



**Universidade Estadual de Goiás  
Unidade Universitária de Ciências Exatas e Tecnológicas  
Programa de Pós-Graduação Stricto Sensu em Ciências  
Moleculares**

**ESTUDO TEÓRICO DA RELAÇÃO ESTRUTURA ATIVIDADE DA  
PIPERINA E SEUS ANÁLOGOS À INIBIÇÃO DO CITOCROMO  
P450 HEPÁTICO.**

Marcos Francisco Novaes Valentino.  
Orientador : Ademir João Camargo.

Anápolis-GO

Maior- 2010.

**Universidade Estadual de Goiás**  
**Unidade Universitária de Ciências Exatas e Tecnológicas**  
**Mestrado em Ciências Moleculares**

**ESTUDO TEÓRICO DA RELAÇÃO ESTRUTURA ATIVIDADE DA  
PIPERINA E SEUS ANÁLOGOS À INIBIÇÃO DO CITOCROMO  
P450 HEPÁTICO.**

Marcos Francisco Novaes Valentino.  
Orientador : Ademir João Camargo.

Anápolis-GO

Maio-2010.

ESTUDO TEÓRICO DA RELAÇÃO ESTRUTURA ATIVIDADE DA PIPERINA E SEUS  
ANÁLOGOS À INIBIÇÃO DO CITOCROMO P450 HEPÁTICO

**MARCOS FRANCISCO NOVAES VALENTINO**

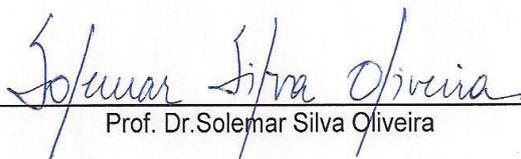
Dissertação apresentada ao corpo docente do Programa de Mestrado em Ciências Moleculares da Universidade Estadual de Goiás apresentada como parte dos requisitos necessários para obtenção do título de Mestre em Ciências Moleculares.

Aprovada por:



---

Prof. Dr. Ademir João Camargo



---

Prof. Dr. Solemar Silva Oliveira



---

Prof. Dr. Rugles César Barbosa

ANÁPOLIS, GO- BRASIL  
MAIO 2010

Só se pode alcançar um grande êxito quando  
nos mantemos fiéis a nós mesmos.

***Friedrich Nietzsche***

**A meu orientador**

**A minha co-orientadora**

**Prof. Dr. Ademir João Camargo**

**Msc. Lilian T. F M. Camargo.**

Pela oportunidade que me ofereceram de realizar esse trabalho.

Atenção e delicadeza fazem-nos perceber quanto algumas pessoas são especiais na forma de ser e como são bem-vindas as suas ações. Muito obrigado!

## Agradecimentos

Ao longo deste caminho encontrei pessoas muito especiais na qual desejo agradecer pela ajuda e pela companhia antes e durante o mestrado.

Primeiramente agradeço, a Deus.

A minha família, que me deram o suporte e incentivo necessário.

A meu orientador, que sabiamente comandou todo o projeto proposto.

A Universidade Estadual de Goiás, como instituição.

Com estima a minha colega e co-orientadora Lilian Tatiane Ferreira de Melo Camargo que não mediu esforços em ajudar-me.

A CAPES (coordenação de aperfeiçoamento de pessoal de nível superior) pelo apoio financeiro.

Aos professores, alunos e ex-alunos do grupo de Químicos Teóricos Estruturais de Anápolis.

## SUMÁRIO

CAPITULO I.....	1
I.1 INTRODUÇÃO.....	1
I.1.1 Atividades farmacológicas das espécies.....	2
I.1.2 Relações estrutura-atividade (SAR).....	5
I.1.3 Biotransformação de drogas e sistema CYP450.....	7
I.1.4 Exemplos de reações, catalisadas pelos microsomas do CYP450.....	9
I.1.5 Reações catalisadas pelo citosol e mitocôndrias do CYP450.....	10
I.2 Piperina e Análogos.....	12
CAPÍTULO II.....	15
II.1 METODOLOGIA.....	15
II.1.1 EQUIPAMENTOS E PROGRAMAS COMPUTACIONAIS.....	16
II.1.2 PROCEDIMENTO COMPUTACIONAL.....	16
II.1.3 Otimização dos compostos.....	17
II.1.4 Descritores geométricos e eletrônicos calculados usando o pacote de programas Gaussian 03, com nível de teoria B3LYP/6-31G*.....	17
II.1.5 ANÁLISE ESTATÍSTICA DOS DADOS .....	17
CAPÍTULO III.....	18-37
III.1 Teoria do Funcional da Densidade.....	18-21
III.2 Métodos estatísticos multivariados de análise.....	21-37
a) Métodos de análise de fatores.....	25
b) Método de classificação.....	26
c) Métodos de partição.....	26
d) Método hierárquico.....	26
III.3 Vetores de médias e matriz de covariância e correlação.....	27
III.4 Escalamento dos dados.....	27
a) Escalamento.....	29
Escalamento por intervalo.....	29
b) Autoescalamento.....	30
III.5 Pesos de Fisher e variância.....	30-32
III.6 Análise de componentes principais (PCA).....	31-33
III.7 Análise hierárquica de cluster (HCA).....	35-37
CAPITULO IV .....	38
IV.1 Análise Química e Conformacional.....	38
IV.2 Análise das propriedades Físico-Químicas e Químicas obtidas.....	38
IV.2.1 Volume molar (V).....	39

IV.2.2 Hidrofobicidade ou Coeficiente de partição (Log p).....	41
IV.2.3 Polarizabilidade ( $\alpha$ ).....	42
IV.2.4 Momento dipolo ( $\mu$ ).....	42
IV.2.5 Refratividade Molar (R.M).....	42
IV.2.6 Energias dos orbitais moleculares de fronteira.....	43
IV.2.7 GAP (HOMO-LUMO).....	45
IV.2.8 Eletronegatividade ( $\chi$ ).....	45
IV.2.9 Dureza ( $\eta$ ) .....	45,46
IV.2.10 Ângulos (A).....	46
IV.2.11 Ângulos diedrais (d).....	47
IV.2.12 Cargas atômicas parciais (C).....	49
IV.2.13 Ordens de ligação (O) .....	52
IV.2.14 Análise Estatística dos dados.....	56
IV.2.15 Análise da atividade dos compostos ao CYP1A1 hepático.....	56-61
Conclusão.....	62
Referências Bibliográficas.....	63,66

## Lista de Figuras

<b>Figura I.1-</b> Isômeros da piperina.....	<b>3</b>
<b>Figura I.2-</b> Gráfico atividade farmacológica versus número de grupamentos metileno.	<b>6</b>
<b>Figura I.3</b> – (-) Norgestrel, (+)Norgestrel.....	<b>6</b>
<b>Figura I.4</b> - Estrutura secundária e terciária do CYP1A2 humano.....	<b>9</b>
<b>Figura I.5</b> - Compostos analisados para identificar suas relações estrutura atividade à inibição ao CYP450 .....	<b>13</b>
<b>Figura I.6</b> - Compostos analisados para identificar suas relações estrutura atividade à inibição ao CYP450 .....	<b>14</b>
<b>FIGURA III.1-</b> O peso de Fisher para a variável $i$ e para as classes $q$ e $p$ .....	<b>30</b>
<b>FIGURA III.2</b> - Sistema de decomposição da matriz $X$ pelo método PCA.....	<b>34</b>
<b>Figura III.3-</b> Exemplo de um cálculo com HCA, cálculo de similaridade dado pela distância entre vizinhos.....	<b>34</b>
<b>Figura III.4</b> – Métodos de agrupamento hierárquico.....	<b>36</b>
<b>Figura IV.1-</b> Estrutura da piperina com enumeração da posição dos átomos.....	<b>40</b>
<b>Figura IV.2.</b> Representação de um diedro .....	<b>48</b>
<b>Figura IV.3</b> Representação gráfica dos scores obtidos usando o programa Einsight....	<b>58</b>
<b>Figura IV.4-</b> Representação Gráfica dos loadings obtidos usando o programa Einsight. ....	<b>59</b>
<b>Figura IV.5</b> – Dendograma dos 22 compostos analisados relativo às variáveis A2, Polarizabilidade e C20.....	<b>61</b>

## Lista de quadro e tabelas

Tabela I.1 –Algumas atividades farmacológicas exibidas por espécies do gênero Piper. .....	<b>2</b>
Tabela I.2 – Distribuição da piperina nas diferentes partes de <i>Piper nigrum</i> .....	<b>3</b>
Quadro 1. Exemplos de reações catalisadas por enzimas do sistema do citocromo P450. .....	<b>11</b>
Tabela II.3: Valores de IC <sub>50</sub> da piperina e seus análogos, em relação á monooxigenase AHH da atividade do CYP1A1.....	<b>16</b>
Tabela III.1 – Métodos multivariados e tipos de variáveis que podem ser analisados pelos mesmos.....	<b>23</b>
Tabela III.2 – Propósitos dos diversos métodos multivariados.....	<b>22</b>
Tabela IV.1. Valores do Volume Molar $V(\text{Å}^3)$ , Coeficiente de Partição, Momento Dipolo $\mu$ (debye), Polarizabilidade $\alpha(\text{Å}^3)$ relativo aos 22 compostos.....	<b>39</b>
Tabela IV.2 Valores de energias dos orbitais moleculares de fronteira, Gap, eletronegatividade( $\chi$ ) e dureza ( $\eta$ ), obtidas a nível de teoria B3LYP/6-31G* . Todos valores são dados em ( eV).....	<b>44</b>
Tabela IV.3 - Ângulos interatômicos (a,b,c) obtidos a nível de teoria B3LYP/6-31G*. Todos os valores estão em graus (°).....	<b>47</b>
Tabela IV.4. Ângulos diedrais (a,b,c,d) obtidas a nível de teoria B3LYP/6-31G*. Todos os valores estão em graus (°).....	<b>48</b>
Tabela IV.5 Cargas atômicas (C),derivadas do potencial eletrostático com método CHELPG, obtidas a nível de teoria B3LYP/6-31G*.....	<b>50</b>
Tabela IV.6 Cargas atômicas (C),derivadas do potencial eletrostático com método CHELPG, obtidas a nível de teoria B3LYP/6-31G*.....	<b>51</b>
Tabela IV.7 Cargas atômicas (C),derivadas do potencial eletrostático com método CHELPG, obtidas a nível de teoria B3LYP/6-31G*.....	<b>52</b>
Tabela IV.7 - Ordens de ligação obtidas a nível de teoria B3LYP/6-31G*.....	<b>53</b>
Tabela IV.8 - Ordens de ligação obtidas a nível de teoria B3LYP/6-31G*.....	<b>54</b>
Tabela IV.9 - Ordens de ligação obtidas a nível de teoria B3LYP/6-31G*.....	<b>55</b>

Tabela IV.10 – Valor do Peso Fisher para as Variáveis.....	<b>54</b>
Tabela IV.11 – Valores dos scores, PC1 e PC2.....	<b>57</b>
Tabela IV.12 - Porcentagem de variância explicada pelas componentes principais obtidos pelo programa Einsight.....	<b>58</b>
Tabela IV.3 - Loading das variáveis para as 2 componentes obtidos pelo programa Einsight.....	<b>59</b>

## RESUMO

Alguns derivados análogos da *piperina* (*Piper nigrum*) são capazes de inibir a ação de enzimas hepáticas metabolizantes de xenobióticos, mais especificamente a enzima CYP1A1 que é uma isoenzima do sistema enzimático CYP450 hepático. Neste trabalho, métodos de estatística multivariada foram usados na construção de um modelo capaz de discriminar os derivados análogos da *piperina* em ativos e inativos baseando-se na correlação da atividade biológica, descritores geométricos e eletrônicos. Dos 70 descritores calculados usando a teoria do funcional da densidade (DFT), com o funcional de troca e correlação B3LYP e o conjunto de base 6-31G\* implementados no programa *Gaussian 03*, foram pré-selecionados baseados no peso de Fischer e, finalmente, três descritores (A2, polarizabilidade e C20) foram selecionados para a construção do modelo de componentes principais. Esse modelo foi capaz de discriminar os análogos em ativos e inativos usando apenas uma componente principal, sendo que esta componente responde sozinha por 77,9% da variância total dos dados. A análise hierárquica de cluster, usando os descritores selecionados pela análise de componentes principais, mostra que os derivados análogos da *piperina* podem ser agrupados em dois grandes grupos: ativos e inativos. Os resultados deste trabalho podem ser utilizados para auxiliar o planejamento de novos derivados de *piperina* com atividade inibidora da isoforma CYP1A1 mais pronunciada.

## Abstract

Some piperine (*Piper nigrum*) analogous derivatives are able to inhibit the action of the hepatic enzymes that metabolize xenobiotics, more specifically the enzyme CYP1A1 that is an isoenzyme of the CYP450 hepatic enzymatic system. In this work, chemometric methods were applied in the search for building discriminant models between active and inactive analogues, based on the correlations among their biological activities and electronic and geometric molecular descriptors. From 70 descriptors calculated with density functional theory (DFT) with the exchange correlation functional B3LYP and the basis set 6-31G\* (Gaussian 03), 8 ones were pre-selected based on their Fisher weights, and finally five descriptors (A2, polarizability, and C20) were selected for a principal component analysis. This analysis was able to discriminate the analogues in active and inactive by using only one principal component, accounting for 77,9 % of the total variance. The cluster hierarchical analysis, using the descriptors selected by principal component analysis, shows that the *piperine* analogues can be grouped into two major groups: active and inactive. The results of this work can be used to help modeling new *piperine* analogues with more pronounced activity inhibition to the isoform CYP1A1.

## CAPÍTULO I

### CONSIDERAÇÕES BÁSICAS

#### I.1 INTRODUÇÃO

As espécies do gênero *Piper* são encontradas em diversas regiões tropicais e subtropicais do planeta e são muito usadas como fitoterápicos na medicina popular e como condimento na área gastronômica.

O nome genérico “pimenta” é derivado do termo espanhol utilizado para designar a pimenta preta (*Piper nigrum*). Em latim tardio, o termo *pigmentum* servia para designar pigmento e também pimenta (em latim tradicional *piper*). Deste termo, origina-se a raiz do inglês *pepper*. Dentre as mais de 1000 espécies descritas no gênero *Piper*, destaca-se a *Piper nigrum* (conhecida no Brasil como pimenta do reino), devido à sua importância econômica, que se deve ao seu sabor e aroma incomparáveis [1-3].

A *Piper nigrum* é comercializada sobre 2 formas principais: a pimenta do reino do tipo preta e a pimenta do reino do tipo branca. Para produzir a pimenta preta, as espigas são colhidas quando os frutos estão completamente desenvolvidos, de coloração verde-clara ou amarelada, debulhadas mecanicamente em debulhadores ou manualmente, ou então são colocadas para secar ao sol, e durante o processo de secagem as drupas secas vão se despreendendo do eixo da espiga. Para preparar a pimenta branca, as espigas são colhidas quando os frutos apresentam a coloração amarelada ou vermelha. As espigas são colocadas em sacos plásticos trançado sem serem debulhadas, e maceradas em tanques. A pimenta produzida por esse processo apresenta coloração branca e o odor característico de pimentado-reino. Após maceração e lavagem, a pimenta é seca ao sol, classificada e embalada [4].

A pimenta do reino contém cerca de 3% de óleo essencial. O princípio picante da pimenta é devido, principalmente, à presença de amidas, que ocorrem na planta em concentrações de até 5%. Dentre estes compostos a piperina (**1**), ver figura I.1, ocorre em maior proporção, principalmente nos frutos da planta. Outro ponto relevante em relação à piperina é a facilidade de sua extração, que se realiza através da preparação do extrato etanólico dos frutos secos de *P. nigrum*, num extrator de Soxhlet fornecendo o produto em sua forma pura, com altos rendimentos. Levando em consideração que o Brasil é um dos grandes produtores mundiais de pimenta do reino, e diante de algumas propriedades farmacológicas observadas na piperina, então pode ser vista como uma matéria-prima potencial para a síntese de outras moléculas de interesse [4,5].

Na medicina popular da Índia, país de onde a *Piper nigrum* é originada, uma mistura de *Piper nigrum*, *Piper longum* e *Zinziber officinallis* (gengibre) constitui o chamado trikatu (em sânscrito, “três acres”), uma formulação usada no tratamento de várias doenças [1]. Também na Índia as raízes da espécie *P. beetle*, são utilizadas como contraceptivo [2]. Na China o princípio pungente da pimenta do reino é indicado no tratamento da epilepsia [6].

São relatadas outras utilizações de espécies do gênero *Piper*, que vão desde o uso como plantas ornamentais à preparação de bebidas alucinógenas [7].

### I.1.1 Atividades farmacológicas das espécies

Como mencionado, a utilização de espécies do gênero *Piper*, com fins medicinais, é prática comum nas culturas chinesa e indiana. Na China, algumas prescrições recomendam o uso de *P. futokasura* no tratamento de arritmias cardíacas. Na Jamaica, dores estomacais são tratadas com uma infusão de *P. aduncum* *P. hispidum* (que também apresentam propriedades repelentes contra insetos). No México e no Brasil, usa-se *P. amalago* para aliviar dores estomacais e no combate a infecções.

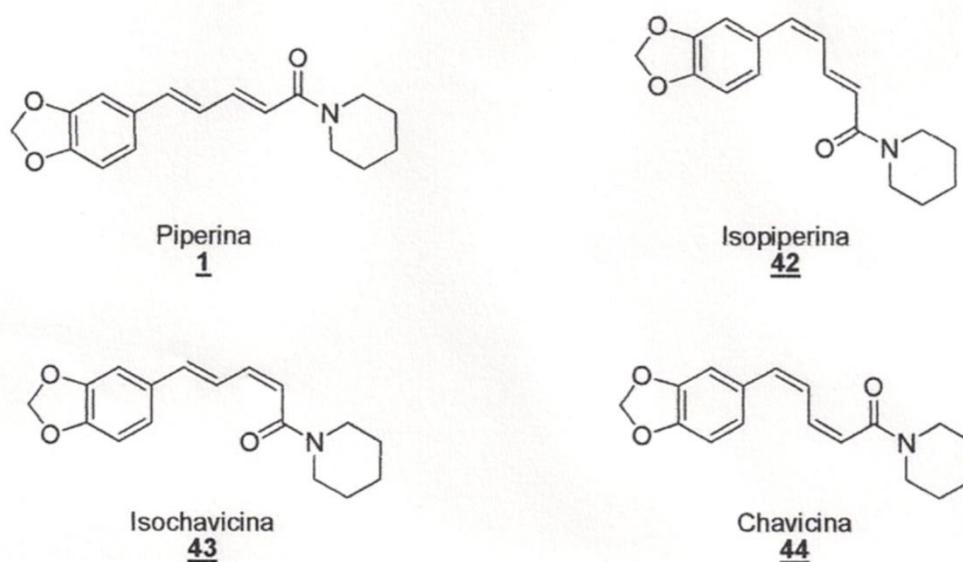
Na Tabela I. 1 estão sumariadas algumas das atividades farmacológicas mais relevantes [7].

**Tabela I.1** –Algumas atividades farmacológicas exibidas por espécies do gênero *Piper*

ATIVIDADE	ESPÉCIE
Antitumoral	<i>P. hookeri</i> , <i>P. clarkii</i> , <i>P. hancei</i> , <i>P. rebesioides</i> , <i>P. wallichii</i> , <i>P. nigrum</i>
Antileucêmica	<i>P. hookeri</i> , <i>P. futokadsura</i> , <i>P. attenuatum</i>
Analgésica	<i>P. arboricola</i> , <i>P. nigrum</i>
Antibacteriana	<i>P. aducum</i> , <i>P. mesthysticum</i>
Anti-reumática	<i>P. guineense</i> , <i>P. nigrum</i>
Inseticida	<i>P. nigrum</i> , <i>P. hispidum</i> , <i>P. longum</i>
Antiinflamatória	<i>P. amalago</i> , <i>P. nigrum</i>
Antiasmática	<i>P. futokadsura</i> , <i>P. chaba</i>
Antipirética	<i>P. chaba</i> , <i>P. nigrum</i>
Anti-hipertensivo	<i>P. beetle</i>
Antifúngica	<i>P. beetle</i>

A amida piperina (**1**) é o constituinte majoritário de *Piper nigrum* e também de várias outras espécies deste gênero. Trata-se do isômero *2E,4E* do esqueleto 5-(benzo[d][1,3]dioxol-5-il)-1-(piperidinil)-2,4-pentadien-1-ona. São também encontradas na planta as formas *2Z,4E* (isopiperina **42**), *2E,4Z* (isochavicina **43**) e *2Z,4Z* (chavicina**44**), sendo que estes 3 isômeros são minoritários [8].

A piperina, devido à sua abundância, foi a primeira substância a ser isolada de espécies de Piper por Oerstedt em 1819, sendo que a sua estrutura só foi determinada quase cem anos depois, por LANDENBURG & SCHOLTZ [9] em 1894 (EPSTEIN et al., 1993) [10].



**Figura I.1-** Isômeros da piperina.

Como pode ser visto na Tabela I.2, a piperina é encontrada principalmente nos frutos, também ocorrendo em menor proporção em outras partes da planta (SEMLER & GROSS, 1998) [11].

**Tabela I.2** – Distribuição da piperina nas diferentes partes de *Piper nigrum*.

Material		Piperina	
		(mg/g pfr) <sup>a</sup>	(mg/g ps) <sup>b</sup>
Raiz		0,057	0,32
Folhas	Verdes	0,0006	0,0049
	Secas	0,0002	0,0007
Galhos	Verdes	0,0087	0,070
	Secas	0,46	1,92
Pimenta preta <sup>c</sup>			
Frutos	Amostra comercial	—	20,7
	Pimenta branca <sup>c</sup>		
	Amostra comercial	—	2,25

a) miligrama por grama de peso fresco de pimenta

b) miligrama por grama de peso seco de pimenta

c) comercial. Valores obtidos por médias

Como a piperina é o principal constituinte, presente em diversas espécies, podemos supor que esta amida natural seja a substância responsável por grande parte das atividades descritas, apesar de podermos esperar também um efeito sinérgico nestas misturas de constituintes presentes nos estudos realizados com extratos.

Em estudos sobre a farmacologia da piperina **1**, avaliada como produto puro, foi demonstrado uma importante atividade, talvez a mais intrigante das atividades, o aumento da biodisponibilidade de outros medicamentos pela piperina **1**, este efeito é observado quando os medicamentos são coadministrados com a piperina **1**, traduzindo-se em maior concentração plasmática dos mesmos. Tal efeito foi constatado por BANO *et al.* (1987;1991) [12,13].

Com o propranolol, teofilina e a fenitoina; por GUPTA *et al* (1998) [14]. Com a nimesulida por SHOBA *et al.* (1998) [15].

Os trabalhos relatam um aumento de 100% na concentração de tais substâncias na corrente sanguínea. As razões levantadas para este efeito, segundo a maioria dos autores são:

- ✓ Alterações na permeabilidade das células epiteliais;
- ✓ Influência na glicuronidação (etapa importante na biotransformação de compostos xenobióticos);
- ✓ Atuação sobre enzimas hepáticas metabolizantes, monooxigenases hepáticas dependentes do citocromo P450.

Sabe-se que após uma droga ser administrada, ela é absorvida pelo sistema digestivo e entra no sistema porta hepático. Antes de atingir a circulação sistêmica, ela é conduzida através da veia porta hepática ao fígado. O fígado metaboliza muitos xenobióticos mais precisamente o sistema de enzimas CYP450, as vezes de tal maneira que somente uma pequena quantidade de droga ativa é lançada a partir do fígado em direção a circulação sistêmica.

Essa primeira passagem pelo fígado diminui significativamente a biodisponibilidade da droga. Talvez uma forma de minimizar o efeito indesejado da diminuição da biodisponibilidade da droga seja inibir a ação de enzimas CYP450 do fígado, e assim permitir uma maior concentração da droga que venha a ser lançada na circulação sistêmica. Seja a utilização combinada de inibidores do CYP450 bem como a piperina e análogos sintéticos. É, portanto, desejável buscar desenvolver potenciais inibidores reversíveis do CYP450 e de baixa toxicidade.

A ação terapêutica de fármacos resulta de interações destes com sistemas biológicos e é dependente de fatores relacionados com sua estrutura química e, conseqüentemente, de suas propriedades físico-químicas. Estes fatores, sejam eles de

caráter eletrônico, hidrofóbico ou estérico, influenciam na interação do fármaco com seu receptor e na distribuição nos compartimentos que compõem o sistema biológico. Assim, dois fármacos com estrutura química semelhantes, diferenciando-se apenas por um átomo ou posição que este ocupa na molécula, podem apresentar diferenças quanto as suas propriedades físico-químicas e, conseqüentemente, quanto à atividade biológica, tanto do ponto de vista quantitativo como qualitativo.

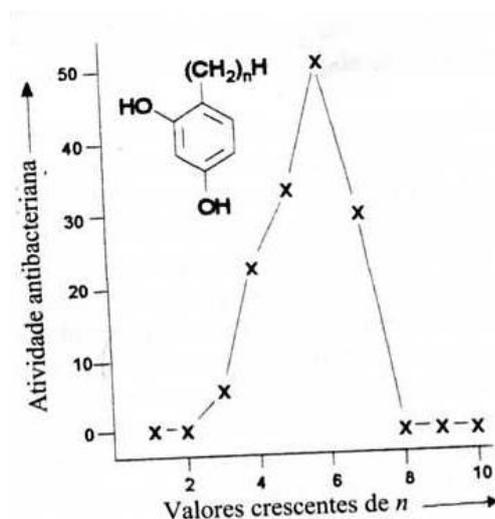
Portanto, modificar a estrutura das moléculas como trocar substituintes, pode ser fundamental para explicar a atividade biológica de um composto, possibilitando a construção de novas drogas que sejam potencialmente mais ativas.

### **I.1.2 Relações estrutura-atividade (SAR)**

Um estudo das SAR de um composto-protótipo e de seus análogos pode ser usado para determinar as partes da estrutura do protótipo que são responsáveis por sua atividade farmacológica, isto é, seu farmacóforo, e também por seus efeitos colaterais adversos. Esta informação é usada para desenvolver um novo fármaco com sua SAR otimizada, com atividade diferente de um fármaco, poucos efeitos colaterais indesejáveis e maior facilidade de administração ao paciente.

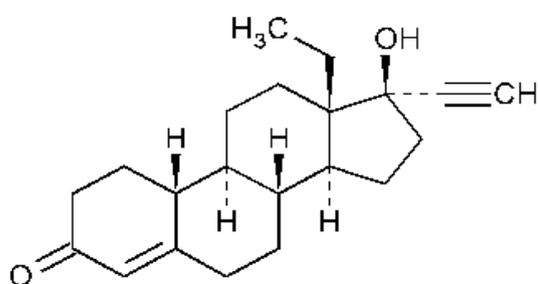
As SAR são determinadas fazendo-se pequenas alterações na estrutura do protótipo (composto líder), seguidas da avaliação do efeito que isto teve sobre a atividade biológica. Essas alterações podem ser classificadas convenientemente como: a dimensão e a conformação do esqueleto do carbono, a natureza e o grau de substituição e a estereoquímica do protótipo.

Por exemplo, ao se aumentar o número de grupamentos metileno numa cadeia ou num anel, aumentam o tamanho e a natureza lipídica (lipofilicidade) do composto. As curvas de resposta biológica associadas com este aumento podem assumir várias formas, conforme observa-ser na figura I.2 .

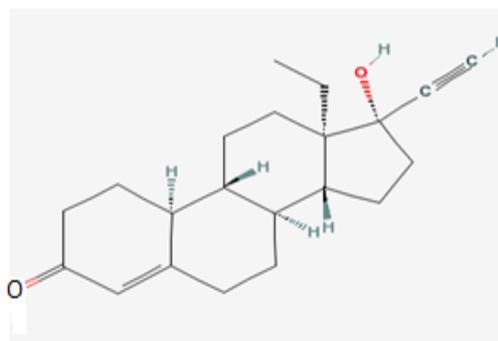


**Figura I.2** - Gráfico da atividade farmacológica versus número de grupamentos metileno.

Importante ressaltar que estereoisômeros possuem conformações diferentes, frequentemente os estereoisômeros biologicamente ativos exibirão diferenças em suas potências e/ou atividades. Tendo como exemplo: o (-) norgestrel é absorvido com uma velocidade duas vezes maior do que o (+) norgestrel através das membranas vaginal e bucal[16].



(-)Norgestrel



(+)Norgestrel

**Figura I.3** – (-) Norgestrel, (+)Norgestrel

### I.1.3 Biotransformação de drogas e sistema CYP450

As enzimas são as responsáveis pelas reações de biotransformação de drogas e encontram-se presentes em todo o organismo (sangue, rins, pulmões, pele, tecido nervoso, intestino delgado e fígado). Embora elas estejam distribuídas em todo o organismo, o fígado é o órgão que mais as concentra. Testes bioquímicos realizados com o tecido hepático, através de centrifugações sucessivas permitiram constatar a presença de enzimas nas diferentes frações denominadas solúveis (desidrogenases, esterases, amidases, transferases), mitocondrial (monoamino oxidases) e microssomal (citocromos P450).

Citocromo P450, componente primordial do sistema enzimático oxidativo, foi assim denominado por Omura & Sato em 1964 [17], porque o complexo formado com o monóxido de carbono apresentava um pico de absorção espectrofotométrica no comprimento de onda 450 nm (nanômetros). Essa enzima apresenta um núcleo pirrólico com o átomo de ferro à semelhança da hemoglobina, sendo considerada uma hemoproteína.

O sistema citocromo P450 faz parte da fração microssômica e é responsável pelas reações de oxidação de inúmeras drogas, desempenhando papel fundamental na biotransformação (Korolkovas e Burckhalter. 1998; DeLucia *et al.*, 1998) [18].

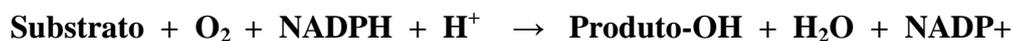
Um sistema de nomenclatura foi desenvolvido para o citocromo P450, sendo que suas isoenzimas são reunidas em subgrupos tendo em vista as semelhanças nas sequências de aminoácidos. O prefixo CYP é usado para designar o sistema citocromo P450. As isoenzimas são classificadas dentro de famílias e subfamílias. Um numeral arábico depois do prefixo CYP indica a família (por exemplo, CYP2). Depois do numeral arábico, há uma letra que representa uma subfamília (por exemplo, CYP2D). O último dígito do sistema de nomenclatura do citocromo P450 é um numeral arábico que designa a isoenzima específica (por exemplo, CYP2D6).

Uma família é constituída por enzimas que compartilham pelo menos 36% da sequência de aminoácidos. As subfamílias são formadas por enzimas com mais de 70% de similaridade na sequência de aminoácidos (Hara e Rocha, 1998; Riesenman, 1995) [19].

Até o momento são conhecidas 11 famílias do citocromo P450 humano, que incluem 30 enzimas ou citocromos diferentes. Apenas as famílias CYP1, CYP2 e CYP3 são importantes na biotransformação de drogas. Dentro dessas famílias, as isoenzimas 1A1, 1A2, 2C9, 2C19, 2D6, e 3A3/4 são reconhecidamente as mais importantes para o metabolismo de drogas.

A reação básica catalisada pelo sistema do citocromo P450 é a monooxigenação, onde o oxigênio molecular é decomposto e um de seus átomos é adicionado a um substrato (um xenobiótico ou uma substância endógena) e o outro reduzido na formação de água, à custa da oxidação do NADPH (Benet et al., 1996; Parkinson 2001) [20].

A reação básica é:



Nesse ciclo, as reações catalisadas pelo citocromo P450 envolvem a hemoproteína do citocromo P450, a NADPH citocromo P450 redutase, o NADPH e o oxigênio molecular, mas devido a rearranjos ocorridos durante as reações, os produtos do sistema do citocromo P450 não se resumem aos álcoois e fenóis e envolvem reações de hidroxilação, N-, O- e S- desalquilações, N-oxidação, sulfoxidação, N-hidroxilação, desaminação, desalogenação e dessulfuração (Benet et al., 1996; Parkinson, 2001).

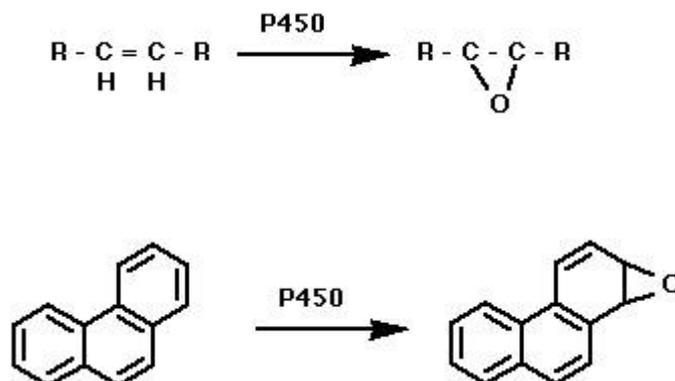
O mecanismo da reação envolve a formação do complexo substrato-citocromo oxidado ( $\text{Fe}^{+3}$ ), que é reduzido ( $\text{Fe}^{+2}$ ) pelo citocromo P450 redutase com o elétron transferido do NADPH. O complexo reduzido reage com o oxigênio molecular e novamente é reduzido pelo mesmo citocromo P450 Redutase com outro elétron doado por outro NADPH e reage também com um próton, passando de  $\text{Fe}^{+2}\text{O}_2$  para  $\text{Fe}^{+2}\text{OOH}$ . A adição de um segundo próton cliva o complexo, originando uma molécula de água e o complexo  $(\text{FeO})^{+3}$  transfere seu oxigênio para o substrato. O substrato oxidado é liberado, reconstituindo o citocromo P450. A interrupção desse ciclo (desacoplamento) após a adição do primeiro elétron origina um ânion superóxido ( $\text{O}^{2-}$ ) e, se desacoplado após a adição do segundo elétron, origina peróxido de hidrogênio (Benet et al., 1996; Parkinson, 2001) [20].

Muitas substâncias exógenas ou endógenas podem ser substratos de isoenzimas P450, ou seja, ser metabolizada por elas. Em geral, um fármaco pode ser substrato de uma única isoenzima CYP450 ou de mais de uma, seja em dado momento ou simultaneamente [21].

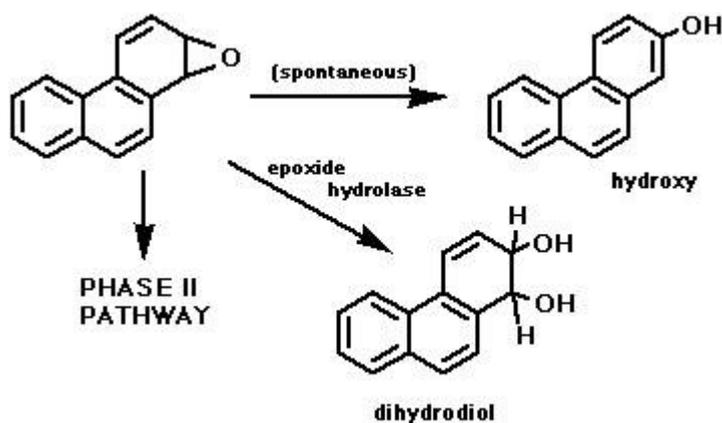
Além disso, pode ser substrato de uma isoenzima CYP450 e atuar como inibidor da mesma. Atuando como inibidor da atividade das isoenzimas, pode provocar interações potenciais com outros fármacos. Um fármaco pode ainda, inibir uma isoenzima CYP450 que não esteja relacionada com o seu processo de biotransformação. Finalmente, uma substância pode induzir um aumento na atividade de certa isoenzima, sendo ou não substrato da mesma (Hara e Rocha, 1998) [19].

### I.1.4 Exemplos de reações, catalisadas pelos microsomas do CYP450.

a) Epoxidação alifática ou aromática.

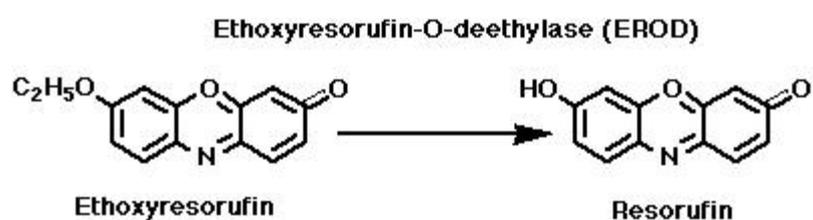


b) Hidroxilação.



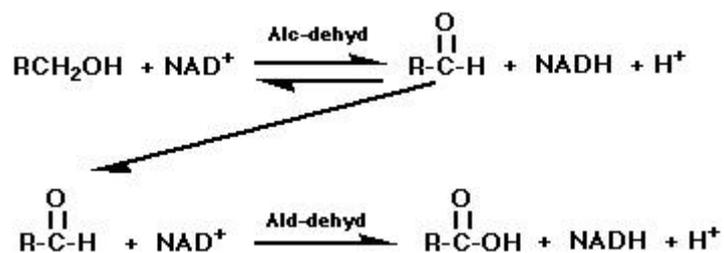
c) Dealquilação.

A dealquilação do etoxiresofurin para resofurin é uma reação específica metabolizada pela atividade da isoforma CYP1A1.



### I.1.5 Reações catalisadas pelo citosol e mitocôncrias do CYP450.

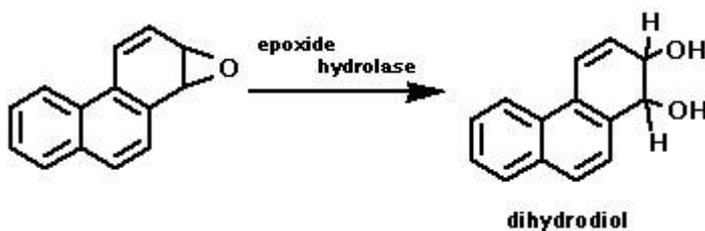
a) Reações de conversões de álcool em aldeído e aldeídos em ácidos carboxílicos.



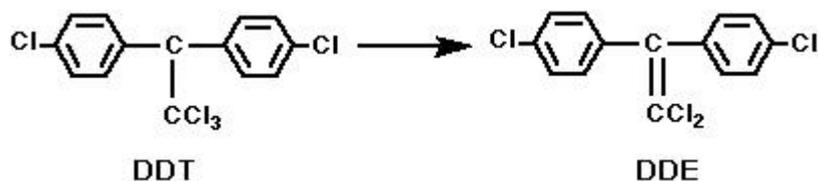
b) Hidrólises.



d) Hidratação de epóxidos.

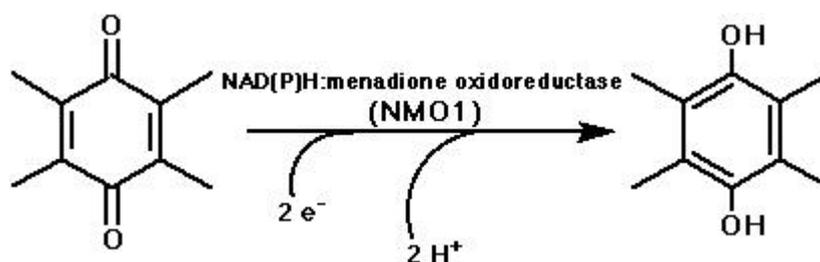


e) Declorinase (DDT a DDE)



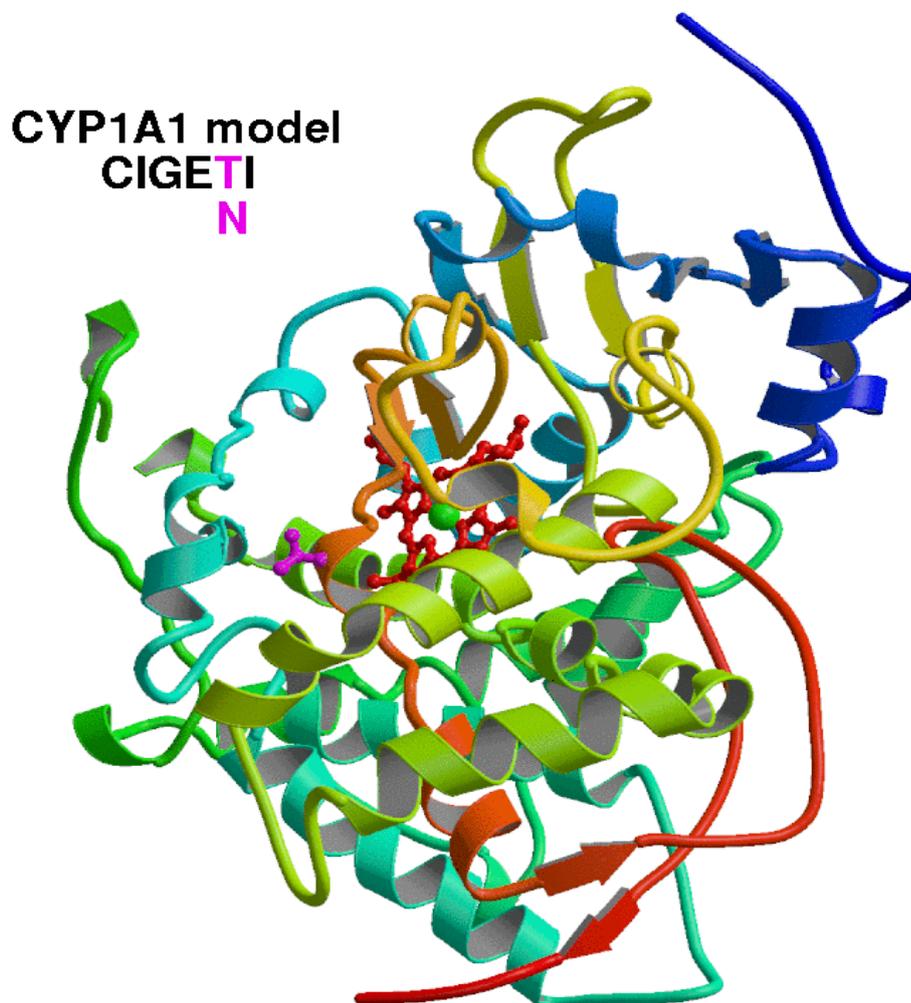
Outras reações de declorinase, também podem ser realizadas pelos microsomas do sistema CYP450.

f) Redução de quinona.



**Quadro 1.** Exemplos de reações catalisadas por enzimas do sistema do citocromo P450.

Reação	CYP450	Substrato	Referência
Hidroxilação alifática	CYP1A1, 2B1	n-Hexano	Oguri. et al., 1994
	CYP3A4, 2B6	Ciclofosfamida	Parkinson, 2001
	CYP3A4	Testosterona	Parkinson, 2001
	CYP2C9	Tolbutamida	Parkinson, 2001
Hidroxilação aromática	CYP2B1	Tolueno	Oguri. et al., 1994
S-oxigenação	CYP3A4	Omeprazol,	Parkinson, 2001
		Lanzoprazol	
N-Oxigenação	CYP2B1, 2B2	Estricicina	Oguri. et al., 1994
Epoxidação	CYP1A1, 1A2	Benzantraceno	Oguri. et al., 1994
		Tricloroetileno	Parkinson, 2001
Formação de dióis	CYP1A1, 1A2,	DMBA	Oguri. et al., 1994
	2B1, 2B2, 2C11		
N-Desalquilação	CYP1A1, 1A2, 2C3	Teofilina	Oguri. et al., 1994
N <sub>1</sub> ,N <sub>7</sub> - desmetilação N <sub>3</sub> -desmetilação	CYP2C19	Diazepan	Oguri. et al., 1994
	CYP2E1	Cafeína	Oguri. et al., 1994
	CYP1A2	Cafeína	Oguri. et al., 1994



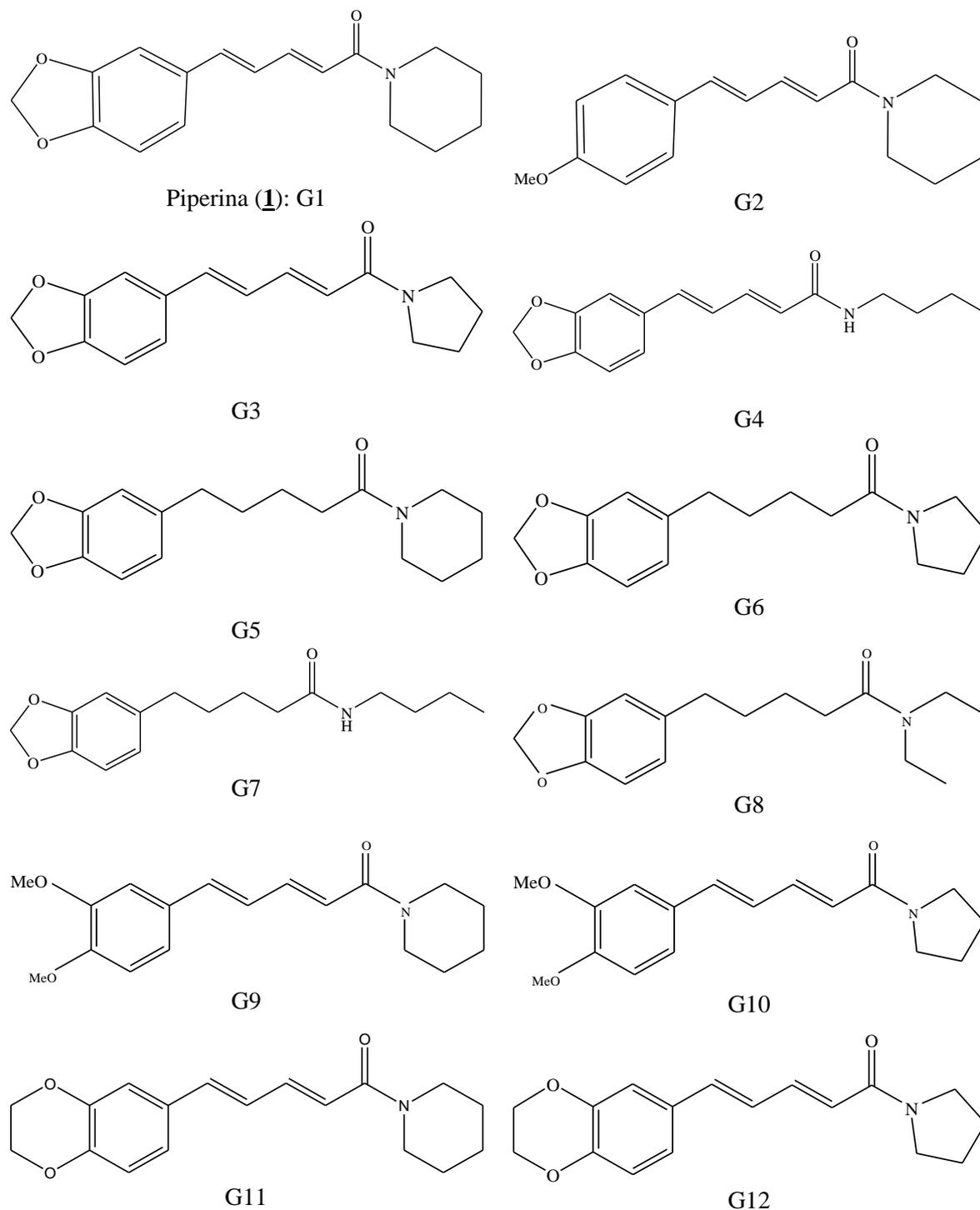
**Figura I.4** - Estrutura tridimensional da isoforma, CYP1A1 humano.

## I.2 Piperina e Análogos

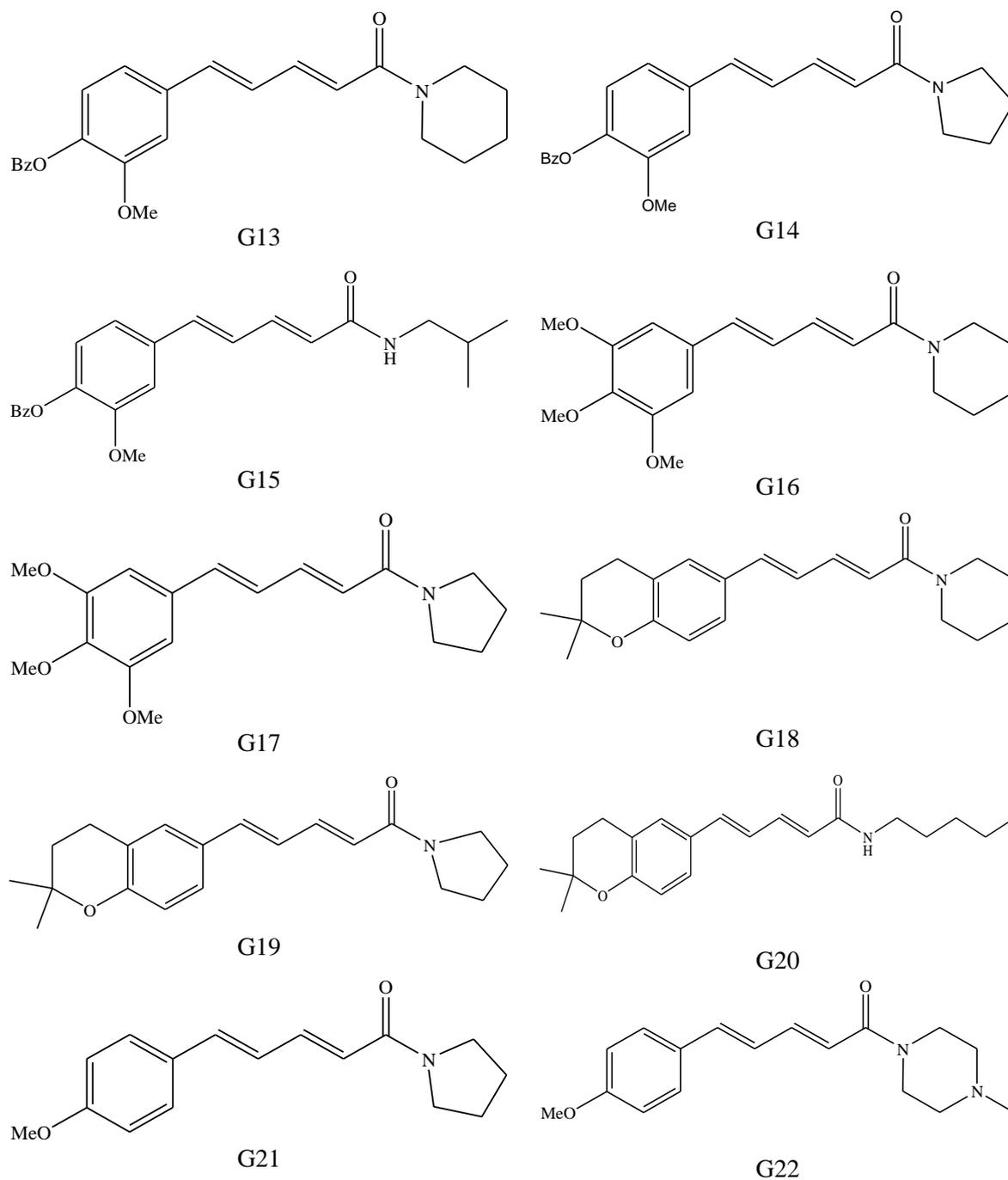
O que causou interesse para os estudos direcionados à esse composto foi, em virtude, de sua atividade inibitória à enzima CYP1A1, que realiza importante papel no metabolismo de xenobióticos. A piperina é constituída por importantes componentes um anel metileno-dioxifenil (MDP), uma fração amida e uma cadeia constituída por 5 átomos de carbono contendo duas insaturações (duplas ligações) que une o anel MDP à fração amida.

Os análogos foram construídos e posteriormente testados *in vitro*, com finalidade de encontrar qual o padrão químico que possivelmente conduz essa atividade inibitória ao CYP1A1.

Foram construídos um total de 21 análogos, em seguida temos a representação das estruturas analisadas.



**Figura I.5** - Compostos analisados para identificar suas relações estrutura atividade à inibição ao CYP450.



**Figura I.6** - Compostos analisados para identificar suas relações estrutura atividade à inibição ao CYP450 hepático.

Dentre os compostos analisados percebe-se que, G1, G3, G4, G5, G6, G7 e G8, são os compostos que possuem o anel MDP, enquanto todos análogos apresentam a cadeia com 5 átomos de carbono e a fração amida [25,26].

## CAPÍTULO II

### II.1 METODOLOGIA

Inicialmente, os pesquisadores postularam que o uso da piperina sob a forma natural de espécies de Piper em várias quantidades, pode ter sido responsável pela melhoria de biodisponibilidade de drogas em função da modulação do metabolismo hepático de drogas na fase 2[27]. Posteriormente, foi mostrado que a piperina melhorou a biodisponibilidade de fenitoína (difenil- hidantoína, agente anticonvulsivante, usado para controlar certos tipos de convulsões, no tratamento da epilepsia) , pois a co-administração da piperina com a fenitoína reduziu a produção de metabólitos mediados pela isoforma hepática CYP1A1[27,28].

Estudos atuais tem sido realizados para relacionar o papel dos componentes funcionais na piperina (**1**) à atividade inibitória de reações químicas mediadas pelo sistema CYP450 hepáticos em ratos [29].

Os testes *in vitro* dos microsomas hepáticos dos ratos relataram, que a piperina (**1**) e análogos relacionados na Figura 4, apresentaram poder inibitório à ocorrência de monooxigenase hepática catalisada pela isoforma CYP1A1, a hidroxilação AHH (*aryl hydrocarbon*, ou benzo [a]pireno hidroxilação)[30]. As influências da ação da piperina (**1**) e análogos sintéticos à ocorrência da monooxigenase (AHH) da atividade do CYP1A1 foram comparadas pelos valores de IC50 ( concentração mínima necessária de substrato para inibir 50% da atividade enzimática). Os valores de IC50 estão sumariados na tabela seguinte.

**Tabela II. 3:** Valores de IC<sub>50</sub> da piperina e seus análogos, em relação á monooxigenase AHH da atividade do CYP1A1.

Compostos	IC <sub>50</sub> em relação ao CYP1A1	Compostos	IC <sub>50</sub> em relação ao CYP1A1
1	35	12	NA
2	>100	13	100
3	80	14	35
4	105	15	>100
5	23	16	>100
6	70	17	>100
7	50	18	105
8	32	19	>100
9	NA	20	100
10	NA	21	NA
11	>100	22	>100

NA: Não possui atividade.

No presente trabalho os compostos com valores de IC<sub>50</sub> menores que 100, foram considerados ativos, enquanto os compostos com valores de IC<sub>50</sub> iguais ou maiores que 100 foram considerados inativos[30].

## II.1.2 Procedimento computacional

Para a realização deste trabalho, tinha-se a disposição um cluster de computadores com 16 nodos e os seguintes softwares: Gaussian 03[31], HyperChem Release 7.5 [32], Einsight e Microsoft Office Excel 2007[33].

Inicialmente, as estruturas dos 22 compostos sob estudo foram construídas usando o programa *HyperChem* Release 7.5 e, em seguida, os compostos foram submetidos à análise conformacional, usando o método semi-empírico PM3 ( *Parametric Method* ) implementado no programa Hyperchem. O objetivo da análise conformacional foi pesquisar as várias conformações possíveis dos derivados análogos e, dentre elas, selecionar a conformação energeticamente mais estável [32].

O confôrmero energeticamente mais estável de cada derivado foi otimizado usando a Teoria do Funcional da Densidade (DFT) com o funcional de troca e correlação híbrido B3LYP [34,35] , e o conjunto de funções de base 6-31G\* [36,37] implementados no programa Gaussian 03.

Com o nível de teoria B3LYP/6-31G\* implementada no programa Gaussian 03 foram calculados os seguintes descritores geométricos e eletônicos: a energia dos

orbitais de fronteira: HOMO (orbital mais alto ocupado), HOMO -1 (próximo orbital abaixo do mais alto orbital ocupado), LUMO ( mais baixo orbital desocupado), LUMO+1 (primeiro orbital acima do mais baixo orbital desocupado), momento de dipolo ( $\mu$ ), ângulos entre átomos (A), ângulos diedrais (d), eletronegatividade ( $\chi$ ), energia do Gap (HOMO-LUMO), dureza ( $\eta$ ), Ordem de ligação (O) e Cargas (C) derivadas do potencial eletrostático (CHELPG). Com o programa Hyperchem 7.5 e o método semiempírico PM3 foram obtidos os coeficientes de partições (LogP), volumes molares (V) e polarizabilidades ( $\alpha$ ).

A correlação estatística dos descritores eletrônicos e geométricos calculados dos derivados análogos da piperina com a atividade relativa à isoforma CYP1A1 foi feita usando os programas *Einsight*, *PLS\_toolbox* e Microsoft Office Excel 2007, seguindo as etapas abaixo:

- ✓ Construiu-se uma tabela de dados em que as linhas representam as amostras e as colunas representam as variáveis (parâmetros) calculados;
- ✓ Fez-se o cálculo dos Pesos de Fisher para ter uma visão preliminar de quais são as variáveis (descritores) 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 e análise hierárquica de *cluster* (HCA).

## CAPÍTULO III

### FUNDAMENTAÇÃO TEÓRICA

#### III.1 Teoria do Funcional da Densidade

A moderna teoria do funcional da densidade (DFT – *Density Functional Theory*) surgiu em meados dos anos 60, sendo atribuída ao cientista austríaco, ganhador do Prêmio Nobel de Química em 1998, Walter Kohn (KOHN, 1964). Esta teoria representa um avanço espetacular nos cálculos de química quântica ao juntar a sua relativa simplicidade com a excelente precisão nos resultados obtidos. O eixo central desta Teoria está no fato de que a Mecânica Quântica deixa de ser baseada em funções de onda, sendo utilizada a densidade eletrônica  $\rho(\mathbf{r})$  na descrição dos estados eletrônicos e na obtenção das propriedades dos sistemas. A densidade eletrônica (Equação 1), que pode ser obtida a partir do quadrado da função de onda integrada sobre as  $N$  coordenada eletrônicas, trás consigo a vantagem de ser obtida na dependência de apenas três coordenadas de espaço, independente do número de elétrons em questão,

$$\rho(\mathbf{r}_1) = \int |\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N.$$

(Equação 1)

Neste trabalho, as equações serão escritas em unidades atômicas e os vetores são representados em negritos. A teoria DFT começou a ter maior divulgação e impulso a partir de trabalhos pioneiros de Kohn e seu aluno, o francês Hohenberg, que através de teoremas demonstraram que a energia total ( $E_0$ ) de um sistema, bem como todas as propriedades do estado fundamental podem ser determinadas pela densidade eletrônica do estado fundamental. A energia é então um *funcional* da densidade eletrônica  $E_0 = E_0(\rho)$  [38,39].

Kohn e Shan (1965) introduziram funções para exprimir a densidade. Os orbitais obtidos a partir destas funções são chamados de orbitais de Kohn-Shan e não possuem o mesmo sentido físico dos orbitais moleculares, sendo usados apenas para a construção da densidade eletrônica total. Deste modo, têm-se as equações de Kohn-Shan (Equação 2),

$$f^{KS} = T + V_{KS}(\mathbf{r}),$$

**(Equação 2)**

onde,  $f^{KS}$  é o operador de Kohn-Shan definido pela soma do operador da energia cinética eletrônica,  $T = -\frac{1}{2}\nabla^2$ , e o operador da energia potencial descrito por um potencial médio efetivo, chamado de potencial de Kohn-Shan  $V_{KS}$ . Sendo este potencial dado pela Equação 3,

$$V_{KS}[\rho(\mathbf{r})] = V_{ext}(\mathbf{r}) + V_{Hartree}[\rho(\mathbf{r})] + V_{xc}[\rho(\mathbf{r})].$$

**(Equação 3)**

Os termos  $V_{ext}(\mathbf{r})$ ,  $V_{Hartree}[\rho(\mathbf{r})]$  e  $V_{xc}[\rho(\mathbf{r})]$  correspondem ao potencial externo (que exprime o potencial de atração núcleo-elétron), potencial de Hartree (campo médio sentido por cada elétron, devido a interação de Coulomb de cada um com todos os outros) e o potencial de troca e correlação (que contém a energia de troca e correlação,  $E_{xc}$ , relacionada a energia de troca,  $E_x$ , imposto pelo princípio da antissimetria, e a energia de correlação,  $E_c$ , necessária para descrever a correlação dinâmica dos movimentos individuais dos elétrons), respectivamente. Estes termos são descritos nas Equações 4, 5 e 6,

$$V_{ext}(\mathbf{r}) = V_{ne}(\mathbf{r}) = -\sum_A \frac{Z_A}{|\mathbf{r}-\mathbf{R}_A|},$$

**(Equação 4)**

$$V_{Hartree} = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|},$$

**(Equação 5)**

$$V_{xc} = \frac{\delta E_{xc}}{\delta \rho}.$$

**(Equação 6)**

O termo de troca e correlação representa o principal problema das equações DFT. Existe na literatura diversos funcionais aproximados que podem ser utilizados para resolvê-lo, sendo este o ponto que diferencia as várias metodologias baseadas na teoria DFT. Uma das aproximações mais simples é a *Aproximação Local da Densidade*, LDA (do inglês *Local Density approximation*) [40] na qual o termo  $E_{xc}$  representa a energia por unidade de volume do modelo de um gás de elétrons, homogêneo, com densidade constante. Assim, a energia pode ser calculada para um determinado ponto espacial conhecendo-se apenas a densidade eletrônica, sem levar em consideração o gradiente. Este modelo, logicamente, traz sérios problemas quando aplicado a sistemas reais, pois não considera de forma precisa as correlações reais entre as partículas. Deste modo, vários funcionais têm sido desenvolvidos [41,42] para contornar este inconveniente, como os funcionais de correlação de Vosko, Wilk e Nusair [43], as Aproximações de Gradiente Generalizados “GGA” [41, ] e os Métodos Híbridos.

A deficiência da aproximação LDA ocorre na medida em que não há uma uniformidade espacial na densidade eletrônica local real. Desta forma, os funcionais GGA incluem o gradiente da densidade eletrônica de forma a corrigir este problema. Basicamente, os funcionais de troca e correlação são construídos adicionando ao funcional LDA um termo de correção dependente do gradiente. Como funcionais de troca obtidos têm-se o funcional de Becke “B” ou “B88” [44]. Para os funcionais de correlação, temos o de Lee, Yang e Parr “LYP” [35].

Para um sistema sem correlação eletrônica, o determinante de Slater dos orbitais de Kohn-Shan corresponde à função de onda exata do operador Hamiltoniano. Neste sistema onde a energia de correlação,  $E_c$ , é igual a zero, tem-se apenas o termo de troca  $E_x$ . Este  $E_x$  pode ser determinado pelo método Hartree-Fock (HF), usando orbitais do tipo KS, já que o método HF inclui nas suas equações o termo de troca. Assim,  $E_{xc}$  pode ser decomposto em uma parte DFT “pura” e numa parte HF. Isto é o que chama-se de Funcional Híbrido (Equação 7),

$$E_{xc}(\text{híbrido}) = (1 - a)E_{xc}(\text{DFT}) + aE_x(\text{HF}),$$

**(Equação 7)**

onde “ $a$ ” é um parâmetro ajustável.

Deste modo, têm-se vários métodos híbridos. Um deles é o B3PW91, que inclui três parâmetros empíricos que ajustam a contribuição dos termos de troca de HF e troca de DFT (LDA e B88) e dois termos de correlação de DFT (VWN e PW91), compondo, assim, a energia total de troca e correlação  $E_{xc}$ . Outro funcional híbrido, bastante utilizado nos dias de hoje devido aos bons resultados obtidos em diversas aplicações, bem como ao relativo baixo custo computacional, quando comparado aos métodos perturbativos (MP2), é o B3LYP [35,44]. Este consiste numa modificação do método anterior, onde o funcional de correlação PW91 é substituído por LYP (Equação 8),

$$E_{xc}(\text{B3LYP}) = (1 - a)E_x(\text{LDA}) + aE_x(\text{HF}) + b\Delta E_x(\text{B88}) + (1 - c)E_c(\text{VWN3}) + cE_c(\text{LYP}),$$

**(Equação 8)**

onde  $a$ ,  $b$  e  $c$  são parâmetros ajustados experimentalmente.

### **III.2 Métodos estatísticos multivariados de análise**

Os métodos de análise multivariada de dados constituem uma poderosa ferramenta para análise de grandes e complexos conjuntos de dados. Infelizmente, dificuldades iniciais no entendimento destes métodos, que se mostram inicialmente confusos, têm afastado muitos profissionais da sua utilização.

A análise de dados pode ser entendida pela consideração de três conceitos: fatos, dados e informação. Fato é uma realidade. Assim, tudo aquilo que tem uma existência real (eventos, pessoas, pedras, etc.) é um fato. A percepção dos fatos se dá através dos dados que são as medidas, notícias, lembranças etc. Os dados constituem a base real na qual são feitos os raciocínios, discussões ou cálculos. Informação é conhecimento obtido da análise dos dados e fatos. Na verdade, conhecimento é quem justifica as alterações de consciência capaz de levar à tomada de decisão, objetivo maior de todo processo de análise de dados.

Os primórdios da análise estatística consistiam, quase que exclusivamente, na disponibilidade de numerosas observações para um número limitado de variáveis [45]. Geralmente, o objetivo maior era a validação de modelo de dependência ou um teste específico de hipótese (comparação de médias ou grupos de experimentos, análise de regressão, etc.). O maior obstáculo encontrado na época era a disponibilidade de aparatos computacionais que pudessem desenvolver cálculos rápidos capazes de analisar variações simultâneas de muitas variáveis.

A classificação dos métodos multivariados pode ser baseada em duas considerações: natureza da informação ou tipo das variáveis usadas e propósito da análise ou o modelo para a estruturação dos dados [46]. Seis tipos de variáveis podem ser definidos:

- 1) Variáveis binárias, as quais podem assumir dois valores: sim/não ou 0/1.
- 2) Classes ordenadas. Estas variáveis podem assumir  $n$  números ou valores alfanuméricos como, por exemplo: baixo, médio e alto, grau 1, grau 2 e grau 3, etc.
- 3) Classes desordenadas. Nestas variáveis inexistente a hierarquia entre as classes, ex: catalisadores diferentes, colunas cromatográficas, etc.
- 4) Variáveis inteiras ou discretas, ex: 1, 2, 3 etc.
- 5) Variáveis contínuas. São variáveis capazes de assumir qualquer valor dentro de certo intervalo definido, ex:  $I = 2,37 \mu\text{A}$  etc.
- 6) Variáveis de categoria. Estas variáveis são usadas em métodos não paramétricos com poucos exemplos da utilização por métodos multivariados.

O tipo de variável disponível em um determinado grupo de variáveis limita a escolha do método. A Tabela III.1 mostra os tipos de variáveis aos quais os modelos melhor se adequam. A Tabela III.2 relaciona alguns métodos estatísticos multivariados e seus principais propósitos. Ressalvas devem ser feitas, neste momento, devido à flexibilização dos conceitos de predição, descrição, agrupamento e classificação. Feinberg (1986) infira que a análise da componente principal (PCA) não seja adequada à predição, Massart (1990) inclui como uma meta do PCA. Isso ocorre porque o limite entre estes dois conceitos é bastante tênue. Ao se fazer uma classificação, de certo modo, faz-se também uma predição e vice-versa. Os conceitos de componentes principais e regressão sobre componentes são métodos complementares diferentes [46].

**Tabela III.1** – Métodos multivariados e tipos de variáveis que podem ser analisados pelos mesmos.

Métodos	C	CO	CD	B
---------	---	----	----	---

Análise das componentes principais (PCA)	+	-	-	-
Regressão sobre componentes principais (PCR)	+	-	-	-
Regressão múltipla (MR)	+	+	-	-
Correlações canônicas (CC)	+	-	-	-
Mínimos quadrados parciais (PLS)	+	+	-	-
Análise de fatores correspondentes (CFA)	-	+	+	+
Análise de fatores (FA)	+	+	-	-
Análise de fatores múltiplos correspondentes (MCFA)	-	+	+	+
Discriminante linear e de fatores (LDA)	+	+	-	-
SIMCA, ALLOC	+	-	-	-
Classificação por partição	+	-	-	-
Classificação hierárquica	+	-	-	-

C, contínua; CO, classe ordenada; CD, classe desordenada; B, binária. Fonte: Feinberg, 1986.

Um problema multivariado pode ser abordado de quatro maneiras diferentes. O objetivo a ser alcançado é quem vai dizer o modelo a ser escolhido. Suponha que uma matriz  $\mathbf{X}$  possa ser subdividida em duas submatrizes  $\mathbf{K}_1$  e  $\mathbf{K}_2$ . As metas a serem alcançadas são:

**1-Predição.** Este tipo de problema relaciona uma variável independente com outras variáveis ou grupo de variáveis para que possa fornecer resultados satisfatórios quando valores experimentais desconhecidos forem introduzidos nas variáveis.

**2-Tomada de decisão** (descrição). Neste caso interessa-se apenas pelo aspecto qualitativo das relações entre as variáveis ou observações. A decisão aparece de um compromisso existente entre estas relações.

**3-Agrupamento** (aprendizagem não supervisionada). Grupos de observações similares são formados e, assim, são definidos grupos homogêneos dentro de uma tabela de dados. A priori não se conhece o número de grupos que podem ser formados a partir de uma tabela de dados.

**4- Discriminação** (aprendizagem supervisionada). Novas observações são classificadas como pertencentes a uma classe pré-definida depois de cálculos probabilísticos. Faz-se necessário um aprendizado prévio de forma a se obter as funções discriminatórias usadas para a classificação.

**Tabela III.2** – Propósitos dos diversos métodos multivariados.

<b>Métodos</b>	<b>P</b>	<b>D</b>	<b>G</b>	<b>C</b>
Análise das componentes principais (PCA)	-	++	+	-
Regressão sobre componentes principais (PCR)	+++	-	-	-
Regressão múltipla (MR)	+++	+	-	-
Correlações canônicas (CC)	++	++	-	-
Mínimos quadrados parciais (PLS)	+++	++	-	-
Análise de fatores correspondentes (CFA)	-	+++	+	-
Análise de fatores (FA)	-	++	+	-
Análise de fatores múltiplos correspondentes (MCFA)	-	+++	+	-
Discriminante linear e de fatores (LDA)	-	+	-	+++
SIMCA, ALLOC	-	+	-	+++
Classificação por partição	-	-	+++	-
Classificação hierárquica	-	-	+++	-

**P**, predição; **D**, descrição para a tomada de decisão; **G**, agrupamento; **C**, Classificação. Fonte, [46].

Geralmente, é impossível atingir diferentes propósitos com a utilização de um único método estatístico. A Tabela III.2 [46] fornece uma idéia mais clara de qual método é mais adequado para um determinado tipo de problema. A dimensionalidade do problema constitui um aspecto importante a ser considerado. Não existe, a priori, um número exato de variáveis necessário para descrever adequadamente as propriedades de um sistema. As características pessoais com que cada profissional aproxima de uma problemática a ser

resolvida é que determina a dimensionalidade de um sistema. A capacidade de selecionar e medir as variáveis relevantes para uma boa descrição do sistema está diretamente relacionada à qualidade das respostas obtidas e das tomadas de decisões.

Ao excluir uma variável de determinada seleção pode ou não influir nos padrões de dados obtidos em determinada análise, depende do quanto essa variável se correlaciona com as demais. Se os padrões de dados observados se comportam de modo impreciso na realização do efeito, será necessária a busca da variável ausente. Se caso eles elucidem o evento, a variável ausente tem baixa correlação com as demais e seu efeito consiste basicamente na variância normal que o sistema apresenta [45].

Os diversos métodos de análise multivariada podem ser subdivididos em dois grandes grupos, descritos a seguir.

#### **a) Métodos de análise de fatores.**

Os métodos de análise de fatores calculam, a partir das variáveis originais, novas variáveis chamadas de fatores, que são combinações lineares das variáveis iniciais. Essas combinações lineares (ou fatores) são usadas porque a matriz de dados original contém várias correlações entre colunas e linhas, o que resulta em redundância. Os fatores são calculados de maneira tal que essas correlações são levadas em conta e são formadas de maneira a não se correlacionarem, isto é, os fatores são ortogonais entre si. Assim, torna-se mais fácil interpretar os dados do que no espaço original. Os padrões de dados gerados são mais resolvidos e relacionados ao fenômeno sob estudo do que quando descritos pelas variáveis medidas.

Para calcular os fatores, deve-se transformar a matriz de dados em uma matriz de fatores. Esta matriz de fatores leva consigo informações suficiente de maneira que os fatores obtidos auxiliem a interpretação dos padrões existentes nos dados.

Como exemplos deste grupo podem ser citados: análise de componentes principais (PCA), análise de correspondência (CFA), análise de discriminante, análise de correlações canônicas, análise de múltiplas correlações canônicas.

#### **b) Método de classificação**

Este método, também conhecido como *clustering* ou método de análise de aglomerados, pode analisar um grupo de dados por estratificação. Neste método considera-se a entrada de apenas uma linha ou coluna de cada vez, diferenciando-se, portanto, dos métodos de análise de fatores, que podem considerar várias colunas ao mesmo tempo.

O método, quando aplicado a colunas ou a linhas, promove um agrupamento dessas entradas em classes que guardam em si uma relação de similaridade o bastante para provocar o agrupamento de semelhança em classes contíguas. Dois passos são decisivos para esse procedimento: a definição de um critério de similaridade, que pode ser uma distância geométrica e a definição de um critério de agregação, o qual é uma extensão do conceito de critério de distância.

Os métodos de classificação podem ser divididos em dois subtipos, descritos a seguir.

#### **c) Métodos de partição**

A partição pode ser definida como sendo uma coleção de subconjuntos não vazios e não coincidentes cuja união é igual ao conjunto inicial. Particionar significa localizar as colunas ou linhas de uma matriz de dados num número pré-definido de classes até que todos tenham sido classificados. O algoritmo assim se inicia com a especificação, pelo usuário, de quantas classes e definições iniciais de propriedades de cada classe (centro da gravidade ou núcleos), sendo que essas propriedades podem ser mudadas no andamento do algoritmo. Estes algoritmos são convergentes (sempre que chegam a uma solução), porém podem não serem ótimos porque seus resultados dependem das escolhas iniciais [46].

#### **d) Método hierárquico**

Este método provoca uma agregação da matriz de dados. Inicia-se pelos dois elementos mais similares (menos distantes) e procedendo de acordo com os critérios de similaridade e agregação selecionados até que apenas uma classe é obtida contendo todo o conjunto original, o qual produz uma série de classes vizinhas formando uma *árvore dendograma*.

### III.3 Vetores de médias e matriz de covariância e correlação

Variáveis simples e randômicas representativas de populações e amostras são frequentemente sumarizadas pelos os seus valores médios e suas variâncias. A média de uma variável é dada por,

$$\bar{x} = \sum_{i=1}^n x_i / n,$$

**(Equação 9)**

onde  $\bar{x}$  representa o valor médio,  $x_i$  é o valor da variável para cada amostra e  $n$  é número total de amostra. A estimativa da variância ( $s^2$ ) é dada por,

$$s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1),$$

**(Equação 10)**

onde  $n$  corresponde ao número total de medidas realizadas sobre a variável  $x$ .

De um modo similar, populações e amostras multivariadas podem ser sumarizadas por *vetores de médias* e matrizes de *covariâncias*. Suponhamos que haja  $p$  variáveis  $x_1, x_2, x_3, \dots, x_p$  e os valores destas para o  $i$ -ésimo indivíduo na amostra são  $x_{i1}, x_{i2}, x_{i3}, \dots, x_{ij}$ , respectivamente. Então a média da variável  $x_j$  na amostra é,

$$\bar{x}_j = \frac{\sum_{i=1}^n x_{ij}}{n};$$

**(Equação 11)**

enquanto que a variância será dada por,

$$\frac{s_j^2 = \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n - 1}.$$

**(Equação 12)**

A covariância entre as variáveis é definida como,

$$c_{ij} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{n-1},$$

**(Equação 13)**

sendo que a covariância significa uma medida da extensão da conectividade entre as duas variáveis, isto é, a tendência que as duas variáveis têm de se desviarem de forma mais ou menos conjunta em relação às respectivas médias [47]. O valor de covariância depende da escala usada para medir  $x_j$  e  $x_k$ , o que torna difícil comparar o grau de associação estatística existente em diferentes pares de variáveis aleatórias. Para eliminar esse problema aplica-se um fator de escala, dividindo-se cada desvio individual pelo desvio padrão da variável correspondente. Obtém-se, então, uma espécie de covariância normalizada, que é chamada de *coeficiente de correlação* das duas variáveis. Devido a esta definição, o coeficiente de correlação de qualquer par de variáveis aleatórias fica obrigatoriamente restrito ao intervalo  $[-1, +1]$ . Deste modo, diferentes pares de variáveis ficam em uma mesma escala adimensional e, portanto, podem ser comparadas. A Equação 14 mostra como calcular o coeficiente de correlação de duas variáveis,

$$r_{jk} = \frac{c_{jk}}{s_j s_k},$$

**(Equação 14)**

em que  $c_{jk}$  é a covariância definida pela Equação 13 e  $s_j$ ,  $s_k$  são os desvios padrões das respectivas variáveis.

Variáveis estatisticamente independentes têm coeficiente de correlação igual a zero, mas a recíproca não é verdadeira. Diferentemente, se  $r_{ij} > 0$ , então quando uma das variáveis do par é maior que sua média, a outra também é maior que sua média. Quando é menor que sua média a outra também é menor que sua respectiva média. Se  $r_{ij} < 0$ , então quando uma variável é maior que sua média a outra é menor que sua respectiva média. Portanto, o coeficiente de correlação mede a *associação linear* existente entre duas variáveis, sendo seu valor situado entre -1 e 1. As equações 13 e 14 deixam claro que  $c_{kj} =$

$c_{jk}$ ,  $r_{kj} = r_{jk}$ ,  $c_{kk} = s_k^2$  e  $r_{kk} = 1$ . Nota-se que tanto a matriz de covariância como a matriz de correlação são simétricas em relação à diagonal principal.

### III.4 Escalamento dos dados

O objetivo do escalamento dos dados é dar a todas as variáveis a mesma importância, ou seja, o mesmo peso na análise dos dados. Suponha-se, por hipótese, que no estudo de um sistema estejam envolvidas variáveis como calor de formação dado em calorias e comprimento de ligação dado em Angstrom. Nota-se que a variância ocorrida no comprimento de ligação torna-se insignificante em relação à variância devido ao calor de formação, por estarmos trabalhando com unidades diferentes. Isto poderia levar a interpretação errônea do fenômeno em estudo. Na ausência de informações a priori, todos os dados devem ser escalados.

Os dois tipos de escalamento mais usados em análise multivariada são descritos a seguir.

#### a) Escalamento por intervalo

É talvez a forma mais familiar de escalamento. É feito colocando-se o valor mínimo de cada variável na origem e dividindo pelo intervalo dos dados. Assim, o maior valor possível para a nova variável é 1.0:

$$x'_{ik} = \frac{x_{ik} - x_k(\min)}{x_k(\max) - x_k(\min)}. \quad 0.0 \leq x'_{ik} \leq 1.0$$

(Equação 15)

Este método é bastante sensível a presença de *outliers* (pontos atípicos).

b)

**Autoescalamento**

O *autoescalamento* remove qualquer inadvertida influência que surge devido a unidades arbitrárias, mas não é tão sensível a *outliers* quanto o escalamento por intervalo. Duas versões de autoescalamento são de práticas comuns na literatura quimiométrica. Ambas centram na média, mas envolvem fatores de escalas diferentes.

- *Autoescalamento por unidade de variância*. Este método centra na média e divide pelo desvio padrão da variável que está sendo analisada,

$$x'_{ik} = \frac{x_{ik} - \bar{x}_k}{s_k},$$

**(Equação 16)**

onde,

$$s_k = \left[ \frac{1}{NP - 1} \sum_{i=1}^{NP} (x_{ik} - \bar{x}_k)^2 \right]^{1/2}.$$

**(Equação 17)**

onde  $NP$  é o número total de medidas da variável.

**III.5 Pesos de Fisher e variância**

A redução de variáveis permite eliminar os valores que não sejam relevantes para a classificação desejada. Uma maneira de se fazer essa redução, consiste em decidir se um valor é mantido ou eliminado da matriz de dados. Pode-se também aplicar um critério mais suave, calculando o *peso de variância* ou o *peso de Fisher* para cada variável.

O *peso de Fisher* [48] para uma variável  $i$  e para amostras das classes  $p$  e  $q$ , pode ser calculado como se segue:

$$W_{pq}(i) = \frac{[\bar{x}_i(p) - \bar{x}_i(q)]^2}{s_i^2(p) + s_i^2(q)},$$

**(Equação 19)**

sendo  $\bar{x}_i(p)$  média da classe dos compostos ativos e  $\bar{x}_i(q)$  média da classe dos compostos inativos e  $s_i^2(p)$  a variância dos compostos ativos e  $s_i^2(q)$  a variância dos compostos inativos.

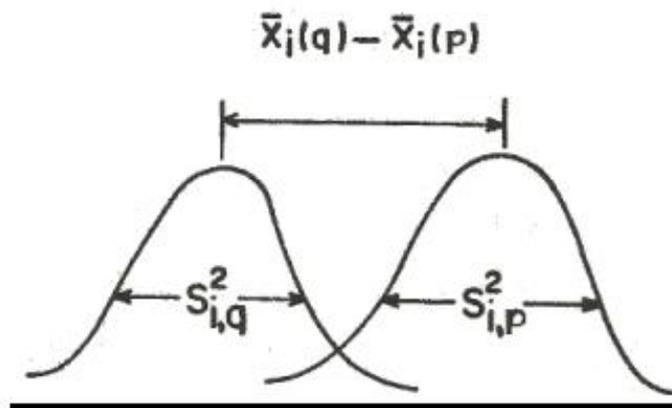
A Figura III.1 ilustra o conceito do peso de Fisher. Seu valor aumenta com o aumento da separação do valor das médias das duas classes. Quando a variância interclasse aumenta diminui o peso de Fisher. Variâncias pequenas corresponde a picos mais estreitos na Figura III.1. Quanto mais estreitos forem os picos maior resolução apresentará a variável na separação das classes ou categorias. Assim, quanto maior o peso de Fisher, maior a habilidade da variável na discriminação das classes.

O peso de variância tem a mesma função que o de Fisher. Para as classes  $p$  e  $q$ , o peso de variância é calculado dividindo a variância interclasse pela soma das variâncias das interclasses. A Equação 20 ilustra o cálculo do peso de variância para uma variável  $i$ .

$$w_{pq}(i) = 2 \frac{(1/N_{i,p}) \sum x_{i,p}^2 + (1/N_{i,q}) \sum x_{i,q}^2 - (2/N_{i,p}N_{i,q}) \sum x_{i,p} \sum x_{i,q}}{\left(\frac{1}{N_{i,p}}\right) \sum (x_{i,p} - \bar{x}_{i,p})^2 + \left(\frac{1}{N_{i,q}}\right) \sum (x_{i,q} - \bar{x}_{i,q})^2}.$$

**(Equação 20)**

$N_{i,p}$  = total de valores da variável  $i$  para a classe  $p$ .  $N_{i,q}$  = total de valores da variável  $i$  para a classe  $q$ .



**Figura III.1-** O peso de Fisher para a variável  $i$  e para as classes  $q$  e  $p$  [48].

O peso de variância é sempre maior ou igual a uma unidade de variância. Ao se trabalhar com um sistema de várias classes, além do cálculo dos pesos para cada par de classes, é desejável fazer a média geométrica de todos os pares  $NJ = \frac{1}{2} NV(NV-1)$ , que é o número de combinações possíveis. Assim, pode-se ter uma idéia global da habilidade da variável na discriminação das categorias [49],

$$W(i) = \left[ \prod_{j=1}^{NJ} W_{i(j)} \right]^{1/NJ}.$$

(Equação 21)

### III.6 Análise de componentes principais (PCA)

A técnica da componente principal foi inicialmente descrita por Karl Pearson (1901) [50]. Ele aparentemente acreditava que essa técnica poderia dar a solução correta para alguns problemas da biomedicina. A descrição prática dos métodos de cálculos só veio muito mais tarde com Hotelling em 1933. Os cálculos, na época, eram extremamente aterrorizantes e eram feitos todos à mão. Somente com advento dos computadores eletrônicos é que a técnica adquiriu uso generalizado. Na química, PCA foi introduzido por Malinowski por volta de 1960 sob o nome de análise de fatores principais e após 1970 um grande número de aplicações já tinha sido publicado [45].

A PCA de um conjunto de  $m$  variáveis originais gera  $m$  novas variáveis, as componentes principais ( $PC_1, PC_2, PC_3, \dots, PC_m$ ), que são combinações lineares das variáveis originais, que carregam em si a variância verificada nas variáveis originais [51].

$$PC_1 = b_{1,1}X_1 + b_{1,2}X_2 + b_{1,3}X_3 + \dots + b_{1,m}X_m = \mathbf{Xb}_1 \quad \text{(Equação 22)}$$

$$PC_2 = b_{2,1}X_1 + b_{2,2}X_2 + b_{2,3}X_3 + \dots + b_{2,m}X_m = \mathbf{Xb}_2 \quad \text{(Equação 23)}$$

.....

$$PC_m = b_{m,1}X_1 + b_{m,2}X_2 + b_{m,3}X_3 + \dots + b_{m,m}X_m = \mathbf{Xb}_m \quad \text{(Equação 24)}$$

As coordenadas de cada vetor ou componente são exatamente as coordenadas dos vetores latentes (autovetores) da matriz de covariância ou correlação oriunda dos dados originais. A variância associada a cada componente são as raízes latentes (autovalores) dessas matrizes. Os autovalores são ordenados de modo decrescentes e o maior autovalor é associado à primeira componente e assim sucessivamente:

$$Var(PC_1) \geq Var(PC_2) \geq \dots \geq Var(PC_m).$$

As componentes não são correlacionadas entre si. Esta é uma importante propriedade, pois, significa que cada componente mede diferentes ‘dimensões’ dos dados originais. Se os dados originais não tiverem nenhuma correlação entre si, a análise perde seu mérito, pois, para explicar toda a variância dos dados seria necessário um conjunto de componentes igual ao número de variáveis.

As componentes são sensíveis às unidades das variáveis originais. Por isso, é recomendável um pré-processamento dos dados no sentido de padronizá-las. Pode-se, também fazer a análise sobre a matriz de resíduos, que goza da propriedade de simetria das matrizes de correlação. No entanto, na interpretação dos mesmos, não se deve esquecer as transformações que deram origem a essa matriz [52].

Uma matriz de dados contendo  $N$  objetos e  $K$  variáveis na verdade representa um conjunto de  $n$  pontos num espaço  $K$ -dimensional. Se o número de variáveis for maior que 3, torna-se difícil a visualização gráfica. Quando os dados originais estão correlacionados, poucas componentes explicam praticamente toda a variância dos dados. Assim, podemos escolher de um a três componentes, desprezando as demais que têm baixo autovalor (variância). Com isso, o espaço que era  $K$ -dimensional fica reduzido a duas ou três dimensões, no qual se pode fazer uma interpretação geométrica. Pode-se dizer que a

técnica de componente principal projeta os objetos de um espaço  $K$ -dimensional para um espaço de dimensão mais baixa.

Uma propriedade importante dos autovalores  $\lambda_i$  [ $Var(PC_i) = \lambda_i$ ] é que sua soma é igual a soma dos elementos da diagonal principal ( $C_{ii}$ ) da matriz de correlação ou variância:

$$\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_m = c_{11} + c_{22} + c_{33} + \dots + c_{mm} .$$

**(Equação 25)**

Isto quer dizer que a soma dos autovalores totaliza toda a covariância ou correlação da matriz de dados. Portanto, as componentes principais levam consigo toda a variação dos dados originais [53,54].

Para uma matriz  $\mathbf{X}$  de dados contendo  $p$  variáveis existem  $p$  médias,  $p$  variâncias,  $[(p^2 - p)/2]$  covariâncias, fornecendo um conjunto total de  $[(p^2 - p)/2 + 2p]$  parâmetros [52]. Para sistemas cujas variáveis não estão correlacionadas existem apenas  $2p$  parâmetros (média e variância) a serem interpretados e estimados. No entanto, são extremamente raras matrizes experimentais não apresentarem intercorrelação, quando poderiam ser tratados pelos métodos tradicionais de regressão simples.

Contudo, a PCA é usada para transformar os dados para duas dimensões e, assim, fazer uma estimativa das similaridades dos dados [55- 57]. O PCA consiste na projeção das medidas representadas por  $m$ -dimensões de uma matriz  $\mathbf{X}$  ( $n$  amostras *versus*  $m$  variáveis) em coordenadas bidimensionais. As  $m$ -dimensões da matriz  $\mathbf{X}$  são decompostas na matriz score  $\mathbf{T}$ , na matriz loading  $\mathbf{P}'$  e na matriz erro residual  $\mathbf{E}$  (Figura 5).

$$\begin{array}{c} m \\ \boxed{\mathbf{X}} \\ n \end{array} = \begin{array}{c} A \\ \boxed{\mathbf{T}} \\ n \end{array} \begin{array}{c} m \\ \boxed{\mathbf{P}'} \\ A \end{array} + \begin{array}{c} m \\ \boxed{\mathbf{E}} \\ n \end{array}$$

**Figura III.2** - Sistema de decomposição da matriz  $\mathbf{X}$  pelo método PCA [51].

Portanto, o PCA consiste essencialmente em reescrever as coordenadas das amostras em outro sistema de eixo mais conveniente para a análise dos dados. Em outras palavras, as  $n$ -variáveis originais geram, através de suas combinações lineares,  $n$ -

componentes principais (PCs), cuja principal característica, além da ortogonalidade, é que são obtidas em ordem decrescente de máxima variância; ou seja, a componente principal 1 (PC1) detém mais informações que componente principal 2 (PC2), que por sua vez tem mais informação estatística que a componente principal 3 (PC3) e assim por diante. Normalmente, valores de autovalores  $> 1$  são considerados de interesse para a interpretação dos dados estatísticos [57,58].

### III.7 Análise hierárquica de cluster (HCA)

A principal meta do aprendizado não supervisionado é encontrar uma propriedade de uma coleção de objetos via medidas feitas sobre os objetos, sem fazer a priori nenhuma suposição sobre os dados. Esta técnica é útil para checar a validade do conjunto de dados, detectar *outliers* e resultados errôneos.

A maioria das técnicas de aprendizado não supervisionado é baseada na idéia de encontrar grupos (*clusters*) no conjunto de dados. Esses grupos são formados pela proximidade ou similaridade de suas amostras. A similaridade ( $S_{ij}$ ) existente entre as amostras de um conjunto de dados pode ser medida pela Equação 26,

$$S_{ij} = 1 - \frac{d_{ij}}{d(max)},$$

(Equação 26)

onde  $d_{ij}$  é a distância entre os pontos  $i$  e  $j$ , e  $d(max)$  é a maior distância. Desta maneira, os dois pontos mais distantes na matriz de distância têm similaridade  $S = 0.0$  e pontos idênticos têm similaridade  $S = 1.0$ . A classificação hierárquica começa assumindo que cada ponto constitui um grupo. Este, então, é ligado ao próximo mais similar. Tiram-se as médias entre eles e conecta-os aos mais próximos. Este processo se repete até que seja atingido um único grupo. O resultado deste procedimento é um diagrama chamado *dendograma*.

Uma medida geral de distância é dada pela Equação 27,

$$d_{ij} = \left[ \sum_{k=1}^{NV} (x_{ik} - x_{jk})^N \right]^{1/N},$$

**(Equação 27)**

onde  $NV$  representa o número total de variáveis,  $i$  e  $j$  representam as amostras  $i$  e  $j$ , respectivamente. Para  $N = 2$ , temos o conceito familiar de distância no espaço euclidiano. Altos valores de  $N$  darão maiores pesos às pequenas distâncias. Para algumas situações em que estão envolvidos somente valores inteiros pode ser mais apropriado usar a Equação 28 para o cálculo da distância [49],

$$d_{ij} = \sum_{k=i}^{NV} |x_{ik} - x_{jk}|.$$

**(Equação 28)**

Uma variedade de maneiras de agrupar pontos tem sido desenvolvido ultimamente. O método *single link* determina a proximidade entre dois grupos pela medida da menor distância. Contrariamente, o método do *complete link* faz uso da distância dos pontos mais distantes.

O método mais rigoroso, mas computacionalmente mais lento, é o *centroid*. Este método calcula o centro de gravidade de cada grupo e, então, a distância entre eles. Todos estes métodos estão ilustrados na Figura III.3.

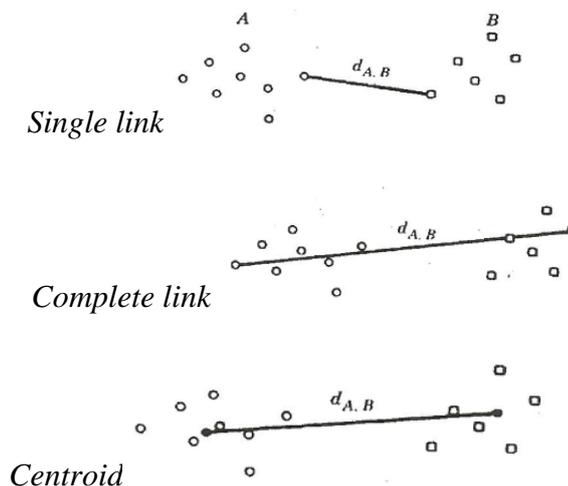
Em certas circunstâncias pode ser mais apropriado obter as similaridades diretamente das amostras. É possível calcular a distância entre estas amostras com a utilização da Equação 29,

$$d_{ij} = 1 - \frac{s_{ij}}{s_{ij} (max)}.$$

**(Equação 29)**

Onde  $d_{ij}$  distancia entre os pontos  $i$  e  $j$ ,  $s_{ij}$  é a similaridade entre os pontos  $i$  e  $j$ .

Desta maneira teremos  $0 \leq d_{ij} \leq 1.0$  e a análise se procede como em qualquer outro método de reconhecimento de padrão.



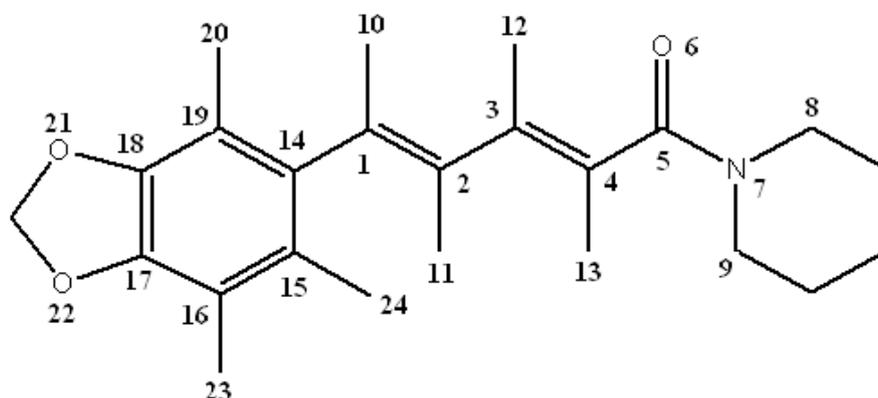
**Figura III.3** – Métodos de agrupamento hierárquico. Fonte [49].

Contudo, a técnica **HCA** é usada para reconhecer padrões (similaridades) de amostras a partir de um conjunto de dados obtidos; ou seja, de acordo com as variáveis escolhidas esta técnica agrupa as amostras entre si. A suposição básica de sua interpretação é que quanto menor for a distância entre os pontos, maior a semelhança entre as amostras. Os dendrogramas são especialmente úteis na visualização de semelhanças entre amostras ou objetos representados por pontos em espaço com dimensão maior do que três, onde a representação de gráficos convencionais não é possível. Existem muitas maneiras de procurar agrupamentos no espaço  $n$ -dimensional [55-57].

## RESULTADOS E DISCUSSÃO

### IV.1 Análise Química e Conformacional

A molécula piperina apresenta, sobretudo, carbonos e hidrogênios, com duas funções éter e uma função amida, conferindo à molécula baixa polaridade e portanto, quase insolúvel em água, de fato é solúvel em solventes apolares. Possui estrutura química quase totalmente planar, devido principalmente às insaturações dos carbonos que estão presentes entre os anéis, como segue-se representado na Fig. (IV.1).



**Figura IV.1-** Estrutura da piperina com enumeração da posição dos átomos.

Na Figura IV.1 e nos demais compostos analisados os átomos nas posições 10, 11, 12, 13, 20, 23, 24, correspondem a átomos de hidrogênio.

### IV.2 Análise das propriedades Físico-Químicas

Desde o estabelecimento dos mecanismos empregados nos estudos da relação estrutura-atividade, muitos trabalhos têm reconhecido que os efeitos eletrônicos transmitidos por grupos substituintes, não raro, influenciam a atividade biológica de drogas. O entendimento da natureza destes efeitos bem como o seu dimensionamento são de importância fundamental para o entendimento das relações entre a estrutura química e a atividade de compostos bioativos.

O estudo do aspecto tridimensional de compostos bioativos, bem como de seus receptores no sistema biológico é muitas vezes decisivo para a compreensão dos mecanismos envolvidos no estabelecimento da atividade biológica desempenhada por estes compostos. Um dos conceitos da teoria de ação de fármacos se baseia no princípio de

ajuste entre a estrutura molecular do fármaco e seu receptor e considera, entre outras propriedades, a flexibilidade presente nas biomacromoléculas envolvidas. Assim, a estereoquímica de compostos bioativos governa os princípios de reconhecimento e discriminação molecular, sendo a estereoseletividade um aspecto importante nesta abordagem.

Com intuito em alcançar quais parâmetros físico-químicos condicionam a relação estrutura-atividade da piperina e seus análogos, sob as isoforma CYP1A1 hepática, que atuam no metabolismo de fármacos, foram calculados 74 propriedades, incluindo diferentes propriedades eletrônicas e estruturais.

#### **IV.2.1 Volume molar (V)**

A ação terapêutica de fármacos resulta das interações desses com sistemas biológicos e é dependente de fatores relacionados com sua estrutura química e, conseqüentemente, de suas propriedades físico-químicas. O volume molar é a medida da dimensão de um mol de substrato. Essa variável ajuda avaliar o efeito estérico provocado pela troca do substituinte. O volume médio da piperina (**1**) e análogos, foi de 962,6414 ( $\text{\AA}^3$ ), como mostra a Tab. IV.1. O desvio padrão médio para esta variável foi de 106,8862 ( $\text{\AA}^3$ ). Os substituintes inseridos nos compostos G13, G14 e G15 contribuíram em maior proporção para aumentar o desvio padrão médio dessa variável, como podem ser vistos nos dados na tabela IV.1.

**Tabela IV.1** - Valores do Volume Molar  $V$  ( $\text{\AA}^3$ ), Coeficiente de Partição, Momento Dipolo  $\mu$ (debye), Polarizabilidade  $\alpha$  ( $\text{\AA}^3$ ) relativo aos 22 compostos.

<b>Compostos</b>	<b><math>V(\text{\AA}^3)</math></b>	<b>LogP</b>	<b><math>\mu</math>(debye)</b>	<b><math>\alpha</math></b>
1	882.11	2.87	3.33	198.95
2	892.13	2.93	4.93	199.168
3	840.72	2.45	3.69	189.154
4	879.18	3.09	3.62	188.304
5	901.93	2.88	3.37	155.55
6	865.69	2.5	3.73	149.73
7	907.48	3.12	3.81	148.779
8	893.65	2.86	3.55	150.136
9	967.31	2.68	3.54	261.137
10	910.46	2.26	2.62	196.299
11	924.68	2.5	3.98	221.668
12	887.28	2.8	4.01	202.026
13	1183.87	4.44	3.97	268.432
14	1145.81	4.04	3.78	264.168
15	1184.69	4.68	3.78	264.105
16	1011.3	2.28	3.76	201.824
17	984.33	1.88	3.55	200.008
18	1046.42	3.8	5.53	233.828
19	1005.61	3.38	3.49	226.013
20	1094.29	4.42	4.19	233.234
21	846.78	2.52	5.90	177.328
22	922.39	1.99	5.75	201.039
Média	962.6414	3.016818	3.99	205.9492
Variância	11424.67	0.636489	0.67	1412.836
Desvio Padrão	106.8862	0.797803	0.82	37.58771

#### IV.2.2 Hidrofobicidade ou Coeficiente de partição (Log p)

A resposta biológica produzida por compostos bioativos está diretamente relacionada com os mecanismos de absorção, distribuição e atividade intrínseca. Os processos de absorção e distribuição destes compostos são regulados basicamente pela sua hidrofobicidade, uma vez que é necessário que se dissolvam e atravessem as membranas biológicas que formam os tecidos e sistemas multi-compartimentados dos seres vivos, até que alcancem seus respectivos sítios de ação.

Os clássicos trabalhos de Overton e de Meyer, juntamente com as inúmeras correlações encontradas entre atividade biológica e a hidrofobicidade de compostos bioativos, demonstraram claramente que esta propriedade desempenha papel fundamental na expressão da bioatividade destes compostos, embora esta seja apenas uma das várias propriedades físico-químicas que, atuando conjuntamente, determinam a atividade biológica destes compostos.

O valor do coeficiente de partição (*Log P*) foi calculado pelo programa Hyperchem 7.5. É um parâmetro frequentemente usado, como medida quantitativa da lipofilicidade de compostos bioativos. Em outras palavras, o coeficiente de partição de uma determinada espécie química é definido como sendo a razão entre concentrações que se estabelecem nas condições de equilíbrio de uma substância química, quando dissolvida em sistema constituído por uma fase orgânica e uma fase aquosa, e está associado à mudança de energia livre provocada pela substância sobre o equilíbrio termodinâmico do sistema. Essa relação pode ser expressa pela relação

$$P = [orgânica]/[aquosa]$$

Em que **P** é o coeficiente de partição do composto analisado; [orgânica] é a concentração do composto na fase orgânica nas condições de equilíbrio; [aquosa] é a concentração da substância na fase aquosa nas condições de equilíbrio. O sistema n-octanol/tampão fosfato pH 7,4 é considerado o sistema preferencial para a determinação do coeficiente de partição. Valores de LogP maior que 1, indica que o composto é mais solúvel em meio orgânico, valores de Log P menor que 1 indica que o composto é mais solúvel em meio aquoso.

O composto G15 é o que apresenta maior valor de LogP (4.68), portanto dentre todos os compostos analisados é o mais solúvel em meio orgânico. Entretanto a molécula 17 apresenta o menor coeficiente de partição (1.88), sendo então o menos solúvel em meio

orgânico, uma análise minuciosa da Tabela IV.1, nos leva a concluir que os compostos são solúveis em meio orgânico.

### IV.2.3 Polarizabilidade ( $\alpha$ )

Polarizabilidade é uma propriedade anisotrópica e está relacionada com a estabilização intramolecular de íons orgânico em solução e pode ser definida como a medida da facilidade relativa da distorção da nuvem eletrônica de um sistema quando exposto a um campo elétrico externo. Quando uma molécula é sujeita a um campo elétrico, a sua geometria pode ser distorcida ou alinhar o seu momento dipolar permanente. O momento dipolar induzido é proporcional à força do campo elétrico. A polarizabilidade, ou o volume de polarizabilidade está correlacionada com a separação HOMO-LUMO e com o raio molecular: moléculas grandes, com muitos elétrons, tendem a ter pequena separação HOMO-LUMO e polarizabilidades grandes, enquanto percebe-se o contrário em moléculas pequenas. Considerando os dados fornecidos na Tabela IV.1, verifica-se que o composto G13 é o que apresenta maior polarizabilidade e este está entre os compostos com maior volume molar.

### IV.2.4 Momento dipolo ( $\mu$ )

O momento dipolo que foi calculado após a otimização das moléculas utilizando o programa computacional Gaussian03, é definido como um vetor que mede a magnitude da dispersão eletrônica ( densidade de carga ), que desloca em uma ligação formada por átomos que possuem diferentes eletronegatividades. A direção do momento dipolo em um composto é estabelecido pelas eletronegatividades dos átomos que o constitui. Sendo que é numericamente dado pelo vetor resultante dos momentos dipolo de cada ligação pertencentes à molécula.

O momento dipolo é um importante descritor, pois, está relacionado com a polaridade do composto; quanto maior for o momento dipolo, mais polar o é. Analisando a Tabela IV.1, a molécula 21 é a que possui maior momento dipolo e portanto é a mais polar. Em contrapartida a molécula 10 é a mais apolar, um vez que possui menor momento dipolo.

### IV.2.5 Refratividade Molar (RM)

O parâmetro de refração molar,  $RM$ , exprime uma propriedade físico-química de caráter constitutiva e aditiva, sendo, portanto, extremamente dependente da estrutura do composto químico. Com base nesta propriedade, os valores de  $RM$  relativos a subestruturas moleculares podem ser determinados experimentalmente pelo emprego da relação de Lorentz-Lorenz, descrita na seguinte equação:

$$RM = \frac{n^2-1}{n^2+2} \cdot \frac{M_W}{d}$$

Aqui,  $n$  é o índice de refração do composto a 20°C;  $d$  é a sua densidade a 20°C e  $M_W$  é a massa molar. Dunn, em 1977, sugeriu a existência de uma relação entre o valor de  $RM$  e a conformação tridimensional da estrutura molecular de fármacos. Assim a refração molar estaria diretamente relacionada ao ajuste da molécula do fármaco ao seu sítio receptor.

#### IV.2.6 Energias dos orbitais moleculares de fronteira

HOMO e LUMO são siglas para os orbitais moleculares de maior energia ocupado e o de menor energia desocupado respectivamente. As energias dos orbitais de fronteira são comumente usadas 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 (orbital de maior energia ocupado) e do LUMO (orbital de menor energia desocupado) corresponde à energia necessária para remover e adicionar elétrons respectivamente. Assim a energia do HOMO é proporcional a energia de ionização e a energia do LUMO a afinidade eletrônica. Além das energias dos orbitais de fronteira foi calculado também a energia do orbital HOMO-1 (um orbital molecular abaixo do mais alto orbital molecular ocupado) e a energia do orbital LUMO+1 (um orbital acima do mais baixo orbital molecular desocupado).

Analisando a Tabela IV.2 pode-se observar que as energias do HOMO dos 22 compostos, possuem pequena variância, isto pode ser percebido por meio do baixo valor do desvio-padrão encontrado.

**Tabela IV.2** - Valores de energias dos orbitais moleculares de fronteira (eV), Gap, eletronegatividade( $\chi$ ) e dureza ( $\eta$ ), obtidas a nível de teoria B3LYP/6-31G\* . Todos valores são dados em ( eV).

Compostos	HOMO	HOMO -1	LUMO	LUMO +1	Gap	$\chi$	$\eta$
1	-0.198	-0.217	-0.060	-0.138	-0.138	0.13	0.07
2	-0.196	-0.223	-0.057	0.016	-0.139	0.13	0.07
3	-0.195	-0.223	-0.060	-0.002	-0.135	0.13	0.07
4	-0.196	-0.234	-0.061	-0.003	-0.135	0.13	0.07
5	-0.200	-0.227	0.003	0.012	-0.203	0.10	0.10
6	-0.198	-0.227	0.006	0.019	-0.204	0.10	0.10
7	-0.198	-0.237	0.005	0.018	-0.203	0.10	0.10
8	-0.202	-0.227	0.001	0.011	-0.203	0.10	0.10
9	-0.191	-0.222	-0.056	0.006	-0.135	0.12	0.07
10	-0.189	-0.221	-0.054	0.011	-0.135	0.12	0.07
11	-0.195	-0.222	-0.057	0.000	-0.138	0.13	0.07
12	-0.195	-0.222	-0.057	0.001	-0.138	0.13	0.07
13	-0.193	-0.224	-0.055	-0.003	-0.052	0.12	0.07
14	-0.192	-0.222	-0.057	-0.003	-0.054	0.12	0.07
15	-0.194	-0.230	-0.058	-0.003	-0.055	0.13	0.07
16	-0.200	-0.218	-0.061	0.003	-0.139	0.13	0.07
17	-0.198	-0.211	-0.057	-0.008	-0.141	0.13	0.07
18	-0.191	-0.221	-0.055	-0.003	-0.136	0.12	0.07
19	-0.191	-0.220	-0.054	-0.002	-0.137	0.12	0.07
20	-0.193	-0.236	-0.004	0.017	-0.189	0.10	0.09
21	-0.197	-0.225	-0.056	-0.005	-0.051	0.13	0.07
22	-0.194	-0.200	-0.056	-0.006	-0.050	0.13	0.07
Média	-0.195273	-0.223136	-0.043636	-0.002818	-0.132273	0.119	0.076
Variância	0.00001154	0.000064	0.000654	0.000986	0.002648	0	0.0002
D. Padrão	0.003	0.008	0.026	0.031	0.051	0.01	0.0136

### IV.2.7 GAP (HOMO-LUMO)

A diferença de energia do orbital HOMO e do LUMO conhecida por Gap é um importante índice de estabilidade das moléculas, que representa a reatividade do composto. Quanto mais reativo menor o Gap, e quanto maior o intervalo de Gap menor a reatividade da molécula, pois as moléculas são mais estáveis. Baixos valores de Gap indicam instabilidade na molécula e, portanto, alta reatividade.

Como mostra a Tabela IV.2, os compostos que apresentaram maiores valores de Gap foram o 22 e 21, os quais não apresentam atividade ao CYP1A1. Um fator que justifica esta inatividade pode estar relacionado com sua baixa reatividade, quando comparada com os demais compostos do grupo.

### IV.2.8 Eletronegatividade ( $\chi$ )

A eletronegatividade é um parâmetro que foi introduzido por Pauling, o qual é definido como a tendência de um átomo em uma molécula de atrair elétrons, ou seja, é a atração que um átomo exerce sobre os elétrons, em uma ligação química. Uma de suas expressões foi dada por Mulliken, em 1934 como sendo a seguinte relação entre o potencial de ionização (PI) e afinidade eletrônica (AE)

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

Uma vez que os compostos apresentam um pequeno desvio padrão, veja Tabela IV.2, pode-se afirmar que há uma grande similaridade entre os valores de eletronegatividade dos 22 compostos estudadas.

### IV.2.9 Dureza ( $\eta$ )

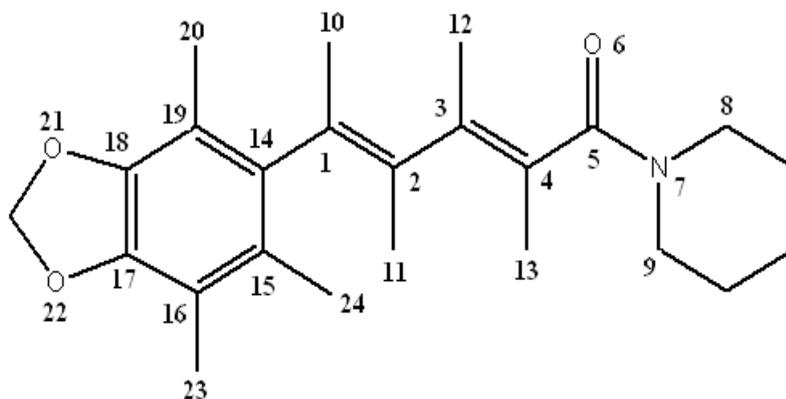
A dureza é um índice que reflete a reatividade química de uma espécie: quanto maior o valor de  $n$ , mais dura é a espécie química e menos reativa ela é, e vice-versa. Segundo a teoria do orbital molecular, a dureza corresponde à facilidade com que os elétrons são rearranjados após uma perturbação. Uma molécula é dita dura se possuir certa resistência ao rearranjo de elétrons após uma perturbação e é mole se tiver facilidade de rearranjo molecular após uma perturbação. A dureza é definida como

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

Os compostos 6, 5 e 7 possuem os maiores valores de dureza e são ativos ao CYP1A1, ou seja, são mais resistentes a um rearranjo eletrônico e, possivelmente, atingem o sítio receptor sem antes sofrer alterações conformacionais. Os valores calculados são mostrados na Tabela IV.2.

#### IV.2.10 Ângulos (A)

Determinar a conformação tridimensional das moléculas é de fundamental importância para o estudo da relação entre atividade farmacológica e estrutura molecular. O ângulo entre três átomos é um parâmetro geométrico que interfere diretamente na atividade da droga com seu específico sítio receptor.



**Figura IV.1-** Estrutura da piperina com enumeração da posição dos átomos.

**Tabela IV.3** - Ângulos interatômicos (a,b,c) obtidos a nível de teoria B3LYP/6-31G\*. Todos os valores estão em graus (°).

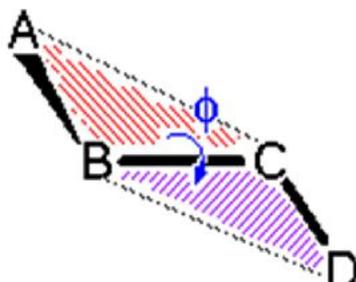
Ângulos entre os átomos a, b, c							
Compostos	A1 (17,18,21)	A2 (18,17,22)	A3 (1,14,15)	A4 (1,2,3)	A5 (4,5,7)	A6 (5,7,8)	A7 (5,7,9)
1	110.46	110.76	123.05	122.88	118.59	117.92	127.87
2	118.60	115.88	123.75	122.79	118.78	118.56	127.72
3	109.59	109.86	123.05	122.73	117.12	119.65	127.81
4	109.59	109.86	123.07	122.73	114.79	121.88	118.86
5	109.63	109.69	120.21	114.22	118.80	126.61	118.99
6	109.58	109.65	121.17	113.84	116.39	127.03	120.32
7	109.58	109.64	121.16	113.86	117.11	128.73	112.60
8	109.76	109.85	120.22	114.07	118.95	117.52	125.11
9	115.07	115.53	119.08	124.94	118.74	127.67	118.59
10	115.61	115.66	118.82	127.04	117.25	119.75	127.62
11	121.74	122.09	123.49	122.82	118.48	118.49	127.10
12	121.75	122.09	123.53	122.74	117.13	119.65	127.81
13	120.27	125.45	122.92	123.25	119.64	125.15	118.14
14	120.27	125.44	122.95	122.83	117.11	119.68	127.79
15	120.26	125.46	122.94	122.82	114.44	123.22	118.50
16	115.23	120.66	123.19	132.15	118.18	118.21	126.61
17	155.55	121.28	118.74	125.50	117.19	119.79	127.65
18	119.79	123.64	119.11	124.87	118.74	127.66	118.61
19	119.95	123.74	123.64	122.78	117.11	119.67	127.79
20	119.92	123.77	123.68	122.67	114.55	123.27	118.42
21	121.12	124.93	123.64	123.20	120.06	127.74	117.74
22	121.12	124.94	123.57	123.26	119.62	125.69	118.43

Nota-se na Tabela IV.3 que os valores dos ângulos interatômicos estão compreendidos em uma faixa de 109.59° à 155.55°. Observou-se que no ângulo (A1), houve o maior desvio padrão enquanto o ângulo (A5) apresentou o menor desvio. A variação relativamente pequena dos ângulos justifica uma parcial rigidez na estrutura dos compostos.

#### IV.2.11 Ângulos diedrais (d)

O ângulo diedral é definido como sendo o ângulo formado por dois planos concorrentes com pode ser visto na Figura IV.2. Neste trabalho, foram calculados cinco diedros, representando a região do composto onde houve um maior número de

substituição. Os dados da Tabela IV.4 mostram que a estrutura dos compostos é muito rígida e houve pequena variação nestes diedros.



**Figura IV.2.** Representação esquemática do ângulo diedral .

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

Compostos	d1	d2	d3	d4	d5
	(21,18,19,14)	(22,17,16,15)	(1,2,3,4)	(6,5,7,8)	(6,5,7,9)
1	180.00	-179.93	179.81	1.60	174.97
2	-179.99	179.96	-179.92	0.02	-177.74
3	178.84	-179.05	179.88	-2.37	177.88
4	178.96	-179.04	179.92	-1.54	-177.05
5	178.21	-178.41	-173.87	1.89	-179.76
6	179.01	-178.92	-178.72	1.82	178.90
7	178.31	-178.14	177.73	-2.99	-172.94
8	179.89	-179.76	-179.97	1.39	-176.12
9	179.96	179.99	-179.95	0.60	-176.39
10	179.10	-179.96	179.65	2.98	178.37
11	179.73	179.49	-179.73	0.89	-168.90
12	-179.79	-179.48	179.94	-2.41	178.25
13	-179.72	179.71	179.52	157.27	-4.27
14	-179.79	179.69	179.94	2.58	-178.46
15	179.74	-179.80	-179.65	0.50	179.09
16	-178.93	-176.40	178.49	-0.41	-165.93
17	178.24	-175.05	-169.75	-1.84	176.34
18	-179.51	178.85	179.87	0.78	-176.46
19	-179.48	179.28	-179.33	2.53	-178.00
20	-179.29	178.97	-179.20	-1.02	179.07
21	-179.98	179.97	-179.32	-161.44	-2.02
22	179.94	-179.93	179.73	-4.47	159.61

#### IV.2.12 Cargas atômicas parciais (C)

As cargas pontuais podem ser avaliadas de diversas formas: como pela análise de população eletrônica dos orbitais (como cargas de Mulliken) ou métodos baseados no ajuste do potencial eletrostático circundante à molécula. Neste trabalho, as cargas obtidas foram derivadas do potencial eletrostático utilizando o procedimento CHELPG [59].

A análise populacional proposta por *Mulliken* divide com equidade as densidades eletrônicas entre dois átomos sem levar em consideração as eletronegatividades dos diferentes átomos. Nesse sentido, as cargas atômicas parciais derivadas da análise populacional de Mulliken não são adequadas para explicar a interação entre o fármaco e o receptor. Já as cargas derivadas do potencial eletrostático são ajustadas de tal modo a reproduzir o potencial eletrostático molecular. Como a interação inicial entre substrato e o respectivo sítio ativo é inicialmente por interação eletrostática as cargas obtidas do potencial eletrostático constituem um importante descritor no estudo da relação estrutura atividade.

No procedimento CHELPG, as cargas atômicas são ajustadas para reproduzir o potencial eletrostático a um número de pontos ao redor da molécula. Como um primeiro passo no processo de ajuste, o potencial é calculado em um certo número de pontos em uma malha (*gridpoints*), espaçados e distribuídos regularmente em um cubo, as dimensões desse cubo são escolhidas de tal forma que a molécula fique localizada em seu centro. A Tabela IV.5, IV.6 e IV.7, mostra os valores calculados para as cargas atômicas parciais.

**Tabela IV.5** - Cargas atômicas (C), derivadas do potencial eletrostático com método CHELPG, obtidas a nível de teoria B3LYP/6-31G\*.

<b>Compostos</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>	<b>C5</b>	<b>C6</b>	<b>C7</b>	<b>C8</b>	<b>C9</b>
1	-0.17	-0.11	0.05	-0.38	0.59	-0.54	-0.33	0.05	0.12
2	-0.19	-0.10	0.02	-0.35	0.57	-0.54	-0.30	0.06	0.09
3	-0.18	-0.11	0.02	-0.32	0.54	-0.54	-0.25	0.05	0.11
4	-0.15	-0.14	0.07	-0.38	0.68	-0.53	-0.65	0.21	0.34
5	0.00	0.05	0.21	-0.25	0.56	-0.53	-0.32	0.02	0.09
6	0.00	0.11	0.12	-0.20	0.55	-0.54	-0.29	0.05	0.10
7	0.01	0.14	0.02	-0.18	0.67	-0.56	-0.66	0.18	0.34
8	-0.09	0.10	0.24	-0.38	0.64	-0.54	-0.40	0.15	0.19
9	-0.16	-0.12	0.03	-0.35	0.57	-0.54	-0.36	0.05	0.07
10	-0.14	-0.15	0.07	-0.34	0.55	-0.54	-0.55	0.05	0.12
11	-0.19	-0.11	0.05	-0.37	0.61	-0.54	-0.34	0.07	0.10
12	-0.19	-0.11	0.05	-0.36	0.57	-0.53	-0.34	0.14	0.15
13	-0.20	-0.09	0.02	-0.31	0.63	-0.54	-0.38	0.21	0.08
14	-0.20	-0.08	0.02	-0.33	0.54	-0.54	-0.25	0.07	0.10
15	-0.17	-0.10	0.03	-0.37	0.69	-0.55	-0.64	0.24	0.31
16	-0.12	-0.15	0.08	-0.33	0.56	-0.53	-0.31	0.06	0.19
17	-0.09	-0.20	0.13	-0.38	0.56	-0.55	-0.26	0.12	0.05
18	-0.19	-0.10	0.03	-0.37	0.58	-0.54	-0.31	0.11	0.08
19	-0.18	-0.12	0.03	-0.31	0.54	-0.54	-0.25	0.04	0.11
20	-0.21	0.11	0.06	-0.38	0.66	-0.52	-0.61	0.31	0.32
21	-0.21	-0.04	-0.03	-0.28	0.60	-0.56	-0.26	0.00	0.10
22	-0.19	-0.11	0.04	-0.37	0.55	-0.53	-0.19	-0.05	-0.03

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

<b>Compostos</b>	<b>C10</b>	<b>C11</b>	<b>C12</b>	<b>C13</b>	<b>C14</b>	<b>C15</b>	<b>C16</b>	<b>C17</b>	<b>C18</b>	<b>C19</b>
1	0.12	0.10	0.09	0.17	0.17	-0.20	-0.20	0.22	0.28	-0.34
2	0.11	0.09	0.10	0.15	0.15	-0.15	-0.23	0.37	-0.23	-0.13
3	0.12	0.10	0.10	0.13	0.20	-0.21	-0.22	0.25	0.28	-0.36
4	0.11	0.12	0.08	0.13	0.16	-0.19	-0.22	0.24	0.28	-0.34
5	0.00	-0.01	-0.05	0.04	0.06	-0.20	-0.22	0.22	-0.39	-0.31
6	0.00	-0.04	-0.03	0.03	0.05	-0.17	-0.23	0.25	0.25	-0.31
7	0.00	-0.05	0.01	0.05	0.07	-0.18	-0.22	0.22	0.28	-0.33
8	0.03	-0.02	-0.08	0.08	0.10	-0.21	-0.24	-0.40	0.25	-0.29
9	0.11	0.09	0.10	0.15	0.18	-0.24	-0.18	0.22	0.26	-0.32
10	0.09	0.10	0.08	0.13	0.17	-0.26	-0.16	0.20	0.25	-0.28
11	0.13	0.09	0.09	0.15	0.20	-0.21	-0.19	0.22	0.24	-0.32
12	0.12	0.09	0.09	0.16	0.21	-0.21	-0.20	0.21	0.25	-0.33
13	0.12	0.10	0.06	0.14	0.17	-0.31	0.23	0.27	-0.22	-0.21
14	0.12	0.09	0.10	0.12	0.19	-0.31	0.22	0.26	-0.21	-0.21
15	0.12	0.10	0.09	0.13	0.16	-0.30	0.24	0.24	-0.20	-0.22
16	0.08	0.10	0.08	0.15	0.25	-0.40	0.32	0.00	0.33	-0.44
17	0.09	0.11	0.06	0.14	0.21	0.21	-0.32	0.24	0.32	-0.39
18	0.12	0.09	0.10	0.16	0.18	0.18	-0.18	0.30	-0.07	-0.22
19	0.12	0.11	0.10	0.12	0.17	-0.18	-0.19	0.30	-0.07	-0.22
20	0.13	0.10	0.07	0.13	0.20	-0.19	-0.18	0.29	-0.05	-0.25
21	0.11	0.08	0.05	0.13	0.20	-0.19	-0.19	0.35	-0.22	-0.16
22	0.12	0.09	0.09	0.16	0.14	-0.13	-0.23	0.38	-0.24	-0.16

**Tabela IV.7** - Cargas atômicas (C), derivadas do potencial eletrostático com método CHELPG, obtidas a nível de teoria B3LYP/6-31G\*.

Compostos	C20	C21	C22	C23	C24
1	0.18	-0.40	-0.39	0.15	0.13
2	0.11	0.13	-0.37	0.13	0.12
3	0.18	-0.39	0.25	0.16	0.13
4	0.18	-0.39	-0.39	0.16	0.12
5	0.16	-0.40	-0.39	0.15	0.13
6	0.17	-0.39	-0.39	0.16	0.12
7	0.18	-0.41	-0.40	0.16	0.12
8	0.16	-0.40	-0.40	0.16	0.13
9	0.15	-0.34	-0.33	0.13	0.13
10	0.13	-0.33	-0.33	0.12	0.14
11	0.15	-0.33	-0.35	0.14	0.13
12	0.15	-0.34	-0.34	0.14	0.13
13	0.12	-0.03	-0.41	-0.29	0.15
14	0.13	0.13	-0.40	-0.28	0.15
15	0.13	0.13	-0.39	-0.29	0.15
16	0.17	-0.27	-0.34	-0.28	0.16
17	0.16	-0.34	-0.29	-0.24	0.13
18	0.13	0.06	0.47	0.13	0.13
19	0.12	0.07	-0.48	0.13	0.12
20	0.12	0.08	-0.47	0.13	0.12
21	0.13	0.12	-0.34	0.11	0.12
22	0.12	0.12	-0.43	0.13	0.11

#### IV.2.13 Ordens de ligação (O)

A ordem de ligação descreve a densidade eletrônica entre dois átomos ligados. Quanto maior a ordem de ligação, maior a densidade eletrônica entre os átomos, e mais estável será a ligação química, o que leva a uma diminuição da distância entre dois átomos. A Tabela IV.7, IV.8 e IV.9, descreve as ordens de ligação entre os átomos que constituem a estrutura base dos 22 compostos analisados.

**Tabela IV.7-** Ordens de ligação obtidas a nível de teoria B3LYP/6-31G\*.

<b>Compostos</b>	<b><math>O_{1.2}</math></b>	<b><math>O_{2.3}</math></b>	<b><math>O_{3.4}</math></b>	<b><math>O_{4.5}</math></b>	<b><math>O_{5.6}</math></b>	<b><math>O_{5.7}</math></b>	<b><math>O_{7.8}</math></b>	<b><math>O_{7.9}</math></b>	<b><math>O_{1.10}</math></b>
1	1.71	1.18	1.73	1.05	1.58	1.16	0.94	0.95	0.91
2	1.71	1.18	1.73	1.05	1.60	1.15	0.94	0.96	0.91
3	1.71	1.18	1.73	1.05	1.60	1.16	0.94	0.95	0.91
4	1.71	1.18	1.73	1.07	1.62	1.17	0.96	0.80	0.91
5	1.00	1.01	1.00	0.98	1.66	1.17	0.94	0.95	0.92
6	1.00	1.01	1.01	0.98	1.66	1.17	0.95	0.94	0.91
7	1.00	1.01	1.01	0.97	1.67	1.18	0.79	0.97	0.91
8	1.00	1.01	1.00	0.98	1.65	1.18	0.94	0.95	0.91
9	1.70	1.18	1.73	1.05	1.60	1.15	0.94	0.96	0.91
10	1.72	1.17	1.74	1.05	1.59	1.16	0.94	0.95	0.91
11	1.71	1.18	1.73	1.05	1.60	1.15	0.94	0.96	0.91
12	1.71	1.18	1.73	1.05	1.60	1.15	0.94	0.96	0.91
13	1.71	1.17	1.75	1.63	1.63	1.15	0.95	0.94	0.91
14	1.70	1.18	1.73	1.05	1.60	1.15	0.94	0.95	0.91
15	1.01	1.18	1.73	1.05	1.62	1.16	0.97	0.80	0.91
16	1.78	1.12	1.78	1.04	1.61	1.15	0.94	0.95	0.91
17	1.73	1.16	1.74	1.05	1.59	1.16	0.95	0.94	0.91
18	1.70	1.18	1.73	1.05	1.60	1.14	0.96	0.94	0.91
19	1.70	1.18	1.73	1.05	1.60	1.15	0.94	0.95	0.91
20	1.70	1.18	1.73	1.05	1.60	1.17	0.96	0.90	0.91
21	1.60	1.25	1.63	1.12	1.61	1.09	0.96	0.95	0.91
22	1.72	1.17	1.75	1.03	1.62	1.15	0.95	0.94	0.91

**Tabela IV.8** - Ordens de ligação obtidas a nível de teoria B3LYP/6-31G\*.

<b>Compostos</b>	<b><math>O_{2.11}</math></b>	<b><math>O_{3.12}</math></b>	<b><math>O_{4.13}</math></b>	<b><math>O_{1.14}</math></b>	<b><math>O_{14.15}</math></b>	<b><math>O_{15.16}</math></b>	<b><math>O_{16.17}</math></b>	<b><math>O_{17.18}</math></b>	<b><math>O_{18.19}</math></b>
1	0,91	0,89	0,91	1,12	1,37	1,42	1,40	1,30	1,44
2	0,91	0,89	0,92	1,13	1,35	1,46	1,36	1,34	1,49
3	0,91	0,89	0,92	1,13	1,37	1,43	1,40	1,29	1,44
4	0,91	0,89	0,92	1,13	1,37	1,43	1,40	1,29	1,44
5	0,91	0,90	0,91	1,01	1,43	1,40	1,42	1,30	1,42
6	0,92	0,90	0,90	1,01	1,45	1,38	1,43	1,29	1,43
7	0,92	0,91	0,90	1,01	1,45	1,38	1,43	1,29	1,43
8	0,91	0,91	0,90	1,01	1,43	1,39	1,42	1,30	1,42
9	0,91	0,89	0,92	1,13	1,39	1,40	1,41	1,24	1,46
10	0,91	0,89	0,92	1,11	1,39	1,41	1,41	1,25	1,44
11	0,91	0,89	0,91	1,12	1,33	1,48	1,36	1,32	1,40
12	0,91	0,89	0,92	1,12	1,33	1,48	1,36	1,32	1,40
13	0,91	0,91	0,90	1,13	1,30	1,46	1,24	1,41	1,40
14	0,91	0,89	0,92	1,13	1,30	1,46	1,24	1,41	1,40
15	0,91	0,89	0,92	1,13	1,29	1,46	1,24	1,41	1,40
16	0,91	0,89	0,91	1,09	1,36	1,39	1,32	1,33	1,37
17	0,91	0,89	0,91	1,10	1,32	1,41	1,29	1,35	1,37
18	0,91	0,89	0,92	1,13	1,33	1,48	1,35	1,32	1,44
19	0,91	0,89	0,92	1,13	1,32	1,50	1,33	1,34	1,42
20	0,91	0,89	0,92	1,13	1,32	1,50	1,33	1,34	1,42
21	0,90	0,90	0,90	1,19	1,29	0,95	1,34	1,36	1,46
22	0,91	0,91	0,90	1,13	1,32	1,51	1,33	1,37	1,44

**Tabela IV.9** - Ordens de ligação obtidas a nível de teoria B3LYP/6-31G\*.

<b>Compostos</b>	<b><i>O</i><sub>19.20</sub></b>	<b><i>O</i><sub>18.21</sub></b>	<b><i>O</i><sub>17.22</sub></b>	<b><i>O</i><sub>16.23</sub></b>	<b><i>O</i><sub>15.24</sub></b>
1	0,91	0,98	0,99	0,91	0,91
2	0,91	0,91	1,02	0,91	0,92
3	0,91	0,98	0,99	0,91	0,91
4	0,91	0,98	0,99	0,91	0,91
5	0,91	0,98	0,98	0,91	0,91
6	0,91	0,98	0,98	0,91	0,91
7	0,91	1,29	0,98	0,91	0,91
8	0,91	0,98	0,98	0,91	0,91
9	0,91	1,01	1,02	0,91	0,91
10	0,90	1,02	1,02	0,91	0,92
11	0,91	0,99	1,00	0,91	0,91
12	0,91	0,99	1,00	0,91	0,91
13	0,92	0,91	1,02	1,01	0,91
14	0,91	0,91	0,00	1,01	0,91
15	0,91	0,91	1,02	1,01	0,91
16	0,91	1,01	0,97	1,01	0,90
17	0,91	1,00	0,97	1,01	0,90
18	0,92	1,01	1,02	0,91	0,91
19	0,91	1,01	1,02	0,91	0,92
20	0,91	1,01	1,02	0,91	0,92
21	0,92	0,91	1,02	0,91	0,91
22	0,92	0,91	1,02	0,91	0,91

#### IV.2.14 Análise Estatística dos dados

Os resultados dos cálculos teóricos foram relacionados com a atividade apresentada pelos compostos em ativos e inativos ao CYP1A1 hepático. Nesta análise foram empregadas as seguintes técnicas: pesos de Fisher, análise de componentes principais (PCA) e análise hierárquica de cluster. Os programas utilizados foram: Einsight e Microsoft Office Excel 2007.

#### IV.2.15 Análise da atividade dos compostos ao CYP1A1 hepático

Inicialmente, calculou-se o peso de Fisher para cada descritor. Desse modo, pode-se ter uma idéia prévia de quais descritores são mais relevantes na separação dos compostos em ativos e inativos. Os descritores com os maiores pesos de Fisher são os mais importantes na discriminação dos compostos.

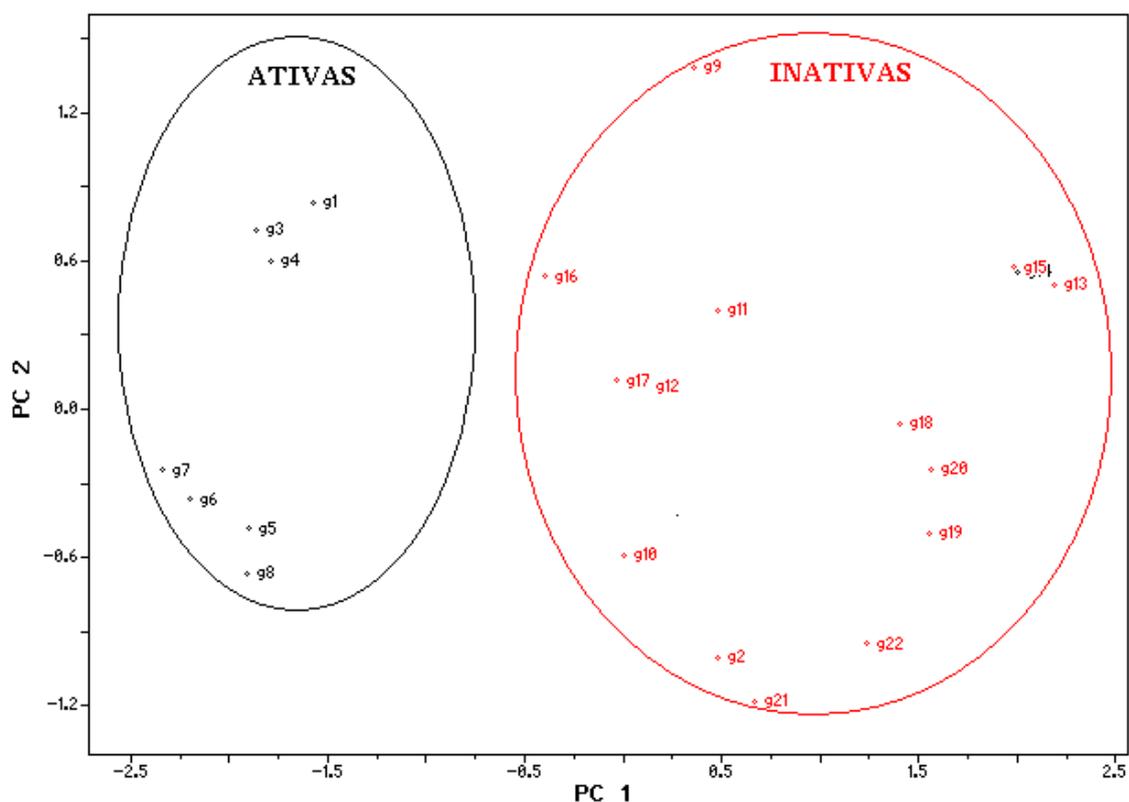
**Tabela IV.10** – Valor do Peso Fisher para as Variáveis.

Variáveis	Peso Fisher	Variáveis	Peso Fisher	Variáveis	Peso Fisher
Área	0.26	<i>d</i> 2	0.57	C18	0.09
V	0.31	<i>d</i> 3	0.08	C19	0.29
LOG p	0.004	<i>d</i> 4	0.00	C20	1.67
R.M	0.50	<i>d</i> 5	0.04	C21	0.78
POLARI.	0.75	C1	0.68	C22	0.00
$\mu$	0.41	C2	0.56	C23	0.15
HOMO	0.65	C3	0.34	C24	0.02
HOMO -1	0.32	C4	0.31	<i>O</i> <sub>1.2</sub>	0.58
LUMO	0.52	C5	0.02	<i>O</i> <sub>2.3</sub>	0.88
LUMO +1	0.17	C6	0.08	<i>O</i> <sub>3.4</sub>	0.99
GAP	0.60	C7	0.02	<i>O</i> <sub>4.5</sub>	0.25
$\chi$	0.42	C8	0.00	<i>O</i> <sub>5.6</sub>	0.55
$\eta$	0.60	C9	0.14	<i>O</i> <sub>5.7</sub>	0.88
A1	1.12	C10	0.73	<i>O</i> <sub>7.8</sub>	0.21
A2	2.55	C11	0.80	<i>O</i> <sub>7.9</sub>	0.00
A3	0.01	C12	0.60	<i>O</i> <sub>1.10</sub>	0.00
A4	1.38	C13	1.01	<i>O</i> <sub>2.11</sub>	0.73
A5	0.05	C14	0.90	<i>O</i> <sub>3.12</sub>	0.29
A6	0.00	C15	0.05	<i>O</i> <sub>4.13</sub>	0.04
A7	0.00	C16	0.07	<i>O</i> <sub>1.14</sub>	0.90
<i>d</i> 1	0.56	C17	0.19	<i>O</i> <sub>14.15</sub>	1.32
<i>O</i> <sub>15.16</sub>	0.00	<i>O</i> <sub>16.17</sub>	0.64	<i>O</i> <sub>17.18</sub>	0.21
<i>O</i> <sub>18.19</sub>	0.07	<i>O</i> <sub>19.20</sub>	0.58	<i>O</i> <sub>18.21</sub>	0.11
<i>O</i> <sub>17.22</sub>	0.21	<i>O</i> <sub>16.23</sub>	0.10	<i>O</i> <sub>15.24</sub>	0.06

De acordo com a Tabela IV.9, os descritores que apresentaram maiores pesos Fisher foram: A2 (2,55), C20 (1,67), A4 (1,38) e A1 (1,12). No entanto, a análise de componentes principais usando estes descritores não foi capaz de discriminar os compostos em ativos e inativos. Decidiu-se incluir na análise de componentes principais os descritores com menor peso de Fisher. Após realizar diversas combinações de descritores, o resultado mais satisfatório encontrado foi a combinação dos seguintes descritores: A2 (ângulo 2), polari. (polarizabilidade) e C20 (carga correspondente ao átomo número 20). A Figura IV.3 mostra uma representação gráfica dos scores relativa às componentes PC1 e PC2.

**Tabela IV.11 – Valores dos scores, PC1 e PC2.**

COMPOSTOS	PC1	PC2
1	-1.5772	0.83705
2	0.48242	-1.0070
3	-1.8680	0.72170
4	-1.7859	0.59747
5	-1.9038	-0.4902
6	-2.2042	-0.3689
7	-2.3414	-0.2498
8	-1.9143	-0.6631
9	0.35645	1.38337
10	0.00025	-0.5886
11	0.47908	0.39973
12	0.11359	0.08979
13	2.19015	0.49920
14	1.99937	0.55732
15	1.98554	0.57202
16	-0.4007	0.53133
17	-0.0368	0.11106
18	1.40614	-0.0564
19	1.55221	-0.4997
20	1.56064	-0.2409
21	0.67014	-1.1843
22	1.23676	-0.9508



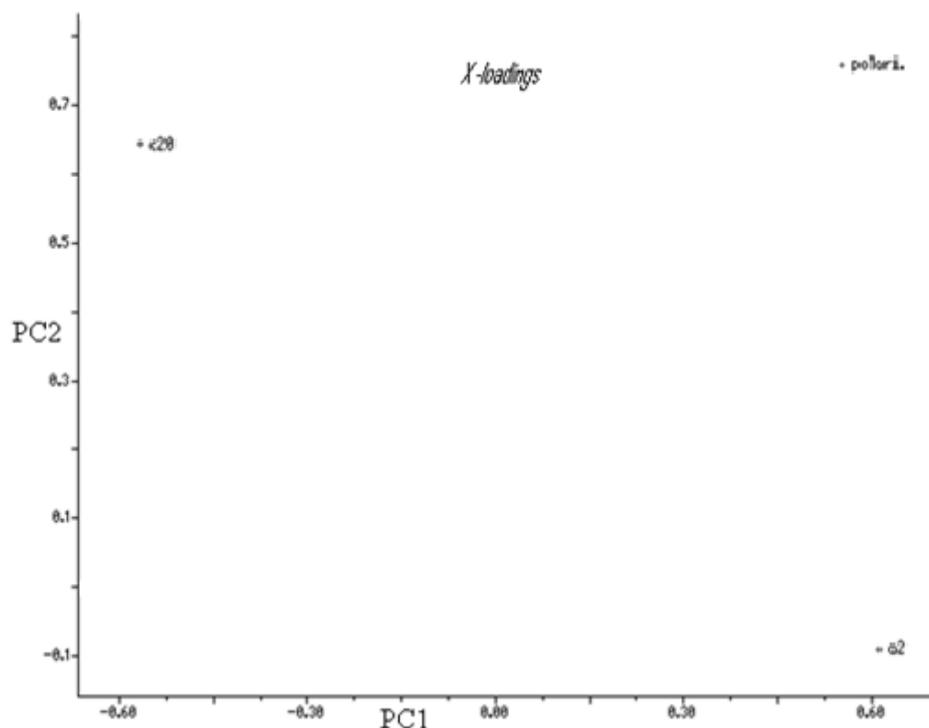
**Figura IV.3** Representação gráfica dos scores obtidos usando o programa Einsight.

Observa-se que a primeira componente é a responsável pela separação dos compostos em ativos e inativos, sendo que a primeira componente (PC1) explica 77,93% da variância total dos dados e a segunda componente (PC2) explica 15,39% da variância total. As duas componentes explicam juntas 93,33% como pode ser visto Tabela IV.12.

**Tabela IV.12** - Porcentagem de variância explicada pelas componentes principais obtidos pelo programa Einsight.

Componentes	Porcentagem individual	Porcentagem Acumulada
PC1	77,9345	77,9345
PC2	15,3946	93,3291

A figura IV.4 representa graficamente o peso das variáveis que separaram as moléculas em dois grupos sendo o ângulo 2 (a2), polarizabilidade (polari.) e a carga correspondente ao átomo 20 (C20), conforme mencionado anteriormente.



**Figura IV.4-** Representação Gráfica dos loadings obtidos usando o programa Einsight.

**Tabela IV.13** - Loading das variáveis para as 2 componentes obtidos pelo programa Einsight.

Variáveis	PC1	PC2
<b>Polari.</b>	0.55066	0.76109
<b>A2</b>	0.6109	-0.0875
<b>C20</b>	-0.5688	0.64271

A Figura IV.4 mostra que a variável A2, não possui peso significativo na segunda componente, no entanto é a variável de maior peso na PC1. A equação da primeira componente, responsável pela separação, pode ser escrita em função dos descritores autoescalados como

$$PC1 = 0,55(Polari.) + 0,61(a2) - 0,56(C20).$$

Analisando a Figura IV.3, juntamente com a equação da PC1, verifica-se que para os compostos serem ativos à isoforma CYP1A1, eles deverão possuir valores da PC1 os mais negativos possíveis; deslocando-se os compostos ativos mais para esquerda do eixo da PC1 e os inativos à direita do eixo da PC1. De acordo com os dados das tabelas IV.1, IV.3 e IV.7, nota-se que as variáveis Polarizabilidade ( $\alpha$ ), ângulo 2 (A2) e carga correspondente ao átomo 20 (C20), possuem valores numéricos positivos. Então,

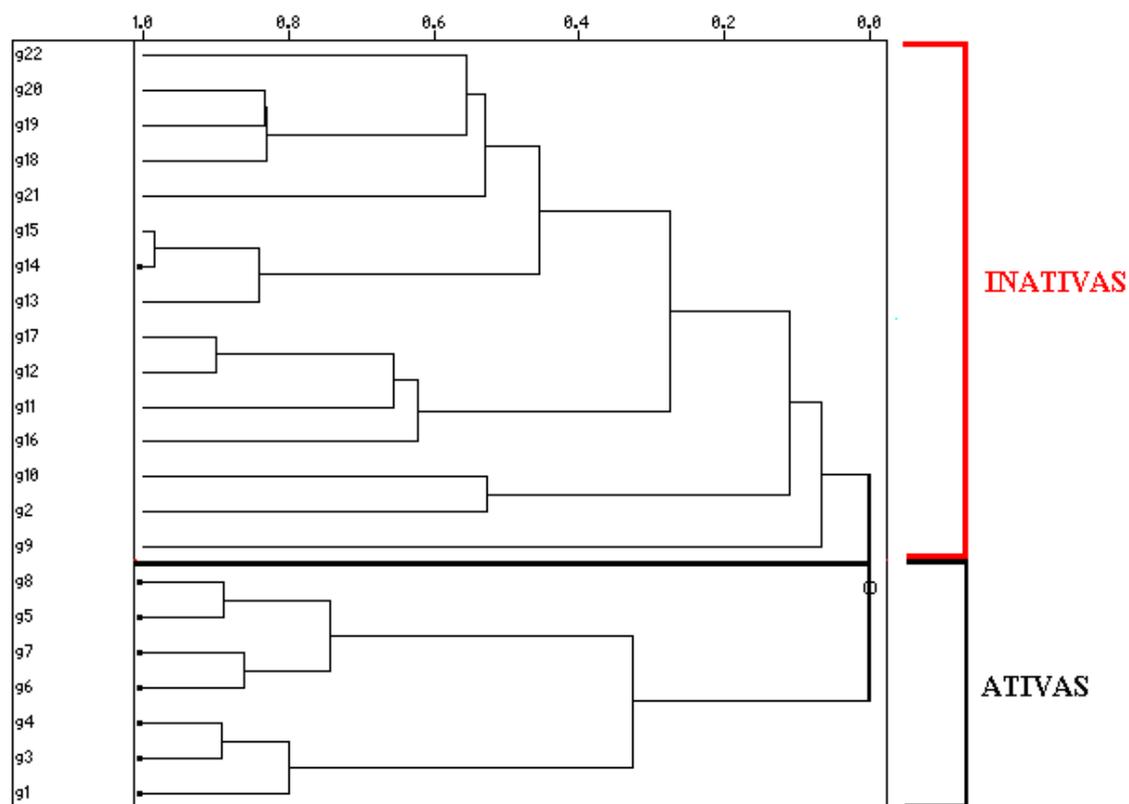
um resultado mais negativo da PC1 será obtido baixando os valores das variáveis Polari. e A2, enquanto que o valor da variável C20 deverá ter seu valor aumentado. Com essas considerações em mente, pode-se relacionar os resultados obtidos relativo as estruturas moleculares com a atividade biológica.

Os compostos G1, G3, G4, G5, G6, G7 e G8 são ativos à isoforma CYP1A1 e possuem o anel metileno-dioxifenil (MDP). Como observado, a atividade dos compostos aumenta com a diminuição da variável A2 e A2 é menor exatamente quando o MDP está presente. Portanto, pode-se postular que a presença do anel MDP é importante para a atividade dos derivados da piperina relativo á isoforma CYP1A1.

Como visto, para que um composto seja ativo devemos diminuir a polarizabilidade. Diminuir a polarizabilidade significa torná-los mais duros, ou seja, menos reativos. Nesse sentido, pode-se postular que o composto ativo não deve reagir com a isoforma CYP1A e sim sofrer apenas um encaixe com o receptor. A polarizabilidade dos compostos derivados pode ser alterada escolhendo grupos substituintes apropriados.

O terceiro descritor selecionado foi a carga atômica parcial sobre o átomo 20 (C20). Nota-se que um composto derivado cai no grupo dos ativos quando esta carga apresenta positivos. Maiores valores desse parâmetro, melhora a separação. Portanto, pode-se postular que a presença de uma região positiva (potencial eletrostático positivo) sobre o átomo 20 é importante para fixar a molécula no sítio receptor e, portanto, inativando a proteína.

A análise hierárquica de cluster (HCA) realizada com as três variáveis selecionadas (A2, Polarizabilidade e C20) tem por objetivo agrupar os compostos e gerar um modelo para futuras predições qualitativas de atividades para novos compostos a serem modelados. O dendograma resultante do cálculo da similaridade dado pela distância entre vizinhos, pode ser visto na Figura IV.5.



**Figura IV.5** – Dendrograma dos 22 compostos analisados relativo às variáveis A2, Polarizabilidade e C20.

Observa-se na figura IV.5 que os compostos ativos foram separados dos inativos e agrupados em grupos distintos. A formação de dois grupos distintos, um para os compostos ativos e outro para os inativos, mostra que os descritores selecionados na análise de componentes principais foram capazes de avaliar corretamente a similaridade que deve existir entre os compostos ativos e entre os inativos. Os dois grupos, ativo e inativo, estão separados por uma linha horizontal cheia. Os compostos pertencentes ao grupo dos ativos são: G1, G3, G4, G6, G7, G5 e G8. Os demais compostos pertencem ao grupo dos inativos. Tanto no HCA quanto no PCA, o composto G14, que tem valor de  $IC_{50}$  menor do que 100, foi classificado no grupo das inativas. É importante salientar que dentre os compostos que possui  $IC_{50}$  menor do que 100, ele é o único composto que não possui em sua estrutura molecular o anel MDP. Isto sugere que talvez o valor experimental publicado na referência [25], para este derivado esteja incorreto.

## Conclusão

A análise de componentes principais (*PCA*) realizada no conjunto de variáveis geométricas e eletrônicas calculadas a nível de teoria B3LYP/6-31G\* mostrou que os descritores A2 (ângulo entre os átomos 18,17 e 22), polarizabilidade e C20 (carga atômica parcial sobre o átomo 20) foram capazes de classificar os compostos derivados da piperina em dois grupos: ativos e inativos. A componente responsável pela classificação foi a *PC1* que explica 77,9% da variância total dos dados. As duas primeiras componentes juntas explicam 93,3% da variância total dos dados. A partir da equação matemática da primeira componente e da tabela de scores, nota-se que, para um composto cair no grupo das ativas, são necessários valores baixos dos descritores A2 e polarizabilidade e valor alto da C20. A análise hierárquica de cluster (*HCA*) realizada nos 22 compostos usando as variáveis selecionadas pela *PCA* reforça o resultado da *PCA*, classificando os compostos em dois grupos distintos. No entanto, o composto 14 (ver Figura I.6) que apresenta valor de  $IC_{50}$  menor que 100, portanto ativo, foi classificado como inativo tanto pela *PCA* quanto pela *HCA*. Este composto é o único entre os ativos que não tem o anel metileno-dioxifenil (MDP), característica dos compostos inativos. Isto sugere que talvez o valor experimental publicado deste composto não esteja correto. Os descritores selecionados na análise da *PCA* sugerem que a interação entre os derivados da piperina e o sítio receptor é de natureza física e não química, sendo que o potencial eletrostático positivo sobre o átomo 20 é importante para prender a molécula no sítio da isoforma CYP1A1. Conseqüentemente, esses descritores podem ser usados na modelagem de novos compostos que sejam ativos na inibição da proteína.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ATAL, C. K.; DHAR, K. L.; SINGH, J. *Chemistry of Indian Piper Species. Lloydia-the Journal of Natural Products*, v. 38, n. 3, p. 256-264, 1975.
- [2] ATAL, C. K.; ZUTSHI, U.; RAO, P. G. *Scientific Evidence on the Role of Ayurvedic Herbals on Bioavailability of Drugs. Journal of Ethnopharmacology*, v. 4, n.2, p. 229-232, 1981.
- [3] BANO, G.; AMLA, V.; RAINA, R. K.; ZUTSHI, U.; CHOPRA, C. L. *The Effect of Piperine on Pharmacokinetics of Phenytoin in Healthy-Volunteers. Planta Medica*, n. 6, p. 568-569, 1987.
- [4] MELO, C. F. M. et al. Alternativas para o aproveitamento industrial da pimenta-do-reino (*Piper nigrum* L.). Belém: EMBRAPA – CPATU, p. 29, 1990.
- [5] IKAN, R. *Natural Products: a Laboratory Guide*. 2nd ed.; San Diego: Academic Press, xiv, 360p, 1991.
- [6] LIU, G. Q.; ALGERI, S.; CECI, A.; GARATTINI, S.; GOBBI, M.; MURAI, S. *Stimulation of Serotonin Synthesis in Rat-Brain After Antiepilepsirine, An Antiepileptic Piperine Derivative. Biochemical Pharmacology*, v. 33, n. 23, p.3883-3886, 1984.
- [7] PARMAR, V. S.; JAIN, S. C.; BISHT, K. S.; JAIN, R.; TANEJA, P.; JHA, A.; TYAGI, O. D.; PRASAD, A. K.; WENGEL, J.; OLSEN, C. E.; BOLL, P. M. *Phytochemistry of the Genus Piper. Phytochemistry*, v. 46, n. 4, p. 597-673, 1997.
- [8] BUDAVARI, S. *The Merck Index: an Encyclopedia of Chemicals, Drugs, and Biologicals*. 12th ed.; Whitehouse Station, N.J: Merck, (various pagings), 1996.
- [9] LANDENBURG, A.; SCHOLTZ, M. *Synthese Der Piperins,,Ure Und Des Piperins. Berichte Der Deutschen Chemischen Gesellschaft*, v. 27, p. 2958, 1894.
- [10] EPSTEIN, W. W.; NETZ, D. F.; SEIDEL, J. L. *Isolation of Piperine From Black Pepper. Journal of Chemical Education*, v. 70, n. 7, p. 598-599, 1993.
- [11] SEMLER, U.; GROSS, G. G. *Distribution of Piperine in Vegetative Parts of Piper Nigrum. Phytochemistry*, v. 27, n. 5, p. 1566-1567, 1988.
- [12] BANO, G.; AMLA, V.; RAINA, R. K.; ZUTSHI, U.; CHOPRA, C. L. *The Effect of Piperine on Pharmacokinetics of Phenytoin in Healthy-Volunteers. Planta Medica*, n. 6, p. 568-569, 1987.
- [13] BANO, G.; RAINA, R. K.; ZUTSHI, U.; BEDI, K. L.; JOHRI, R. K.; SHARMA, S. C. *Effect of Piperine on Bioavailability and Pharmacokinetics of Propranolol and Theophylline in Healthy-Volunteers. European Journal of Clinical Pharmacology*, v. 41, n. 6, p. 615-617, 1991.

- [14] GUPTA, S. K.; VELPANDIAN, T.; SENGUPTA, S.; MATHUR, P.; SAPRA, P. *Influence of Piperine on Nimesulide Induced Antinociception. Phytotherapy Research*, v. 12, n. 4, p. 266-269, 1998.
- [15] SHOBA, G.; JOY, D.; JOSEPH, T.; MAJEED, M.; RAJENDRAN, R.; SRINIVAS, P. S. S. R. *Influence of Piperine on the Pharmacokinetics of Curcumin in Animals and Human Volunteers. Planta Medica*, v. 64, n. 4, p. 353-356, 1998.
- [16] G. THOMAS, Química medicinal. Uma Introdução, Guanabara Koogan, Rio de Janeiro, 2003.
- [17] OMURA T, SATO R. *The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. J. Biol. Chem.* P. 2379-2385, 1964.
- [18] KOROLKOVAS, A.; Burckhalter, J. H.; *Química Farmacêutica*, Ed. Guanabara: Rio de Janeiro, cap. 2, 1982.
- [19] HARA, C.; Rocha, F.L. Interações medicamentosas: antidepressivos novos e o sistema citocromo P450. *J. Bras. Psiqu.*, 47(1):9-18, 1998.
- [20] BENET LZ, Kroetz DL, Sheiner LB. Farmacocinética: a dinâmica da absorção, distribuição e eliminação dos fármacos. In: Hardman JG, Limbird LE, editores-chefes. Goodman & Gilman: As bases farmacológicas da terapêutica. Santiago: McGraw-Hill; p.3-20, 1996.
- [21] KLASSEN, C.D (Ed.) *Casarett & Doull's Toxicology. The basic science of poisons.* 5<sup>th</sup>.ed. New York: McGraw-Hill, 1996.
- [22] SINGH, J.; Reen, R. K.; Atal, C. K. *J. Pharm. Exptl Therap*, p. 236,488, 1986.
- [23] SINGH, J.; Reen, R. K.; Wiebel, F. J. *Cancer Letters*, p. 86, 195, 1994.
- [24] REEN, R. K.; Jamwal, D. S.; Taneja, S. C.; Koul, J. L.; Dubey, R. K.; Wiebel, F. J.; *J. Biochem. Pharmacol*, p. 46, 229, 1993.
- [25] SURRINDER Koul, JAWAHIR L. Koul, SUBHASH C. Taneja, KANAYA L. Dhar, DESHVIR S. Jamwal, KULDEEP Singh, RASHMEET K. Reen e JASWANT *Natural Products Chemistry Division, Regional Research Laboratory (CSIR), Jammu-tawi 180001, India Biochemistry Laboratory, Division of Pharmacology, Regional Research Laboratory (CSIR), Jammu-tawi, India*, 1999.
- [26] ORITIZ de Montellano, P. R.; Reich, N. O. *Inhibition of cytochrome P450 enzymes In Cytochrome P450, Structure, Mechanism and Biochemistry*; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, p. 273±314, 1986.
- [27] NEBERT, D. W. *Critical Reviews in Toxicol*, p. 137. 1989.
- [28] BARIK, B. R.; Kundu, A. B.; DEY, A. K. *Phytochemistry* p. 26, 1987.
- [29] REEN, R. K.; Ramakanth, S.; Wiebel, F. J.; Jain, M. P.; *Singh, J. Analyt. Biochem*, p. 24, 1991.

- [30] Surrinder Koul, Jawahir L. Koul, Subhash C. Taneja, Kanaya L. Dhar, Deshvir S. Jamwal, Kuldeep Singh, Rashmeet K. Reen and Jaswant Singh, Natural Products Chemistry Division, Regional Research Laboratory (CSIR), Jammu-tawi, India Biochemistry Laboratory, Division of Pharmacology, Regional Research Laboratory (CSIR), Jammu-tawi, India Received 16 July 1999; accepted 25 September, 1999.
- [31] 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. POPLER, Gaussian 03, Revision B.04, gaussian, inc., pittsburgh pa, 2003.
- [32] HYPERCHEM. *Molecular Visualization and Simulation*. Ontario, Hypercub, 1995.
- [33] EINSIGHT 3.0, *Infometrix, Seattle, USA*, 1991.
- [34] A. D. BECKE, *Phys. Rev. A* 38, 3098, 1988.
- [35] C. LEE, W. YANG, and R.G. PARR *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev.*, 1988.
- [36] A. PETERSSON AND M. A. AL-LAHAM, *J. Chem. Phys.* 94, 6081, 1991.
- [37] G. A. PETERSSON, A. BENNETT, T. G. TENSFELDT, M. A. AL-LAHAM, W. A. SHIRLEY, AND J. MANTZARIS, *J. Chem. Phys.* 89, 2193, 1988.
- [38] B. H. BRANDOW, *rev. Mod. Phys.* 39, 771, 1967.
- [39] PARR.R.G.;YANG.W.*Density – Funcional Theory of atoms and molecules, General, Oxford University Press. New York. Claredon Press Oxford, 338p,1989.*
- [40] O. GUNNARSSON and B.I. LUNDQVIST: *Exchange and Correlation in Atoms, Molecules and Solids by the Spin-Density-Functional Formalism. Phys. Rev. B*13, 4274, 1976.

- [41] J.P. PERDEW *Density-functional approximation for the correlation energy of the inhomogeneous electron gas*, *Phys. Rev. B* 33 8822-8824, 1986.
- [42] A.D. BECKE, *Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals*, *J. Chem. Phys.* 107, 8554-8560, 1997.
- [43] S.H. VOSKO, L. WILK, and M. NUSAIR *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, *Can. J. Phys.* 58, 1200-1211, 1980.
- [44] A.D. BECKE, *Density-functional exchange-energy approximation with correct asymptotic behaviour*, *Phys. Rev. A* 38, p. 3098-3100, 1988.
- [45] MASSART, D.L. *Chemometrics tutorial*. Amsterdam, Elsevier, 470p, 1990.
- [46] FEINBERG, M. *Anal. Chim. Acta.* V.191, p.75-85.1986.
- [47] NETO, B.B.; SCARMINIO, I.S.; BRUNS, R.E. *Planejamento e otimização de experimentos*. Campinas, UNICAMP, 299p, 1995.
- [48] BRUNS, R.E.; FAIGLE, J.F.G.; "Quimiometria", *Quím. Nova* 8:84, 1985.
- [49] SHARAF, M.A.; ILLMAN, D.L.; KOLWALSKI, B. R. *Chemometrics*. New York, John Wiley, 332p, 1986.
- [50] BRYAN J. F. MANLY, *Métodos de estatística multivariados – Uma Introdução - 3ª Ed*, 1990.
- [51] HARRIS, R.J. *A primer of multivariate statistic*. New York, Academic Press, 332p, 1975.
- [52] COOLEY, W. W. and LOHNES, P. R. *Multivariate Data Analysis*. Wiley, New York, NY, 1971.
- [53] PRESS, S.J. *Applied multivariate statistics*. New York, Holts, Rinehart and Winston, 345p, 1972.
- [54] BARNETT, *Advanced Level Studies; Statistics, Statistical Background V.2*. Center for Statistical Education, University of Sheffield, Sheffield, U.K. p. 114-132, 1985.
- [55] MOITA NETO, J.M.; MOITA, G.C. *Uma introdução à análise exploratória de dados multivariados*. *Química Nova*, 21 (4), 467-469, 1998.
- [56] GOLOBOÈANIN, D.D.; ŠKIRIBAE, B.D.; MILJEVIAE, N.R. 2004. *Principal component analysis for soil contamination with PAHs*. *Chemometrics and Intelligent Laboratory Systems*, p. 219-223, 1972.
- [57] YUDEL, Y.; DEMIR, C. *Principal component analysis and cluster analysis for the characterization of marbles by capillary electrophoresis*. *Talanta*, P. 63, 451, 2004.

[58] DANIELSSON, A.; CATO, I.; CARMAN, R.; RAHM, L. *Spatial clustering of metals in the sediments of the Skagerrak/Kattegat*, *Applied Geochemistry*, 14:689-706, 1999.

[59] BRENNEMAN, C. ., and WILBERG, K. B., *Determining atom-centered monopoles from molecular electrostatic potentials. The Need for high Sampling Densit in Formamide Conformational Analysis* *J. Comp. Chem.*, p. 361, 1990.

