0,20	0,33
16	-0,63	-0,51	0,17	-0,12	0,42	-0,43	-0,12	-0,22	0,34
17	-0,62	-0,52	0,23	-0,21	0,47	-0,43	-0,20	-0,09	0,26
18	-0,64	-0,51	0,29	-0,18	0,40	-0,63	-0,19	-0,08	-0,11
19	-0,64	-0,52	0,29	-0,12	0,24	-0,63	-0,20	-0,08	-0,13
20	-0,64	-0,51	0,21	-0,09	0,36	-0,43	-0,17	-0,06	-0,14
21	-0,68	-0,52	0,19	-0,10	0,33	-0,33	-0,16	-0,06	-0,15
22	-0,64	-0,51	0,20	-0,05	0,15	-0,20	-0,16	-0,07	-0,13
23	-0,67	-0,82	0,26	-0,17	0,25	-0,33	-0,17	-0,08	-0,13
24	-0,66	-0,52	0,22	-0,15	0,34	-0,36	-0,16	-0,06	-0,14
25	-0,64	-0,51	0,20	-0,08	0,24	-0,26	-0,17	-0,06	-0,12
26	-0,67	-0,52	0,23	-0,13	0,24	-0,32	-0,16	-0,09	-0,11
27	-0,67	-0,52	0,25	-0,15	0,27	-0,31	-0,18	-0,06	-0,14
28	-0,66	-0,52	0,22	-0,12	0,23	-0,31	-0,16	-0,07	-0,14
29	-0,67	-0,53	0,26	-0,14	0,26	-0,32	-0,18	-0,07	-0,14
30	-0,64	-0,51	0,26	-0,12	0,25	-0,33	-0,16	0,10	0,01
31	-0,65	-0,51	0,25	-0,12	0,24	-0,32	-0,17	-0,11	0,03
32	-0,64	-0,50	0,22	-0,05	0,14	-0,20	-0,17	-0,10	0,02
33	-0,74	-0,52	0,33	-0,29	0,28	-0,33	-0,29	0,06	-0,19
34	-0,77	-0,53	0,37	-0,26	0,23	-0,30	-0,31	0,08	-0,23

Indolo	C <sub>19</sub>	C <sub>20</sub>	C <sub>21</sub>	C <sub>22</sub>	C <sub>23</sub>	C <sub>24</sub>	C <sub>25</sub>	C <sub>26</sub>	C <sub>27</sub>
1	-0,04	0,11	0,10	0,14	0,10	0,15	0,10	0,10	0,11
2	-0,05	0,00	0,06	0,18	0,10	0,18	0,12	0,14	0,13
3	-0,03	0,64	0,12	0,16	0,13	0,15	0,10	0,11	0,11
4	-0,04	-0,38	0,14	0,18	-0,23	0,16	0,11	0,11	0,11
5	-0,03	-0,80	0,15	0,15	0,14	0,15	0,10	0,11	0,11
6	-0,08	0,11	0,10	0,15	0,11	0,15	0,13	0,62	0,14
7	-0,19	0,10	0,98	0,14	0,10	0,15	0,14	-0,81	0,15
8	-0,20	0,11	0,11	0,14	0,10	0,22	-0,38	-0,39	0,18
9	-0,20	0,10	0,10	0,15	0,10	0,15	0,13	-0,36	0,15
10	-0,16	0,11	0,10	0,15	0,09	0,14	0,12	-0,21	0,14
11	-0,03	0,11	0,10	0,15	0,10	0,16	0,10	-0,12	0,12
12	-0,17	0,11	0,10	0,14	0,10	0,15	0,14	-0,69	0,14
13	-0,15	0,11	0,10	0,15	0,10	0,16	0,14	-0,65	0,15
14	-0,16	0,11	0,10	0,15	0,10	0,15	0,13	-0,67	0,14
15	-0,22	0,11	0,10	0,15	0,09	0,15	0,18	-0,57	0,16
16	-0,02	0,11	0,10	0,15	0,10	0,14	0,18	-0,51	0,16
17	-0,10	0,11	0,10	0,15	0,10	0,18	0,10	-0,57	0,13
18	-0,11	0,11	0,09	0,14	0,09	0,15	0,10	0,10	0,10
19	-0,07	0,10	0,09	0,13	0,10	0,16	0,10	0,10	0,09
20	-0,06	0,11	0,10	0,14	0,10	0,15	0,10	0,11	0,10
21	0,00	0,11	0,09	0,15	0,09	0,15	0,10	0,11	0,06
22	-0,06	0,11	0,10	0,15	0,09	0,15	0,10	0,11	0,08
23	-0,01	0,10	0,10	0,14	0,10	0,15	0,10	0,11	0,05
24	0,00	0,10	0,10	0,14	0,10	0,17	0,10	0,11	0,06
25	-0,03	0,10	0,10	0,15	0,10	0,15	0,10	0,11	0,08
26	-0,03	0,10	0,10	0,14	0,10	0,15	0,10	0,10	0,05
27	-0,01	0,11	0,09	0,14	0,10	0,15	0,10	0,11	0,07
28	-0,03	0,10	0,10	0,14	0,10	0,15	0,10	0,11	0,07
29	-0,02	0,11	0,10	0,15	0,09	0,15	0,10	0,11	0,06
30	-0,05	0,11	0,11	0,14	0,10	0,16	0,14	0,61	0,07
31	-0,05	0,11	0,10	0,14	0,10	0,16	0,14	-0,60	0,06
32	-0,06	0,11	0,10	0,15	0,10	0,16	-0,15	0,60	0,08
33	0,14	0,13	0,11	0,18	0,12	0,19	0,10	-0,05	0,10
34	0,17	0,13	0,11	0,18	0,12	0,20	0,10	0,04	0,03

Observa-se na Tabela IV.2 que apenas as moléculas 4 e 5 têm cargas positivas sobre o átomo 1 ( FiguralV.1). É interessante notar que os substituintes destas duas moléculas (OMe e NH<sub>2</sub>, respectivamente) estão ligados a este átomo (1), sendo provável que a presença destes grupos estejam diminuindo a densidade eletrônica do anel aromático. Pode se perceber também que as moléculas (21 a 34) que têm o oxigênio (átomo 15) na estrutura base, substituído pelo nitrogênio, apresentam menores densidade

eletrônica sobre este átomo, o que é razoável uma vez que o nitrogênio é menos eletronegativo que o oxigênio.

#### *IV.2.6 Ordens de ligação (b)*

A ordem de ligação nos fornece uma perspectiva da densidade eletrônica entre dois átomos. Quanto maior a ordem de ligação, maior a densidade eletrônica entre os átomos, e mais forte será a ligação, o que leva a uma diminuição da distância interatômicas. A Tabela IV.3 descreve as ordens de ligação entre os átomos que constituem a estrutura base das 34 moléculas analisadas [2].

A ligação é classificada como simples se o valor da ordem de ligação estiver em torno de um, dupla se este valor for aproximadamente dois e de caráter ressonante se for em torno de 1,5 e tripla se o valor for de aproximadamente 3[2].

Tabela IV.3 - Descrição de ordens de ligação obtidas a nível de teoria B3LYP/6-31G\*.

Indolo	b <sub>1 2</sub>	b <sub>1 6</sub>	b <sub>1 20</sub>	b <sub>2 3</sub>	b <sub>2 21</sub>	b <sub>3 4</sub>	b <sub>3 7</sub>	b <sub>4 5</sub>
1	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
2	1,44	1,35	1,05	1,35	0,89	1,29	1,05	1,32
3	1,42	1,33	0,93	1,36	0,88	1,29	1,05	1,32
4	1,46	1,24	0,99	1,33	0,89	1,29	1,07	1,32
5	1,36	1,29	1,15	1,39	0,90	1,27	1,05	1,31
6	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
7	1,47	1,39	0,91	1,36	0,90	1,29	1,06	1,34
8	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
9	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
10	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
11	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
12	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
13	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
14	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
15	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
16	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
17	1,47	1,38	0,91	1,36	0,90	1,29	1,06	1,33
18	1,48	1,38	0,91	1,34	0,90	1,30	1,07	1,49
19	1,48	1,38	0,91	1,34	0,90	1,30	1,07	1,33
20	1,48	1,38	0,91	1,34	0,90	1,29	1,07	1,33
21	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33
22	1,48	1,38	0,91	1,35	0,90	1,29	1,04	1,33
23	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33
24	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33
25	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33
26	1,48	1,37	0,91	1,35	0,90	1,29	1,07	1,32
27	1,48	1,37	0,91	1,35	0,90	1,29	1,07	1,32
28	1,48	1,37	0,91	1,35	0,90	1,29	1,07	1,32
29	1,48	1,37	0,91	1,35	0,90	1,29	1,07	1,32
30	1,48	1,38	0,91	1,35	0,90	1,28	1,07	1,33
31	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33
32	1,48	1,38	0,91	1,35	0,90	1,28	1,07	1,33
33	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33
34	1,48	1,38	0,91	1,35	0,90	1,29	1,07	1,33

Indolo	b4 10	b5 22	b5 6	b7 11	b7 8	b8 9	b8 12	b9 10	b9 14
1	1,15	0,90	1,48	1,67	1,02	1,06	1,00	1,66	0,96
2	1,16	0,90	1,48	1,67	1,03	1,06	1,03	1,66	0,96
3	1,16	0,90	1,49	1,68	1,02	1,08	0,99	1,65	0,95
4	1,16	0,90	1,43	1,66	1,02	1,07	1,00	1,65	0,96
5	1,17	0,90	1,52	1,66	1,04	1,05	1,00	1,65	0,97
6	1,15	0,90	1,48	1,67	1,02	1,06	1,00	1,67	0,96
7	1,15	0,90	1,48	1,68	1,01	1,05	1,02	1,67	0,96
8	1,15	0,90	1,48	1,66	1,03	1,07	0,99	0,13	0,96
9	1,15	0,90	1,48	1,66	1,03	1,06	1,00	1,67	0,95
10	1,15	0,90	1,48	1,66	1,03	1,07	0,99	1,66	0,96
11	1,15	0,90	1,48	1,67	1,02	1,06	1,00	1,67	0,96
12	1,15	0,90	1,48	1,67	1,02	1,06	1,00	1,66	0,96
13	1,15	0,90	1,48	1,67	1,02	1,06	1,00	1,66	0,96
14	1,15	0,91	1,48	1,67	1,02	1,06	1,00	1,66	0,96
15	1,15	0,90	1,48	1,67	1,02	1,06	1,00	1,66	0,96
16	1,15	0,90	1,48	1,67	1,03	1,07	0,99	1,66	0,96
17	1,15	0,90	1,48	1,67	1,03	1,06	0,99	1,66	0,96
18	1,14	0,91	1,49	1,66	1,01	1,08	0,99	1,67	0,95
18	1,14	0,91	1,49	1,66	1,02	1,08	0,99	1,67	0,95
20	1,14	0,91	1,49	1,66	1,02	1,08	0,99	1,68	0,93
21	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,63	1,03
22	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,03
23	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,63	1,03
24	1,16	0,91	1,49	1,66	1,02	1,06	1,00	1,63	1,03
25	1,16	0,91	1,49	1,66	1,02	1,07	1,00	1,63	1,03
26	1,17	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,04
27	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,04
28	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,04
29	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,04
30	1,16	0,91	1,49	1,67	1,01	1,05	1,03	1,63	1,04
31	1,16	0,91	1,49	1,67	1,01	1,05	1,03	1,63	1,04
32	1,15	0,90	1,48	1,67	1,00	1,05	1,03	1,63	1,04
33	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,04
34	1,16	0,91	1,49	1,66	1,02	1,06	1,01	1,62	1,04

Indolo	b <sub>12 13</sub>	b <sub>13 14</sub>	b <sub>12 16</sub>	b <sub>13 19</sub>	b <sub>14 15</sub>	b <sub>16 17</sub>	b <sub>16 24</sub>	b <sub>17 25</sub>	b <sub>17 18</sub>
1	1,30	1,05	1,40	1,38	1,78	1,42	0,89	0,91	1,43
2	1,30	1,05	1,40	1,38	1,78	1,42	0,89	0,91	1,43
3	1,30	1,04	1,41	1,38	1,79	1,41	0,89	0,91	1,43
4	1,30	1,05	1,40	1,38	1,78	1,42	0,89	0,91	1,43
5	1,30	1,04	1,40	1,38	1,77	1,42	0,89	0,91	1,43
6	1,31	1,04	1,39	1,38	1,79	1,42	0,89	0,90	1,40
7	1,29	1,03	1,38	1,40	1,80	1,43	0,88	0,89	1,37
8	1,30	1,04	1,39	1,39	1,78	1,44	0,89	0,91	1,34
9	1,32	1,06	1,37	1,31	1,76	1,39	0,88	1,02	1,27
10	1,32	1,04	1,37	1,36	1,78	1,46	0,89	0,91	1,34
11	1,31	1,04	1,39	1,38	1,79	1,42	0,89	0,90	1,39
12	1,31	1,04	1,38	1,38	1,79	1,44	0,89	0,91	1,36
13	1,31	1,04	1,39	1,38	1,79	1,44	0,89	0,91	1,36
14	1,30	1,04	1,39	1,39	1,79	1,43	0,89	0,91	1,43
15	1,30	1,04	1,39	1,39	1,79	1,43	0,89	0,88	1,36
16	1,30	1,04	1,39	1,39	1,78	1,43	0,89	0,88	1,35
17	1,31	1,04	1,38	1,38	1,79	1,44	0,89	0,91	1,35
18	1,33	0,98	1,40	1,43	0,99	1,42	0,89	0,91	1,45
19	1,33	0,98	1,40	1,43	0,98	1,42	0,89	0,91	1,45
20	1,33	0,97	1,40	1,42	0,93	1,42	0,89	0,91	1,44
21	1,30	1,08	1,40	1,38	1,71	1,42	0,89	0,91	1,43
22	1,29	1,08	1,40	1,38	1,68	1,41	0,89	0,91	1,43
23	1,30	1,08	1,40	1,38	1,70	1,42	0,89	0,91	1,43
24	1,30	1,08	1,40	1,38	1,71	1,42	0,89	0,91	1,43
25	1,29	1,08	1,41	1,37	1,71	1,41	0,89	0,91	1,43
26	1,30	1,08	1,40	1,38	1,69	1,42	0,89	0,91	1,43
27	1,30	1,08	1,40	1,38	1,69	1,42	0,89	0,91	1,43
28	1,30	1,08	1,40	1,38	1,69	1,42	0,89	0,91	1,43
29	1,30	1,08	1,40	1,38	1,69	1,42	0,89	0,91	1,43
30	1,28	1,07	1,39	1,40	1,68	1,43	0,88	0,89	1,37
31	1,28	1,07	1,40	1,40	1,69	1,43	0,88	0,89	1,37
32	1,28	1,08	1,39	1,39	1,68	1,43	0,88	0,89	1,37
33	1,29	1,08	1,40	1,38	1,69	1,41	0,89	0,90	1,40
34	1,29	1,08	1,40	1,38	1,69	1,31	0,89	0,90	1,40



Indolo	b <sub>18 19</sub>	b <sub>18 26</sub>	b <sub>19 27</sub>	b <sub>6 23</sub>
1	1,44	0,91	0,90	0,91
2	1,44	0,91	0,90	0,90
3	1,44	0,91	0,90	0,89
4	1,44	0,91	0,90	1,03
5	1,44	0,91	0,90	0,91
6	1,41	1,05	0,89	0,91
7	1,38	0,93	0,88	0,91
8	1,36	1,13	0,91	0,91
9	1,45	0,99	0,89	0,91
10	1,39	1,02	0,90	0,91
11	1,41	1,06	0,89	0,91
12	1,38	1,05	0,90	0,91
13	0,38	1,04	0,90	0,91
14	1,38	1,04	0,90	0,91
15	1,38	1,04	0,90	0,91
16	1,38	1,06	0,90	0,91
17	1,39	1,06	0,90	0,91
18	1,42	0,91	0,89	0,91
19	1,42	0,91	0,91	0,91
20	1,42	0,91	0,91	0,91
21	1,43	0,91	0,90	0,91
22	1,43	0,91	0,90	0,91
23	1,43	0,91	0,90	0,91
24	1,43	0,91	0,90	0,91
25	1,44	0,91	0,90	0,91
26	1,43	0,91	0,90	0,91
27	1,43	0,91	0,90	0,91
28	1,43	0,91	0,90	0,91
29	1,43	0,91	0,90	0,91
30	1,37	0,94	0,88	0,91
31	1,37	0,94	0,88	0,91
32	1,37	0,94	0,88	0,91
33	1,41	1,04	0,89	0,91
34	1,41	1,04	0,89	0,91

A ordem de ligação entre os átomos 14 e 15, para os compostos: 18,19 e 20 é aproximadamente 1, enquanto para todos demais compostos é entorno de 1,7. Tendo como base os resultados teóricos, conclui-se que as moléculas (18,19 e 20) possuem ligação simples entre os átomos 14 e 15, e as outras 31 moléculas possuem dupla ligação entre estes átomos, o que pode ser

verificado observando a fórmula estrutural das moléculas descritas nas Figuras 1.2 e 1.3.

#### *IV. 2.7 Ângulos ( $\alpha$ )*

O ângulo formado entre três átomos é um parâmetro estérico de importância relevante no estudo da relação estrutura atividade da molécula. Os ângulos formados entre os átomos estão diretamente relacionado com a geometria da molécula, e esta geometria interfere na atividade da droga com o seu receptor [37].

Tabela IV.4 Ângulos obtidas a nível de teoria B3LYP/6-31G\*. Todos os valores estão em graus (°).

Indolo	a1 (6,4,23)	a2 (23,6,1)	a3 (6,1,20)	a4 (20,1,4)	a5 (3,7,11)
1	119,70	119,78	119,93	119,94	125,49
2	120,59	119,76	119,05	119,65	125,41
3	121,71	119,22	118,93	120,28	125,46
4	124,57	115,67	121,21	119,06	125,80
5	119,87	119,13	120,26	121,01	125,19
6	119,72	119,76	119,90	119,94	125,71
7	119,71	119,77	119,91	119,92	125,83
8	119,74	119,76	119,90	119,97	125,57
9	119,71	119,77	119,92	119,97	125,53
10	119,73	119,76	119,90	119,97	125,61
11	119,72	119,76	119,90	119,94	125,71
12	119,72	119,77	119,91	119,95	125,71
13	119,71	119,77	119,91	119,95	125,70
14	119,72	119,77	119,90	119,95	125,71
15	119,73	119,77	119,91	119,95	125,66
16	119,73	119,77	119,91	119,96	125,58
17	119,72	119,77	119,91	119,95	125,67
18	119,66	119,71	119,97	120,09	125,61
19	118,66	119,72	119,97	120,09	125,60
20	119,66	119,72	119,96	120,06	125,62
21	119,66	119,72	119,97	120,07	125,60
22	124,57	115,67	121,21	119,06	125,81
23	119,66	119,72	119,97	120,07	125,62
24	119,67	119,71	119,96	120,07	125,59
25	119,67	119,73	119,96	120,05	125,60
26	119,66	119,71	119,98	120,09	125,60
27	119,66	119,71	119,98	120,08	125,60
28	119,66	119,71	119,98	120,09	125,59
29	119,66	119,71	119,98	120,09	125,58
30	119,64	119,71	119,98	120,06	125,90
31	119,63	119,73	119,97	120,03	125,99
32	119,63	119,73	119,97	120,03	125,99
33	119,66	119,71	119,98	120,61	125,79
34	119,66	119,70	119,97	120,09	125,79

Indolo	a <sub>6</sub> (11,7,10)	a <sub>7</sub> (16,17,25)	a <sub>8</sub> (25,17,18)	a <sub>9</sub> (17,18,26)	a <sub>10</sub> (26,18,19)
1	122,33	118,52	119,29	119,77	120,09
2	122,54	118,56	119,29	121,85	119,83
3	122,55	118,52	119,32	119,75	120,05
4	122,16	118,55	119,27	119,77	120,11
5	122,54	118,59	119,27	119,74	120,11
6	122,14	119,42	119,31	119,16	119,50
7	122,00	120,60	118,63	118,78	118,90
8	122,28	118,65	118,69	120,46	120,81
9	122,31	119,97	119,92	119,77	119,71
10	122,22	118,13	120,18	124,52	115,60
11	122,16	119,53	119,23	119,19	119,47
12	122,16	118,73	118,49	118,15	122,84
13	122,16	118,87	118,48	118,49	122,31
14	122,15	118,41	119,21	122,40	118,54
15	122,21	119,78	118,77	123,02	117,23
16	122,28	119,98	118,42	122,79	117,70
17	122,19	118,72	119,05	120,23	120,43
18	121,99	118,89	119,60	119,92	119,81
19	122,01	118,90	119,61	119,92	119,80
20	122,04	118,87	119,63	119,96	119,71
21	122,01	118,81	119,57	119,77	119,53
22	122,16	118,55	119,27	119,77	120,11
23	122,00	118,80	119,57	119,78	119,54
24	122,03	118,80	119,57	119,77	119,54
25	122,04	118,76	119,53	119,79	119,48
26	122,01	118,82	119,59	119,77	119,54
27	122,00	118,82	119,59	119,77	119,54
28	122,01	118,82	119,59	119,77	119,53
29	122,02	118,82	119,59	119,76	119,54
30	121,72	120,80	118,97	118,77	118,48
31	121,67	120,78	118,91	118,77	118,49
32	121,67	120,78	118,91	118,77	118,49
33	121,85	119,70	119,68	119,10	118,92
34	119,70	119,71	119,67	119,08	118,93

Como pode ser visto na Tabela IV.4 os ângulos obtidos estão compreendidos em uma faixa de 119° a 126°. Contudo, a variabilidade observada em um mesmo ângulo, comparando as 34 moléculas, não ultrapassa a 2°, esta pequena variação se justifica pela rigidez da estrutura base.

#### IV. 2.8 Ângulos diedrais (d)

Outro parâmetro estérico calculado foram os ângulos diedrais, que são ângulos externos formados por 4 átomos [37]. Apenas 5 destes ângulos foram calculados, aqueles localizados onde há maior número de substituição na molécula, os valores estão relacionados na Tabela IV.11. Como a estrutura é muito rígida não observou uma grande variação nestes ângulos.

Tabela IV.5 Ângulos Diedrais obtidas a nível de teoria B3LYP/6-31G\*. Todos os valores estão em graus (°).

Indolo	d1	d2	d3	d4
1	-180,00	0,00	0,00	0,03
2	-180,00	0,00	0,00	0,00
3	180,00	0,00	0,00	0,01
4	-178,42	2,96	0,01	0,03
5	-179,68	-2,51	0,00	0,01
6	180,00	0,00	0,00	0,01
7	180,00	0,00	0,00	-0,01
8	177,09	-0,44	0,00	0,00
9	-180,00	0,00	-0,01	-0,12
10	180,00	0,00	0,00	0,00
11	180,00	0,01	0,00	0,05
12	177,93	0,62	0,00	1,34
13	-179,99	-0,02	2,72	-0,01
14	179,99	-0,01	0,14	-0,60
15	-180,00	0,00	0,00	0,00
16	-179,99	-0,01	0,23	0,03
17	179,99	0,00	0,42	0,14
18	179,97	-0,08	-0,11	54,01
19	179,95	-0,06	-0,12	52,65
20	179,97	-1,07	-0,24	-63,96
21	180,00	0,00	0,00	0,00
22	-178,43	2,96	0,01	0,00
23	179,98	0,00	-0,02	0,01
24	179,98	0,00	0,16	-0,03
25	-180,00	0,00	0,00	-0,01
26	-179,99	-0,01	0,00	0,17
27	179,99	-0,01	0,09	0,07
28	179,97	0,01	0,02	0,18
29	179,99	0,00	0,00	-0,12
30	179,99	0,00	0,03	-0,34
31	179,99	0,00	-0,06	-0,08
32	-179,99	0,01	0,04	1,25
33	179,98	0,00	0,03	0,10
34	179,98	0,01	-0,07	0,09

#### *IV.2.9 Volume molar (V)*

O volume molar ajuda avaliar o efeito estérico causado pela troca do substituinte. O volume médio encontrado foi de 886.0394 , como mostra a Tabela IV.6. O desvio padrão para esta variável foi 141.1285 , valor alto. Conclui-se que o volume das moléculas varia em uma ampla faixa e, portanto, o tamanho dos substituintes é bastante variável [37].

#### *IV.2.10 Coeficiente de partição (Log P)*

A lipofilicidade de uma droga é medida pelo coeficiente de partição, que é a razão da concentração de equilíbrio da droga em meio orgânico ( $C_{org}$ ) e ( $C_{aq}$ ) e em meio aquoso [37].

$$\text{LogP} = C_{org} / C_{aq}$$

O solvente orgânico mais utilizado é o octanol devido a sua similaridade com substâncias biológicas, ambos possuem uma longa cadeia alquil e grupos polares [38], além de que grande parte dos compostos é solúvel em octanol e existe na literatura valores de LogP para inúmeros drogas considerando este solúvel orgânico[5].

Valores de LogP maior que 1, indica que a droga é mais solúvel em meio orgânico, valores de LogP menor que 1 indica que a droga é mais solúvel em meio aquoso.

O composto 25 é o que apresenta maior valor de LogP ( 5,02), portanto é o mais solúvel em meio orgânico. Já o composto 16 apresenta o menor coeficiente de partição (0,76), sendo, portanto o mais solúvel em meio aquoso, uma análise minuciosa da Tabela IV.6 nos leva a concluir que é o único composto mais solúvel em meio aquoso do que em meio orgânico, no grupo das 34 moléculas.

#### *IV.2.11 Momento de dipolo (*

O momento de dipolo pode ser definido como um vetor que mede a magnitude da densidade de carga que desloca em uma ligação formada por

átomos diferentes. A direção do momento de dipolo em uma molécula depende da eletronegatividade dos átomos presentes. Sendo que seu valor é obtido pelo vetor resultante dos momentos de dipolo de cada ligação presente na molécula.

Compostos com substituintes de diferentes eletronegatividade resultam em momento de dipolo diferente. Como consequência alteram as propriedades químicas da substância [37].

Tabela IV.6 Momento de dipolo ( $\mu$ ); obtidas a nível de teoria B3LYP/6-31G\*. O Volume(V), coeficiente de partição(LogP) e Polarizabilidade( $\alpha$ ) foram calculadas pela rotina do pacote Hyperchem7.5.

Indolo	V ( $\text{\AA}^3$ )	LogP	(debye)	( $\text{\AA}^3$ )
1	677,71	2,26	2,01	160,50
2	740,97	3,05	1,81	178,10
3	736,01	2,22	4,79	178,20
4	830,20	1,76	1,92	194,50
5	712,66	1,48	4,25	181,60
6	736,77	2,22	6,25	180,10
7	713,72	1,48	2,87	178,20
8	743,89	1,95	2,72	180,50
9	753,90	2,01	3,26	179,23
10	731,12	2,73	2,09	171,45
11	722,47	2,78	3,09	175,46
12	818,03	1,11	3,85	204,06
13	922,19	2,14	3,76	218,04
14	873,23	1,74	3,90	207,23
15	863,62	1,68	0,57	209,65
16	1103,70	0,76	2,77	252,83
17	864,37	2,05	2,47	218,49
18	741,31	2,43	2,46	162,80
19	893,07	3,69	2,55	183,26
20	1053,21	4,65	2,03	231,90
21	713,52	2,92	1,84	174,85
22	842,89	2,51	1,99	209,03
23	913,10	2,50	1,47	209,04
24	1073,06	3,39	2,97	260,72
25	1083,25	5,02	3,26	274,75
26	1020,97	2,98	2,54	224,36
27	954,37	2,93	2,22	215,72
28	1033,30	3,26	2,39	228,44
29	1086,27	3,31	3,31	236,91
30	1013,24	2,88	4,64	235,22
31	1080,37	2,94	4,95	242,79
32	902,24	2,46	3,71	217,51
33	1032,28	3,72	2,67	233,96
34	1144,33	4,44	2,82	254,91
Média	886,04	2,63	2,95	207,77
Variância	19917,26	0,87	1,22	881,72
Desvio Padrão	141,13	0,93	1,11	29,69

O momento de dipolo está relacionado com a polaridade, quanto maior for o momento, mais polar é a molécula. Observando pela Tabela IV.6, a molécula 6 é a que possui maior momento de dipolo e portanto é a mais polar.



A molécula 15 é a mais apolar, uma vez que possui o menor momento de dipolo.

#### *IV.2.12 Polarizabilidade ( )*

A polarizabilidade mede a facilidade com que uma nuvem eletrônica de um átomo ou molécula é distorcida quando um campo elétrico externo é aplicado. A propriedade mais significativa da polarizabilidade molecular é a sua relação com o volume ou tamanho da molécula, sendo que quanto maior a molécula maior será sua polarizabilidade [5].

O composto 1 é o mais rígido uma vez que apresenta o menor valor de polarizabilidade, o que é facilmente explicável uma vez que este composto é formado apenas por anéis planares.

### *IV. 3. ANÁLISE ESTATÍSTICA DOS DADOS*

Os resultados dos cálculos teóricos serão relacionados com a atividade apresentada pelos compostos no combate ao câncer de mama. Nesta análise serão empregadas as seguintes técnicas: pesos de Fisher e análise de componentes principais, os *softwares* utilizados foram: Unscrambler 9.1 e Microsoft Office Excel 2007.

#### *IV. 3.1 Análises da atividade contra o câncer de mama*

O cálculo do peso Fisher é um método que relaciona a molécula com sua atividade, esta técnica foi a primeira a ser empregada nas análises, uma vez que permite ter-se uma visão preliminar de quais variáveis apresentam maior significado na separação das moléculas em ativas e inativas.

Tabela IV.7 Pesos de Fisher

Variáveis	Peso Fisher	Variáveis	Peso Fisher	Variáveis	Peso Fisher	Variáveis	Peso Fisher
						LUMO	
C1	0,18	C22	0	b <sub>9 10</sub>	0,09	+1	0,05
C2	0,01	C23	0,06	b <sub>9 14</sub>	1,9		0,3
C3	0,01	C24	0,02	b <sub>12 13</sub>	1,83		0,3
C4	0,07	C25	0,08	b <sub>12 16</sub>	0,98		0,01
C5	0	C26	0	b <sub>13 14</sub>	0,03	a <sub>1</sub>	0,04
C6	0,02	C27	0,31	b <sub>13 19</sub>	0	a <sub>2</sub>	0,03
C7	0	b <sub>1 2</sub>	0,32	b <sub>14 15</sub>	0,18	a <sub>3</sub>	0,03
C8	0,01	b <sub>1 6</sub>	0,27	b <sub>16 17</sub>	0,01	a <sub>4</sub>	0
C9	0,01	b <sub>1 20</sub>	0,28	b <sub>16 24</sub>	0,03	a <sub>5</sub>	0,52
C10	0,21	b <sub>2 3</sub>	0,06	b <sub>17 25</sub>	0,22	a <sub>6</sub>	0,37
C11	0,03	b <sub>2 21</sub>	0,26	b <sub>17 18</sub>	0	a <sub>7</sub>	0,31
C12	0	b <sub>3 4</sub>	0,07	b <sub>18 19</sub>	0,09	a <sub>8</sub>	0,04
C13	0	b <sub>3 7</sub>	0,12	b <sub>18 26</sub>	0,06	a <sub>9</sub>	0,09
C14	0,55	b <sub>4 5</sub>	0,07	b <sub>19 27</sub>	0,24	a <sub>10</sub>	0,01
C15	1,12	b <sub>4 10</sub>	0,25	b <sub>6 23</sub>	0,04	a <sub>11</sub>	0
C16	0,05	b <sub>5 22</sub>	0,43	V	0,79	a <sub>12</sub>	0,76
C17	0,07	b <sub>5 6</sub>	0,04	LogP	0,01	d <sub>1</sub>	0,02
C18	0,01	b <sub>7 11</sub>	0		0,04	d <sub>2</sub>	0,02
C19	0,13	b <sub>7 8</sub>	0,21	HOMO	0,16	d <sub>3</sub>	0,07
				HOMO			
C20	0,07	b <sub>8 9</sub>	0,3	-1	0,25	d <sub>4</sub>	0,02
C21	0,04	b <sub>8 12</sub>	0,36	LUMO	0,02	d <sub>5</sub>	0,02
	1,03						

Como descrito na Tabela IV.7, a variável que apresentou maior valor de peso Fisher (1,9) foi a ordem de ligação entre os átomos 9 e 14 (b<sub>9 14</sub>). Como a maior parte das variáveis apresentou baixos valores de peso Fisher (inferior a 0,1), neste trabalho foi adotado o seguinte critério: apenas as variáveis com peso de Fisher superior a 0,2 serão consideradas significativas na discriminação das moléculas em ativas e inativas. Assim foram selecionadas as seguintes variáveis: C10, C14, C15, C27, b<sub>1\_2</sub>, b<sub>1\_6</sub>, b<sub>2\_21</sub>, b<sub>1\_20</sub>, b<sub>4\_10</sub>, b<sub>5\_22</sub>, b<sub>7\_8</sub>, b<sub>8\_9</sub>, b<sub>8\_12</sub>, b<sub>9\_14</sub>, b<sub>12\_13</sub>, b<sub>12\_16</sub>, b<sub>14\_15</sub>, b<sub>17\_25</sub>, b<sub>19\_27</sub>, V, HOMO-1, , , a<sub>5</sub>, a<sub>6</sub>, a<sub>7</sub> e a<sub>12</sub>. Desta forma a matriz de dados que inicialmente continha 84 variáveis foi reduzida para 28 variáveis. Utilizando a técnica de PCA várias combinações foram testadas, a fim de encontrar os descritores que separassem as moléculas em duas classes. O resultado mais

satisfatório obtido foi utilizando os descritores: C15, b12\_13, b17\_25 e, entretanto algumas moléculas inativas ficaram misturadas no grupo das ativas.

Para obter uma separação mais nítida entre as duas classes (inativas e ativas) alguns descritores que não tinham sido pré - selecionados pela técnica de pesos Fisher foram incluídos na tabela de dados. O critério utilizado para inclusão destas variáveis na tabela de dados, na qual estava sendo aplicada a técnica de PCA, foi a observação das estruturas das moléculas. Foram incluídos descritores que estivessem relacionados com locais onde a molécula base sofreu maior número de substituições, (região do átomo 15 e 18). Como já havia sido escolhido um descritor na região do átomo 15, foram adicionados descritores relacionados ao átomo 18. Foram testados: a ordem de ligação entre os átomos 18 e 26 (b18\_26), a carga sobre o átomo 18 (C18) e a carga sobre o átomo 26 (C26). Desta forma foi obtido a separação das moléculas em ativas e inativas adicionando a ordem de ligação entre os átomos 18 e 26 no grupo das outras 4 variáveis (C15, b12\_13, b17\_25 e Gap) que já tinham sido pré - selecionadas.

A Figura IV.2 mostra a composição da PC1 e PC2 em função das amostras.

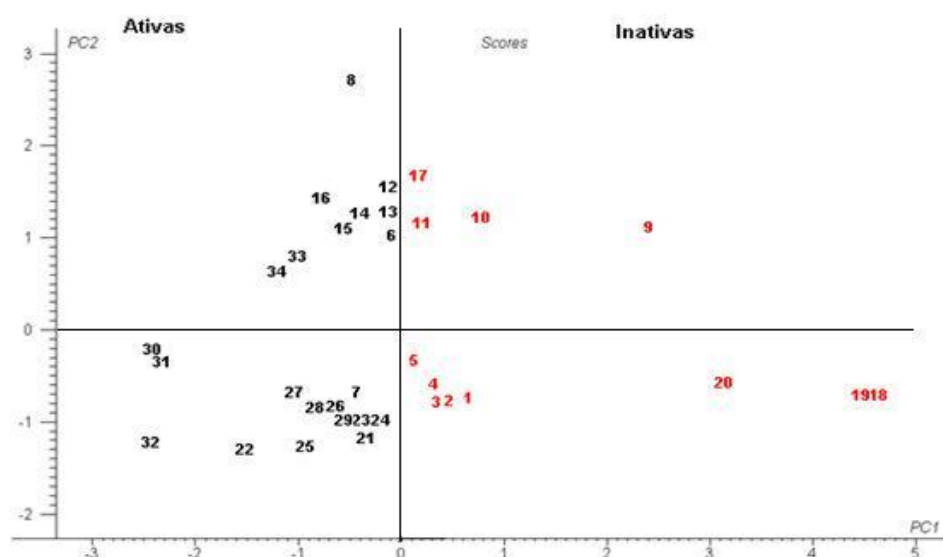


Figura IV.2 Representação gráfica dos scores obtidos usando o programa Unscrambler 9.1.

A primeira componente explica 50.020% (Tabela IV.8) da variância total dos dados. Considerando as duas primeiras componentes a variância explicada foi de 74.512%, considerando as três primeiras componentes 93.740%, e acumulando as quatro primeiras componentes 97.940%.

Tabela IV.8 Porcentagem de variância explicada pelas componentes principais obtidos pelo programa Unscrambler 9.1.

Componentes	Porcentagem Individual	Porcentagem acumulada
PC_01	50,02	50,02
PC_02	24,49	74,51
PC_03	19,23	93,74
PC_04	4,20	97,94
PC_05	2,06	100,00

A Figura IV.3 mostra que a variável b<sub>18 26</sub>, não possui peso significativo na primeira componente, no entanto é a variável de maior peso na segunda PC.

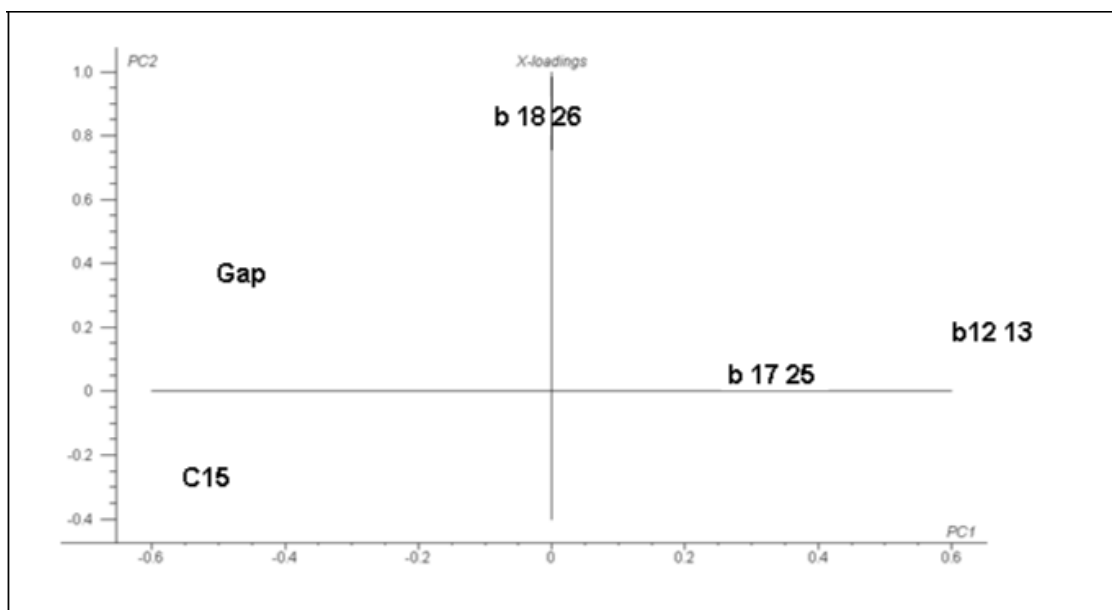


Figura IV.3 Representação gráfica dos loading obtidos usando o programa Unscrambler 9.1.

A combinação linear da primeira componente que é a responsável pela separação, pode ser escrita como:

$$7 \quad 0,08 \quad (IV.1)$$

A variável C15 que representa a carga sobre o átomo 15, possui peso (loading) significativo (Tabela IV.9) e a carga do átomo 15 também é negativa, o que permite afirmar que devem aumentar (no sentido positivo) os valores numéricos desta variável para que as moléculas façam parte do grupo das ativas.

Os descritores b12 13 e b17 25 que representam respectivamente a ordem de ligação entre os átomos 12 e 13 e a ordem de ligação entre os átomos 17 e 25, têm pesos mais significativos na primeira componente. Como os valores de ordem de ligação são positivos e seus coeficientes também são positivos ( Figura IV.3 e Tabela IV.9), estas variáveis devem ter seus valores diminuídos para que as moléculas faça parte do grupo das ativas.

Apesar da variável b18 26 (ordem de ligação entre os átomos 18 e 26), apresentar baixo peso para a primeira componente, esta propriedade contribui para discriminação das moléculas, a não inclusão deste descritor torna inseparável o grupo das ativas e inativas. Como o coeficiente desta variável é negativo ( Figura IV.3) para a primeira componente e valores de ordem de ligação são positivos, para as moléculas faça parte do grupo das ativas precisam ter seu valor aumentado.

A variável Gap (que tem valor negativo) possui coeficiente negativo também (Figura IV.3 e Tabela IV.9), para as moléculas faça parte do grupo das ativas, precisa aumentar ( no sentido positivo) o valor desta variável na primeira componente.

Tabela IV.9 Loading das variáveis para as 2 primeiras componentes obtidos pelo programa Unscrambler 9.1.

Variáveis	PC_01	PC_02
c15	-0,55	-0,26
b12.13	0,60	0,18
b17.25	0,27	0,04
b18.26	-0,08	0,87
Gap	-0,50	0,38

Analisando o gráfico dos scores nas duas primeiras componentes pode-se destacar alguns grupos formado por moléculas que possuem similaridades, as quais serão explicadas detalhadamente. Grupo G1 formado pelos

compostos: 7,21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,33 e 34. É interessante observar que estas moléculas (exceto a 7) tem o oxigênio (átomo 15) substituído pelo nitrogênio. Uma vez que o nitrogênio é menos eletronegativo que o oxigênio, a densidade eletrônica sobre o átomo 15 para este grupo de molécula é menor que dos demais compostos analisados. Como foi visto na análise anterior, um dos descritores responsável pela separação das moléculas em ativas e inativas é a carga sobre o átomo 15, sendo importante observar que para as moléculas façam parte do grupo das ativas, esta variável precisa ter seu valor numérico aumentado, isto equivale a diminuir a densidade eletrônica sobre este átomo. Possivelmente esta é uma das regiões da droga que esteja interagindo com o receptor biológico.

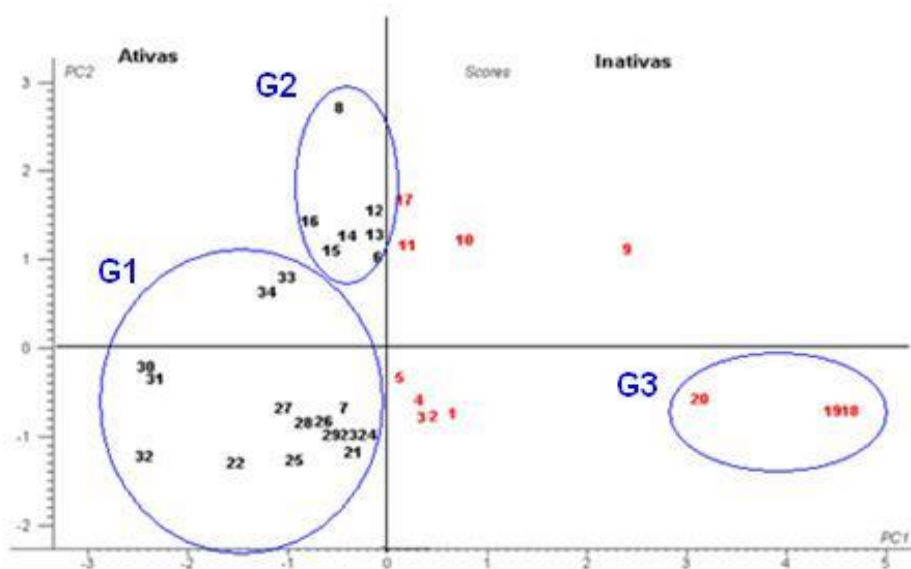


Figura IV.4 Representação gráfica dos scores, usando o programa Unscrambler 9.1.

O grupo formado pelos compostos 30, 31 e 32, pode ser explicado pelo substituinte NO<sub>2</sub> que estas três moléculas têm em comum na posição R1 (ligado ao átomo 18). O que pode explicar também a presença da molécula 7 neste quadrante do gráfico. É importante salientar que o átomo 15 do composto 7 é um oxigênio, isto é, este átomo não foi substituído por um nitrogênio como nas demais moléculas deste grupo. Mesmo assim, a densidade eletrônica entre

os átomos 12 e 13 é menor para molécula 7 do que para as moléculas que não fazem parte de grupo G1. Podendo inferir a hipótese de que o grupo retirador de elétrons (NO<sub>2</sub>) esteja contribuindo para diminuir a densidade eletrônica entre estes átomos desta molécula.

Uma semelhança dos compostos 33 e 34 é a presença do bromo ligado ao átomo 18. Esta característica pode ser responsável pelo fato destas moléculas estarem mais acima do grupo de compostos: 7, 21 a 32.

Os compostos: 18,19 e 20 encontram-se bem afastados do grupo das moléculas ativas (Figura IV.4), destacando sua pronunciada inatividade no combate ao câncer de mama. Estas três moléculas diferem dos demais compostos analisados, por não apresentar o grupo carbonila formado pelos átomos 14 e 15. O ataque de um grupo nucleofílico a carbonila transformou-a em um álcool, observa-se que o ataque nucleofílico torna as moléculas mais inativas.

As moléculas do grupo G2 têm em comum, o substituinte ligado ao átomo 18 ou ao átomo 26, sendo que estes átomos estão interligados. Com exceção da molécula 8, todas demais substâncias deste conjunto (G2), o substituinte possui uma carbonila.

Partindo da mesma base estrutural investigada no trabalho, conclui-se que para que a moléculas seja uma possível candidata a um novo fármaco contra o câncer de mama, devem-se diminuir o valor das variáveis: b<sub>12 13</sub>, b<sub>17 25</sub> e aumentar os valores das variáveis: C<sub>15</sub>, Gap e b<sub>18 26</sub>. A variável b<sub>18 26</sub> apresentou baixa influência para separação das moléculas em ativas e inativas, porém a presença desta variável é indispensável para nítida discriminação das amostras.

Não é fácil prever a interação que pode haver entre a droga e seu receptor uma vez que não conhecemos o sítio do seu receptor alvo, mas com as informações até aqui obtidas pode se inferir algumas hipóteses. A análise de PCA indica que para a molécula se tornar mais ativa precisa ter os valores numéricos de carga no átomo 15 aumentado, ou seja, a densidade eletrônica sobre este átomo deve ser baixa, os valores das ordens de ligação entre os átomos 12-13 e 17-25 também devem ser baixos, isto também significa menor densidade eletrônica nestas regiões da molécula. Ao adicionar um grupo retirador de elétrons no átomo 18, a densidade eletrônica sobre e o átomo 15

diminui. Pelos dados da Tabela IV.3 que refere as ordens de ligação, pode-se prever um caráter ressonante entre as ligações que vai do átomo 13 ao 18. O que permite uma especulação, uma vez que analisando a mesma tabela observa-se que os átomos (7,31 e 32) que possuem o grupo retirador de elétrons (NO<sub>2</sub>) ligado ao átomo 18, são os que apresentam menor ordem de ligação entre os átomos 12-13, isto significa que possuem maior caráter de ligação simples, quando comparado com as demais moléculas estudadas. Supondo assim a formação de um carbocátion no átomo 13 (Figura IV.5), que pode estar contribuindo para diminuir a densidade eletrônica sobre o átomo 15, através da transferência de carga do átomo 15 para o 13. Assim pode-se inferir a hipótese de que o grupo farmacofórico esteja na região da molécula que compreende os átomos 13 e 15. E possivelmente a interação entre a droga e o receptor ocorre através da transferência de cargas ou por formação de ligação, isto se o sítio do receptor tiver alta densidade eletrônica.

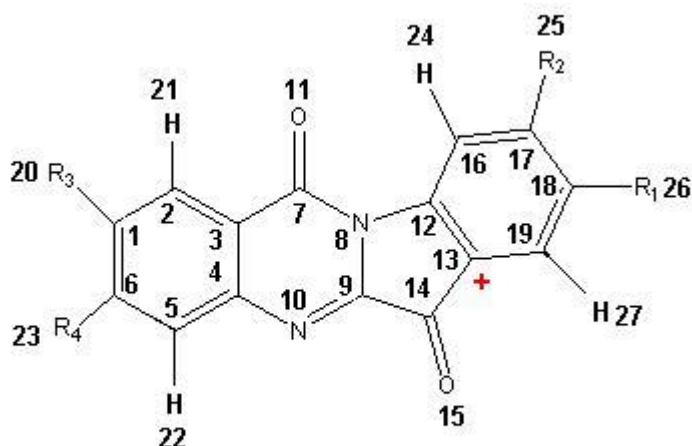


Figura IV.5 Representação da formação de carbocátion devido a adição de um grupo retirador de elétrons, obtidos pelo programa Unscrambler 9.1.



## **CAPÍTULO V**

### **CONCLUSÃO**

Os resultados dos cálculos teóricos usando a teoria do funcional da densidade(DFT) com o funcional de troca e correlação híbrido B3LYP e o conjunto de base 6-31G\* mostram-se adequados na determinação dos descritores estéricos e eletrônicos dos derivados análogos da Indolo[2,1b]quinazoline. Usando a Análise de Componentes Principais (PCA), foi possível discriminar os descritores que estão correlacionados com a atividade contra o câncer de mama.

Os descritores que melhor discriminaram as moléculas em duas classes (ativas e inativas) foram a carga sobre o átomo 15, Gap (diferença de energia do orbital mais alto ocupado e do orbital mais baixo desocupado) e as ordens de ligação entre os átomos 12-13,17-25 e 18-26. Esta última variável (b18 26) apresenta pouca influência na discriminação das moléculas, contudo, sua presença no grupo de descritores responsáveis pela separação é indispensável.

Com base nos resultados alcançados, é possível propor compostos com a mesma estrutura base e que sejam mais ativos do que os estudados, sendo necessário aumentar o valor numérico das variáveis(C15, Gap e b18 26) e diminuir o valor numérico das variáveis(b12 13 e b17 25 que foram significativas na separação das moléculas quanto à sua atividade. Vale lembrar que no processo de descoberta de um novo fármaco, os testes experimentais levam aos estudos químicos-quântico-quimiométricos que melhoram o entendimento sobre determinados compostos, propondo uma nova molécula mais eficiente, que novamente deve se submeter a testes experimentais e, assim, o processo deve ser interativo.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BERBERICE, S. **Biotech sector adapting to times**, GAZZET.NET, 2006
- [2] CAMARGO, J.C. **Estudo químico-quântico da relação estrutura atividade de compostos neolignânicos e derivados análogos contra E. coli e P.brasiliensis**. São Carlos. Dissertação de mestrado. Instituto de Química de São Carlos, Universidade de São Paulo, 147 p, 1998.
- [3] TAVARES, L.C.QSAR: A abordagem de Hansh. **Química Nova**. v.27, p.631-639, 2004
- [4] HANSH, C. Quantitative Approaches to pharmacological Structure-Activity Relationships. In: CAVALITO, C.J. **Structure-Activity Relationships**. Oxford, Pergamon, V.1, p.239, 1973.
- [5] MATI, K.; LABONOV, V.S. Quantum-Chemical Descriptors in QSAR/QSPR Studies. **Journal American Chemical Society**. v.96.p.1027-1043,1996
- [6] SHARMA, V.M.; PRASANNA, P.; SESHU, K.V.A.; RENUKA, B. Novel Indolo [ 2,1b ] quinazoline analogues as cytostatic agents synthesis, biological evaluation and structure-activity relationship. **Bioorganic & Medicinal Chemistry**. v.12, p. 2303-2307, 2002.
- [7] WORLD HEALTH ORGANIZATION, **New guide on palliative care services for people living with advanced cancer**, 51p, 2007.
- [8] ALMEIDA, V.L.; LEITÃO, A.; BARRET, L.C.; MONTANARI, C.A.; DONACI, C.L.; LOPES, M.T.P. Câncer e agente antineoplástico ciclo-celular específicos e ciclo celular não específicos que interagem com o DNA: Uma Introdução. **Química Nova**, v.28 N° 1, 118-129; 2005.
- [9] PINHO, V. F. S.; COUTINHO, E.S.F. Risk factors for breast cancer: a systematic review of studies with female sample among the general population in Brasil. **Casa de Saúde Pública**. v. 21, p. 351-360, 2005.
- [10] JUNIOR, R.F.; CURADO, M. P.; SOUZA, A.A. A situação do câncer de mama no Goiás, no Brasil e no mundo: tendências atuais para incidência e mortalidade. **Revista Brasileira de Saúde Materno Infantil**. v.3, p. 17-24, 2003.
- [11] TEMMINK, O. H.; PRINS, H. J, VAN GELDEROP, E., PETERS, G. J. The hollow fibre assay as a model for in vivo pharmacodynamics of fluoropyrimidines in colon cancer cells. **British journal of cancer**. v.96, p.61-66, 2007
- [12] Hyperchem. Molecular Visualization and Simulation. Ontario, Hypercub, 1995
- [13] A. D. BECKE, **Phys. Rev. A** **38**, 3098, 1988.

- 
- [14] C. LEE, W. YANG, AND R. G. PARR, **Phys. Rev. B** **37**, 785 1988.
- [15] A. PETERSSON AND M. A. AL-LAHAM, **J. Chem. Phys.** **94**, 6081, 1991.
- [16] G. A. PETERSSON, A. BENNETT, T. G. TENSFELDT, M. A. AL-LAHAM, W. A. SHIRLEY, AND J. MANTZARIS, **J. Chem. Phys.** **89**, 2193, 1988.
- [17 ] M. J. FRISCH, G. W. TRUCKS, H. B. SCHLEGEL, G. E. SCUSERIA, M. A. ROBB, J. R. CHEESEMAN, J. A. MONTGOMERY, JR., T. VREVEN, K. N. KUDIN, J. C. BURANT, J. M. MILLAM, S. S. IYENGAR, J. TOMASI, V. BARONE, B. MENNUCCI, M. COSSI, G. SCALMANI, N. REGA, G. A. PETERSSON, H. NAKATSUJI, M. HADA, M. EHARA, K. TOYOTA, R. FUKUDA, J. HASEGAWA, M. ISHIDA, T. NAKAJIMA, Y. HONDA, O. KITAO, H. NAKAI, M. KLENE, X. LI, J. E. KNOX, H. P. HRATCHIAN, J. B. CROSS, C. ADAMO, J. JARAMILLO, R. GOMPERTS, R. E. STRATMANN, O. YAZYEV, A. J. AUSTIN, R. CAMMI, C. POMELLI, J. W. OCHTERSKI, P. Y. AYALA, K. MOROKUMA, G. A. VOTH, P. SALVADOR, J. J. DANNENBERG, V. G. ZAKRZEWSKI, S. DAPPRICH, A. D. DANIELS, M. C. STRAIN, O. FARKAS, D. K. MALICK, A. D. RABUCK, K. RAGHAVACHARI, J. B. FORESMAN, J. V. ORTIZ, Q. CUI, A. G. BABOUL, S. CLIFFORD, J. CIOSLOWSKI, B. B. STEFANOV, G. LIU, A. LIASHENKO, P. PISKORZ, I. KOMAROMI, R. L. MARTIN, D. J. FOX, T. KEITH, M. A. AL-LAHAM, C. Y. PENG, A. NANAYAKKARA, M. CHALLACOMBE, P. M. W. GILL, B. JOHNSON, W. CHEN, M. W. WONG, C. GONZALEZ, AND J. A. POPLE, Gaussian 03, Revision B.04, gaussian, inc., pittsburgh pa, 2003.
- [18] CAMARGO,A.J.; NAPOLITANO, H.B.; OLIVEIRA.S.S. Método de Química Quântica Computacional Aplicado a Modelagem Molecular. **Estudos**. V. 32, p.1223-1252, 2005.
- [19] SZABO, A.; OSTULAND, n. s.; **Modern quantum chemistry: introduction to advanced electronic structure theory**. NewYork, Macmillan, 446p, 1982
- [20] CAMARGO,A.J.; **Estudo Químico-Quântico ab initio e Semi Empírico deCompostos Inorgânicos e orgânicos com possíveis aplicações tecnológicas**. São Carlos. Tese ( Doutorado )- Instituto de Química de São Carlos, Universidade de São Paulo, 221p, 2001.
- [21] LEVINE, I.N. **Quantum chemistry**. 5.th. New York, Prentice Hall,709p, 1991
- [22] SLATER,J.C. **Phisycs Rev**. v.35, p.210-211,1930.
- [23] KONSTANTINOVA, E. **Novas Fases de Carbono. Abordagem Computacional**. Rio de Janeiro, Tese ( Doutorado ) – Centro Brasileiro de Pesquisas Físicas, 130p, 2005.
- [24 ] THOMAS,L.H. The calculation of atomics fields. **Proc.Cam.Phil.Soc.**,v.23, p.542-548,1927.

- 
- [25 ] FERMI, E. Astatical method for the determination of some atomic properties and the application of this method to the periodic system of element. **Z.Phys.**, v.48, p.73-79,1928.
- [26] PARR,R.G.;YANG,W.**Density – Funcional Theory of atoms and molecules, General**, Oxford University Press. New York. Claredon Press Oxford, 338p,1989.
- [27] LUNDQVIST,S.; MARCH,N.H. **Theory of the inhomogeneous electron gas, Physics of solid and liquids**.New York, Plenum, 237p, 1983.
- [28] HONENBERG,P.;KOHN,W. Inhomogeneous electron gas. **Physics Rev.**,v.136,p.B864,1964.
- [29] KOHN, W.;SHAM, L.J. **Physics Rev.**,v.136, p.A1133-A1138,1965
- [30 ]FEINBERG, M. **Anal. Chim. Acta**. V.191, p.75-85.1986
- [31] MASSART, D.L. **Chemometrics tutorial**. Amsterdam, Elsevier, 470p, 1990.
- [32] NETO,B.B.;SCARMINIO,I.S.;BRUNS,R.E. **Planejamento e otimização de experimentos**. Campinas,UNICAMP, 299p, 1995.
- [33]SHARAF, M.A.; ILLMAN, D.L.; KOLWALSKI, B. R. **Chemometrics**.New York, JonhWiley, 332p, 1986.
- [34] PRESS, S.J. **Applied multivariate statistics**. New York, Holts, Rinehart and Winston, 345p 1972.
- [35] HARRIS, R.J. **A primer of multivariate statistic**. New York, Academic Press, 332p, 1975.
- [36] SENA, M.M.; FRIGHETTO, R.T.S.; VALARINI,P.J.; TOKHESHI,H.; POPPI, R.J. Discrimination of management effects on soil parameters by using principal component analysis: a multivariate analysis case study. **Soil & Tillage Research**. v.67, p. 171-181, 2002.
- [37] COSTA, M. C. A.; **Relação entre a estrutura química de Neolignanas e a sua atividade anti-Leishmaniose, envolvendo cálculo de Solvatação**. Campinas, Tese ( Doutorado)-Universidade Estadual de Campinas, 183p, 1998.
- [38]BORGES, E. G. **Estudo da Relação estrutura química e Atividade biológica em compostos anti-úlcera**. Campinas Dissertação (Mestrado), Universidade Estadual de Campinas, 207p. 1998