



**UNIVERSIDADE FEDERAL DE PERNAMBUCO**  
**Centro de Ciências Exatas e da Natureza**  
**Departamento de Química Fundamental**  
**Programa de Pós-Graduação em Química**

## **Dissertação de Mestrado**

# **Aplicações de Metodologias Teóricas em Estudos de QSAR-3D de Pirimidinas Substituídas**

**Maíra de Almeida Carvalho**

**Recife-PE Brasil**

**Novembro / 2004**



**UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA  
DEPARTAMENTO DE QUÍMICA FUNDAMENTAL  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

# **Aplicações de Metodologias Teóricas em Estudos de QSAR-3D de Pirimidinas Substituídas**

**Maíra de Almeida Carvalho\***

Dissertação apresentada ao  
Programa de Pós-Graduação em  
Química da UFPE como parte  
dos requisitos para a obtenção  
do título de Mestre em Química.

**Orientador: Prof. Dr. João Bosco Paraíso da Silva**

**Co-orientador: Prof. Dr. Marcelo Zaldini Hernandes**

**\*Bolsista CNPq**

**Recife-PE Brasil  
Novembro / 2004**

**Dissertação submetida ao Corpo Docente do Programa de Pós-Graduação em Química do Departamento de Química Fundamental da Universidade Federal de Pernambuco como parte dos requisitos necessários para a obtenção do Grau de Mestre em Química.**

**Aprovada:**

---

**Prof. Dr. João Bosco Paraíso da Silva**  
Departamento de Química Fundamental da UFPE  
(Orientador)

---

**Prof. Dr. Alfredo Mayall Simas**  
Departamento de Química Fundamental da UFPE

---

**Prof. Dr. Roy Edward Bruns**  
Instituto de Química da Universidade Estadual de Campinas

## **"Aplicações de Metodologias Teóricas em Estudos de QSAR-3D de Pirimidinas Substituídas"**

**por**

**MAÍRA DE ALMEIDA CARVALHO**

Departamento de Química Fundamental  
Centro de Ciências Exatas e da Natureza  
Universidade Federal de Pernambuco  
Recife-PE-Brasil  
30 de novembro de 2004

A todos aqueles que acreditaram e torceram por mim. Especialmente ao meu pai, à minha mãe, aos meus irmãos Joca, Luiz e Bruno, às minhas irmãs Cláisse e Rita e às minhas primas Nina e Ana por todo o carinho e amizade.

## **AGRADECIMENTOS**

Agradeço a todos aqueles que de alguma forma contribuíram para o desenvolvimento desta dissertação, seja pela amizade irrestrita nos bons e maus momentos, seja pela ajuda em científica propriamente dita, durante todos esses anos de graduação e mestrado.

Agradeço ao meu pai, Eduardo, por ter me proporcionado estabilidade e segurança durante todos esses anos da minha vida. A minha mãe, Melania, pelo amor, pela amizade, pelo carinho e principalmente pela paciência e compreensão em raramente poder visitá-la devido os diversos atropelos da vida. Aos meus irmãos, Joca, Luiz e Bruno, pelas horas de conversa fiada, pelas longas risadas antes de dormir e pelas muitas horas que deixaram o computador disponível só para mim. A minha irmã, Clarisse que mesmo de longe torceu e acreditou em mim, que mesmo me encontrando apenas no celular, passava horas conversando comigo, a minha irmã Rita, que vê em mim, não sei como nem porque, a pessoa mais inteligente do mundo (acho que é porque sou mais velha...). À minha madrasta, Norma, por ter me recebido de braços abertos na sua casa, e por ter me ensinado a olhar para frente quando temos um grande problema ou pequenas dificuldades. Ao meu padrasto Evandro, pela confiança depositada em mim. Agradeço também a André, meu cunhado, pela torcida pelo meu sucesso. A Prazeres por ajudar a organizar minha vida e cuidar de mim. Agradeço a todos ele por terem, principalmente, acreditado em mim.

A Prof. Bosco, orientador e amigo, por estes anos de amizade e ensinamento. Pelos conhecimentos que me passou, pela paciência e confiança em mim.

A Prof. Marcelo Zaldini pela co-orientação, o qual esteve sempre presente nestes dois anos, sempre me ajudando, acreditando e confiando em mim.

A todos os professores do DQF, especialmente os professores do Grupo de Química Teórica e Computacional (GQTC) pelas horas de orientação e de gerenciamento de um trabalho em equipe. Pelos churrascos de confraternização do grupo, e por tê-los como professores e amigos.

Aos meus heróis, o professor Oscar Malta e o meu amigo Hélcio Batista, pelo socorro, ajuda, paciência, carinho, dedicação e tudo mais o que não consigo expressar

em palavras, no difícil momento do nosso acidente na volta da XII SBQT, os quais passaram a ser pessoas muito especiais para mim!

Agradeço a Rosa por estes anos de acompanhamento, onde pude amadurecer e aprender a encarar a vida de uma maneira mais agradável.

As minhas grandes e maravilhosas amigas, que vou citar por ordem alfabética para não ficarem com ciúmes, Ana, Amanda, Babie, Carol, Elisa, Karina, Karla, Nina, Patty e Zinha e aos meus amigos Guga, Mathias e Ricardo Freire, por todos os nossos momentos de estudo, balada, companheirismo, ginástica, etc, ou seja, por estarmos juntos em todos os momentos, rindo ou chorando.

Agradeço a Klaus, uma ótima pessoa que tive a oportunidade de conhecer e que sem ele essa dissertação não teria sido o que é. Com as implementações em tempo hábil para resolver algum problema ou necessidades que foram surgindo, ele me ajudou de uma maneira indescritível.

A todos os meus amigos, àqueles que conheci dentro e fora da universidade, que também me ajudaram, cada um a sua maneira, a desenvolver o meu projeto. Cito alguns deles, aqueles que sou capaz de me lembrar agora: Mário, Fred, Andrea, Giovannia, Paula, Flavia, Claudia, Vivianni, Mozart, Juliana, Sidney, Paula Coutinho, Marcus, Silvania, Marcio, Joacy, Gerd, Pequeno, Ênio, Rodrigo, Artur, Diego Araújo, Felipe (Bolinho), Cecília Meira, Felippe, Carlinhos, Henrique, Marconi, Gilvan, Pedro Hugo, Paulo Henrique, Paulo (LQTM) e Érika (LQTM).

Ao CNPq pela bolsa de mestrado.

Este trabalho tem em vista um estudo sistemático para a determinação de um modelo QSAR-3D para uma série homóloga de pirimidinas substituídas, cujas atividades anti-inflamatórias são conhecidas. Estudos preliminares de QSAR clássica, incluindo parâmetros calculados pelo método AM1, mostraram a importância de parâmetros eletrônicos na previsão, bem sucedida, de um novo composto mais ativo.

Devido à dependência dos parâmetros eletrônicos dos modelos QSAR com a geometria e função de onda utilizada, foi realizada uma investigação sistemática do nível de cálculo e do conjunto de base na etapa inicial deste trabalho.

A molécula não substituída (**5h**) foi selecionada como modelo para os demais compostos da série e, inicialmente, cálculos de orbitais moleculares em níveis AM1, HF e DFT e MP2 foram realizados para as funções de base com e/ou sem funções difusas e de polarização através do programa G98. A análise quimiométrica e comparações com dados de cristalografia de raio-X de sistemas similares sugeriram o emprego do método DFT-BP86/6-31G\*\* para o estudo das demais moléculas desta série.

Após a determinação do método e da função de base a serem utilizados, fez-se um estudo conformacional para determinação das geometrias de equilíbrio das pirimidinas substituídas e então foi avaliado o efeito da otimização completa ou parcial da geometria destas moléculas sobre a distribuição de cargas CHelpG, utilizando-se o método quimiométrico PCA.

Por último, foi o estudo da SPE desta série para as moléculas totalmente otimizadas. Nesta etapa foram utilizados os níveis de cálculo AM1 e BP86/6-31G\*\* para se avaliar o efeito da metodologia de cálculo na estimativa da SPE a partir de uma mesma geometria de equilíbrio. Utilizou-se o programa ALINHAMOL para estabelecer a geometria sob a qual seriam realizados os cálculos da SPE com o programa G98. Em seguida, o programa CAMOL foi utilizado para gerar uma matriz de dados do PE com a qual se desenvolveram os procedimentos de QSAR-3D. Foi feita uma análise sistemática da influência das variáveis de execução deste programa (valores de corte, etc.) na qualidade dos modelos quimiométricos (PCA e PLS) obtidos posteriormente.

Nossos resultados indicaram que há um grande comprometimento entre o método/função de base e a SPE, uma vez que o padrão do gráfico dos escores de PC1 *versus* PC2 dependerá de maneira significativa da escolha da metodologia de cálculo. Já a análise de regressão, mostrou que não há grande influência das variáveis do programa CAMOL (VCDS e VCIS) com a qualidade de previsão dos modelos obtidos.

The aim of this work is determining a 3D QSAR model for a homologous series of substituted pyrimidines with known anti-inflammatory activities. Previous results from classic QSAR (which include AM1 calculated parameters) have shown the importance of electronic parameters in the successful prediction of a novel and more active compound.

The first step was to make an evaluation of the calculation of level and the basis set due to the dependence between electronic parameters of QSAR models and the geometry and the base function used.

The non substituted molecule (**5h**) was selected as a model compound and, initially, molecular orbital calculations at AM1, HF, DFT and MP2 levels were used within and/or without base function that contain diffuse and/or polarization functions using the G98 program. A chemometrics analyses and a comparison of x-ray crystallographic data of related systems suggested the DFT-BP86/6-31G\*\* method to be used for the study oh the other molecules of this series.

After determinations of which method and basis set should be used, a conformation analysis and the evaluation of the effect of a partial and a full geometry optimization procedure over the atomic ChelpG charges was performed using the chemometric method of PCA.

Then an EPS study of this series was performed with the complete by optimized molecule. At this step AM1 and BP86/6-31G\*\* levels were used to assess the effect of the method on the EPS of the same equilibrium geometry. The ALINHAMOL program was used to align the molecules to the EPS calculation with the G98 program. After that, the CAMOL program was used to generate a data matrix of PE with which was performed the 3D QSAR procedure. Then, a systematic study of the influence of the variables of this program (cutoff values, etc) on the quality of the chemometric models (PCA and PLS) was carried out.

Our results had indicated a high relationship between method/base function and the EPS, because the pattern of PC1 versus PC2 scores plots will depend in a significant way on the choice of the calculation methodology. On the other hand the regression analysis showed that Camol's variables (VCDS and VCIS) do not have a great influence on the prediction quality of the obtained models.

## ÍNDICE GERAL

	Página
<b>Capítulo 1 - Introdução</b>	1
1.1. A Química Medicinal e o Desenvolvimento de Novos Fármacos	2
1.2. Interação fármaco-receptor	3
1.3. Estudos baseados na estrutura do bioreceptor	4
1.4. Estudos baseados na estrutura do ligante	4
1.5. Relação estrutura-atividade	5
1.6. Escolha e determinação dos descritores moleculares	7
1.7. O sistema de interesse	8
1.7.1. Teste da atividade biológica	9
1.7.2. Resultados Teóricos Prévios	10
<b>Capítulo 2 - Objetivos</b>	11
2.1. Objetivo Geral	12
2.2 Objetivos Específicos	12
<b>Capítulo 3 – Metodologia</b>	13
3.1. Cálculos de Química Quântica	14
3.2. Otimização da geometria molecular	16
3.2.1. Otimização da geometria molecular da pirimidina não substituída ( <b>5h</b> )	16
3.2.2. Otimização da geometria molecular dos substituintes	16
3.2.3. Otimização da geometria molecular total das pirimidinas substituídas	18
3.3. Estudo Conformacional	19
3.4. Planejamento Fatorial de Dois Níveis	20
3.5. Análise de Componentes Principais	21
3.6. Regressão pelo método dos mínimos quadrados parciais (PLS)	23
3.7. Alinhamento molecular e determinação da caixa 3D.	25
3.7.1. Alinhamento com o programa ALINHAMOL	26
3.7.2. Alinhamento com o programa MIPSIM	27
3.7.3. Determinação da caixa 3D	28
3.8. Escolha da Partição de Cargas	30

3.9. Cálculo da Superfície de Potencial Eletrostático e o programa CAMOL	31
3.10. Teste da Atividade Antiinflamatória	37
3.11. Infra-estrutura	38
<b>Capítulo 4 – Resultados e Discussões</b>	39
4.1. Estudo da Estrutura Padrão: a Molécula 5h	40
4.1.1. Análise conformacional da pirimidina 5h	40
4.2. Pirimidinas Substituídas	49
4.2.1. Análise conformacional	49
4.3. Comparação das pirimidinas substituídas	54
4.3.1. Comparações entre os ângulos diédricos C23C12C2N3 e C14C13C4N3	54
4.3.2. Comparação das cargas CHelpG em função da otimização da geometria	60
4.4. Alinhamento Molecular	66
4.5. Regressões da SPE das moléculas alinhadas no ALINHAMOL e calculadas no CAMOL	68
4.5.1. Determinação do Valor de Corte Independente do Sinal (VCIS)	69
4.5.2. Determinação do Valor de Corte Dependente do Sinal (VCDs)	75
4.5.3. Análise de componentes principais da SEP das pirimidinas totalmente otimizadas	81
4.5.4. Comparação da análise de componentes principais da SEP das pirimidinas semi-otimizadas com as totalmente otimizadas	89
4.5.5. PLS da SEP das pirimidinas totalmente otimizadas – Determinação de um modelo de QSAR-3D	91
<b>Capítulo 5 – Conclusões</b>	99
<b>Capítulo 6 – Perspectivas</b>	102
6.1. Cálculo da SPE das pirimidinas alinhadas com o programa MIPSIM	103

---

6.2. Estrutura de hidratação	103
6.3. Outras sondas moleculares	103
6.4. Outras QSAR-3D	104
<b>Capítulo 7 – Referências</b>	105
<b>Anexo A</b>	111
<b>Anexo B</b>	118
<b>Anexo C</b>	129

## ÍNDICE FIGURAS

	Página	
<b>Figura 1.1</b>	Esquema de informação no estudo da ação dos fármacos baseado no ligante.	4
<b>Figura 1.2</b>	Pirimidinas Substituídas.	8
<b>Figura 3.1</b>	Moléculas de benzeno substituídos, calculados com o método BP86/6-31G**.	17
<b>Figura 3.2</b>	Esquema da transferência dos parâmetros geométricos dos substituintes no sistema benzeno substituído para a pirimidina não substituída otimizada, gerando as pirimidinas substituídas semi-otimizadas.	18
<b>Figura 3.3</b>	Numeração dos átomos envolvidos na barreira de rotação do anel benzênico na molécula <b>5h</b> .	19
<b>Figura 3.4</b>	Esquema da rotação do grupo metoxi no: (a) metoxibenzeno e (b) nas pirimidinas substituídas.	19
<b>Figura 3.5</b>	(a) Esquema hierárquico de variáveis e (b) projeção dos objetos sobre as PC's.	23
<b>Figura 3.6</b>	Superposição das estruturas provenientes do alinhamento das pirimidinas <b>5a-q</b> em relação a molécula <b>5h</b> feita com o programa ALINHAMOL.	27
<b>Figura 3.7</b>	Esquema do número de pontos onde será calculado o potencial eletrostático para uma caixa com 2 Å além do extremo da molécula para cada uma das direções x, y e z, com resolução de (a) 0,5 Å e (b) 1,0 Å.	28
<b>Figura 3.8</b>	Esquema da comparação do tamanho das caixas 3D e posicionamento da molécula.	30
<b>Figura 3.9</b>	Esquema dos pontos onde será calculada a SPE.	31
<b>Figura 3.10</b>	Número de pontos onde é calculado o potencial eletrostático com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.	32

<b>Figura 3.11</b>	Esquema para exclusão dos pontos da SPE de acordo com o VCDS e o VCIS.	33
<b>Figura 3.12</b>	Pontos remanescentes da SPE após utilização dos valores de corte do CAMOL com VCDS = 100% e VCIS = (a) 0% (b) 50% e (c) 90%.	34
<b>Figura 3.13</b>	Pontos remanescentes da SPE após utilização dos valores de corte do CAMOL com VCIS = 20% e VCDS = (a) 30% (b) 50% e (c) 100%.	36
<b>Figura 4.1</b>	Molécula Padrão. Pirimidina substituída com $R_1 = R_2 = R_3 = H$ .	40
<b>Figura 4.2</b>	Superfície de energia da molécula <b>5h</b> calculada pelo método AM1 através da variação simultânea dos ângulos diédricos C23C12C2N3 e C14C13C4N3.	41
<b>Figura 4.3</b>	Histograma da freqüência de ocorrência dodiédro C14C13C4N3.	44
<b>Figura 4.4</b>	Gráfico das cargas de (a) Mulliken e (b) CHelpG da molécula <b>5h</b> em função do método, calculadas com os métodos AM1, HF e DFT (B3LYP e BP86) com funções de base 6-31G, 6-31G**, 6-31+G** e 6-31++G** e MP2 com funções de base 6-31G e 6-31G**.	45
<b>Figura 4.5</b>	Gráfico dos escores de PC1 x PC2 da PCA das cargas CHelpG da molécula <b>5h</b> , calculados com os métodos AM1, HF B3LYP e BP86 com funções de base 6-31G, 6-31G**, 6-31+G** e 6-31++G**, MP2 com funções de base 6-31G e 6-31G**.	47
<b>Figura 4.6</b>	Gráfico dos (a) Escores e (b) <i>Loadings</i> de PC1 x PC2 das cargas CHelpG da molécula <b>5h</b> , calculados com os métodos HF e DFT (B3LYP e BP86) com funções de base 6-31G, 6-31G**, 6-31+G** e 6-31++G**, MP2 com funções de base 6-31G e 6-31G**.	48
<b>Figura 4.7</b>	Pirimidina Substituída.	50

<b>Figura 4.8</b>	Gráficos das barreiras de energia relativa (em relação à conformação mais estável) (a) do metoxi-benzeno, e (b) da molécula <b>5f</b> calculadas com os métodos BP86/6-31G** (preto) e com AM1 (vermelho).	51
<b>Figura 4.9</b>	Conformações mais estáveis das pirimidinas substituídas com um grupo metoxi com as conformações (a) “dentro” (b) “fora” e com dois metoxi com as conformações (c) “dentro/dentro”, (d) “dentro/fora”, e (f) “fora/dentro” e (f) “fora/fora”.	53
<b>Figura 4.10</b>	Comparação entre as geometrias semi-otimizadas (azul) e totalmente otimizadas (vermelho) das moléculas <b>5a-n</b> obtidas pelo método BP86/6-31G**.	56
<b>Figura 4.11</b>	Comparação das geometrias de equilíbrio das moléculas <b>5a-q</b> totalmente otimizadas obtidas com os métodos BP86/6-31G** (vermelho) e AM1 (amarelo).	59
<b>Figura 4.12</b>	Gráfico das cargas CHelpG das pirimidinas <b>5a-q</b> : (a) com a geometria semi-otimizada e (b) com a geometria totalmente otimizada.	61
<b>Figura 4.13</b>	Gráfico dos escores de PC1 x PC2 das cargas CHelpG das moléculas <b>5a-q</b> com a geometria (a) semi-otimizada e (b) totalmente otimizada; e gráficos dos <i>loadings</i> de PC1 x PC2 das cargas CHelpG das moléculas <b>5a-q</b> com a geometria (c) semi-otimizada e (d) totalmente otimizada.	63
<b>Figura 4.14</b>	Gráfico dos escores de PC1 x PC2 das cargas CHelpG do conjunto de 22 moléculas das pirimidinas substituídas <b>5a-q</b> com a geometria (a) semi-otimizada e (b) totalmente otimizada; e gráficos dos <i>loadings</i> de PC1 x PC2 das cargas CHelpG das moléculas <b>5a-q</b> com a geometria (c) semi-otimizada e (d) totalmente otimizada.	65
<b>Figura 4.15</b>	Alinhamento molecular das pirimidinas substituídas <b>5a-q</b> obtido com o programa (a) ALINHAMOL com a molécula <b>5h</b> como referência e com o programa MIPSIM com as moléculas de referência (b) <b>5h</b> , (c) <b>5a<sub>d</sub></b> e (d) <b>5i<sub>dd</sub></b> .	67

<b>Figura 4.16</b>	Gráfico da função resposta R em função do VCIS, da caixa 5 com VCDS = 100% e $10\% < \text{VCIS} < 90\%$ .	70
<b>Figura 4.17</b>	Superposição dos gráficos dos escores de (a) PC1 x PC2 e (b) PC1 x PC3 da SPE (BP86/6-31G**) das pirimidinas <b>5a-q</b> totalmente otimizadas com $0\% < \text{VCIS} < 90\%$ e VCDS = 100% para a caixa 5.	70
<b>Figura 4.18</b>	Gráficos da caixa 8 da (a) da função resposta R em função do VCIS com VCDS = 100% e $10\% < \text{VCIS} < 90\%$ , e dos escores de (b) PC1 x PC2 e (c) PC1 x PC3 da SPE com VCDS = 100% e $0\% < \text{VCIS} < 90\%$ .	73
<b>Figura 4.19</b>	Gráfico da função resposta R em função do VCDS, com VCIS = 20% e $40\% < \text{VCDS} < 100\%$ da (a) caixa 5 e (b) caixa 8.	76
<b>Figura 4.20</b>	Gráficos superpostos dos escores de PC1 x PC2 e de PC1 x PC3 da SPE das Pirimidinas Substituídas: (a) da caixa 5 e de (b) da caixa 8, com $30\% < \text{VCDS} < 100\%$ e VCIS = 20%.	77
<b>Figura 4.21</b>	Esquema da redução no número de pontos da SPE da molécula <b>5i<sub>dd</sub></b> , com VCDS = 50% e VCIS = 20% para o conjunto completo de moléculas.	80
<b>Figura 4.22</b>	Gráfico dos escores da PCA da SEP (BP86/6-31G**) das pirimidinas substituídas 5a-q totalmente otimizadas de PC1 x PC2 e PC1 x PC3 com (a) VCDS = VCIS = 0% e (b) VCDS = 50% VCIS = 20%.	81
<b>Figura 4.23</b>	Gráfico dos <i>loadings</i> de PC1 x PC2 e PC1 x PC3 da PCA da SEP (BP86/6-31G**) das pirimidinas substituídas 5a-q totalmente otimizadas com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.	83
<b>Figura 4.24</b>	Localização dos pontos da SPE (BP86/6-31G**) das pirimidinas substituídas <b>5a-q</b> totalmente otimizadas com o valor dos <i>loadings</i> superiores a 60% do maior e menor valor em PC1, PC2 e PC3 (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.	84

<b>Figura 4.25</b>	Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da PCA da SEP (AM1) das pirimidinas substituídas <b>5a-q</b> totalmente otimizadas de (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.	85
<b>Figura 4.26</b>	Gráfico dos pesos de PC1 x PC2 da PCA da SEP (AM1) das pirimidinas substituídas <b>5a-q</b> totalmente otimizadas com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.	87
<b>Figura 4.27</b>	Localização dos pontos da SPE (AM1//BNP86/6-31G**) das pirimidinas substituídas <b>5a-q</b> totalmente otimizadas com o valor dos <i>loadings</i> e PC1, PC2 e PC3 superiores a 60% do maior e menor valores em cada PC (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.	88
<b>Figura 4.28</b>	Sobreposição dos gráficos dos escores de PC1 x PC2 e PC1 x PC3 da SEP (caixa 5) das pirimidinas substituídas com a geometria semi-otimizada e totalmente otimizada com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS=20%.	89
<b>Figura 4.29</b>	Gráfico da atividade experimental <i>versus</i> a prevista da PLS com VCDS = VCIS = 0% das pirimidinas com o grupo metoxi (a) dentro, (b) fora/dentro, (c) dentro/fora e (d) fora.	92
<b>Figura 4.30</b>	Mapa dos coeficientes da PLS entre o PE e a atividade antiinflamatória das pirimidinas <b>5a-q</b> , com VCDS = VCIS = 0% para as conformações (a) dentro, (b) dentro/fora, (c) fora/dentro e (d) fora.	93
<b>Figura 4.31</b>	Gráfico da atividade experimental <i>versus</i> a prevista da regressão PLS com VCDS = 50% e VCIS = 20% das pirimidinas (a) dentro, (b) fora/dentro, (c) dentro/fora e (d) fora.	95
<b>Figura 4.32</b>	Mapa dos coeficientes da PLS entre o PE e a atividade antiinflamatória das pirimidinas <b>5a-q</b> do PLS com VCDS = 50% e VCIS = 20% para as conformações (a) dentro, (b) dentro/fora, (c) fora/dentro e (d) fora.	96
<b>Figura A1</b>	Estruturas de raios-X, CSD, com o diedro C14C13C4N3.	112

<b>Figura A2</b>	Estrutura de raios-X, CSD, com o diedro C23C12C2N3.	117
<b>Figura B1</b>	Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas <b>5a-q</b> totalmente otimizadas com VCDS = 100% e 0% < VCIS < 90% da caixa 5.	119
<b>Figura B2</b>	Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas <b>5a-q</b> totalmente otimizadas com VCDS = 100% e 0% < VCIS < 90% da caixa 8.	124
<b>Figura C1</b>	Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas <b>5a-q</b> totalmente otimizadas com VCIS = 20% e 30% < VCDS < 100% da caixa 5.	130
<b>Figura C2</b>	Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas <b>5a-q</b> totalmente otimizadas com VCIS = 20% e 30 % < VCIS < 100% da caixa 8.	134

## ÍNDICE TABELAS

	Página
<b>Tabela 1.1</b> Pirimidinas substituídas e atividade antiinflamatória.	9
<b>Tabela 3.1</b> Matriz de Codificação de um Planejamento Fatorial 2 <sup>3</sup>	21
<b>Tabela 3.2</b> Parâmetros e nomenclatura das caixas 3D.	29
<b>Tabela 4.1</b> Ângulos diédricos C23C12C2N3 e C14C13C4N3, em graus, da pirimidina <b>5h</b> otimizada com os métodos AM1, HF, BP86 e B3LYP e funções de base 6-31G, 6-31++G, 6-31G**, 6-31+G** e 6-31++G**, e MP2 com funções de base /6-31G, 6-31G**.	42
<b>Tabela 4.2</b> Efeitos principais e de interação de um Planejamento Fatorial de Dois Níveis com três fatores, 2 <sup>3</sup> .	43
<b>Tabela 4.3</b> Pirimidinas substituídas.	50
<b>Tabela 4.4</b> Ângulos diédricos C23C12C2N3 e C14C13C4N3 das pirimidinas substituídas <b>5a-q</b> totalmente otimizadas com o método BP86/6-31G**.	55
<b>Tabela 4.5</b> Ângulos diédricos C23C12C2N3 e C14C13C4N3 das pirimidinas substituídas <b>5a-q</b> totalmente otimizadas com o método AM1.	58
<b>Tabela 4.6</b> Número de pontos utilizados e o tempo de CPU na PCA da SPE da caixa 5 das pirimidinas <b>5a-q</b> com o VCDS = 100% e 0% < VCIS < 90% .	72
<b>Tabela 4.7</b> Tempo de CPU utilizado na PCA da SPE da caixa 8 das pirimidinas <b>5a-q</b> com o VCDS = 100% e 0% < VCIS < 90% .	74
<b>Tabela 4.8</b> Tempo de CPU utilizado na PCA da SPE da caixa 5 das pirimidinas <b>5a-q</b> com 30% < VCDS < 100% e VCIS = 20%.	79
<b>Tabela 4.9</b> Tempo de CPU utilizado na PCA da SPE da caixa 8 das pirimidinas <b>5a-q</b> com 30% < VCDS < 100% e VCIS = 20%.	79

## Siglas e Notações

AM1	<i>Austin Model 1</i>
HF	Método de Hartree-Fock
MP2	Moller-Plesset em Segunda Ordem
DFT	Teoría do Funcional da Densidade
PFDN	Planejamento Fatorial de Dois Níveis
PCA	Análise de componentes principais
SPE	Superfície de Potencial Eletrostatico
VCIS	Valor de corte independente do sinal
VCDS	Valor de corte dependente do sinal
Atv	Atividade biológica
PLS	<i>Partial least square</i>
QSAR	<i>Quantitative structure-activity relationships.</i>
CHelpG	<i>Charge from Eletrostatic Potencial Grid based</i>
PCn	n-ésima componente principal

# CAPÍTULO 1

## INTRODUÇÃO

## 1.1. A Química Medicinal e o Desenvolvimento de Novos Fármacos

De um ponto de vista histórico, as plantas foram, e ainda hoje continuam sendo, uma importante fonte de identificação de compostos com algum tipo de atividade biológica. Entretanto, no início, a escolha das plantas para aplicações farmacêuticas baseava-se em aspectos mais empíricos que propriamente científicos. Atualmente, a busca por compostos líderes, bem como os estudos visando modificações estruturais em prol da melhoria da atividade ou diminuição de efeitos indesejáveis como a toxicidade dos compostos, representam dois dos principais propósitos da Química Medicinal (QM)<sup>1</sup>.

Apenas nos anos 40 com o surgimento da equação de Hammett<sup>2</sup>, começaram a surgir as primeiras contribuições da Físico-Química Orgânica que serviram de base para o desenvolvimento de novos fármacos. Foi na década de 60, com os trabalhos pioneiros de Hansch<sup>3</sup> relacionando a atividade biológica a parâmetros moleculares como a constante eletrônica de Hammett, que se inaugurou um novo período para a QM. Já a partir da década de 70 a Química Teórica (QT) passou a contribuir para os estudos de QSAR, entretanto foi na década de 80, como o advento de programas e computadores cada vez mais rápidos e mais velozes<sup>4</sup>, que a QT passou a auxiliar de forma mais sistemática a QM<sup>5</sup>.

Como consequência do avanço computacional, muitas ferramentas foram geradas e passaram a ser amplamente utilizadas para construir representações e modelos de estrutura molecular, bem como o cálculo de diversas propriedades moleculares<sup>6,7,8</sup>. O comportamento e as características de uma determinada molécula ou conjuntos das mesmas podem ser simulados através de ferramentas quânticas e clássicas<sup>9,10</sup>, tendo a química quântica um caráter especial na obtenção de descritores que podem, a princípio, expressar as propriedades energéticas, geométricas e eletrônicas de uma molécula, e suas interações com o ambiente. Estas simulações fornecem várias propriedades que podem ser úteis em QM, pois podem ajudar a explicar conexões ou relações entre a estrutura e a atividade biológica<sup>11</sup>. São de importância nesta área a mecânica quântica, a mecânica molecular, a análise conformacional, a teoria de gráficos moleculares, o planejamento de moléculas auxiliado por computador e as relações quantitativas de estrutura-atividade<sup>12</sup>.

Independentemente da técnica adotada para o desenvolvimento de novos fármacos, sabe-se que pequenas modificações na estrutura de uma molécula podem modificar suas propriedades físico-químicas, podendo, assim, revelar efeitos que estavam latentes ou mascarados no composto matriz de partida (ou composto líder)<sup>13</sup>. Então, a busca por relações quantitativas entre as variáveis (ou dekritores) físico-químicos associadas a uma dada estrutura química (representada por uma ou algumas configurações e/ou conformações moleculares) e a resposta biológica é o objetivo do estudo da relação quantitativa estrutura atividade (do inglês QSAR). Estas relações ajudam a entender e explicar a direção das forças que agem por trás da ação do fármaco, contribuindo para o desenvolvimento de novos compostos com as propriedades biológicas desejadas, conforme está explicado no item 1.5. A resposta biológica é resultado do tipo de interação entre o fármaco e o receptor (biomacromolécula), a qual pode ser dividida, principalmente, em estruturalmente inespecífica (dependem única e exclusivamente de suas propriedades físico-químicas para promover o efeito biológico) e específica.

## 1.2. Intereração fármaco-receptor

A interação de um fármaco com o seu sítio de ação no sistema biológico define sua fase farmacodinâmica de ação<sup>14</sup> que envolve nesta interação, forças intermoleculares como interações eletrostáticas, hidrofóbicas, polares e estéricas.

Os fármacos estruturalmente específicos compreendem a maioria dos medicamentos, e seu efeito biológico deve-se a interação com determinada biomacromolécula, denominada receptor ou bioreceptor. O reconhecimento do fármaco (micromolécula ou ligante) pela biomacromolécula, é dependente da estrutura do fármaco, incluindo o arranjo espacial dos seus grupamentos funcionais, que devem ser complementares ao sítio de ligação localizado no receptor e também aos aspectos eletrônicos como as forças eletrostáticas, as forças de dispersão, as interações hidrofóbicas, as ligações de hidrogênio, a lipofilicidade entre outras.

Assim, pode-se utilizar duas aproximações para o desenvolvimento de novos

fármacos i) uma análise direta, baseada na estrutura do receptor ou ii) uma análise baseada apenas no ligante, como explicadas a seguir.

### 1.3. Estudos baseados na estrutura do bioreceptor

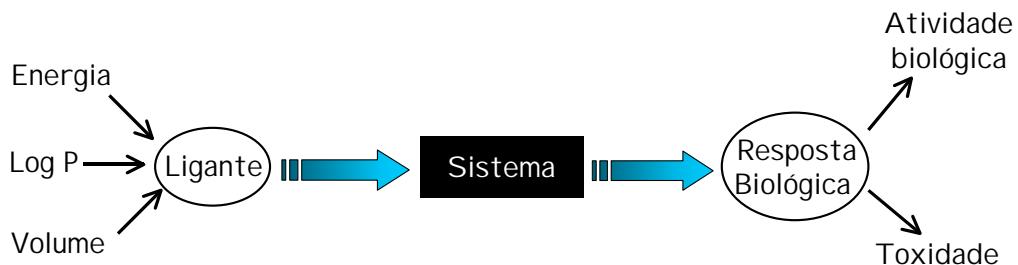
O conhecimento da estrutura do bioreceptor pode tornar o desenvolvimento de novas drogas mais simples, pois desta forma pode-se, em tese, desenvolver moléculas ligantes com características necessárias para uma maior complementaridade com o bioreceptor. Porém, a determinação da estrutura tridimensional do bioreceptor (biomacromolécula responsável ou associada a atividade) pode ser algo bastante complexo, tendo que passar por etapas como: extração, purificação, crescimento de cristal e difração de raios-x. Após estas, ainda é necessário um estudo para a determinação do sítio de interação do fármaco com o receptor e, finalmente, uma análise de complementaridade.

Logo, por estas razões, em muitos casos, o desenvolvimento de novos fármacos não utiliza a estrutura do bioreceptor para o planejamento de novos ligantes.

### 1.4. Estudos baseados na estrutura do ligante

Por outro lado, quando a estrutura do receptor não é conhecida (situação mais freqüente), o conhecimento da estrutura do ligante e de propriedades físico-químicas a ela associadas, pode trazer valiosas informações a cerca do funcionamento do fármaco.

Em geral, a situação mais freqüentemente deparada, quando apenas se conhecem informações sobre o ligante, pode ser representada no esquema da Figura 1.1.



**Figura 1.1:** Esquema de informação no estudo da ação dos fármacos baseado no ligante.

Pode-se notar, neste esquema, que o planejamento do ligante pode ser realizado a partir de diversas propriedades físico-químicas, como por exemplo, o Log P da molécula (coeficiente de partição da molécula, em geral medido pela relação octanol/água) a energia do sistema, etc. Porém, uma vez no organismo biológico, nem sempre é possível se determinar, de forma simples, o processo metabólico deste fármaco, sendo, então, este processo análogo a uma caixa preta, onde apenas se conhece a resposta de saída, ou seja, a resposta biológica provocada pelo fármaco após seu transporte e metabolismo. Essa resposta biológica pode representar informações distintas como a atividade biológica contra algum tipo de enfermidade, ou mesmo a toxicidade da molécula.

De um ponto de vista mais específico, a correspondência entre propriedades do ligante e do receptor, ou mesmo do ligante e da resposta biológica, pode ser estudada através de uma análise QSAR. Quando dados são escassos ou imprecisos, pode-se realizar simplesmente uma Relação Estrutura-Atividade (do inglês SAR)<sup>15</sup>

## 1.5. Relação estrutura-atividade

O desenvolvimento da QSAR está mudando o perfil de investigação na busca por novos fármacos que antes baseava-se puramente na experiência e na intuição química, para uma maneira matematicamente mais fundamentada, por exemplo, ajustando-se curvas e prevendo-se novas atividades através de equações envolvendo a relação entre a estrutura química e a atividade biológica. Uma vez determinada a relação entre uma estrutura ou propriedade com a atividade desejada, qualquer composto, inclusive os ainda não sintetizados, podem ser avaliados, preliminarmente à sua síntese, no computador. Assim é possível escolher, por exemplo, os compostos mais e menos promissores para então serem sintetizados, testados e assim, avaliar os modelos de QSAR.

Porém, para se realizar uma QSAR de sucesso, pelo menos três critérios são fundamentais: i) possuir valores de atividade confiáveis, ii) possuir uma boa variância na resposta biológica e iii) determinar e identificar descritores físico-químicos relevantes para os cálculos de regressão, sendo este último, um permanente desafio no

desenvolvimento de novos fármacos. Após a contribuição fundamental de Hansch<sup>3</sup>, o qual demonstrou que é possível obter funções que correlacionavam descritores físico-químicos com as propriedades biológicas de compostos, vários outros descritores vem sendo projetados e usados no planejamento de fármacos, e mais especificamente, na relação quantitativa estrutura-atividade.

A quantificação racional da relação estrutura-atividade de uma série de compostos, pode ser feita através de um estudo de QSAR clássica<sup>16</sup>, onde são usados apenas descritores unidimensionais, como a constante de Hammett, o coeficiente de partição Log P, a energia de orbitais, o momento de dipolo, enfim, por um conjunto de descritores que são associados a molécula como um todo. Em 1985, Goodford<sup>17</sup> introduziu o conceito de campo de interação molecular (MIF) e o trabalho de Cramer e colaboradores introduziu a idéia do uso de parâmetros tridimensionais (3D) de compostos<sup>18</sup>, desenvolvendo, portanto, o conceito de QSAR-3D, onde descritores tridimensionais são usados para descrever a molécula e, ainda, através de QSAR-4D e 5D, onde são levados em conta os graus de liberdade do ligante e do receptor, respectivamente.

A análise comparativa do campo de interação molecular (do inglês CoMFA)<sup>18</sup> é uma das ferramentas mais poderosa em QSAR-3D, onde as atividades biológicas podem ser correlacionadas com os campos moleculares, como por exemplo, os campos eletrostáticos e/ou estéricos ao redor da molécula do ligante. Um aspecto importante que não se pode perder de vista, é que esses descritores moleculares (propriedades de interesse) devem ser escolhidos de modo que sua propriedade tenha um significado físico de fácil interpretação<sup>19</sup>. Estes campos geram mapas de contorno da superfície ao redor da molécula e são obtidos através da utilização de sondas que mimetizam grupos químicos presentes tipicamente no sítio de ligação ou sítio ativo, como por exemplo, carga de prova (sonda eletrostática), a molécula de água (sonda hidrofilica), grupos com átomos doadores e aceptores de ligação de hidrogênio, etc.

A QSAR-3D é feita utilizando o método de campo de interação via métodos estatístico<sup>20</sup>, como as análises multivariadas Análise de Componentes Principais (PCA), ou mínimos quadrados parciais (PLS)<sup>21</sup>, que podem ser usados para estabelecer o melhor conjunto de variáveis independentes. Estes métodos vêm sendo amplamente utilizado em desenvolvimento de novos fármacos, uma vez que não apresentam a eventual deficiência do método de regressão linear múltipla de colinearidade dos

descritores, podendo, portanto, fornecer um modelo quantitativo da relação estrutura-atividade biológica.

Para obter boas correlações entre uma propriedade ou estrutura e a sua atividade biológica é fundamental que se utilizem descritores apropriados, indiferentemente se eles vêm de origem teórica, empírica, ou derivados de medidas experimentais obtidas dos ligantes<sup>22</sup>. Muitos dos descritores refletem propriedades moleculares simples, e podem ajudar na compreensão da natureza físico-química da atividade/propriedade sob consideração e, assim, viabilizar a proposição da modificação estrutural que leve a mudança de atividade/propriedade desejada.

Neste trabalho, foram usados alguns descritores de origem da mecânica quântica, obtidos através de cálculos teóricos, como as cargas atômicas e a superfície do potencial eletrostático (SPE).

Vale ainda ressaltar que, atualmente, a maioria das pesquisas em QM no Brasil ainda depende de programas computacionais comerciais, que além de caros (da ordem de milhares de dólares), são pacotes fechados, onde, em geral, não é possível a implementação de novas modificações para a resolução de possíveis problemas que surjam durante o desenvolvimento de um projeto, limitando assim, o desenvolvimento das pesquisas na maioria das universidades e centros de pesquisa brasileiros.

Então, neste contexto, foi particularmente importante o desenvolvimento de alguns programas computacionais na nossa instituição, como por exemplo, o ALINHAMOL<sup>23</sup> para o desenvolvimento de pesquisas em QSAR-3D. Em sintonia com estes esforços, os estudos que se desenvolveram nesta dissertação serviram, inclusive, como parâmetro de referência para modificações dos próprios códigos computacionais. Este sinergismo foi de extrema importância para o desenvolvimento deste trabalho, sendo também possível, através da interação com o grupo de síntese e teste de atividade, a proposição de um novo composto planejado. Ainda, vale a pena enfatizar que todos os grupos de pesquisa envolvidos fazem parte da UFPE. Este foi um aspecto importante para o desenvolvimento desta dissertação.

## 1.6. Escolha e determinação dos descritores moleculares

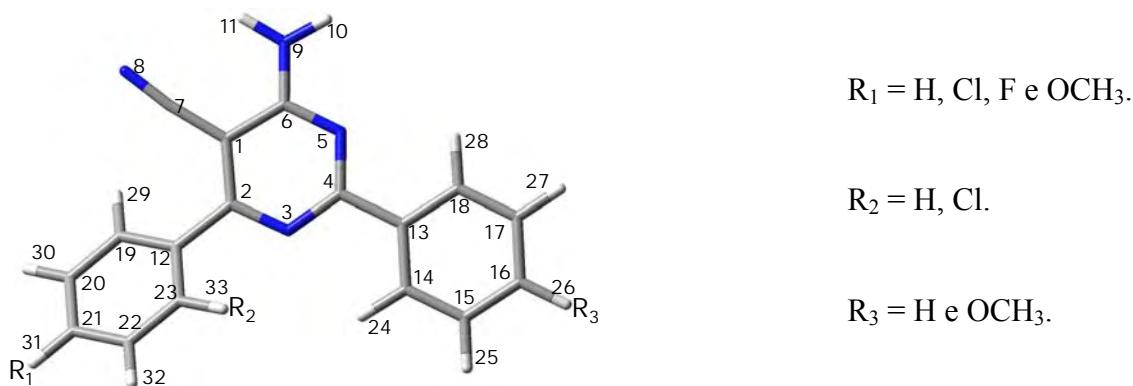
A maioria das metodologias de planejamento de drogas começa com uma

descrição das entidades moleculares envolvidas, ligantes e receptores, através da utilização de diversas sondas químicas, isto é, de grupos químicos que mimetizam as interações entre o ligante e o bioreceptor, no cálculo de superfícies moleculares. Normalmente, relaciona-se a interação ligante-receptor como consequência de interações de natureza eletrostática, sendo esta, portanto, uma das interações mais importante entre os ligantes e a biomacromolécula, onde, alguns programas computacionais vêm sendo desenvolvidos no intuito de se calcular o campo molecular do ponto de vista do Potencial Eletrostático (PE), como por exemplo o programa CAMOL<sup>24</sup>.

Tendo em vista a dependência computacional entre o nível de cálculo e o conjunto de funções de base e cálculo do campo de interação molecular<sup>25</sup>, optou-se por analisar, inicialmente, as cargas atômicas, uma vez que envolve uma menor quantidade de variáveis. Entretanto, como as cargas atômicas mostram-se bastante sensíveis ao nível de teoria empregado no seu cálculo como, por exemplo, o conjunto de base utilizado e a introdução de correlação eletrônica<sup>26</sup>, bem como surgem diferença em seu valor de acordo com a partição utilizada, neste trabalho foram investigadas duas partições de carga, as cargas de Mulliken<sup>27</sup> e as cargas CHelpG<sup>28</sup>, derivada do potencial eletrostático, em diversos níveis de teoria, conforme pode ser visto na seção 3.8 da metodologia, página 30.

## 1.7. O sistema de interesse

O sistema de interesse é uma série homóloga de pirimidinas substituídas (Figura 1.2), as quais foram sintetizadas e tiveram sua atividade biológica<sup>29</sup> determinada no Departamento de Antibióticos desta universidade.

**Figura 1.2:** Pirimidinas Substituídas.

Os compostos pirimidínicos são heterocíclos amplamente distribuídos na natureza, estando presentes em importantes moléculas biológicas como, por exemplo, nos ácidos nucléicos. Estes compostos apresentam numerosas aplicações no que diz respeito a química farmacêutica, como por exemplo: atividade antifúngica<sup>30</sup>, antiinflamatória<sup>31</sup>, antitumoral<sup>32</sup>, protetora da mucosa gástrica<sup>33</sup> entre outras.

A atividade antiinflamatória dos compostos sintetizados pode ser vistas na Tabela 1.1.

**Tabela 1.1:** Pirimidinas substituídas e atividade antiinflamatória.

Molécula	$R_1$	$R_2$	$R_3$	Atv* (%)
<b>5e</b>	Cl	Cl	$OCH_3$	4,9
<b>5f</b>	F	H	$OCH_3$	10,9
<b>5a</b>	Cl	H	$OCH_3$	17,1
<b>5d</b>	Cl	Cl	H	20,3
<b>5b</b>	Cl	H	H	20,7
<b>5h</b>	H	H	H	38,7
<b>5c</b>	F	H	H	47,3
<b>5i</b>	$OCH_3$	H	$OCH_3$	47,8
<b>5l<sup>#</sup></b>	$OCH_3$	H	H	52,8

Atv\* Atividade antiinflamatória em % da inibição do edema da pata de rato, induzido por carragenina.

<sup>#</sup> Molécula proposta por Ramos e col. Baseado em modelos de QSAR.

Recentemente, Ramos e colaboradores<sup>34</sup> realizaram cálculos AM1 para obter parâmetros que foram utilizados numa análise de Regressão Múltipla contra a atividade

biológica (QSAR linear), com a qual foi possível propor, com sucesso, a estrutura do composto **5I** (Tabela 1.1). Todavia, o modelo obtido por estes autores não foi amplamente explorado nem confirmado por outro modelo teórico.

### 1.7.1. Teste da atividade biológica

No processo de modelagem molecular, a atividade biológica é geralmente determinada para um conjunto de compostos com estruturas homólogas. Este conceito consiste no fato que a similaridade química é também refletida na atividade biológica, ou seja, compostos quimicamente parecidos ou análogos podem ser relacionados no modo de ação.

O procedimento dos testes de atividade antiinflamatória encontra-se descrito em detalhes na tese de doutorado de Falcão<sup>29</sup>, entretanto faremos na seção 3.10 da metodologia, um breve resumo, a título de situar o leitor deste trabalho, sobre o que representa a função resposta que foi utilizada.

### 1.7.2. Resultados Teóricos Prévios

Após o teste das atividades antiinflamatórias dos compostos desta série, Ramos e colaboradores<sup>34</sup> propuseram um modelo teórico de regressão para a previsão de um composto mais ativo do que os inicialmente testados. O modelo mostrou-se bastante satisfatório com a previsão de um composto mais ativo o qual foi sintetizado e testado, comprovando assim o modelo proposto. O novo composto sugerido foi a molécula 5I (ver Tabela 1.1). Neste estudo, contudo, utilizou-se o método AM1 para determinação dos descriptores moleculares, e sendo este um método semi-empírico e o principal desritor sugerido foram descriptores eletrônicos, torna-se interessante a avaliação do modelo a partir de um nível de teoria mais elevado.

Desta forma, um estudo teórico destas moléculas a fim de se estabelecer uma QSAR-3D para a previsão de novos compostos desta série, com atividade mais elevada, foi realizado utilizando-se métodos *ab initio* sem e com correlação eletrônica e métodos do funcional densidade, cuja metodologia adotada está descrita no Capítulo 3 e os resultados obtidos encontram-se no Capítulo 4.

# CAPÍTULO 2

## OBJETIVOS

## 2.1. Objetivo Geral

O principal objetivo desta dissertação é o desenvolvimento de uma QSAR-3D de uma série homóloga de pirimidinas substituídas, as quais possuem suas atividades antiinflamatórias determinadas.

## 2.2 Objetivos Específicos

Como objetivos específicos desta dissertação tem-se:

- Aplicar técnicas multivariadas para avaliar a metodologia de cálculo da geometria molecular;
- Analisar a influência da geometria molecular no modelo de QSAR-3D,
- Avaliar o modelo de alinhamento molecular;
- Determinar os valores de corte do programa CAMOL mais adequados para esse sistema no cálculo da Superfície do Potencial Eletrostático;
- Avaliar padrões nas cargas atômicas em função da metodologia de cálculo, geometrias utilizadas e dos substituintes presentes nas pirimidinas;
- Obter um modelo de QSAR-3D através do modelo de regressão dos mínimos quadrados parciais.

# CAPÍTULO 3

## METODOLOGIA

### 3.1. Cálculos de Química Quântica

Os métodos de química quântica são, atualmente, amplamente utilizados na compreensão e na previsão do comportamento de propriedades moleculares. Eles estão baseados, na sua essência, na resolução da equação de Schrödinger independente do tempo, dada pela equação 3.1

$$H\Psi = E\Psi \quad \text{Equação 3.1,}$$

onde H é o operador hamiltoniano não relativístico, E é a energia total e  $\Psi$  é a função de onda que depende das coordenadas cartesianas e de spin de todas as partículas.

Porém, como esta equação não pode ser resolvida de maneira exata para sistemas multieletrônicos, então, algumas aproximações são utilizadas para viabilizar o estudo desses sistemas.

Uma das aproximações utilizadas é a Aproximação Born-Oppenheimer (B.O.A.), a qual considera o desacoplamento dos movimentos eletrônicos e nucleares. Como consequência desse desacoplamento, cria-se um cenário onde elétrons se adaptam nas posições definidas pelos núcleos. Portanto, a equação 3.1, que antes dependia das coordenadas dos núcleos e dos elétrons, torna-se uma equação paramétrica nas coordenadas dos elétrons dependentes da posição do núcleo ponto a ponto.

Uma outra aproximação bastante utilizada para resolver a equação de Schrödinger independente do tempo assume que a interação entre os elétrons do sistema pode ser substituída por um potencial médio de interação onde um elétron interage com uma distribuição de carga gerada pelos elétrons restantes, e a função de onda do sistema pode ser representada como um produto de funções de um elétron antissimetrizada representada pelo determinante de Slater. Esta é a chamada aproximação da partícula independente, que definem as bases do método Hartree-Fock (HF)<sup>35</sup>.

A aplicação do método HF para moléculas ganhou um grande impulso quando, em 1951 Roothaan propôs a representação dos orbitais em termos de uma expansão em funções de base, originando o chamado método de Hartree-Fock-Roothaan. O objetivo deste método é obter, a partir de um conjunto de orbitais de um elétron, os orbitais moleculares do sistema minimizando a energia de uma dada configuração eletrônica

sujeita, normalmente, a condições de ortonormalidade entre os orbitais.

O teorema variacional é utilizado na expressão do valor esperado do hamiltoniano, fornecendo as equações HF. A resolução destas equações é feita de maneira iterativa até que os autovalores e autovetores tenham atingido a auto-consistência, ou seja, quando os orbitais gerados forem idênticos aos utilizados para formar o hamiltoniano que os gerou.

Contudo, a função de onda HF não é uma representação completa da estrutura eletrônica de uma molécula, devido ao fato da interação, ou correlação efetiva, entre elétrons não acontecer por meio de um potencial médio. Este fato impõe um limite máximo de aprimoramento no valor da energia, mesmo que se utilize a melhor função de onda HF. A diferença entre a energia exata e a energia HF é conhecida como a energia de correlação. Apesar de pequena, a inclusão da mesma é de fundamental importância para descrever adequadamente diversas propriedades moleculares. Nesta dissertação, foram utilizados o método da Teoria do Funcional da Densidade (DFT)<sup>36,37</sup> e a Teoria de Perturbação de Muitos Corpos de Moller-Plesset em Segunda Ordem (MP2)<sup>38</sup>. A vantagem do método MP2 está no fato deste recuperar cerca de 75% da energia de correlação<sup>39,40,41</sup>, porém, tem uma demanda computacional que cresce com  $N^5$ , onde N é o número de funções de base. Por sua vez, o método DFT exige uma menor demanda computacional, onde, em geral, nestes métodos a demanda computacional aumenta na ordem de  $N^3 - N^4$ , portanto, bem próximo do método HF que apresenta uma dependência com  $N^4$ .

Tendo em vista que o método DFT, além da demanda computacional comparável com a do método HF, as implementações atuais incluem, de maneira efetiva, os efeitos da correlação eletrônica, e são comparáveis ao método MP2 para várias propriedades, particularmente, estrutura molecular e freqüências vibracionais.

As funções de base empregadas nesta dissertação foram as funções tipo gaussianas, que descrevem, de maneira satisfatória, as propriedades atômicas, perto ou longe do núcleo.

Este conjunto de base usado pode ser dividido em três tipos:

- Bases mínimas, que são aquelas que incluem um número mínimo de funções de base requerido para acomodar os elétrons do sistema,

- Duplo-zeta, que são aquelas cujo número de funções é o dobro das utilizadas na base mínima, e,
- Estendidos, que são todos com mais do que o dobro das funções de base mínima.

Estas últimas funções são, em geral, essenciais para a descrição das propriedades eletrônicas moleculares.

Além destas, funções difusas e de polarização podem ser utilizadas para uma correta descrição das propriedades moleculares.

Em particular, utilizou-se o conjunto de funções de base desenvolvidas pelo grupo do Pople, presentes no programa Gaussian 98 (G98)<sup>42</sup>, onde as regiões de caroço e de valência são descritas por um conjunto composto por um número diferente de funções gaussianas contraídas, cujos valores dos coeficientes foram previamente otimizados para minimizar as energias atômicas HF.

Neste trabalho, serão utilizados os métodos semi-empírico, *ab initio* e DFT, para se avaliar a influência do método na descrição de parâmetros energéticos, geométricos e eletrônicos do sistema investigado. Já para o estudo da superfície do potencial eletrostático foram utilizados diferentes métodos na previsão do potencial eletrostático. Gradientes analíticos e segundas derivadas estão disponíveis para este nível de teoria, portanto, pode-se explorar as superfícies de energia potencial nestes níveis.

### 3.2. Otimização da geometria molecular

O uso mais comum e de maior sucesso da mecânica quântica em sistemas moleculares é a reprodução e a previsão da estrutura de moléculas<sup>41,43</sup>, particularmente dos estados estacionários. Em muitos casos, cálculos teóricos de otimização de geometria molecular conseguem reproduzir dados de cristalografia com grande exatidão, onde os comprimentos de ligação e ângulos de torção podem, eventualmente, apresentar erros inferiores a 0,01Å e 1,0°, respectivamente, apesar das diferenças entre a fase condensada e aquela na qual os cálculos teóricos são realizados. Desta forma, a escolha adequada do conjunto de bases (em geral um conjunto de bases grande, especialmente aqueles estendidos e com funções de polarização) e com a inclusão de

correlação eletrônica nos cálculos de otimização de geometria é de fundamental importância para reprodução dos dados experimentais.

Neste trabalho os cálculos de otimização da geometria molecular foram feitos com o programa G98, utilizando-se os quatro critérios de convergência internos do programa.

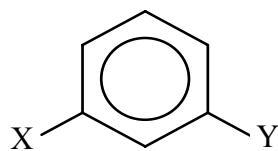
Dividiu-se o processo de otimização apenas da geometria molecular em três fases: i) otimização da geometria da molécula **5h** (molécula não substituída apresentada na Figura 1.2, página 8), ii) otimização da geometria molecular do substituinte (molécula (X,Y)benzeno, Figura 3.1) para formação das estruturas as quais chamaremos de semi-otimizada e, iii) otimização da geometria da molécula completa, ou seja, otimização da geometria das pirimidinas substituídas.

### *3.2.1. Otimização da geometria molecular da pirimidina não substituída (**5h**)*

Inicialmente, fez-se a otimização da geometria molecular apenas da molécula não substituída, a qual chamaremos de molécula padrão ou de referência (Figura 1.2) com os métodos HF e DFT com os funcionais B3LYP e BP86, em ambos os casos com as funções de base 6-31G, 6-31G\*\*, 6-31++G, 6-31+G\*\*, 6-31++G\*\*, com o método MP2 e funções de base 6-31G, 6-31G\*\*, e com o método semi-empírico AM1<sup>44</sup>, todos implementados no programa Gaussian 98.

### *3.2.2. Otimização da geometria molecular dos substituintes*

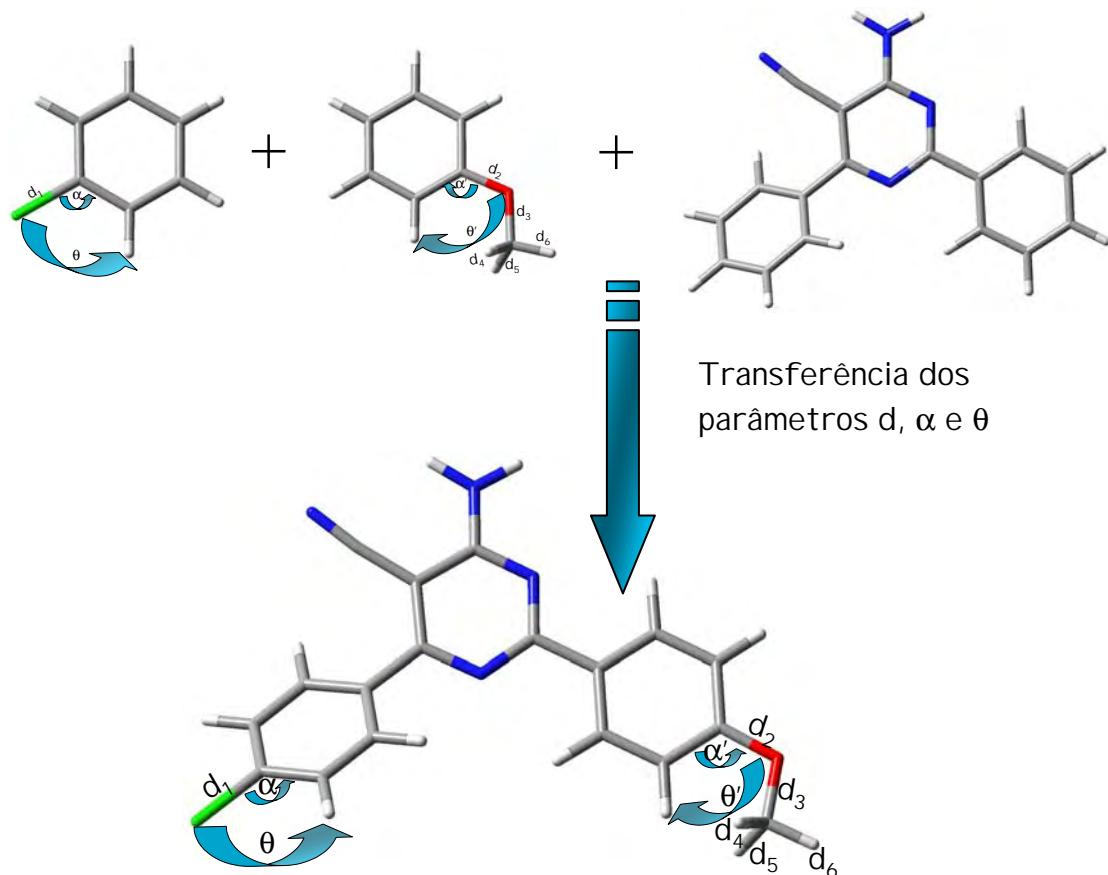
Otimizou-se a geometria dos substituintes a partir da molécula do benzeno substituído, (X,Y)-benzeno (Figura 3.1). Utilizou-se, nesta etapa, apenas o método BP86/6-31G\*\*.



$$(X, Y) = (H, H), (Cl, H), (F, H), (NH_2, H), (NO_2, H), (CF_3, H) \text{ e } (OCH_3, H) \text{ e } (Cl, Cl)$$

**Figura 3.1 :** Moléculas de benzeno substituídos, calculados com o método BP86/6-31G\*\*.

A geometria dos substituintes Cl, F, NH<sub>2</sub>, NO<sub>2</sub>, CF<sub>3</sub> e OCH<sub>3</sub> foi introduzida na molécula **5h** (previamente otimizada com o método BP86/6-31G\*\*), através da transferência dos parâmetros de comprimento de ligação (d), ângulo de ligação ( $\alpha$ ) e ângulo diédrico ( $\theta$ ) dos substituintes otimizados na molécula do (X,Y) benzeno, conforme pode ser visto no esquema da Figura 3.2. Essas moléculas foram chamadas de pirimidinas semi-otimizadas.



**Figura 3.2:** Esquema da transferência dos parâmetros geométricos dos substituintes no sistema benzeno substituído para a pirimidina não substituída otimizada, gerando as pirimidinas substituídas semi-otimizadas.

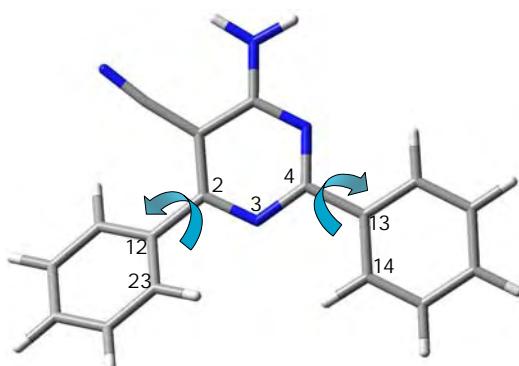
### 3.2.3. Otimização da geometria molecular total das pirimidinas substituídas

Também foi realizada a otimização completa da geometria molecular das pirimidinas substituídas com os métodos DFT-BP86/6-31G\*\* e semi-empírico AM1 conforme será explicado na seção 4.3 dos resultados. Neste caso, utilizaram-se as

estruturas semi-otimizadas (descritas anteriormente) como geometria inicial para o cálculo de otimização.

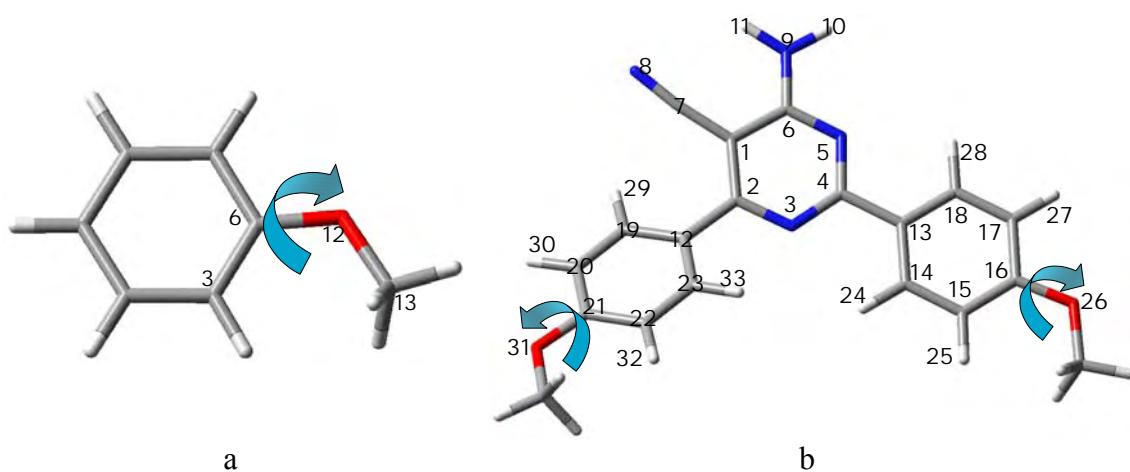
### 3.3. Estudo Conformatacional

Inicialmente, um estudo conformacional foi realizado através do cálculo AM1 da barreira de energia para os dois anéis benzênicos da molécula **5h** (Figura 3.3).



**Figura 3.3:** Numeração dos átomos envolvidos na barreira de rotação do anel benzênico na molécula **5h**.

Em seguida, foi realizado um outro estudo conformacional, AM1, com relação ao grupo metoxi tanto na molécula do metoxibenzeno (Figura 3.4a) quanto no anel benzênico das pirimidinas contendo este substituinte (Figura 3.4b).



**Figura 3.4:** Esquema da rotação do grupo metoxi no: a) metoxibenzeno e b) nas pirimidinas substituídas.

Para a análise da barreira de energia do anel benzênico (ângulos diedricos C23C12C2N3 e C14C13C4N3), fez-se o cálculo da energia da molécula variando-se estes dois ângulos diédricos simultaneamente (Figura 3.3) no programa G98, com o método AM1 e com um passo de 10 graus e otimização do restante da geometria molecular.

Em seguida realizou-se a otimização da geometria utilizando-se o método BP86/6-31G\*\* a partir da conformação de mínimo AM1, devido ao elevado tempo de cálculo requerido pelo método DFT.

Ainda, realizou-se o estudo da barreira de rotação do grupo metoxi com ambos os métodos AM1 e BP86/6-31G\*\*. Nesta etapa utilizou-se a opção *scan* onde se avaliou a barreira de energia do metoxibenzeno e da molécula 5f, ambos com um passo de 10 graus, ou seja, fixando-se o grupo metoxi de 10 em 10 graus e calculando-se as novas geometrias de equilíbrio.

### 3.4. Planejamento Fatorial de Dois Níveis

Quando uma função resposta depende de vários fatores, ou variáveis, torna-se importante como cada fator influencia, bem como se existem, ou não, efeitos sinergísticos ou antagonísticos entre os fatores e a resposta. A avaliação da contribuição de cada fator passa necessariamente pela execução de experimentos planejados. Uma técnica quimiométrica multivariada muito utilizada para esse fim é o Planejamento Fatorial de Dois Níveis (PFDN)<sup>45,46,47,48,49</sup>, que associa o efeito de cada fator (efeito principal) com a propriedade investigada, além de poder se quantificar os chamados efeitos de interação (novo efeito que surge a partir da interação de dois ou mais fatores).

Nesta técnica, inicialmente são definidos dois níveis para cada fator que pode influenciar uma resposta investigada. Em seguida, são realizados experimentos para todas as  $2^n$  possíveis combinações de n fatores e 2 níveis.

Neste trabalho, em particular, três fatores e dois níveis serão estudados: i) a presença ou ausência de funções de polarização (Pol), ii) a presença ou ausência de funções difusas (Dif) e iii) o uso ou não de correlação eletrônica, via a Teoria do Funcional Densidade, (Corr).

Na Tabela 3.1, está apresentada a matriz de codificação correspondente às  $2^3$  funções de onda resultantes dessas combinações. Para o cálculo dos efeitos principais e de interação, aplicam-se os sinais das colunas da Tabela 3.1 aos valores da propriedade investigada e então, se toma a média dos valores calculados no nível superior (+) menos a média dos resultados no nível inferior (-). Aqui vale ressaltar que os sinais dos efeitos de interação são obtidos diretamente pelo produto dos sinais dos efeitos principais, como pode ser visto na Tabela 3.1.

**Tabela 3.1:** Matriz de Codificação de um Planejamento Fatorial  $2^3$ .

Função de Onda	Efeitos							
	Principais			Interação				
	Dif*	Pol <sup>#</sup>	Corr <sup>‡</sup>	Dif-Pol	Dif-Corr	Pol-Corr	Dif-Pol-Corr	
HF/6-31G	-	-	-	+	+	+	-	
HF/6-31++G	+	-	-	-	-	+	+	
HF/6-31G**	-	+	-	-	+	-	+	
HF/6-31++G**	+	+	-	+	-	-	-	
DFT- BP86/6-31G	-	-	+	+	-	-	+	
DFT- BP866-31++G	+	-	+	-	+	-	-	
DFT- BP866-31G**	-	+	+	-	-	+	-	
DFT- BP86/6-31++G**	+	+	+	+	+	+	+	

\* Dif: função difusa (- → ausente, + → presente);

# Pol: função de polarização (- → ausente, + → presente);

‡ Corr: correlação electrônica (- → cálculos RHF, + → cálculos DFT-BP86/6-31G\*\*).

### 3.5. Análise de Componentes Principais

A técnica quimiométrica Análise de Componentes Principais (do inglês PCA)<sup>50,51,52</sup> é um dos tipos de análise multivariada mais utilizados uma vez que permite a redução dimensional do conjunto de dados com a perda mínima de informação, permitindo, em muitos casos, a observação de possíveis agrupamentos dos objetos que definem a estrutura do conjunto de dados.

O objetivo principal da análise de componentes principais é a obtenção de um

número pequeno de novos eixos, as Componentes Principais (PC), que na verdade são combinações lineares dos eixos originais e que, espera-se, retenham o máximo possível da informação original. Desta forma, um pequeno número de componentes pode ser usado em lugar das variáveis originais e, assim, espera-se que eventuais padrões nos dados possam ser melhores visualizados.

Vale a pena ressaltar que a PCA nem sempre reduz um grande número de variáveis originais a um pequeno número de variáveis transformadas. De fato, se as variáveis originais não são correlacionadas, a PCA não produz o efeito desejado. Os melhores resultados são obtidos quando as variáveis originais são positivamente ou negativamente correlacionadas. Nesse caso, por exemplo, as variáveis originais podem ser representadas adequadamente por poucas componentes principais, deixando clara a existência de grande redundância nas variáveis originais.

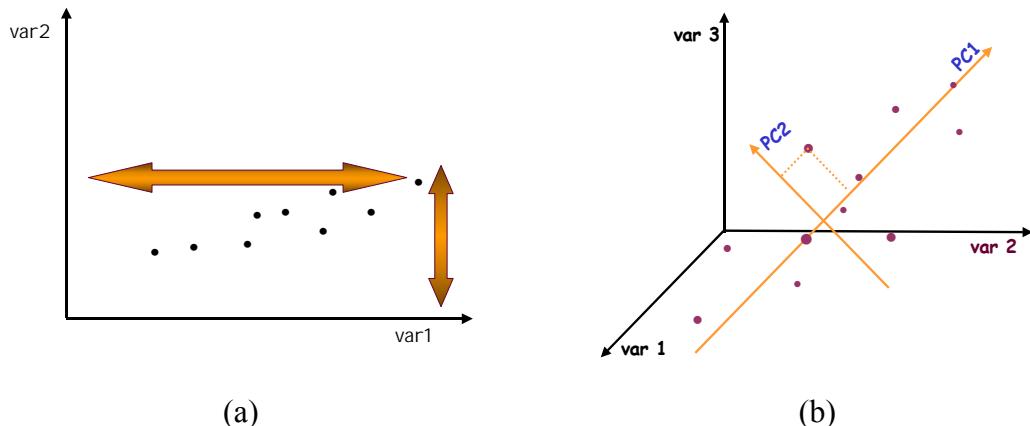
Na prática, isto é feito através da diagonalização da matriz de covariância ( $\mathbf{X}^t \mathbf{X}$ ) onde  $\mathbf{X}^t$  é a transposta da matriz original dos dados,  $\mathbf{X}$ , na qual os objetos são definidos nas linhas e as variáveis nas colunas. Os pesos, ou *loadings* são os autovetores obtidos desta diagonalização, e representam a contribuição com que cada eixo original entra na formação dos novos eixos, ou seja, de cada PC. Os autovalores representam a variância descrita em cada PC. Anteriormente a diagonalização, os dados da matriz  $\mathbf{X}$  são, em geral, pré-processados (centrados na média, ou seja, cada elemento na matriz de dados  $\mathbf{X}$  foi subtraído da média de todos os elementos pertencentes a coluna correspondente ou ainda auto-escalados).

A Figura 3.5 mostra que para um conjunto de dados definido num espaço de mais de uma variável, em geral, existe uma hierarquia de direções nas quais os pontos estão mais espalhados (a variável var 1). Então, os novos eixos (PC's) podem ser formados a partir da combinação linear dos eixos originais, de forma a preservar a maior quantidade de informação original (Figura 3.5b). A primeira componente principal (PC1) será a componente com maior variância, ou seja, com a maior parte da informação. As variâncias das demais PC's serão sempre menores que a variância das PC's anteriores e todas elas serão ortogonais entre si. O peso, ou contribuição, com que cada variável entra para a formação dos novos eixos pode ser vista no gráfico dos *loadings*. Já o gráfico dos escores mostra as coordenadas que cada objeto possui ao longo de cada componente principal. Este processo é descrito pela equação 3.2.

$$t_i = X_a l_i$$

Equação 3.2

onde  $t_i$  são os *scores* para  $i$ -ésima componente principal,  $X_a$  é a matriz obtida após o pré-processamento dos dados e  $l_i$  são os *loadings* (pesos) de cada variável em relação à  $i$ -ésima componente principal.



**Figura 3.5:** (a) Esquema hierárquico de variáveis e (b) projeção dos objetos sobre as PC's.

O estudo multivariado de PCA desta série de pirimidinas foi feito com os dados centrados na média (técnica de pré-processamento dos dados) uma vez que os parâmetros investigados em cada caso, as cargas atômicas calculadas com a partição CHelpG<sup>53</sup> e a superfície do potencial eletrostático, possuem, em cada caso, a mesma grandeza (unidade).

Nesta etapa foram adotados dois critérios para as análises de componentes principais das cargas atômicas e da SPE das pirimidinas substituídas: i) uma análise das pirimidinas com a geometria semi-otimizada e ii) uma análise das pirimidinas substituídas totalmente otimizadas. Nos dois casos foram feitas PCA's com as 22 estruturas, ou seja, as moléculas **5b**, **5c**, **5d** e **5h** incluindo as duas conformações mais estáveis das moléculas **5a**, **5e**, **5f**, **5l**, **5n**, **5p** e **5q** e as quatro conformações mais estáveis da molécula **5i** (ver Figura 4.7 dos resultados) e outra apenas com as moléculas mais estáveis (12 compostos ou objetos).

### 3.6. Regressão pelo método dos mínimos quadrados parciais (PLS)

O método dos mínimos quadrados parciais (do inglês PLS)<sup>21</sup> é uma ferramenta quimiométrica para a regressão multivariada a qual se mostra sensível as variáveis correlacionadas e pode lidar com várias funções respostas simultaneamente.

O PLS tem entre seus objetivos a determinação de um modelo que estabeleça uma relação entre as matrizes **X** (matriz das variáveis) e **Y** (matriz da resposta, neste caso, a resposta biológica, ou seja, a atividade antiinflamatória) estimando a correlação entre estas matrizes **X** e **Y**.

A principal suposição deste método é que a variável  $y$ , dada pela equação 3.3, possa ser descrita como combinação linear das informações da matriz **X** como na análise de componentes principais. A diferença, neste caso, é que as componentes principais são “otimizadas” para melhor descrever a relação entre a matriz **X** e o vetor  $y$  simultaneamente (e por isso chamadas variáveis latentes). Os coeficientes de regressão,  $\beta$ , nos diz quais descriptores são importantes na modelagem da resposta biológica.

$$y = \mathbf{X}\beta \quad \text{Equação 3.3}$$

Existem vários métodos de validação para determinação do número de componentes principais. Neste trabalho, utilizou-se o método de validação cruzada, o qual retira uma amostra  $i$  do conjunto de dados (isto é, um objeto) para construção de um modelo que será usado para prever a propriedade da amostra removida. Este procedimento foi usado para todos os objetos. O erro associado a cada previsão é dado pela diferença entre o valor previsto ( $\hat{Y}_i$ ) e o valor experimental da atividade biológica ( $Y_i$ ), somando-se em seguida os quadrados dos erros da previsão. O desvio padrão ( $d$ ) da validação cruzada é dado pela equação 3.4.

$$d = \sqrt{\frac{\sum_i (Y_i - \hat{Y}_i)^2}{n}} \quad \text{Equação 3.4}$$

Este modelo é repetido tantas vezes quantas sejam necessárias para que se obtenha o

menor erro possível.

Para o PLS da SPE desta série homóloga de pirimidinas substituídas, dividiu-se o conjunto de moléculas em quatro grupos diferentes de acordo com a conformação do grupo metoxi (ver Figura 3.4 para numeração dos átomos) conforme explicado a seguir:

- conformação “dentro” formada pelas moléculas que possuem o grupo metoxi voltado para o interior da molécula, isto é, com o diângulo C15C16O26C34 igual a zero grau, indicada pelo subscrito ‘d’ (moléculas **5a<sub>d</sub>**, **5e<sub>d</sub>**, **5f<sub>d</sub>**, **5i<sub>dd</sub>**, **5l<sub>d</sub>**, **5n<sub>d</sub>**, **5p<sub>d</sub>** e **5q<sub>d</sub>**).
- conformação “fora” formada pelas moléculas que possuem o grupo metoxi voltado para o lado externo da molécula, isto é, com o diângulo C15C16O26C34 igual a 180 graus, indicado pelo subscrito ‘f’ (moléculas **5a<sub>f</sub>**, **5e<sub>f</sub>**, **5f<sub>f</sub>**, **5i<sub>ff</sub>**, **5l<sub>f</sub>**, **5n<sub>f</sub>**, **5p<sub>f</sub>** e **5q<sub>f</sub>**).
- conformação “dentro/fora” formada pelas moléculas que possuem o grupo metoxi voltado para o lado externo da molécula, isto é, com o diângulo C15C16O26C34 igual a 180 graus e a molécula **5i<sub>df</sub>** possui o diângulo C22C21O31C38 igual a zero grau.
- conformação o “fora/dentro” formada pelas moléculas que possuem o grupo metoxi voltado para o interior da molécula, isto é, com o diângulo C15C16O26C34 igual a zero grau e a molécula **5i<sub>fd</sub>** possui o diângulo C22C21O31C38 igual a 180 graus.

### 3.7. Alinhamento molecular e determinação da caixa 3D.

O alinhamento molecular é, provavelmente, um dos aspectos mais importantes em análises de QSAR-3D, uma vez que, alinhar, pré-supõe um modelo de alinhamento. Vale ainda ressaltar que um alinhamento deficitário pode resultar em um modelo estatístico inadequado<sup>54</sup>.

O alinhamento molecular pode ser feito de várias maneiras diferentes, como por exemplo, através do esqueleto molecular (isto é, das coordenadas nucleares), se o conjunto de moléculas apresentarem esqueletos moleculares similares, ou através dos

grupos de ligação similares. O alinhamento pode ainda ser feito através do campo de interação 3D<sup>54</sup>.

Neste trabalho o alinhamento molecular foi feito de duas maneiras diferentes, i) baseado na posição dos núcleos utilizando o programa ALINHAMOL<sup>23</sup> e ii) através dos campos moleculares com o programa MIPSIM<sup>55</sup>. Em ambos programas, durante o processo de alinhamento, uma molécula é “fixada” (i. e., é usada como molécula de referência) enquanto as outras, chamadas de moléculas móveis, estão livres para alterarem sua posição relativa (através de rotação e translação) em relação àquela de referência, sem alteração interna da sua própria geometria. Ao final do processo de otimização do alinhamento, cada molécula móvel terá a melhor posição relativa em relação a molécula de referência.

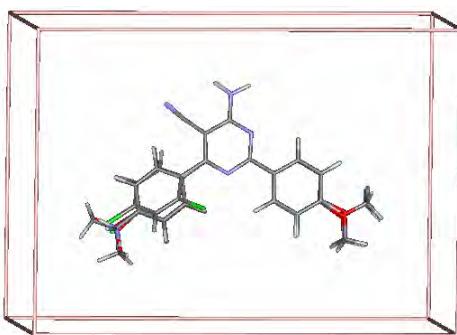
### 3.7.1. Alinhamento com o programa ALINHAMOL

O alinhamento molecular obtido com o programa ALINHAMOL está baseado na posição dos núcleos. Este programa permite a escolha da molécula de referência (previamente orientada na caixa de acordo com os 3 eixos cartesianos) e do conjunto de pares de átomos que serão utilizados no alinhamento de cada uma das moléculas móveis em relação à referência.

O ALINHAMOL também calcula o tamanho e a resolução da caixa 3D, de tal modo que todas as moléculas terão caixas do mesmo tamanho e, assim, o mesmo número de pontos. Este programa também calcula o grau de similaridade entre a molécula de referência e a alinhada, através do Raiz Média Quadrática (do inglês RMS, *Root Mean Square*) da posição espacial dos pares de átomos equivalentes que são usados para definir o alinhamento. Outra vantagem é a possibilidade de se gerar automaticamente os arquivos de entrada (“inputs”) necessários para o cálculo da Superfície do Potencial Eletrostático (SPE) no programa Gaussian 98, sendo possível a escolha do tamanho da caixa, a resolução, ou seja, a distância, ou espaçamento, entre cada ponto onde será calculado o Potencial Eletrostático (PE), e o método e a função de base que serão utilizados no cálculo da SPE .

Em particular, neste trabalho, o alinhamento foi feito com a molécula não substituída **5h** (padrão) como molécula de referência, utilizando-se apenas os átomos do anel central (C1C2N3C4N5C6C7N8N9H10H11, ver Figura 3.6), ou seja, os átomos do

grupo farmacofórico (isto é, grupo químico que produz a resposta biológica desejada).



**Figura 3.6:** Superposição das estruturas provenientes do alinhamento das pirimidinas **5a-q** em relação a molécula **5h** feita com o programa ALINHAMOL.

### 3.7.2. Alinhamento com o programa MIPSIM

O alinhamento feito pelo programa MIPSIM (exploração automática de similaridade biomolecular baseado nos potenciais de interação molecular<sup>55</sup>) é baseado numa avaliação de similaridade da série de compostos em relação aos valores de máximo e mínimo do potencial eletrostático molecular (MEP).

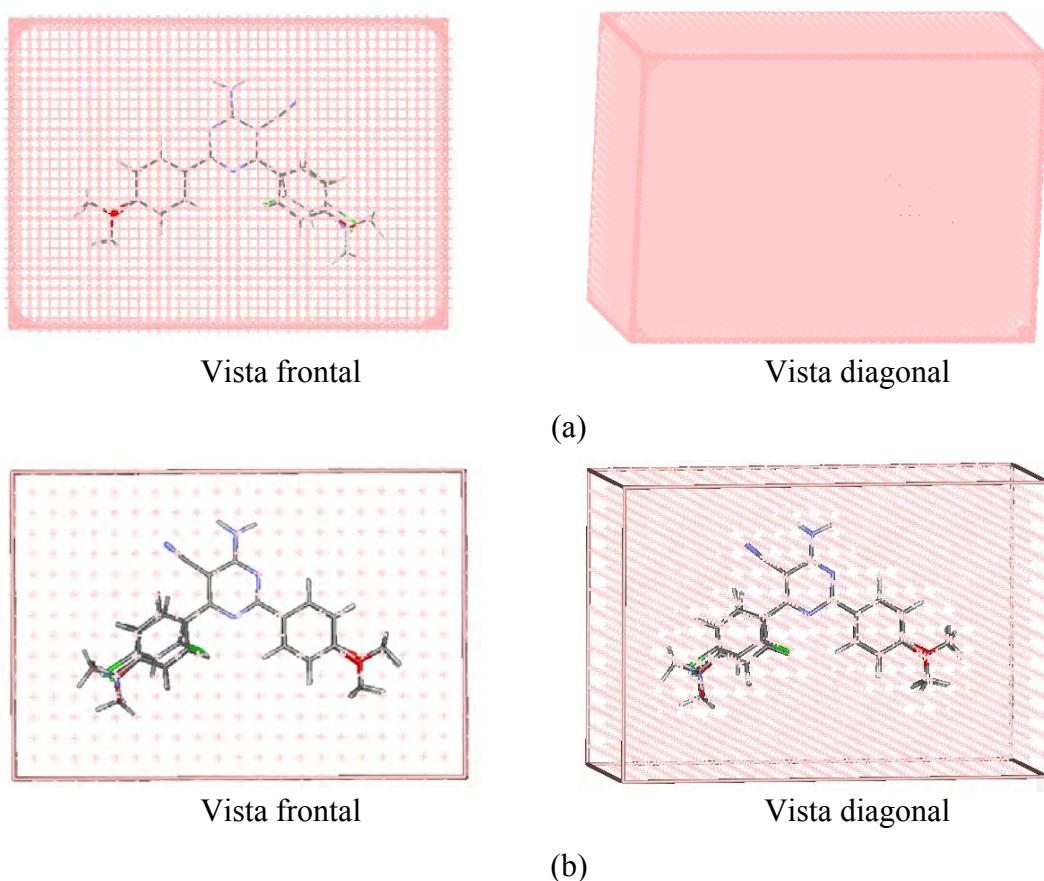
Neste trabalho foi utilizado o módulo MIPCOMP, presente neste pacote, responsável por fazer uma comparação das várias estruturas químicas no contexto da distribuição do Potencial de Interação Molecular (MIP) ao redor das moléculas dando o índice de correlação para mostrar a similaridade entre ambas distribuições. Este módulo computa o coeficiente de similaridade entre a distribuição MIP das moléculas estudadas.

Utilizou-se, inicialmente, como molécula de referência para o alinhamento a molécula não substituída **5h**. Posteriormente, analisou-se o efeito da molécula de referência na qualidade do alinhamento, utilizando então, as conformações mais estáveis das moléculas **5a** e **5i**. As diferenças podem ser vistas no item 4.6 dos resultados, página 66.

### 3.7.3. Determinação da caixa 3D

O tamanho da caixa, assim como a resolução entre os pontos da grade, estão associados a demanda computacional, bem como a introdução ou a exclusão de variáveis com pouca ou muita informação relevante para a posterior análise quimiométrica, respectivamente. Um espaçamento (resolução) adequado entre os pontos

também é importante para fornecer um maior detalhamento da propriedade eletrônica calculada ao redor das moléculas. Na Figura 3.7 pode-se observar as diferença no número de pontos onde será calculada a propriedade de interesse, neste caso, o potencial eletrostático, quando se utiliza uma resolução de 0,5 Å e 1,0 Å. Desta forma, torna-se importante a escolha de valores adequados para esses parâmetros. Os valores utilizados neste estudo estão apresentados na Tabela 3.2.



**Figura 3.7:** Esquema do número de pontos onde será calculado o potencial eletrostático para uma caixa com 2 Å além do extremo da molécula para cada uma das direções x, y e z, com resolução de (a) 0,5 Å e (b) 1,0 Å.

Nesta figura, pode-se perceber que a quantidade de pontos, onde será calculado o potencial eletrostático, está relacionado com a resolução da grade. Assim, um espaçamento maior ou menor entre os pontos, ou uma menor ou maior resolução, variará o número de pontos, tornando a malha menos ou mais densa, respectivamente. A Figura 3.7a mostra que com uma resolução de 0,5 Å, a malha torna-se bastante densa,

sendo maior o número de pontos e, assim, dificultando a visualização da molécula no interior da caixa. Já a Figura da 3.7b, mostra que um maior espaçamento entre os pontos da grade torna o malha menos densa, sendo menor o número de pontos onde será calculado o PE.

**Tabela 3.2:** Parâmetros e nomenclatura das caixas 3D.

SPE BP86/6-31G**			
Nomenclatura	Tamanho da caixa 3D ( $\text{\AA}$ ) <sup>†</sup>	Espaçamento entre os pontos ( $\text{\AA}$ )	Número de Pontos
Caixa 1	2	1,00	3667
Caixa 2	3	1,00	5469
Caixa 3	4	1,00	7719
Caixa 4	6	1,00	13755
Caixa 5	2	0,5	28623
Caixa 6	3	0,5	42807
Caixa 7	4	0,5	60543
Caixa 8	6	0,5	108207

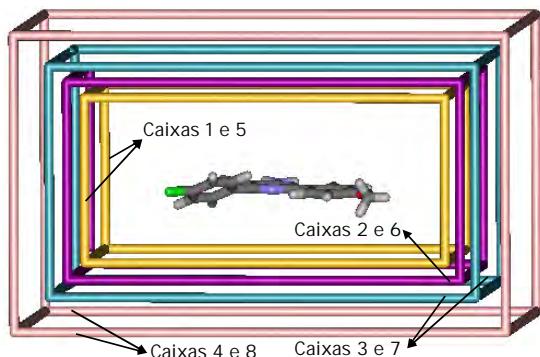
  

SPE AM1			
Nomenclatura	Tamanho da caixa 3D ( $\text{\AA}$ ) <sup>†</sup>	Espaçamento entre os pontos ( $\text{\AA}$ )	Número de Pontos
Caixa 9	2	1,00	3667
Caixa 10	3	1,00	5469
Caixa 11	4	1,00	7719
Caixa 12	6	1,00	13755
Caixa 13	2	0,5	28623
Caixa 14	3	0,5	42807
Caixa 15	4	0,5	60543
Caixa 16	6	0,5	108207

<sup>†</sup> O tamanho da grade refere-se a distância entre a borda da caixa e ponto mais extremo de todas as moléculas (nas direções x, y e z), identificado com o programa ALINHAMOL.

Neste ponto, vale a pena ressaltar que a molécula de referência é posicionada no

centro da caixa, conforme pode ser visto na Figura 3.8 e as demais são alinhadas em relação a ela, nesta posição.



**Figura 3.8:** Esquema da comparação do tamanho das caixas 3D e posicionamento da molécula.

Estas caixas foram utilizadas apenas para as moléculas alinhadas com o ALINHAMOL (seção 3.7.1).

### 3.8. Escolha da Partição de Cargas

A análise populacional de Mulliken é, provavelmente, o método mais popular de partição de carga entre os químicos. Este método baseia-se na teoria de orbitais moleculares (O.M.), onde um conjunto de orbitais moleculares é definido pela combinação linear dos orbitais atômicos, sendo estes definidos em termos das chamadas funções de bases, cujos coeficientes nos O.M. são obtidos pelo método de Hartree-Fock.

Os motivos para a ampla difusão do método populacional de Mulliken foram a intensa aplicação dos métodos que empregam a teoria do orbital molecular, a facilidade com a qual pode ser calculada e pelo fato da análise populacional ser obtida de forma direta, sem custo computacional adicional. Porém, este possui algumas arbitrariedades, como a divisão igualitária das populações de recobrimento para as contribuições atômicas, negligenciando, por exemplo, a separação de cargas entre dois átomos de diferentes eletronegatividades.

Desta forma, para se corrigir algumas imperfeições das distribuições de cargas que a análise populacional de Mulliken fornece, surgiram os métodos derivados do

potencial eletrostático, como por exemplo, o método Chelp<sup>56</sup> e o método CHelpG (*Charge from Eletrostatic Potencial Grid based*), que será utilizado neste trabalho dentre varias razões por ser bastante utilizado nos estudos de QSAR.

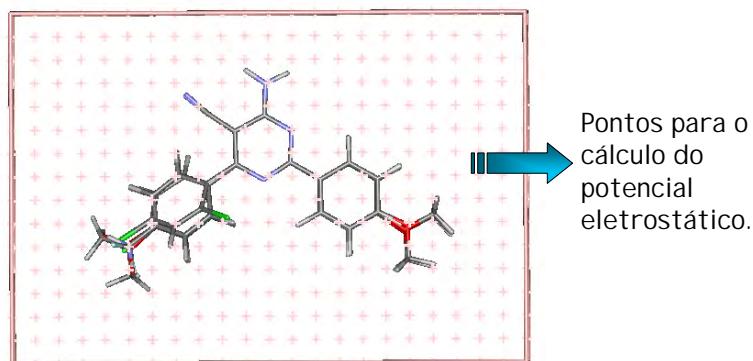
### 3.9. Cálculo da Superfície de Potencial Eletrostático e o programa CAMOL

A superfície de Potencial Eletrostático (SPE) foi calculada através da expressão da equação 3.5

$$\phi(\vec{r}) = \phi_{\text{nuc}}(\vec{r}) + \phi_{\text{ele}}(\vec{r}) = \sum_{\alpha=1}^{N_{\text{nuc}}} \frac{Z_{\alpha}e}{|\vec{r} - \vec{R}_{\alpha}|} + \sum_{j=1}^{N_{\text{ele}}} \int \frac{-e\Phi_j^* \Phi_j}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad \text{Equação 3.5,}$$

onde o termo nuclear depende do número de partículas,  $N_{\text{nuc}}$ , do número atômico,  $Z_{\alpha}$ , e da posição,  $\underline{\vec{R}_{\alpha}}$ , enquanto o termo eletrônico depende do número total de elétrons,  $N_{\text{ele}}$ , e da densidade eletrônica  $-e\Phi_j^* \Phi_j d\vec{r}'$  do  $j^{\text{th}}$  elétron no estado  $\Phi_j$ .

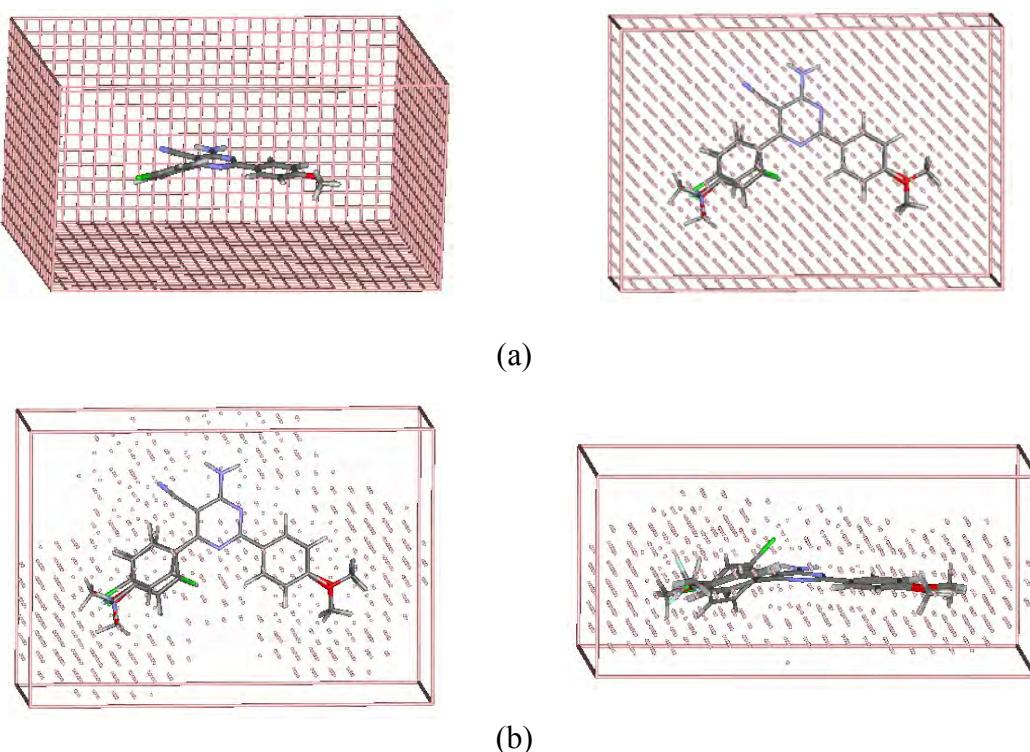
A propriedade eletrônica, neste caso o potencial eletrostático molecular, será calculado em todos os pontos da caixa tridimensional, conforme pode ser visto no esquema da Figura 3.9, gerando uma matriz de dados (disponível nos arquivos “cub”) que será interpretada com o auxílio do programa CAMOL e, posteriormente, através de ferramentas quimiométricas.



**Figura 3.9:** Esquema dos pontos onde será calculada a SPE.

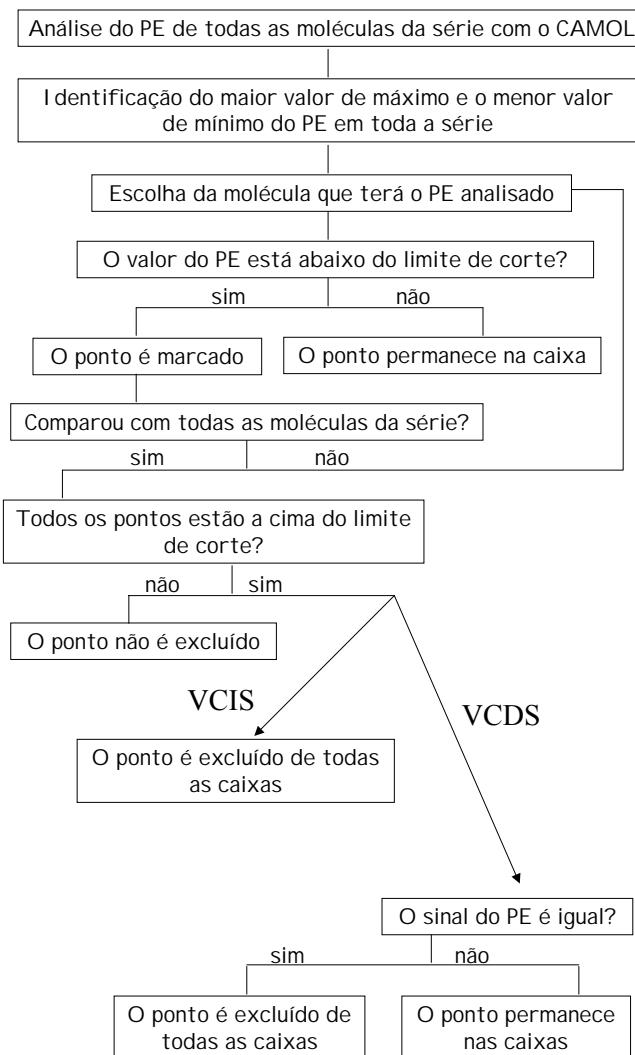
O programa CAMOL exclui os pontos da grade que apresentam uma

superposição com os átomos da molécula definidos pelo seus respectivos raios de van der Waals. É possível também a exclusão antecipada de pontos da SPE que apresentem valores inferiores a uma porcentagem do valor do máximo e/ou do mínimo do PE. Estes valores são definidos pelo usuário, denominados de valores de corte, e então a SPE deixa de ter a forma e tamanho da caixa utilizada e passa a ter a forma de um novo sólido determinada pelos pontos remanescentes na caixa 3D. Neste trabalho utilizaram-se os parâmetros da caixa apresentados na Tabela 3.2 para os métodos AM1 e BP86/6-31G\*\*. Uma figura esquemática desta nova superfície está apresentada na Figura 3.10b.



**Figura 3.10:** Número de pontos onde é calculado o potencial eletrostático com  
(a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.

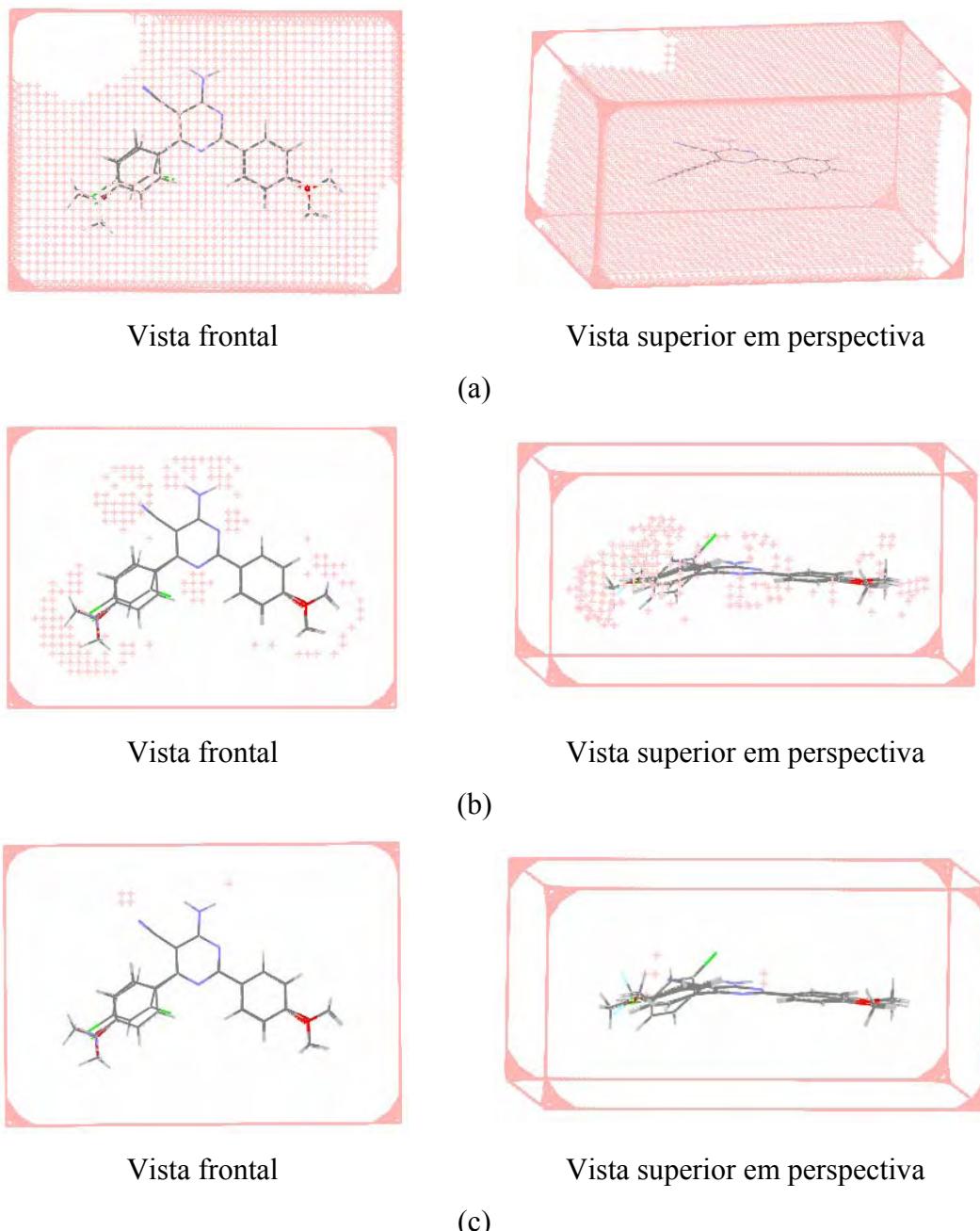
De uma forma mais detalhada, este novo sólido obtido a partir da exclusão de alguns pontos da superfície, é feita de duas maneiras, quais sejam, através da exclusão de pontos com valores de PE inferiores ao chamado valor de corte dependente do sinal (VCDS) do PE (+/-) e para valor de corte independente do sinal (VCIS). A Figura 3.11 mostra, de maneira esquemática, o procedimento adotado pelo CAMOL para exclusão dos pontos.



**Figura 3.11:** Esquema para exclusão dos pontos da SPE de acordo com o VCDS e o VCIS.

Percebe-se então, que o VCIS excluirá uma quantidade maior de pontos, uma vez que considera apenas o valor do potencial, isto é, desconsidera o sinal deste PE no dado ponto. Já VCDS considera estas duas restrições para exclusão dos pontos, sendo então mais rigorosa e, assim, excluindo uma quantidade menor de pontos. Desta forma, o valor adotado para o VCIS deverá ser sempre inferior a 100% (pois  $VCIS = 100\%$  ocasionaria a exclusão de todos os pontos da superfície) e menor que o do VCDS. Já o VCDS pode adotar valores grandes, chegando até a 100% do valor do máximo e do mínimo do PE, pois a restrição no sinal do PE no ponto garantirá a não exclusão de todos os pontos. Uma vez que há uma maior seleção dos pontos a serem excluídos, o

VCDS será responsável pela permanência dos pontos que tornam as superfícies mais dissimilares entre elas, ressaltando as regiões de interesse. As Figuras 3.12 e 3.13 mostram os pontos remanescentes após a utilização de três valores distintos do VCIS e VCDS, respectivamente.

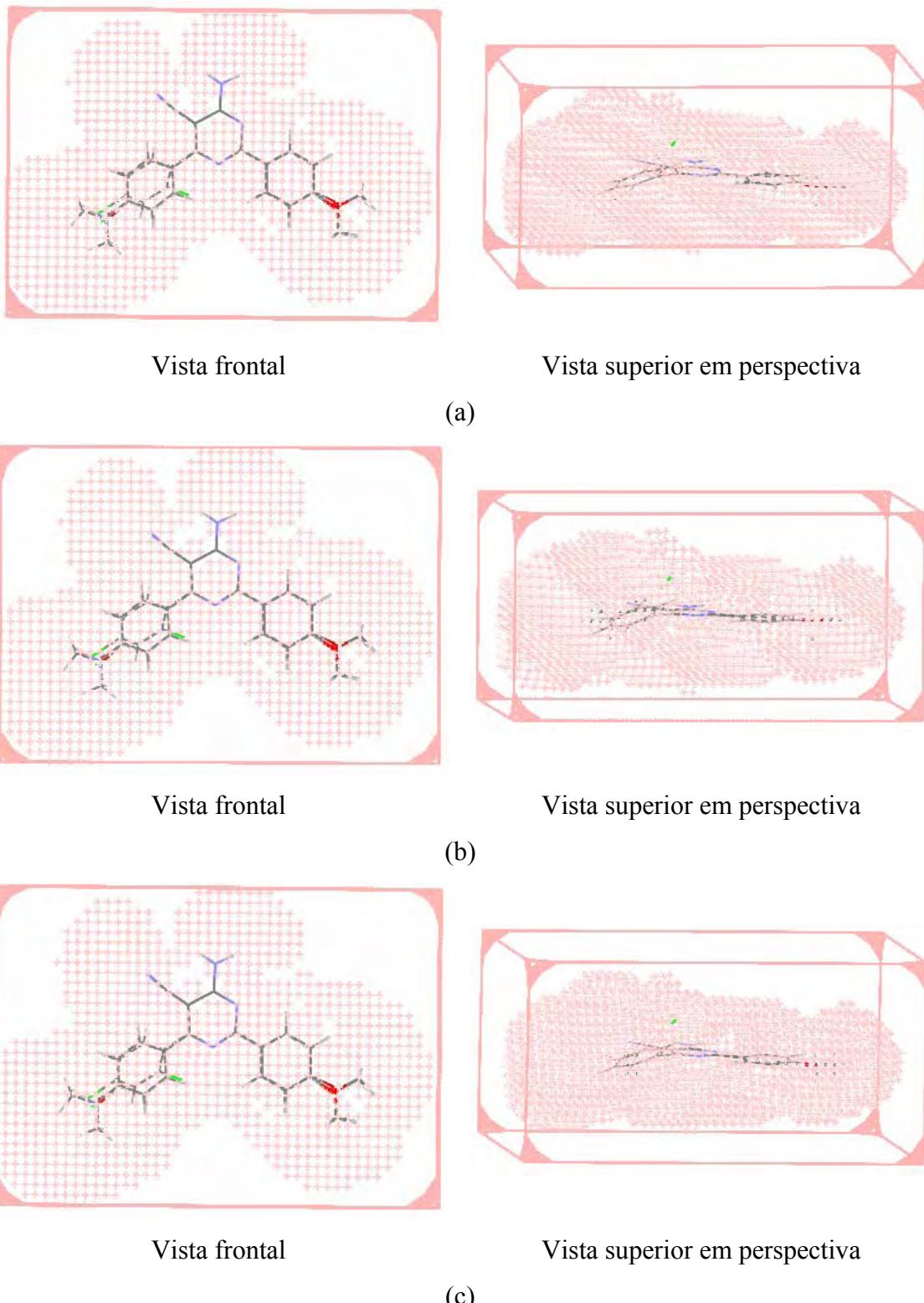


**Figura 3.12:** Pontos remanescentes da SPE após utilização dos valores de corte do CAMOL com VCDS = 100% e VCIS = (a) 0% (b) 50% e (c) 90%.

Pode-se perceber através da Figura 3.12 que há uma redução significativa no

número de pontos da SPE de acordo com o VCIS utilizado, onde um VCIS = 0% serão apenas excluídos os pontos que obedecam a restrição de VCDS = 100% (Figura 3.12a). Já VCIS = 50% há uma grande redução no número de variáveis (Figura 3.12b) e para VCIS = 90% praticamente todos os pontos da SPE foram excluídos (Figura 3.12c), devido a pouca restrição imposta no processo de identificação e exclusão dos pontos. Desta forma, nota-se a importância da escolha de um valor adequado para esta variável, para não haver exclusão de pontos importantes para a descrição das propriedades e, assim, não ser gerado um modelo de QSAR-3D com pouco poder de previsão. Por outro lado, nota-se também, que a utilização deste valor de corte possibilita a exclusão de pontos de pouca informação, ampliando a viabilidade do estudo de sistemas grandes.

A Figura 3.13 mostra a dependência no número de pontos restantes na SPE com o valor do VCDS utilizado. Percebe-se que há pouca variação na quantidade de pontos da SPE para os três valores utilizados (VCDS = 30%, 50% e 100%), sendo as Figuras 3.13a, b e c praticamente iguais. Logo, pode-se perceber o VCDS irá ser responsável pela exclusão de pontos mais próximos da molécula, fazendo que as diferenças mais importantes entre as moléculas analisadas sejam ressaltadas.



**Figura 3.13:** Pontos remanescentes da SPE após utilização dos valores de corte do CAMOL com VCIS = 20% e VCDS = (a) 30% (b) 50% e (c) 100%.

Vale a pena ressaltar que o emprego destes parâmetros foi feito com o propósito de saber o impacto dos mesmos sobre os resultados obtidos nos modelos de QSAR-3D. Para esse estudo manteve-se o “valor de corte dependente do sinal” (VCDS) constante em 100% e variou-se o “valor de corte independente do sinal” (VCIS) de 0 a 100% para se avaliar a influência destes parâmetros na exclusão dos pontos. Após a identificação de uma mudança no padrão da PCA realizada com esses dados, variou-se o “valor de corte dependente do sinal” (VCDS) de 30% a 100%. Esses resultados podem ser vistos na seção 4.5, página 68.

Para a avaliação da SPE das pirimidinas foram testados cálculos semi-empíricos AM1 e DFT, BP86/6-31G\*\* utilizando, em ambos os casos, a geometria otimizada pelo método BP86/6-31G\*\* para que, desta forma, as diferenças apresentadas pela SPE fossem resultados dos métodos/funções de onda utilizados e não devido, também, as diferenças das geometrias.

O cálculo da SPE da molécula do ligante, neste caso as pirimidinas substituídas, foi realizado, através do programa G98 (arquivos “cub”) para diferentes caixas 3D, como descrito na seção 3.7.3, página 27.

### 3.10. Teste da Atividade Antiinflamatória

A atividade antiinflamatória desta série de pirimidinas foi analisada pelo modelo da indução do edema na pata de rato, método bastante usado na literatura<sup>57</sup>. A ação inflamatória foi induzida pela injeção intraepidérmica da carragenina 0,1% de concentração salina (inflamação não específica) na pata posterior direita de camundongos, em grupos de dez animais. Como o edema é sensível a drogas antiinflamatórias não esteroidais, NSAIDs, a carragenina vem sendo aceita e usada como agente de inflamação para estudo de novas drogas com este tipo de atividade<sup>58</sup>. Estes modelos prevêem agentes antiinflamatórios como resultado da inibição da prostaglandina<sup>59</sup>.

O processo para determinação da atividade antiinflamatória inicia-se após 30 minutos da indução da inflamação, com a administração dos compostos por via intraperitoneal no camundongo. O antiinflamatório padrão utilizado é o ácido acetil

salicico (AAS) que, apesar da descoberta de várias novas drogas, continua sendo o analgésico, antipirético e antiinflamatório mais prescrito na medicina e é um padrão para a avaliação de novas drogas em ensaios farmacológicos. Foram utilizadas as mesmas doses de AAS e dos compostos testados e sob a mesma metodologia. Após 4h da administração da droga, o animal é sacrificado por inalação com éter e suas patas posteriores são amputadas e pesadas e, então, o peso é comparado com a média de peso das patas do grupo teste com aquela obtida com o grupo que recebeu a solução salina, sendo então avaliada a ação antiinflamatória.

### 3.11. Infra-estrutura

Para este trabalho, foram utilizados microcomputadores pessoal, de uso comum do Laboratório de Química Teórica e Computacional (LQTC) do Departamento de Química Fundamental da UFPE, do Laboratório de Química Teórica Medicinal (LQTM) do Departamento de Ciências Farmacêuticas da UFPE, e do Centro Nacional de Processamento de Alto Desempenho em São Paulo (CENAPAD-SP).

# CAPÍTULO 4

## RESULTADOS

E

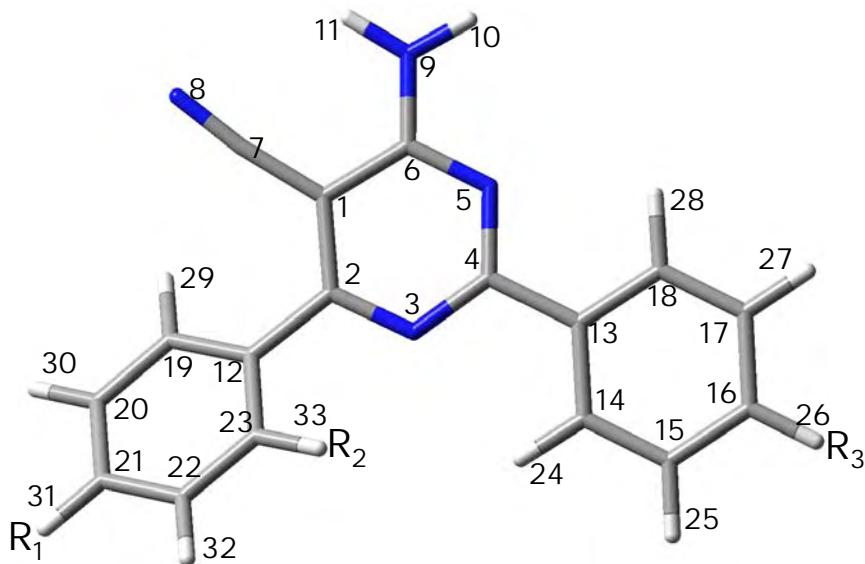
## DISCUSSÕES

A determinação de um modelo de QSAR-3D requer previamente várias etapas, como, por exemplo, a escolha dos parâmetros a serem analisados, os programas que serão utilizados, e a metodologia de cálculo que será empregada nos cálculos das variáveis, entre outras. Então, devido ao interesse de se estabelecer um modelo de QSAR-3D para a série homóloga de pirimidinas substituídas, buscou-se, inicialmente, avaliar qual a metodologia de cálculo que será utilizada neste trabalho. Os resultados obtidos estão apresentados a seguir.

#### 4.1. Estudo da Estrutura Padrão: a Molécula 5h

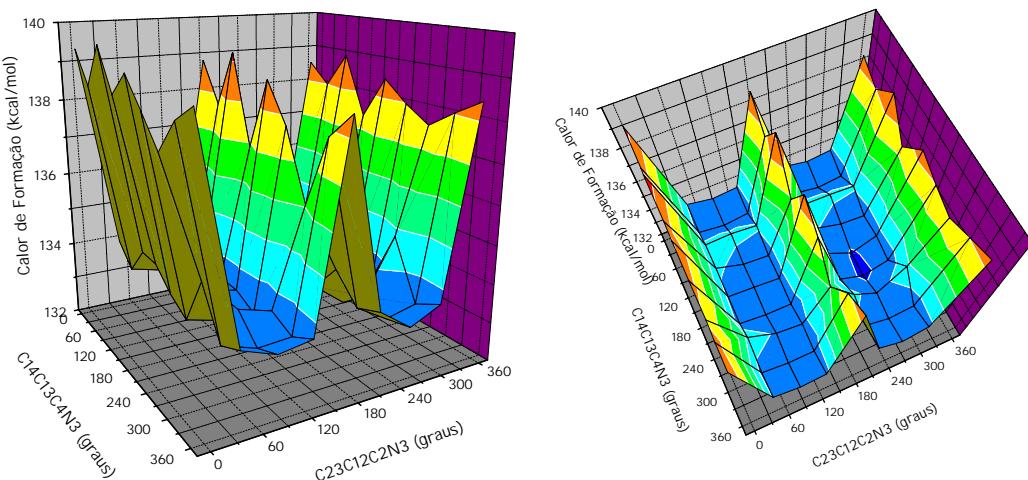
##### 4.1.1. Análise conformacional da pirimidina 5h

Inicialmente, fez-se uma análise conformacional, com o método AM1, dos anéis benzênicos presentes na molécula **5h** ( $R_1 = R_2 = R_3 = H$ ), ligados aos átomos C2 e C4, (Figura 4.1).



**Figura 4.1:** Molécula Padrão. Pirimidina substituída com  $R_1 = R_2 = R_3 = H$ .

O gráfico da Figura 4.2 mostra a superfície de energia calculada com o método AM1, gerada pela rotação destes dois anéis benzênicos, ou seja, pela variação dos ângulos diédricos C23C12C2N3 e C14C13C4N3.



**Figura 4.2:** Superfície de energia da molécula 5h calculada pelo método AM1 através da variação simultânea dos ângulos diédricos C23C12C2N3 e C14C13C4N3.

Percebe-se através do gráfico da Figura 4.2 que existem pontos de máximo (identificados pelas cores próximas à vermelha) e de mínimo (identificados pelas cores próximas ao azul) na superfície. Pode-se perceber também que a rotação do diedro C23C12C2N3 não provoca uma variação, significativa, no calor de formação da molécula sendo a maioria dos pontos da superfície, pontos de mínimo. Porém, a rotação do diedro C14C13C4N3 provoca grandes variações na energia do sistema, chegando a uma diferença de energia de até aproximadamente 30 kcal.mol<sup>-1</sup>. Desta forma, determinou-se que a conformação mais provável do anel benzênico terá com o diedro C23C12C2N3 próximo a 50° e o diedro C14C13C4N3 próximo a 40°. Posteriormente, utilizou-se o método BP86/6-31G\*\* para uma avaliação comparativa com os resultados AM1. Os resultados obtidos pelo método BP86/6-31G\*\* são diferentes daqueles obtidos pelo AM1 com os diedros C23C12C2N3 e C14C13C4N3 iguais a 28,2° e 3,2°, respectivamente. Como uma maior ou menor coplanaridade destes anéis para com o da porção farmacofórica (anel pirimidínico) pode estar associado com o efeito dos substituintes (via efeito eletrônico) para a mudança na atividade antiinflamatória destes compostos, decidiu-se ampliar a análise da dependência destes ângulos com o nível de cálculo e funções de base utilizadas. Para este fim, utilizou-se a estratégia de um Planejamento Fatorial de Dois Níveis (PFDN), onde foi avaliado o efeito da introdução de funções difusas, de funções de polarização e da correlação eletrônica usando o método DFT com dois tipos de funcionais (B3LYP e BP86). A título comparativo, dois

cálculos MP2 foram realizados. Na Tabela 4.1 apresentam-se os valores dos ângulos diédricos C23C12C2N3 e C14C13C4N3, de todos os cálculos realizados. Já na Tabela 4.2 são mostrados os efeitos principais e de interação entre os fatores (Pol, Dif e Corr, ver seção 3.4 da metodologia).

**Tabela 4.1:** Ângulos diédricos C23C12C2N3 e C14C13C4N3, em graus, da pirimidina **5h** otimizada com os métodos AM1, HF, BP86 e B3LYP e funções de base 6-31G, 6-31++G, 6-31G\*\*, 6-31+G\*\* e 6-31++G\*\*, e MP2 com funções de base /6-31G, 6-31G\*\*.

Função de Onda	Diédro C23C12C2N3 (graus)	Diédro C14C13C4N3 (graus)
<b>HF/6-31G</b>	36,5	2,4
<b>HF/6-31++G</b>	39,2	2,8
<b>HF/6-31G**</b>	38,5	3,3
<b>HF/6-31+G**</b>	40,1	3,9
<b>HF/6-31++G**</b>	40,0	3,7
<b>BP86/6-31G</b>	26,4	2,7
<b>BP86/6-31++G</b>	30,1	2,5
<b>BP86/6-31G**</b>	28,2	3,2
<b>BP86/6-31+G**</b>	31,0	3,6
<b>BP86/6-31++G**</b>	31,0	3,5
<b>B3LYP/6-31G</b>	29,2	2,1
<b>B3LYP/6-31++G</b>	32,5	2,4
<b>B3LYP/6-31G**</b>	31,0	3,1
<b>B3LYP/6-31+G**</b>	33,7	3,3
<b>B3LYP/6-31++G**</b>	33,6	3,3
<b>MP2/6-31G</b>	39,8	13,7
<b>MP2/6-31G**</b>	37,3	9,0
<b>AM1</b>	49,7	38,9

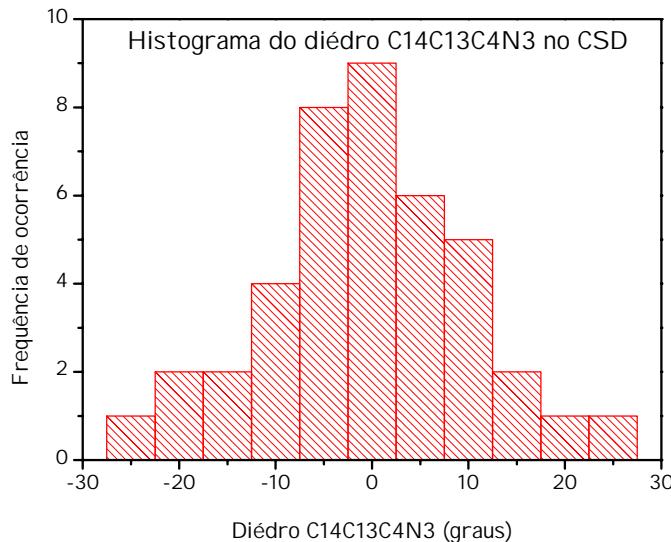
**Tabela 4.2:** Efeitos principais e de interação de um Planejamento Fatorial de DoisNíveis com três fatores,  $2^3$ .

Efeito	HF e BP86		HF e B3LYP	
	C23C12C2N3	C14C13C4N3	C23C12C2N3	C14C13C4N3
	(graus)	(graus)	(graus)	(graus)
<b>Dif</b>	2,68	0,23	2,5	0,33
<b>Pol</b>	1,38	0,83	1,40	0,93
<b>Corr</b>	-9,63	-0,08	-7,00	-0,33
<b>Dif- Pol</b>	-0,53	0,13	-0,45	0,03
<b>Dif-Corr</b>	0,58	-0,18	0,45	-0,08
<b>Pol-Corr</b>	-0,03	-0,08	0,05	0,03
<b>Pol-Dif-Corr</b>	0,08	0,13	0,10	-0,03
<b>Média</b>	33,74	3,01	35,08	2,89

Dos resultados da Tabela 4.2 nota-se que o diedro C23C12C2N3 possui um valor médio próximo a  $34^\circ$  e, que os efeitos principais são maiores que os de interação. Dentre os efeitos principais a introdução de correlação eletrônica tende a diminuir, em média, o valor deste diedro em  $10^\circ$  e  $7^\circ$  quando os funcionais BP86 e B3LYP são utilizados, respectivamente. Já o diângulo C14C13C4N3 é bem menor, em torno de  $3^\circ$  para ambos os funcionais. Neste caso, os efeitos principais são também, em geral, maiores que os efeitos de interação, destacando-se o efeito de polarização que tende a aumentar este diângulo em apenas 0,8 graus. Os demais efeitos principais e de interação, apresentam pouca influência na média deste diedro. Logo, estes resultados indicam importância da introdução de correlação eletrônica na otimização da geometria.

Em vista a fortalecer o estudo conformacional, foi realizado uma busca por moléculas semelhantes às pirimidinas substituídas no banco de dados cristalográficos *Cambridge Structural Database*<sup>60,61,62</sup> (CSD). Nesta pesquisa, procurou-se por estruturas que continham o anel central e ambientes químicos similares nas posições C2 e C4 das pirimidina. Foram localizadas 21 arquivos contendo o diângulo C14C13C4N3 e apenas 1 estrutura contendo o diângulo C23C12C2N3. As estruturas localizadas podem ser vistas na Figura A1 e A2, páginas 111 e 116, respectivamente. Na Figura 4.3

apresenta-se um histograma da ocorrência do diedro C14C13C4N3 encontrados nesta pesquisa.

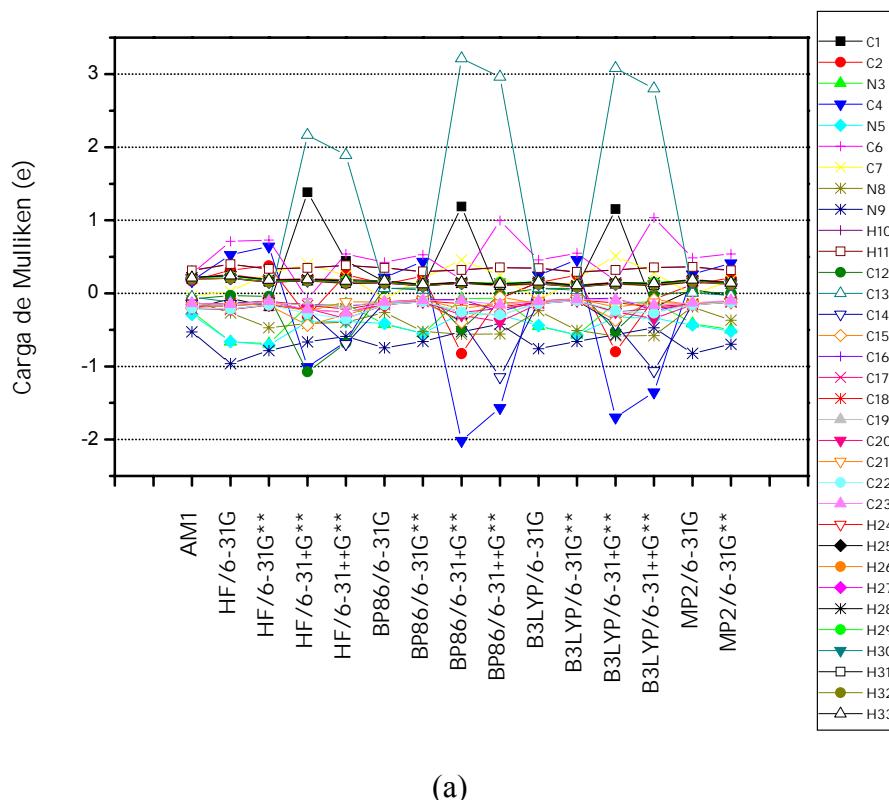


**Figura 4.3:** Histograma da freqüência de ocorrência do diédro C14C13C4N3.

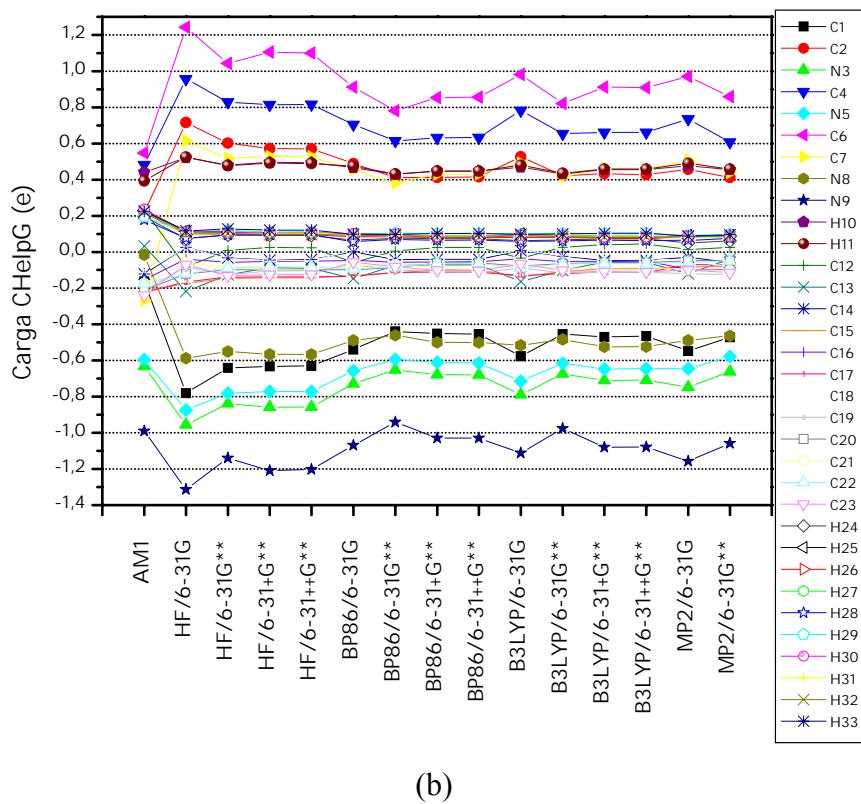
Esta figura mostra que o diédr C14C13C4N3 surge em valores de  $-25,87^\circ$  a  $25^\circ$ . Entretanto, a maioria dos valores deste diedro encontram-se próximos a  $0^\circ$ . Com relação ao arquivo da estrutura com o diédro C22C12C2N3<sup>63</sup> (contendo duas estruturas cristalográficas similares da mesma molécula) os valores do diedro valem  $45,8^\circ$  e  $30,5^\circ$ . Apesar de se saber que em estado sólido a geometria molecular pode ser relativamente diferente daquela em fase gasosa (a que mais se aproxima destes cálculos) os valores calculados para os diédros C14C13C4N3 mostram-se bastante próximos destes experimentais, em especial para os resultados obtidos com o método DFT.

Como se pretende realizar um estudo de QSAR-3D, utilizando, num primeiro instante, a sonda eletrostática, é interessante se avaliar como o nível de cálculo e o conjunto de funções de base influenciam a SPE. Entretanto, foi observado recentemente<sup>25</sup>, que a análise do PE para um número elevado de pontos na grade (como também é o caso para os sistemas em investigação nessa dissertação) em função de diferentes níveis de cálculo e funções de base utilizadas pode não ser uma tarefa trivial. Assim, uma forma alternativa para avaliar a distribuição eletrônica é através da análise da dependência da carga atômica com as características da função de onda utilizada. Neste sentido, foi realizada uma análise das cargas atômicas, dessa molécula,

calculadas com as partições de Mulliken e CHelpG e uma PCA das cargas atômicas (CHelpG) conforme explicado na seção 3.5 da metodologia, página 21, em todos os níveis de cálculo estudado, ou seja, HF e DFT (funcionais BP86 e B3LYP) ambos com funções de onda 6-31G, 6-31G\*\*, 6-31+G\*\* e 6-31++G\*\*, MP2 com as funções de onda 6-31G e 6-31G\*\* e AM1. Os gráficos das cargas em função da metodologia de cálculo e os resultados da PCA para as duas primeiras componentes principais podem ser vistos nas Figuras 4.4 e 4.5, respectivamente.



(a)

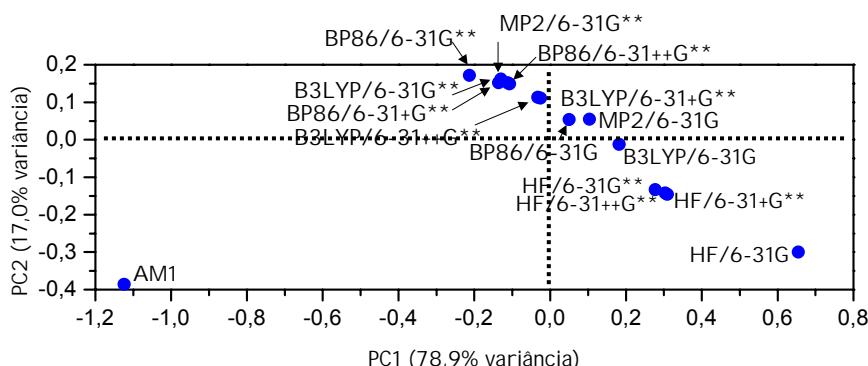


**Figura 4.4:** Gráfico das cargas de (a) Mulliken e (b) CHelpG da molécula **5h** em função do método, calculadas com os métodos AM1, HF e DFT (B3LYP e BP86) com funções de base 6-31G, 6-31G\*\*, 6-31+G\*\* e 6-31++G\*\* e MP2 com funções de base 6-31G e 6-31G\*\*.

Pode-se perceber através da Figura 4.4a que há grande variação nas cargas de Mulliken da molécula **5h** calculada pelos diferentes métodos, já o gráfico das cargas CHelpG, Figura 4.4b, mostra uma distribuição mais uniforme. Em particular, o gráfico da Figura 4.4a mostra que há uma variação, positiva ou negativa, de até três elétrons nas cargas de alguns átomos quando calculadas com os métodos HF, BP86 e B3LYP, com as funções de base incluindo funções difusas, o que é, obviamente, inconsistente com a expectativa química para a variação de cargas atômicas. Entretanto, no gráfico da Figura 4.4b nota-se que estes excessos são menos freqüentes, a exceção para os átomos C6 e N9. Para estes átomos vários cálculos prevêem cargas com mais que  $\pm 1$  elétron. Duas exceções são os cálculos BP86/6-31G\*\* e B3LYP/6-31G\*\*. Assim, não foram utilizadas funções difusas nos cálculos de otimização de geometria e cálculo das cargas e da SPE do sistema sob investigação. Percebe-se ainda que o valor das cargas CHelpG

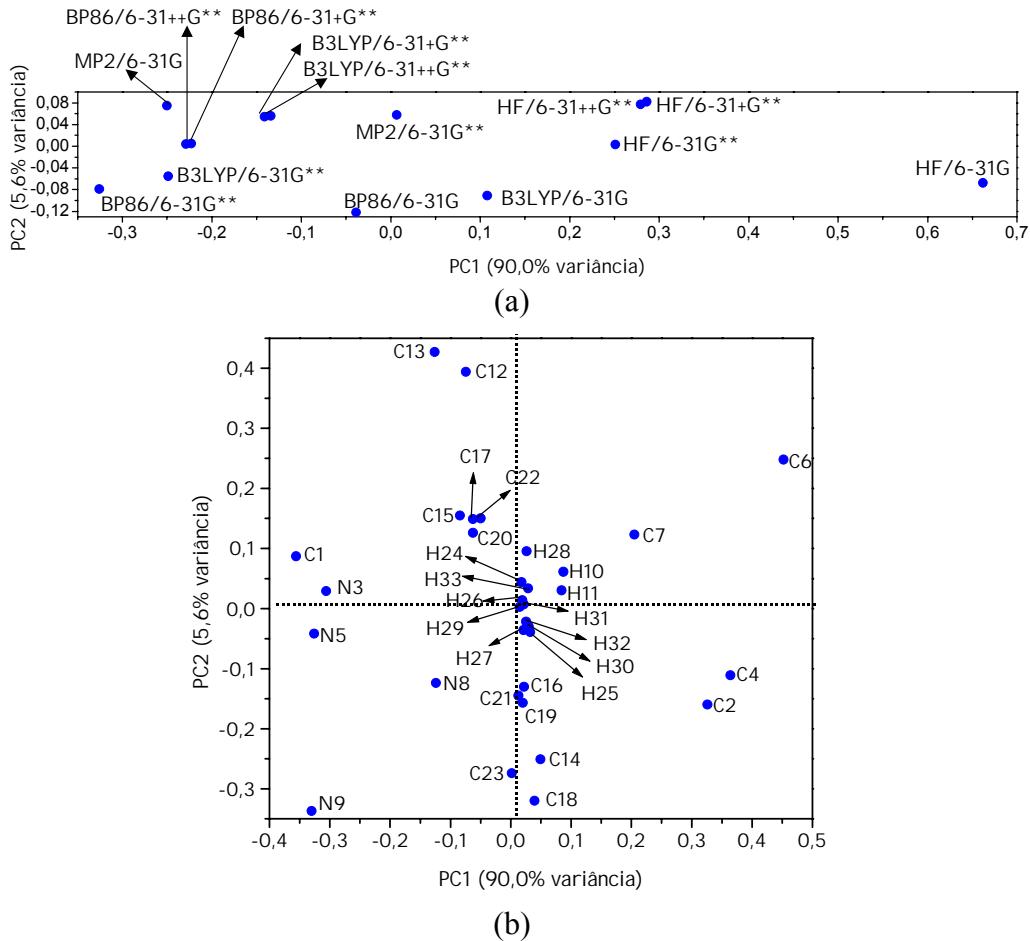
obtidas com método AM1 é sempre, em módulo, inferior as cargas calculadas pelos métodos *ab initio* e DFT, ou ainda, que as cargas atômicas obtidas por este método são sempre mais positivas ou mais negativas do que as cargas obtidas pelos demais métodos utilizados.

Então, ainda no intuito da escolha da metodologia de cálculo para o estudo das demais moléculas da série, foi feita uma PCA das cargas CHelpG da molécula **5h** para os diferentes níveis de cálculo e funções de base, conforme pode ser visto nas Figuras 4.5 e 4.6.



**Figura 4.5:** Gráfico dos escores de PC1 x PC2 da PCA das cargas CHelpG da molécula **5h**, calculados com os métodos AM1, HF B3LYP e BP86 com funções de base 6-31G, 6-31G\*\*, 6-31+G\*\* e 6-31++G\*\*, MP2 com funções de base 6-31G e 6-31G\*\*.

Pode-se perceber, através dos gráficos da Figura 4.5 que a primeira componente principal, ou seja, PC1 (com aproximadamente 79% da variância), separa os níveis de cálculo em dois grandes grupos, sendo um formado pelo método semi-empírico AM1 em escore negativo e outro pelos métodos *ab initio* e DFT em escores positivos. Já a segunda componente principal (PC2), com aproximadamente 17% da informação contida nos eixos originais, separa a presença ou ausência de funções de polarização e difusas. Uma análise mais detalhada do gráfico dos escores de PC1 x PC2 (Figura 4.5) mostra o método AM1 bastante diferente dos métodos *ab initio* e DFT, sugerindo que o método AM1 comporta-se como um *outlier* em relação aos demais objetos. Por esta razão, foi feita uma nova PCA das cargas CHelpG apenas para as moléculas calculadas com os métodos *ab initio* e DFT. Os resultados obtidos nesta nova análise estão apresentados nos gráficos da Figura 4.6.



**Figura 4.6:** Gráfico dos (a) Escores e (b) Loadings de PC1 x PC2 das cargas CHelpG da molécula **5h**, calculados com os métodos HF e DFT (B3LYP e BP86) com funções de base 6-31G, 6-31G\*\*, 6-31+G\*\* e 6-31++G\*\*, MP2 com funções de base 6-31G e 6-31G\*\*.

Após a exclusão do método AM1, a nova PCA (gráficos da Figura 4.6) mostrou um padrão semelhante ao apresentado pelos gráficos da Figura 4.5, porém, as diferenças das cargas atômicas CHelpG obtidas pelos métodos HF e pós-HF (DFT) estão mais ressaltadas. Nota-se neste último gráfico dos escores que PC1, com aproximadamente 90% da informação contida nos eixos originais (Figura 4.6a), representa o efeito da introdução de correlação eletrônica, estando os cálculos pós-HF em valores mais negativos de PC1. O principal efeito observado em PC2, com 5,6% da variância, é a presença de funções difusas onde os objetos que possuem essas funções encontram-se em valores mais positivos de PC2. Porém, os resultados MP2 não seguem o mesmo

padrão, uma vez que a presença destas funções leva esse objeto a um valor mais negativo de PC2. Esta diferença apresentada pelo método MP2 no gráfico dos escores da Figura 4.6 pode ser resultado da quebra, ou diminuição, da ressonância dos elétrons do grupo farmacofórico e do anel benzênico ligado ao carbono C4, o qual se encontra mais rotacionado em relação ao plano molecular (ver Tabela 4.1).

Através dos gráficos dos *loadings* das duas primeiras componentes principais (Figura 4.6b) pode-se ver que os átomos responsáveis pela formação do padrão dos gráficos dos escores são os átomos do anel central, isto é, o grupo farmacofórico (C1, C2, N3, C4, N5 e C6), dos substituintes (C7 e N8 e N9), e os átomos de carbono C12 e C13 que ligam os anéis benzênicos ao anel pirimidínico. Assim, uma diminuição no valor da carga dos átomos C1, N9, N5, N3, C13, N8 levam os objetos para valores mais negativos de PC1, ou seja, para a inclusão de efeitos de correlação eletrônica, já, um aumento no valor da carga dos átomos C7, C2, C4 e C6 levará os objetos para a região mais próxima dos métodos HF. A diminuição na carga dos átomos N9, C18, C23, C14, C2, C19, C21, C16, N8, C4 e um aumento positivo nas cargas dos átomos C7, C20, C17, C22, C15, C6, C12 e C13 leva os objetos para valores mais positivos de PC2, descrevendo, então, os efeitos dados pelas funções de polarização e difusas. Percebe-se também que os átomos estão separados em valores positivos e negativos de PC1 e PC2 segundo serem átomos com cargas positivas ou negativas, conforme pode ser visto no gráfico da Figura 4.4b.

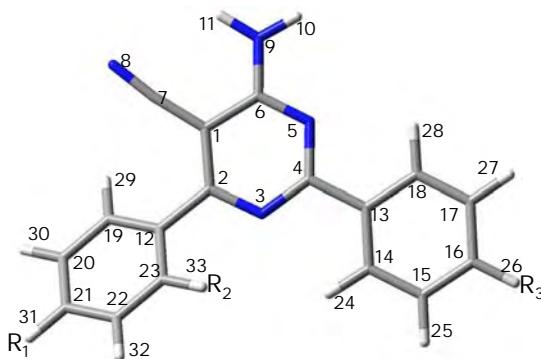
Foi, portanto, da análise conjunta entre o padrão observado no gráfico dos escores da PCA da Figura 4.6a, da variação das cargas atômicas apresentadas nos gráficos da Figura 4.4 e da comparação das estruturas de raios-X obtidas no CSD, que se selecionou o método DFT com a função de base 6-31G\*\* como sendo o mais adequado para as etapas subsequentes deste trabalho, qual seja, a otimização de geometria e cálculo da SPE das demais moléculas da série (pirimidinas substituídas **5a-q**). Em particular, o funcional BP86 foi escolhido tendo em vista que o funcional B3LYP não descreve, de maneira satisfatória, ligações de hidrogênio, e este tipo de ligação será importante nos estudos futuros das estruturas de hidratação destes compostos.

## 4.2. Pirimidinas Substituídas

### 4.2.1. Análise conformacional

Após a escolha da metodologia de cálculo, passou-se a estudar a série homóloga das pirimidinas substituídas nas posições R<sub>1</sub>, R<sub>2</sub> e R<sub>3</sub>. Estas moléculas estão novamente apresentadas na Figura 4.7 e Tabela 4.3, abaixo.

**Tabela 4.3:** Pirimidinas substituídas.

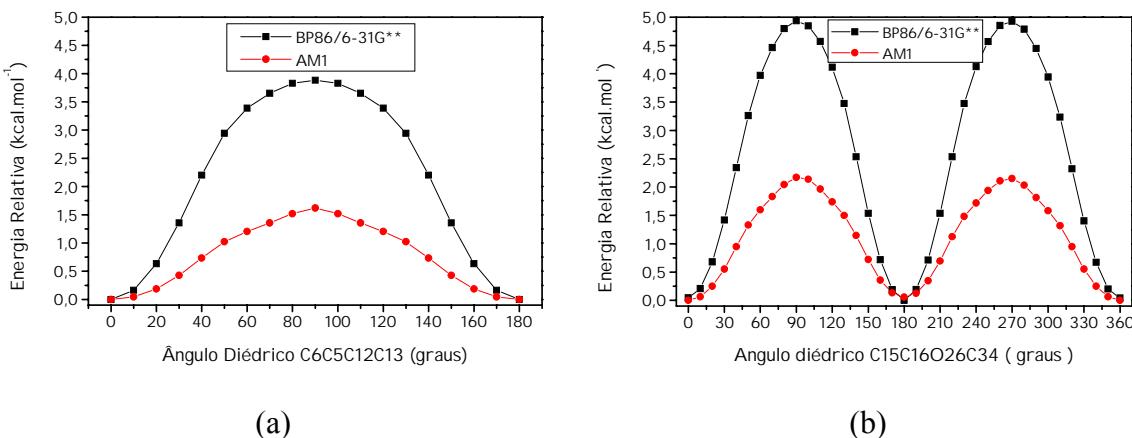


**Figura 4.7:** Pirimidina Substituída.

Molécula	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
<b>5a</b>	Cl	H	OCH <sub>3</sub>
<b>5b</b>	Cl	H	H
<b>5c</b>	F	H	H
<b>5d</b>	Cl	Cl	H
<b>5e</b>	Cl	Cl	OCH <sub>3</sub>
<b>5f</b>	F	H	OCH <sub>3</sub>
<b>5h</b>	H	H	H
<b>5i</b>	OCH <sub>3</sub>	H	OCH <sub>3</sub>
<b>5l</b>	OCH <sub>3</sub>	H	H
<b>5n</b>	NH <sub>2</sub>	H	OCH <sub>3</sub>
<b>5p</b>	NO <sub>2</sub>	H	OCH <sub>3</sub>
<b>5q</b>	CF <sub>3</sub>	H	OCH <sub>3</sub>

Tendo em vista a presença do grupo metoxi em várias das moléculas analisadas (**5a**, **5e**, **5f**, **5i**, **5l**, **5n**, **5p** e **5q**), é importante que seja feita uma análise conformacional deste grupo funcional destas moléculas para se identificar à conformação mais estável. Neste intento, fez-se o cálculo da barreira de energia de rotação do grupo CH<sub>3</sub> em torno da ligação C-O (opção *scan* no G98) de duas maneiras distintas: i) apenas para o metoxi-benzeno e ii) para uma pirimidina substituída, como explicado na seção 3.3 da metodologia, página 19, nos níveis AM1 e BP86/6-31G\*\*.

Os gráficos das barreiras para o metoxi-benzeno e para uma pirimidina substituída (**5f**), variando-se o ângulo diédrico do grupo metoxi com um passo de 10 graus, entre 0° e 180° e entre 0° e 360°, respectivamente, como pode ser visto na Figura 4.8.



**Figura 4.8:** Gráficos das barreiras de energia relativa (em relação à conformação mais estável) (a) do metoxi-benzeno, e (b) da molécula **5f** calculadas com os métodos BP86/6-31G\*\* (preto) e com AM1 (vermelho).

O gráfico da barreira de energia para o metoxi-benzeno (Figura 4.8a), em ambas as metodologias de cálculo, mostra que há dois pontos de mínimo de energia equivalentes nesta molécula, um a  $0^\circ$  e outro a  $180^\circ$ , sendo que o ponto de máximo de energia encontra-se a  $90^\circ$ . Obviamente, a presença de dois mínimos equivalentes neste gráfico pode ser explicada pela simetria da molécula. Já a presença do ponto de máximo em  $90^\circ$  pode ser resultado de uma quebra de ressonância entre os elétrons  $\pi$  do anel benzênico e os pares de elétrons do oxigênio do grupo metoxi. O valor calculado da barreira de energia do metoxi benzeno é de  $3,88 \text{ kcal.mol}^{-1}$  para o método BP86/6-31G\*\* e de  $1,62 \text{ kcal.mol}^{-1}$  para o método AM1. A diferença no valor da energia entre ambos os métodos, pode ser explicada pela tendência que o método AM1 tem em subestimar os valores das barreiras de energia<sup>65</sup>.

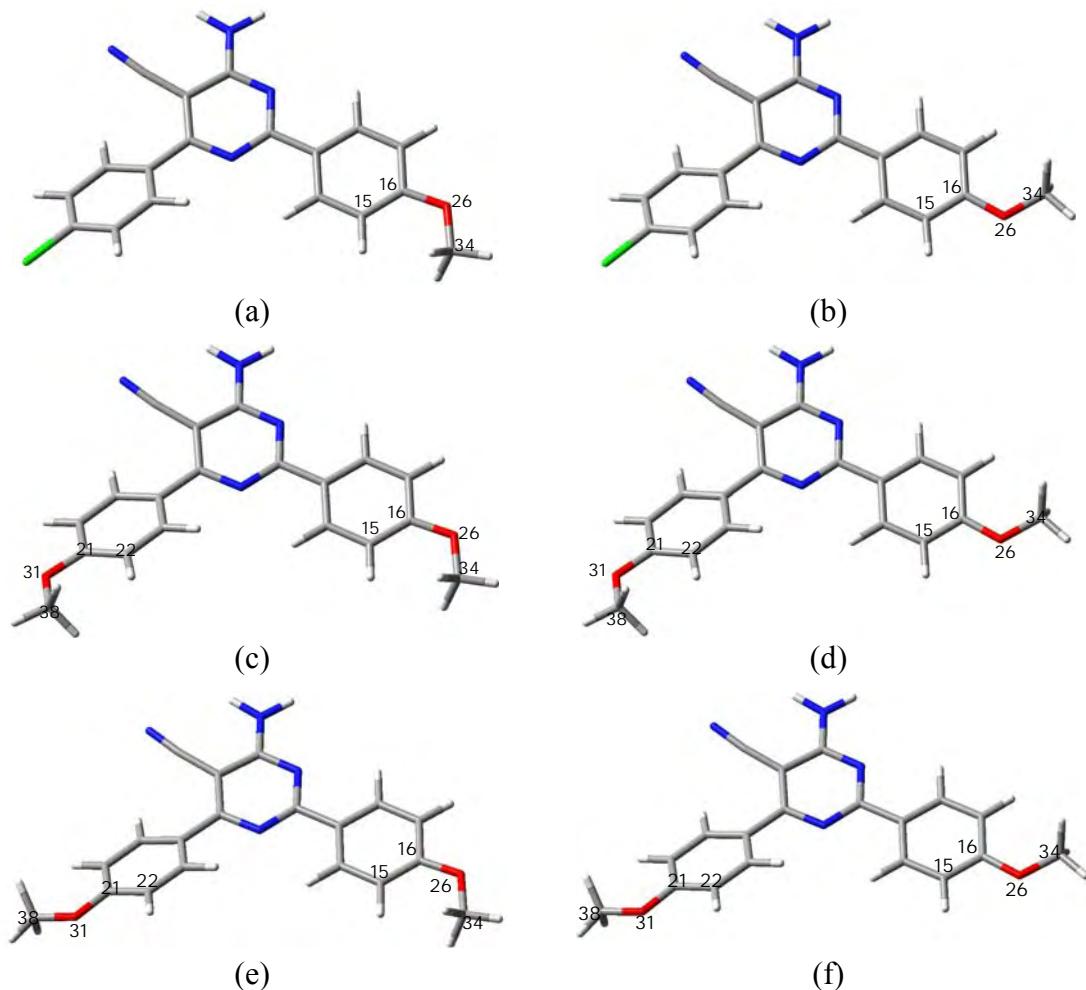
Contudo, para a molécula **5f** pode-se perceber que há dois pontos de máximo e dois pontos de mínimo, muito próximos, porém não equivalentes; calculados com os dois métodos (Figura 4.8b). Neste caso, o ponto de mínimo mais estável é aquele no qual o grupo metoxi encontra-se orientado para a parte interna da molécula, ou seja, com um ângulo de  $0^\circ$  em relação ao átomo C15 (ângulo diédrico C15C16O26C34 igual a  $0^\circ$ ), estas conformações foram chamadas, neste trabalho, de conformações “dentro”, como descrito na seção 3.6, página 23. O segundo ponto de mínimo de energia aparece com o grupo metoxi orientado para a parte externa da molécula, formando um ângulo diédrico de  $180^\circ$  com o átomo C15 (ângulo diédrico C15C16H26C34 igual a  $180^\circ$ ) e,

neste caso, as conformações foram chamadas de conformações “fora”. Os pontos de máximo encontram-se a 90 e 270 graus em relação a C14 devido a possível quebra de ressonância entre os pares de elétrons do oxigênio e a molécula, como observado anteriormente para a molécula do metoxibenzeno. Pode-se perceber, também, neste gráfico que o valor da barreira de energia da molécula **5f** é de aproximadamente 5,08 kcal.mol<sup>-1</sup> e 2,17 kcal.mol<sup>-1</sup> para os métodos BP86/6-31G\*\* e AM1, respectivamente. Nesta molécula, também, pode-se perceber, que o método AM1 subestima o valor da barreira de energia, porém, novamente mostra-se satisfatório na identificação das conformações mais estáveis. Além disto, é interessante notar que a diferença entre a altura da barreira de rotação do grupo metoxi da molécula **5f** e no metoxibenzeno é de 1,2 kcal.mol<sup>-1</sup> e 0,55 kcal.mol<sup>-1</sup> para os métodos BP86/6-31G\*\* e AM1, respectivamente. Esta diferença pode ser atribuída a efeitos de ressonância para com o anel central da pirimidina em **5f** sendo possivelmente menor no caso AM1 devido a maior perda de planaridade do anel benzênico quando comparado ao método BP86/6-31G\*\* (Tabela 4.1), o que seria uma demonstração interessante do conceito amplamente utilizado que, a mudança estrutural numa série homóloga leva ao ajuste fino das propriedades físico-químicas do grupo farmacofórico.

Neste ponto, vale a pena ressaltar que o estudo da barreira de energia foi feito apenas para uma molécula da série homóloga de pirimidinas substituídas, a molécula **5f**, devido a alta demanda computacional necessária para este cálculo, com um custo de, aproximadamente, 12 dias de CPU (microcomputador Pentium 4 2,6 GHz, dois HD's de 80GB de HD, 1GB de memória RAM), sendo, então, este resultado extrapolado para as demais moléculas da série, as quais tiveram as geometrias otimizadas e a energia total determinada apenas nos dois pontos de mínimo, ou seja, nas duas conformações mais estáveis. A Figura 4.9 mostra as duas conformações mais estáveis que serão utilizadas em todas as moléculas da série. Em particular, a molécula **5i** apresenta quatro conformações de baixa energia, uma vez que possui dois grupos metoxi como substituintes (Figura 4.9c-f). Neste caso, a conformação mais estável é aquela que possui os diedros C15C16O26C34 e C22C21O31C38 iguais a 0°, isto é, a conformação “dentro/dentro”.

É importante lembrar que devido à presença de dois grupos metoxi na molécula **5i** as conformações mais estáveis receberam uma nomenclatura especial, como descrito

na seção 3.6, página 23. A molécula que possui os dois metoxis com ângulos diédricos C22C21O31C38 e C15C16O26C34 iguais a  $0^\circ$ , foi chamada de molécula “dentro/dentro” e recebeu o subscrito ‘dd’ (Figura 4.9c). Quando os diedros C22C21O31C38 e C15C16O26C34 foram iguais a  $0^\circ$  e  $180^\circ$ , respectivamente a molécula recebeu o nome de “dentro/fora” e o subscrito ‘df’ (Figura 4.9d), quando estes diedros forem de  $180^\circ$  e  $0^\circ$  respectivamente, a molécula foi chamada de “fora/dentro” e recebeu o subscrito ‘fd’ (Figura 4.9e) e quando os dois diedros C22C21O31C38 e C15C16O26C34 foram iguais a  $180^\circ$  a molécula foi chamada de molécula “fora/fora” e recebeu o subscrito ‘ff’ (Figura 4.9f).



**Figura 4.9:** Conformações mais estáveis das pirimidinas substituídas com um grupo metoxi com as conformações (a) “dentro” (b) “fora” e com dois metoxi com as conformações (c) “dentro/dentro”, (d) “dentro/fora”, (e) “fora/dentro” e (f) “fora/fora”.

Desta forma, obteve-se dois conjuntos, um formado por 22 moléculas, incluindo-se as duas (ou quatro), conformações mais estáveis (ou seja, as conformações “dentro” e “fora” das pirimidinas **5a**, **5e**, **5f**, **5l**, **5n**, **5p** e **5q**, e as conformações “dentro/dentro”, “dentro/fora”, “fora/dentro” e “fora/fora” da molécula **5i**), além das moléculas **5b**, **5c**, **5d** e **5h** que possuem apenas uma conformação. O outro conjunto é composto pelas 12 moléculas mais estáveis, ou seja, aquelas com conformação “dentro” e “dentro/dentro”.

### 4.3. Comparação das pirimidinas substituídas

A introdução do substituinte na pirimidina **5h** (molécula padrão) foi feita de duas maneiras distintas como descrito nas seções 3.2.2 e 3.2.3 da metodologia, página 17 e 18, respectivamente. Nesta seção serão apresentados, de forma comparativa, os resultados obtidos pelos dois métodos de otimização de geometria, isto é, os resultados obtidos para as moléculas semi-otimizadas e totalmente otimizadas.

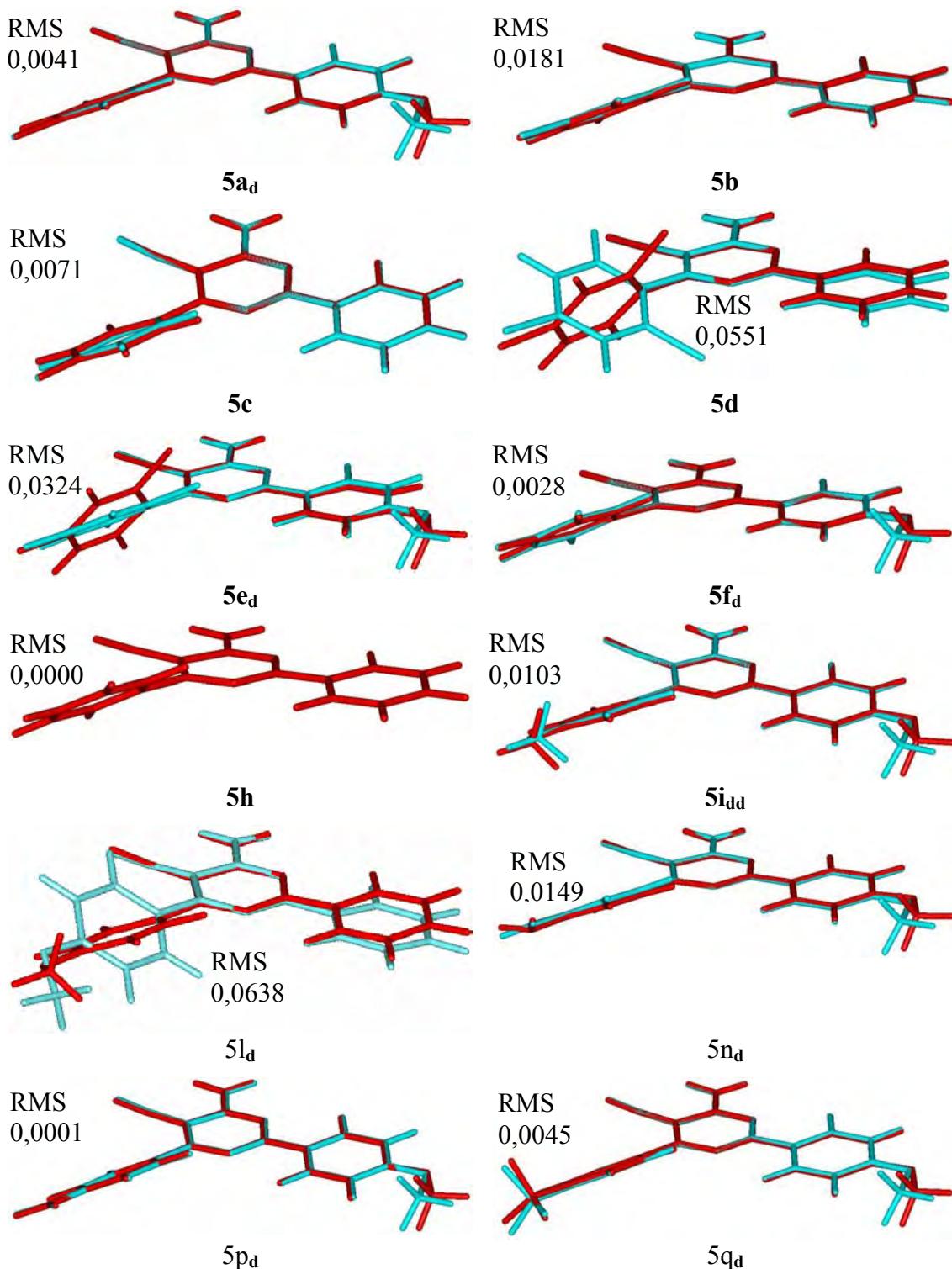
#### 4.3.1. Comparações entre os ângulos diédricos C23C12C2N3 e C14C13C4N3

Inicialmente, fez-se uma comparação dos ângulos diédricos dos anéis laterais (os anéis benzênicos) C23C12C2N3 e C14C13C4N3 das pirimidinas substituídas entre as geometrias semi e totalmente otimizadas (ver seção 3.3 da metodologia, página 19). Neste estudo foram consideradas apenas as moléculas que possuem o grupo metoxi na posição mais estável, ou seja, com os diédros C22C21O31C38 e C15C16O26C34 iguais a 0° para aquelas moléculas que possuem este grupo como substituinte em uma ou ambas posições. Os valores destes ângulos para as moléculas totalmente otimizadas estão apresentados na Tabela 4.4. Os diedros das pirimidinas com a geometria semi-otimizada são, em todos os casos, os mesmos da molécula 5h totalmente otimizada (molécula padrão), conforme explicado na metodologia.

**Tabela 4.4:** Ângulos diédricos C23C12C2N3 e C14C13C4N3 das pirimidinas substituídas **5a-q** totalmente otimizadas com o método BP86/6-31G\*\*.

	Diédro C23C12C2N3 (graus)	Diédro C14C13C4N3 (graus)		Diédro C23C12C2N3 (graus)	Diédro C14C13C4N3 (graus)
<b>5a</b>	26,51	2,74	<b>5h</b>	28,16	3,19
<b>5b</b>	26,43	2,33	<b>5i</b>	24,60	2,61
<b>5c</b>	26,05	3,73	<b>5l</b>	24,23	1,53
<b>5d</b>	52,68	-1,17	<b>5n</b>	22,20	2,84
<b>5e</b>	53,30	-1,01	<b>5p</b>	28,56	2,76
<b>5f</b>	26,35	2,89	<b>5q</b>	27,54	3,26

Em seguida, fez-se um estudo da similaridade, com o programa ALINHAMOL, da geometria semi-otimizada e totalmente otimizada destas moléculas, as quais estão apresentadas na Figura 4.10.



**Figura 4.10:** Comparação entre as geometrias semi-otimizadas (azul) e totalmente otimizadas (vermelho) das moléculas **5a-n** obtidas pelo método BP86/6-31G\*\*.

Pode-se perceber através do valor do RMS apresentado na Figura 4.10 que o efeito da otimização completa sobre a geometria é pequeno. Um RMS igual a zero indica que as moléculas são idênticas e um aumento no valor do RMS indica uma diminuição na similaridade. Para esta série de compostos, a molécula **5h** possui um RMS = 0,0000, mostrando não haver diferença nas geometrias da molécula semi e totalmente otimizadas, uma vez que se trata da mesma estrutura. Como explicado anteriormente, esta molécula foi utilizada como a molécula padrão ou de referência. As moléculas **5a**, **5b**, **5c**, **5f**, **5p** e **5q** apresentam o RMS na ordem de  $10^{-3}$  indicando alto grau de similaridade entre as moléculas semi e totalmente otimizadas. Já as moléculas **5d**, **5e**, **5i** e **5n** apresentam um RMS da ordem de  $10^{-2}$  indicando que estas moléculas apresentam diferenças conformacionais mais significativas que as anteriores.

A análise da sobreposição das estruturas obtidas pelos dois processos de otimização de geometria indica que o efeito da otimização da geometria completa das pirimidinas é significativo apenas para as moléculas **5d** e **5e**, ou seja, para as moléculas que possuem substituição, com o cloro, na posição R<sub>2</sub>. As demais moléculas possuem uma geometria molecular bastante similar com aquela da molécula **5h**. Provavelmente, há uma torção do anel fenila do lado esquerdo, ou seja, há uma variação do diedro C23C12C2N3 nas moléculas **5d** e **5e** para diminuir a repulsão eletrônica entre o átomo de Cl e o par de elétrons do nitrogênio N3 do anel pirimidínico (aproximadamente 26° a mais que a rotação deste diedro na molécula **5h** otimizada no mesmo nível de teoria). Como as demais moléculas desta série apresentam apenas o átomo de H como substituinte da posição R<sub>2</sub> o efeito da otimização da geometria completa ser pequeno (desprezível) para este diedro.

O processo de otimização da geometria molecular exigido pelo método DFT-BP86/6-31G\*\* requer um alto consumo de tempo e demanda computacional, sendo necessário no mínimo 6h e 5 minutos e no máximo 1 dia e 25 minutos de tempo de CPU (em microprocessadores Pentium 4 2.8GHz, 2,0GB MB de RAM DDR DUAL CHANNEL 400 2 HD's de 80 GB e sistema linux RED HAT 9) para a otimização completa da geometria molecular das pirimidinas substituídas. A falta de espaço em disco ou de memória necessários para estes cálculos também foram barreiras encontradas durante a otimização da geometria completa, tendo sido necessário, em muitos casos, a re-submissão dos cálculos.

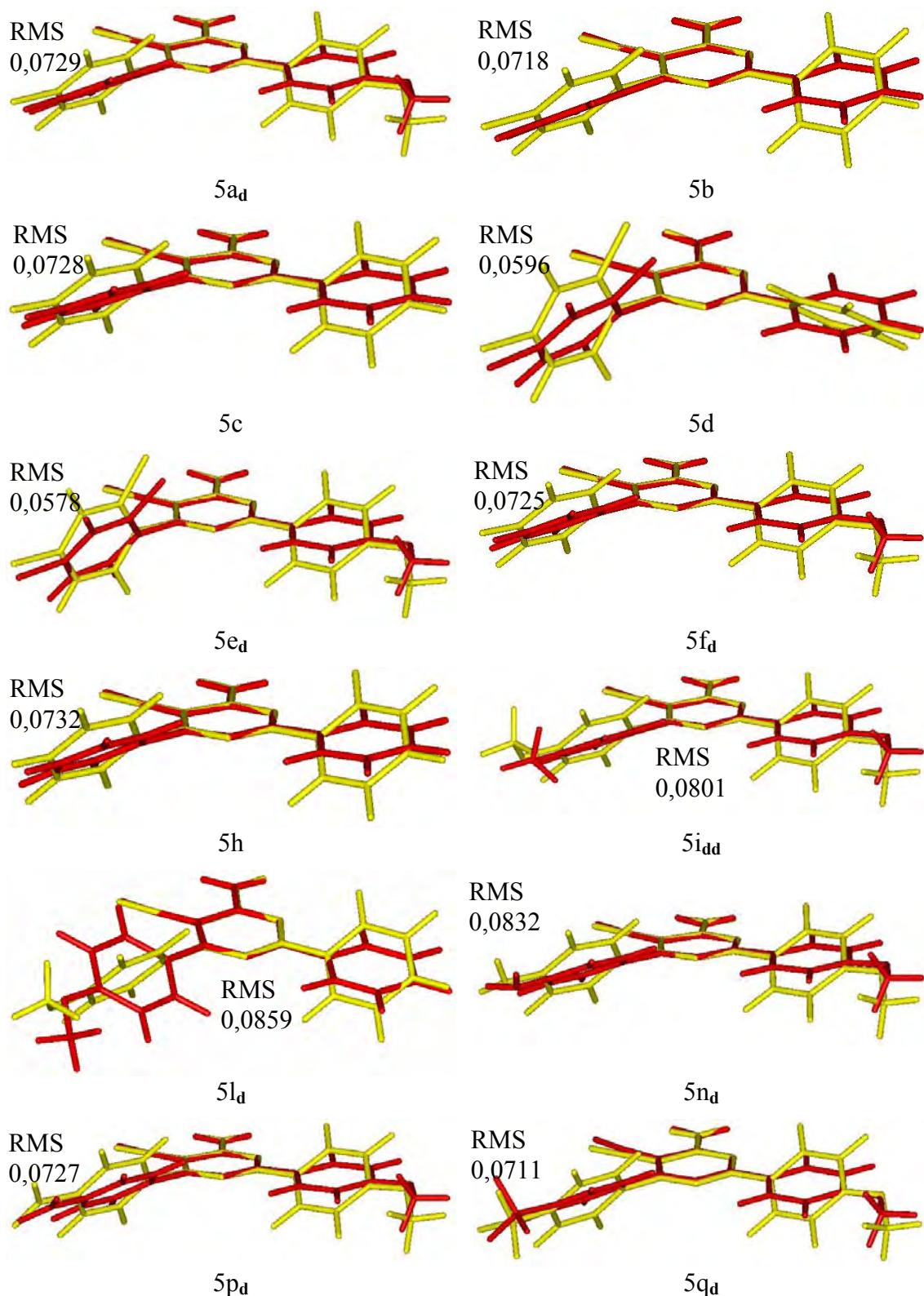
É importante enfatizar, que neste processo de otimização de geometria, utilizou-se o esqueleto molecular das pirimidinas semi-otimizadas, ou seja, partiu-se de uma molécula cuja geometria havia tido seu esqueleto molecular previamente otimizado, a molécula **5h** (1 dia, 2h e 42 minutos de tempo de CPU, Atlon 1,2GHz, 40,0 GB de HD, 1,0 GB de memória RAM DDR, sistema linux RED HAT 9), e substituintes previamente otimizados (ver seção 3.2.2 da metodologia, página 17), levando a uma grande economia de tempo de CPU.

Desta forma, a fim de se avaliar o uso de uma geometria padrão, porém com uma maior economia de tempo computacional, decidiu-se utilizar o método AM1, (que requer apenas alguns minutos de tempo de CPU). A Tabela 4.5 apresenta os ângulos diédicos C23C12C2N3 e C14C13C4N3 obtidos com o método AM1 para esta série.

**Tabela 4.5:** Ângulos diédricos C23C12C2N3 e C14C13C4N3 das pirimidinas substituídas **5a-q** totalmente otimizadas com o método AM1.

	Diédro C23C12C2N3 (graus)	Diédro C14C13C4N3 (graus)		Diédro C23C12C2N3 (graus)	Diedro C14C13C4N3 (graus)
<b>5a</b>	49,78	34,91	<b>5h</b>	49,75	38,85
<b>5b</b>	49,61	38,73	<b>5i</b>	48,43	35,46
<b>5c</b>	49,14	38,80	<b>5l</b>	50,80	35,88
<b>5d</b>	84,04	-38,26	<b>5n</b>	47,29	35,26
<b>5e</b>	80,10	34,27	<b>5p</b>	51,65	34,44
<b>5f</b>	49,35	34,95	<b>5q</b>	50,96	34,45

Percebe-se na Tabela 4.5 que os ângulos diédricos das pirimidinas substituídas otimizadas com o AM1 são maiores, em módulo, do que aqueles obtidos com o método BP86/6-31G\*\*, porém, apresentam o mesmo sentido de variação, isto é, há um aumento ou diminuição nos diédros C23C12C2N3 e C14C13C4N3 para as mesmas moléculas quando compararam-se os dois métodos. Uma comparação das geometrias totalmente otimizadas, das pirimidinas substituídas mais estáveis, obtidas com os métodos DFT-BP86/6-31G\*\* e AM1 está apresentada na Figura 4.11.



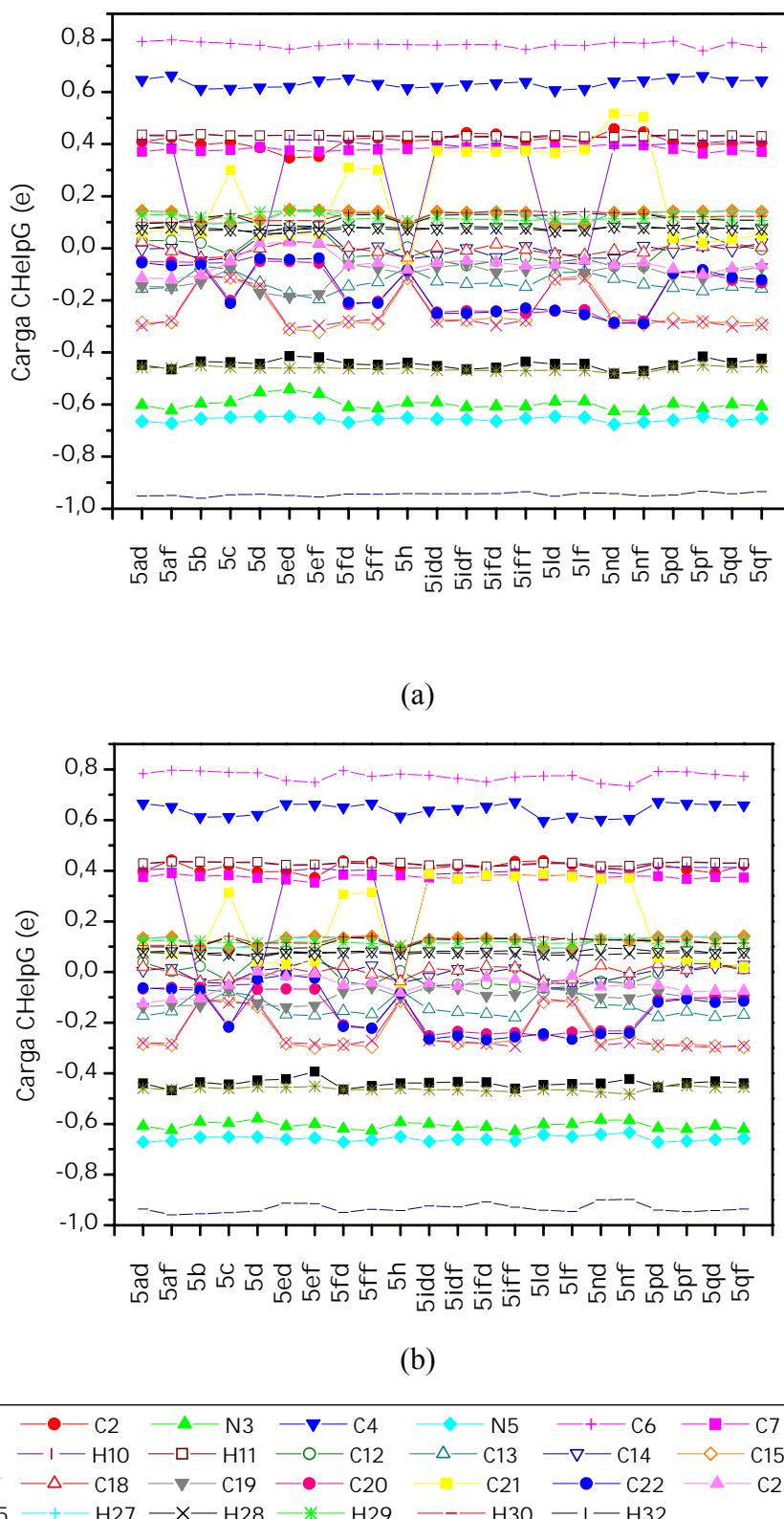
**Figura 4.11:** Comparação das geometrias de equilíbrio das moléculas **5a-q** totalmente otimizadas obtidas com os métodos BP86/6-31G\*\* (vermelho) e AM1 (amarelo).

Como observado para a molécula **5h** (seção 4.1 página 40), as demais pirimidinas substituídas calculadas com o método AM1 apresentaram uma geometria molecular bastante diferente das geometrias obtidas com método BP86/6-31G\*\*. Essa diferença é mais significativa para o anel benzila do lado direito da molécula, ou seja, a diferença é maior para o diângulo C14C13C4N3 que para o diângulo C23C12C2N3, conforme pode ser visto na Tabela 4.5 e na Figura 4.11. As moléculas **5d** e **5e**, que se mostraram bastantes sensíveis ao efeito da otimização total da geometria completa, quando agora se utilizou o método DFT, também apresentaram uma maior torção no ângulo diédrico C23C12C2N3, em aproximadamente o dobro da torção apresentada por este diângulo na molécula **5h**, quando otimizadas com o método AM1. Logo, percebe-se que o efeito da repulsão eletrônica entre o átomo de cloro da posição R<sub>2</sub> e o par de elétrons de N3 estão sendo contemplados, também, quando se utiliza o método AM1. Entretanto, a rotação do diângulo C14C13C4N3 traz a necessidade de uma investigação mais detalhada da utilização do método AM1 no cálculo das propriedades eletrônicas, como por exemplo, a carga atômica, calculada com a partição CHelpG destas moléculas, para que se possa inferir se é necessário tão somente a utilização de um método e conjunto de base mais sofisticados para um estudo de QSAR-3D.

Então, devido ao alto custo computacional requerido, pelo método DFT, no processo de otimização de geometria, analisou-se as diferenças nas cargas de todos os átomos e, posteriormente, da SPE desta série de compostos para a geometria semi-otimizada e totalmente otimizada através da comparação da análise de PCA (itens 4.5.2 e 4.5.3, respectivamente), conforme pode ser visto a seguir.

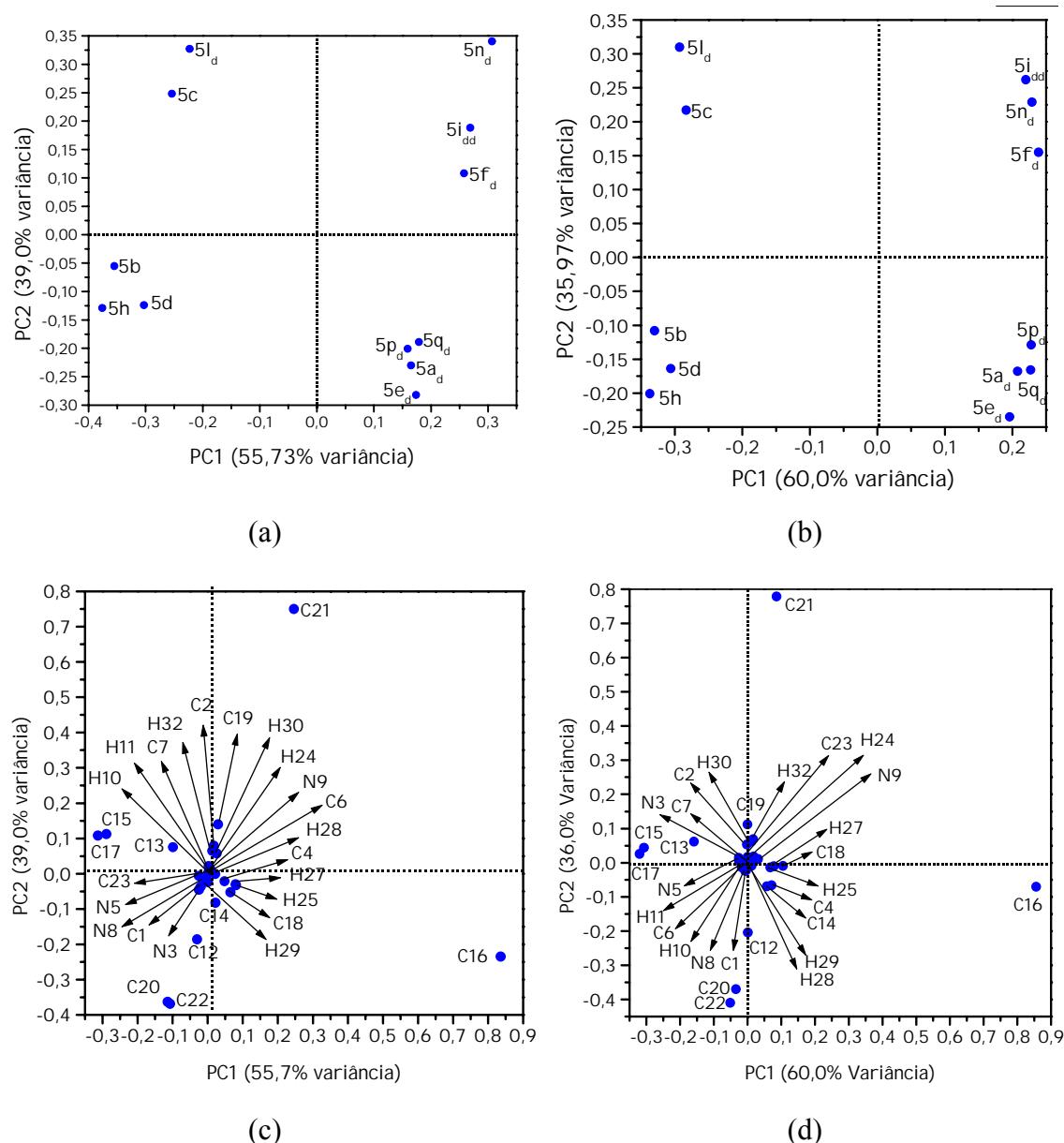
#### 4.3.2. Comparação das cargas CHelpG em função da otimização da geometria

De forma semelhante à realizada para a molécula **5h**, foi feito, inicialmente, um estudo das cargas CHelpG das pirimidinas substituídas **5a-q** (Tabela 4.3) as quais apresentaram o comportamento mostrado nos gráficos da Figura 4.12.



**Figura 4.12:** Gráfico das cargas CHelpG das pirimidinas **5a-q**: (a) com a geometria semi-otimizada e (b) com a geometria totalmente otimizada.

Percebe-se nos gráficos da Figura 4.12 que as cargas CHelpG das 22 estruturas comportam-se de modo semelhante para as duas metodologias utilizadas, onde apenas os átomos ligados diretamente aos substituintes, ou próximos a ele, terão um padrão diferente (para cada molécula) de acordo com as características do substituinte, por exemplo, o comportamento, ou sinal, da carga dos átomos C15, C16, C21 e C22 (átomos próximos aos substituintes), será diferente caso o substituinte seja um grupo doador ou retirador de elétrons, mostrando que, neste caso, o efeito do substituinte será apenas um efeito local na distribuição de cargas, não influenciando, de maneira significativa, a distribuição de cargas do grupo farmacofórico. Então, a fim de se identificar um padrão nesta série do ponto de vista da distribuição de carga, foi feita posteriormente, uma Análise de Componentes Principais das cargas CHelpG deste conjunto de moléculas considerando duas situações distintas: i) o conjunto de 12 moléculas e ii) o conjunto de 22 moléculas.



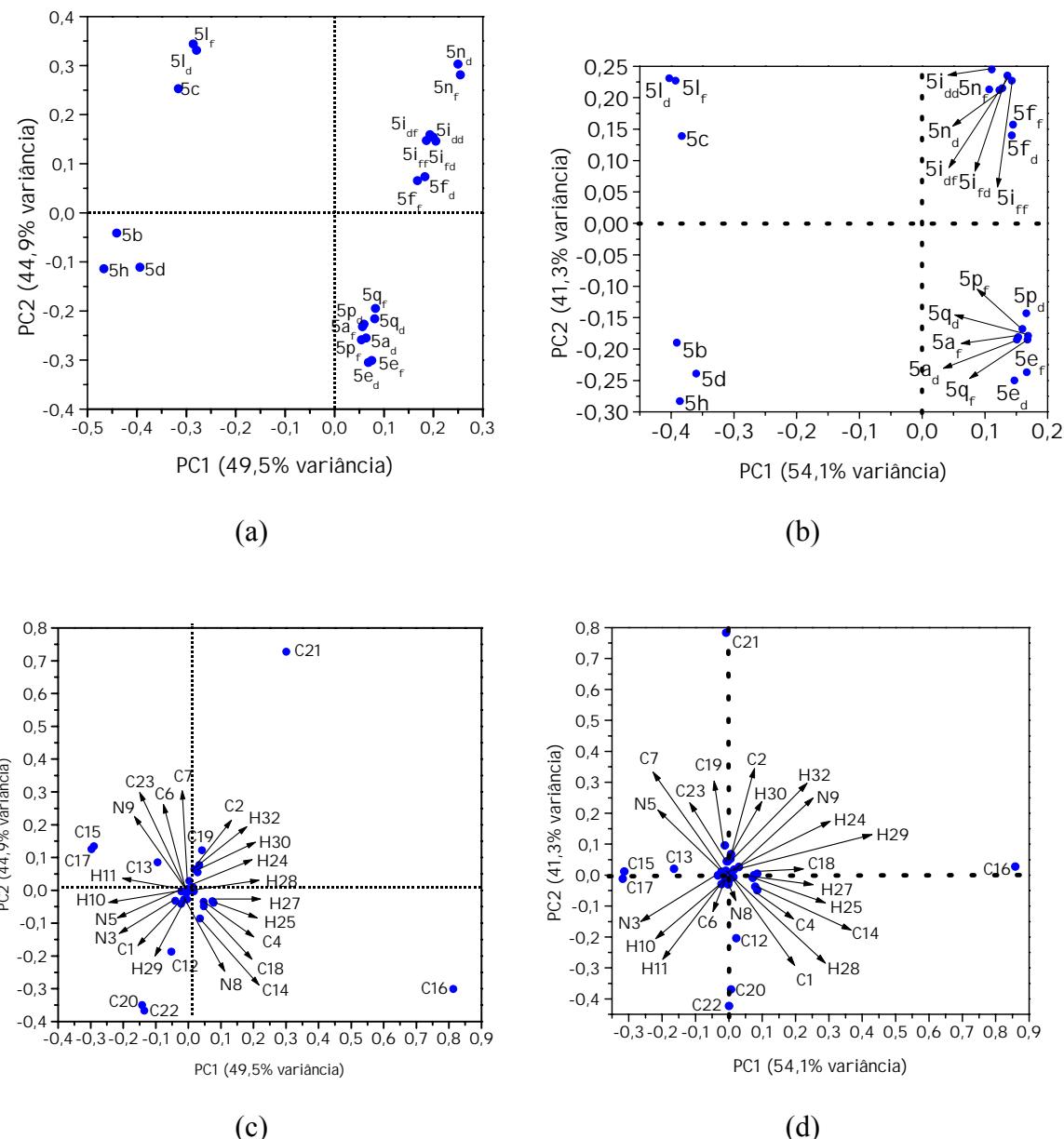
**Figura 4.13:** Gráfico dos escores de PC1 x PC2 das cargas CHelpG das moléculas **5a-q** com a geometria (a) semi-otimizada e (b) totalmente otimizada; e gráficos dos *loadings* de PC1 x PC2 das cargas CHelpG das moléculas **5a-q** com a geometria (c) semi-otimizada e (d) totalmente otimizada.

Pode-se perceber nos gráficos dos escores da Figura 4.13a e b que PC1 (55,7% e 60,0% da variância, respectivamente) está separando as moléculas segundo a substituição em R<sub>3</sub>, ou seja, a presença ou ausência do grupo metoxi nesta posição. Por sua vez, PC2 (39,0% e 35,9% da variância, respectivamente) está sendo descrita pela

natureza do substituinte da posição R<sub>1</sub>, ou seja, se estes substituintes são grupos para-retiradores de elétrons (Cl, NO<sub>2</sub> e CF<sub>3</sub>), as moléculas **5a**, **5b**, **5d**, **5e**, **5h**, **5p** e **5q**, ou grupos para-doadores de elétrons (OCH<sub>3</sub>, NH<sub>2</sub>, F e H) as moléculas **5c**, **5f**, **5i**, **5l** e **5n**, havendo assim a formação de quatro grupos distintos (no gráfico dos escores de PC1 x PC2), com i) PC1 e PC2 positivos (moléculas **5f**, **5i** e **5n**), ii) PC1 positivo e PC2 negativo (moléculas **5a**, **5e**, **5p** e **5q**), iii) PC1 negativo e PC2 positivo (moléculas **5c** e **5l**) e iv) PC1 e PC2 negativos (moléculas **5b**, **5d** e **5h**). Neste ponto vale a pena ressaltar que o flúor (moléculas **5c** e **5f**) apesar de ser um grupo retirador de elétrons através das ligações- $\sigma$ , também pode agir como um doador, através de uma retrodoação de cargas pelas ligações- $\pi$ , compensando, então, o efeito anterior e atuando como um grupo doador de elétrons, quando em posição *p*-substituído. Uma forma alternativa de se avaliar esse efeito é através dos valores da constante de Hammett ( $\sigma$ ) que é igual a 0,34 e 0,06 para o meta-fluor-benzeno e para-fluor-benzeno, respectivamente<sup>65</sup>.

Os gráficos dos *loadings* de PC1xPC2 (Figura 4.13c e d respectivamente) mostram a contribuição de cada variável na formação dos padrões observados nos gráficos dos escores (Figuras 4.13a e b). Logo, pode-se perceber que os átomos C13, C15, C16 e C17 são os principais responsáveis pela separação das moléculas ao longo de PC1, enquanto os átomos C12, C16, C20, C21 e C22 são os principais responsáveis pela separação apresentada em PC2.

Posteriormente, repetiu-se este estudo introduzindo-se no conjunto de dados as pirimidinas com o grupo metoxi nas conformações “fora”. Os gráficos dos escores e dos *loadings* das duas primeiras componentes principais podem ser vistos nos gráficos da Figura 4.14.



**Figura 4.14:** Gráfico dos escores de PC1 x PC2 das cargas CHelpG do conjunto de 22 moléculas das pirimidinas substituídas **5a-q** com a geometria (a) semi-otimizada e (b) totalmente otimizada; e gráficos dos *loadings* de PC1 x PC2 das cargas CHelpG das moléculas 5a-q com a geometria (c) semi-otimizada e (d) totalmente otimizada.

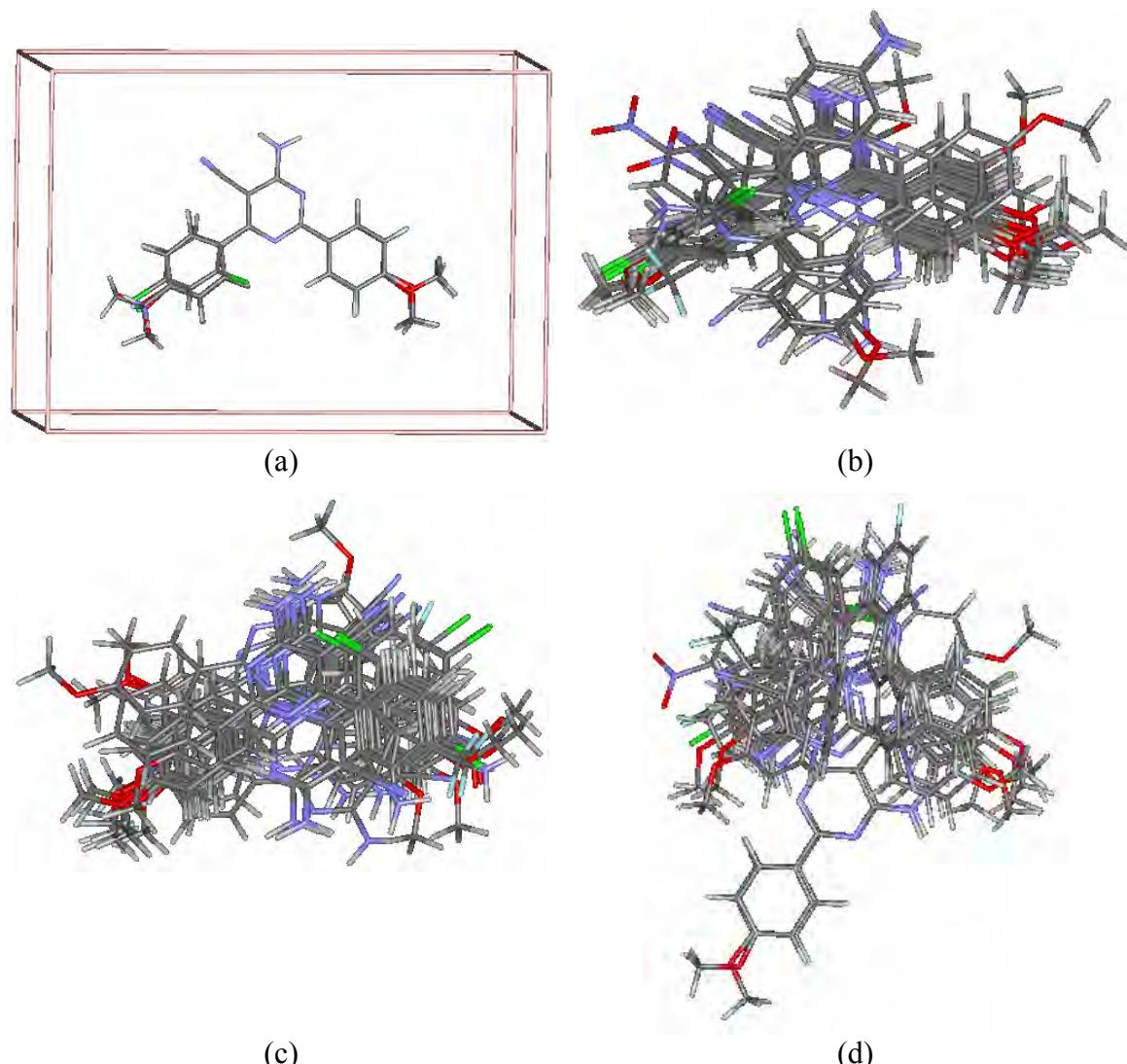
Nota-se no gráfico dos escores da Figura 4.14 que o padrão observado é similar ao apresentado pelos gráficos da Figura 4.13. Neste novo gráfico dos escores, observa-se uma separação das moléculas com a mesma formação de quatro grupos, tal como na figura anterior.

Os gráficos dos *loadings* (Figura 4.14c e d) também apresentam o mesmo padrão dos gráficos da Figura 4.13c e d, onde os átomos C13, C15, C16 e C17 são os principais responsáveis pela separação das moléculas ao longo de PC1, enquanto que os átomos C12, C16, C20, C21 e C22 são os principais responsáveis pela separação apresentada em PC2.

Desta forma, pode-se concluir que a introdução das moléculas com a segunda conformação mais estável, ou seja, as moléculas com conformação “fora”, não modifica de maneira significativa o padrão da PCA e, assim, podem ser incluídas nas análises posteriores. Conhecer esta informação foi importante para os nossos estudos de regressão pois, estas moléculas quando sintetizadas, pode haver a formação de misturas destas conformações, ou ainda, pode haver interconversão das estruturas quando em meio biológico, uma vez que a barreira de energia entre essas conformações é relativamente pequena ( $5,1 \text{ kcal.mol}^{-1}$ , conforme apresentado na seção 4.1.1 e mesma pode ser ainda menor em meio biológico devido a ação do efeito solvente). Outro fator importante a ser considerado é a necessidade ou não da otimização da geometria completa das moléculas. Percebeu-se por estas análises que não é necessária a otimização completa da geometria, reduzindo-se assim, de maneira significativa, o tempo e a demanda computacional requerida por estes estudos e, também, por tornar possível a investigação de moléculas maiores em níveis de teoria mais elevados, garantindo a qualidade do modelo obtido. Porém, vale destacar que neste trabalho foram consideradas as geometrias totalmente otimizadas, uma vez que estas já haviam sido calculadas.

#### 4.4. Alinhamento Molecular

O alinhamento molecular das pirimidinas substituídas **5a-q** foi feito de duas maneiras distintas como explicado na seção 3.7 da metodologia, página 25, sendo uma com o programa ALINHAMOL, desenvolvido em nosso grupo, e outra com o programa MIPSIM, desenvolvido pelo grupo do Sanz. Os resultados obtidos para ambos alinhamentos estão apresentados na Figura 4.15.



**Figura 4.15:** Alinhamento molecular das pirimidinas substituídas **5a-q** obtido com o programa (a) ALINHAMOL com a molécula **5h** como referência e com o programa MIPSIM com as moléculas de referência (b) **5h**, (c) **5ad** e (d) **5idd**.

Pode-se perceber que há bastante diferença entre o alinhamento obtido com os programas ALINHAMOL e o MIPSIM (Figura 4.15a e Figuras 4.15b-d, respectivamente). Essas diferenças observadas são resultados das metodologias de cada programa para obtenção do alinhamento.

O alinhamento obtido com o programa ALINHAMOL (Figura 4.15.a) mostrou-se bastante satisfatório devido a sua baixa complexidade de utilização e o perfil do alinhamento obtido, onde todas as moléculas mostraram ter uma geometria de equilíbrio bastante próxima umas das outras e com a molécula de referência. Vale a pena ressaltar

que a utilização desta metodologia de alinhamento molecular deve ser utilizada para grupo de moléculas estruturalmente semelhantes, uma vez que este alinhamento se dá através da posição dos átomos.

O alinhamento realizado com o programa MIPSIM (Figuras 4.15b-d) mostrou que o alinhamento obtido através dos campos moleculares, além de requerer uma alta demanda computacional, é bastante sensível a escolha da molécula de referência utilizada. Pode-se perceber que cada escolha da molécula de referência, obteve-se um padrão distinto de alinhamento. Devido às grandes diferenças espaciais entre as moléculas, a comparação da SPE torna-se bastante complexa, uma vez que as caixas necessárias para cálculo do PE serão bastante grandes, para que contenha toda a informação relevante necessária para descrever o sistema, havendo desta forma um aumento na demanda computacional requerida.

Após o alinhamento molecular foram feitas análises quimiométricas utilizando-se a análise de componentes principais e a regressão pelo método dos mínimos quadrados parciais, das pirimidinas substituídas 5a-q alinhadas com o programa ALINHAMOL.

#### **4.5. Regressões da SPE das moléculas alinhadas no ALINHAMOL e calculadas no CAMOL**

Após o estudo do comportamento das cargas atômicas das pirimidinas substituídas, fez-se o cálculo da Superfície do Potencial Eletrostático (SPE) destas moléculas com diferentes tamanhos de caixa, espaçamento entre os pontos (resolução da caixa) e valores de corte da SPE definidos no programa CAMOL, conforme explicados nos itens 3.7.3. e 3.8 da metodologia, respectivamente.

As caixas utilizadas para o cálculo da SPE estão apresentadas na Tabela 3.2. O tamanho indica a distância entre a borda da caixa e ponto mais extremo das moléculas nas direções x, y e z após o alinhamento molecular com o programa ALINHAMOL. O espaçamento da caixa corresponde à distância entre os pontos da grade onde serão realizados os cálculos do Potencial Eletrostático. O método BP86/6-31G\*\* empregado foi determinado segundo análise dos resultados obtidos para a molécula **5h**,

apresentados na seção 4.1 página 40.

#### *4.5.1. Determinação do Valor de Corte Independente do Sinal (VCIS)*

Os valores de corte do programa CAMOL podem facilitar o estudo de regressão da SPE através da redução do número de variáveis, como explicado no item 3.8 da metodologia, página 30. Neste ponto, deve-se enfatizar que é importante a escolha de um valor de corte adequado para que não haja perda de informação relevante, ou seja, para que não sejam excluídos pontos importantes da superfície, evitando, desta forma, a obtenção de modelos QSAR com baixa qualidade. A vantagem da utilização destes valores de corte é a redução no número de variáveis a serem analisadas, uma vez que há grande dependência entre o tamanho do sistema, o número de variáveis e a demanda computacional requerida para analisar e processar estes dados. Assim, a eliminação prévia de variáveis pouco importantes para a descrição do sistema possibilita o estudo de sistemas maiores ou a utilização de caixas com uma melhor resolução entre os pontos.

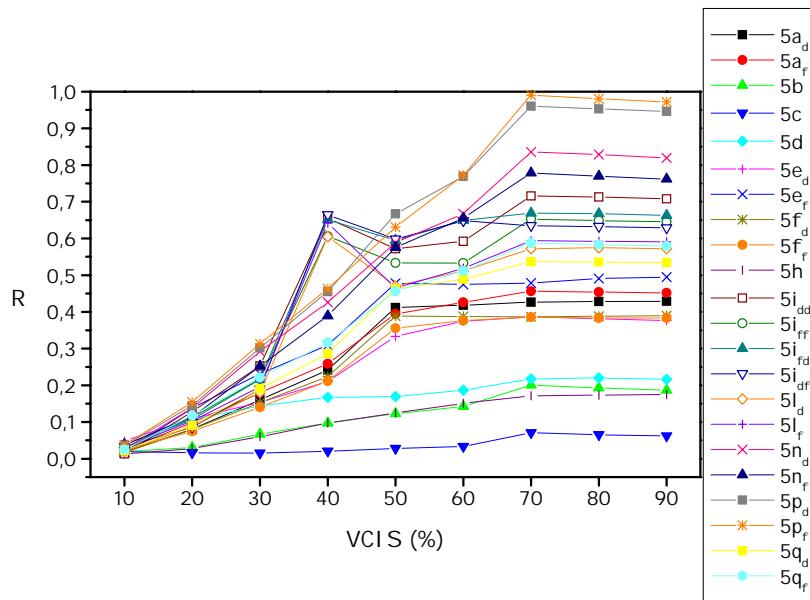
Para se escolher o valor de corte da SPE das pirimidinas utilizado no programa CAMOL, fez-se, inicialmente, um estudo da influência das diferenças provocadas por estas variáveis fixando-se o valor de corte dependente do sinal (VCDS) em 100% e variando-se o valor de corte independente do sinal (VCIS) entre 0 e 100%. A razão pela escolha do parâmetro VCDS=100% ficará claro na seção 4.5.2 a seguir. Esta variável foi analisada a partir de uma função resposta, R, dada pela equação 4.1.

$$R = \sqrt{(PC1_0 - PC1_i)^2 + (PC2_0 - PC2_i)^2 + (PC3_0 - PC3_i)^2} \quad \text{equação 4.1,}$$

onde  $PC1_0$ ,  $PC2_0$  e  $PC3_0$  são os escores nas PC's com  $VCIS = 0\%$  e  $PC1_i$ ,  $PC2_i$  e  $PC3_i$  são os escores nas PC's com  $VCIS$  variando de 10% a 90%.

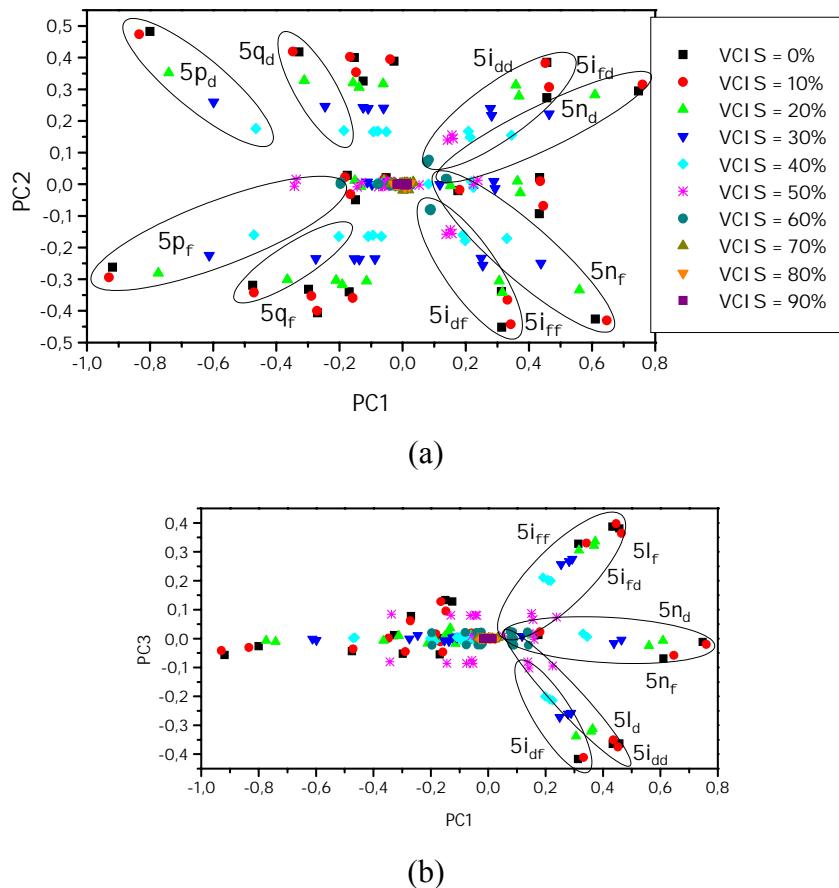
O gráfico da Figura 4.16 mostra a variação da função resposta, R, em função do valor do VCIS da SPE calculada na caixa 5 (ver Tabela 3.2). Esta função resposta, R, representa a distância euclidiana dos escores de PC1, PC2 e PC3 calculados a partir da SPE das pirimidinas substituídas 5a-q com o  $VCIS = 0\%$  e aqueles escores variando  $10\% < VCIS < 90\%$ . Um  $VCIS = 100\%$  seria a exclusão de todos os pontos do PE, uma vez que o programa exclui todos os pontos que tenham o valor do PE, logo não deve ser

considerado.



**Figura 4.16:** Gráfico da função resposta  $R$  em função do VCIS, da caixa 5 com VCDS = 100% e 10% < VCIS < 90%.

Observa-se na Figura 4.16, que as moléculas que possuem o grupo metoxi em  $R_1$  (anel da esquerda, diedro C23C12C2N3) apresentam um ponto de máximo local na função resposta em torno de 40% do VCIS. Haja vista que esses valores de corte podem ocasionar diferenças nos modelos de regressões, fez-se uma análise de PCA da SPE para cada um dos VCIS conforme apresentados no gráfico (comparativo) da Figura 4.17, para se entender melhor os efeitos da variação dos valores de corte através da mudança de padrões nesta figura. Os gráficos (individuais) dos escores de cada PCA para cada VCIS encontram-se na Figura B.1, página 116.



**Figura 4.17:** Superposição dos gráficos dos escores de (a) PC1 x PC2 e (b) PC1 x PC3 da SPE (BP86/6-31G\*\*) das pirimidinas **5a-q** totalmente otimizadas com  $0\% < \text{VCIS} < 90\%$  e  $\text{VCDS} = 100\%$  para a caixa 5.

Pode-se perceber então, na Figura 4.17 que o padrão dos gráficos dos escores desta série de moléculas muda de acordo com a variação do VCIS para PC1 x PC2 e para PC1 x PC3, conforme previsto pela análise do gráfico da função resposta R (Figura 4.16). Vê-se que há duas inversões no padrão do gráfico: i) quando se utiliza um VCIS  $> 40\%$  e ii) quando se utiliza um VCIS  $> 70\%$  (ver Figura B1 página 117). Estas diferenças no padrão são esperadas tendo em vista que a exclusão de uma quantidade maior de pontos, com o VCIS, pode tornar a SPE das moléculas mais parecidas entre si (ou menos dissimilares). Em particular, para as duas primeiras componentes principais, há uma modificação na separação das estruturas onde o aumento no valor de corte (VCIS) diminui a separação entre os objetos (as moléculas) tanto nos escores de PC1 como de PC2, dificultando assim, ser estabelecida uma relação para a separação destes objetos, conforme pode ser observado no gráfico da Figura 4.17a. Já a terceira

componente (gráfico da Figura 4.17b) foi bastante alterada com o aumento do número de pontos excluídos como pode ser melhor observado nos gráficos da Figura B1 (página 117). A terceira componente principal, que inicialmente está separando as moléculas segundo a presença ou ausência do grupo metoxi na posição R<sub>1</sub>, posteriormente (após VCIS = 40%) passa a descrever a presença ou ausência do grupo metoxi na posição R<sub>3</sub> de substituição da molécula e, por último, após um VCIS = 70%, passa a descrever as moléculas segundo a presença do átomo de cloro na posição R<sub>2</sub>.

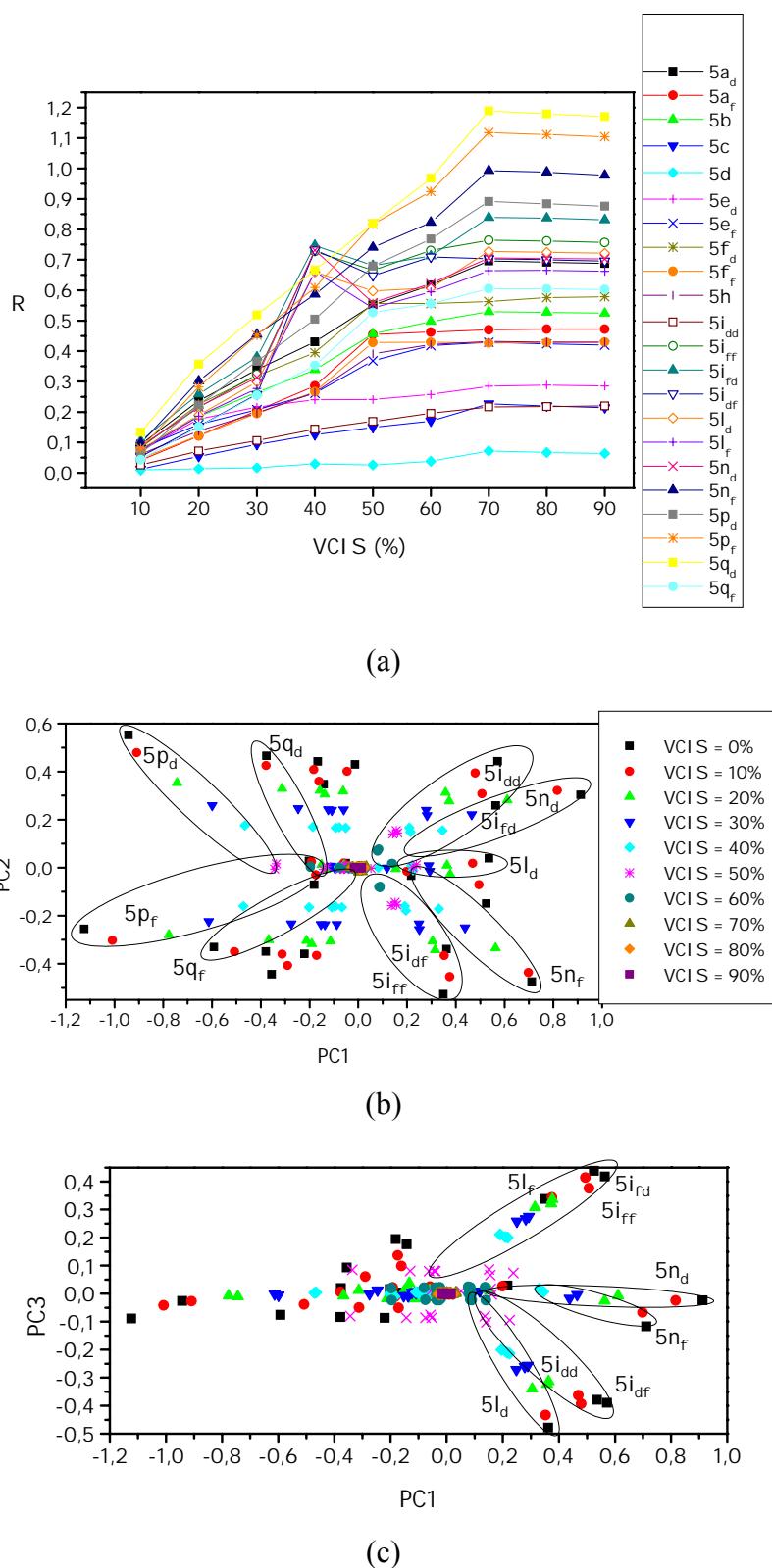
É importante observar que após a diminuição do número de variáveis o tempo gasto na análise quimiométrica também foi reduzido, como pode ser visto na Tabela 4.6. Então, para sistemas muito grandes ou para caixas com uma alta resolução, a utilização do VCIS pode alterar de maneira significativa o tempo de cálculo e, assim, tornar o estudo mais viável.

**Tabela 4.6:** Número de pontos utilizados e o tempo de CPU na PCA da SPE da caixa 5 das pirimidinas **5a-q** com o VCDS = 100% e 0% < VCIS < 90% .

VCIS (%)	Número de Pontos	Tempo de CPU (s)*	VCIS(%)	Número de Pontos	Tempo de CPU (s)
0	18626	156	50	271	1
10	17150	138	60	90	< 1
20	5434	43	70	26	< 1
30	1966	7	80	16	< 1
40	744	3	90	6	< 1

\* Cálculos feitos em um microcomputador pessoal, Altron XP 2,8 MHz, 512MB de memória ram e 40GB de HD.

A fim de avaliar se o padrão obtido para a caixa 5 é particular do pequeno limite da caixa em relação à borda da molécula ou se independe do tamanho da caixa, fez-se a mesma análise para a caixa 8 (maior caixa e melhor resolução utilizada neste trabalho). Os gráficos da Figura 4.18 mostram que não há modificação na relação entre a função resposta R e o valor do VCIS utilizado para esta segunda caixa, bem como para o padrão observado no gráfico dos escores das três primeiras componentes principais da SPE das pirimidinas. Os gráficos individuais dos escores de PC1, PC2 e PC3 podem ser vistos na Figura B2, página 123.



**Figura 4.18:** Gráficos da caixa 8 da (a) da função resposta  $R$  em função do VCIS com  $\text{VCDS} = 100\%$  e  $10\% < \text{VCIS} < 90\%$ , e dos escores de (b) PC1 x PC2 e (c) PC1 x PC3 da SPE com  $\text{VCDS} = 100\%$  e  $0\% < \text{VCIS} < 90\%$ .

Logo, pode-se perceber que o efeito provocado pela variação no VCIS é comum a diferentes valores de tamanho de caixa, de onde se conclui viável, do ponto de vista computacional, a utilização de uma caixa grande devido à remoção de uma grande quantidade de variáveis com pouca informação relevante para descrever o sistema. Observe que não é possível para o usuário prever, a priori, com segurança para qualquer sistema molecular, o tamanho mínimo dessas caixas, porém o uso desses valores mostra-se uma maneira interessante de minimizar esse problema.

A Tabela 4.7 mostra que há um grande aumento no tempo de CPU durante o cálculo quimiométrico da SPE das pirimidinas 5a-q para a caixa 8, quando comparados com os tempos requeridos pela caixa 5 (Tabela 4.6) e que a utilização do valor de corte reduz de maneira significativa o número de pontos. Assim, o tempo gasto durante o cálculo da PCA, caiu de aproximadamente 10 minutos para menos de 1 segundo.

**Tabela 4.7:** Tempo de CPU utilizado na PCA da SPE da caixa 8 das pirimidinas **5a-q**  
com o VCDS = 100% e 0% < VCIS < 90% .

VCIS (%)	Número de Pontos	Tempo de CPU (s)*	VCIS (%)	Número de Pontos	Tempo de CPU (s)
0	74453	640	50	271	1
10	22425	192	60	90	< 1
20	5494	46	70	25	< 1
30	1966	8	80	15	< 1
40	745	3	90	6	< 1

\*Cálculos feitos em um microcomputador pessoal, Altlon XP 2,8MHz, 512MB de memória ram e 40GB de HD.

Percebe-se também nos gráficos da Figura 4.18 e no tempo de cálculo (Tabela 4.7) que a escolha do valor do VCIS pode gerar matrizes de variáveis equivalentes para diferentes tamanhos de caixa, uma vez que pode haver a exclusão de grande quantidade de pontos.

Os nossos resultados não apresentam um valor único do VCIS a ser utilizado, entretanto, eles indicam como razoável, para este sistema (as pirimidinas substituídas), um VCIS igual a 20%, uma vez que este valor aponta para uma relação, satisfatória, entre o número de variáveis (conseqüentemente a demanda computacional) e os

padrões dos gráficos dos escores da PCA das Figuras 4.17 e 4.18, sendo este valor, anterior as modificações ocorridas nos gráficos dos escores.

Em seguida foi adotado o mesmo procedimento para determinação do valor de VCDS, conforme pode ser visto a seguir.

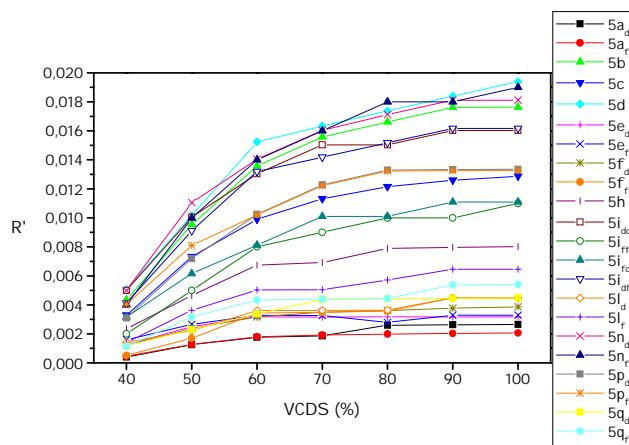
#### *4.5.2. Determinação do Valor de Corte Dependente do Sinal (VCDS)*

Como explicado na seção 3.9 da metodologia, página 31, a redução no número de variáveis é menor para o Valor de Corte Dependente do Sinal (VCDS), isto é, o VCDS é menos efetivo na redução do número de variáveis que o VCIS, uma vez que o VCDS considera o sinal do PE no ponto comparado, enquanto que o VCIS não. Porém, também é necessária a escolha de um valor máximo que não provoque perda de informação relevante para a descrição do sistema. Então, se verificou o comportamento da função resposta,  $R'$  (equação 4.2), em função do VCDS (Figura 4.19) e do gráfico dos escores das três primeiras componentes principais da SPE das pirimidinas 5a-q (Figura 4.20) para as caixas 5 e 8 (ver Tabela 3.2).

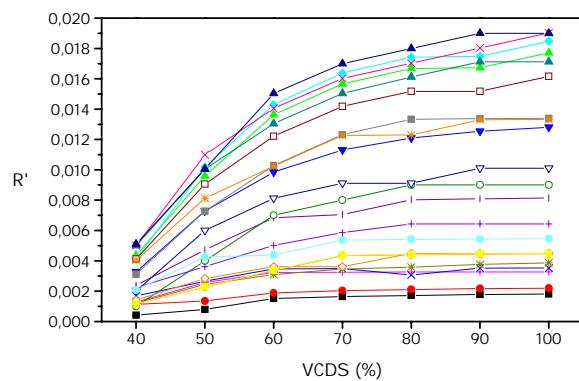
$$R' = \sqrt{(PC1_{30} - PC1_i)^2 + (PC2_{30} - PC2_i)^2 + (PC3_{30} - PC3_i)^2} \quad \text{equação 4.2,}$$

onde  $PC1_{30}$ ,  $PC2_{30}$  e  $PC3_{30}$  são os escores das PC's com VCDS = 30% e  $PC1_i$ ,  $PC2_i$  e  $PC3_i$  são as PC's com VCDS variando entre 40% e 100%.

Pode-se notar que o valor inicial do VCDS é de 30%, pois, como explicado anteriormente, este valor deve ser sempre maior que o valor do VCIS que, como visto anteriormente, foi indicado ser igual a 20%, logo, o primeiro ponto do gráfico da função resposta,  $R'$ , deve ser com VCDS = 40%.



(a)

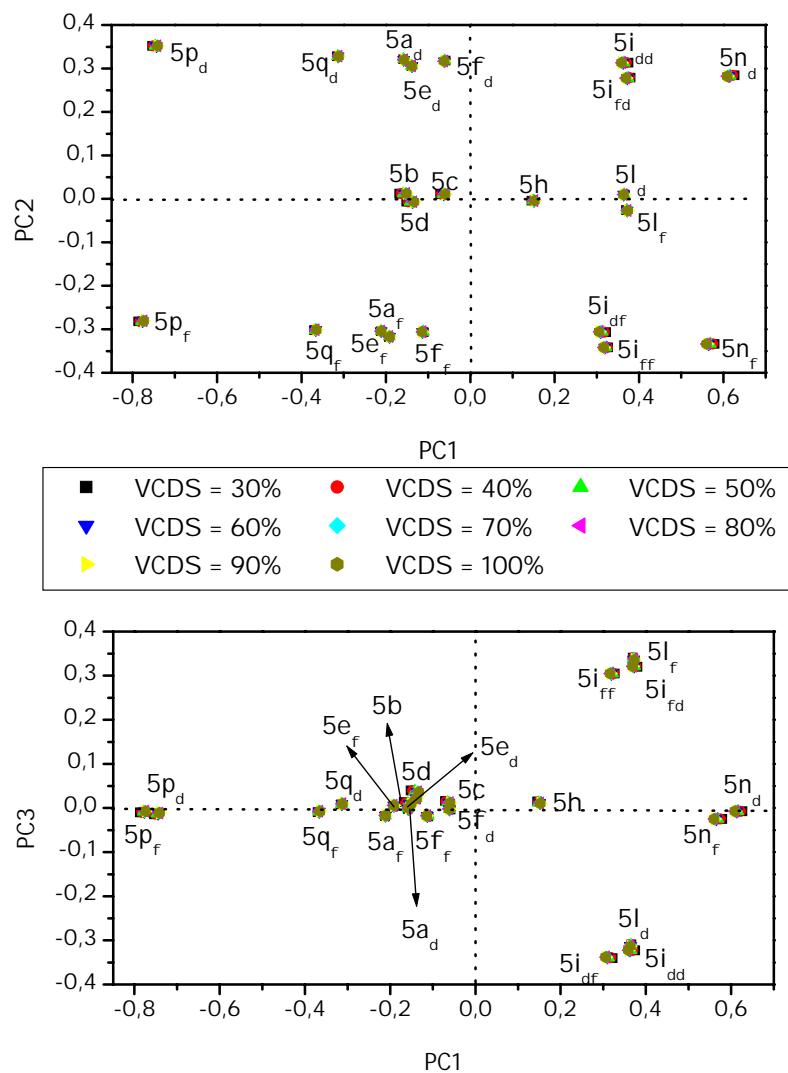


(b)

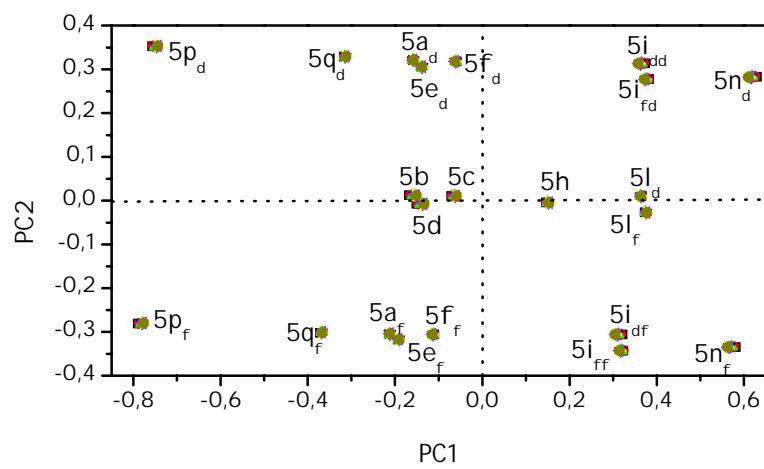
**Figura 4.19:** Gráfico da função resposta  $R'$  em função do VCDS, com VCIS = 20% e  $40\% < \text{VCDS} < 100\%$  da (a) caixa 5 e (b) caixa 8.

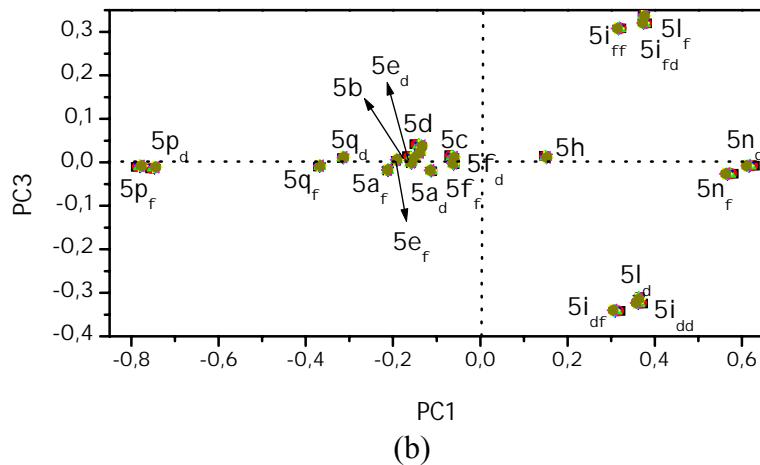
Pode-se perceber nos gráficos da Figura 4.19 que a função  $R'$  apresenta, em geral, um comportamento crescente e regular em função dos diferentes valores VCDS utilizados. Já na Figura 4.20 pode-ser ver os gráficos dos escores de PC1, PC2 e PC3 para as caixas 5 e 8.

Então, como feito para o VCIS, fez-se uma análise comparativa dos gráficos dos escores de PC1, PC2 e PC3 (Figura 4.20) para verificar se este desvio da regularidade da função  $R'$  modifica de maneira significativa o padrão dos gráficos dos escores.



(a)





**Figura 4.20:** Gráficos superpostos dos escores de PC1 x PC2 e de PC1 x PC3 da SPE das Pirimidinas Substituídas: (a) da caixa 5 e de (b) da caixa 8, com  $30\% < \text{VCDS} < 100\%$  e  $\text{VCIS} = 20\%$ .

Os gráficos (individuais) dos escores de cada PCA estão apresentados na Figura C1 e C2, páginas 128 e 131, respectivamente. A figura acima mostra a sobreposição dos gráficos dos escores das três primeiras PC's para os diferentes valores de VCDS. Desta, é possível notar que as variações provocadas no gráfico dos escores da SPE das pirimidinas 5a-q, pelos diferentes valores de VCDS, são muito menores que as variações provocadas pelos diferentes valores do VCIS (Figura 4.17 e 4.18b-c). Em outras palavras, a variação no VCDS não modifica, de maneira significativa, o padrão dos gráficos nesta análise quimiométrica. Este resultado está de acordo com o esperado uma vez que este valor de corte considera o sinal do potencial em cada ponto da superfície e, então, uma menor quantidade de pontos será excluída. Logo, os padrões observados em PC1 x PC2 e PC1 x PC3 mostram-se bastante similares para todos os valores testados. Vale a pena ressaltar que se iniciou a variação deste valor em 30% uma vez que o VCDS deve ser sempre maior que o VCIS, caso contrário os pontos excluídos serão resultado do VCIS e não do VCDS. Desta forma, quando for necessária a utilização de uma caixa grande, ou seja, no estudo de moléculas grandes, ou um pequeno espaçamento entre os pontos, isto é, uma alta resolução, a utilização de um valor grande para este parâmetro pode ser útil para a eliminação de variáveis pouco importantes para a descrição do sistema, tornando o estudo mais barato do ponto de vista computacional.

Esta diminuição no número de pontos da SPE varia pouco com o aumento do VCDS, como explicado anteriormente, e assim a variação no tempo de cálculo da PCA será pequena ou inexistente, como pode ser observado na Tabela 4.8.

**Tabela 4.8:** Tempo de CPU utilizado na PCA da SPE da caixa 5 das pirimidinas 5a-q com  $30\% < \text{VCDS} < 100\%$  e  $\text{VCIS} = 20\%$ .

VCDS (%)	Número de Pontos	Tempo de CPU (s)*	VCDS (%)	Número de Pontos	Tempo de CPU (s)*
30	6166	10	70	5496	9
40	5884	10	80	5465	9
50	5677	9	90	5446	9
60	5556	9	100	5434	9

\* Cálculos feitos em um microcomputador pessoal, Altron XP 2,8 MHz, 512MB de memória ram e 40GB de HD.

O mesmo pode ser observado quando se avalia o tempo de cálculo para a caixa 8, conforme pode ser observado na Tabela 4.9.

**Tabela 4.9:** Tempo de CPU utilizado na PCA da SPE da caixa 8 das pirimidinas 5a-q com  $30\% < \text{VCDS} < 100\%$  e  $\text{VCIS} = 20\%$ .

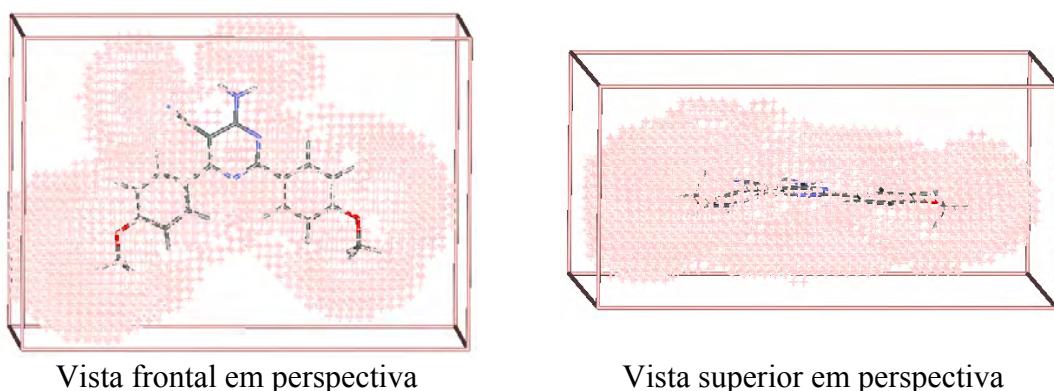
VCDS (%)	Número de Pontos	Tempo de CPU (s)*	VCDS (%)	Número de Pontos	Tempo de CPU (s)*
30	6225	55	70	5555	46
40	5944	51	80	5525	46
50	5737	49	90	5506	46
60	5616	47	100	5494	46

\* Cálculos feitos em um microcomputador pessoal, Altron XP 2,8MHz, 512MB de memória RAM e 40GB de HD.

Pode-se perceber também neste caso, que a redução no número de variáveis é pequena quando se compararam os tempos de CPU do cálculo para o mesmo valor do VCDS nas duas caixas, onde a caixa 8 terá sempre um gasto maior de tempo para realizar toda a análise, enquanto, a variação no VCIS após um certo limite gasta o

mesmo tempo para as duas caixas, como explicado no item anterior. Logo, tendo em vista as variações ocorridas nos gráficos das funções resposta R e R' (Figuras 4.16, 4.18a e 4.19, respectivamente) e nos gráficos dos escores de PC1 x PC2 (Figuras 4.18b-c e 4.20), avaliou-se que um valor de 50% para o VCDS seria um valor apropriado para esta variável.

Desta forma, conclui-se que a escolha do VCIS é mais importante do que a escolha do VCDS para poderem ser ressaltadas as maiores diferenças da SPE. Porém, é importante a utilização de um VCDS adequado para não se perder informação relevante e também para a exclusão de um número adicional de pontos, caso isto se faça necessário. Nas etapas seguintes deste trabalho utilizamos VCDS = 50% e VCIS = 20% para as análises quimiométricas de PCA e PLS, uma vez que observamos uma significativa diminuição no número de variáveis, sobretudo nas caixas maiores ou com melhor resolução. A Figura 4.21 ilustra a redução no número de pontos, ou seja, no número de variáveis que será utilizada na regressão, quando se utilizam os valores de corte do CAMOL para a molécula **5i<sub>dd</sub>**.



**Figura 4.21:** Esquema da redução no número de pontos da SPE da molécula **5i<sub>dd</sub>**, com VCDS = 50% e VCIS = 20% para o conjunto completo de moléculas.

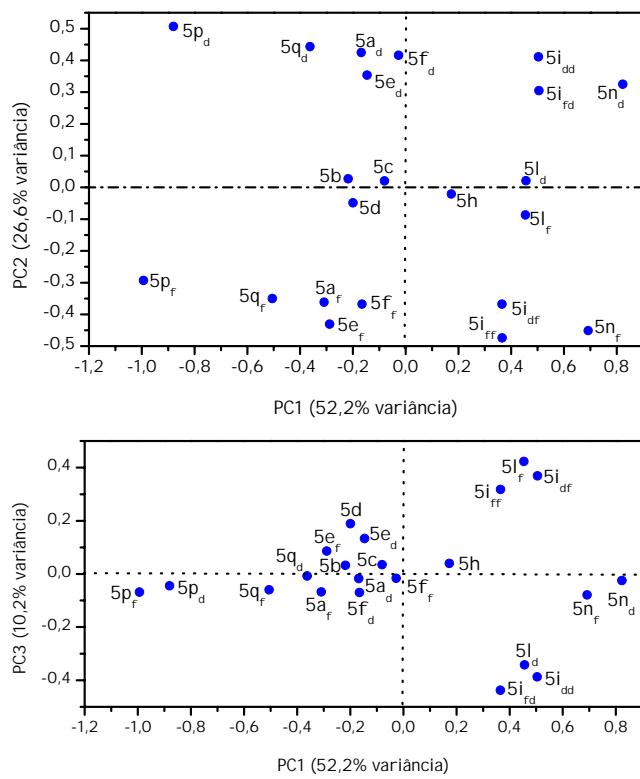
Então, foram feitas análises de componentes principais da SPE e regressões do PE contra a atividade biológica, pelo método dos mínimos quadrados parciais, das moléculas para as 16 caixas (Tabela 3.2) com VCDS = 0% e 50% e VCIS = 0% e 20%. Estes resultados estão apresentados nas próximas seções.

Nas seções 4.7.3 e 4.7.4 serão apresentados apenas os gráficos referentes as caixas 5 e 13, obtidas com os métodos BP86/6-31G\*\* e AM1, respectivamente, que

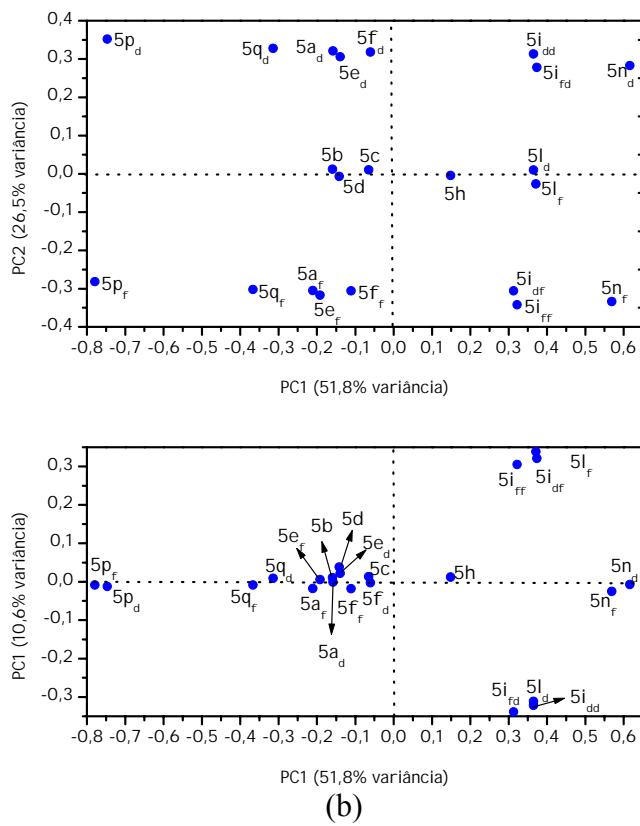
apresentaram o melhor padrão na separação dos objetos. Entretanto, é importante enfatizar que as variações nos parâmetros tamanho da caixa, espaçamento entre os pontos da grade e os valores de corte do CAMOL (seções 3.9 da metodologia página 31) não alteraram, de maneira significativa, o padrão apresentado nos gráficos dos escores, dos *loadings* e da atividade prevista, e todas as análises e conclusões obtidas a partir destes gráficos podem ser extrapoladas para todos os demais.

#### 4.5.3. Análise de componentes principais da SEP das pirimidinas totalmente otimizadas.

A análise de componentes principais da SPE das moléculas **5a-q** revelou o padrão apresentado na Figura 4.22, onde, se pode perceber a existência de alguns grupos distintos ao longo de PC1 e PC2, respectivamente.



(a)

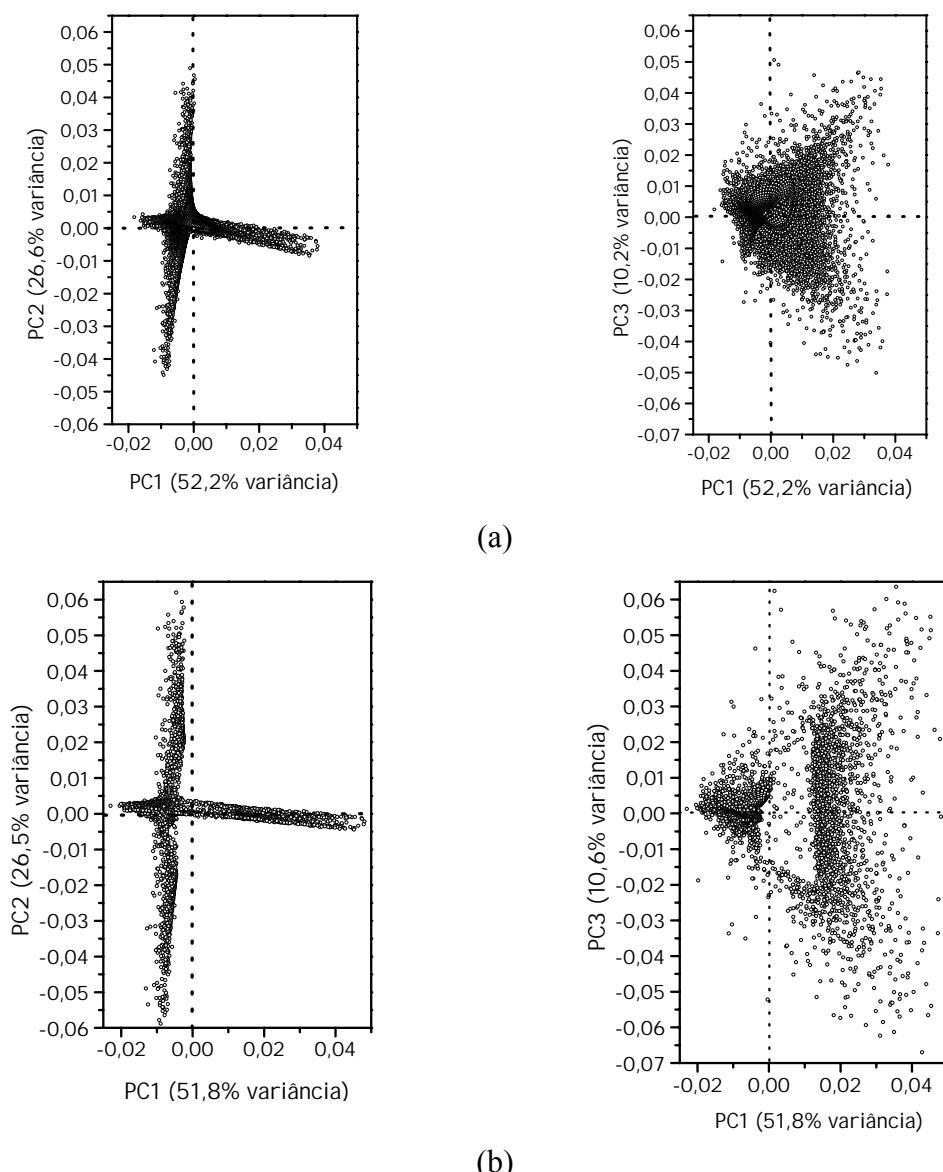


**Figura 4.22:** Gráfico dos escores da PCA da SEP (BP86/6-31G\*\*) das pirimidinas substituídas 5a-q totalmente otimizadas de PC1 x PC2 e PC1 x PC3 com (a) VCDS = VCIS = 0% e (b) VCDS = 50% VCIS = 20%.

Observa-se na Figura 4.22a que há uma separação das moléculas ao longo de PC1 (com aproximadamente 52% da variância) de acordo com o substituinte presente na molécula, onde os substituintes retiradores de elétrons (moléculas **5a**, **5b**, **5c**, **5d**, **5f**, **5e**, **5p** e **5q**, ver Tabela 4.3) encontram-se mais a esquerda, ou seja, em valores negativos de PC1, e os substituintes doadores de elétrons (moléculas **5i**, **5l** e **5n**) e a molécula **5h**, encontram-se mais a direita do gráfico, ou seja, em valores mais positivos de PC1. Percebe-se ainda a presença de algumas moléculas em PC1 próximo a zero (**5a** e **5c**), sendo estas moléculas substituídas por átomos fracamente retiradores de elétrons. Já ao longo de PC2 (aproximadamente 26% da variância) há a separação das moléculas segundo a conformação apresentada pelo grupo metoxi (moléculas **5a**, **5e**, **5f**, **5i**, **5n**, **5p** e **5q**), sendo o efeito da conformação mais significativo quando o grupo metoxi está na posição R<sub>3</sub>. Desta forma, as conformações mais estáveis, isto é, aquelas com a conformação “dentro” (seção 3.6 da metodologia, página mm) estão em valores

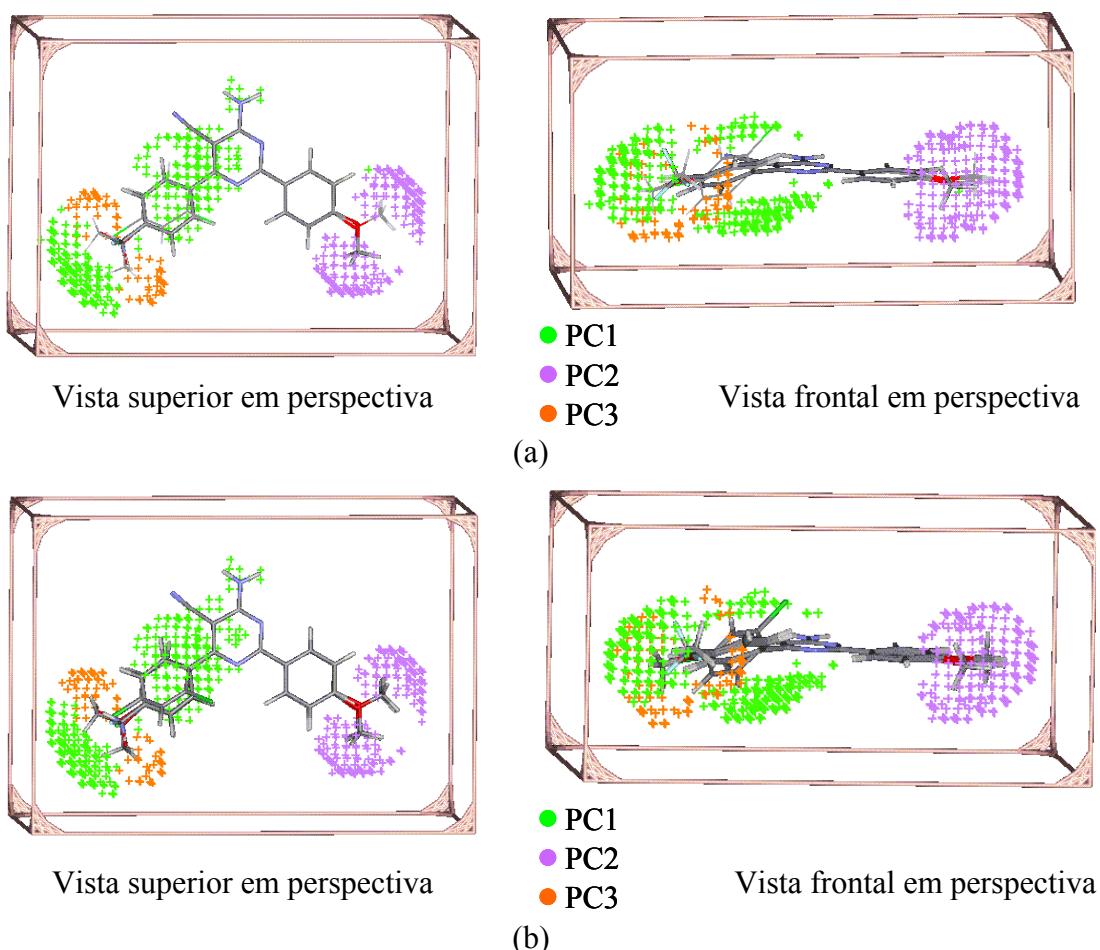
positivos de PC2. A terceira componente principal, PC3, com variância de aproximadamente 10%, também está separando as moléculas segundo a conformação do grupo metoxi, porém, quando este está presente na posição R<sub>1</sub> (**5i** e **5l**); sendo este efeito menos influente na formação do padrão e consequentemente nas diferenças entre as moléculas.

O gráfico dos *loadings* de PC1 x PC2 e PC x PC3 para os dois conjuntos de raio de corte utilizados estão apresentados na Figura 4.23.



**Figura 4.23:** Gráfico dos *loadings* de PC1 x PC2 e PC1 x PC3 da PCA da SEP (BP86/6-31G\*\*) das pirimidinas substituídas 5a-q totalmente otimizadas com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.

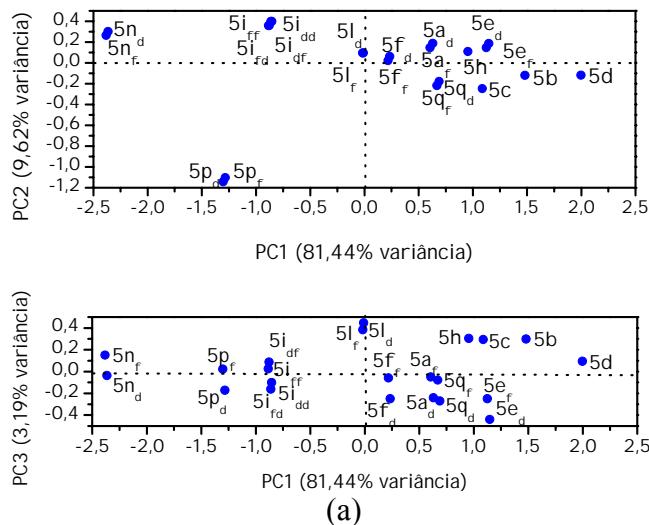
Percebe-se através da Figura 4.23 que o padrão no gráfico dos *loadings* permanece, essencialmente, o mesmo para os dois valores de corte, VCDS e VCIS, utilizados, porém, os gráficos da Figura 4.23b possuem menos pontos próximos do ponto  $(PC1, PC2, PC3) = (0,0,0)$ . Isto está de acordo com o esperado, devido à retirada prévia de pontos da SPE com a utilização dos valores de corte, os quais possuem um pequeno valor do PE, ficando claro então, que estes pontos trazem consigo pouca informação relevante para a descrição do sistema. Na Figura 4.24 estão apresentados os pontos da grade cujos *loadings* são até 60% do maior valor de *loading* (positivo e negativo) em cada componente.



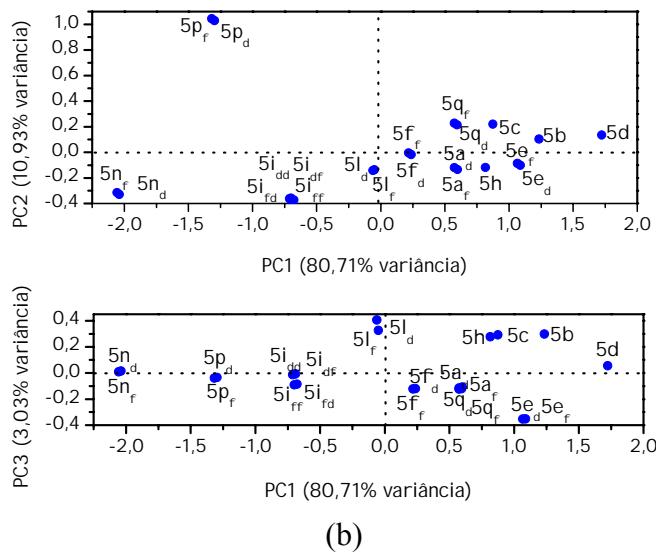
**Figura 4.24:** Localização dos pontos da SPE (BP86/6-31G\*\*) das pirimidinas substituídas **5a-q** totalmente otimizadas com o valor dos *loadings* superiores a 60% do maior e menor valor em PC1, PC2 e PC3 para (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.

Pode-se perceber nessa figura quais as regiões mais importantes na formação dos padrões dos gráficos dos escores. Percebe-se também que não há diferença significativa na quantidade de pontos para descrever estes padrões com a utilização dos valores de corte do CAMOL (Figura 4.24b), enfatizando assim, que não houve perda de informação relevante para formação do padrão com a exclusão, antecipada, de pontos da superfície feitas pelos valores de corte. Observe que este procedimento de corte de pequenos valores do PE poderá ser bastante útil no estudo de sistemas com grande número de pontos. Ainda, pode-se perceber que PC1 está sendo descrita pelo anel benzênico que se encontra fora do plano e pela presença ou ausência de substituintes na posição R<sub>1</sub>. PC2 mostra-se mais sensível à presença e a conformação do grupo metoxi na posição R<sub>3</sub>, enquanto que PC3 mostra-se sensível às conformações do grupo metoxi da posição R<sub>1</sub>, enfatizando as diferenças explicadas anteriormente.

Um outro aspecto investigado nesta dissertação diz respeito à metodologia de cálculo do potencial eletrostático. Neste sentido, as geometrias obtidas com o método DFT-BP86/6-31G\*\* foram novamente utilizadas, porém, agora, com o PE calculado pelo método AM1. A Figura 4.25 mostra os gráficos dos escores de PC1 *versus* PC2 e PC1 *versus* PC3



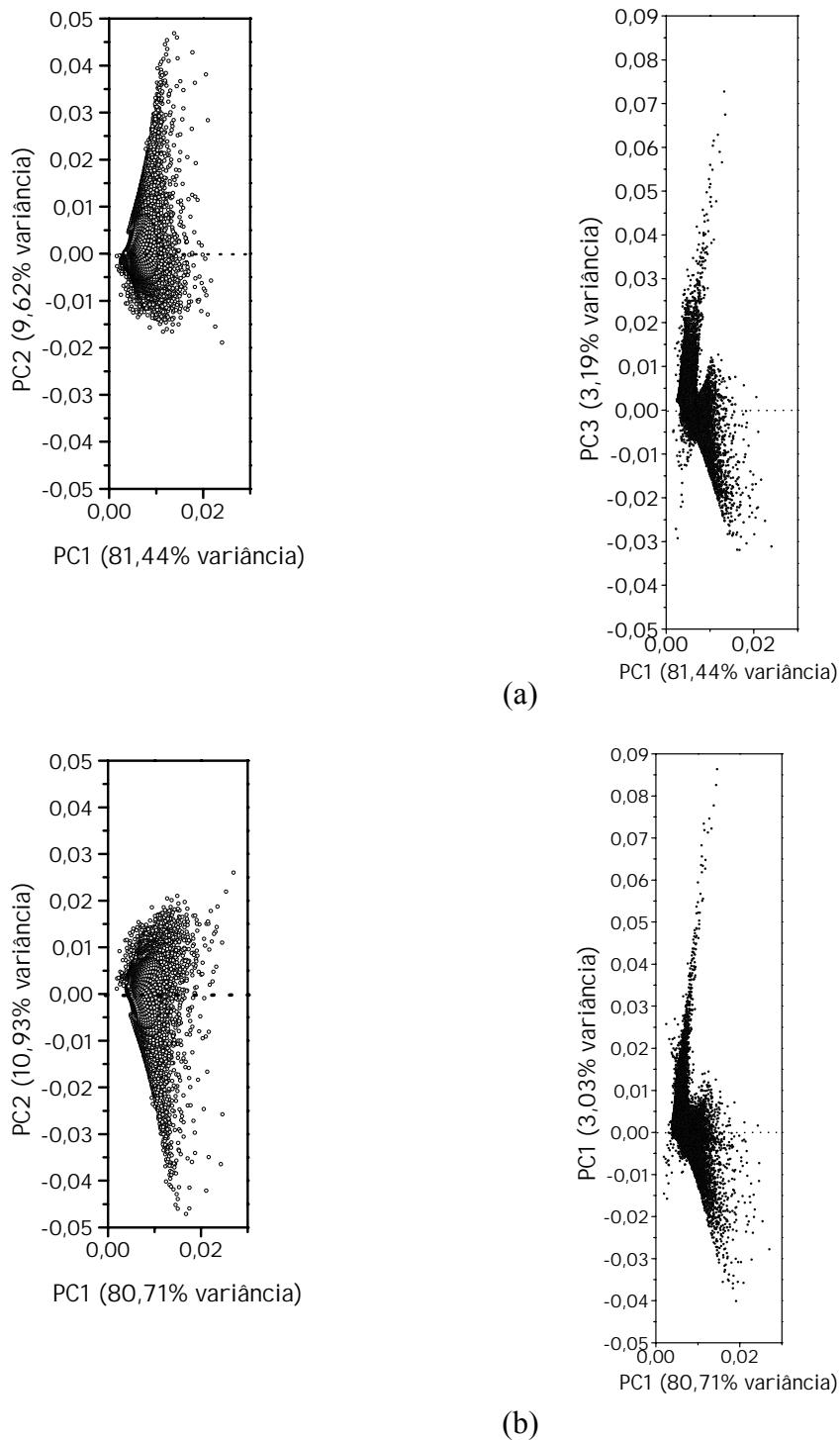
(a)



**Figura 4.25:** Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da PCA da SEP (AM1) das pirimidinas substituídas **5a-q** totalmente otimizadas de (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.

A análise da SPE das pirimidinas, calculada com o método AM1//BP86/6-31G\*\*, mostra que o principal efeito na separação das moléculas são as propriedades eletrostáticas dos substituintes, onde as moléculas que possuem grupos retiradores de elétrons na posição R<sub>1</sub> (**5a**, **5b**, **5d**, **5c**, **5e**, **5f**, e **5q**) e a molécula **5h** encontram-se em valores positivos de PC1 (aproximadamente 81% da variância) com exceção da molécula **5p**, e as moléculas com grupo doadores de elétrons (**5i**, **5l** e **5n**) na mesma posição, encontram-se em valores negativos de PC1. A maior diferença observada quando comparamos os gráficos da Figura 4.22 com os gráficos da Figura 4.25 é a separação provocada em PC2 (aproximadamente 10% da variância), onde, neste caso, não temos mais a separação das conformações do grupo metoxi na posição R<sub>3</sub>, mas sim a natureza do substituinte da posição R<sub>1</sub>, ainda que de uma maneira menos efetiva, onde as moléculas com grupos retiradores de elétrons encontram-se em valores mais negativos de PC2. Já PC3 (aproximadamente 3% da variância) não define uma separação clara entre os objetos. A inversão de fase, ocorrida no padrão dos gráficos dos escores da Figura 4.25a e b não invalida, ou contradiz, os resultados, uma vez que é compensada pela inversão no gráfico dos *loadings*, garantindo, assim, que os resultados são equivalentes, conforme pode ser visto a seguir.

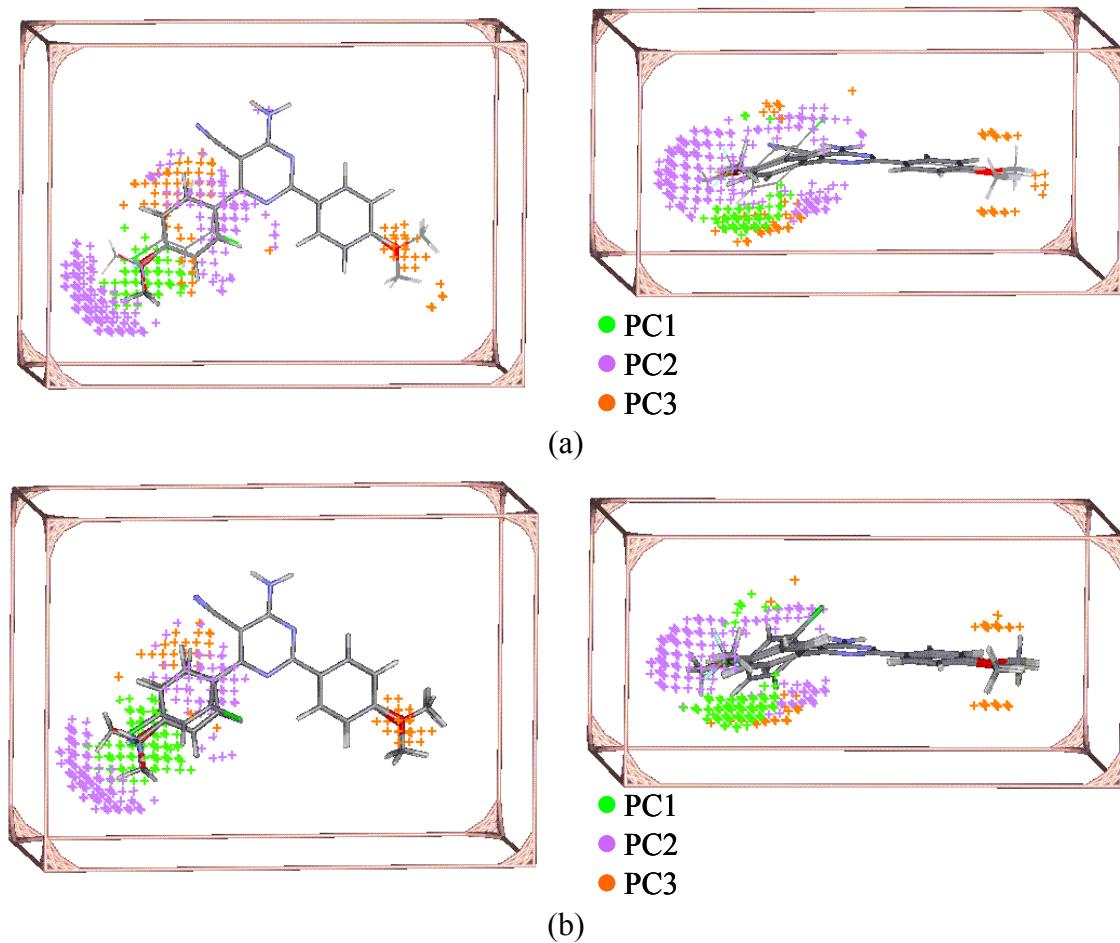
Os gráficos dos *loadings* associados aos gráficos dos escores da Figura 4.25 podem ser observados na Figura 4.26 abaixo.



**Figura 4.26:** Gráfico dos pesos de PC1 x PC2 da PCA da SEP (AM1) das pirimidinas substituídas 5a-q totalmente otimizadas com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.

Nota-se nesses gráficos que a utilização dos VCDS e VCIS, não modifica, de maneira significativa, os gráficos dos *loadings* e dos escores do PE DFT (Figuras 4.20 e 4.23). Porém, a mudança de método, como por exemplo, do DFT-BP86/6-31G\*\* para o AM1 leva a resultados que são bastante diferentes quando comparados com os gráficos dos *loadings* de PC1 x PC2 e PC1 x PC3 da Figura 4.23. É importante ressaltar que os *loadings* em PC1, neste caso, apresenta apenas valores positivos, identificando assim que não há diferença significativa nos objetos.

A Figura 4.27 mostra a região na superfície molecular onde se encontram os principais pontos da SPE na análise de PCA, ou seja, os pontos que possuem até 60% do maior valor de *loadings* (positivo e negativo) nas três primeiras componentes principais



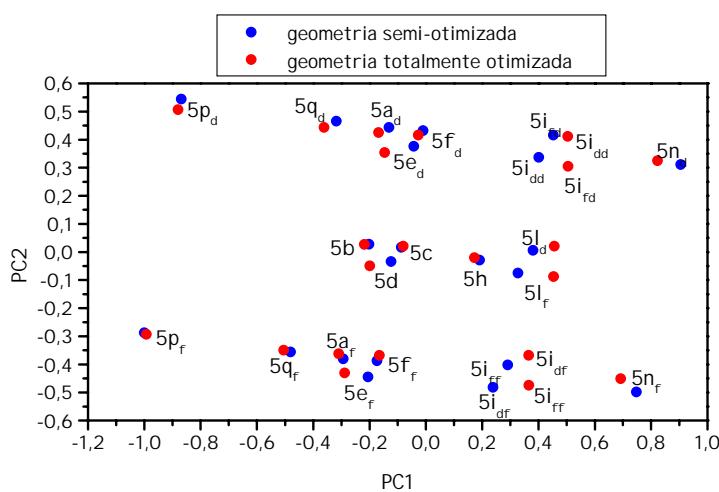
**Figura 4.27:** Localização dos pontos da SPE (AM1//BNP86/6-31G\*\*) das pirimidinas substituídas **5a-q** totalmente otimizadas com o valor dos *loadings* e PC1, PC2 e PC3 superiores a 60% do maior e menor valores em cada PC para (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS = 20%.

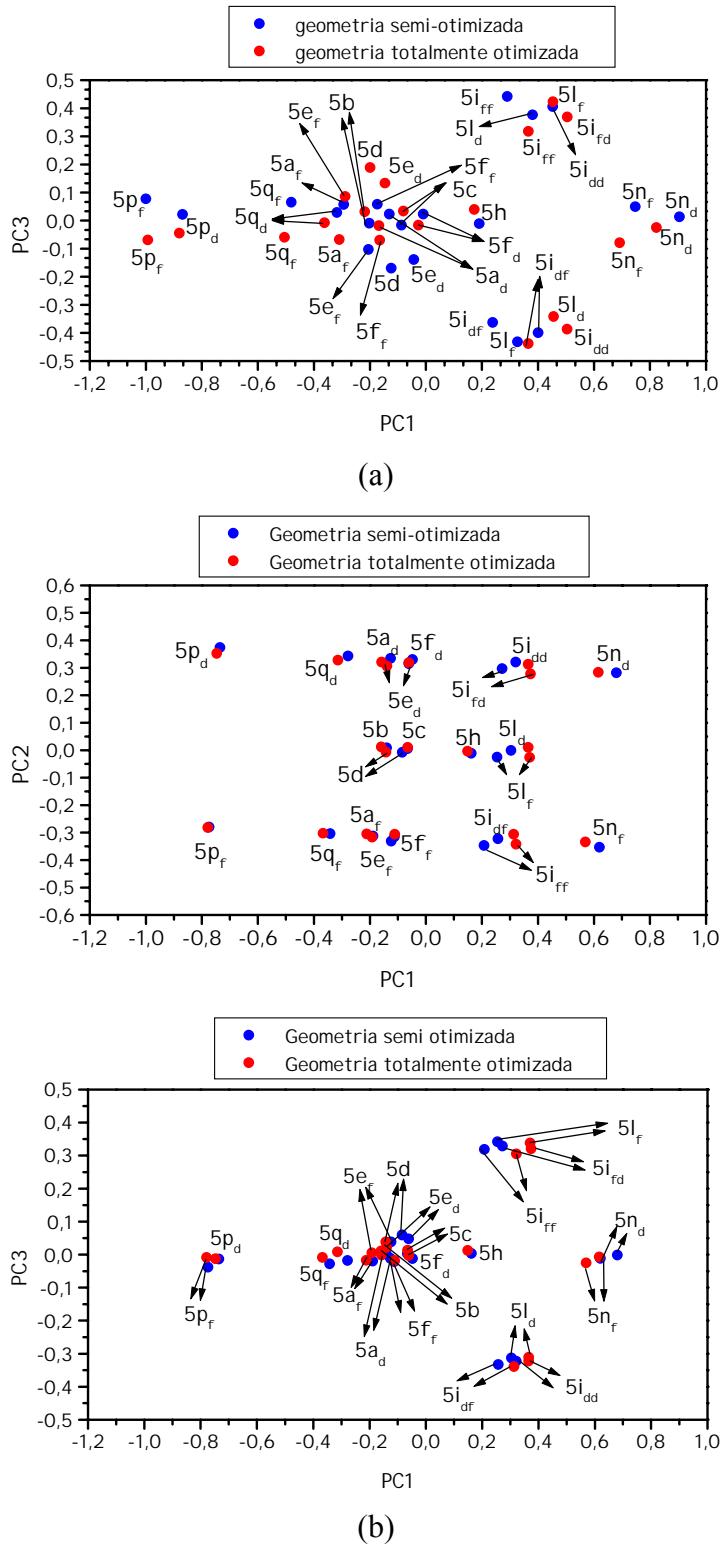
Percebe-se na Figura 4.27 que há pouca diferença na localização dos pontos da SPE em função dos vaores de corte, tal como ocorrido nas Figuras 4.26. Porém, como apresentado nas análises anteriores, o padrão da superfície é bastante diferente quando comparado com as superfícies da Figura 4.24.

Desta forma, pode-se perceber, que há dependência significativa na estimativa da propriedade SPE com o nível de cálculo empregado. Dentre os dois métodos empregados no estudo da SPE na série de compostos pirimidínicos, o método que melhor está separando os objetos é o método BP86/6-31G\*\*. Este padrão já havia sido observado no estudo da pirimidina **5h**. Então, posteriormente, foi feita uma comparação da PCA das SPE para as pirimidinas semi-otimizadas e totalmente otimizadas, conforme pode ser visto no item 4.5.4, para verificar se há grande diferença no padrão dos objetos como previsto na comparação da PCA das cargas CHelpG (seção 4.3.2).

#### 4.5.4. Comparação da análise de componentes principais da SEP das pirimidinas semi-otimizadas com as totalmente otimizadas

Para se verificar a influência da otimização da geometria completa fez-se a comparação da PCA da SPE das pirimidinas **5a-q** com a geometria semi-otimizada e totalmente otimizada, conforme pode ser visto na Figura 4.28.





**Figura 4.28:** Sobreposição dos gráficos dos escores de PC1 x PC2 e PC1 x PC3 da SEP (caixa 5) das pirimidinas substituídas com a geometria semi-otimizada e totalmente otimizada com (a) VCDS = VCIS = 0% e (b) VCDS = 50% e VCIS=20%.

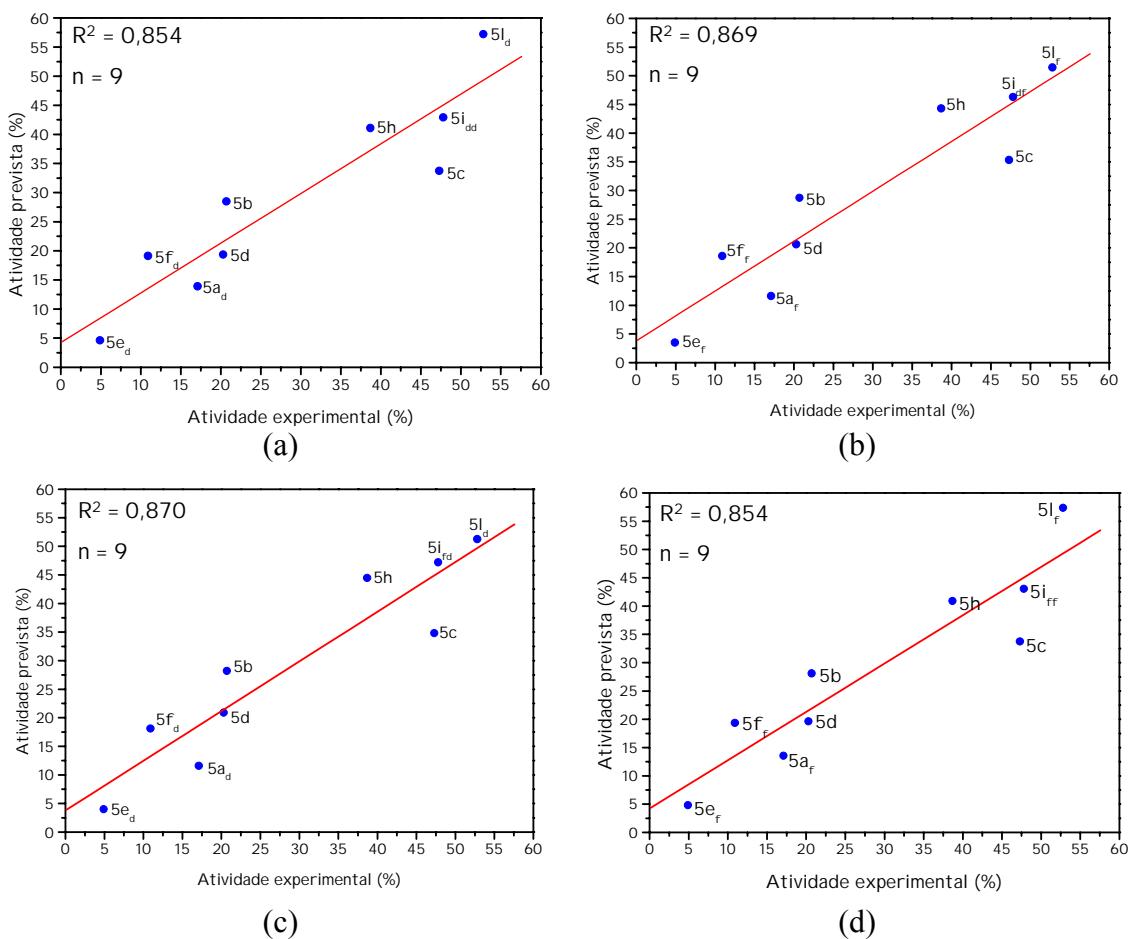
Pode-se verificar na Figura 4.28 que os gráficos dos escores de PC1 x PC2 e PC1 x PC3 para ambos valores do VCDS e VCIS são semelhantes, não havendo diferença significativa após a otimização da geometria completa, uma vez que os objetos aparecem aos pares ou em valores muito próximos em cada componente. Desta forma, concluiu-se que o efeito da otimização completa da geometria molecular também não modifica de maneira significativa a SPE da molécula, podendo, quando necessário, então ser adotado o procedimento de otimização parcial de moléculas, ou seja, otimização do esqueleto molecular e posterior adição dos substituintes otimizados separadamente, como explicado no item 3.2.2. É importante enfatizar que este procedimento de otimização parcial da geometria pode ser útil no estudo de moléculas grandes, as quais não poderiam ter a sua geometria totalmente otimizada com métodos *ab initio* (principalmente os correlacionados) ou DFT, devido ao alto custo computacional como, por exemplo, em moléculas de proteínas ou estruturas de ligantes interagindo com o sítio do seu bioreceptor.

Por fim, foi feita uma regressão PLS da SPE contra a atividade biológica destas moléculas para a determinação de um modelo de QSAR-3D, apresentada a seguir.

#### *4.5.5. PLS da SEP das pirimidinas totalmente otimizadas – Determinação de um modelo de QSAR-3D*

A Regressão pelo método dos mínimos quadrados parciais, PLS, da SPE das pirimidinas **5a-I** contra a atividade antiinflamatória, foi feita para quatro grupos distintos, isto é, de acordo com a conformação do grupo metoxi destas moléculas (ver seção 3.6, páginas 23).

Os gráficos da Figura 4.29 mostram a atividade medida (experimental) versus à atividade prevista para estes quatro grupos.

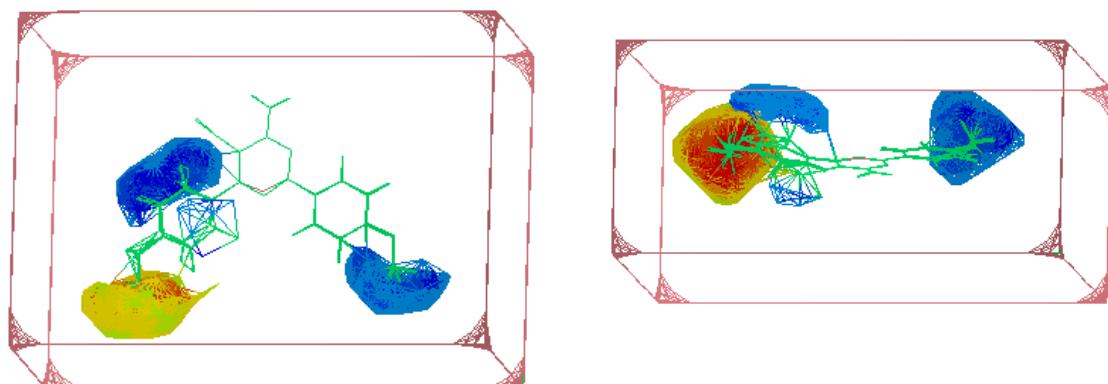


**Figura 4.29:** Gráfico da atividade experimental *versus* a prevista da PLS com VCDS = VCIS = 0% das pirimidinas com o grupo metoxi (a) dentro, (b) fora/dentro, (c) dentro/fora e (d) fora.

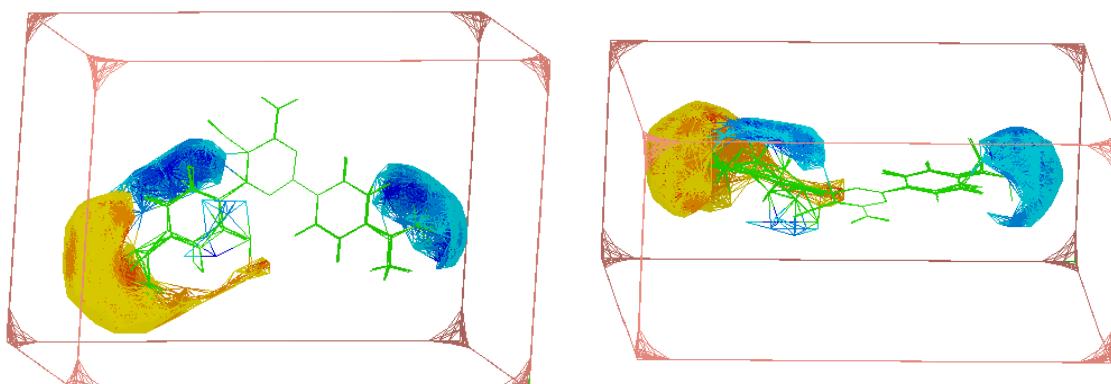
Pode-se perceber através da análise da regressão, que a atividade biológica prevista pelo modelo, aponta para o conjunto “fora/dentro” como o modelo mais adequado para se prever a atividade biológica desta série, com uma variância de aproximadamente 87%. Porém, percebe-se que para os quatro conjuntos analisados, os modelos de regressão possuem uma variância bastante próxima, indicando que os modelos obtidos são equivalentes, ou seja, que o modelo obtido não depende, de maneira significativa, da conformação destas moléculas. Em seguida, analisaram-se os mapas dos coeficientes obtidos da PLS entre o PE e a atividade antiinflamatória (“campos moleculares”). Na verdade, estes campos foram gerados utilizando-se um gradiente de cores associado aos pontos da grade cujos coeficientes são positivos (ou máximos) indicados pelas cores “quentes” (vermelho, laranja e amarelo) e negativos (ou

mínimos) indicados pelas cores "frias" (azul escuro, azul claro e verde). Em particular, foram utilizados valores de coeficientes maiores que 60% do maior coeficiente positivo ou negativo, conforme é apresentado na Figura 4.30.

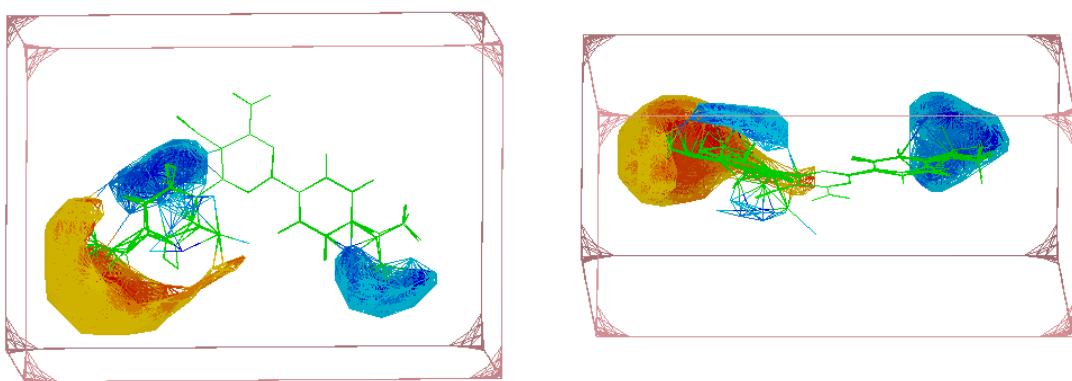
Neste ponto, vale a pena ressaltar que as regiões identificadas como importantes, as quais são visualizadas nestes mapas, são as regiões onde modificações estruturais devem ser feitas para se alterar a atividade biológica.

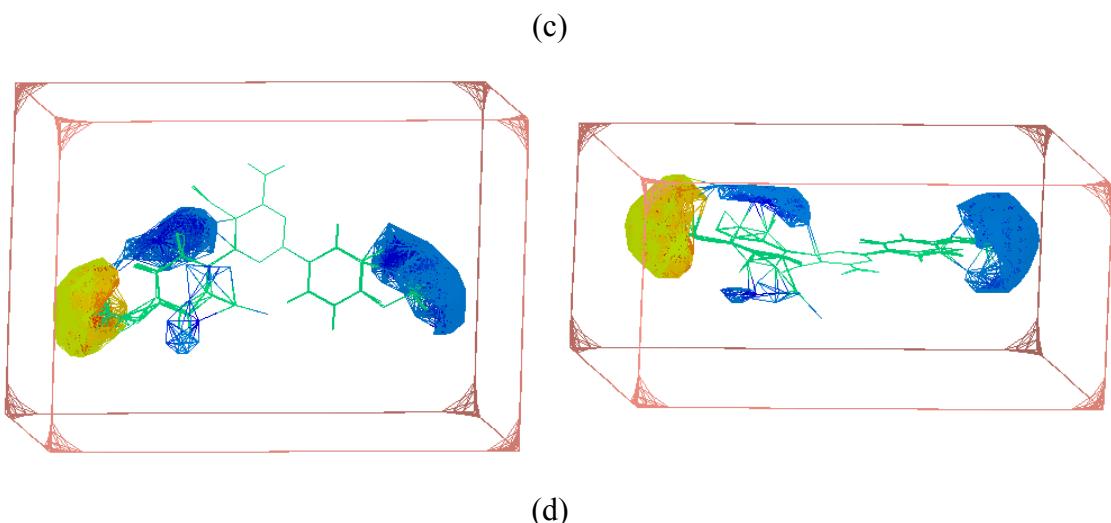


(a)



(b)

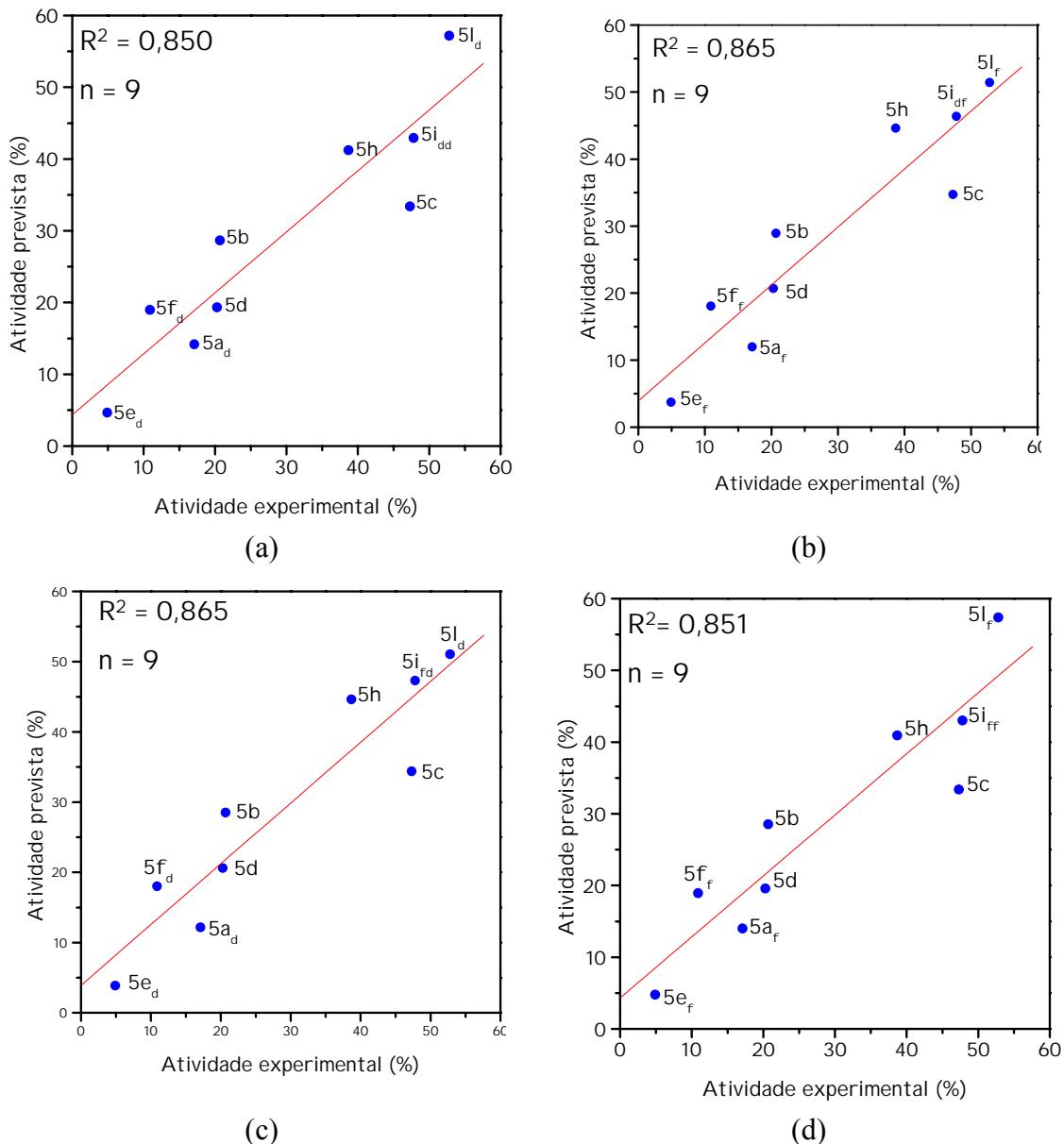




**Figura 4.30:** Mapa dos coeficientes da PLS entre o PE e a atividade antiinflamatória das pirimidinas **5a-q**, com VCDS = VCIS = 0% para as conformações (a) dentro, (b) dentro/fora, (c) fora/dentro e (d) fora.

Nota-se na Figura 4.30 que o mapa do PE é similar para os quatro conjuntos de moléculas. Pode-se, também, observar a identificação de três regiões importantes (uma de máximo e duas de mínimo). A região de máximo está localizada próxima ao substituinte  $R_1$ , e identificada pela variação de cores de vermelha à amarela. Já as duas regiões de mínimo, identificadas pela cor azul, próximas ao substituinte próximo a  $R_3$  e a nuvem- $\pi$  nos carbonos C19 e C20. Percebe-se nesta figura, que a região de máximo próxima a posição  $R_1$ , indica que o aumento do potencial nesta região irá produzir um aumento na atividade biológica do composto. Já a região de mínimo, próxima ao substituinte da posição  $R_3$ , e a nuvem-p em C19 e C20, indica que deverá haver uma diminuição no potencial destas regiões, para assim, haver um aumento na atividade antiinflamatória. Estas variações são alcançadas quando se utilizam substituintes doadores de elétrons na posição  $R_1$ , onde estes podem entrar em ressonância com os demais elétrons do anel benzênico, como é o caso do substituinte  $OCH_3$ , e quando se utilizam substituintes retiradores de elétrons da posição  $R_3$ .

Uma outra etapa desta análise foi verificar a influência dos valores de corte do CAMOL no modelo de regressão. Os gráficos da atividade prevista x atividade experimental utilizando-se um VCDS = 50% e um VCIS = 20% estão apresentados na Figura 4.31.



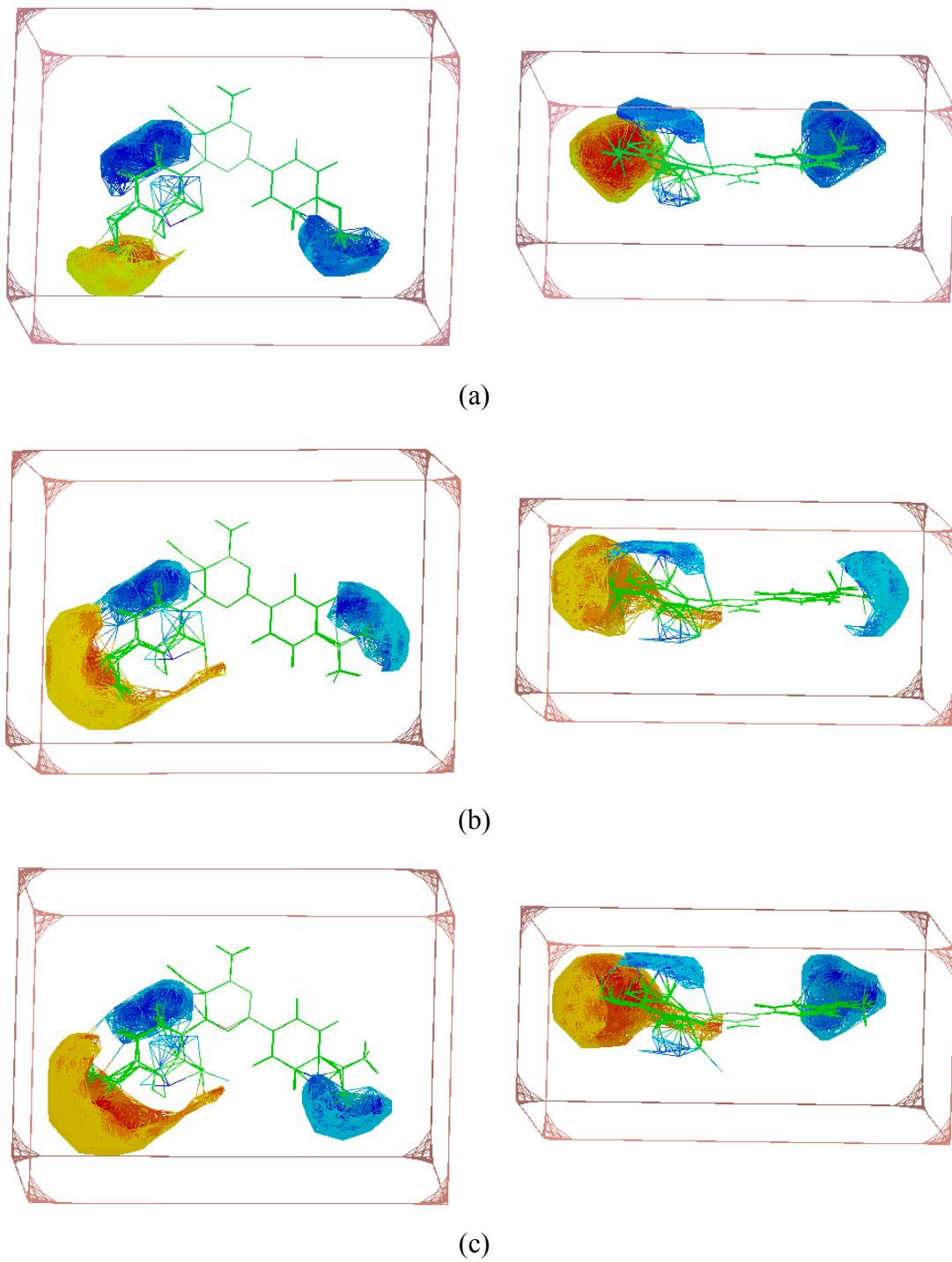
**Figura 4.31:** Gráfico da atividade experimental *versus* a prevista da regressão

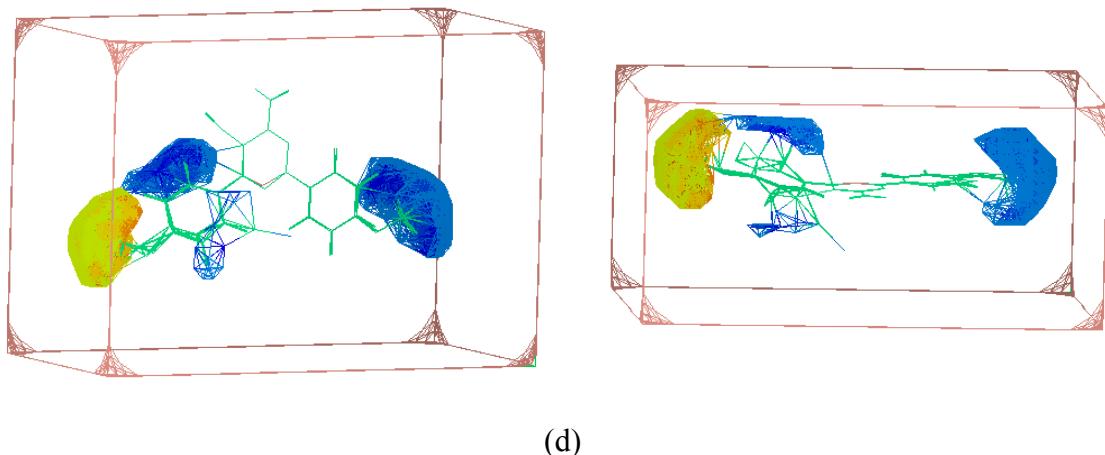
PLS com VCDS = 50% e VCIS = 20% das pirimidinas (a) dentro, (b) fora/dentro, (c) dentro/fora e (d) fora.

Percebe-se na Figura 4.31 que, de modo semelhante ao obtido anteriormente (com VCDS = VCIS = 0%, Figura 4.29), o modelo teórico prevê que os conjuntos contendo as moléculas **5i<sub>df</sub>** e **5i<sub>fd</sub>** fornecem o melhor resultado, com uma variância de aproximadamente 87%, porém os quatros conjuntos de moléculas possuem um valor de  $R^2$  bastante próximos, sendo então, equivalentes entre si. Desta forma, conclui-se que a utilização dos valores de corte do programa CAMOL reduziu o número de variáveis a

serem analisadas e que as diferentes conformações do grupo metoxi nestas moléculas não irão comprometer a qualidade dos modelos obtidos.

Os campos moleculares obtidos para este novo conjunto de variáveis podem ser vistos na Figura 4.32, na qual utilizou-se o mesmo gradiente de cores dos campos moleculares da Figura 4.30.





**Figura 4.32:** Mapa dos coeficientes da PLS entre o PE e a atividade antiinflamatória das pirimidinas **5a-q** do PLS com VCDS = 50% e VCIS = 20% para as conformações (a) dentro, (b) dentro/fora, (c) fora/dentro e (d) fora.

Nota-se na Figura 4.32 que os mapas moleculares são similares tanto para os quatro conjuntos de moléculas quanto para os mapas apresentados na Figura 4.30, com a identificação de três regiões importantes. A região de coeficientes positivos (cor vermelha) está novamente localizada próximo ao substituinte da posição R<sub>1</sub> e a região de mínimo (cor azul) está também localizada próxima ao substituinte da posição R<sub>3</sub> e átomos C19 e C20.

Desta forma, como observado nos gráficos da Figura 4.31, percebe-se que exclusão antecipada de pontos da SPE, através da utilização dos valores de cortes do CAMOL, não modifica, de maneira significativa, o modelo obtido podendo então ser utilizado no desenvolvimento de uma QSAR-3D.

Neste momento, gostaríamos de enfatizar que os nossos resultados estão em concordância com aqueles obtidos com o modelo de regressão linear da atividade antiinflamatória destas mesmas pirimidinas com parâmetros calculados pelo método AM1, por Ramos e colaboradores<sup>34</sup>. No modelo de Ramos e colaboradores, o aumento da atividade se dá com o aumento da carga em C12, a diminuição desta em C13 e a diminuição do momento de dipolo elétrico molecular. Observe que, conforme prevê o nosso modelo, a introdução de um substituinte doador de elétrons em R<sub>1</sub> aumentará a carga em C12 se este puder exercer este efeito doador de elétrons em longo alcance através das ligações-π. Este é o caso do grupo metoxi. Note que na medida que os pares de elétrons do oxigênio entram em ressonância com o anel fenila, por um lado o

potencial próximo aos átomos de hidrogênios do metóxi cresce, enquanto que este tende a diminuir na região próximo aos carbonos C19 e C20 (envolvidos na ligação- $\pi$ ) devido ao aumento da densidade de carga nessa região. Ainda, tanto pelo nosso modelo, quanto pelo de Ramos e colaboradores, a presença de um substituinte metoxi em R<sub>3</sub> indica que uma diminuição da atividade. De fato, uma inspeção na Tabela 1.1 (página 9) mostra que os compostos com grupos retiradores de elétrons em R<sub>1</sub> e doadores de elétrons em R<sub>3</sub> têm as menores atividades. Entretanto, a atividade anti-inflamatória aumenta quando o grupo metoxi está presente em R<sub>1</sub>. Por exemplo as atividades dos compostos **5a**, **5i** e **5l**, valem 17,1%, 47,8% e 52,8%, respectivamente.

Então, a fim de testar esse modelo, esforços foram feitos no sentido de se sintetizar e testar novas moléculas planejadas. Tais moléculas já foram incluídas nas análises se PCA, moléculas **5n**, **5p** e **5q**, (seção 4.3.2), e os resultados obtidos indicaram que a molécula **5p** (com o substituinte NO<sub>2</sub> na posição R<sub>1</sub> e OCH<sub>3</sub> na posição R<sub>3</sub>) deverá ser a molécula de menor atividade antiinflamatória e a molécula com a maior atividade deverá ser a molécula **5n** (com o substituinte NH<sub>2</sub> na posição R<sub>1</sub> e OCH<sub>3</sub> na posição R<sub>3</sub>).

Vale enfatizar que a molécula **5p** já teve sua atividade antiinflamatória determinada, sendo esta igual a 47,42%, indo, portanto, em discordância com o nosso modelo. Na verdade, sua atividade foi menor, apenas, que a atividade da molécula **5l**, proposta por Ramos e colaboradores. Este alto valor na atividade pode, entretanto, ser resultado de uma possível biorredução deste composto para o correspondente amino em R<sub>1</sub>, onde então, passa-se, possivelmente, a ter mistura dos compostos **5p** e **5n**. Entretanto, a molécula **5n** ainda não teve sua atividade determinada para se avaliar a presença deste efeito, assim como para comprovação do modelo. Desta forma, em sentido a avaliar o modelo do ponto de vista de uma estrutura com menor atividade, foi proposta a molécula **5q**, porém esta ainda não foi sintetizada.

# CAPÍTULO 5

## CONCLUSÕES

O Planejamento Fatorial de Dois Níveis da molécula **5h** mostrou que a correlação eletrônica é o efeito mais destacado para o diedro C22C12C2N3, já para o diedro C14C13C4N3, o efeito mais destacado é a inclusão de funções de polarização. Por sua vez, a comparação entre as geometrias moleculares obtidas teoricamente e os dados de cristalografia de raios-X apontaram para resultados próximos dos obtidos com os métodos correlacionados. Assim como se deu para a geometria molecular na molécula **5h**, a partição de carga CHelpG, indicou os métodos correlacionados, em particular, aqueles com funções de polarização, como os mais adequados para descrição da distribuição de carga. Desta forma, decidiu-se utilizar o método BP86/6-31G\*\* para as etapas de cálculo do PE das pirimidinas substituídas **5a-l**.

Observou-se que o efeito da otimização completa da geometria molecular, com o método BP86/6-31G\*\*, não altera o potencial eletrostático, de maneira significativa, quando se utiliza a geometria molecular semi-otimizada. Portanto, o uso de geometrias semi-otimizadas pode vir a ser uma boa aproximação em sistemas envolvendo um grande número de átomos, quando for necessária a utilização um nível de teoria mais sofisticado ou mesmo de efeitos não considerados, como, por exemplo, o efeito solvente.

Tendo em vista que o alinhamento molecular é importante para a comparação dos potenciais moleculares de diferentes moléculas, o emprego de programa MIPSIM revelou que este é bastante sensível a escolha da molécula de referência. Por outro lado, o programa ALINHAMOL mostrou-se como uma opção interessante por fornecer resultados consistentes de alinhamento molecular.

De um ponto de vista quimiométrico, conclui-se, através da Análise de Componentes Principais, que há uma dependência significativa no padrão dos gráficos dos escores da SPE segundo a metodologia de cálculo utilizada, onde o método DFT-BP86/6-31G\*\* mostrou-se mais adequado do que o método AM1, para a descrição da SPE. Por sua vez, os modelos PLS de regressão da atividade antiinflamatória contra o potencial eletrostático, mostraram que as diferentes conformações do grupo metoxi não comprometem a qualidade do modelo obtido. Este fato foi confirmado através dos campos moleculares, pois estes são equivalentes para as quatro conformações do grupo metoxi. É importante notar, também, que devido à alta demanda computacional requerida pela grande quantidade de variáveis, a redução, antecipada, de pontos da SPE com a utilização do VCIS e VCDS (20% e 50%, respectivamente) do programa

CAMOL, mostrou que os modelos obtidos são equivalentes àqueles contendo todas as variáveis.

Finalmente, pode-se, ainda, confirmar as principais conclusões apontadas no trabalho de Ramos e colaboradores, onde modificações estruturais nas posições R<sub>1</sub> e R<sub>3</sub>, em particular, a introdução de grupos doadores e retiradores de elétrons, respectivamente, sugerem uma maior atividade antiinflamatória para esta série de pirimidinas substituídas. Novas moléculas foram sugeridas a partir deste trabalho e esforços estão sendo realizados no intuito de se sintetizar e testar essas novas moléculas.

# CAPÍTULO 6

## PERSPECTIVAS

Neste capítulo, são apresentadas algumas perspectivas de continuação deste trabalho que se abrem a partir dos resultados até aqui obtidos.

### **6.1. Cálculo da SPE das pirimidinas alinhadas com o programa MIPSIM**

Como apresentado nos resultados, fez-se o alinhamento molecular desta série de compostos com o programa MIPSIM, o qual mostrou-se bastante sensível à escolha da molécula de referência (seção 4.6 dos resultados, página 66). Apesar disto, seria interessante o cálculo da SPE, com o mesmo nível de teoria utilizado para as pirimidinas alinhadas com o programa ALINHAMOL, isto é, utilizando-se o método DFT-BP86/6-31G\*\*, para verificar os possíveis modelo de QSAR-3D que poderão vir a ser obtidos e compará-los com os obtidos a partir do alinhamento baseados no programa ALINHAMOL.

### **6.2. Estrutura de hidratação**

Devido à natureza biológica da ação destes compostos ser, em meio condensado, torna-se importante à investigação do efeito do solvente, em particular da água na determinação das propriedades dos compostos. Uma forma de avaliação desse efeito pode se dar através de estudos das chamadas águas de hidratação sobre a molécula. Neste sentido é importante ressaltar que neste trabalho os cálculos foram realizados com a molécula sem interação com nenhum tipo de solvente.

Neste estudo pode-se utilizar o programa AGOA 2.0 também desenvolvido em nosso grupo<sup>66</sup> onde as moléculas do solvente são posicionadas e orientadas uma a uma, baseadas na SPE previamente calculada<sup>67</sup>.

### **6.3. Outras sondas moleculares**

Outra abordagem do ponto de vista de desenvolvimento de uma QSAR-3D pode

ser o a utilização de novas sondas como, por exemplo, a hidrofílica (água) via MC em um conjunto limitado de pontos previamente selecionado ou via AGOA para inclusive reavaliar a própria sonda eletrostática, e gerar o campo molecular do ponto de vista hidrofílico.

#### **6.4. Outras QSAR-3D**

Seria interessante utilizar as aproximações aqui empregadas para estudar outras séries homólogas com atividade biológica bem determinada.

# CAPÍTULO 7

## REFERÊNCIAS

1. Montanari, C.A. e Bolzani, V.S. *Quim. Nova* **2001**, 24 (1), 105.
2. Hammett, L. P. *Physical Organic Chemistry*, **1940**, 1st Ed.; Mc Graw Hill; USA.
3. Hansch, C. *Acc. Chem. Res.* **1969**, 2, 232.
4. Pulay, P. *J. Mol. Struct.* **1995**, 347, 293.
5. Perun, T. J.; Prost, C. L.; *Computer-Aided Drug Design: Methods and Applications*; Dekker; USA, 1989.
6. Martin, Y.C. *Persp. Drug Discov. Design* **1998**, 12-14, 3.
7. Cruciani, G., Clementi, S.; Pastor, M. *Persp. Drug Discov. Design* **1998**, 12-14, 71.
8. Dunn III, W. J.; Hopfinger, A. J. *Persp. Drug Discov. Design* **1998**, 12-14, 167.
9. Beck, B., Clark, T. *Persp. Drug Discov. Design* **1998**, 9-11, 131.
10. Andreoni, W. *Persp. Drug Discov. Design* **1998**, 9-11, 161.
11. Carbó-Dórca, R.; Amat, L.; Besalú, E.; Robert. D. *J. Mol. Struct. -THEOCHEM* **2000**, 504, 181..
12. Silverman, R. B.; *The Organic Chemistry of Drug Design and Drug Action*; Academic Press Inc.; USA, 1992..
13. Korolkovas, A.; Buckhalter, J.H., *Química Farmacêutica*, **1982**, Ed. Guanabara Dois: Rio de Janeiro
14. Barreiro, E.J. e Carlos Alberto Manssour Fraga, *Química Medicinal, As Bases Moleculares da Ação dos Fármacos*, **2001**, Editora ArtMed.
15. Dunn III, W.J. *Chemom. Intell. Lab. Syst.* **1989**, 6, 181.
16. Kubinyi, H. *Drug Disc. Today* **1997**, 2, 457.
17. Goodford, P. J. *J. Med. Chem.* **1985**, 28, 849.
18. Cramer III, R. D.; Patterson, D. E.; Bunce, J. *J. Am. Chem. Soc.* **1988**, 110, 5959.
19. Goodford P.A. *J. Chemom.* **1996**, 10, 107.
20. Cruciani, G.; Watson, K. *J. Med. Chem.* **1994**, 37, 2589.
21. Hoskuldsson, A. *J. Chemom.* **1988**, 2, 211.
22. Karelson, M.; Lobanov, V.S.; Katritzky, A.R. *Chem. Reviews* **1996**, 96, 1027.
23. ALINHAMOL, Marcelo Zaldini Hernandes, Depto. Ciências Farmacêuticas, zaldini@ufpe.br, 2004.

24. CAMOL, Marcelo Zaldini Hernandes, Depto. Ciências Farmacêuticas, zaldini@ufpe.br, 2003.
25. Carvalho, M. A.; da Silva, J. B. P.; Hernandes, M. Z. *Int J. Quantum Chem.* **2005**, 101 (*In Press*).
26. Bruns, R. E.; Haiduke, R. L. A.; do Amaral, A. T *J. Braz. Chem. Soc.* **2002**, 13 (6), 800.
27. Mulliken, R.S. *J. Chem. Phys.* **1955**, 23, 1833.
28. Breneman, C.M.; Wiberg, K.B. *J. Comp. Chem.* **1990**, 11, 361
29. Falcão, E. P. S., Tese de Doutorado, Pós-Graduação em Ciências Biológicas, Universidade Federal de Pernambuco, **2002**.
30. Anderson, G. W.; Halverstadt, I. F.; Miller, W. H.; Rohling, R. O. *J. Am. Chem. Soc* **1945**, 67, 2197.
31. Hepworth, W.; Thompson, T. W. *Chem. Abstracts* **1968**, 69, 7724.
32. Benneche, T.; Strande, P.; Oftebro, R.; Undheim, K. *Eur. J. Med. Chem.* **1993**, 28, 463.
33. Sanfilippo, P.J.; Urbanski, M.; Williams, L.; Press, J. B.; Katz, L. B.; Shriver, D. A.; Fernandez, J. A.; Shatynski, D.; Offord, S. J. *Eur. J. Med. Chem.*, **1992**, 27, 655.
34. Ramos, M. N.; Falcão, E. P. S; daSilva, J. B. P. Comunicação pessoal.
35. Roothaan, C. C. *J. Rev. Mod. Phys.* **1951**, 23, 69.
36. Hohenberg, P.; Kohn, W. *Phys Review* **1964**, 136, B864.
37. Kohn, W.; Sham, L. J. *Phys Review* **1965**, 140, A1133.
38. Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, 46, 618.
39. Levine, I.; *Quantum Chemistry*, **2000**, 5<sup>a</sup> ed., Prentice Hall. Upper Saddle River.
40. Cramer, Christopher J.; *Essentials of Computational Chemistry*, 2002, Wiley, New York.
41. Foresman, J. B.; *Exploring Chemistry With Electronic Structure Methods*, **1993** Gaussian Inc., Pittsburgh.
42. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E. ; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery Jr., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S; Millam, J. M; Daniels, A. D.; Kudin, K. N.; Strain, M. C; Farkas, O; Tomasi, J; Barone, V; Cossi, M; Cammi, R.; Mennucci, B.;

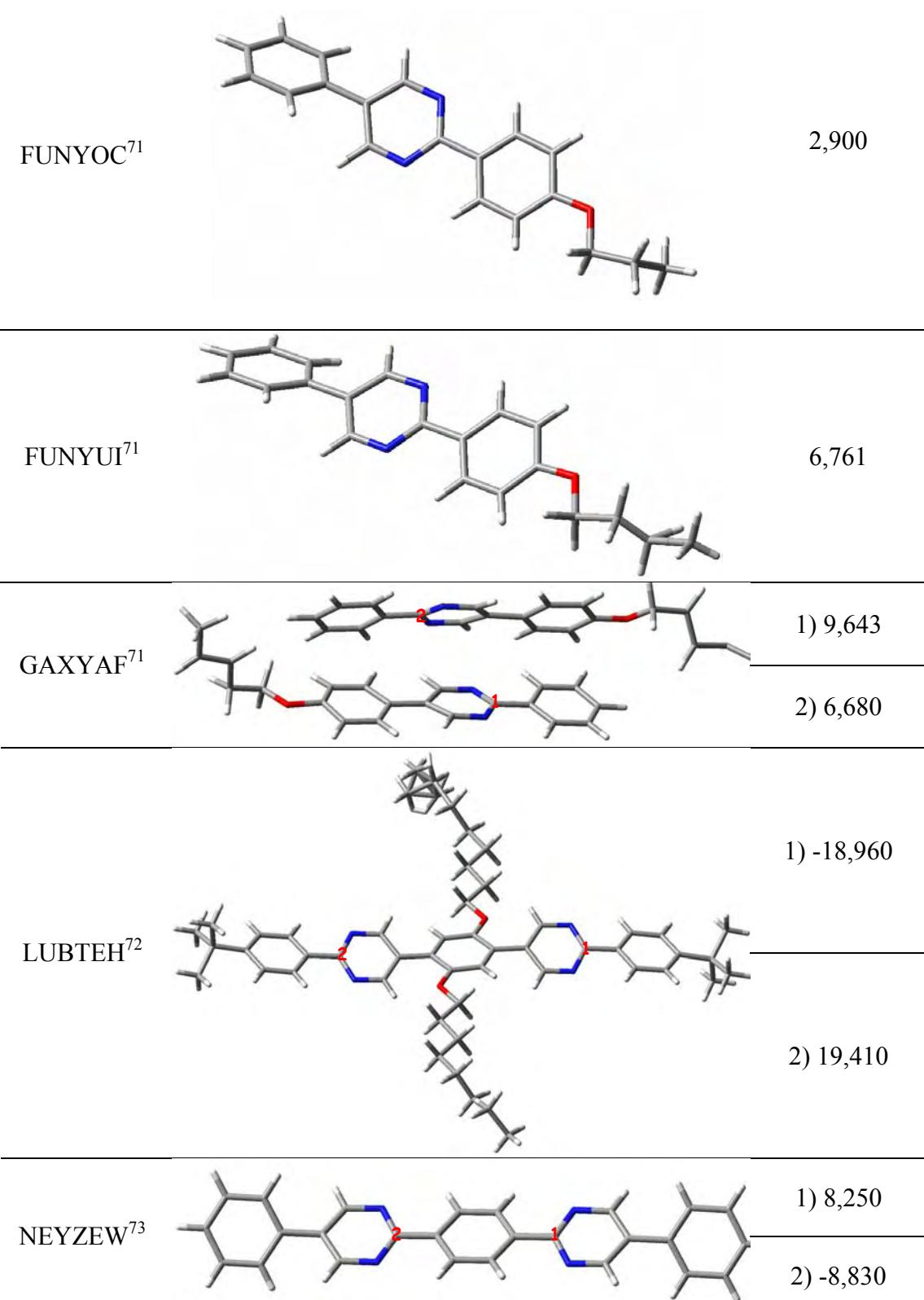
- Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A.; GAUSSIAN 98, Rev.A.11.2, Gaussian, Inc., Pittsburgh PA, 1998)
43. Szabo, A.; Ostlund, N. S.; *Modern Quantum Chemistry*, **1996**, Dover Publications. New York.
44. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.*, **1985**, 107, 3902.
45. de Azevedo, A. L. M. S.; Neto, B. B.; Scarminio, I.; de Oliveira, A. E.; Bruns, R. E. *J. Comp. Chem*, **1996**, 17, 167.
46. Bruns, R. E.; Guadagnini, P. H.; Scarminio, I.; Neto, B. B. *J. Mol. Struct. - Theochem*, **1996**, 394, 197.
47. da Silva, J. B. P.; Ramos, M. N.; Bruns, R. E. *Spectrochim. Acta A*, **1997**, 53, 733.
48. da Silva, J. B. P.; Neto, B. B.; Ramos, M. N.; Bruns, R. E. *Spectrochim. Acta A*, **1997**, 53, 1563.
49. Neto, B.B.; Scarmínio, I.S.; Bruns, R. E. em “Como Fazer Experimentos”, **2001**, Unicamp, Campinas.
50. Kachigan, S.K., "Statistical Analysis – An Interdisciplinary Introduction to Univariate & Multivariate Methods", **1996**, Radius Press.
51. Mardia, K. V.; Kent, J. T.; Bibby, J. M.; *Multivariate Analysis*, Academic **1979**, Chapter 8, Press: London.
52. Suto, E.; Ferreira, M. M. C.; Bruns, R. E. *J. Comp. Chem.* **1991**, 12, 885.
53. Breneman, C. M.; Wiberg, K. B. *J. Comp. Chem.* **1990**, 11, 361.
54. Cosentino, U., Moro, G., Bonalumi, D., Bonati, L., Lasagni, M., Todesquini, R., Pitea, D. *Chemom. Intell. Lab. Syst* **2000**, 52, 183.
55. de Cáceres, M.; Villà, J.; Lozano, J. J.; Sanz, F. *Bioinformatics*, **2000**, 16, 568.
56. Chirlian, L. E.; Franch, M. M. *J. Comp. Chem.* **1987** 8, 894.

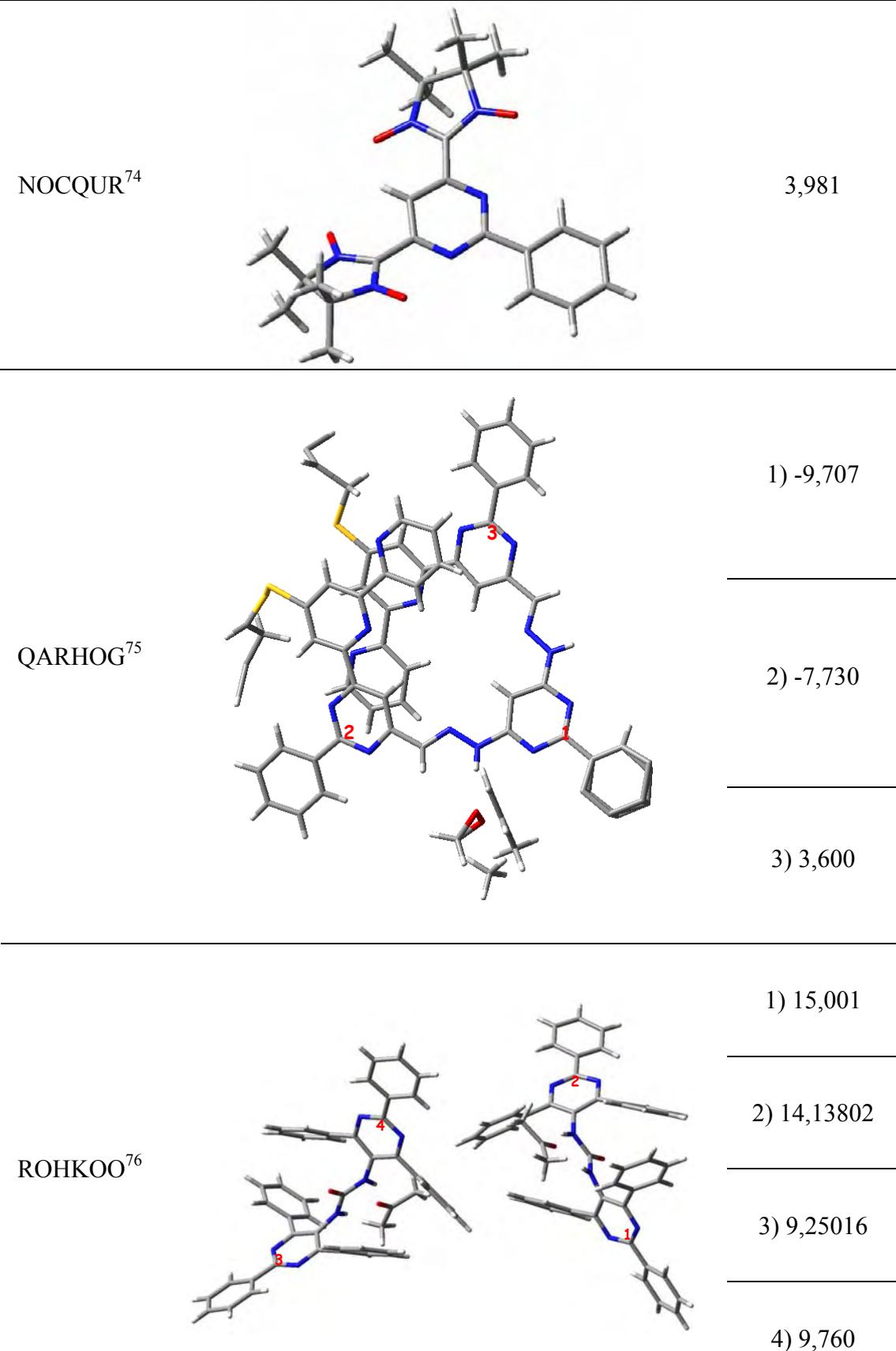
57. Hadjipavlou-Litina, D; Geronikaki, A; Mgonzo, R.; Doytchinova, I.; Drug Development Research, **1999**, 48, 53.
58. Kuroda, T.; Suzuki, F.; Ohmori, K.; Hosoe H. *J. Med. Chem.* **1992**, 35, 1130.
59. Vinegar R, Truax Jf, Selph Jl Quantitative Studies Of Pathway To Acute Carrageenan Inflammation Federation Proceedings 35 (13): 2447-2456 1976
60. Allen, F.H., Acta Crystallogr. B, **2002**, 58, 380.
61. Bruno, I.J.; Cole, J.C.; Edgington, P.R.; Kessler, M.; Macrae, C.F.; McCabe, P.; Pearson, J.; Taylor, R., Acta Crystallogr. B, **2002**, 58, 389
62. Allen, F.H.; Motherwell, W.D.S., Acta Crystallogr. B, **2002**, 58, 407.
63. Yonemoto, K.; Shibuya, I.; Honda, K. *Bull.Chem.Soc.Jpn.* **1989**, 62, 1086.
64. Buemi G, Zuccarello F, Raudino A *J. Mol. Struct. – THEOCHEM* **1988**, 41, 379
65. Marilene P. Bastos- Nneviva, Química Nova, 1984, Julho, 141.
66. AGOA v2.0, Marcelo Zaldini Hernandes, Depto. de Ciências Farmacêuticas, UFPE, [zaldini@ufpe.br](mailto:zaldini@ufpe.br), 2004]
67. Hernandes, M. Z.; da Silva, J.B. P.; Longo, R. L. *J. Braz. Chem. Soc.* **2002**, 13, 36.
68. Rybalova, T.V.; Sedova, V.F.; Bagryanskaya, I.Yu.; Gatilov, Yu.V. Shkurko, O.P. *Chem.Hetero.Compnd* 2001, 229.
69. Rapthel, I.; Hartung, H.; Richter, R.; Jaskolski, M. *J.Prakt.Chem.-Chem.-Zeitung* **1983**, 325, 489.
70. Tarzia, G.; Panzone, G.; Zerilli, L.; Lanfranchi, M.; Pelizzi, G. *Gazz.Chim.Ital.* **1984**, 114, 431.
71. Winter, G.; Hartung, H.; Jaskolski, M. *Mol.Cryst.Liq.Cryst.* **1987**, 149, 17
72. Wong, Ken-Tsung; Hung, Tsung Shi; Lin, Yuting; Wu, Chung-Chih; Lee, Gene-Hsiang; Peng, Shie-Ming; Chou, Chung Hsien; O.Su, Yuhlong *Organic Letters* **2002**, 4, 513.
73. Gompper, R.; Mair, H.-J., Polborn, K. *Synthesis* **1997**, 696.
74. Oshio, H.; Umeno, M.; Fukushi, T.; Ito T.; *Chem.Lett.* **1997**, 1065
75. Gardinier, K.M.; Khoury, R.G.; Lehn J.-M.; *Chem.-Eur.J.* **2000**, 6, 4124.
76. Bolte, M. *Acta Crystallogr.,Sect.C:Cryst.Struct.Commun.* **1997**, 53, 9700015.
77. Korbonits, D.; Simon, K.; Kolonits, P. *Chem.Ber.* **1991**, 124, 111.
78. Mandal, P.; Paul, S.; Stam, S.H.; Schenk, H. *Mol.Cryst.Liq.Cryst.*

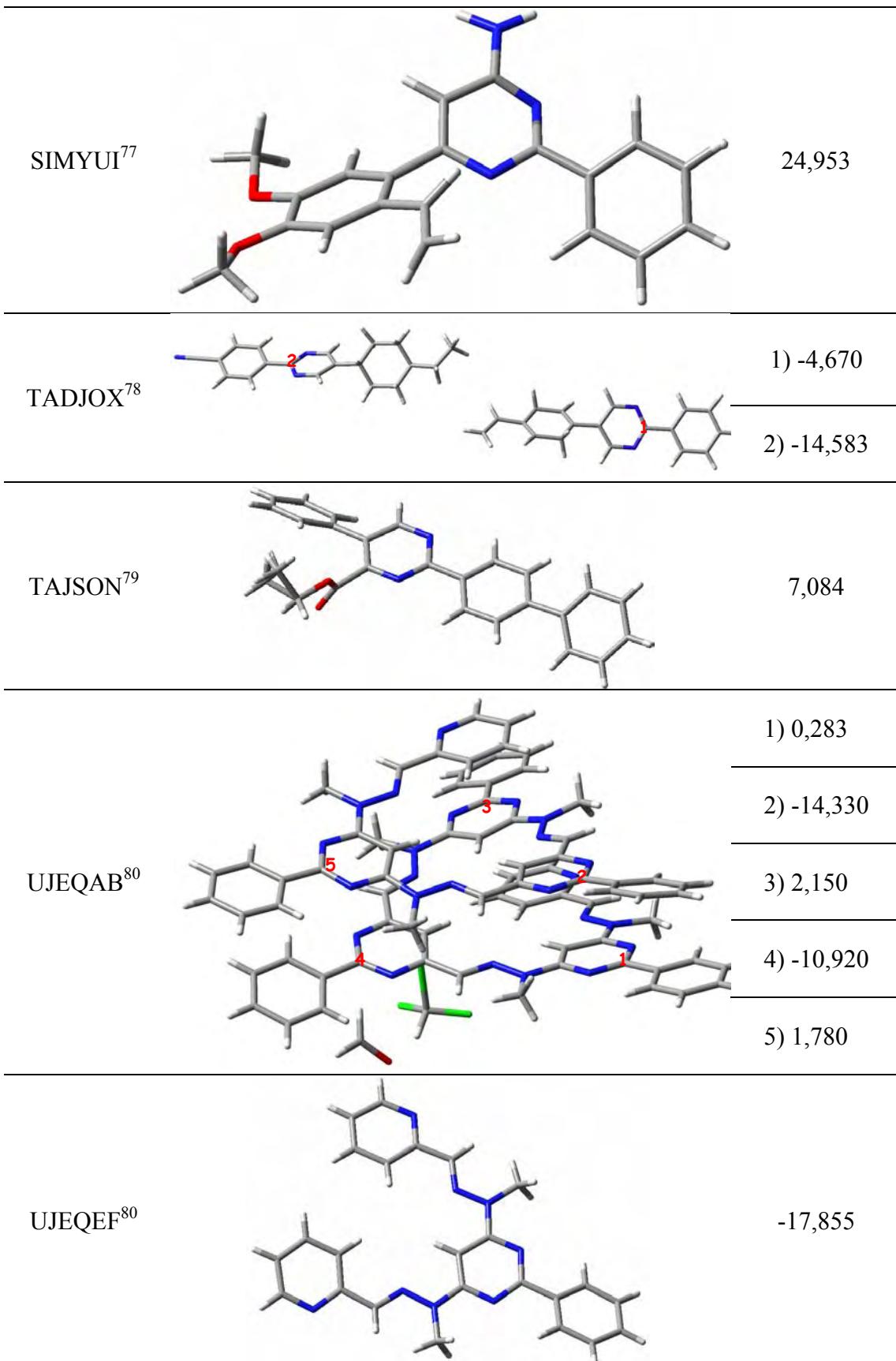
- 1990,'180B',369.**
79. Huschka, M.; Gompper, R.; Polborn, K.; *Private Communication*, **2003**.
80. Schmitt, J.-L.; Stadler, A.-M.; Kyritsakas, N.; Lehn, J.-M. *Helv.Chim.Acta* **2003**,*86*,1598.
81. Mandal, P.; Majumdar, B.; Paul, S.; Schenk, H.; Goubitz, K. *Mol.Cryst.Liq.Cryst.* **1989**,*168*,135.
82. Hernandez, R.P.; Rodriguez, J.D.; Trimino, M.I.G.; de Armas, H.N.; Toscano, R.A. *Acta Crystallogr.,Sect.C:Cryst.Struct.Commun.* **1995**,*51*,1392.
83. Meier, C.; Gompper, R.; Polborn, K. *Private Communication* **2002**.

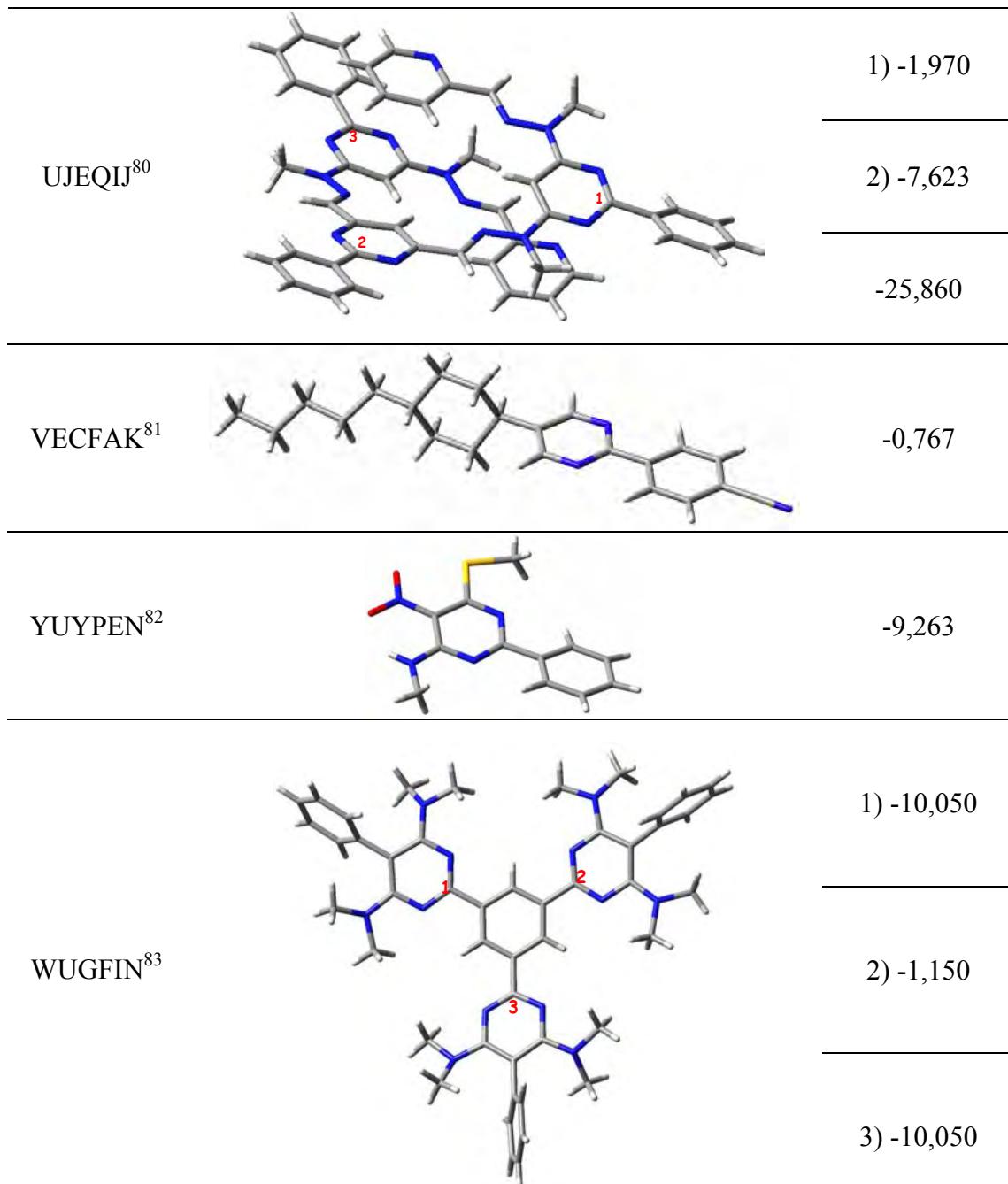
## **ANEXO A**

<b>Arquivos</b>	<b>Estrutura de raios-x</b>	<b>C14C13C6N3</b>
ACAYOS <sup>68</sup>		-6,988
ACAYUY <sup>68</sup>		-0,839
CEDWOX <sup>69</sup>		-1,990
		5,931
DANNEL <sup>70</sup>		-4,772

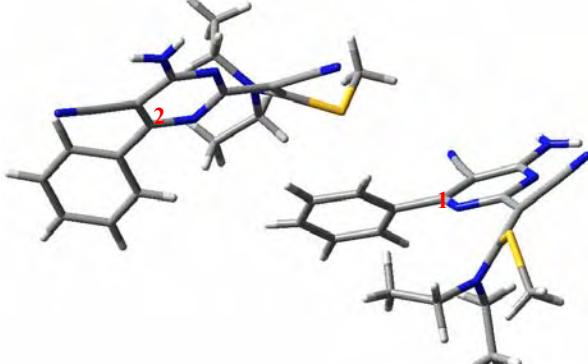






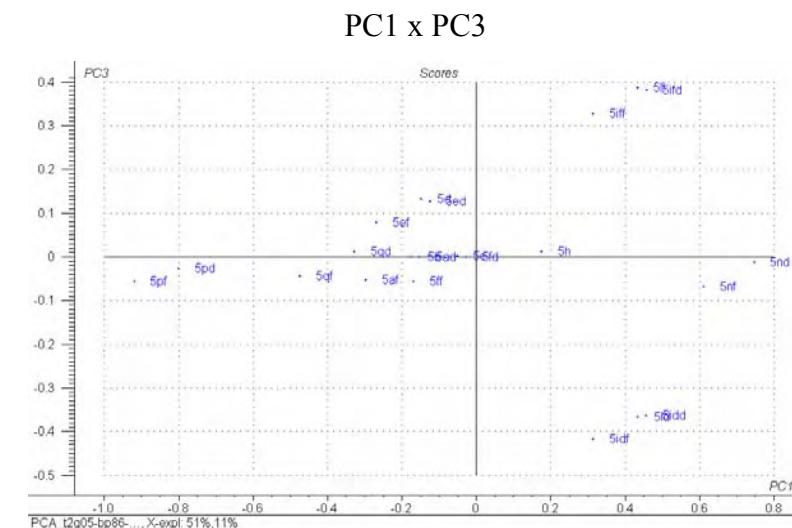
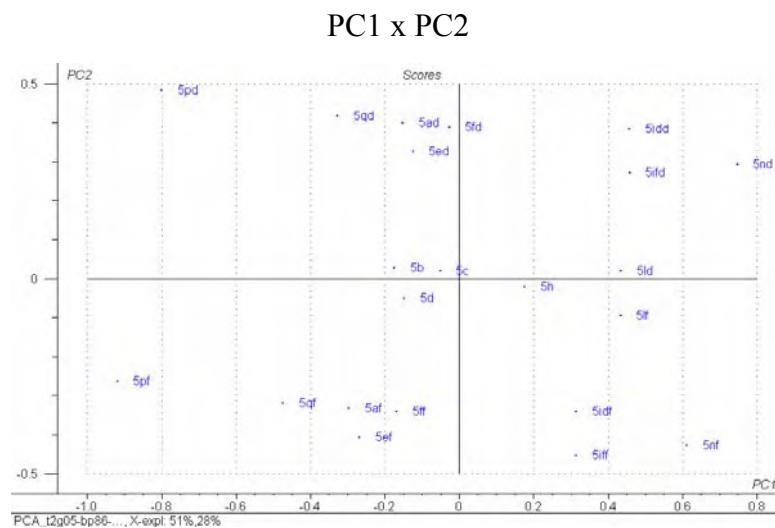


**Figura A1:** Estruturas de raios-X, CSD, com o diedro C14C13C4N3.

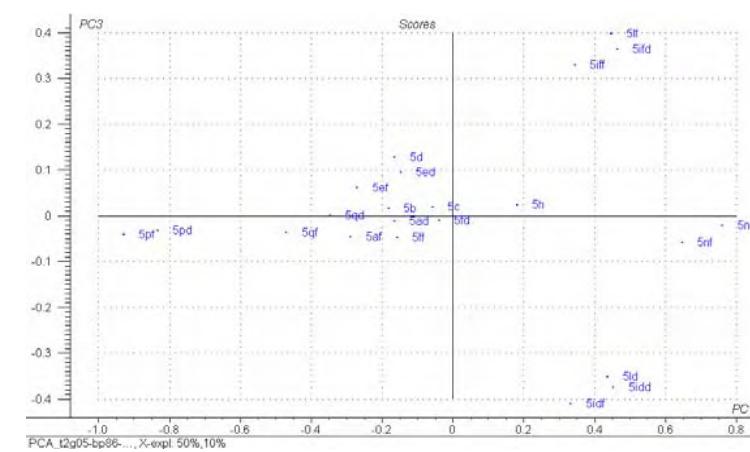
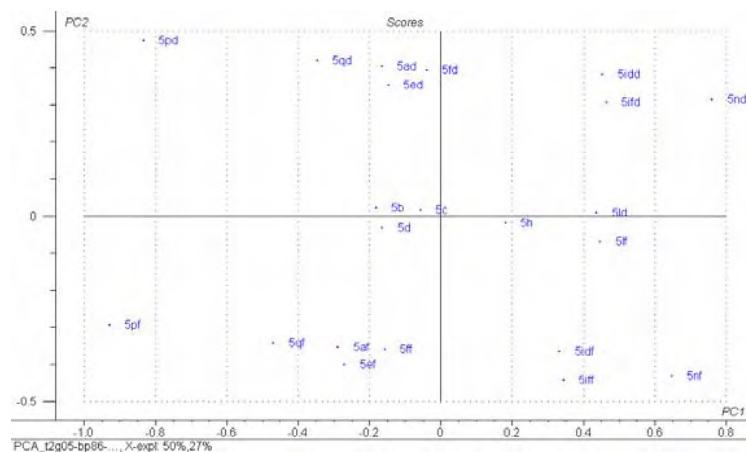
Arquivos	Estrutura de raios-x	C23C12C2N3
KERDIU <sup>63</sup>		1) 45,826 2) 30,450

**Figura A2:** Estrutura de raios-X, CSD, com o diedro C23C12C2N3.

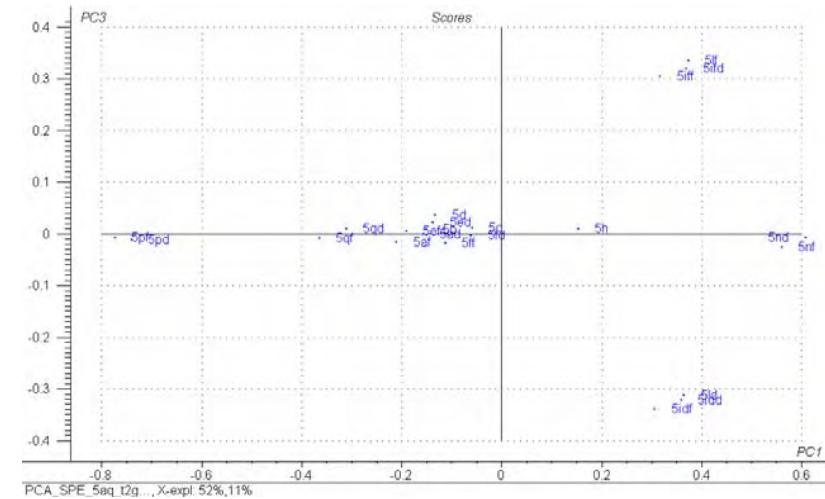
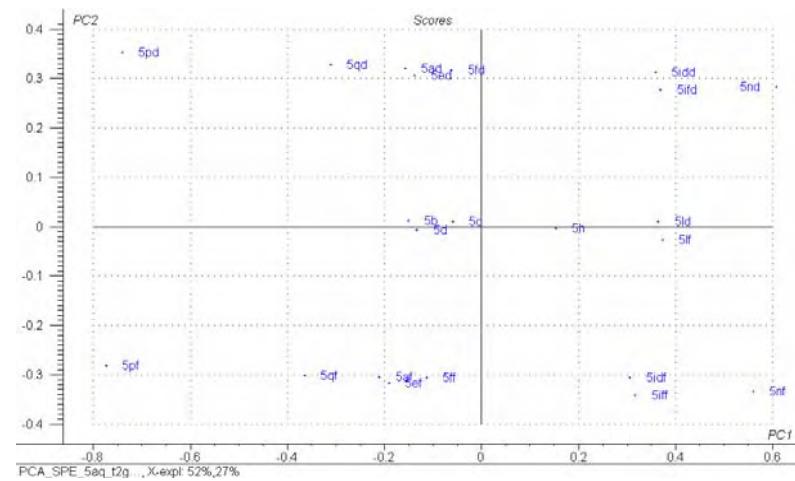
## **ANEXO B**



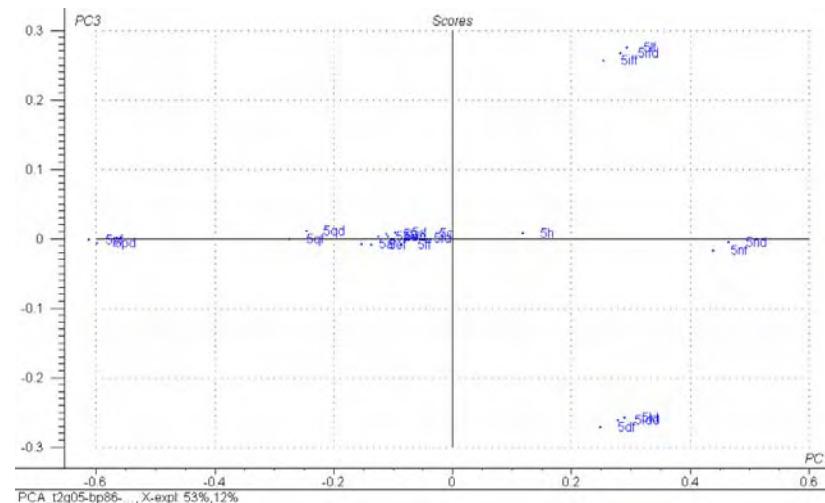
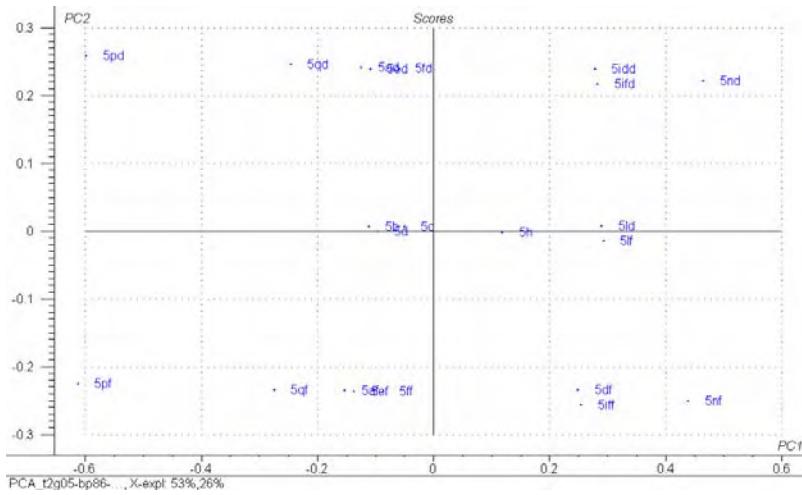
VCIS = 0%



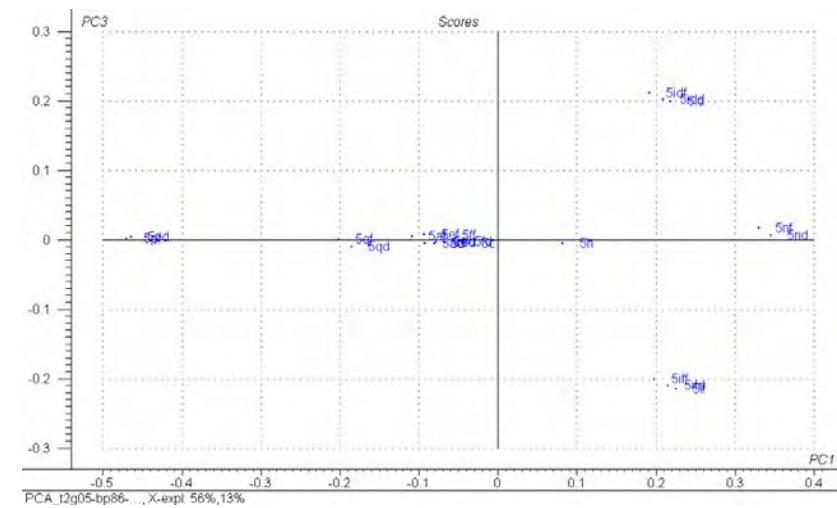
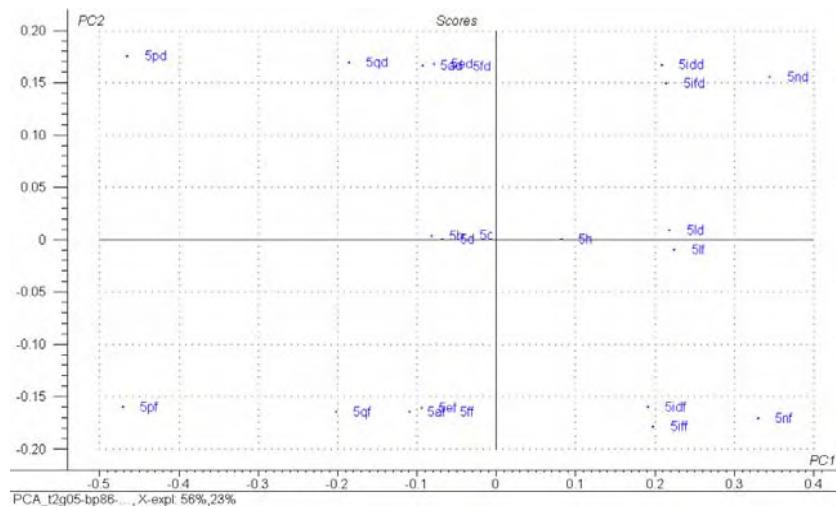
VCIS = 10%



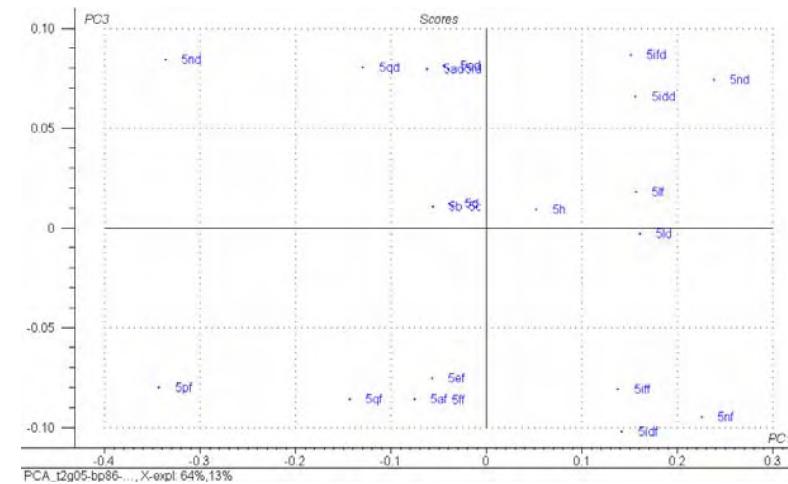
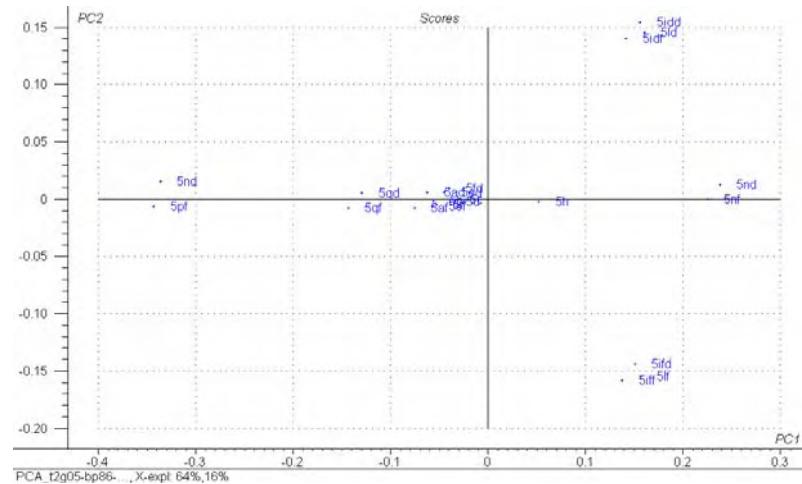
VCIS = 20%



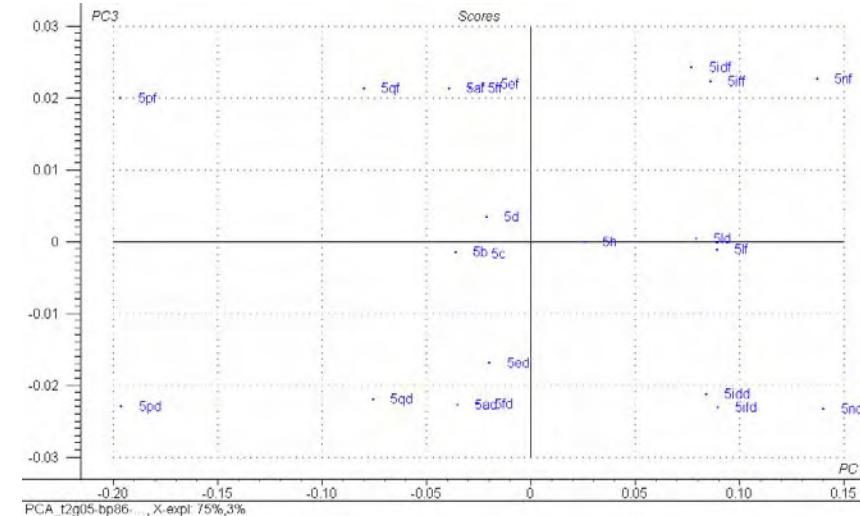
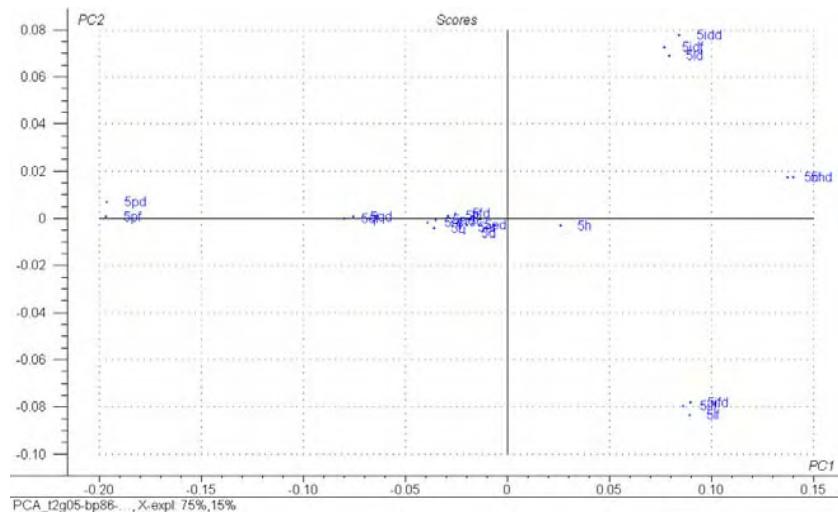
VCIS = 30%



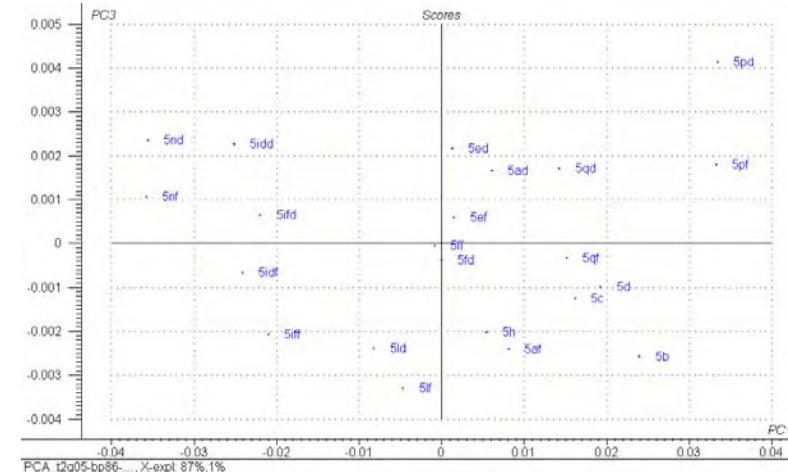
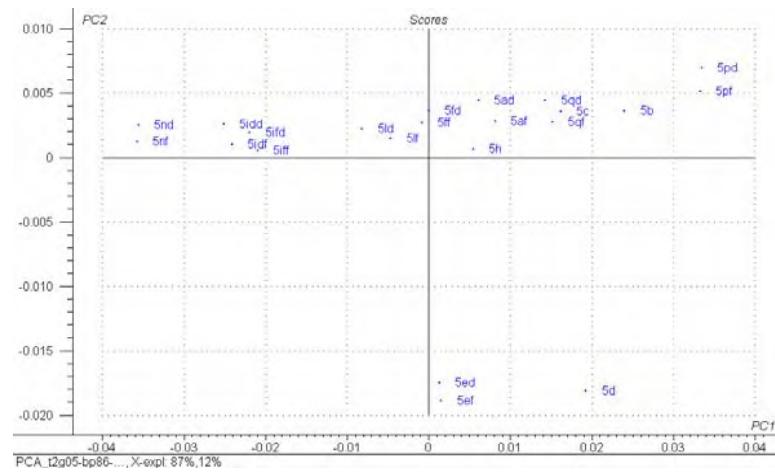
VCIS = 40%



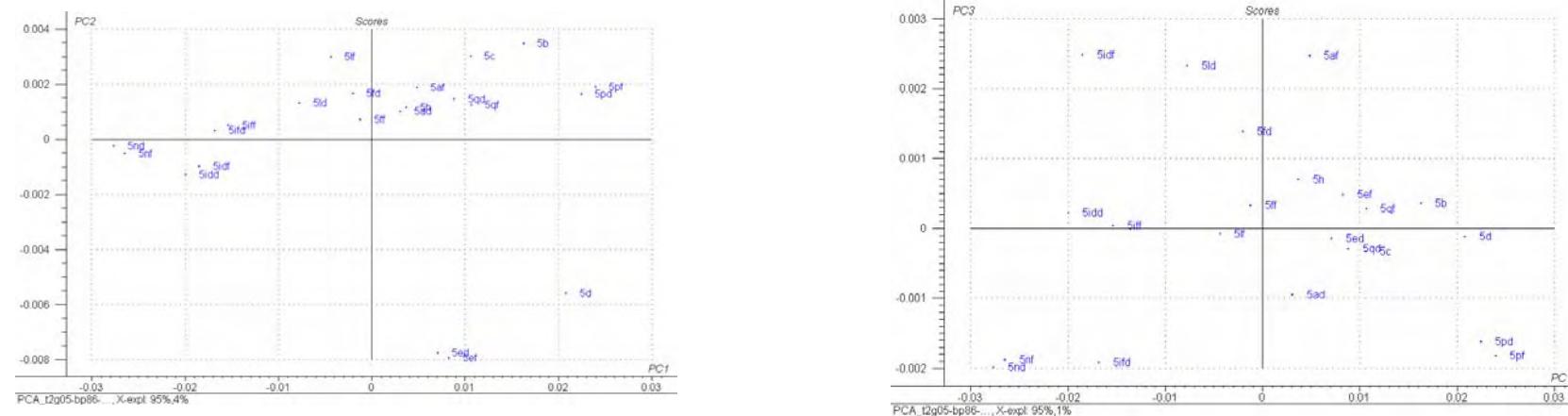
VCIS = 50%



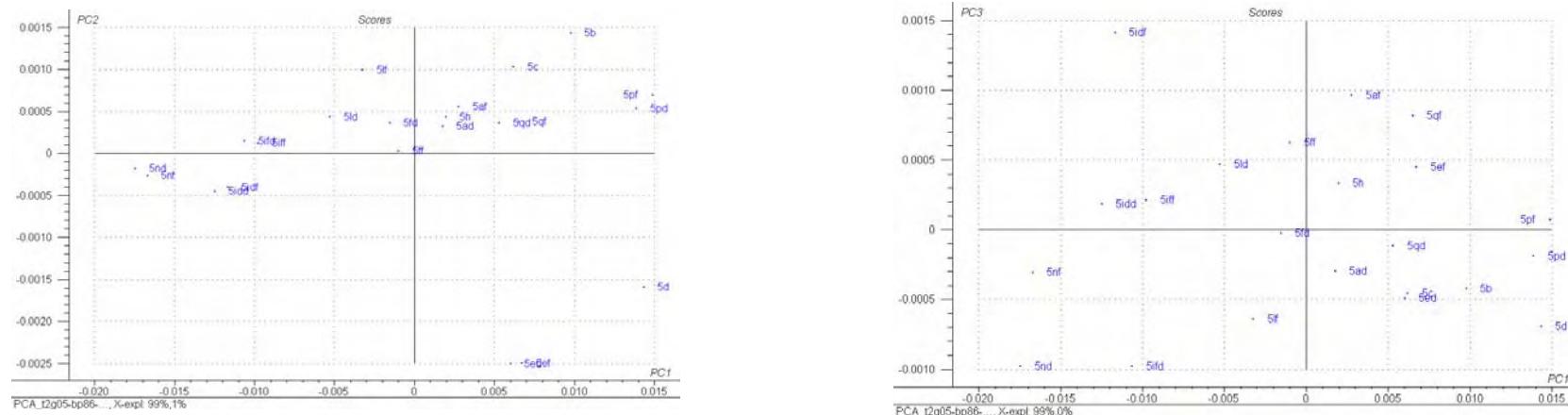
VCIS = 60%



VCIS = 70%

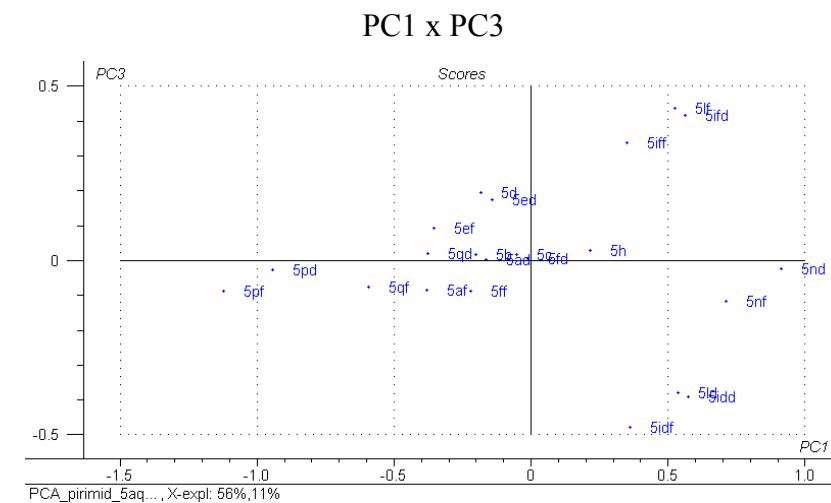
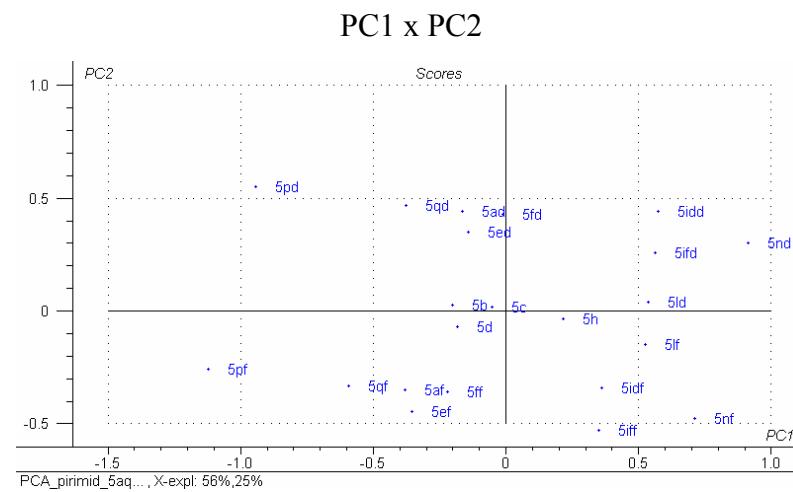


VCIS = 80%

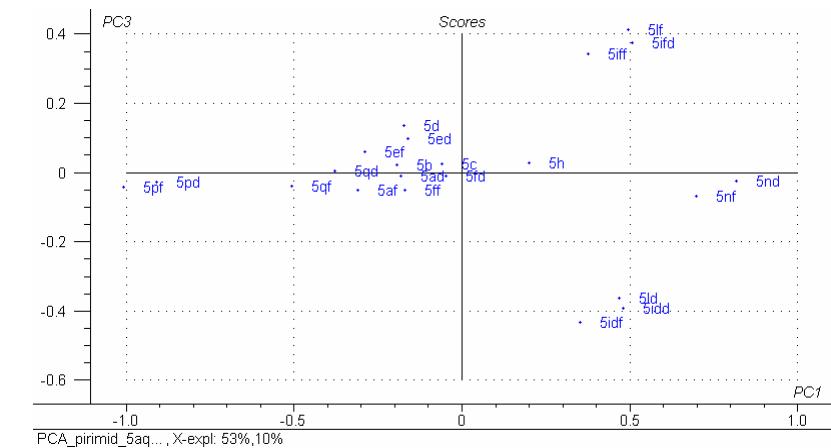
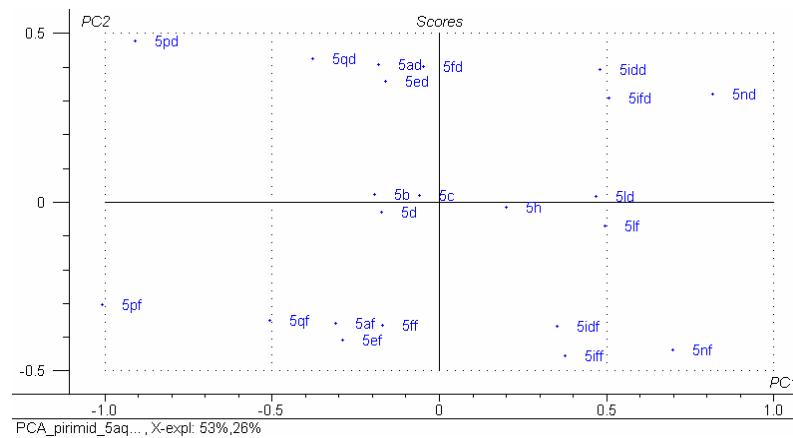


VCIS = 90%

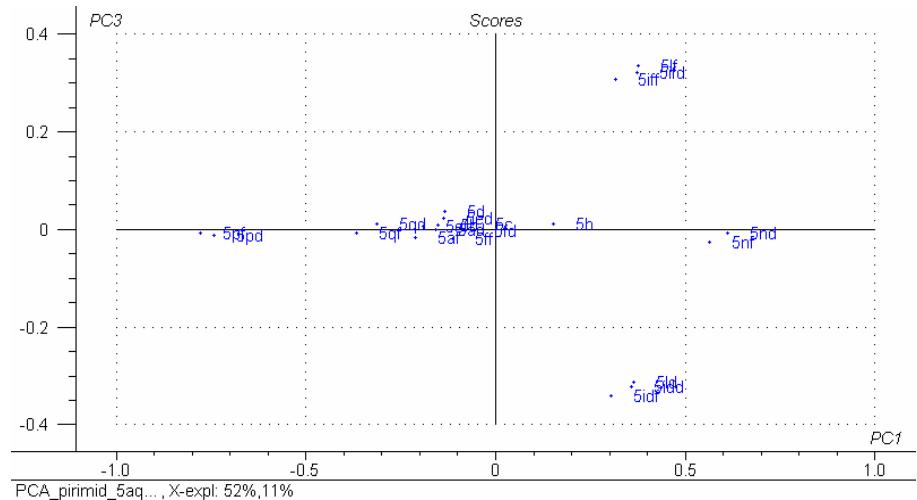
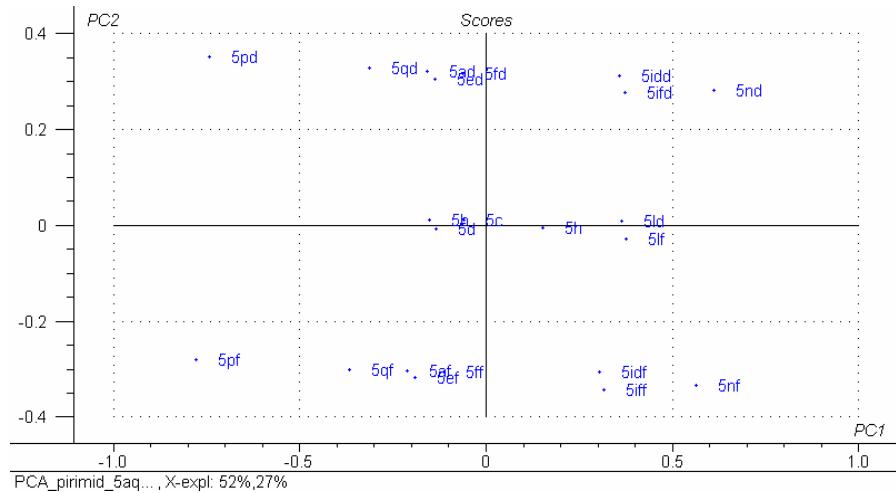
**Figura B1:** Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas 5a-q totalmente otimizadas com VCDS = 100% e  $0\% < \text{VCIS} < 90\%$  da caixa 5.



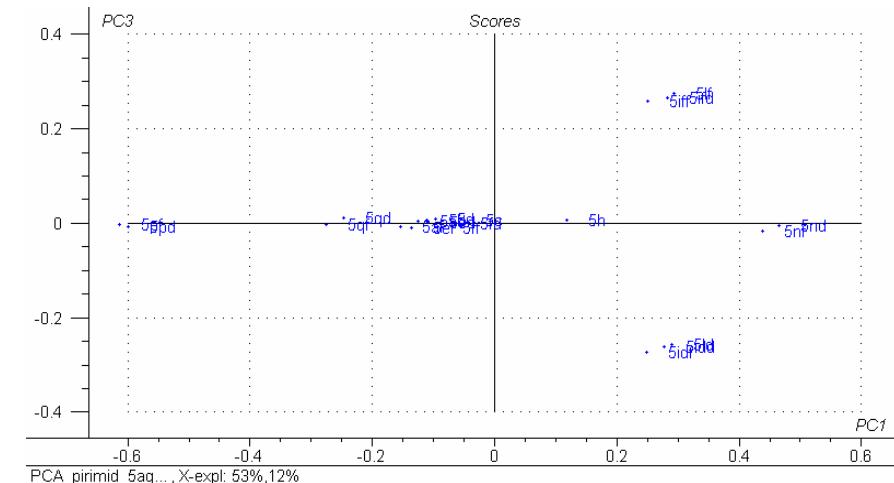
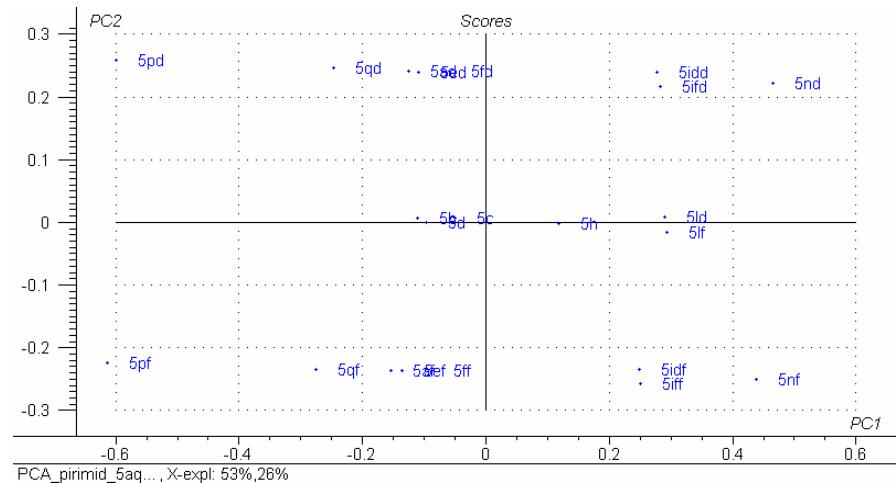
VCIS = 0%



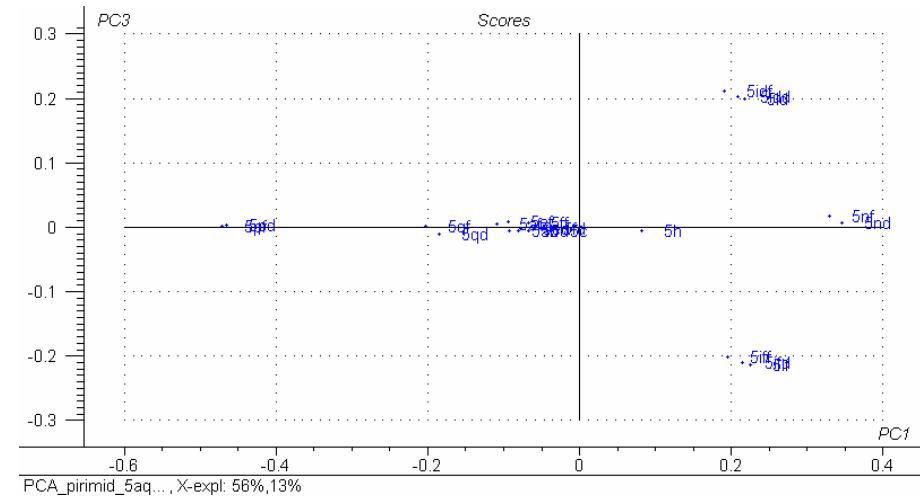
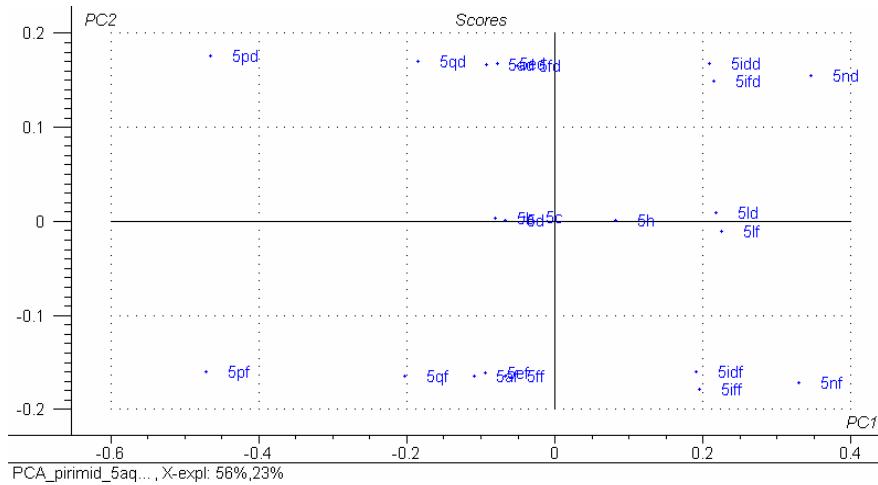
VCIS = 10%



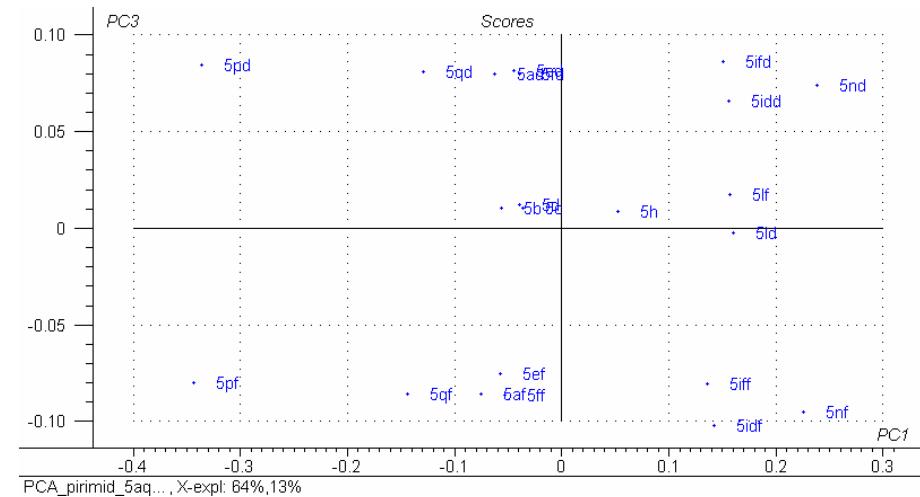
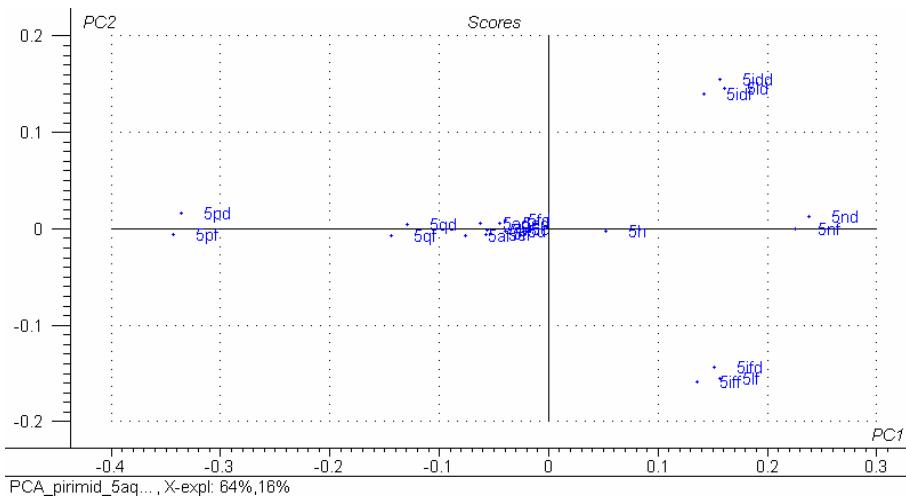
VCIS = 20%



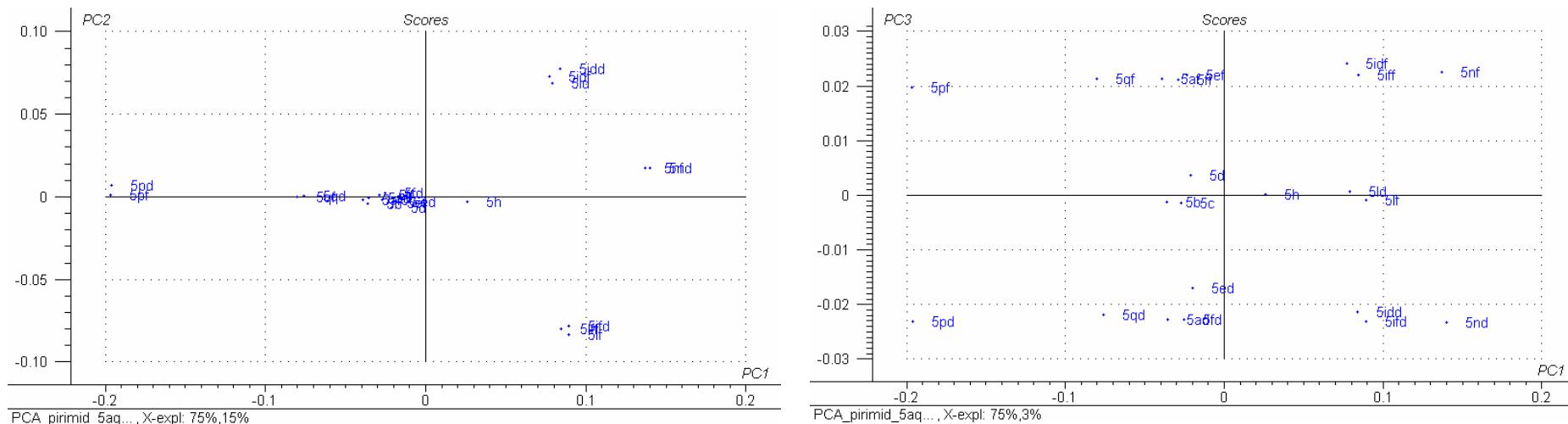
VCIS = 30%



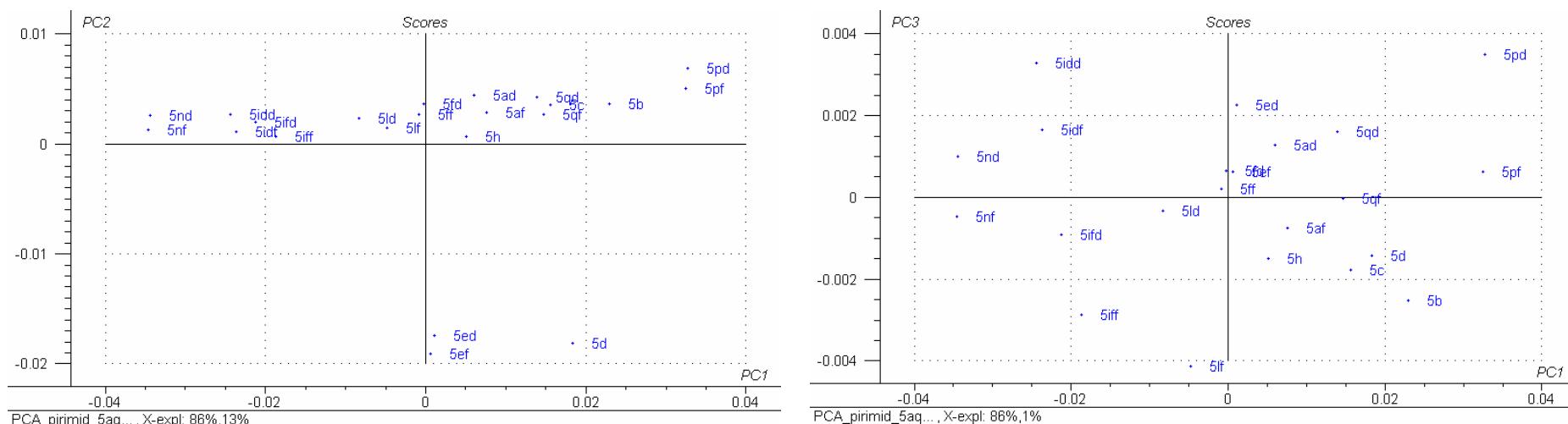
VCIS = 40%



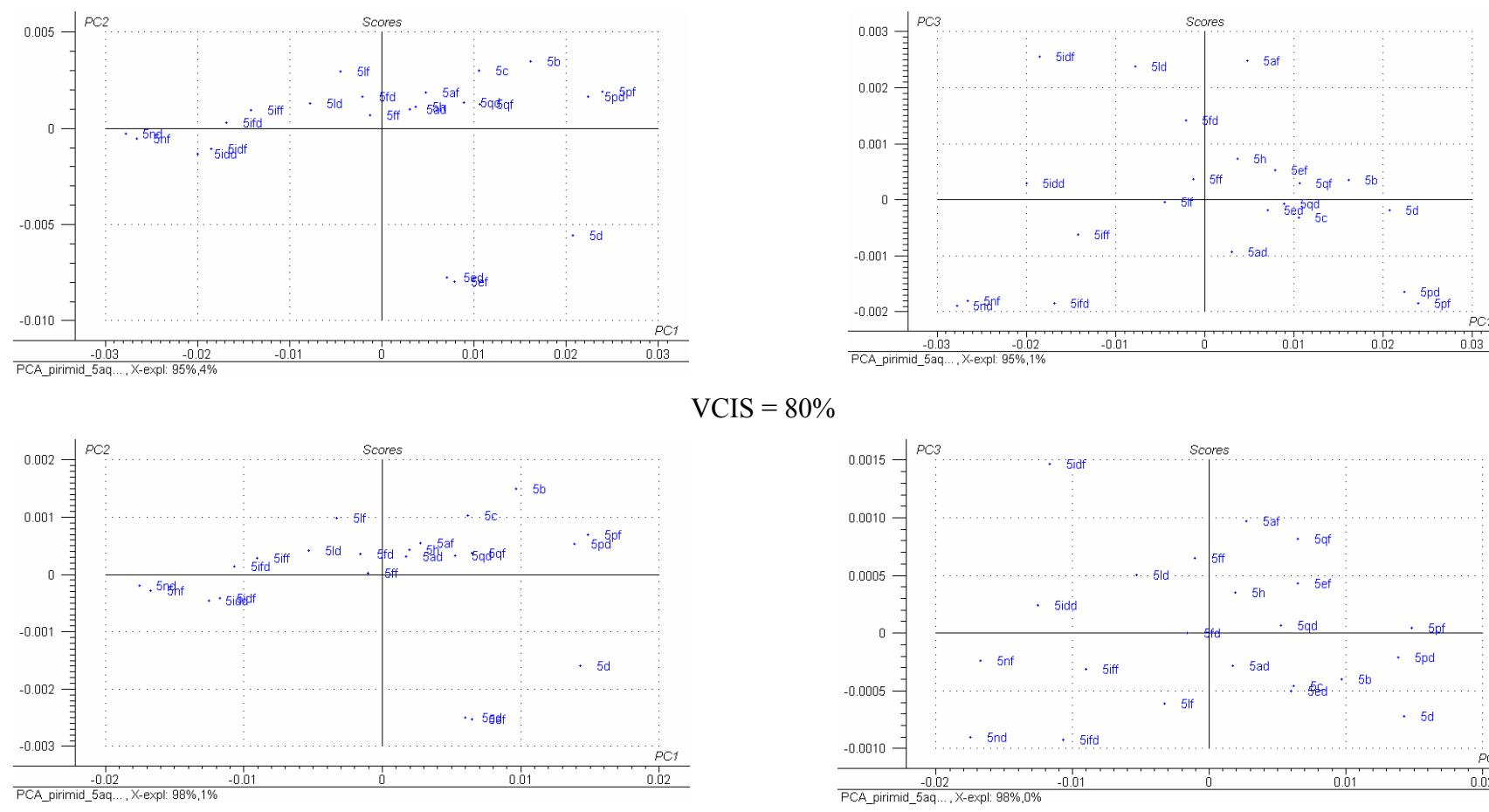
VCIS = 50%



VCIS = 60%



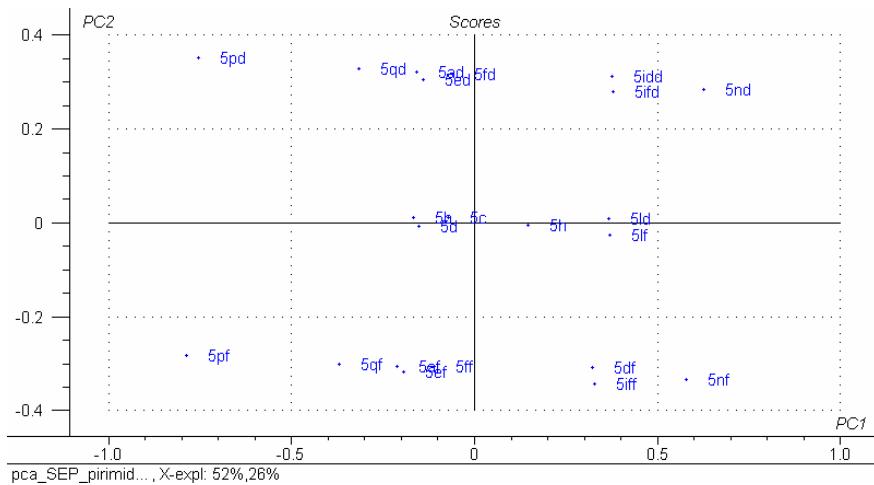
VCIS = 70%



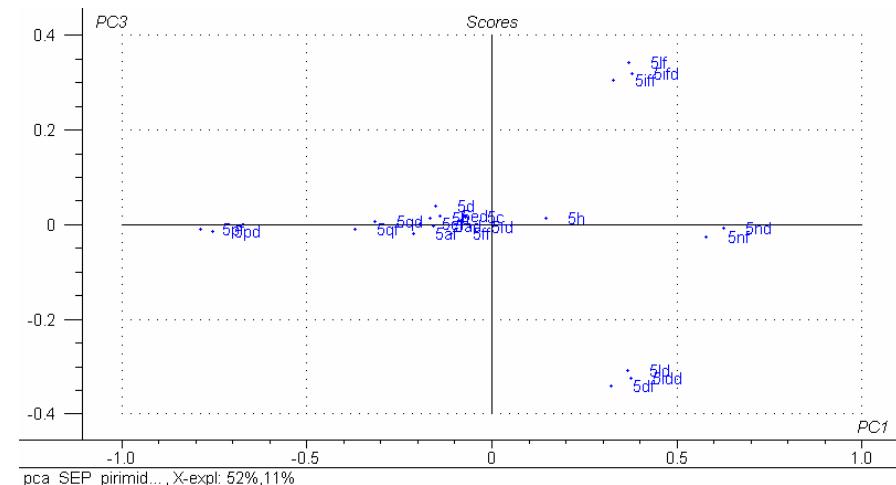
**Figura B2:** Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas 5a-q totalmente otimizadas com VCDS = 100% e  $0\% \leq \text{VCJS} \leq 90\%$  da caixa 8

## **ANEXO C**

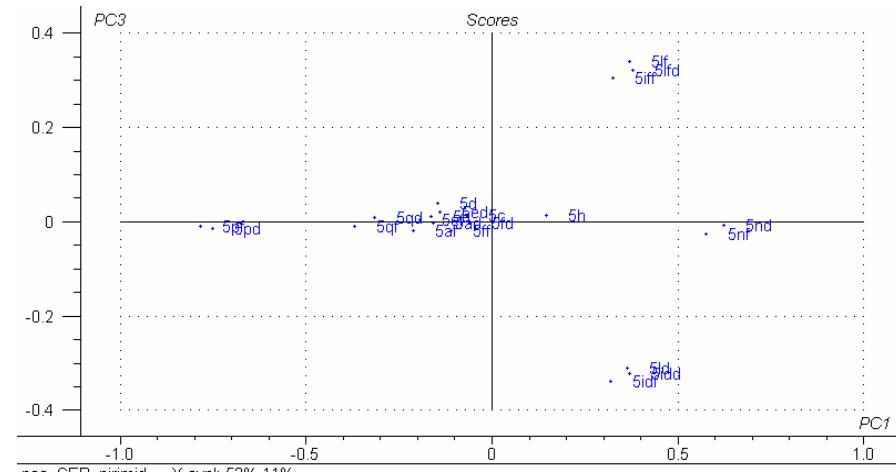
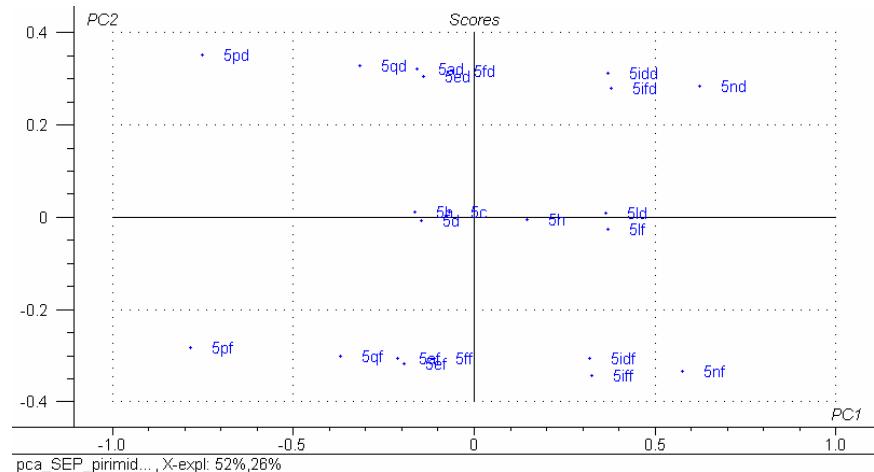
PC1 x PC2



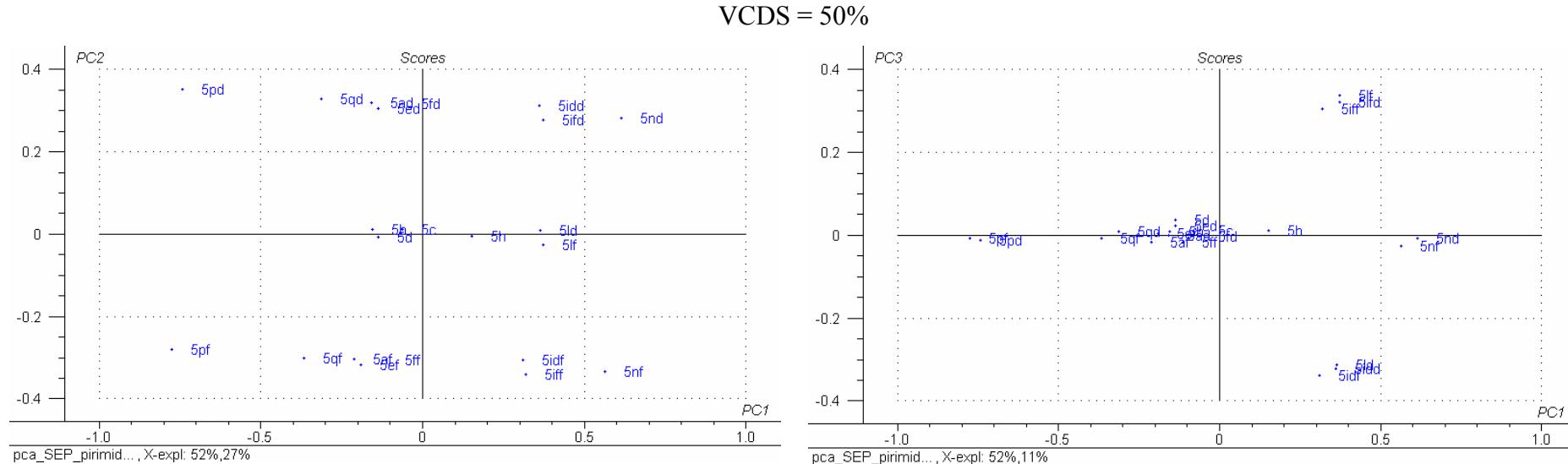
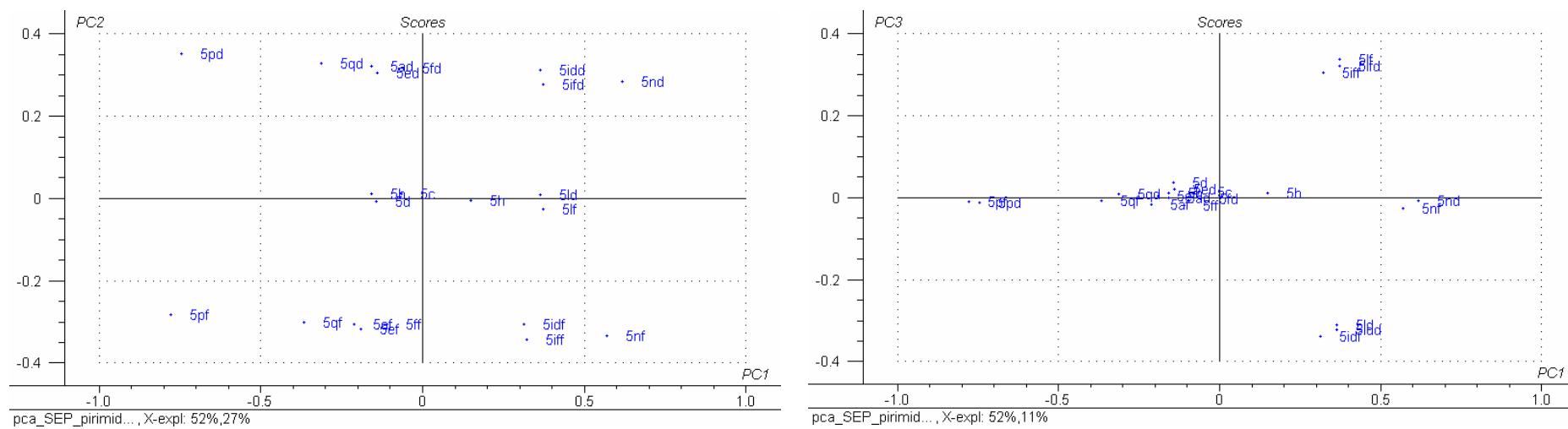
PC1 x PC3

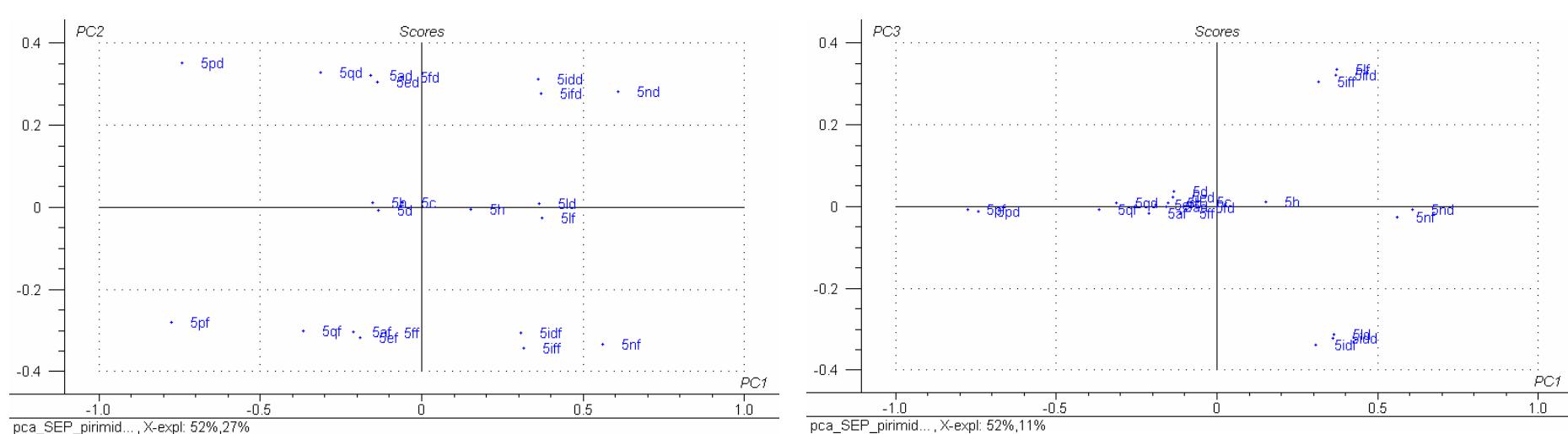
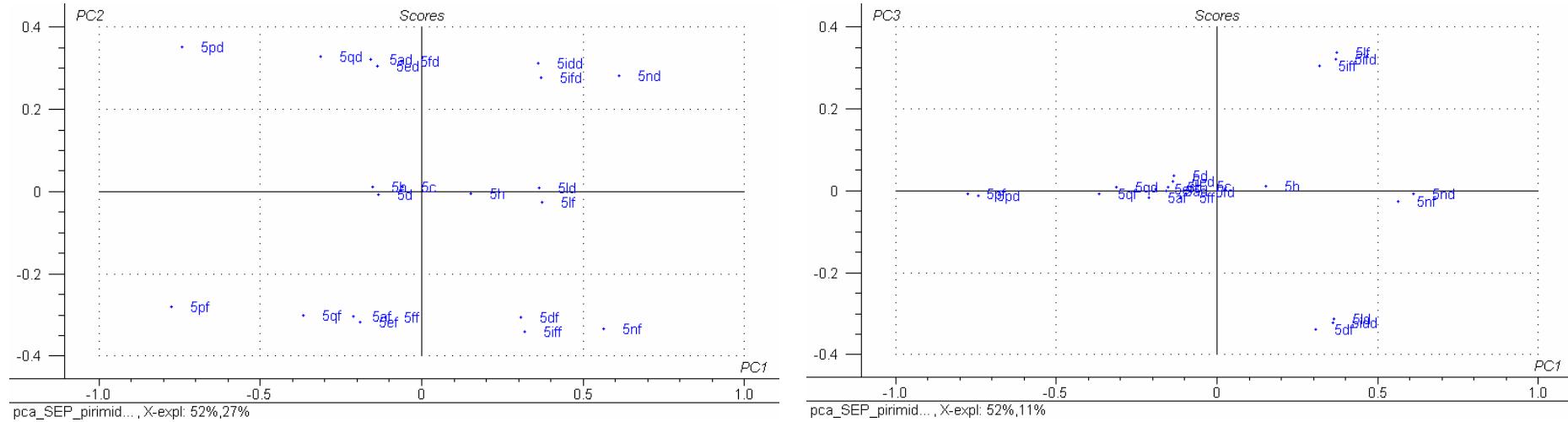


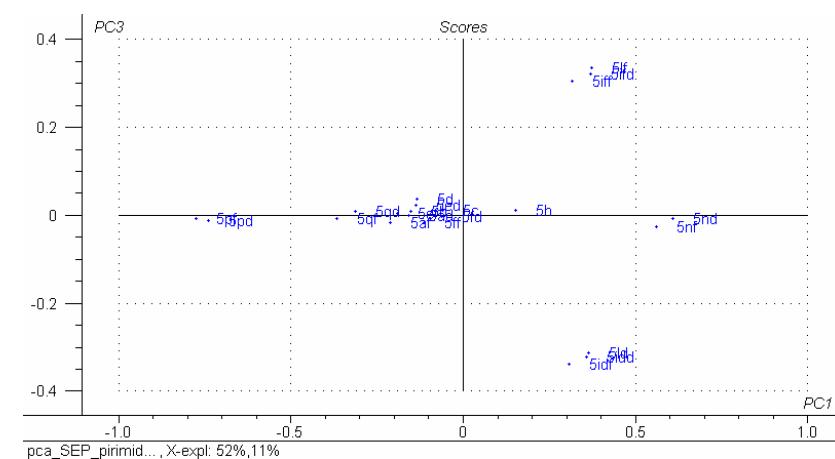
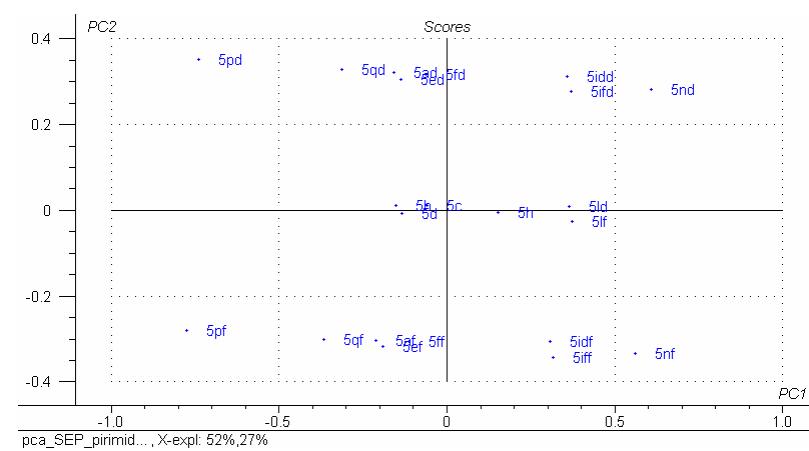
VCDS = 30%



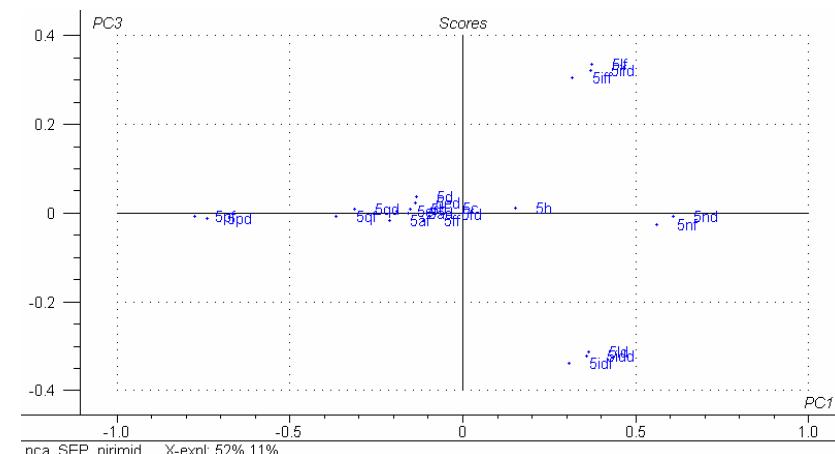
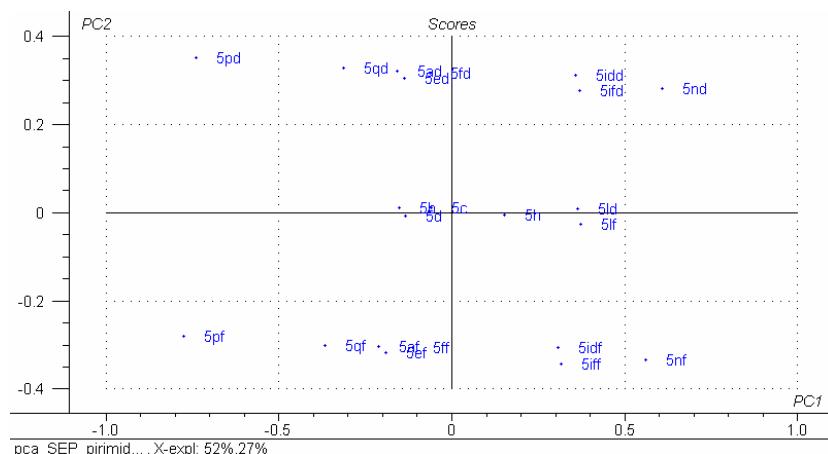
VCDS = 40%







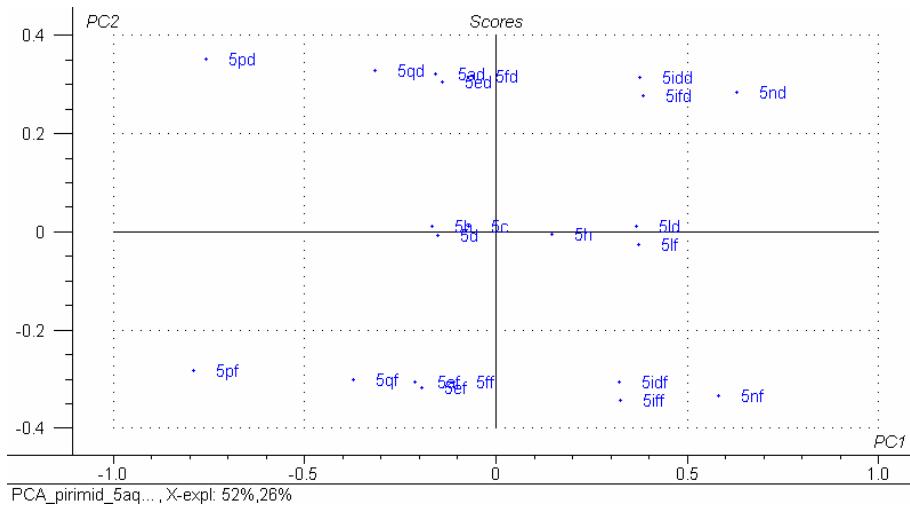
VCDS = 90%



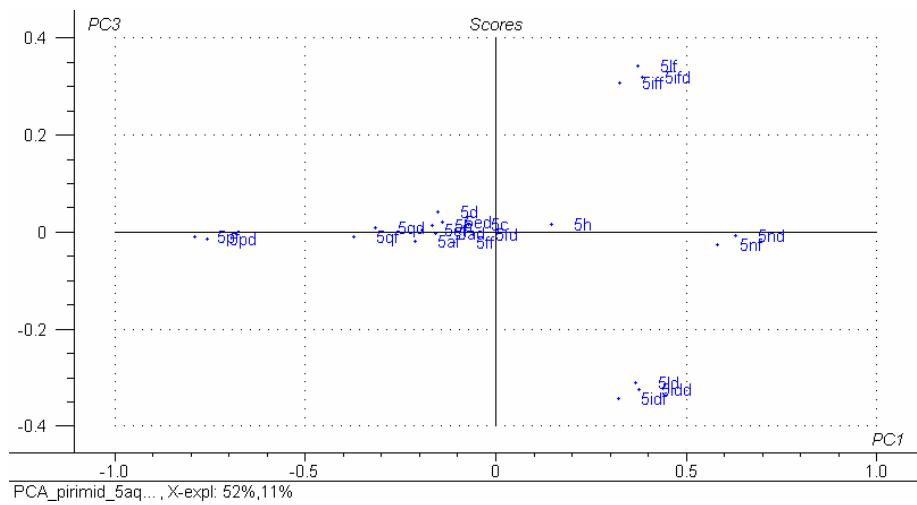
VCDS = 100%

**Figura C1:** Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas 5a-q totalmente otimizadas com VCIS = 20% e  $30\% < \text{VCDS} < 100\%$  da caixa 5.

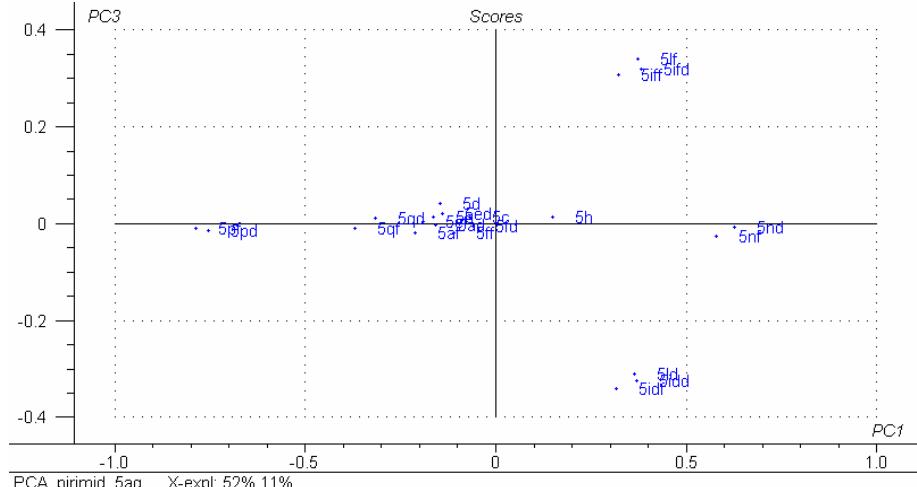
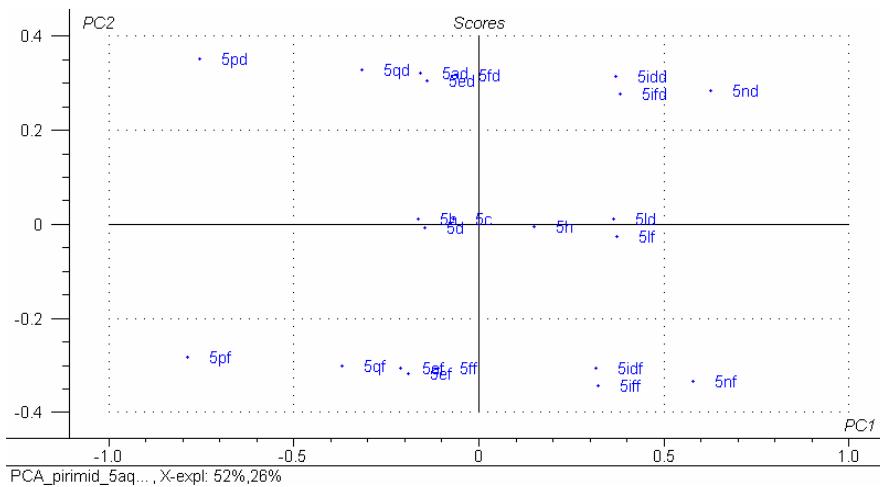
PC1 x PC2



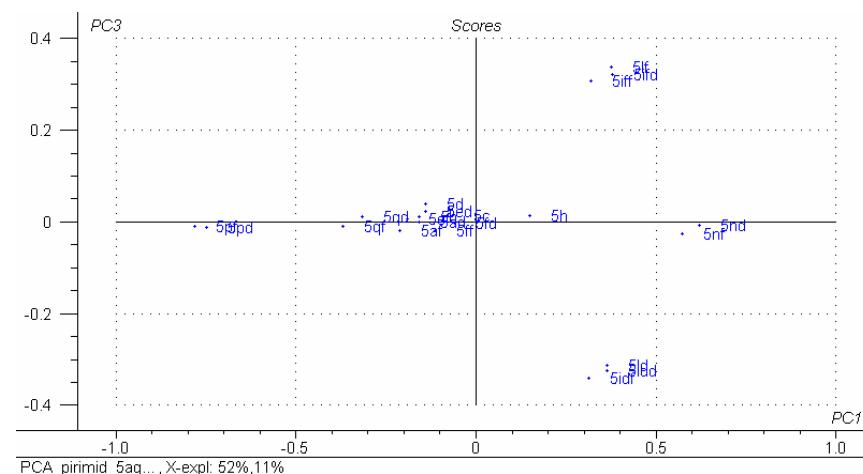
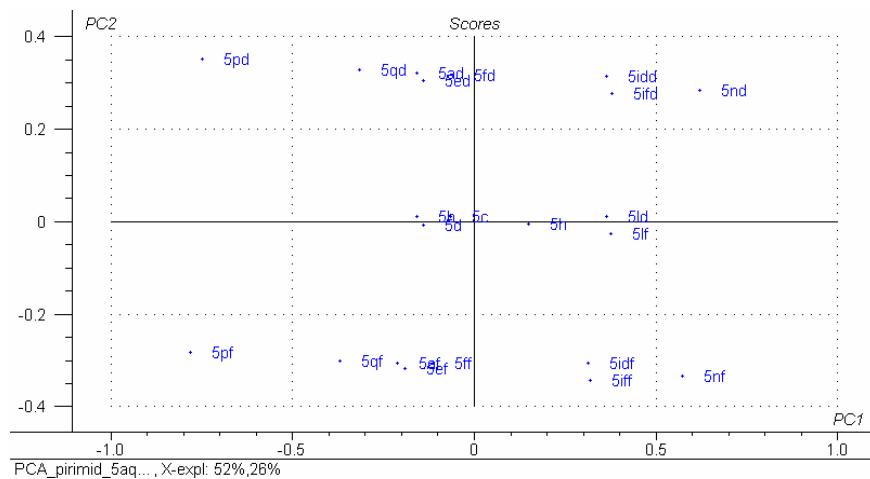
PC1 x PC3



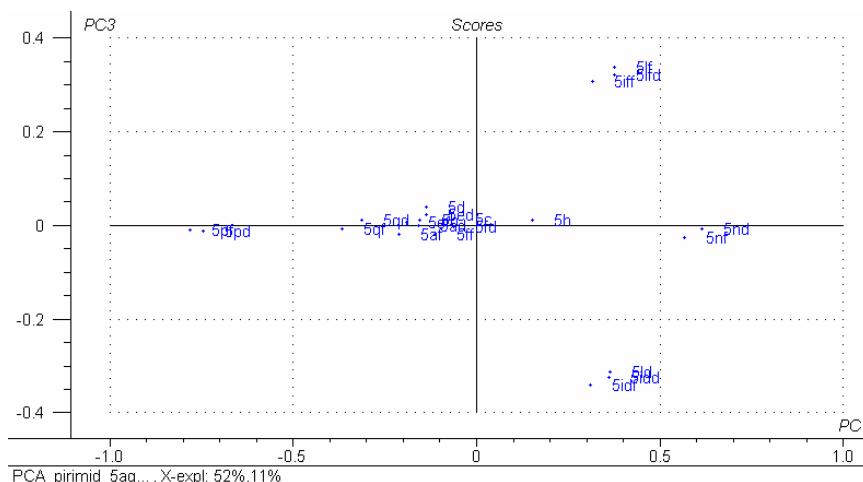
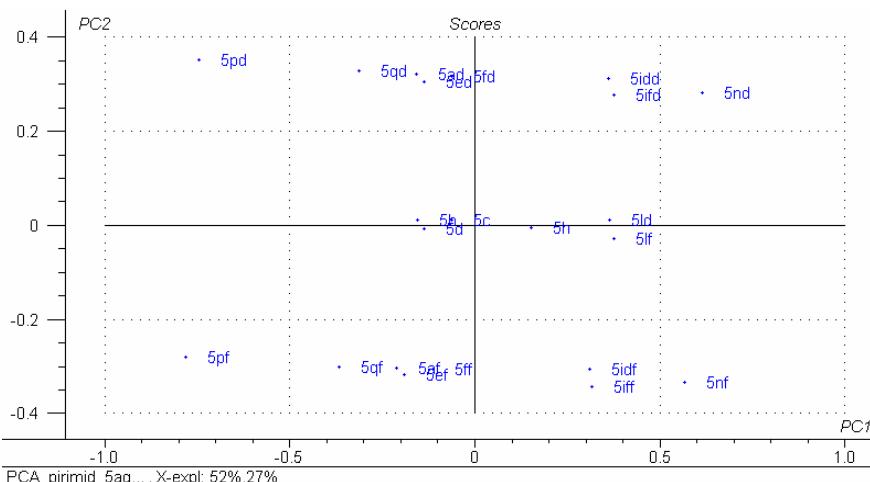
VCDS = 30%



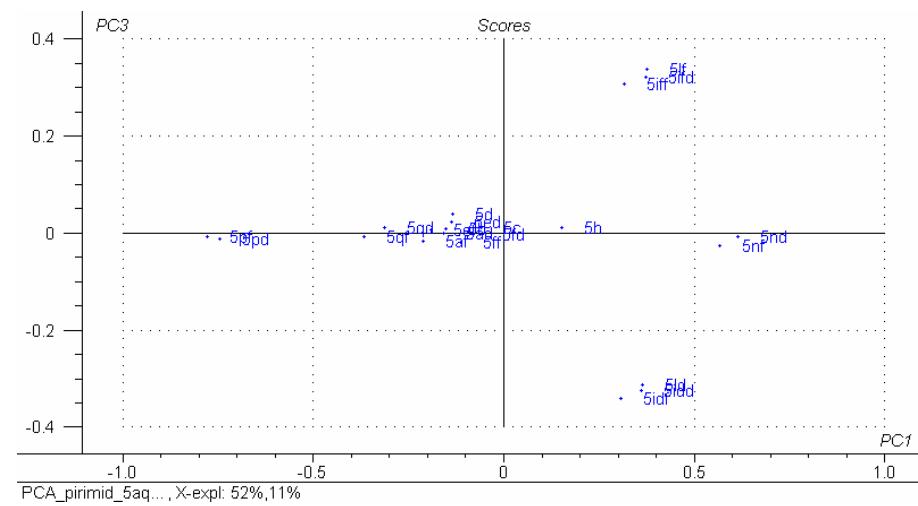
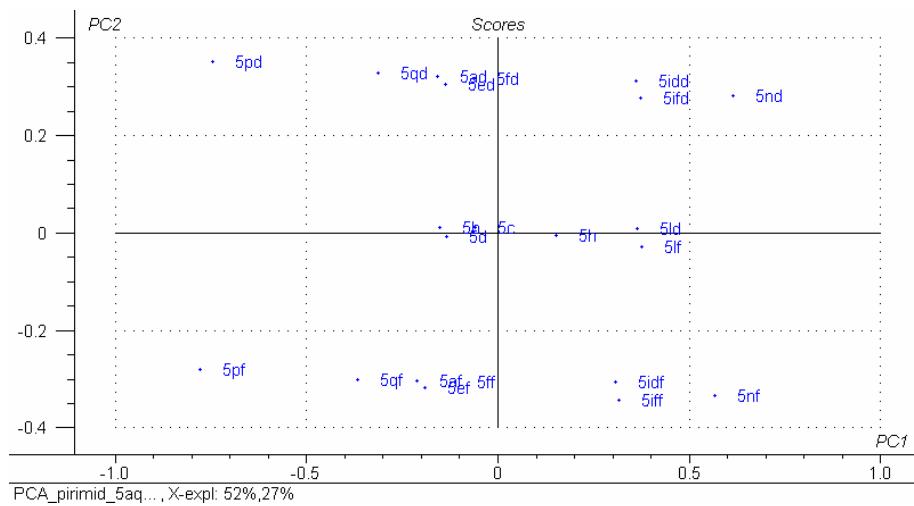
VCDS = 40%



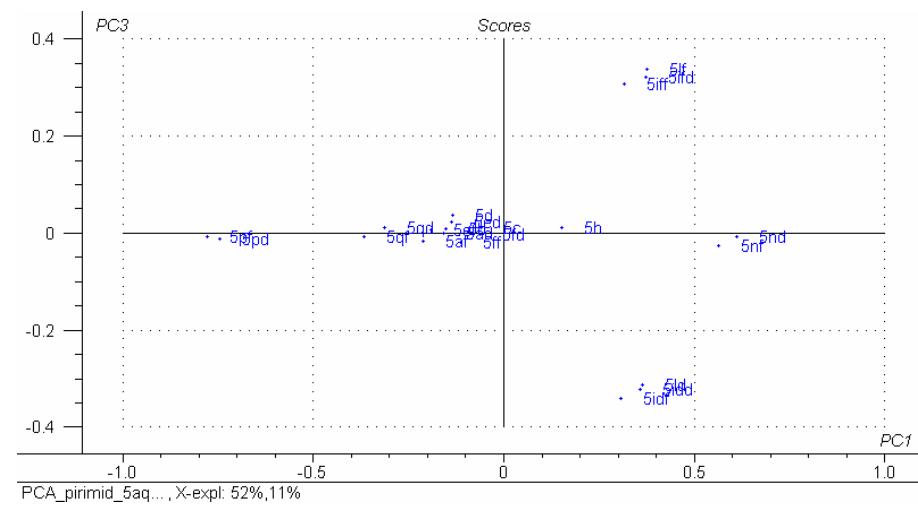
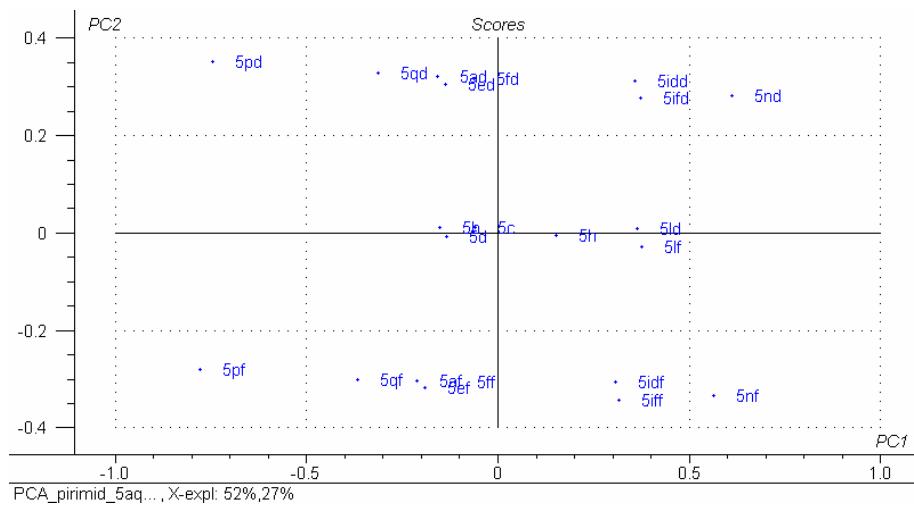
VCDS = 50%



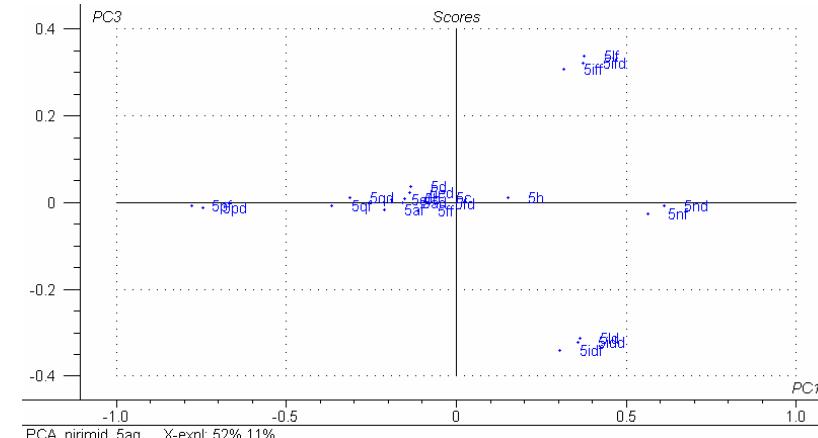
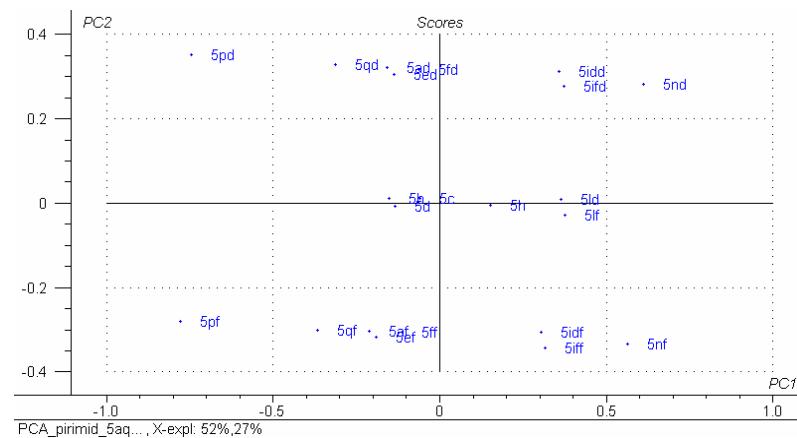
VCDS = 60%



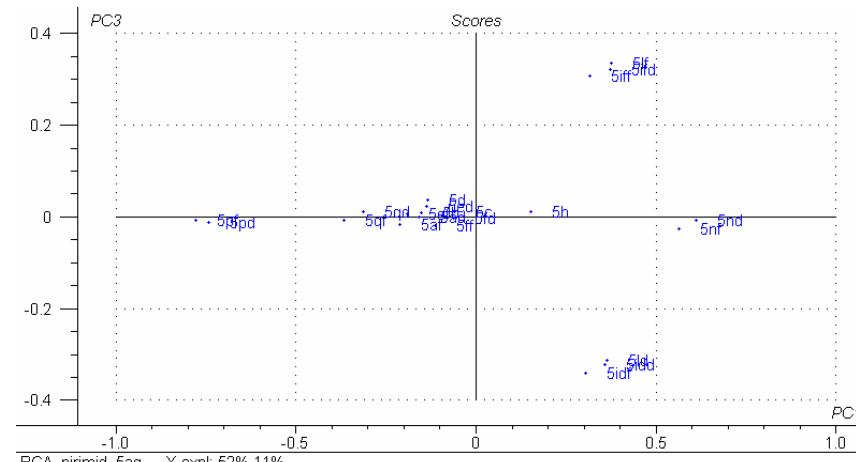
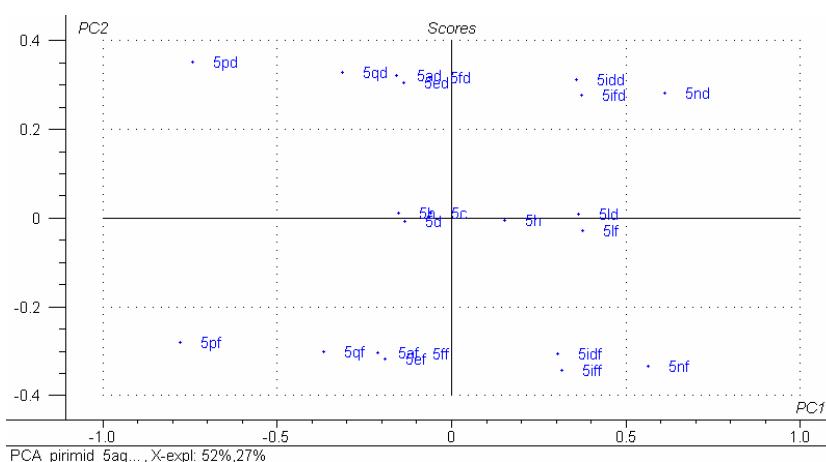
VCDS = 70%



VCDS = 80%



VCDs = 90%



VCDs = 100%

**Figura C2:** Gráfico dos escores de PC1 x PC2 e PC1 x PC3 da SPE das pirimidinas 5a-q totalmente otimizadas com VCIS = 20% e  $30\% < \text{VCIS} < 100\%$  da caixa 8.

