

Universidade Federal de Juiz de Fora
Pós-Graduação em Ciências Exatas
Mestrado em Física

Mônica Abreu Silva

**Propriedades Estruturais e Eletrônicas de análogos
da Acridina**

Juiz de Fora
Dezembro de 2013

Mônica Abreu Silva

Propriedades Estruturais e Eletrônicas de análogos da Acridina

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Física, da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Mestre.

Orientador: Prof. Dr. Paulo M.V.B. Barone

Juiz de Fora
Dezembro de 2013

Propriedades Estruturais e Eletrônicas de análogos da Acridina

Por

Mônica de Abreu Silva

Dissertação apresentada ao Instituto de Física da Universidade Federal de Juiz de Fora como parte dos requisitos necessários para a obtenção do título de Mestre em Física.

Aprovada por:

Prof. Dr. Paulo M. V. B. Barone

(Orientador) IF-UFJF

Prof. Dr. Sócrates de Oliveira Dantas

IF-UFJF

Prof. Dr. Marcone Augusto Leal de Oliveira

IQ-UFJF

Juiz de Fora,

Prof. Dr. Sócrates de Oliveira Dantas

Coordenador de Pós-Graduação

Instituto de Física

Universidade Federal de Juiz de Fora

Agradecimentos

Primeiramente à Deus pelo dom da vida, e por ser meu refúgio e a minha fortaleza.

Aos meus pais Ilarino Gomes da Silva e Euronice Rosa de Abreu que me ensinaram o valor do respeito, da justiça e da dignidade, e que com dedicação e carinho me apresentaram o amor com tudo que ele é capaz de abarcar. As minhas irmãs Adriana de Abreu Silva e Fernanda de Abreu Silva, pelo carinho, motivação e confiança; pelas conversas, brincadeiras e sonhos compartilhados.

Ao professor Paulo Barone, pela orientação, confiança, amizade e paciência. Alguém que admiro muito profissionalmente e como pessoa, com quem tive a alegria e a honra de trabalhar.

Ao professor Fernando Sato, pelo trabalho realizado com grandeza na disciplina de Métodos Computacionais, pela dedicação exercida com o laboratório, com os computadores, pra que tudo funcione, pela ajuda com os programas e com a parte de simulação, que me ajudou muito no início do meu trabalho de pesquisa, e pela amizade.

Aos professores do departamento de Pós-graduação em Física com quem cursei disciplinas que foram essenciais na minha formação (Sidney, Pablo, Maria Cristina).

A CAPES pelo apoio financeiro, a UFJF e ao Departamento de Pós-graduação em Física pelo excelente espaço físico e quadro de profissionais. A Domingos Lopes pela seriedade e eficiência com que realiza seu trabalho.

Aos colegas do curso de pós-graduação em física, Jefferson, Leonarde (Geral), Danilo (que me ajudou muito com o Latex), Poliane, Mônica, Monaliza, Lucélia, Dante, Cleber, Leandro (me auxiliando com os programas nessa última etapa do trabalho), Allison (os dois), Tatiane, Vanessa, Wesley, a vocês todos pelas conversas sempre muito produtivas, com exceção daquelas que acabavam em piadinhas sobre baiano, e pelos momentos de descontração que me proporcionaram. Aos colegas de curso, Max, João, Chico, Camila, Aninha, Isis.

A Sheila, Floriano, Flávia e Lorenza, pelo carinho e cuidado que tiveram comigo, sempre me trazendo palavras de coragem e de amor.

A Larissa e Gabriel, que me acolheram com carinho e respeito, pela paz e tranquilidade que me passaram, pelas risadas, pipocas e pelo apoio nessa última fase.

Aos amigos queridos Góe, Dane, Ledo, Dai, Dinha, Andressa, Larissa, Maíra, Fábio, Ênio, Rosinha, Nádia, Dobes, Rose, Rene, Ilza, Vilma, Adê, Zete, Élide, Cal, Geni, Marília, Tony por cultivarem com tamanha leveza e verdade a nossa relação de amizade.

Aos meus familiares, especialmente a Gerson, Lita, Dete, Elza, Ilda, Eleni, Nalva,

Em especial, a Robson Sampaio, pelo companheirismo, compreensão, força e amor que tem me dedicado.

*“A ciência existe, antes de tudo, para aliviar a nossa perplexidade natural
e desenvolver novas faculdades, em nossa vida cada vez mais intensa,
para evitar o nocivo e facilitar o útil. ”*
— *Johann Wolfgang von Goethe*(1749 - 1832)

Resumo

O câncer é um problema mundial crescente, em vista do gradual aumento da expectativa de vida e de mudanças no estilo de vida em grande parte dos países, entre outros fatores. Informações acerca de mecanismos carcinogênicos configuram-se importantes instrumentos na prevenção e controle do câncer. São também relevantes o desenvolvimento de novos fármacos para a quimioterapia do câncer, assim como a compreensão dos seus mecanismos de ação. O entendimento de relações estrutura-atividade de compostos que apresentam atividade biológica tem perfil chave na investigação do comportamento de compostos carcinogênicos no nível molecular e no desenvolvimento de fármacos com atividade otimizada. Este trabalho apresenta um estudo das propriedades eletrônicas e das relações estrutura-atividades de compostos análogos da acridina. Esses compostos foram anteriormente estudados por Cuny e colaboradores que, utilizando a técnica TR-FRET (Time-Resolved Fluorescence Resonance Energy Transfer), os classificaram como potentes inibidores de duas enzimas, Haspin e DYRK2, que participam do processo de mitose celular. O objetivo principal do presente trabalho foi estabelecer uma relação entre os dados obtidos através dos cálculos de propriedades eletrônicas e os resultados obtidos experimentalmente, com a finalidade de classificar estes compostos quanto a sua atividade biológica, por meio da Metodologia de Índices Eletrônicos (MIE). A otimização de geometria, bem como os cálculos de estrutura eletrônica, foram realizados pelos métodos semiempíricos AM1 e PM3, e pelo método ab initio DFT. Para estes cálculos utilizamos o software Spartan, como ferramenta visual e para análise conformacional. Os cálculos de estrutura eletrônica foram realizados com o programa GAMESS. O programa Chem2Pac foi utilizado para calcular a densidade local de estados, com o objetivo de estudar as relações estrutura-atividade por meio da MIE. Os métodos estatísticos multivariados PCA e HCA foram também utilizados na análise, por meio do programa Einsigth. Utilizando os parâmetros da MIE, estabelecemos regras simples, com as quais conseguimos identificar com precisão os compostos ativos e inativos. Os parâmetros eletrônicos encontrados pela MIE mostraram-se importantes para as outras metodologias de reconhecimento de padrões para esta família de moléculas.

Abstract

The Cancer is a growing worldwide problem and has shown an epidemiological profile. Informations about the carcinogenic mechanisms configure themselves as important tools in cancer control and prevention. The understanding of structure-activity relationships in compounds who present biological activity is highly relevant on the investigation of the carcinogenic compounds behavior in molecular systems and in the activity-optimized drugs developing. We performed an study of the eletronic properties from 14 acridine similar compounds. These compounds were former studied by Cuny et al and within an esperimental study with the TR-FRET thec-nique (Time-Resolved Florescence Resonance Energy Transfer), and were classified as potent inhibitors of two enzymes, Haspin and DYRK2, who had part on the conclusion mitosis process in several cells, inclusive cancerigen cells. The main objective of the present work was establish a relationship between obtained data from the calculations of eletronic properties and the experimentally obtained data with the aim of classify these compounds regarding it's biological activite, within the eletronic indices motodology (EIM). The geometric otimization, as the eletronic structure calculations, were made by the semi-empiric methody AM1 and PM3, by the DFT ab initio methody. For these calculations, we used firstly the software Spartan as visual tool and for the conformational analisys, then, the calculations were made with the GAMESS program. After the geometry was optimized, we used the Chem2pac in order to calculate the state local density. The PCA and HCA studies were performed by the Eeinsigth program. Th-rough the EIM parameters we established simple rules, with which we identified precisely the active and inactive compounds. In the investigations EIM, PCA and HCA, the acridine analo-gous activity was correlated with the theoretical parameters. The eletronic parameters found by EIM showed themselves largely important for others pattern recognition metodologies.

Lista de Figuras

Figura 1.1 - Exossomos de células de melanoma (verde) incorporados por células de vasos sanguíneos de pulmão (núcleos das células em azul)	20
Figura 1.2 - Divisão celular e crescimento desordenado de células cancerígenas. . .	20
Figura 1.3 - Processo de metastase, quando células cancerígenas se reproduzem e se espalham pelo corpo, formando novos tumores.(Foto: Instituto Nacional do Câncer-EUA)	21
Figura 1.4 - Estrutura do DNA sendo deformada pela inserção da cistoplatina (agente alquilante) que impede a transcrição levando a morte celular ou apoptose.	23
Figura 1.5 - Estrutura da Acridina	23
Figura 1.6 - Estrutura da Guanina.	24
Figura 1.7 - Esqueleto molecular dos compostos análogos da acridina.	25
Figura 2.1 - Resultados obtidos para MIE e por outros métodos, e algumas classes estudadas por Barone <i>et al.</i> [1].	36
Figura 2.2 - Visualização gráfica da variância das PCs	41
Figura 2.3 - Dendrograma ilustrativo construído com variáveis de 1 a 10 e relacionados em ordem aleatória	42
Figura 3.1 - Estrutura da molécula 1, derivado da acridina.	43
Figura 3.2 - Estrutura da molécula 2, derivado da acridina.	43
Figura 3.3 - Estrutura da molécula 3, derivado da acridina.	44

Figura 3.4 - Estrutura da molécula 4, derivado da acridina.	44
Figura 3.5 - Estrutura da molécula 5, derivado da acridina.	44
Figura 3.6 - Estrutura da molécula 6, derivado da acridina.	45
Figura 3.7 - Estrutura da molécula 7, derivado da acridina.	45
Figura 3.8 - Estrutura da molécula 8, derivado da acridina.	45
Figura 3.9 - Estrutura da molécula 9, derivado da acridina.	46
Figura 3.10 - Estrutura da molécula 10, derivado da acridina.	46
Figura 3.11 - Estrutura da molécula 11, derivado da acridina.	46
Figura 3.12 - Estrutura da molécula 12, derivado da acridina.	47
Figura 3.13 - Estrutura da molécula 13, derivado da acridina.	47
Figura 3.14 - Estrutura da molécula 14, derivado da acridina.	47
Figura 3.15 - Estrutura da molécula 1 depois de otimizada pelo método PM3.	52
Figura 3.16 - Estrutura da molécula 1 depois de otimizada pelo método DFT.	53
Figura 3.17 - Estrutura da molécula 2 depois de otimizada pelo método PM3.	53
Figura 3.18 - Estrutura da molécula 2 depois de otimizada pelo método DFT.	54
Figura 3.19 - Estrutura da molécula 7 depois de otimizada pelo método PM3.	55
Figura 3.20 - Estrutura da molécula 7 depois de otimizada pelo método DFT.	55
Figura 3.21 - Estrutura da molécula 6 depois de otimizada pelo método PM3.	56
Figura 3.22 - Estrutura da molécula 6 depois de otimizada pelo método DFT.	56
Figura 3.23 - Esqueleto molecular dos compostos análogos da acridina, especificando as regiões para o estudo MIE.	57
Figura 3.24 - Legenda com os nomes das região e suas respectivas representações na figura 3.23.	57

Figura 3.25 - Gráfico da análise da variável CHW, para a enzima Haspin com o método AM1.	71
Figura 3.26 - Gráfico da análise binária, com os parâmetros CL1W e CHW para a enzima Haspin com o método AM1.	73
Figura 3.27 - Gráfico da análise da variável CHP, para a enzima Haspin com o método PM3.	75
Figura 3.28 - Gráfico da análise binária, com os parâmetros LUMO+1 e CHK para a enzima Haspin com o método PM3.	77
Figura 3.29 - Gráfico da análise da variável HOMO, para a enzima DYRK2 com o método AM1.	79
Figura 3.30 - Gráfico da análise binária, com os parâmetros CHK e LUMO para a enzima DYRK2 com o método AM1.	81
Figura 3.31 - Gráfico da análise da variável CLW, para a enzima DYRK2 com o método PM3.	82
Figura 3.32 - Gráfico da análise binária, com os parâmetros HOMO-1 e CHP para a enzima DYRK2 com o método PM3.	84
Figura 3.33 - Gráfico loadings da PCA, para enzima Haspin, com o método AM1, mostrando a distribuição dos parâmetros no plano PC1 X PC2.	87
Figura 3.34 - Gráfico scores da PCA, para enzima Haspin, com o método AM1, mostrando a distribuição das moléculas no plano PC1 X PC2.	88
Figura 3.35 - Gráfico loadings da PCA, para enzima Haspin, com o método PM3, mostrando a distribuição dos parâmetros no plano PC1 X PC2.	89
Figura 3.36 - Gráfico scores da PCA, para enzima Haspin, com o método PM3, mostrando a distribuição das moléculas no plano PC1 X PC2.	90

Figura 3.37 - Gráfico loadings da PCA, para enzima DYRK2, com o método AM1, mostrando a distribuição dos parâmetros no plano PC1 X PC2.	91
Figura 3.38 - Gráfico scores da PCA, para enzima DYRK2, com o método AM1, mostrando a distribuição das moléculas no plano PC1 X PC2.	92
Figura 3.39 - Gráfico loadings da PCA, para enzima DYRK2, com o método PM3, mostrando a distribuição dos parâmetros no plano PC1 X PC2.	94
Figura 3.40 - Gráfico scores da PCA, para enzima DYRK2, com o método PM3, mostrando a distribuição das moléculas no plano PC1 X PC2.	95
Figura 3.41 - Gráfico com o resultado para HCA da enzima Haspin, pelo método AM1, especificando as moléculas inativas pela cor vermelha e as ativas, preto.	95
Figura 3.42 - Gráfico com o resultado para HCA da enzima Haspin, pelo método PM3, especificando as moléculas inativas pela cor vermelha e as ativas, preto.	96
Figura 3.43 - Gráfico com o resultado para HCA da enzima DYRK2, pelo método AM1, especificando as moléculas inativas pela cor vermelha e as ativas, preto.	96
Figura 3.44 - Gráfico com o resultado para HCA da enzima DYRK2, pelo método PM3, especificando as moléculas inativas pela cor vermelha e as ativas, preto.	97

Lista de Tabelas

Tabela 1.1 - Tabela com os compostos estudados, especificando os radicais e as posições.	25
Tabela 2.1 - Tabela Booleana para classificação da atividade dos compostos, $+=\Delta$ e η para compostos analisados $> \eta_c$ e Δ_c e $-=\Delta$ e η para compostos analisados $< \eta_c$ e Δ_c .	36
Tabela 3.1 - Tabela com os valores do Calor de Formação dos 14 compostos para os métodos AM1 e PM3.	49
Tabela 3.2 - Tabela com os resultados para as energias dos orbitais de fronteira e suas respectivas diferenças para o método PM3.	51
Tabela 3.3 - Tabela com os resultados para as energias dos orbitais de fronteira e suas respectivas diferenças para o método DFT.	52
Tabela 3.4 - Tabela com os resultados para as energias dos orbitais de fronteira, suas respectivas diferenças, o gap de energia e o calor de formação (C.F.), para o método PM3.	58
Tabela 3.5 - Tabela com os resultados para a região P.	63
Tabela 3.6 - Tabela com os resultados para a região W.	65
Tabela 3.7 - Tabela com os resultados para a região K.	66
Tabela 3.8 - Tabela com os valores do IC_{50} para a classificação da atividade biológica dos 14 compostos para as enzimas Haspin e DYRK2.	68
Tabela 3.9 - Tabela com os resultados de análise com uma variável para enzima Haspin, utilizando o método AM1.	69

Tabela 3.10 - Tabela booleana da análise com uma varável, para enzima Haspin, utilizando o método AM1.	70
Tabela 3.11 - Tabela com os melhores resultados de análise binária para enzima Haspin, utilizando o método AM1.	71
Tabela 3.12 - Tabela booleana para análise binária feita com as variáveis CL1W e CHW para a enzima Haspin, utilizando o método AM1, com resultados da predição comparados com os obtidos experimentalmente.	72
Tabela 3.13 - Tabela com os resultados de análise com uma variável para enzima Haspin, utilizando o método PM3.	74
Tabela 3.14 - Tabela booleana da análise com uma varável, para enzima Haspin, utilizando o método PM3.	75
Tabela 3.15 - Tabela com os resultados de análise binária para enzima Haspin, utilizando o método PM3.	76
Tabela 3.16 - Tabela booleana para análise binária feita com as variáveis LUMO+1 e CHK para a enzima Haspin, utilizando o método PM3, com resultados da predição comparados com os obtidos experimentalmente.	77
Tabela 3.17 - Tabela com os resultados de análise com uma variável para enzima DYRK2, utilizando o método AM1.	78
Tabela 3.18 - Tabela booleana da análise com uma varável, para enzima DYRK2, utilizando o método AM1.	79
Tabela 3.19 - Tabela com os resultados de análise binária para enzima DYRK2, utilizando o método AM1.	80
Tabela 3.20 - Tabela booleana para análise binária feita com as variáveis CHK e LUMO para a enzima DYRK2, utilizando o método AM1, com resultados da predição comparados com os obtidos experimentalmente.	80

Tabela 3.21 - Tabela com os resultados de análise com uma variável para enzima DYRK2, utilizando o método PM3.	81
Tabela 3.22 - Tabela booleana da análise com uma varável, para enzima DYRK2, utilizando o método PM3.	82
Tabela 3.23 - Tabela com os resultados de análise binária para enzima DYRK2, utilizando o método PM3.	83
Tabela 3.24 - Tabela booleana para análise binária feita com as variáveis HOMO-1 e CHP para a enzima Haspin, utilizando o método PM3, com resultados da predição comparados com os obtidos experimentalmente.	83
Tabela 3.25 - Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método AM1, para a enzima Haspin.	86
Tabela 3.26 - Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método PM3, para a enzima Haspin.	88
Tabela 3.27 - Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método AM1, para a enzima DYRK2.	90
Tabela 3.28 - Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método PM3, para a enzima DYRK2.	93

Lista de Abreviaturas

<i>INCA</i>	Instituto Nacional do Câncer
<i>HAPs</i>	Hidrocarbonetos Aromáticos Policíclicos
<i>HAH</i>	Hidrocarbonetos Aromáticos Heterocíclicos
<i>DNA</i>	Ácido desoxirribonucleico
<i>SAR</i>	Structure Activity Relationship (Relação de Estrutura e Atividade)
<i>MIE</i>	Metodologia de Índices Eletrônicos
<i>HOMO</i>	“Highest Occupied Molecular Orbital” (Energia do último orbital molecular ocupado)
<i>HOMO-1</i>	(Energia do penúltimo orbital molecular ocupado)
<i>LUMO</i>	“Lowest Unoccupied Molecular Orbital” (Energia do primeiro orbital virtual ocupado)
<i>LUMO+1</i>	(Energia do segundo orbital virtual ocupado)
<i>PCA</i>	Principal Component Analysis (Análise de Componentes Principais)
<i>HCA</i>	Hierarchical Cluster Analysis (Análise de Agrupamentos Hierárquicos)
<i>Haspin</i>	Gene Codificador de Proteína quinase nuclear específica de CG haplóides
<i>DYRK2</i>	Dual-specificity tyrosine-(Y)-phosphorylation regulated
<i>DFT</i>	“Density Functional Theory” (Teoria do Funcional de Densidade)
<i>B3LYP</i>	Funcional híbrido de troca de Becke (B3) combinado com o funcional de correlação de LYP (Lee, Yang, Parr)
<i>SCF</i>	“Self Consistent Field” (Campo Auto Consistente)
<i>RHF</i>	“Restricted Hartree-Fock” (Hartree Fock Restrito)
<i>PM3</i>	“Parametric Method Number 3” (Método Paramétrico Número 3)
<i>DOS</i>	Densidade de Estados
<i>LDOS</i>	Densidade Local de Estados
<i>LCAO</i>	Combinação Linear de Orbitais Atômicos

Sumário

1	INTRODUÇÃO	17
1.1	Carcinogênese	19
1.2	Quimioterápicos Antineoplásicos	21
1.3	Acridina e seus Análogos	22
2	METODOLOGIA	27
2.1	Métodos <i>ab initio</i>	27
2.1.1	Teoria do Funcional da Densidade	27
2.1.2	Funcional de Troca e Correlação	30
2.2	Métodos Semi-empíricos	32
2.3	Metodologia de índices Eletrônicos (MIE)	33
2.4	Métodos Estatísticos Multivariados	37
2.4.1	Análise de Componentes principais (PCA)	37
2.4.2	Análise de Agrupamento Hierárquico (HCA)	41
3	RESULTADOS E DISCUSSÕES	43
3.1	Processo de Otimização	48
3.2	Estudo das Relações Estruturais dos análogos da Acridina	55
3.2.1	Estudo MIE para a enzima Haspin	68

3.2.2	Estudo MIE para a enzima DYRK2	77
3.2.3	Estudo PCA para a enzima Haspin	84
3.2.4	Estudo PCA para a enzima DYRK2	90
4	CONCLUSÃO	98
4.1	PERSPECTIVAS FUTURAS	100
	Apêndice A – O MÉTODO HARTREE-FOCK	101
A.1	A Teoria do Orbital Molecular	102
A.2	Equação de Hartree-Fock	108
A.3	A equação de Hartree-Fock-Roothaan	111
	Apêndice B – FORMALISMO DOS MÉTODOS SEMI-EMPÍRICOS	114
B.1	Equações MNDO	115
	Referências	118

1 INTRODUÇÃO

Segundo o Instituto Nacional de Câncer (INCA), dados do ano de 2011, computados pelo Ministério da Saúde (MS), apontam o câncer como a segunda causa de óbitos no Brasil, com propensão de crescimento para os próximos anos. Essa doença configura-se um dos problemas de saúde pública mais complexos enfrentados, atualmente, pelo sistema de saúde brasileiro, principalmente pelo perfil epidemiológico assumido por essa doença [2].

O surgimento de câncer está associado a vários fatores como hereditariedade, estilo de vida e exposição a agentes físicos, biológicos e químicos, como radiação, substâncias químicas, vírus, entre outros. Uma das causas do câncer é a exposição aos carcinógenos químicos, que compreendem substâncias químicas facilmente encontradas na natureza. Esses carcinógenos podem ter ação carcinogênica direta ou indireta. Os agentes alquilantes são carcinogênicos diretos bem conhecidos; no entanto, a maior parte das substâncias químicas tem efeito indireto, necessitando passar por modificações químicas no organismo para ativá-las. Uma grande quantidade de enzimas solúveis ou ligadas a membranas, como a P-450, considerada uma das mais importantes, é responsável por metabolizar os carcinógenos e torná-los eletrofílicos [1, 3].

Os Hidrocarbonetos Aromáticos Policíclicos (HAP) representam uma família conhecida de agentes carcinogênicos. Quando submetidos à ação metabólica, os HAP podem induzir tumores em diversos tecidos e espécies, como tumores cutâneos (quando aplicados sobre a pele), sarcomas (se injetados via subcutânea) e causar tumores bem localizados (se introduzidos num órgão específico). Os HAPs são produzidos na fumaça de cigarro, a partir da gordura animal no processo de cozimento de carnes e encontram-se também em carnes defumadas e peixes [4].

Os HAP são compostos orgânicos que apresentam na cadeia principal um ou mais anéis benzênicos ligados de forma linear, angular ou agrupados. Essas substâncias bem como seus derivados, podem ser nitradas ou oxigenadas, constituídos quando um carbono do anel é substituído por um heteroátomo, dos hidrocarbonetos aromáticos heterocíclicos, estes são encontrados como constituintes de misturas complexas de diferentes concentrações em vários compartimentos ambientais.

A descoberta de compostos que poderiam causar cancro data do ano de 1775 em Londres, pelo cirurgião do hospital St. Bartholomew Percival Pott, ao observar que a maior incidência de cancro na cavidade nasal e na pele dos testículos ocorria, quase exclusivamente, em pacientes que trabalhavam como limpadores de chaminés, geralmente garotos órfão ou senhores que quase nus e com o copro coberto de óleo limpavam as chaminés. Pott foi o primeiro a associar o câncer com um fator ambiental, ele ligou o fato do aparecimento de câncer à exposição crônica à fuligem. Mas, como na época ainda não havia recursos tecnológicos para identificar quais compostos químicos eram responsáveis pela formação do cancro, somente em 1920 é que dois japoneses Lchikawa e Yamagina confirmaram a descoberta de Pott, ao observar que ratos submetidos a exposição contínua do alcatrão da hulha apresentavam tumor de pele. No alcatrão foi isolado o primeiro hidrocarboneto puro cancerígeno, o dibenzo[a, h]antraceno DahA [5, 6].

A emissão de HAPs se dá por duas fontes, as naturais e as antropogênicas (pela ação do homem)- que representa o principal processo de reprodução de HAPs. As fontes naturais restringem-se a erupções vulcânicas e queima espontâneas de florestas devido a efeitos naturais, petróleo (já que são constituintes naturais), carvão, alguns HAPs como benzo(a) antraceno podem estar presentes na biossíntese de bactérias, em algumas algas e vegetais superiores. Uma fonte predominante de HAP é a combustão incompleta de matéria orgânica, que diz respeito à pirólise de madeira para a produção de carvão e de querosene para a obtenção do tolueno e benzeno, incineração de resíduos domésticos, queima de combustíveis fósseis, emissão de fumaça originada de cigarro, água contaminada, cozimento de alimentos, exaustão de automóveis, etc. Devido a grande concentração destes compostos no meio ambiente, ocorre uma grande exposição do homem aos HAP. Para o homem as vias de penetração podem ser diver-

sas, como inalação (devido a presença destas substâncias na atmosfera, ingestão de alimentos, devido a contaminação do solo por meio de atividades industriais ou depósito de lixo urbano ou ainda de fontes hídricas contaminadas. [7, 8, 9]

Os HAPs constituem uma das principais famílias de agentes cancerígenos, e uma das classes mais investigadas. Tanto os HAPs quanto seus derivados estão associados a diversos tipos de neoplasias e processos hiper proliferativos. O fato de que a exposição do homem aos HAPS pode levar ao desenvolvimento de neoplasias fez com que houvesse um grande interesse pela comunidade científica em desenvolver metodologias analíticas que com agilidade e perfeição atuem no sentido de identificar os bioindicadores das concentrações desses compostos, presentes nos sítios de ação biológica. Devido à complexidade dos HAPs e as suas baixas concentrações, tem-se exigido que estes métodos analíticos sejam extremamente seletivos e sensíveis [10, 11].

1.1 Carcinogênese

Os HAPs são altamente lipossolúveis e podem ser absorvidos pelos pulmões, intestinos e pela pele. No entanto, os processos metabólicos fazem com que estes sofram uma rápida degradação; então os níveis encontrados no organismo, para uma pequena exposição, não são significativos. O processo de passagem dos HAPs pelo organismo se dá da seguinte forma: primeiramente eles são absorvidos pelas células e ativados após metabolizados por uma super-família de Citocromo P450. E são metabolicamente ativados, depois de sofrerem reações de oxidação, redução, hidrólise e de conjugação, tornando-se reativos a grupos nucleofílicos de macromoléculas celulares, formando adutos de DNA, tais adutos são essenciais para o desenvolvimento da carcinogênese [12].

A formação de um tumor acontece em vários estágios. No estágio de iniciação, as células encontram-se modificadas geneticamente; no segundo estágio é a promoção, as células alteradas sofrem uma seleção de células neoplásicas e competitivas e a expansão, e podem ser transformadas em células malignas de forma gradual (ver figura 1.2); no estágio final é a progressão

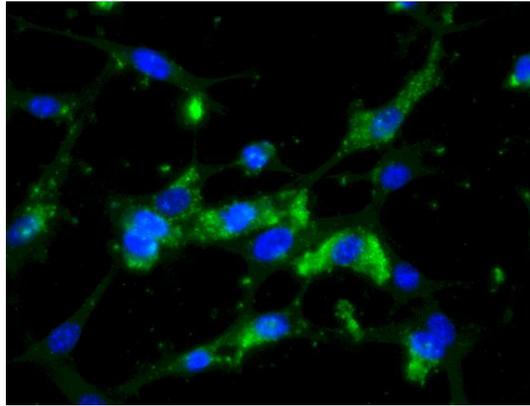


Figura 1.1 – Exossomos de células de melanoma (verde) incorporados por células de vasos sanguíneos de pulmão (núcleos das células em azul)

Fonte: Revista Nature Medicine, 2012

de uma lesão benigna ou pré-maligna, alterando o DNA novamente, agora nos genes supressores de tumor, nesta fase há um descontrole de crescimento celular, uma invasão de tecidos, metástase (ver figura 1.3). [13, 14].

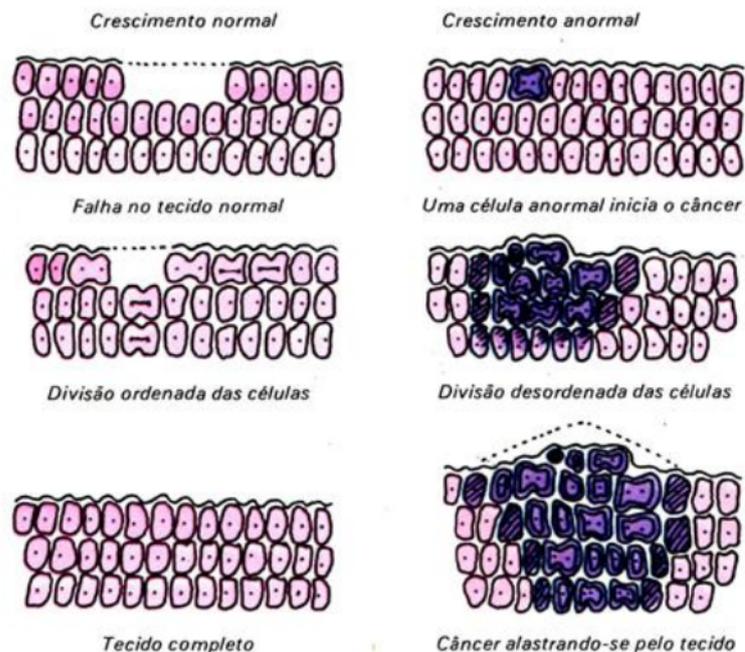


Figura 1.2 – Divisão celular e crescimento desordenado de células cancerígenas.

Aproximadamente 10% dos cânceres são hereditários, que originam-se de alterações nos genes, tornando o organismo mais suscetível a agentes químicos, físicos ou virais. Mas a maior parte dos cânceres é adquirida devido à exposição humana a agentes indutores de câncer, e

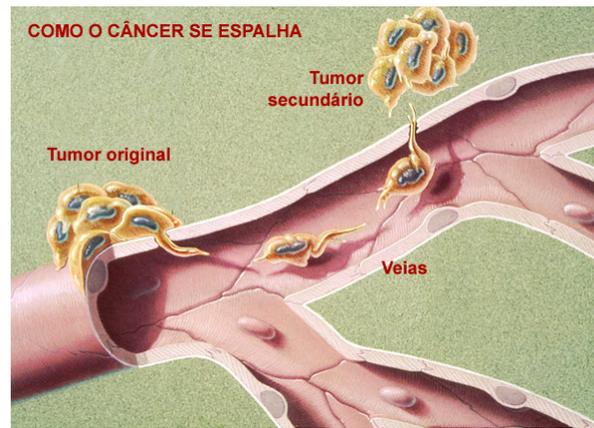


Figura 1.3 – Processo de metastase, quando células cancerígenas se reproduzem e se espalham pelo corpo, formando novos tumores. (Foto: Instituto Nacional do Câncer-EUA)

acontece de forma esporádica. Muitos genes estão associados ao surgimento ou desenvolvimento de diversos cânceres, entre estes os carcinomas, sarcomas, linfomas, leucemia, milomas, melanomas e neuroblastomas [15, 16].

Se o tumor estiver ainda no estágio inicial, tratamentos como cirurgias e terapias de radiação ionizantes são indicadas. A prática da cirurgia leva à remoção de tumores com sucesso, se encontrados na iniciação. Geralmente junto com a cirurgia, utiliza-se a radioterapia, com raios gama, radioisótopos, ou outras técnicas. Caso contrário, se estiver espalhado, mesmo que em uma região delimitada, a cirurgia torna-se inviável, e o tratamento mais indicado, neste caso, é a quimioterapia. A radiação ionizante é usada para tratar vários tipos de câncer; ionizando os tecidos a medida que o feixe incide sobre o mesmo. Na ionização das camadas superficiais de tecido dependendo da quantidade da dose ionizante usada, a célula pode morrer ou sofrer mudanças genéticas [13, 14].

1.2 Quimioterápicos Antineoplásicos

Considerando a importância da quimioterapia, o estudo sistemático dos agentes químicos antitumorais é extremamente importante, dos pontos de vista da classificação química, estrutural, e dos grupos funcionais diferentes presentes. Os quimioterápicos antineoplásicos são fármacos heterogêneos, em grande maioria de natureza tóxica. Devido a isso um dos grandes

desafios é encontrar compostos antineoplásicos com um grau mínimo de toxicidade [17].

Esses agentes são classificados quanto a sua estrutura química e sua função a nível celular. Os agentes alquilantes são um dos quimioterápicos utilizados no tratamento de câncer, fazem uma interação com o DNA, causando uma alteração irreversível, ao impedir a síntese do novo material genético. Fazem parte desse grupo as mostardas nitrogenadas, alquilsulfonados, nitrosuréias, triazenos. São compostos altamente eletrofílicos, que reagem com nucleófilos e formam ligações covalentes. Na molécula de DNA há vários grupos nucleófilos, como os nitrogênios presentes na base da adenina, na citosina e na guanina.[18].

Agentes intercalantes são moléculas que têm a aptidão de encaixar-se entre os pares de base da molécula de DNA, deformando a sua estrutura de dupla hélice, comprometendo a replicação e a transcrição. Esses agentes podem causar inserções ou deleções e ainda intercalarse em fita simples. A intercalação geralmente acontece com os pares de base da CG(citosina e guanina) e é uma interação do tipo π , que existe entre o DNA e os antineoplásicos. Os agentes intercalantes possuem anéis aromáticos e alta densidade eletrônica, que modificam a estrutura do DNA aumentando o espaçamento dos pares de base CG de 3,4 Å para 7Å, devido ao volume molecular, formando um ângulo de 90 graus aproximadamente, em relação ao eixo do DNA. Dentre as substâncias com propriedades intercalantes, sobressaem-se algumas, como as antraciclinas, dexorrubicina, daunorrubicina, actinomicinas e as acridinas [13, 19].

As acridinas fazem parte de um grupo de antineoplásicos que interagem com os pares da bases nitrogenadas por meio de interações de van der Waals, eletrostáticas e ligações de hidrogênio, conformações do tipo iso-hélicas. As acridinas bem como seus análogos são o grupo de compostos nos quais estamos interessados[13].

1.3 Acridina e seus Análogos

Acridinas são compostos orgânicos cuja estrutura química é formada por três anéis aromáticos condensados, similar à estrutura do antraceno, com um carbono, geralmente o do anel central, substituído por um heteroátomo de nitrogênio (ver figura 1.5).

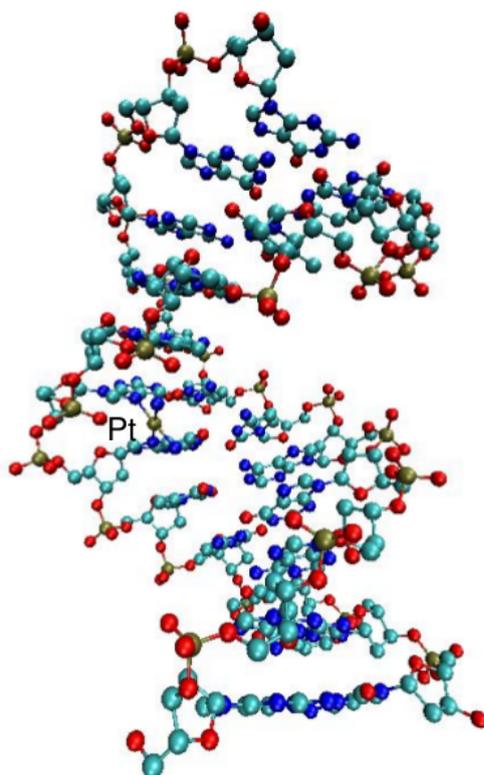


Figura 1.4 – Estrutura do DNA sendo deformada pela inserção da cistoplatina (agente alquilante) que impede a transcrição levando a morte celular ou apoptose.

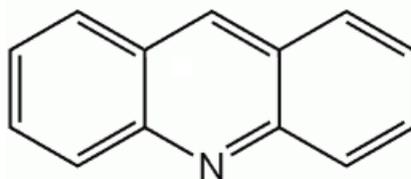


Figura 1.5 – Estrutura da Acridina

As acridinas possuem propriedades antitumorais que não se devem somente ao fato dessas moléculas interagirem com o DNA, mas também à sua habilidade de interagir com algumas enzimas indispensáveis à proliferação celular, como por exemplo, as telomerasas. Esses compostos podem reagir com o DNA ou intercalar-se na estrutura desta molécula, impedindo a replicação e a transcrição do DNA [20, 21].

Ao contrário da acridina, seus análogos formam adutos, que ligam-se principalmente com o grupo guanina (figura 1.6) da dupla hélice do DNA, conferindo a essas substâncias a característica de serem reativas, com um alto poder de redução no grupo nitro.

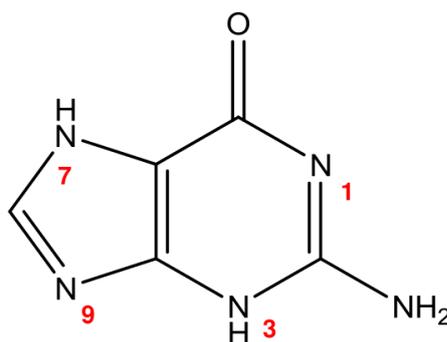


Figura 1.6 – Estrutura da Guanina.

Um grupo de moléculas derivadas da acridina foi estudado por Finlay et al., comprovando a preferência dessas moléculas em se ligarem ao par de bases CG na molécula de DNA [22].

A atividade carcinogênica das acridinas e alguns de seus análogos está relatada na literatura [13, 21, 23, 24]. O trabalho realizado por Cuny e colaboradores, desenvolveu uma investigação da atividade biológica de alguns derivados da acridina com as proteínas Haspin e DYRK2. O Haspin é uma quinase indispensável para a conclusão da mitose (processo de divisão celular que dá origem a uma nova célula eucariótica). Tanto a Haspin quanto a DYRK2 (Dual- specificity tyrosine-(Y)-phosphorylation regulated) fosforilam o H3T3 (uma das principais proteínas envolvidas na estrutura da cromatina em células eucarióticas) encontrando-se em vários órgãos e tecidos como baço, testículos e medula óssea, inclusive na população de células cancerígenas, que resulta em diversos tipos de neoplasias, como alguns linfomas [25].

Cuny, com a finalidade de compreender o comportamento desses compostos derivados da acridina (figuras 1.7 e tabela 1.1), no que se refere a atividade do Haspin e DYRK2 na mitose e em outros processos celulares e seu potencial inibidor em processos de câncer, realizou um estudo experimental utilizando a técnica TR-FRET (Time-Resolved Fluorescence Resonance Energy Transfer).

Experimentalmente observou-se que o esqueleto molecular planar é importante para que os compostos insiram-se no DNA, sem no entanto alterar consideravelmente a atividade biológica dos derivados. Os grupos metoxi ligados à estrutura foram decisivos para potencia-

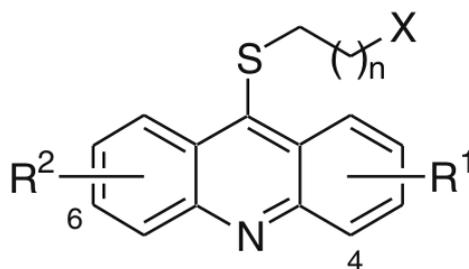


Figura 1.7 – Esqueleto molecular dos compostos análogos da acridina.

Tabela 1.1 – Tabela com os compostos estudados, especificando os radicais e as posições.

Composto	X	R ¹	R ²	n
1	NH ₂	2-OMe	7-OMe	2
2	NH ₂	H	H	2
3	NH ₂	2-OMe	H	2
4	NH ₂	2-OMe	7-Me	2
5	NH ₂	2-OMe	7-Cl	2
6	NH ₂	2-OH	7-OH	2
7	NH ₂	2-OMe,3-OMe	H	2
8	NH ₂	2-OMe,3-Cl	H	2
9	NH ₂	2-OMe	7-OMe	2
10	NH ₂	2-OMe	7-OMe	1
11	NHMe	2-OMe	7-OMe	3
12	NHMe ₂	2-OMe	7-OMe	2
13	Phth	2-OMe	7-OMe	2
14	Phth	2-OMe	H	2

lizar a inibição, uma vez que ao serem removidos a atividade inibitória caiu drasticamente. O cloro também mostrou-se importante quando substituiu o metoxi, mantendo o composto ainda como um potente inibidor. Curiosamente a transposição de um grupo metoxi para outra posição na estrutura, resultou na perda da atividade. A adição de grupos metileno interferiu pouco nos resultados da atividade do composto, mas a incorporação de um átomo de nitrogênio num anel ftalimida resultou numa perda significativa da atividade.

Os testes de atividade biológica são demorados e os custos financeiros para realizá-los são elevados, além da complexidade envolvida no processo de síntese destes compostos. Estas são algumas barreiras encontradas para a compreensão e desenvolvimento de medicamentos destinados ao tratamento de neoplasias. Assim, os estudos teóricos baseados em propriedades moleculares, tornam-se uma importante contribuição para o desenvolvimento de novos

fármacos.

Neste trabalho, realizamos uma investigação teórica sistemática de estrutura atividade biológica para esse conjunto de 14 compostos análogos da acridina analisados experimentalmente por Cuny com o objetivo de encontrar relações de estrutura e atividade (SAR, do inglês Structure Activity Relationships) inibitória das quinases Haspin e DYRK2. O estudo considera parâmetros estruturais, eletrônicos e físico-químicos que podem estar relacionados com os processos de interação desses derivados acridínicos com alvos moleculares no meio biológico.

Nesse estudo, utilizamos a metodologia denominada MIE (Metodologia de Índices Eletrônicos), que relaciona índices eletrônicos com a atividade biológica. A MIE é baseada nos conceitos de densidade local de estados eletrônicos (LDOS, do inglês Local Density of States) e total, e com os orbitais de fronteira HOMO (Highest Occupied Molecular Orbital), HOMO-1, LUMO (Lowest Unoccupied Molecular Orbital) e LUMO+1. A análise LDOS permite a identificação de níveis moleculares relevantes, correlacionados com a atividade biológica. A MIE utiliza-se somente de dois parâmetros, relacionados com a LDOS e com os orbitais de fronteira [1, 26, 27, 28].

O trabalho reporta o estudo utilizando a MIE para um grupo de 14 compostos derivados da acridina, que possuem atividade inibitória para as enzimas Haspin e DYRK2. O estudo da MIE foi feito separadamente para as duas enzimas. Utilizamos também duas outras metodologias estatísticas multivariadas, que permitem o reconhecimento de padrões, a Análise de Componentes Principais (PCA, Principal Component Analysis) e a Análise de Agrupamentos Hierárquica (HCA, Hierarchical Cluster Analysis) para analisar de forma sistemática os resultados obtidos com a MIE.

O trabalho está organizado da seguinte forma. No capítulo 2, descrevemos as metodologias utilizadas nos estudos realizados, que são relatados e discutidos no capítulo 3. As conclusões são apresentadas no capítulo 4. Os Apêndices contêm o detalhamento do método Hartree-Fock e dos métodos semiempíricos.

2 ***METODOLOGIA***

2.1 **Métodos *ab initio***

Os cálculos de propriedades eletrônicas de moléculas consistem na solução aproximada da equação de Schrödinger para sistemas de muitos elétrons e muitos centros atômicos. Os métodos *ab initio* são baseados em simplificações matemáticas de equações de primeiros princípios, como a aproximação de Hartree-Fock, apresentada em detalhe no Apêndice A. Outra possibilidade consiste em descrever a densidade eletrônica por meio da Teoria do Funcional da Densidade (DFT, Density Functional Theory)[27, 29, 30].

Por outro lado, os métodos semiempíricos utilizam as mesmas aproximações que os métodos Hartree-Fock, acrescidas de outras simplificações e parametrizações baseadas em resultados de experimentos. Estes métodos, cujo custo computacional é menor, permitindo o seu uso em sistemas moleculares maiores, são apresentados no apêndice B.

2.1.1 **Teoria do Funcional da Densidade**

A teoria do funcional da densidade (DFT, do inglês "Density Functional Theory") é uma forma de solucionar o problema para sistemas de muitos elétrons, utilizando a densidade eletrônica total $\rho(r)$. A equação de Schrödinger de N elétrons com uma função de onda de com $3N$ variáveis pode ser escrita como uma equação da densidade eletrônica com três variáveis.

Historicamente a utilização da densidade eletrônica $\rho(r)$ como variável fundamental para a descrição de um sistema eletrônico data de 1927 quando Llewellyn H. Thomas e Enrico Fermi

propuseram a Aproximação de Thomas-Fermi, uma alternativa para a resolução da equação de Schrödinger, usando a densidade eletrônica $\rho(r)$ ao invés da função de onda.

Segundo Coutinho *et al.* [31] a densidade eletrônica foi usada por Drude, desde o início do século XX, para desenvolver a teoria sob condução térmica e elétrica, quando aplicou a teoria dos gases a um metal, considerando como um gás homogêneo de elétrons. E a partir de então, outros modelos como os de Sommerfeld, Thomas, Fermi e Dirac foram propostos e aperfeiçoados. [31]

A solução exata, com o uso da densidade eletrônica, foi reconhecida em 1964 através da publicação de dois teoremas por Hohenberg e Kohn, e ficou conhecida como Teoria do Funcional da Densidade-DFT. Este trabalho resultou num prêmio Nobel de Química dado a Walter Kohn em 1998. A aplicação da Teoria do Funcional da Densidade em átomos, moléculas e sólidos aumentou consideravelmente, e vem se mostrando o método mais eficiente para o cálculo de propriedades estruturais e eletrônicas, para sistema no estado fundamental. A partir do formalismo de Hohenberg e Kohn elementos importantes para a descrição química de compostos são descritos bem melhor pela Teoria do Funcional da Densidade [31, 32, 33].

Trabalhos anteriores à Teoria do Funcional da Densidade foram importantes na formulação dos teoremas de Hohenberg e Kohn. Os primeiros trabalhos foram publicados por Thomas(1927) e Fermi(1928). Usando a densidade eletrônica ($\rho(\mathbf{r})$) como variável básica na descrição do sistema eletrônico. O modelo é hoje conhecido como modelo de Thomas-Fermi-Dirac (TFD). Em 1964 foi legitimada por Hohenberg e Kohn. Em 1965 Kohn e Sham apresentaram melhores resultados para a energia cinética, o método KS [32].

O primeiro teorema de Hohenber-Kohn (HK) estabelece que o potencial externo é um funcional único de $\rho(r)$ além de uma constante aditiva. A energia do sistema é calculada mediante resolução da equação de Schrodinger, $\hat{H}_{BO}\Psi = E\Psi$, sendo determinada pela densidade eletrônica $\rho(\mathbf{r})$.

$$E = E_v[\rho] \tag{2.1}$$

O segundo teorema estabelece que, havendo qualquer aproximação da densidade eletrônica, $\tilde{\rho}(r)$, de modo que $\tilde{\rho}(r) \geq 0$ e $\int \tilde{\rho}(r)dr = N$, a energia total será sempre maior ou igual a energia exata do sistema, ou seja, $E\tilde{\rho}(r) \geq E[\rho] = E_0$.

Kohn e Sham propuseram um modelo de partículas não interagentes, no qual o funcional energia é dada por:

$$E_{KS}[\rho] = T_0[\rho] + U_H[\rho] + V_{ext}[\rho] + V_{XC}[\rho] \quad (2.2)$$

onde $T_0[\rho]$ é a energia cinética média de um sistema não interagente de elétrons cuja densidade é igual a de um sistema interagente, $U_H[\rho]$ é a energia coulombiana média entre os elétrons, $V_{ext}[\rho]$ é o potencial efetivo externo e $V_{XC}[\rho]$ é o termo de troca e correlação, contendo correções para energia cinética e energia de Hartree [34, 35].

O método KS estabelece a repulsão elétron-elétron de Coulomb e uma nova função universal $G[\rho]$:

$$E_v[\rho] = G[\rho] + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 + \int \rho(r)v(r)dr, \quad (2.3)$$

em que

$$G[\rho] = T_s[\rho] + E_{xc}[\rho]$$

onde T_s é a energia cinética dos elétrons não interagentes e E_{xc} é o potencial de troca e correlação, que inclui todos os termos de interação entre as partículas.

Tendo uma expressão para a energia, equação (2.3), e utilizando o princípio variacional estabelecido por Hohenberg e Kohn, onde aplica-se o método de multiplicadores de Lagrange de maneira a minimizar a energia, é possível encontrar equações que descrevam o estado do sistema, chamadas equações de Kohn-Sham. Os orbitais KS são obtidos a partir da equação de

Schrödinger de um elétron como segue:

$$\left(-\frac{1}{2}\nabla^2 + v_{ef}\right)\psi_i^{KS} = \varepsilon_i\psi_i^{KS} \quad (2.4)$$

onde ε_i são os autovalores e ψ_i^{KS} são as autofunções e v_{ef} é o potencial local efetivo para um sistema de referência de elétrons não interagentes. A energia cinética é calculada de maneira autoconsistente. O potencial efetivo é um funcional da densidade $v_{ef}=v_{ef}[\rho]$ e é dado por:

$$v_{ef}(r) = v(r) + \int \frac{\rho(r_1)}{|r-r_1|} dr_1 + v_{xc}(r), \quad (2.5)$$

em que

$$v_{xc}(r) = \frac{\delta E_{XC}[\rho]}{\delta \rho(r)}. \quad (2.6)$$

A grande questão para a DFT é a dificuldade em encontrar boas aproximações para este termo, já que funcionais exatos para troca e correlação não são conhecidos, com exceção do gás de elétrons livres. Contudo existem aproximações que definem o potencial de troca e correlação. Uma vez calculado este termo, a teoria torna-se consistente. Com isso o cálculo de propriedades do sistema é feito de maneira precisa, aproximando-se dos dados obtidos experimentalmente, conferindo confiabilidade ao método.

2.1.2 Funcional de Troca e Correlação

O método de Aproximação do Gradiente Generalizado (*GGA- Generalized- Gradiente Approximation*) utiliza aproximação não local. O funcional de troca e correlação E_{XC}^{GGA} é escrito como a soma de duas partes, uma para a troca e outra para a correlação,

$$E_{XC}^{GGA} = E_X^{GGA} + E_C^{GGA}. \quad (2.7)$$

Um dos funcionais de troca, com densidade corrigida pelo gradiente, mais utilizados é o de Becke de 1988 conhecido como B88, escrito como segue

$$E_X^{B88} = E_X^{LSDA} - b \sum_{\sigma=\alpha,\beta} \int \frac{(\rho^\sigma)^{\frac{4}{3}} \chi_\sigma^2}{1 + 6b \chi_\sigma \sinh^{-1} \chi_\sigma} d\mathbf{r}, \quad (2.8)$$

onde $\chi_\sigma \equiv |\nabla \rho^\sigma| / (\rho^\sigma)^{\frac{4}{3}}$, $\sinh^{-1} = \ln[x + (x^2 + 1)^{\frac{1}{2}}]$ e b é um parâmetro semiempírico que vale $0,0042u.a.$ obtido após cálculo de energia de troca pelo método HF realizado e ajustado para vários átomos, onde E_X^{LSDA} é escrito como,

$$E_X^{LSDA} = -\frac{3}{4} \left(\frac{6}{\pi} \right)^{\frac{1}{3}} \int \left[(\rho^\alpha)^{\frac{4}{3}} + (\rho^\beta)^{\frac{4}{3}} \right] d\mathbf{r}. \quad (2.9)$$

Dentre os funcionais de correlação corrigidos pelo gradiente mais utilizados estão o de Lee-Yang-Parr (LYP) e os funcionais de Becke. Contudo os funcionais híbridos são muito empregados. Esses funcionais são uma combinação linear dos funcionais DFT puros e o termo de troca calculado pelo método Hartree-Fock (E_X^{HF}). Para sistemas de camada fechada o termo de troca é escrito como:

$$E_X = -\frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \left\langle \phi_i^{KS}(1) \phi_j^{KS}(2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \phi_j^{KS}(1) \phi_i^{KS}(2) \right\rangle. \quad (2.10)$$

O funcional híbrido empregado nesta pesquisa foi o B3LYP, que possui 3 parâmetros de Becke e mais os funcionais de correlação LYP. O B3LYP é definido como:

$$E_{XC}^{B3LYP} = (1 - a_0 - a_x) E_X^{LSDA} + a_0 E_X^{HF} + a_x E_X^{B88} + (1 - a_C) E_C^{VWN} + a_C E_C^{LYP}, \quad (2.11)$$

onde o termo de correlação E_C^{VWN} provém do processo de aproximação da densidade local-LSDA (Local-Spin-Density Aproximação) desenvolvido por Vosko-Wilk-Nusair. Os demais parâmetros $a_0 = 0.20$, $a_x = 0.72$ e $a_C = 0.81$ foram escolhidos de maneira a imprimir os resultados para a energia de otimização e energias atômicas totais. No presente trabalho optou-se por utilizar o método fundamentado na Teoria do Funcional da Densidade, já que para otimização

de geometria apresenta resultados muito próximos à outros métodos *ab initio*, além de possuir uma similaridade nos valores dos parâmetros geométricos, o DFT, dentre os métodos *ab initio* é computacionalmente um dos mais vantajosos. Funcionais híbridos, tal como B3LYP, escolhido como funcional de troca e correlação para calcular a energia eletrônica mínima, melhoram a precisão dos parâmetros fornecidos pelo método [36, 37]. Uma outra maneira para resolver a equação de Schrödinger eletrônica é conhecida como métodos semi-empíricos, descrita na próxima seção.

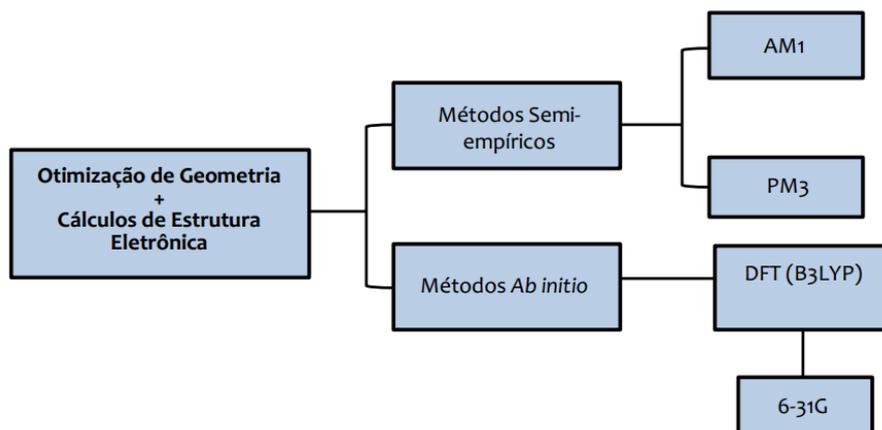
2.2 Métodos Semi-empíricos

Os métodos semi-empíricos procuram encontrar soluções aproximadas para a equação de Schrödinger eletrônica pela inserção de restrições matemáticas e pelo emprego de parâmetros empíricos[38]. Em geral, esses métodos são utilizados para tratar sistemas grandes, como biomoléculas, por exigir um pequeno esforço computacional.

As restrições desses métodos são incorporadas por aproximações para resolução de integrais, utilizando-se de expressões para resolver várias parcelas de integrais que fazem parte da equação de Schrodinger, e no que diz respeito a um procedimento de parametrização, utilizando valores obtidos experimentalmente ou através de outros cálculos teóricos. Estas aproximações são inseridas com o objetivo de diminuir o custo computacional. Isto faz com que os métodos semi-empíricos sejam um dos mais relevantes para o tratamento de sistemas de biomoléculas e sistemas condensados, uma vez que estes métodos conseguem descrever sistemas moleculares relativamente complexos com grande precisão [39].

Neste trabalho utilizamos os métodos semi-empíricos AM1 (Austin Method One) e PM3 (Parametric Method 3). O formalismo dos métodos semi-empíricos encontra-se melhor detalhado no Apêndice B.

O quadro resumo dos métodos utilizados para cálculos de otimização de geometria bem como os cálculos de estrutura eletrônica segue abaixo:



Os cálculos de estrutura eletrônica nos fornece como resultado as propriedades eletrônicas e geométricas dos sistemas estudados. Uma maneira de estabelecer correlações entre estas propriedades é a utilização de métodos que realizam uma análise qualitativa de relações de estrutura-atividade. A Metodologia de Índices Eletrônicos (MIE) é uma das metodologias adotadas nesse estudo para classificar os compostos quanto a sua atividade biológica, e a mesma será descrita na seção que se segue.

2.3 Metodologia de índices Eletrônicos (MIE)

A metodologia de índices eletrônicos foi desenvolvida, em 1996, por Barone *et al.* com o objetivo inicial de investigar a atividade carcinogênica dos hidrocarbonetos aromáticos policíclicos (HAPs), classificando-os em ativos ou inativos [28, 27, 1].

A MIE classifica, por meio de regras simples, compostos em ativos e inativos de maneira eficaz, utilizando apenas dois descritores quânticos moleculares (parâmetros puramente eletrônicos). Essa metodologia está baseada no conceito de densidade local de estados (LDOS- *Local Density of States*) e nos valores críticos da diferença de energia eletrônica entre os níveis moleculares dos orbitais de fronteira [29, 40].

A densidade de estados (DOS) é definida como o número de estados eletrônicos por unidade de energia. O conceito relacionado LDOS é a DOS calculada sobre uma região molecular

específica. A LDOS descreve a distribuição espacial de estados específicos sobre determinadas regiões moleculares.

Considerando a combinação linear dos orbitais atômicos (LCAO), em que a contribuição de cada átomo para um nível eletrônico específico é dada pelo quadrado do coeficiente do orbital atômico, que é a densidade de probabilidade correspondente ao nível em um sítio escolhido, a LDOS é encontrada somando-se as contribuições dos orbitais atômicos selecionados (n_i a n_f):

$$LDOS(E_i) = 2 \sum_{m=n_i}^{n_f} |C_{mi}|^2, \quad (2.12)$$

onde

– E_i é a energia do orbital atômico;

– n_i e n_f são os sítios atômicos, inicial e final, respectivamente;

– C_{mi} é o coeficiente de expansão do orbital molecular i , em termo dos orbitais atômicos m ;

–O fator 2 presente na equação acima, foi introduzido na equação por respeitar o princípio da exclusão de Pauli(máximo de dois elétrons por nível eletrônico);

A MIE utiliza-se de dois parâmetros puramente eletrônicos, o primeiro η , é relacionado com as diferenças das contribuições LDOS de níveis eletrônicos específicos, geralmente os níveis de fronteira HOMO (*Highest Occupied Molecular Orbital*) e HOMO-1, ou LUMO (*Lowest Unoccupied Molecular Orbital*) e LUMO+1.

$$\eta = 2 \sum_{m=n_i}^{n_f} (|C_{mnivel1}|^2 - |C_{mnivel2}|^2) \quad (2.13)$$

O descritor η contém informações locais que refletem a distribuição de cargas sobre a molécula, e está associado com as diferenças entre as densidades de estados local de dois orbitais de fronteira, podendo ser entre o último e o penúltimo dos orbitais ocupados, HOMO e

HOMO-1, para os quais:

$$\eta_H = LDOS(E_{HOMO}) - LDOS(E_{HOMO-1}) \quad (2.14)$$

ou entre o primeiro e o segundo orbital não ocupado, LUMO e LUMO+1:

$$\eta_L = LDOS(E_{LUMO+1}) - LDOS(E_{LUMO}) \quad (2.15)$$

O segundo parâmetro da MIE é o Δ , definido como a diferença de energia entre dois níveis moleculares;

$$\Delta = E_{nível2} - E_{nível1} \quad (2.16)$$

O parâmetro Δ nos fornece informações globais sobre a molécula, já que o mesmo envolve os autovalores de energia de um determinado nível molecular.

A partir dos parâmetros η e Δ apresentados acima, a MIE constrói regras para classificar compostos quanto a sua atividade biológica. Um exemplo é a regra típica, como:

Se $\eta > \eta_c$ e $\Delta < \Delta_c$, a molécula será ativa, caso contrário, inativa.

Os parâmetros η_c e Δ_c são os valores críticos extraídos da classificação de atividade dos compostos, dados pelos índices experimentais, depois de uma rigorosa análise que considera estes índices. Portanto, η_c e Δ_c são valores limites por meio dos quais classificamos os compostos como ativos e inativos.

As regras de classificação dos compostos quanto à sua bioatividade podem ser encontradas durante a busca dos valores críticos dos parâmetros determinados pelos orbitais atômicos (HOMO e HOMO-1) e (LUMO e LUMO+1) que melhor representam a atividade biológica, ou por meio de tabelas Booleanas de predição de atividade, como se vê na tabela 2.1 a seguir.

Barone *et al.* desenvolveram uma metodologia que consegue identificar, a nível qua-

Tabela 2.1 – Tabela Booleana para classificação da atividade dos compostos, $+\Delta$ e η para compostos analisados $> \eta_c$ e Δ_c e $-\Delta$ e η para compostos analisados $< \eta_c$ e Δ_c .

(η)	(Δ)	Atividade Biológica (Predição)
+	+	Ativa
+	-	Inativa
-	+	Inativa
-	-	Inativa

litativo, compostos ativos e inativos de classes de moléculas com diferentes ações biológicas. Ainda que não permita uma interpretação direta do porquê dos parâmetros utilizados, a MIE tem se mostrado cada vez mais eficaz na classificação e na predição da atividade de compostos, por isso vem sendo muito utilizada e identificada como uma metodologia importante nos estudos de estrutura-atividade para o desenvolvimento e aprimoramento de novos fármacos [27, 10, 28, 1], como mostrado na figura 2.1

Figura 2.1 – Resultados obtidos para MIE e por outros métodos, e algumas classes estudadas por Barone et al. [1].

Classe de Composto (Método de cálculo)	Ação Biológica	# moléculas	Δ_c (eV)	η_c	Acerto (%)	
					PCA	MIE
HAP(Hückel)	Carcinogênico	78	0,408	0	80,8	78,2
HAP(PM3)	Carcinogênico	78	0,309	0	78,2	79,5
Aza-HAPs(PM3)	Carcinogênico	32	0,545 (LUMO)	0,282 (CLK)	94	91
Ciclopenta-HAPs(PM3)	Carcinogênico	36	0,3	-	-	80
Lapachol(PM3)	Anti-tumoral Antibiótico Anti-malarial	25	-	0	-	96
Ácido naftóico(PM3)	Inibidor da Integrase do HIV	6	-	0,0(LUMO)	-	100
Talidomida(PM3)	Anti-inflamatório	32	0,15	0,1	-	85
Talidomida(AM1)	Anti-inflamatório	32	0,15	0,1	-	90
Esteróides(PM3)	Afinidade de ligação ao CBG	31	0,41	0,156	100	100
Aps(PM3)	Atividade Progesterona	21	0,8	0	85,7	90,5
Taxóides(PM3)	Anti-tumoral	20	0,15	0	100	85
Benzo[c]quinolizina-3-onas(AM1)	Inibidor de 5alfa-reductase	23	-	1	-	96
Benzo[c]quinolizina-3-onas(PM3)	Inibidor de 5alfa-reductase	23	-	1,3	-	100

2.4 Métodos Estatísticos Multivariados

Após utilizarmos os métodos químico-quânticos para os descritores eletrônicos, obtivemos uma gama de parâmetros moleculares, e torna-se necessária a utilização de métodos que permita a análise simultânea de todos os parâmetros encontrados. Assim, métodos multivariados são ferramentas úteis, já que a interpretação de dados é difícil, uma vez que temos uma elevada quantidade de variáveis a ser analisadas[41].

Para a análise dos resultados escolhemos o método de análise multivariada Análise de Componentes Principais (PCA), umas das técnicas mais utilizadas que procura estabelecer uma possível correlação entre descritores teóricos e a atividade biológica [42].

Outros métodos multivariados como a Modelagem Independente de Analogia de Classes (SIMCA) que é um método de classificação, e Regressão em Componentes Principais(PCR) que é um método de calibração, ou regressão por mínimos quadrados(PLS), derivam do PCA. Estes métodos estatísticos de análise de dados, quando aplicados em problemas químicos, são chamados de métodos quimiométricos [43, 44, 45].

2.4.1 Análise de Componentes principais (PCA)

A análise de componentes principais (PCA- *Principal Components Analysis*) é um método estatístico multivariado que por meio da redução de dados, busca o reconhecimento de padrões.

O método abrevia-se em fatorar a matriz de dados iniciais X, escoltando-a de maneira a transformar os descritores para serem comparados numa escala igual, mesmo que suas unidades sejam diferentes. Então, temos uma matriz com as variáveis de dados e uma segunda matriz com as amostras.

Os elementos da matriz de dados iniciais possuem unidades de medida diferentes, e a comparação entre elementos de natureza diferente não é praticável. Para tornar essa comparação possível é preciso fazer com que estes elementos (variáveis) sejam adimensionais, com média nula e variância unitária. E um termo que represente todas as variáveis modificadas é escrito da

seguinte maneira:

$$y_{nm} = \frac{x_{nm} - \bar{x}_m}{\sigma_m} \quad (2.17)$$

onde os índices mn , representam, respectivamente, m -ésimo descritor e n -ésima molécula. E

$$\bar{x}_m = \frac{\sum_{n=1}^N x_{nm}}{N} \quad (2.18)$$

é a média aritmética dos valores calculados para o m -ésimo descritor e

$$\sigma_m = \sqrt{\frac{\sum_{n=1}^N (x_{nm} - \bar{x}_m)^2}{N - 1}} \quad (2.19)$$

é o desvio padrão do m -ésimo descritor.

Tem-se agora uma nova matriz de variáveis simétrica de ordem $m \times m$, a matriz de covariância C , dada por:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mm} \end{pmatrix}$$

Com o propósito de diminuir a correlação entre as variáveis, os elementos fora da diagonal principal na matriz C , precisam ser minimizados, enquanto os termos da diagonal principal maximizados [46, 47]. Para que isso ocorra utiliza-se a diagonalização da matriz de covariância C , de maneira a obter uma matriz P que equivale aos autovetores de C , cujo autovalores associados é a variância de cada componente principal. Assim, determinar a componente principal, aquela que tem maior relevância, resume-se em encontrar o autovetor ao qual esteja associado

o maior autovalor, então

$$CP = VP \quad (2.20)$$

onde a matriz de autovalores é chamada de V.

De forma geral, PC1 é formada de maneira que tenha a maior parte da variância. PC2 é retirada da variância restante, e da mesma maneira gera-se a PC3, até que a PCn seja gerada. PC2 é então uma combinação linear de variáveis observadas não correlacionadas com a primeira coluna (PC1) que tenha máximo de resto da variação total. A matriz de variância pode então ser vista da seguinte maneira:

$$Var = \begin{pmatrix} \text{Variância da PC1} & 0 & \cdots & 0 \\ 0 & \text{Variância da PC2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{Variância da PCn} \end{pmatrix}$$

Numa visão geométrica, os objetos são representados por uma nuvem de N pontos (são as amostras) em um espaço multidimensional, com um eixo para cada um dos descritores (variáveis que dizem respeito às amostras), o centro dos pontos é definido pela média de cada amostra, e a variância é a média dos quadrados da diferença dos N pontos com relação à média de cada amostra.

$$V_i = \frac{1}{n-1} \sum_{m=1}^n (X_{im} - \bar{X}_i)^2. \quad (2.21)$$

Os autovetores podem ser representados matricialmente da seguinte forma:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

As componentes principais, obtidas em ordem decrescente de relevância, são escritas como uma combinação linear de variáveis modificadas

$$PC_j = a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n. \quad (2.22)$$

O grau com que cada variável é linearmente correlacionada é representado pela sua covariância:

$$C_{ij} = \frac{1}{n-1} \sum_{m=1}^n (X_{im} - \bar{X}_i)(X_{jm} - \bar{X}_j) \quad (2.23)$$

onde,

– C_{ij} = covariância de variáveis i e j ;

– \sum = soma sobre todos os n objetos;

– X_{im} = valor da variável i no objeto m ;

– \bar{X}_i = média de variável i .

O objetivo do PCA é rotacionar rigidamente os eixos desse espaço p -dimensional para novas posições, que são os eixos principais, de forma que o eixo principal 1 tenha maior variância que o eixo 2, e assim por diante, até o último eixo que tenha menor variância. Os eixos principais não são correlacionados, o PCA realiza rotações dos eixos cartesianos de forma que o antigo eixo horizontal contenha a maior informação sobre as distribuições das amostras, e o PC2 a segundo maior. Numa visualização gráfica desses pontos no espaço multidimensional

pode ser visto na figura 2.2 que mostra a variância das PCs.

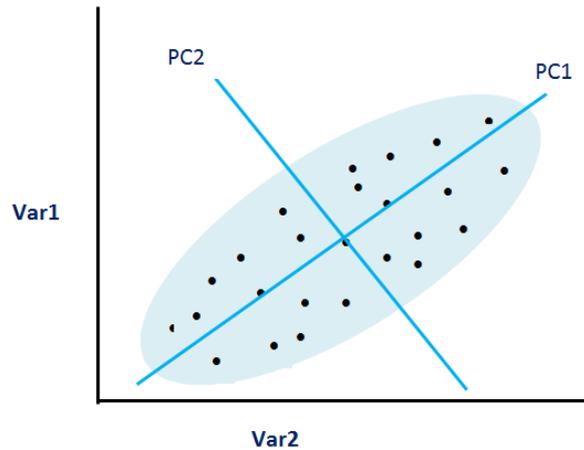


Figura 2.2 – Visualização gráfica da variância das PCs

Como o problema é bidimensional, os padrões podem ser visualizados plotando as duas primeiras componentes principais, que contêm a maior parte da variância, caso essas não sejam suficientes para exibir as correlações, a terceira componente pode ser plotada auxiliando na identificação de similaridades.

2.4.2 Análise de Agrupamento Hierárquico (HCA)

A HCA (Análise de Agrupamento Hierárquico, do inglês, *Hierarchical Cluster Analysis*) agrupa as amostras em categorias, fundamentalmente, através de semelhanças entre as amostras (compostos) que fazem parte da mesma categoria e diferenças entre elementos que fazem parte de categorias distintas. Essa análise permite a visualização gráfica (dendograma) de todo o conjunto de dados, apresentando a maior quantidade de informações a respeito desse conjunto, tornando possível a interpretação dos resultados de reconhecimento de padrões, o que não é tão simples, uma vez que a quantidade de variáveis analisadas (parâmetros) é elevada. As informações contidas nos dendogramas, expressam a maneira como as amostras estão inter-relacionadas, com a interpretação de que quanto menor a distância entre os pontos, maior a similaridade entre as amostras, o que norteia a interpretação multivariada do comportamento das amostras [48, 10, 49].

Uma forma de calcular a distância entre dois pontos a e b , que para o seguinte estudo são representados pelas amostras no espaço n -dimensional, é a distância euclidiana X_{ab} dada por:

$$X_{ab}^2 = \sum_{j=1}^n (d_{aj} - d_{bj})^2 \quad (2.24)$$

Os pontos que estão mais próximos são agrupados com base na distância euclidiana, e substituído por um novo ponto localizado na metade da distância entre eles, e iterado várias vezes até agrupar todos os pontos em um só ponto, essa ordenação hierárquica e crescente é representada pelo dendograma.

Os dendogramas são muito úteis na visualização das similaridades das amostras, já que o espaço de representação pode ter dimensão maior do que três, que gráficos tradicionais permitem [50].

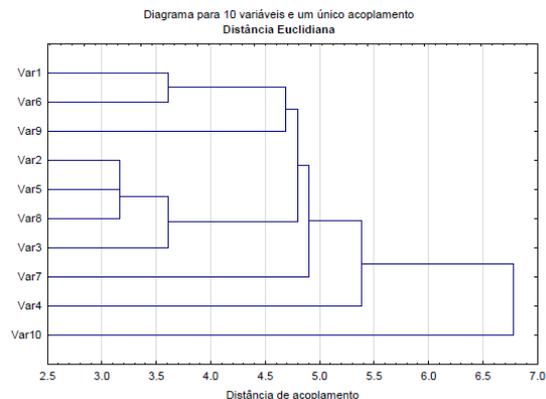


Figura 2.3 – Dendrograma ilustrativo construído com variáveis de 1 a 10 e relacionados em ordem aleatória

3 *RESULTADOS E DISCUSSÕES*

Os 14 compostos análogos da Acridina foram submetidos a uma busca conformacional detalhada e sistemática, averiguando os comprimentos de ligação entre átomos, os ângulos de ligações e seus ângulos diedrais. As otimizações foram realizadas utilizando os métodos semi-empíricos AM1 e PM3 e um método *ab initio* DFT, juntamente com os softwares GAMESS e Spartan, em todos eles realizamos a otimização de geometria para os três métodos, uma vez que buscamos uma boa geometria final para os compostos estudados.

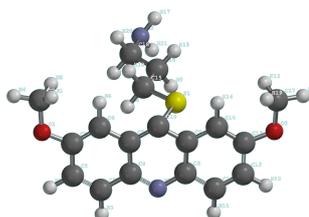


Figura 3.1 – Estrutura da molécula 1, derivado da acridina.

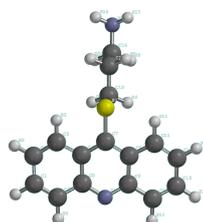


Figura 3.2 – Estrutura da molécula 2, derivado da acridina.

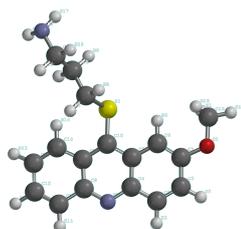


Figura 3.3 – Estrutura da molécula 3, derivado da acridina.

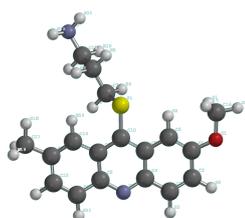


Figura 3.4 – Estrutura da molécula 4, derivado da acridina.



Figura 3.5 – Estrutura da molécula 5, derivado da acridina.

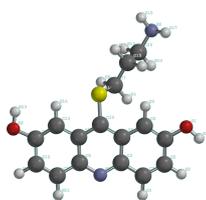


Figura 3.6 – Estrutura da molécula 6, derivado da acridina.

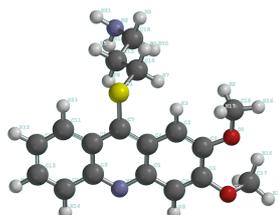


Figura 3.7 – Estrutura da molécula 7, derivado da acridina.

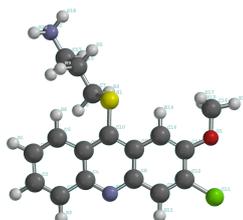


Figura 3.8 – Estrutura da molécula 8, derivado da acridina.

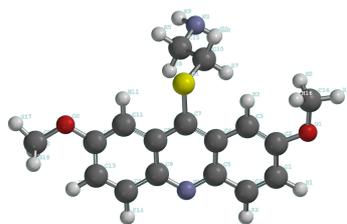


Figura 3.9 – Estrutura da molécula 9, derivado da acridina.

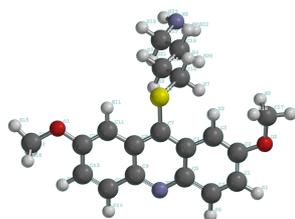


Figura 3.10 – Estrutura da molécula 10, derivado da acridina.

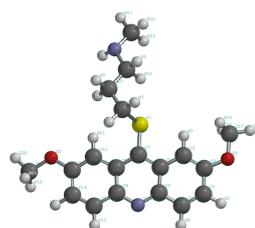


Figura 3.11 – Estrutura da molécula 11, derivado da acridina.

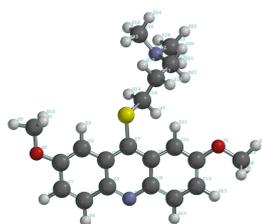


Figura 3.12 – Estrutura da molécula 12, derivado da acridina.

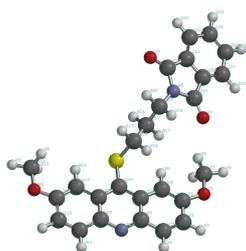


Figura 3.13 – Estrutura da molécula 13, derivado da acridina.

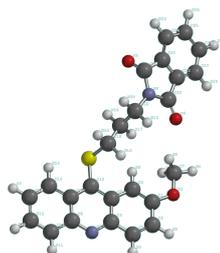


Figura 3.14 – Estrutura da molécula 14, derivado da acridina.

3.1 Processo de Otimização

A busca conformacional é uma importante ferramenta para o planejamento de fármacos fornecendo uma visualização tridimensional das moléculas e suas variações espaciais dessas moléculas com relação aos ângulos diedrais. Foi realizada uma busca conformacional sistemática, utilizando o Spartan como interface gráfica para a construção das moléculas, variando os diedrais formados por dois carbonos no anel central do esqueleto molecular com o elemento enxofre. Escolhemos outro conjunto de átomos na parte superior que compreende aos carbonos ligados a este enxofre. Para o primeiro diedro a variação foi de 30 graus, e para cada variação do primeiro o segundo variava também de 30 graus, gerando um grande número de conformémeros, exigindo um grande esforço computacional e conseqüentemente demandando muito tempo de dedicação.

Iniciamos a nossa análise com o busca da geometria pela conformação que a apresentasse o menor valor para o calor de formação de cada um dos compostos. Realizamos a otimização de geometria estabelecendo algumas considerações, como a de que as moléculas encontram-se no vácuo e estejam no estado fundamental. As novas posições assumidas pelos átomos após a otimização, são então a fonte primária de enformações teóricas.

Para cada molécula geramos 144 conformémeros, os diedros foram girados de 0 a 360 graus, em passos de 30 graus. Em seguida foi realizado o refinamento dos cálculos e desta forma chegamos ao conformémero de menor energia.

Para cada uma das conformações, inicialmente, foi realizado um cálculo de otimização de geometria no Spartan, com os métodos AM1 e PM3, este cálculos nos forneceu valores de calor de formação. A partir desses valores, selecionamos os conformémeros que nos indicaram um mínimo de energia. Os conformémeros com as novas posições dos átomos nas moléculas fora utilizados como arquivo de entrada para que o calculo de otimização de energia fosse novamente realizado no GAMESS, e pudéssemos obter parâmetros que estivessem relacionados com a atividade biológica dessas moléculas.

Os valores para o calor de formação obtidos para os métodos semi-empíricos estão na tabela 3.1 abaixo. Esta tabela contém os valores para o calor de formação dos análogos acridínicos, após realizada uma busca de um mínimo de energia para cada composto.

Tabela 3.1 – Tabela com os valores do Calor de Formação dos 14 compostos para os métodos AM1 e PM3.

Molécula	C.F.(Kcal/mol)	
	AM1	PM3
1	-6.0250	-4.9726
2	71.2003	72.7355
3	32.5205	33.0253
4	24.6512	24.5533
5	27.5955	28.2851
6	-17.3810	-19.1674
7	-2.0944	-2.2765
8	29.1638	28.9787
9	3.7114	1.8550
10	-10.1388	-9.7376
11	-1.5461	-6.2044
12	3.8192	-9.9828
13	-18.7800	-44.5190
14	19.8622	-5.5714

Os compostos estudados tem o esqueleto molecular principal, formados por três anéis aromáticos, o que lhes confere uma geometria planar, no entanto, os radicais ligados nas duas laterais, que podem ser radicais como o grupo metoxi; assim como os radicais ligados ao enxofre na parte superior dos composto, cujos elementos relacionados são em sua maior parte carbonos (2 ou 3) ligados a um grupo metil; devido a repulsões entre átomos de hidrogênio contidos na ponta dos anéis do esqueleto molecular, ou mesmo a repulsão destes hidrogênio com os hidrogênios dos radicais laterais(o que depende do angulo de torção); estes compostos tendem a sair do plano.

As ligações de hidrogênio são as mais importantes, as interações realizadas por estes elementos são do tipo não covalente, são elas que mantêm as conformações bioativas das moléculas com as proteínas, que geralmente ocorre entre heteroátomos, nesse caso é preciso salientar a importância do heteroátomo de nitrogênio no anel central dos análogos de acridina aqui investigados.

Os 14 compostos análogos da acridina são moléculas compostas por muitos átomos e apresenta grupos flexíveis. Como temos uma grande quantidade de conformeros gerados, em consequência da flexibilidade de ângulos diedrais, torna-se indispensável uma análise detalhada dos vários conformeros tratáveis na descoberta daquele que após geometria otimizada, nos forneça o menor valor para o calor de formação, o mais estável.

Para o cálculo de minimização da energia foram realizados cálculos mecânicos-quânticos com o AM1 e o PM3 (melhores descritos no apêndice B) e um nível de método *ab initio* baseado na teoria do funcional densidade-DFT (descrito no apêndice A), com um funcional híbrido B3LYP e base 6-31G e utilizamos os resultados obtidos por esse método como mais um parâmetro de comparação para análise feita como os métodos AM1 e PM3. Como o DFT é um cálculo que exige grande esforço computacional, a busca sistemática ficou restrita somente aos métodos semi-empíricos. Utilizamos o DFT, depois de encontramos um mínimo global (M.G.) para obtermos parâmetros teóricos, como propriedades físico-químicas, estruturais, químicas. Esses parâmetros que dizem respeito as estruturas químicas, também chamados de descritores moleculares, são muito importantes para estabelecermos uma analogia entre estrutura química com a atividade biológica, o estudo SAR [51].

Com base nos resultados para as energia dos orbitais de fronteira, obtido para o DFT e PM3, após a otimização de geometria, fizemos uma comparação entre estes dois métodos e percebemos que, os valores das energias para os orbitais de fronteira e suas respectivas diferenças tem valores muito próximos, ou iguais, para ambos os métodos, como pode ser visto nas tabelas 3.2 e 3.3.

Para o método DFT os valores para essas energias foram:

Tabela 3.2 – Tabela com os resultados para as energias dos orbitais de fronteira e suas respectivas diferenças para o método PM3.

Molécula	<i>HOMO</i> – 1	<i>HOMO</i>	ΔH	<i>LUMO</i>	<i>LUMO</i> + 1	ΔL
1	-8.9107	-8.4883	0.4225	-1.2881	-0.3147	0.97339
2	-8.9974	-8.6034	0.3941	-1.3366	-0.3580	0.97857
3	-9.0102	-8.5050	0.5052	-1.3490	-0.3747	0.97428
4	-8.9288	-8.4908	0.4380	-1.3025	-0.3376	0.96488
5	-9.1520	-8.4280	0.7240	-1.4124	-0.4817	0.93067
6	-8.9742	-8.4882	0.4861	-1.3950	-0.4366	0.95841
7	-8.9563	-8.5649	0.3914	-1.3142	-0.3309	0.98324
8	-9.1202	-8.5592	0.5609	-1.4546	-0.5016	0.95300
9	-8.8469	-8.4170	0.4299	-1.3001	-0.3314	0.96873
10	-8.8540	-8.2855	0.5685	-1.2374	-0.2698	0.96764
11	-8.7617	-8.4209	0.3409	-1.2699	-0.2964	0.97346
12	-8.7441	-8.2581	0.4859	-1.1853	-0.2212	0.96411
13	-8.8505	-8.3010	0.5495	-1.1899	-1.1812	0.00865
14	-8.9313	-8.3435	0.5877	-1.2129	-1.1760	0.03691

Outra comparação pode ser feita baseada nas estruturas otimizadas pelos dois métodos. Escolhemos duas estruturas de moléculas que fossem ativas para as duas enzimas, e duas outras que fossem inativas, e comparamos as estruturas depois de otimizadas para os métodos em questão.

A molécula 1 é ativa para as enzimas Haspin e DYRK2, a estrutura dessa molécula muda muito pouco para as duas otimizações, somente o grupo de radicais ligados ao enxofre do anel central sofre uma mudança de posição de átomos. Numa visão mais geral, vemos que a mudança básica entre esses dois métodos para a molécula 1 é apenas uma mudança na posição dos átomos.

Tabela 3.3 – Tabela com os resultados para as energias dos orbitais de fronteira e suas respectivas diferenças para o método DFT.

Molécula	HOMO – 1	HOMO	ΔH	LUMO	LUMO + 1	ΔL
1	-8.9107	-8.4883	0.4225	-1.2881	-0.3147	0.9734
2	-8.9974	-8.6034	0.3941	-1.3366	-0.3580	0.9785
3	-9.0101	-8.5050	0.5052	-1.3490	-0.3747	0.9743
4	-8.9288	-8.4908	0.4379	-1.3025	-0.3376	0.9649
5	-9.1520	-8.4280	0.7240	-1.4124	-0.4817	0.9306
6	-8.9742	-8.4881	0.4861	-1.3950	-0.4366	0.9584
7	-8.9563	-8.5649	0.3915	-1.3141	-0.3309	0.9832
8	-9.1201	-8.5592	0.5609	-1.4546	-0.5016	0.9530
9	-8.8469	-8.4170	0.4298	-1.3001	-0.3314	0.9688
10	-8.8540	-8.2855	0.5685	-1.2374	-0.2697	0.9676
11	-8.7617	-8.4209	0.3409	-1.2699	-0.2964	0.9734
12	-8.7441	-8.2581	0.4859	-1.1853	-0.2212	0.9641
13	-8.8505	-8.3010	0.5495	-1.1899	-1.1812	0.0086
14	-8.9313	-8.3435	0.5877	-1.2129	-1.1760	0.0369

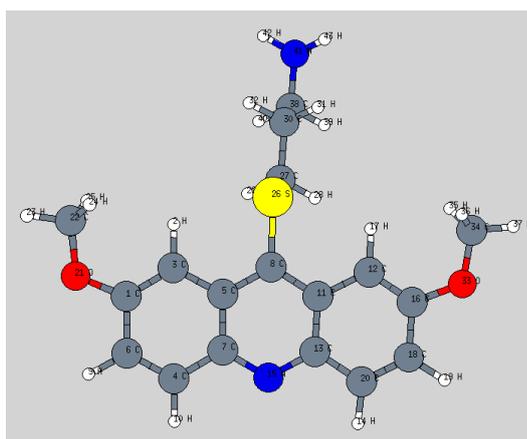


Figura 3.15 – Estrutura da molécula 1 depois de otimizada pelo método PM3.

A molécula 2 é inativa para as duas enzimas, não apresenta radicais ligados ao anel lateral, e portanto, sofreu mudanças somente nas posições dos átomos, carbonos e hidrogênios, ligados ao enxofre do anel central.

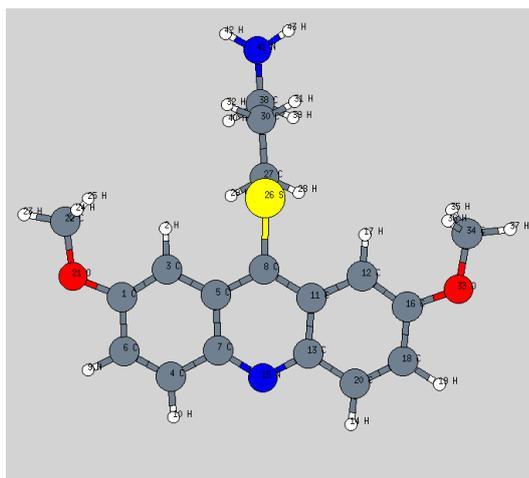


Figura 3.16 – Estrutura da molécula 1 depois de otimizada pelo método DFT.

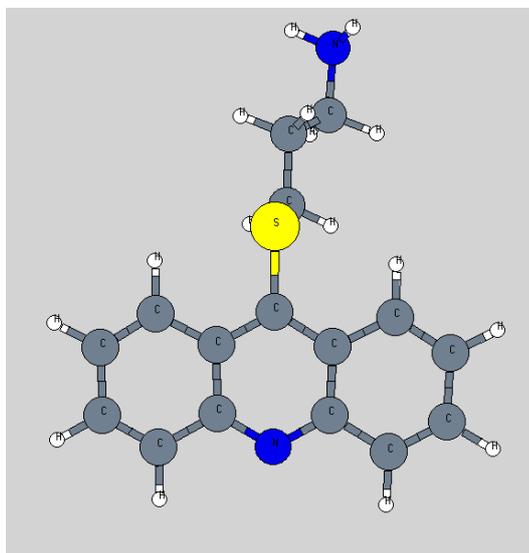


Figura 3.17 – Estrutura da molécula 2 depois de otimizada pelo método PM3.

A molécula 7 é inativa para os dois grupos, e apresenta radicais ligados em um dos anéis laterais. Percebemos que além da mudança nas posições dos átomos ligados ao enxofre do anel central, o radical ligado ao anel lateral na parte inferior, sofreu uma rotação. No entanto, não houve mudanças nos valores das energias dos orbitais de fronteira.

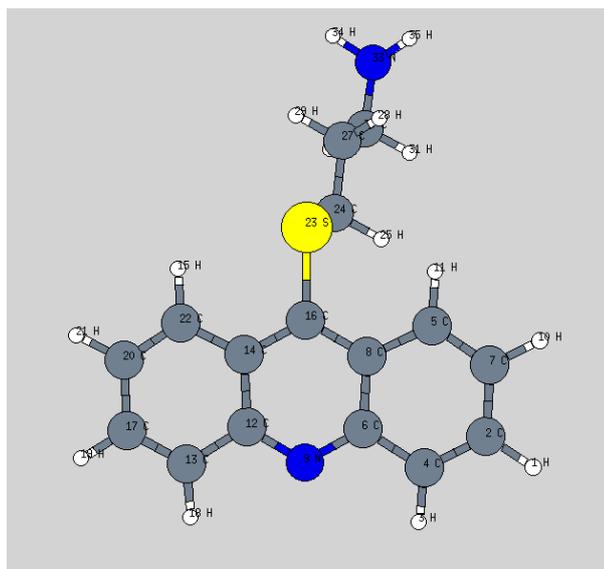


Figura 3.18 – Estrutura da molécula 2 depois de otimizada pelo método DFT.

A molécula 6 é ativa para as enzimas, assim como a molécula 1 sofreu mudanças apenas nas posições dos átomos ligados ao anel central. Apesar de possuir radicais ligados aos anéis laterais, estes não sofreram grandes mudanças. A energias dos orbitais de de fronteira também permaneceram identicos para os cálculos realizados com os dois métodos.

O próximo passo do nosso estudo, é realizar o estudo da atividade biológica dos compostos pela Metodologia de índices Eletrônicos, que utiliza-se dos valores das energias dos orbitais de fronteira. Para esse estudo utilizamos o programa Einsight, para o qual os resultados obtidos para a otimização de geometria, será utilizado como entrada para realizamos os cálculos LDOS. Devido ao fato de que os valores das energias dos orbitais de fronteira terem sido iguais para o PM3 e o DFT, e do programa Einsigth, para gerar valores de energia para as regiões, realizar um ciclo com cálculos a nível PM3 e AM1, restringimos a nossa análise MIE para os métodos AM1 e PM3.

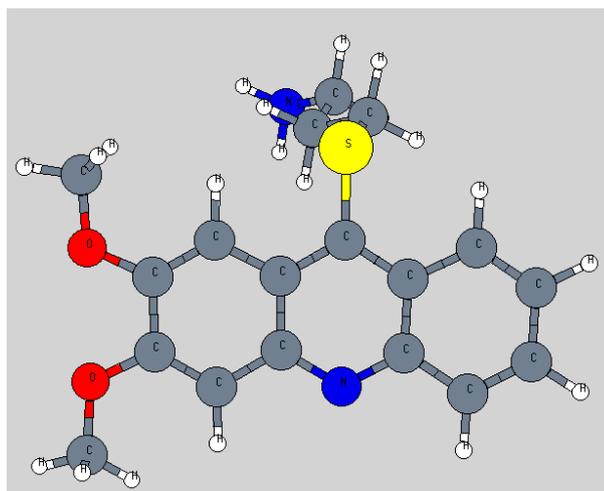


Figura 3.19 – Estrutura da molécula 7 depois de otimizada pelo método PM3.

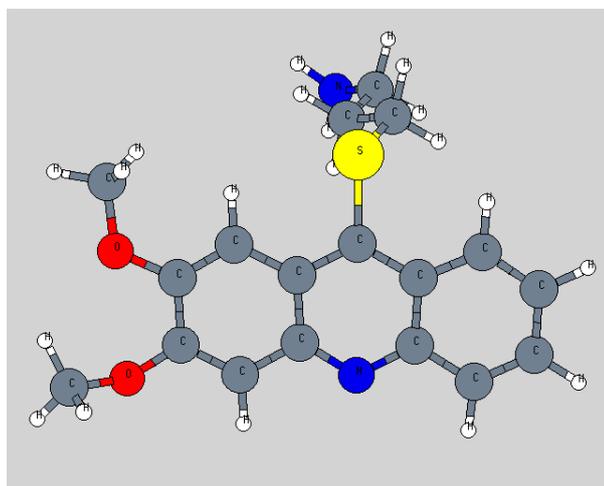


Figura 3.20 – Estrutura da molécula 7 depois de otimizada pelo método DFT.

3.2 Estudo das Relações Estruturais dos análogos da Acridina

Iniciamos a nossa investigação da atividade biológica dos compostos derivados da acridina e sua correlação com os descritores teóricos pela metodologia de índices eletrônicos, que como descrito no capítulo 3, utiliza-se somente de dois parâmetros ($\Delta e\eta$) para descrever a atividade biológica. O parâmetro Δ é a diferença entre as energias de dois estados de fronteiras vizinhos, este parâmetro nos fornece uma informação global sobre as propriedades eletrônicas do composto em questão. O outro parâmetro, η , é a diferença entre contribuições de átomos

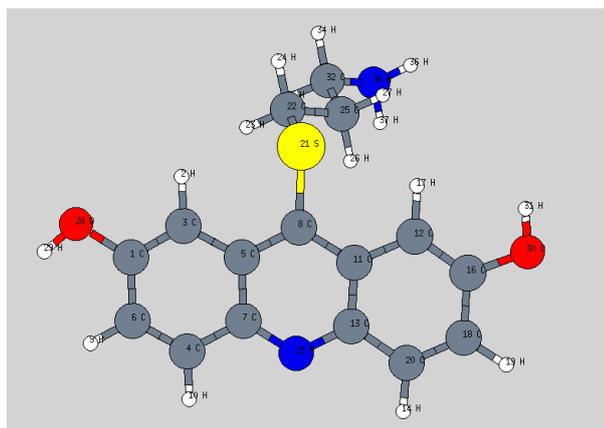


Figura 3.21 – Estrutura da molécula 6 depois de otimizada pelo método PM3.

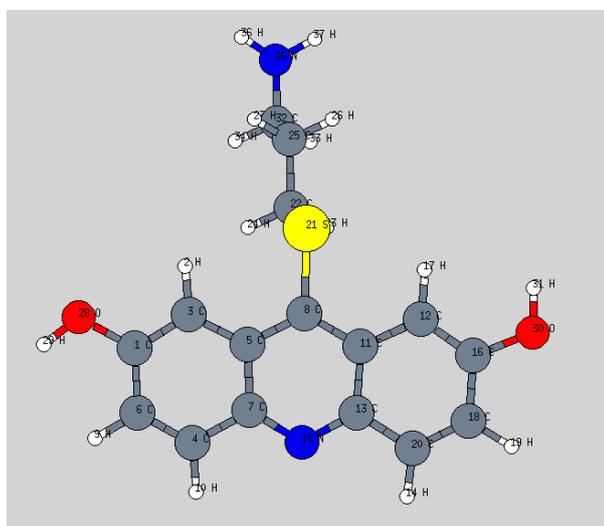


Figura 3.22 – Estrutura da molécula 6 depois de otimizada pelo método DFT.

para uma região específica para a formação dos estados de fronteira vizinhos, fornecendo assim, uma informação local de regiões atômicas.

Para realizarmos o estudo da MIE é necessário dividirmos as estruturas moleculares em regiões, já que a MIE representa um estudo de propriedades locais. Este estudo foi realizado para porções individuais dos compostos acridínicos, e consideramos para cada uma dessas porções suas contribuições individuais, que juntas levam a densidade total de estados dos compostos analisados. Realizamos este estudo utilizando o programa EInsight. O conjunto de descritores foi construído com os valores coletados do cálculo LDOS para os orbitais de fronteira.

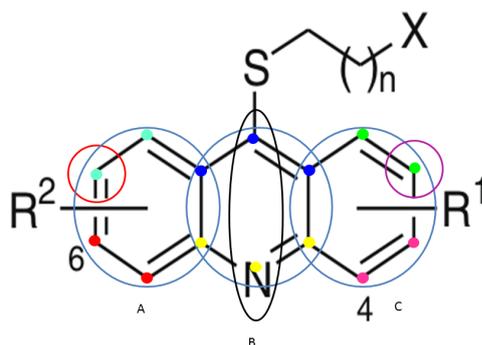


Figura 3.23 – Esqueleto molecular dos compostos análogos da acridina, especificando as regiões para o estudo MIE.

Delimitação	Região
Anel lateral esquerdo	A
Anel central	B
Anel lateral direito	C
Átomos	F
Átomos	X
Átomos	R
Átomos	P
Átomos	W
Átomos	D
Átomos	V
Átomos	S
Átomos	K

Figura 3.24 – Legenda com os nomes das região e suas respectivas representações na figura 3.23.

Primeiramente analisamos parâmetros que nos fornece uma informação global, que não estão relacionados a uma região molecular específica. Os dados analisados são: as energias dos orbitais de fronteira vizinhos (HOMO-1, HOMO, LUMO e LUMO+1, calor de formação(CF), a diferença de energia entre os orbitais HOMO e HOMO-1 (ΔH), a diferença entre LUMO+1 e LUMO (ΔL), momento de dipolo(D) e a diferença entre as energias LUMO e HOMO($\Delta \epsilon$), que corresponde ao gap de energia.

Tabela 3.4 – Tabela com os resultados para as energias dos orbitais de fronteira, suas respectivas diferenças, o gap de energia e o calor de formação (C.F.), para o método PM3.

Molécula	HOMO - 1	HOMO	ΔH	LUMO	LUMO + 1	ΔL	$\Delta \epsilon$	C.F.
1	-8.9107	-8.4883	0.4225	-1.2881	-0.3147	0.97339	7.2002	-4.9726
2	-8.9974	-8.6034	0.3941	-1.3366	-0.3580	0.97857	7.2668	72.7355
3	-9.0102	-8.5050	0.5052	-1.3490	-0.3747	0.97428	7.1560	33.0253
4	-8.9288	-8.4908	0.4380	-1.3025	-0.3376	0.96488	7.1883	24.5533
5	-9.1520	-8.4280	0.7240	-1.4124	-0.4817	0.93067	7.0156	28.2851
6	-8.9742	-8.4882	0.4861	-1.3950	-0.4366	0.95841	7.0931	-19.1674
7	-8.9563	-8.5649	0.3914	-1.3142	-0.3309	0.98324	7.2507	-2.2765
8	-9.1202	-8.5592	0.5609	-1.4546	-0.5016	0.95300	7.1047	28.9787
9	-8.8469	-8.4170	0.4299	-1.3001	-0.3314	0.96873	7.1169	1.8550
10	-8.8540	-8.2855	0.5685	-1.2374	-0.2698	0.96764	7.0481	-9.7376
11	-8.7617	-8.4209	0.3409	-1.2699	-0.2964	0.97346	7.1510	-6.2044
12	-8.7441	-8.2581	0.4859	-1.1853	-0.2212	0.96411	7.0728	-9.9828
13	-8.8505	-8.3010	0.5495	-1.1899	-1.1812	0.00865	7.1112	-44.5190
14	-8.9313	-8.3435	0.5877	-1.2129	-1.1760	0.03691	7.1306	-5.5714

De acordo com os dados apresentados nesta tabela, verificamos que as moléculas enumeradas na primeira coluna por 1,6,7,9,10,11,12 e 13 apresentam os menores valores para o calor de formação (CF), esses valores estão compreendidos entre -19 e 4 Kcal/mol. Essas moléculas apresentam na sua estrutura molecular o elemento oxigênio (O) e um grupo metil ou ainda um grupo metoxi, ligados aos anéis aromáticos laterais nas posições 2 e 7, com exceção da molécula 7 que apresenta os dois oxigênios e seus grupos metil, ligados no mesmo anel, nas posições 2 e 3. As moléculas que, dentro desse grupo, apresentam os maiores CF, são as moléculas 9 e 12 (3,71 e 3,81 Kcal/mol respectivamente) cujas estruturas moleculares são muito parecidas, diferenciam-se nos radicais ligados ao elemento enxofre (S) na parte superior do anel aromático central. Na molécula 12, além de um número maior de carbonos(C), temos também a

inclusão de um nitrogênio (N), o que nos leva a acreditar que modificações neste radical, no que se refere a análise feita aqui para o CF dessas duas estruturas, não contribui significativamente. Sobretudo, quando comparamos as estruturas 12 e 13, que se diferenciam pelo acréscimo de uma ftalimida a estrutura 13, o valor do CF cai de 3,81 para $-18,779$ kcal/mol demonstrando a importância dessa adição na energia.

O outro grupo, com as moléculas enumeradas por 2, 3, 4, 5, 8 e 14, com CF variando de 19 a 71,5 Kcal/mol, apresentam nenhum ou somente um grupo metoxi ligado a um dos anéis laterais. As comparações pra este grupo de moléculas foram feitas par aquelas que apresentava CF com valores próximo.

As moléculas 4 e 5 apresentam uma diferença de 3 Kcal/mol no CF, o que diferenciam-lhes estruturalmente é o grupo metoxi na molécula 4 se liga na posição 7, enquanto na molécula 5 ele está ligado na posição 2; além disso, a molécula 5 apresenta um elemento cloro (Cl) ligado na posição 2, o que significa que, o acréscimo de um Cl, bem como a mudança do grupo metoxi de lugar, acarretou no aumento do CF. As moléculas 3 e 8, para as quais a diferença no CF é de aproximadamente 3 Kcal/mol, e assim como as moléculas 4 e 5 diferenciam-se por um elemento Cl. A molécula 3 tem um grupo metoxi ligado na posição 7, na molécula 8 esse grupo está ligado na mesma posição, contudo apresenta um Cl ligado na posição 6. A adição do Cl, no mesmo anel aromático, representou uma queda de 3 kcal/mol no CF da molécula 8. A comparação pode ser feita também com as moléculas 5 e 8, já que como relacionado acima, as duas apresentam grupos metoxi e Cl, mas em posições bem diferentes. A molécula 5 tem o Cl e o metoxi em anéis diferentes, mas a molécula 8 tem esse elementos no mesmo anel, o que foi representativo, já que a molécula 8 tem um acréscimo de 1,5 kcal/mol em seu CF.

A molécula 14 pertence a esse segundo grupo por conter em sua estrutura somente um grupo metoxi ligado a anel lateral direito, mas ela se diferencia estruturalmente das demais moléculas, pertencentes a esse grupo, aqui comparadas, por apresentar um radical ftalimida ligado ao enxofre do anel central. A molécula 2, porém, não apresenta algum grupo metoxi ou Cl ligado aos anéis laterais, nem o ftalimida como a molécula 14; todavia, é a molécula que,

dentro desse grupo e dentre as analisadas neste trabalho, apresenta o maior valor para o CF.

Depois de discutirmos como as diferenças estruturais influenciavam no valor do CF, analisaremos a seguir os outros parâmetros apresentados na tabela. De acordo com os dados apresentados para os valores do orbital atômico HOMO-1, vemos que para as moléculas enumeradas por 1 e 2 a energia nesse orbital tem uma diferença de $-0,1429$ eV, a molécula 1, apresenta radicais metoxi ligados nas posições 2 e 7, enquanto a molécula 2 não apresenta grupos de radicais ligados aos anéis laterais, a molécula 3 se diferencia por apresentar um grupo metoxi na posição 7 e um grupo metil na posição 2, o que não tem significado no valor da energia, já que a diferença é de $-0,0084$ eV.

Verificamos que a energia do orbital atômico HOMO-1 tem uma diferença muito pequena entre os compostos acridínicos estudados. Percebemos que um pequeno aumento de energia para esse orbital acontece para as moléculas 2, 3, 5 e 8; essas moléculas apresentam somente um grupo metoxi ligados aos anéis laterais, com exceção da molécula 2 que não apresenta quaisquer radicais. A molécula 14, apesar de também apresentar somente um grupo metoxi como radical ligado na posição 7, não sofreu aumento no valor da energia, ainda que pequeno, comparado aos compostos separados neste grupo. Lembrando que, dentro do grupo que apresenta somente um metoxi ligado nos anéis laterais, a molécula 14 se diferencia por apresentar um grupo ftalimida na cadeia ligada ao enxofre do anel central do esqueleto molecular. As diferenças de energia no orbital HOMO-1, em sua totalidade, sofre pequenas alterações no que se refere a inserção de radicais metoxi nos anéis aromáticos desses compostos.

O orbital atômico HOMO apresenta diferenças menores ainda que o orbital HOMO-1, com relação a inserção de grupos metoxi. Observando as moléculas 5 e 8, para o orbital atômico HOMO, ambas possuem os radicais metoxi e Cl ligados aos anéis laterais, mas em posições diferentes e ainda assim apresentam valores de energia, para esse orbital, muito próximos, $-8,4511$ e $-8,4165$ eV, respectivamente. A molécula 2 que não apresenta radicais ligados as suas extremidades possui o menor valor da energia para esse orbital. Enquanto a molécula 12, com dois grupos metoxi ligados nas posições 2 e 7, possui o maior valor de energia $-8,02$

eV para o orbital HOMO dentre as moléculas.

Para o orbital LUMO, as diferenças de energia, entre os compostos que apresentam ou não radicais metoxi, oxigênio ou cloro ligados aos seus anéis laterais, são muito pequenas. Se compararmos a diferença de energia entre as moléculas que apresentam radicais ligados nos dois anéis laterais, como a molécula 1, com moléculas que não apresentam esses radicais, como a molécula 2, a diferença é muito pequena. Para essas duas moléculas, 1 e 2, a diferença é de 0,0501 eV. A molécula 12 é a que, dentre as moléculas, apresenta o maior valor para a energia LUMO de $-1,1709$ eV.

O orbital LUMO+1 diferentemente dos orbitais HOMO-1, HOMO e LUMO, apresenta os menores valores para as moléculas 13 e 14, com energias $-1,1714$ e $-1,1933$ eV respectivamente. A remoção de um grupo metoxi na posição 7 da molécula 14, fez com que a energia do orbital LUMO+1 diminuísse de $-0,0219$ eV. Observamos também que a presença de um grupo metoxi e do cloro, nas moléculas 5 e 8, fizeram com que a energia, ainda que pouco, diminuísse. O que pode ser visto ao compararmos os valores das energias das moléculas 5 e 8, com os demais compostos, exceto com as moléculas 13 e 14, que já foram analisadas e que possuem diferenças estruturais. O comportamento do orbital atômico LUMO+1 é diferente dos demais orbitais estudados, pois os valores de energia sofreram diferenças maiores para as moléculas que apresentavam o grupo ftalimida inserido na cadeia ligada ao enxofre do anel central. A adição ou remoção de grupos metoxi, ou a mudança de posição desses radicais, não alteraram muito os valores das energias dessas duas moléculas nesse orbital.

Para a diferença de energia entre os orbitais ocupado ΔH (HOMO-(HOMO-1)), observamos que os valores apresentados para as moléculas não nos permite uma separação destes compostos, apenas relacionando e comparando os radicais metoxi ligados aos anéis de suas estruturas. Os valores estão bem misturados, de maneira que não podemos conferir-lhes uma padronização.

A diferença de energia entre os orbitais atômicos desocupados ΔL ((LUMO+1)-LUMO) apresentadas na tabela nos mostra que os menores valores para esta diferença é conferido as

moléculas 13 e 14. Mas as demais diferenças não nos fornece um padrão com relação aos radicais presentes nesses compostos.

O gap de energia dado por $\Delta\epsilon$, que é a diferença entre as energias dos orbitais atômicos HOMO e LUMO, com base nos dados expostos na tabela, apresenta um padrão, pois, a presença de radicais ligados aos anéis laterais faz com que o valor do gap decresça.

Para realizarmos o estudo da MIE é necessário dividirmos as estruturas moleculares em regiões, já que a MIE representa um estudo de propriedades locais. Este estudo foi realizado para porções individuais dos compostos acridínicos, e consideramos para cada uma dessas porções suas contribuições individuais, que juntas levam a densidade total de estados dos compostos analisados. Realizamos este estudo utilizando o programa Einsight. O conjunto de descritores foi construído com os valores coletados do cálculo LDOS para os orbitais de fronteira.

Para o cálculo de densidade de estados, dividimos os compostos em várias regiões comum a todos. A escolha dessas regiões podem ser feitas considerando a densidade de carga, a maior localização de orbitais de fronteira HOMO ou LUMO. Estudos foram realizados no sentido de identificar algumas dessas regiões, que para os HAPs e seus derivados, apresentavam maior densidade eletrônica, e portanto, estas eram mais favoráveis a reações eletrofílicas. Ou ainda algumas regiões preferencias, pelas quais os derivados acridínicos possam reagir com as bases do DNA [1, 24, 17].

Além de dividirmos os compostos em regiões também fizemos uma discriminação das moléculas quanto a sua atividade biológica atribuídas por meio de resultados experimentais. O estudo feito por Cuny e colaboradores classificou que os 14 compostos, derivados da acridina, em ativos ou inativos para dois grupos de enzimas, as DYRK2 e as Haspin.

Os parâmetros eletrônicos obtidos pelo PM3 descritos acima foram, o calor de formação (C.F.), as energias para os orbitais de fronteira HOMO, HOMO-1, LUMO, LUMO+1, ΔH [HOMO-(HOMO-1)], ΔL [(LUMO+1)-LUMO], $\Delta\epsilon$ (LUMO-HOMO). A contribuição de cada um dos orbitais e suas respectivas diferenças para a densidade local de estados LDOS,

estes parâmetros foram definidos para uma região qualquer que chamamos de Q. Ao valor do LDOS na região Q produzida pelo orbital HOMO-1, chamamos de CH1Q. Este valor dado a esta região pelo orbital HOMO, chamamos de CHQ. O valor da LDOS para esta região dada pelo orbital LUMO é a CLQ, para o orbital LUMO+1 é CL1Q. As diferenças entre essas contribuições regionais, chamamos de η_{HQ} a diferença entre CHQ e CH1Q e de η_{LQ} a diferença entre CL1Q e CLQ.

A distribuição espacial dos orbitais atômicos HOMO e LUMO, para duas moléculas são mostradas na figura abaixo. Essas duas moléculas escolhidas representam o grupo estudado e são ambas ativas e inativas para as duas enzimas.

Continuamos os nosso estudo, iniciando a investigação da correlação entre descritores teóricos e atividade biológica através da metodologia de índices eletrônicos (MIE).

Os resultados obtidos para a LDOS em algumas regiões são apresentados nas tabelas a seguir.

Tabela 3.5 – Tabela com os resultados para a região P.

Molécula	CH1P	CHP	CLP	CL1P	η_{HP}	η_{LP}
1	0,0338	0,1697	0,2439	0,1501	0,1359	-0,0938
2	0,0905	0,2406	0,2172	0,1561	0,1501	-0,0610
3	0,0308	0,1387	0,2507	0,1385	0,1079	-0,1122
4	0,0249	0,1525	0,2522	0,1320	0,1276	-0,1202
5	0,1179	0,1391	0,2276	0,2041	0,0212	-0,0236
6	0,0559	0,1402	0,2749	0,1661	0,0843	-0,1088
7	0,2307	0,2352	0,2356	0,1586	0,0045	-0,0770
8	0,0399	0,1423	0,2917	0,1690	0,1024	-0,1227
9	0,0345	0,1596	0,2488	0,1487	0,1252	-0,1001
10	0,0545	0,1493	0,2637	0,1581	0,0948	-0,1057
11	0,0095	0,1792	0,2485	0,1469	0,1696	-0,1016
12	0,0638	0,1332	0,2577	0,1574	0,0694	-0,1003
13	0,1241	0,1401	0,2446	0,0005	0,0160	-0,2441
14	0,0982	0,2077	0,2093	0,0000	0,1095	-0,2092

Essa tabela nos fornece informações sobre as contribuições dos parâmetros, orbitais de fronteira e suas respectivas diferenças, de cada região para a LDOS. Começamos pela região que chamamos de região P, que na literatura é conhecida como região k. Nesta região verifica-

mos que a contribuição do orbital HOMO-1 para os análogos acridínicos apresentam valores parecidos para alguns grupos de moléculas. As moléculas 1, 3, 4, 8 e 9, apresentam valor de aproximadamente 0,030 (em unidades arbitrárias), o que não confere nenhum padrão de distribuição destes compostos, quanto aos radicais ligados à eles, e suas respectivas posições. As moléculas 2 e 14 são estruturalmente muito diferentes, mas o valor da contribuição LDOS para esta região é próximo de 0,09, para ambas. O que também ocorre para as contribuições das moléculas 5 e 3, apresentando um valor próximo a 0,11 e para as moléculas 6, 10 e 12, com valor aproximado de 0,05. A molécula 7 possui o maior valor de contribuição desse orbital para essa região, de 0,2307; e o menor valor é apresentado pela molécula 11, de 0,009. O orbital CH1P apresentou a menor contribuição para a LDOS entre todas as outras contribuições de orbitais nessa região. Verificamos que a presença do grupo metoxi, nas posições 6 e 7, no caso da molécula 7, contribuiu para o aumento da LDOS, a presença de dois metoxi em posições 2 e 7, que é o caso da molécula 11, fez com que a contribuição nesse orbital atômico diminuísse drasticamente.

Para a contribuição CHP as moléculas 2, 7 e 14 apresentam os maiores valores, as outras moléculas apresentam valores muito próximos. Observando a maior contribuição para esse orbital que é da molécula 2, comparando com a contribuição da molécula 1, notamos que a ausência de radicais ligados aos anéis laterais contribuiu para que o valor dessa contribuição aumentasse consideravelmente. O que não acontece quando comparamos as moléculas 2 e 7, pois a molécula 7 possui dois grupos metoxi ligados ao primeiro anel e ainda assim possui uma das maiores contribuições para a LDOS, nesse orbital.

A contribuição CLP não apresentou nenhum padrão de distribuição dos compostos acridínicos. A contribuição CL1P nos mostra que a maior contribuição para esse orbital é dada pela molécula 5, que apresenta um cloro ligado na posição 7 e um metoxi ligado na posição 2. Outra molécula que apresenta esses dois grupos é a 8, porém, eles estão ligados no primeiro anel nas posições 6(cloro) e 7(metoxi), no entanto essa molécula apresenta um valor muito próximo ao das demais, para esse orbital. Percebemos que a presença de cloro e metoxi em anéis diferentes, faz com que a contribuição para esse orbital aumente. As moléculas 13 e 14, apresentam radicais

ftalimida(C8H5O2N) ligados a cadeia de carbonos ligados ao enxofre do anel central, e notamos que a presença desses radicais fez com que a contribuição dessas moléculas para esse orbital caísse drasticamente, aproximando-se de zero. De modo geral, o orbital CL1P apresentou a maior contribuição para a LDOS entre todos os orbitais dessa região, com exceção das contribuições das moléculas 13 e 14 para esse orbital.

Tabela 3.6 – Tabela com os resultados para a região W.

Molécula	CH1W	CHW	CLW	CL1W	η_{HW}	η_{LW}
1	0,0160	0,1667	0,4180	0,6154	0,1507	0,1974
2	0,0438	0,2182	0,3957	0,6957	0,1745	0,3000
3	0,0661	0,2019	0,4073	0,6560	0,1358	0,2487
4	0,0434	0,1949	0,4066	0,6544	0,1515	0,2478
5	0,2573	0,1865	0,3878	0,6621	-0,0709	0,2743
6	0,0370	0,1575	0,4175	0,5905	0,1205	0,1730
7	0,1857	0,2235	0,3849	0,6759	0,0378	0,2910
8	0,0792	0,1968	0,4008	0,6203	0,1176	0,2195
9	0,0273	0,1643	0,4192	0,6000	0,1369	0,1808
10	0,1366	0,1599	0,4158	0,6016	0,0233	0,1858
11	0,0078	0,1707	0,4212	0,5989	0,1629	0,1777
12	0,0485	0,1619	0,4147	0,6015	0,1134	0,1868
13	0,6542	0,1601	0,4150	0,0008	-0,4941	-0,4143
14	0,0954	0,1964	0,4072	0,0000	0,1010	-0,4072

A região W compreende a região que contém o heteroátomo de carbono no esqueleto molecular. O orbital CH1W apresentou a menor contribuição para a LDOS entre todos os orbitais dessa região. O maior valor dentro desse orbital é o da molécula 13, que possui radicais ftalimida ligados ao anel central, o que mais visivelmente a diferencia das demais moléculas, mas isso não confere um padrão, já que a molécula 14 também possui esse radical e apresenta um valor muito baixo se comparado a molécula 13. As moléculas 7 e 5 possuem os outros maiores valores, seguidos da molécula 13. Mas isso não confere algum padrão de distribuição de compostos no que se refere a metilação ou adição de grupos metoxi e cloro aos anéis aromáticos desses compostos.

No orbital CHW as moléculas 2, 3 e 7 apresentam os maiores valores para a contribuição, essas moléculas são justamente as que não possuem radicais ligados aos anéis laterais, contém

somente um radical metoxi ligado ao primeiro anel ou dois radicais metoxi ligados ao primeiro anel, respectivamente. Verificamos que moléculas que possuem radicais somente no primeiro anel ou que não possuem radicais apresentam maior contribuição para a LDOS no orbital CHW. As demais moléculas que possuem radicais (metoxi, metil, cloro) ligados aos dois anéis laterais apresentam valores muito próximos para a contribuição da LDOS no orbital CHW.

O orbital CLW não apresentou qualquer padrão e os valores para a contribuição da LDOS nessa região foi de aproximadamente 0,41 (em unidades arbitrárias) para todas as moléculas. O orbital CL1W apresentou a maior contribuição para LDOS entre todos ou outros orbitais dessa região, com exceção das moléculas 13 e 14 que para esse orbital apresentam valores próximos de zero. Verificamos que compostos que não apresentam grupos metoxi ligados aos seus anéis laterais, molécula 2, ou que possuem esses grupos ligados somente em um anel, e nenhum ligante no outro anel lateral, molécula 7; exibem os maiores valores para a contribuição LDOS no orbital CL1W.

Tabela 3.7 – Tabela com os resultados para a região K.

Molécula	CH1K	CHK	CLK	CL1K	η_{HK}	η_{LK}
1	0,0396	0,1627	0,2501	0,1574	0,1230	-0,0927
2	0,1037	0,2216	0,2239	0,1662	0,1179	-0,0578
3	0,0813	0,2082	0,2170	0,1765	0,1269	-0,0405
4	0,0643	0,2161	0,2112	0,1734	0,1518	-0,0378
5	0,0495	0,1185	0,2404	0,1144	0,0690	-0,1260
6	0,0586	0,1441	0,2541	0,1514	0,0854	-0,1027
7	0,0385	0,2527	0,2082	0,1625	0,2142	-0,0457
8	0,1140	0,1974	0,2038	0,1705	0,0834	-0,0332
9	0,0367	0,1640	0,2715	0,1641	0,1273	-0,1074
10	0,0833	0,1307	0,2545	0,1564	0,0474	-0,0981
11	0,0141	0,1850	0,2710	0,1648	0,1709	-0,1062
12	0,0454	0,1558	0,2656	0,1571	0,1105	-0,1085
13	0,1215	0,1338	0,2511	0,0005	0,0122	-0,2507
14	0,0360	0,1233	0,2623	0,0000	0,0873	-0,2623

Para a região k, a análise de CH1K mostrou que esse orbital apresentou a menor contribuição para a LDOS entre todos os orbitais dessa região. Os maiores valores para essa contribuição são dados pelos compostos 2, 8 e 13, que do ponto de vista estrutural são bem diferentes, e portanto

não observamos um padrão de distribuição dos compostos quanto a radicais ligados aos anéis laterais. As demais moléculas também não exibiram uma regularidade para esse orbital.

Na contribuição CHK, o composto 7 apresentou o maior valor para a contribuição LDOS nessa região, essa molécula tem dois grupos metoxi ligados no primeiro anel de sua estrutura molecular nas posições 6 e 7. O valor caiu bastante para a molécula 8, que com relação a molécula 7, substituiu o metoxi d posição 6 por um cloro. A mudança de posição de um grupo metoxi, da posição 6 para a posição 2, molécula 9, fez com que o valor da contribuição para a LDOS nesse orbital diminuísse de 0,25 para 0,16 (em unidades arbitrárias). As moléculas 3 e 4 apresentam valores muito próximos, mostrando que a adição de um metil na posição 2 (molécula 4) não exerceu influência sobre o valor da contribuição para a LDOS, dessas moléculas, nessa região. As moléculas 5 e 14 apresentaram os menores valores, mas não mostram regularidade alguma, assim como as demais moléculas.

O orbital CLK não exprime regularidade na distribuição de compostos quanto a semelhança ou disparidade de radicais ligados ao esqueleto molecular. Mas esse orbital apresentou a maior contribuição para a LDOS entre todos os orbitais dessa região. O orbital CL1K também não apresenta uma regularidade, os valores para todas as moléculas fica em torno de 0,16 (em unidades arbitrárias), com exceção das moléculas 13 e 14 que apresentam valores muito pequenos para essa contribuição, próximo de zero, e a molécula 5 que apresenta valor de 0,11. Verificamos que a adição de radicais na parte superior do anel central, no caso das moléculas 13 e 14 é um grupo ftalimida ligado aos carbonos que está ligado ao enxofre no anel central, faz com que o valor da contribuição para a LDOS, dessas moléculas nessa região, reduza de modo drástico.

Partimos agora para a próxima etapa que é, utilizando os dados encontrados, procurar um padrão de identificação da atividade carcinogênica. A classificação da atividade dos compostos que utilizamos está baseada no trabalho de Cuny et al. [25], onde os 14 compostos estudados foram divididos em dois grupos: ativas ($ID_{50} < 5\mu M$) e inativas ($ID_{50} \geq 5\mu M$). O índice biológico in vitro ID_{50} indica a dose de droga necessária para diminuir em 50% a multiplicação das células cancerígenas.

Tabela 3.8 – Tabela com os valores do IC_{50} para a classificação da atividade biológica dos 14 compostos para as enzimas Haspin e DYRK2.

Molécula	$IC_{50}(\mu M)$	
	Haspin	DYRK2
1	0.010 ²⁴	0.018
2	9.8	7.7
3	0.017	0.19
4	0.070	0.43
5	0.055	9.9
6	0.044	0.17
7	~7	> 10
8	0.079	9.4
9	0.20	0.29
10	0.040	0.036
11	0.008	0.12
12	0.046	0.23
13	2.1	0.39
14	8.8	1.3

O estudo realizado por Cuny et al. [25] apresenta a atividade desses compostos para dois grupos de enzimas, o Haspin e o DYRK2. O estudo baseado nos parâmetros da MIE será realizado separadamente para estes dois grupos, visto que os resultados apresentados por Cuny mostram que a mesma molécula pode ser ativa para uma enzima e inativa para a outra.

3.2.1 Estudo MIE para a enzima Haspin

Primeiramente realizaremos a análise da MIE para a enzima Haspin. Começamos nossa análise pela energia dos orbitais de fronteira na procura de um modelo para a atividade cancerígena dos compostos estudados. Com base na tabela 3.2 verificamos que compostos que apresentam valores de energia para o orbital HOMO-1 superiores a -8,93 eV (HOMO-1 > -8,93 eV) apresentam propriedade inibitória dessa enzima, são consideradas biologicamente ativas. No entanto apresentou algumas exceções para os compostos 3, 5 e 8, de forma que o resultado não é satisfatório, pois o número de exceções é bem significativo.

Para o orbital HOMO, verificamos que compostos que apresentam valores de energia maior ou igual a -8,55 são ativos para a enzima Haspin, com exceção da molécula 14. Consideramos a identificação da atividade carcinogênica para esse orbital um resultado bom, já que

o número de exceções é pequeno.

O orbital LUMO não apresentou nenhum valor padrão para a identificação da atividade dos compostos, uma vez que o número de exceções para esse orbital foi muito elevado, e ele não foi levado em consideração. Já o orbital LUMO+1 apresentou um limiar maior ou igual a $-0,5$ para o qual as moléculas são biologicamente ativas, com exceções das moléculas 2, 7 e 13. O que não podemos considerar um dos melhores resultados, porque as exceções foram para duas moléculas, 2 e 7, que segundo os resultados experimentais obtidos por Cuny et al. [25] essas moléculas são inativas para essa enzima, tornando o erro muito grande para esse resultado.

A análise para as diferenças entre as energias desses orbitais no mostrou que para a diferença entre os dois últimos orbitais ocupados ΔH [HOMO-(HOMO-1)] apresentou um limiar de $0,42$ eV, moléculas que apresentaram valores maiores ou igual a este limiar, exibem atividade biológica na inibição da enzima Haspin, com exceções das moléculas 11 e 14. As diferenças ΔL e $\Delta \epsilon$ não mostraram um valor padrão para a atividade.

Analisando a contribuições para a LDOS dos orbitais, nas regiões estudadas, algumas delas apresentam parâmetros com os quais podemos prever a atividade biológica dos compostos.

Para o método AM1 as regiões P e W foram as que mais contribuíram para a identificação da atividade carcinogênica das moléculas. Os melhores resultados para análise com uma variável foram:

Tabela 3.9 – Tabela com os resultados de análise com uma variável para enzima Haspin, utilizando o método AM1.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
CHP	$CHP < 0.15$	7
CHW	$CHW < 0.185$	8
CH1X	$CH1X > 0.06$	7 e 5
η_{LK}	$\eta_{LK} < -0.10$	3,4,8 e 13
η_{LP}	$\eta_{LP} < -0.12$	5 e 13
η_{HP}	$\eta_{HP} < 0.015$	3,4,6 e 8
η_{LW}	$\eta_{LW} < 0.32$	7 e 13

Para análise a partir de uma única variável, escolhemos a variável CHW que apresenta somente um erro, a molécula 8, que para essa enzima é ativa. A variável CHP também é um bom parâmetro, conforme mostra a tabela 3.10, esse parâmetro indica compostos ativos com apenas uma exceção, no entanto, essa exceção é a molécula 7, que faz parte do grupo de moléculas inativas, e o erro para esse parâmetro é maior que para CHW.

Tabela 3.10 – Tabela booleana da análise com uma variável, para enzima Haspin, utilizando o método AMI.

Molécula	$CHW < 0.185$	Atividade	Classificação experimental
1	Sim	Ativo	ATIVO
2	Não	Inativo	INATIVO
3	Sim	Ativo	ATIVO
4	Sim	Ativo	ATIVO
5	Sim	Ativo	ATIVO
6	Sim	Ativo	ATIVO
7	Não	Inativo	INATIVO
8	Não	Inativo	ATIVO
9	Sim	Ativo	ATIVO
10	Sim	Ativo	ATIVO
11	Sim	Ativo	ATIVO
12	Sim	Ativo	ATIVO
13	Sim	Ativo	ATIVO
14	Não	Inativo	INATIVO

Para a análise envolvendo a contribuição CHW, obtemos a seguinte regra para a atividade:

Regra: Se $CHW < 0.185$ a molécula será ativa, caso contrário será inativa.

Com esta regra nota-se que apenas a molécula 8 é classificada erroneamente quanto a atividade biológica. Isso pode ser melhor visualizado por meio do gráfico 3.25 abaixo.

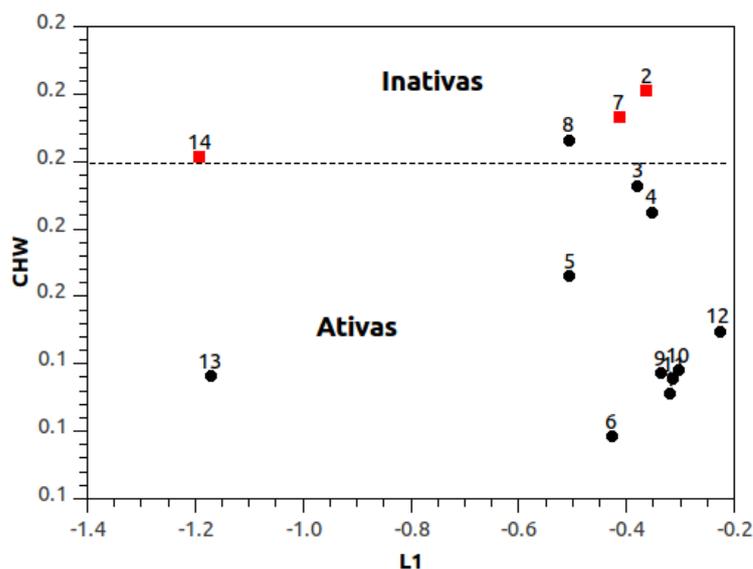


Figura 3.25 – Gráfico da análise da variável CHW, para a enzima Haspin com o método AM1.

Após analisar a predição para a atividade biológica dos compostos utilizando uma variável, partimos para uma identificação dessa atividade baseada na análise de duas variáveis, a análise binária. Por meio dessa análise encontramos padrões de separação de atividade, das contribuições LDOS para um mesmo orbital, por combinações de regiões diferentes.

Os melhores resultados obtidos para análise com duas variáveis, para esta enzima pelo método AM1, estão relacionados abaixo.

Tabela 3.11 – Tabela com os melhores resultados de análise binária para enzima Haspin, utilizando o método AM1.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
H e CHC	$H > -8.55$ e $CHC < 0.48$	5 e 7
ΔL e CHV	$\Delta L > 0.8$ e $CHV \geq 0.319$	7 e 13
CHW e CL1W	$CHW < 0.185$ e $CL1W \geq 0.40$	

Para análise binária escolhemos as contribuições CL1W (contribuição para o orbital LUMO+1) e CHW (contribuição para o orbital HOMO), ambas na região W. As regiões V e C também se mostraram importantes para essa análise, mas apresentam 2 erros, por isso não foram escolhidas para regra. A tabela 3.12 mostra a análise para as variáveis escolhidas.

Tabela 3.12 – Tabela booleana para análise binária feita com as variáveis CL1W e CHW para a enzima Haspin, utilizando o método AM1, com resultados da predição comparados com os obtidos experimentalmente.

Molécula	$CL1W \geq 0.40$	$CHW < 0.185$	Atividade	Classificação experimental
1	Sim	Sim	Ativo	ATIVO
2	Sim	Não	Inativo	INATIVO
3	Sim	Sim	Ativo	ATIVO
4	Sim	Sim	Ativo	ATIVO
5	Sim	Sim	Ativo	ATIVO
6	Sim	Sim	Ativo	ATIVO
7	Sim	Não	Inativo	INATIVO
8	Sim	Sim	Ativo	ATIVO
9	Sim	Sim	Ativo	ATIVO
10	Sim	Sim	Ativo	ATIVO
11	Sim	Sim	Ativo	ATIVO
12	Sim	Sim	Ativo	ATIVO
13	Sim	Sim	Ativo	ATIVO
14	Não	Sim	Inativo	INATIVO

Para a análise binária, envolvendo as contribuições CHW e CL1W, obtemos a seguinte regra do tipo *e* para a atividade:

Regra: Se $CHW < 0.185$ e $CL1W \geq 0.40$ a molécula será ativa, caso contrário será inativa.

Verificamos que o acerto dessa regra MIE é de 100%, identifica corretamente a atividade inibitória dos compostos em relação a enzima Haspin. Esse critério pode ser visualizado no gráfico 3.26, que mostra a separação do grupo de moléculas ativas e inativas para essas duas variáveis.

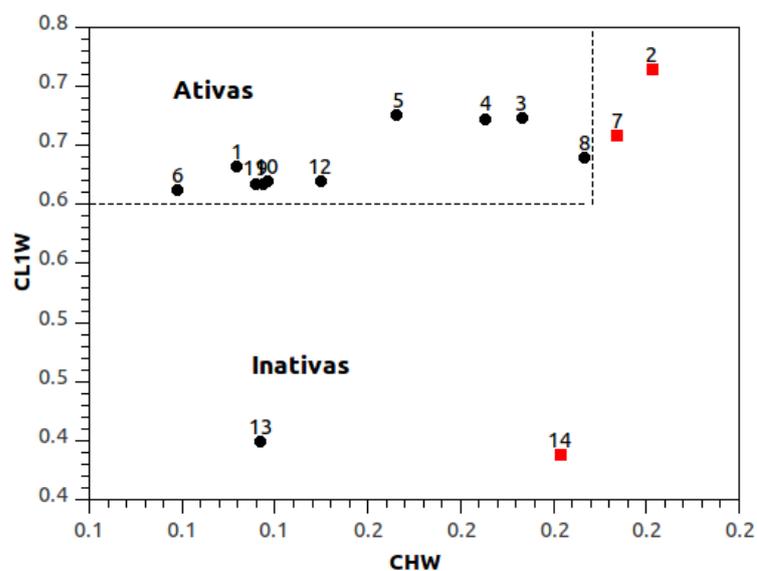


Figura 3.26 – Gráfico da análise binária, com os parâmetros CLIW e CHW para a enzima Haspin com o método AMI.

Para o método PM3 as regiões A, B, C e D não apresentaram nenhum padrão de separação entre as moléculas ativas e inativas. As regiões S e X apresentaram separações com muitos erros. Os melhores resultados obtidos, a partir da análise das contribuições para a LDOS, que apresenta um padrão para a classificação utilizando somente uma variável, para este método foram:

Tabela 3.13 – Tabela com os resultados de análise com uma variável para enzima Haspin, utilizando o método PM3.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
CHP	$CHP < 0.190$	
CHK	$CHK < 0.213$	14
CLP	$CLP > 0.24$	5
CHC	$CHC > 0.35$	3,5 e 8
CLV	$CLV < 0.57$	14
CHW	$CHW < 0.208$	14
η_{LD}	$\eta_{LD} > -0.08$	5,12 e 13
η_{LF}	$\eta_{LF} > -0.735$	4 e 3
η_{LK}	$\eta_{LK} < -0.08$	3,4,8 e 14
η_{LP}	$\eta_{LP} < -0.08$	7 e 13
η_{HV}	$\eta_{HV} > 0.22$	5,8 e 13
η_{LV}	$\eta_{LV} > 0.1$	13
η_{LW}	$\eta_{LW} < 0.28$	14

Resultado para análise com uma variável para a enzima Haspin com o método PM3, a variável que melhor discriminou as moléculas em grupos de ativas e inativas foi CHP. A densidade local de estados na região P exibiu um valor limite para a identificação da atividade carcinogênica, o que está de acordo com os estudos teóricos anteriores [26]. Ao analisarmos os compostos análogos da acridina, identificamos a contribuição da densidade para o orbital Homo na região P, como um bom parâmetro para identificação da atividade biológica destas moléculas, pois essa variável não apresenta exceções na predição.

Para essa análise, a regra que define as moléculas como ativas é:

Regra: Se $CHP < 0.190$, a molécula será ativa, caso contrário será inativa.

Essa regra mostra que apenas a variável CHP é suficiente para distinguir os compostos ativos. A figura 3.27 mostra como se dá essa divisão.

Tabela 3.14 – Tabela booleana da análise com uma variável, para enzima Haspin, utilizando o método PM3.

Molécula	$CHP < 0.190$	Atividade	Classificação experimental
1	Sim	Ativo	ATIVO
2	Não	Inativo	INATIVO
3	Sim	Ativo	ATIVO
4	Sim	Ativo	ATIVO
5	Sim	Ativo	ATIVO
6	Sim	Ativo	ATIVO
7	Não	Inativo	INATIVO
8	Sim	Ativo	ATIVO
9	Sim	Ativo	ATIVO
10	Sim	Ativo	ATIVO
11	Sim	Ativo	ATIVO
12	Sim	Ativo	ATIVO
13	Sim	Ativo	ATIVO
14	Não	Inativo	INATIVO

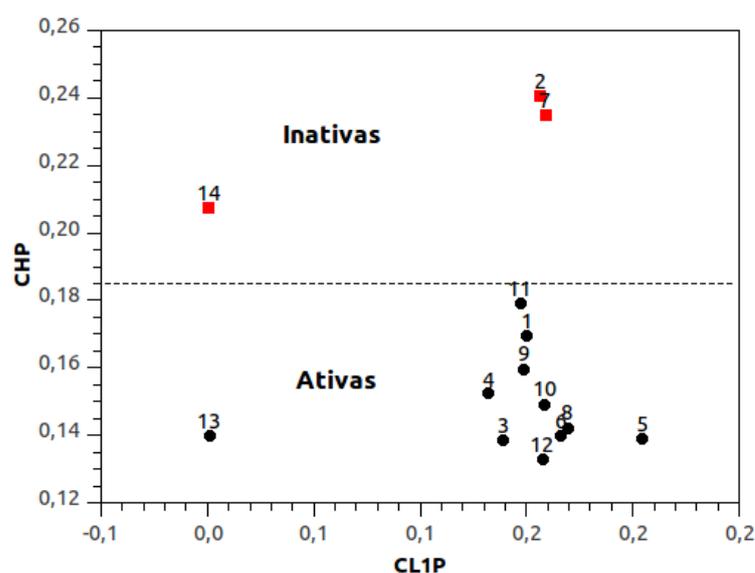


Figura 3.27 – Gráfico da análise da variável CHP, para a enzima Haspin com o método PM3.

Na análise binária as regiões mais relevantes foram K, P, V e W. Os orbitais LUMO, LUMO+1 e HOMO também foram importantes na identificação dos compostos ativos. A tabela 3.15 contém os melhores resultados obtidos pelo método PM3 para a enzima Haspin.

Tabela 3.15 – Tabela com os resultados de análise binária para enzima Haspin, utilizando o método PM3.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
CL1W e CHW	$CL1W \geq 0.5$ e $CHW < 0.208$	13
L1 e CHW	$L1 > -0.6$ e $CHW < 0.213$	13
CL1K e CHK	$CL1K > 0.10$ e $CHK < 0.208$	13
CHK e L1	$CHK < 0.213$ e $L1 > -0.6$	14
CL1R e H	$CL1R > 0.10$ e $H > -8.55$	13
CL1V e CLV	$CL1V > 0.5$ e $CLV < 0.57$	13
ΔL e CHW	$\Delta L > 0.9$ e $CHW < 0.207$	13
CHW e CL1W	$CHW < 0.216$ e $CL1W > 0.5$	13
L1 e CHV	$L1 > -0.6$ e $CHV > 0.32$	7 e 13

De acordo com os resultados obtidos para o método PM3 as regiões P e K, mostraram-se eficientes ao expressarem um valor limite para a identificação da atividade carcinogênica dos compostos. Essas regiões são conhecidas como região K que possui a característica de facilitar a junção de outras estruturas ao esqueleto molecular da estrutura em questão. A densidade local de estados na região P, mostrou-se importante com as contribuições no orbital HOMO (CHP) que consegue separar as moléculas em grupos de ativas e inativas, sem exceções. A outra contribuição nessa região para a LDOS (CHK) apresentou somente uma exceção. O orbital LUMO também apresenta resultados significantes para essa região, a contribuição CLP é uma boa variável para identificação do grupo de compostos ativos com apenas um erro. Outras contribuições para o orbital LUMO na região V (CLV) e para o orbital HOMO na região W mostraram-se interessantes por apresentarem uma exceção.

Empregando a MIE, a regra que melhor distingue as moléculas ativas é representada a seguir:

Regra: Se $L1 > -0.6$ e $CHK < 0.213$ a molécula será ativa, caso contrário será inativa.

Verificamos que, com essa regra a molécula 13 não é classificada corretamente quanto a sua atividade biológica. A predição da MIE produz um resultado de 92,86% de acerto para as moléculas estudadas. O gráfico 3.28 mostra a disposição das moléculas para análise com as dois parâmetros em questão.

Tabela 3.16 – Tabela booleana para análise binária feita com as variáveis $LUMO+1$ e CHK para a enzima Haspin, utilizando o método PM3, com resultados da predição comparados com os obtidos experimentalmente.

Molécula	$L1 > -0.6$	$CHK < 0.213$	Atividade	Classificação experimental
1	Sim	Sim	Ativo	ATIVO
2	Sim	Não	Inativo	INATIVO
3	Sim	Sim	Ativo	ATIVO
4	Sim	Sim	Ativo	ATIVO
5	Sim	Sim	Ativo	ATIVO
6	Sim	Sim	Ativo	ATIVO
7	Sim	Não	Inativo	INATIVO
8	Sim	Sim	Ativo	ATIVO
9	Sim	Sim	Ativo	ATIVO
10	Sim	Sim	Ativo	ATIVO
11	Sim	Sim	Ativo	ATIVO
12	Sim	Sim	Ativo	ATIVO
13	Não	Sim	Inativo	ATIVO
14	Não	Sim	Inativo	INATIVO

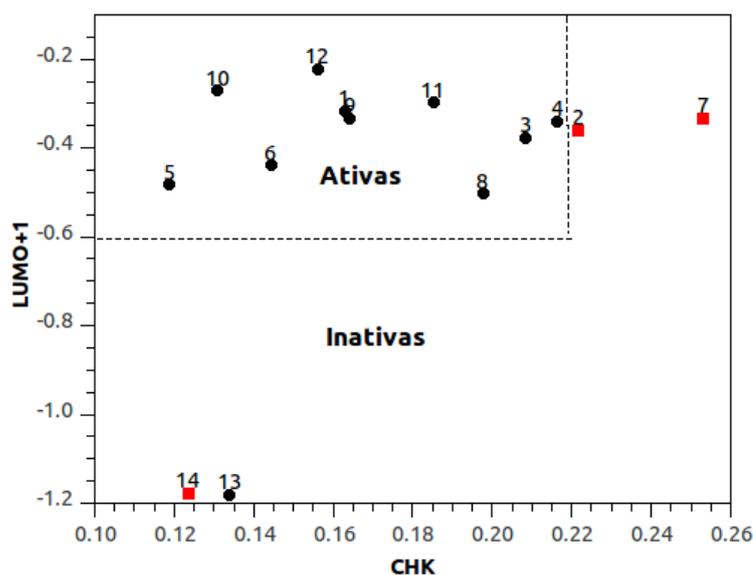


Figura 3.28 – Gráfico da análise binária, com os parâmetros $LUMO+1$ e CHK para a enzima Haspin com o método PM3.

3.2.2 Estudo MIE para a enzima DYRK2

A análise com os índices eletrônicos para a enzima DYRK2 procede da mesma maneira que para a enzima Haspin, no entanto, para essa enzima os parâmetros envolvidos na predição da atividade biológica são diferentes, visto que o comportamento de algumas moléculas se

diferenciam para essa enzima. De acordo com os estudos de Cuny *et al.* [25], para a DYRK2 as moléculas 2, 5, 7 e 8 mostraram-se incapazes de inibi-la. E, portanto, foram classificadas como inativas.

Para a investigação usando o método AM1, baseado na análise de uma variável, encontramos os resultados apresentados na tabela 3.17.

Tabela 3.17 – Tabela com os resultados de análise com uma variável para enzima DYRK2, utilizando o método AM1.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
H	$H > -8.55$	3
H1	$H1 > -9.00$	3
CHK	$CHK < 0.15$	3 e 5
CHP	$CHP > 0.13$	8 e 14
CHC	$CHC > 0.331$	-
L	$L > -1.40$	2 e 7
ηHW	$\eta HW < -0.15$	3,8,9 e 11
ηLP	$\eta LP < -0.01$	6,8 e 13
ηLS	$\eta LS > -0.08$	9
ηLV	$\eta LV > 0.15$	8 e 9

As regiões K, P, S e V são as que mostram-se mais eficientes na separação dos grupos de moléculas ativas e inativas, apresentando poucos erros. Os orbitais LUMO, HOMO e HOMO-1 mostram bons resultados na distinção de moléculas ativas, apresentando 1 erro (LUMO) e 2 erros (HOMO e HOMO-1) somente. Apesar do orbital LUMO apresentar menos erros na predição, o orbital HOMO aproxima mais as moléculas ativas.

A regra baseada na análise com o parâmetro HOMO, e expressada por:

Regra: Se $H > -8.55$ a molécula será ativa, caso contrário será inativa.

Tabela 3.18 – Tabela booleana da análise com uma variável, para enzima DYRK2, utilizando o método AM1.

Molécula	$H > -8.55$	Atividade	Classificação experimental
1	Sim	Ativo	ATIVO
2	Não	Inativo	INATIVO
3	Não	Inativo	ATIVO
4	Sim	Ativo	ATIVO
5	Não	Inativo	INATIVO
6	Sim	Ativo	ATIVO
7	Não	Inativo	INATIVO
8	Não	Inativo	INATIVO
9	Sim	Ativo	ATIVO
10	Sim	Ativo	ATIVO
11	Sim	Ativo	ATIVO
12	Sim	Ativo	ATIVO
13	Sim	Ativo	ATIVO
14	Sim	Ativo	ATIVO

Com essa regra, observamos que a molécula 3 é classificada erroneamente. O gráfico 3.29 mostra o comportamento dos compostos para o parâmetro HOMO.

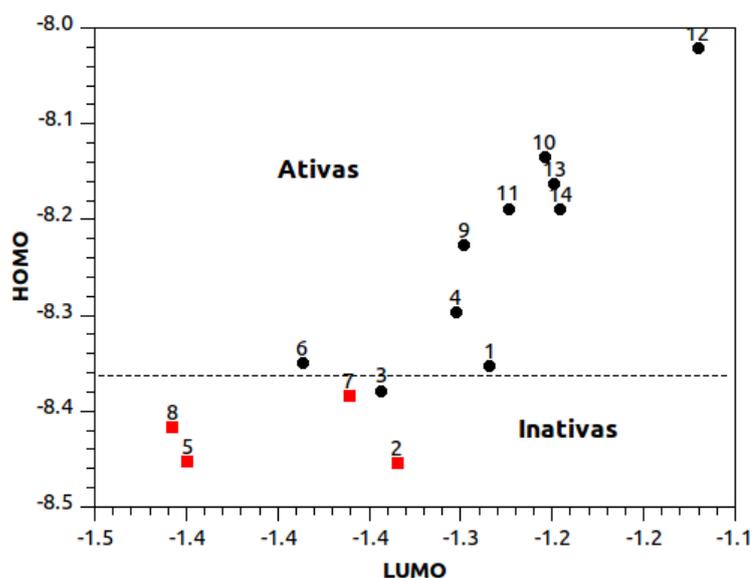


Figura 3.29 – Gráfico da análise da variável HOMO, para a enzima DYRK2 com o método AM1.

Os melhores resultados para o método AM1, com duas variáveis, para a enzima DYRK2, são apresentados na tabela 3.19. Nesta análise somente as regiões K e P exibiram separações entre moléculas ativas e inativas, com poucos erros. Os orbitais HOMO e LUMO também

apresentaram um padrão para a atividade dos compostos estudados. Escolhemos a contribuição para o orbital HOMO na região K (CHK) e o orbital LUMO, pois apresenta somente um erro, com a molécula 3, que é ativa para essa enzima.

Tabela 3.19 – Tabela com os resultados de análise binária para enzima DYRK2, utilizando o método AM1.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
CHK e L	$CHK < 0.15$ e $L > -0.140$	3
CHP e L	$CHP < 0.17$ e $L > -0.140$	7
ΔH e L	$\Delta H > 0.61$ e $L > -0.140$	1 e 7

Os padrões obtidos com os valores críticos dos parâmetros CHK e LUMO, são exibidos na tabela Booleana 3.20, para cada um dos compostos.

Tabela 3.20 – Tabela booleana para análise binária feita com as variáveis CHK e LUMO para a enzima DYRK2, utilizando o método AM1, com resultados da predição comparados com os obtidos experimentalmente.

Molécula	$CHK < 0.15$	$L > -0.140$	Atividade	Classificação experimental
1	Sim	Sim	Ativo	ATIVO
2	Não	Sim	Inativo	INATIVO
3	Não	Sim	Inativo	ATIVO
4	Sim	Sim	Ativo	ATIVO
5	Sim	Não	Ativo	INATIVO
6	Sim	Sim	Ativo	ATIVO
7	Não	Sim	Inativo	INATIVO
8	Sim	Não	Inativo	INATIVO
9	Sim	Sim	Ativo	ATIVO
10	Sim	Sim	Ativo	ATIVO
11	Sim	Sim	Ativo	ATIVO
12	Sim	Sim	Ativo	ATIVO
13	Não	Sim	Inativo	ATIVO
14	Não	Sim	Inativo	ATIVO

A regra para atividade, com os parâmetros MIE escolhidos, é expressada por:

Regra: Se $CHK < 0.15$ e $L > -0.140$ a molécula será ativa, caso contrário será inativa.

Para essa regra a molécula 3 é classificada erroneamente. Assim, a predição utilizando a MIE apresentou 92,86% de acerto na classificação quanto a atividade biológica dos compostos derivados de acridina, estudados. O gráfico 3.30 apresenta a distribuição dos compostos nesta análise.

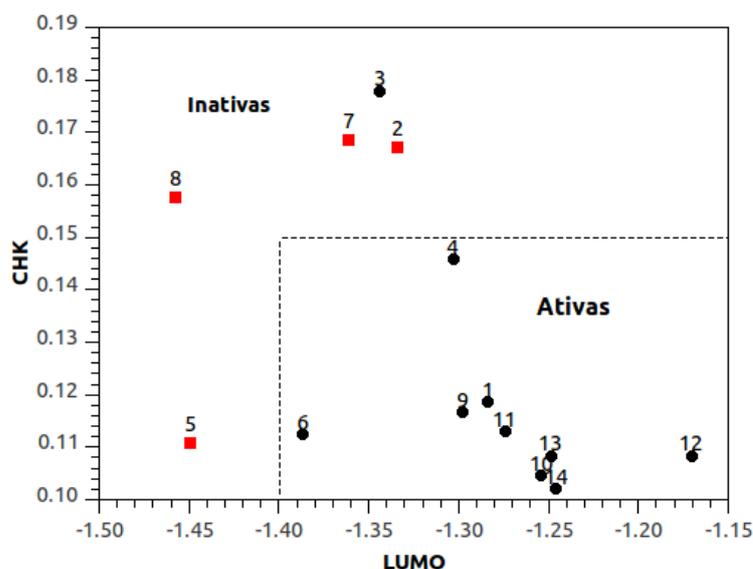


Figura 3.30 – Gráfico da análise binária, com os parâmetros CHK e LUMO para a enzima DYRK2 com o método AM1.

Os melhores resultados para análise com uma variável para a enzima DYRK2, pelo método PM3, é apresentado na tabela 3.21. Para essa análise as regiões A, C, P, V e W apresentaram um padrão de separação dos compostos ativos. Os orbitais HOMO e HOMO-1 apresentaram-se como bons parâmetros para distinção dos compostos ativos e inativos. Para essa análise selecionamos a contribuição para o orbital LUMO na região W (clw), pois este não apresenta restrições em sua predição.

Tabela 3.21 – Tabela com os resultados de análise com uma variável para enzima DYRK2, utilizando o método PM3.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
H	$H > -8.55$	5
H1	$H1 > -9.05$	2 e 7
CLA	$CLA < 0.606$	2
CLC	$CLC > 0.562$	2
CHC	$CHC > 0.34$	-
CLW	$CLW > 0.405$	-
η_{HC}	$\eta_{HC} > 0.3$	3,10 e 13
η_{LP}	$\eta_{LP} < -0.08$	8
η_{HV}	$\eta_{HV} > 0.22$	13 e 14

A tabela Booleana 3.22 construída com base no valor limite para CLW, mostra a eficiência do mesmo, na classificação dos compostos ativos.

Tabela 3.22 – Tabela booleana da análise com uma variável, para enzima DYRK2, utilizando o método PM3.

Molécula	$CLW > 0.405$	Atividade	Classificação experimental
1	Sim	Ativo	ATIVO
2	Não	Inativo	INATIVO
3	Sim	Ativo	ATIVO
4	Sim	Ativo	ATIVO
5	Sim	Inativo	INATIVO
6	Sim	Ativo	ATIVO
7	Não	Inativo	INATIVO
8	Sim	Inativo	INATIVO
9	Sim	Ativo	ATIVO
10	Sim	Ativo	ATIVO
11	Sim	Ativo	ATIVO
12	Sim	Ativo	ATIVO
13	Sim	Ativo	ATIVO
14	Não	Ativo	ATIVO

A regra baseada na análise com a contribuição CLW, e expressada por:

Regra: Se $CLW > 0.405$ a molécula será ativa, caso contrário será inativa.

Essa regra não apresenta exceções. A distribuição das moléculas pode se vista no gráfico 3.31.

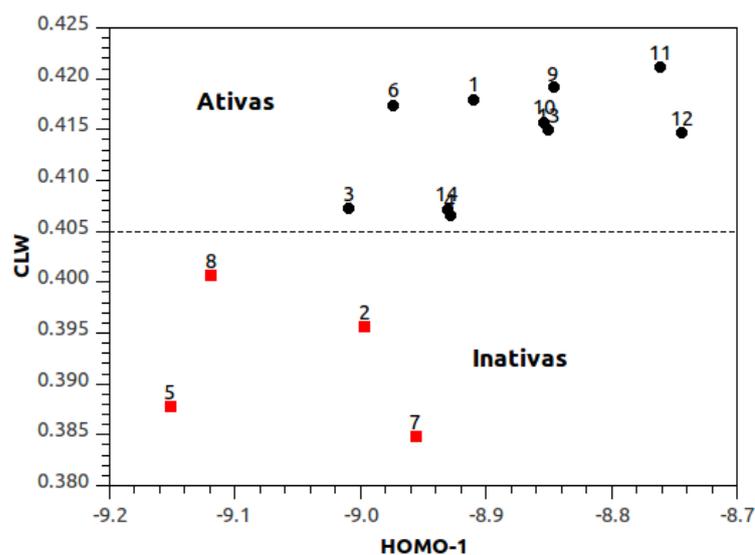


Figura 3.31 – Gráfico da análise da variável CLW, para a enzima DYRK2 com o método PM3.

Os resultados obtidos com a análise binária mostram que as regiões C, D, F, K, P e V

apresentam um padrão de separação para as moléculas ativas. Os orbitais HOMO, HOMO–1 e LUMO mostraram-se bons parâmetros para distinguir os compostos ativos. Dentre estes parâmetros selecionamos a contribuição para o orbital HOMO na região P (CHP), e o orbital HOMO–1, pois dentre as separações esta combinação foi a que melhor agrupou os compostos ativos.

Tabela 3.23 – Tabela com os resultados de análise binária para enzima DYRK2, utilizando o método PM3.

Variáveis	Regras para atividade carcinogênica dos compostos	Restrições
CLC e H	$CLC > 0.562$ e $H > -8.55$	-
CLD e H1	$CL1D < 0.23$ e $H1 > -9.05$	2
CL1D e H	$CL1D < 0.23$ e $H > -8.55$	-
CL1F e H	$CL1F < 0.0015$ e $H > -8.55$	-
ΔH e H	$\Delta H < 0.60$ e $H > -8.55$	-
CHK e H1	$CHK < 0.215$ e $H1 > -9.05$	2
CHP e H1	$CHP < 0.220$ e $H1 > -9.05$	-
L e CHP	$CHP < 0.220$ e $L > -1.40$	-
H e CL1P	$H > -8.55$ e $CL1P < 0.15$	5
H1 e CHV	$H1 > -9.05$ e $CHV > 0.34$	-

A tabela 3.24 apresenta a análise para classificação por meio de uma tabela Booleana, para os valores críticos dos parâmetros CHP e HOMO–1.

Tabela 3.24 – Tabela booleana para análise binária feita com as variáveis HOMO-1 e CHP para a enzima Haspin, utilizando o método PM3, com resultados da predição comparados com os obtidos experimentalmente.

Molécula	$H1 > -0.6$	$CHP < 0.215$	Atividade	Classificação experimental
1	Sim	Sim	Ativo	ATIVO
2	Sim	Não	Inativo	INATIVO
3	Sim	Sim	Ativo	ATIVO
4	Sim	Sim	Ativo	ATIVO
5	Não	Sim	Inativo	INATIVO
6	Sim	Sim	Ativo	ATIVO
7	Sim	Não	Inativo	INATIVO
8	Não	Sim	Inativo	INATIVO
9	Sim	Sim	Ativo	ATIVO
10	Sim	Sim	Ativo	ATIVO
11	Sim	Sim	Ativo	ATIVO
12	Sim	Sim	Ativo	ATIVO
13	Sim	Sim	Ativo	ATIVO
14	Sim	Sim	Ativo	IATIVO

A regra para atividade, com os parâmetros MIE selecionados, é expressada por:

Regra: Se $H1 > -0.6$ e $CHP < 0.215$ a molécula será ativa, caso contrário será inativa.

A predição de atividade, utilizando os parâmetros MIE apresentou 100% de acerto na classificação quanto a atividade biológica dos compostos derivados de acridina, estudados. O gráfico 3.32 apresenta a distribuição dos compostos nesta análise.

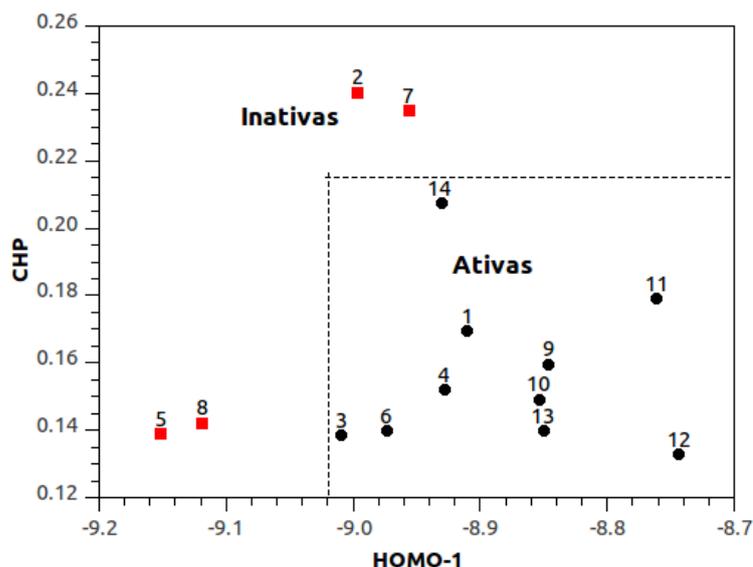


Figura 3.32 – Gráfico da análise binária, com os parâmetros HOMO-1 e CHP para a enzima DYRK2 com o método PM3.

3.2.3 Estudo PCA para a enzima Haspin

A metodologia de índices eletrônicos (MIE), é uma ferramenta para classificar os compostos quanto a sua atividade, ativos ou inativos, (qualitativa) e não tem a finalidade, ou não foi criada com a finalidade de avaliar os compostos quanto à veemência de sua atividade. Contudo, os parâmetros eletrônicos, através de padrões mais apurados podem fornecer uma indicação da potência da atividade dos compostos. Nesse aspecto, os parâmetros MIE demonstraram mais uma vez um grande poder na descrição da atividade biológica. O estudo MIE apresenta para mais da capacidade de classificação dos compostos, numa análise binária, da atividade; uma informação mais completa dos compostos estudados. O próximo passo para que os parâmetros da MIE sejam utilizados numa investigação de potencialidade da atividade, é utilizar os descritores teóricos e correlacioná-los com a atividade biológica, através da análise de componentes

principais (PCA) [28]

A PCA condensa dentro de uma mesma análise, dezenas de parâmetros, estabelecendo combinações lineares entre parâmetros que exprimem a maior parte da variância e efetua um estudo de correlação entre os descritores teóricos e a atividade biológica. Para o estudo com a metodologia PCA consideramos parâmetros calculados pelo método AM1 e PM3, dentro de um programa Einsight. Fizemos a coleta dos descritores teóricos, para os dois métodos envolvidos, já que a predição para a MIE apresenta um grande número de acertos e os descritores teóricos sofreram algumas mudanças quando calculados para os métodos semi-empíricos AM1 e PM3.

A partir de 78 variáveis, conseguimos obter uma separação entre as moléculas ativas e inativas. Os descritores foram obtidos com o programa Chem2pac, para o qual utilizamos os arquivos de saída do programa GAMESS em coordenadas internas no estilo mopac. De acordo com a metodologia de PCA a matriz inicial utilizada, nesse estudo, foi a matriz de dimensão 14 (compostos ou amostras) x 78 (número de variáveis). A seleção de variáveis que separasse os compostos em dois grupos de moléculas, ativas e inativas, foi realizado com o programa Einsigth. A escolha desses parâmetros partiu do estudo, feito anteriormente, com os descritores MIE para a classificação dos compostos. Os descritores MIE nos deram indícios de que parâmetros (um mínimo de três parâmetros é permitido pelo programa Einsigth para que se realize o estudo com a metodologia de PCA) seriam suficientes para distinguir os compostos em ativos e nativos.

Em sua maior parte, os parâmetros que exibiram melhor resultado na metodologia de PCA foram os parâmetros eletrônicos relacionados com os descritores da MIE. Uma análise baseada na aleatoriedade na escolha de parâmetros, que não necessariamente estariam envolvidos com os descritores MIE, foi realizada. No entanto, com as combinações entre tais parâmetros, não encontramos um resultado igual ou melhor ao daqueles que utilizavam descritores MIE, já que a quantidade de moléculas classificadas erroneamente era bem maior para estas escolhas, além disso, a porcentagem do primeiro componente caiu consideravelmente com a combinação de parâmetros, diferentes daqueles analisados com MIE, escolhidos aleatoriamente.

A metodologia PCA foi realizada para os dois métodos AM1 e PM3. Os melhores resultados encontrados para a enzima Haspin estão descritos abaixo. Começamos com os resultados obtidos para o método AM1.

O resultado na tabela 3.25 para a variância, obtidos para a enzima Haspin, com o método AM1, distingue as moléculas em dois grupos, ativas e inativas, com a exceção da molécula 3. Essa separação envolve três regiões: B, F, P e W.

Tabela 3.25 – Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método AM1, para a enzima Haspin.

AM1 – PCA			
EIG	Variância	Percentual	Acumulativo
PC1	6.20092	99.8067	99.8067
PC2	0.00895	0.14403	99.9507
PC3	0.00268	0.04315	99.9939

As equações para as variáveis (loadings) PCA-AM1, são escritas para uma regra do tipo e , como segue:

$$PC1 = 0.75049(CHB) + 0.58430(CHF) + 0.18684(CHP) + 0.24588(CHW) \quad (3.1)$$

e

$$PC2 = -0.1550(CHB) - 0.2545(CHF) + 0.85173(CHP) + 0.43094(CHW) \quad (3.2)$$

No gráfico 3.33 podemos verificar o padrão fornecido pelo PCA, para a contribuição de cada variável nas componentes principais. As duas primeiras componentes, possuem juntas, 99,95% da variância total do dados, como mostrado na tabela 3.25.

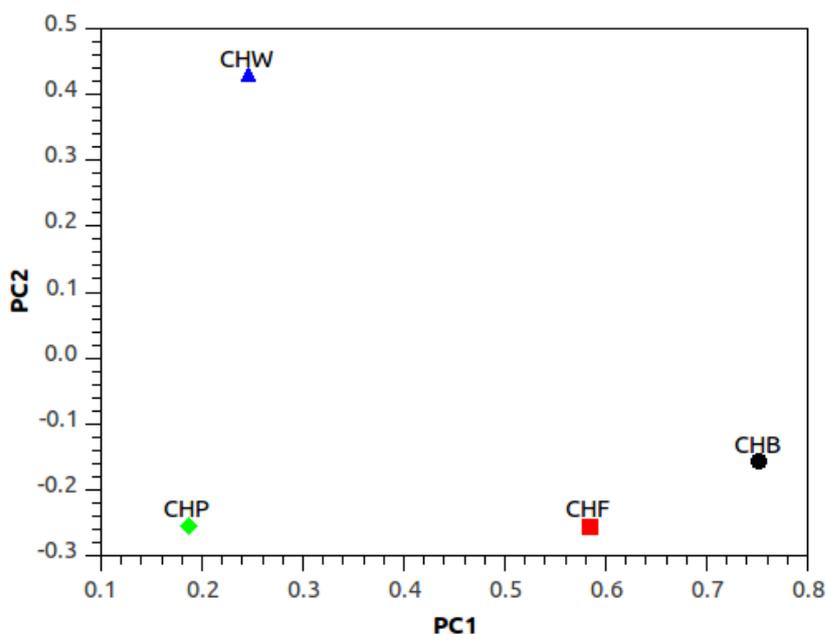


Figura 3.33 – Gráfico loadings da PCA, para enzima Haspin, com o método AM1, mostrando a distribuição dos parâmetros no plano PC1 X PC2.

No gráfico scores 3.34 e nas equações 3.13.2, verificamos que as variáveis CHB, CHF, CHP e CHW loadings em 3.33 foram suficientes para separar os compostos ativos. Podemos, a partir do resultado da PCA, obter uma relação parecida com a que foi construída para análise binária. A PCA, pode ser relacionada a partir de uma regra do tipo *e*:

Regra: Se $PC1 < 0.673$ e $PC2 < 0.9$, as moléculas são ativas.

ou

Regra: Se $PC1 > 0.673$ e $PC2 > 0.9$, as moléculas são inativas.

Observamos através do gráfico 3.34 que, com essa regra, os compostos são classificados corretamente sem exceções. Obtemos, então com a PCA 100% de acerto na classificação dos compostos. É importante considerarmos que dois dos parâmetros responsáveis pelo resultado na análise PCA, foram identificados pela MIE, tornando notória essa metodologia.

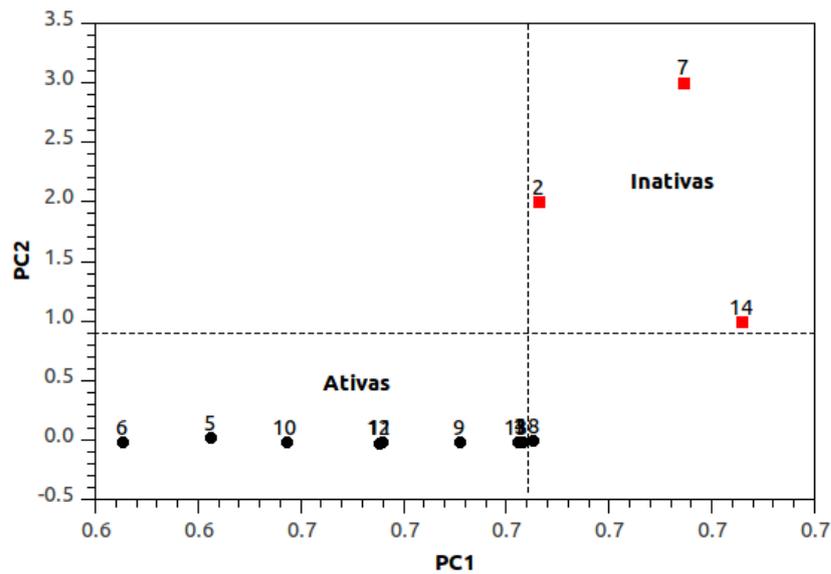


Figura 3.34 – Gráfico scores da PCA, para enzima Haspin, com o método AM1, mostrando a distribuição das moléculas no plano PC1 X PC2.

Os resultados obtidos para a separação das moléculas ativas e inativas para a enzima Haspin, utilizando o método PM3, envolve as regiões B, P e V. Os valores para a variância para este caso, são apresentados na tabela 3.26.

Tabela 3.26 – Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método PM3, para a enzima Haspin.

PM3 – PCA			
EIG	Variância	Percentual	Acumulativo
PC1	17.3040	99.9022	99.9022
PC2	0.01550	0.08950	99.9917
PC3	0.00142	0.00822	100.000

As equações para as variáveis (loadings) PCA-PM3, são escritas para uma regra do tipo e , como segue:

$$PC1 = 0.85798(CLB) + 0.15010(CHP) + 0.49126(CLV) \quad (3.3)$$

e

$$PC2 = -0.2499(CL B) + 0.95750(CHP) + 0.14397(CLV) \quad (3.4)$$

No gráfico 3.35 podemos verificar o padrão fornecido pelo PCA, para a contribuição de cada variável nas componentes principais. As duas primeiras componentes, possuem juntas, 99,99% da variância total do dados, como mostrado na tabela 3.26.

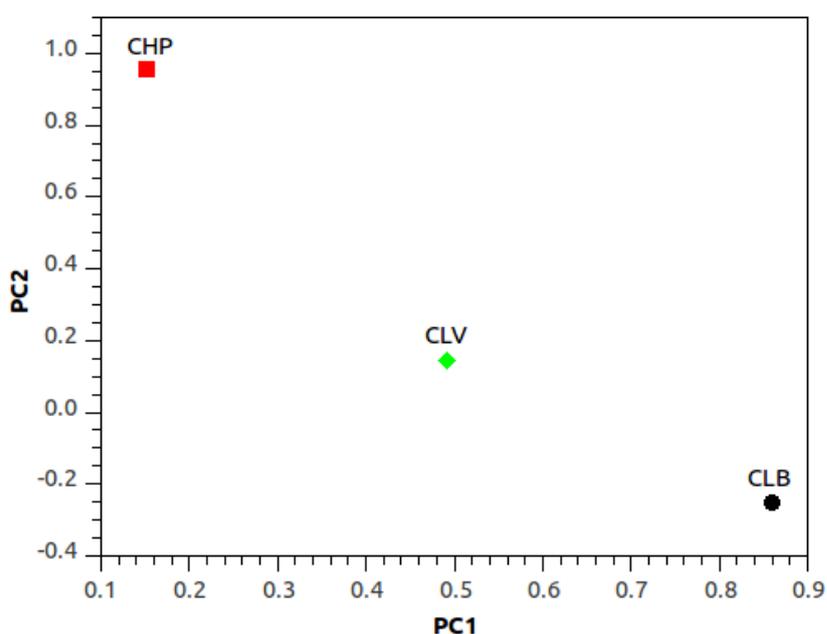


Figura 3.35 – Gráfico loadings da PCA, para enzima Haspin, com o método PM3, mostrando a distribuição dos parâmetros no plano PC1 X PC2.

No gráfico scores 3.36 e nas equações 3.3 3.4, verificamos que as variáveis CLB, CHP e CLV, loadings em 3.35 foram suficientes para separar os compostos ativos. Podemos, a partir do resultado da PCA, obter uma relação parecida com a que foi construída para análise binária. A PCA, pode ser relacionada a partir de uma regra do tipo *e*:

Regra: Se $PC1 < 1.137$ e $PC2 < 0.4$, as moléculas são ativas.

ou

Regra: Se $PC1 > 1.137$ e $PC2 > 0.4$, as moléculas são inativas.

Observamos através do gráfico 3.36 que, com essa regra, os compostos são classificados corretamente sem exceções. Obtemos, então com a PCA 100% de acerto na classificação dos compostos, também com método PM3. É importante considerarmos o parâmetro CHP, um dos responsáveis pelo resultado na análise PCA, foi um grande contribuinte na MIE, evidenciando a importância e validade de aplicar essa metodologia.

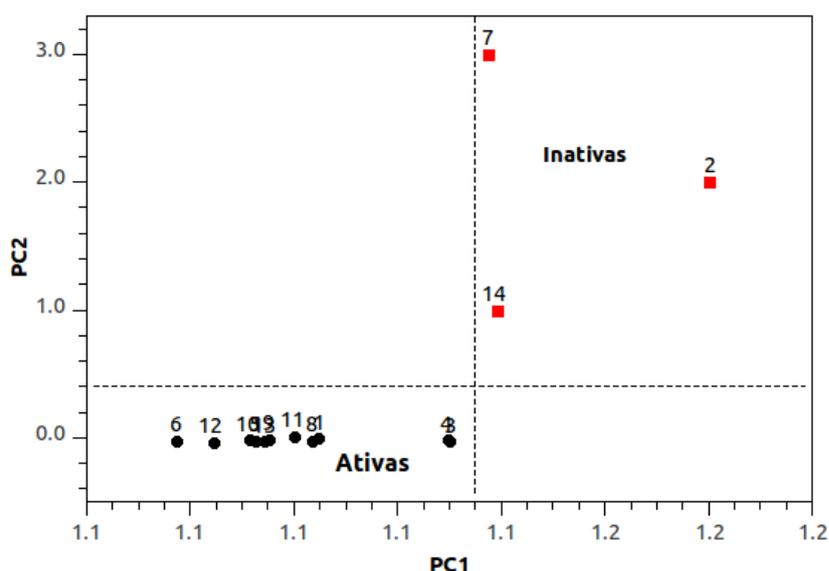


Figura 3.36 – Gráfico scores da PCA, para enzima Haspin, com o método PM3, mostrando a distribuição das moléculas no plano PC1 X PC2.

3.2.4 Estudo PCA para a enzima DYRK2

Os resultados obtidos para a separação das moléculas ativas e inativas para a enzima DYRK2, utilizando o método AM1, envolve apenas duas regiões, V e W, e o orbital HOMO. Os valores para a variância para este caso, são apresentados na tabela 3.27.

Tabela 3.27 – Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método AM1, para a enzima DYRK2.

AM1 – PCA			
EIG	Variância	Percentual	Acumulativo
PC1	963.336	99.9991	99.9991
PC2	0.00689	0.00072	99.9998
PC3	0.00165	0.00017	100.000

As equações para as variáveis (loadings) PCA-AM1, para a enzima DYRK2, são escritas para uma regra do tipo e , como segue:

$$PC1 = 0.99899(H) - 0.0404(CHV) - 0.0197(CHW) \quad (3.5)$$

e

$$PC2 = 0.02022(H) + 0.79463(CHV) - 0.6067(CHW) \quad (3.6)$$

No gráfico 3.37 podemos verificar o padrão fornecido pelo PCA, para a contribuição de cada variável nas componentes principais. As duas primeiras componentes, possuem juntas, 99,9998% da variância total do dados, como mostrado na tabela 3.27.

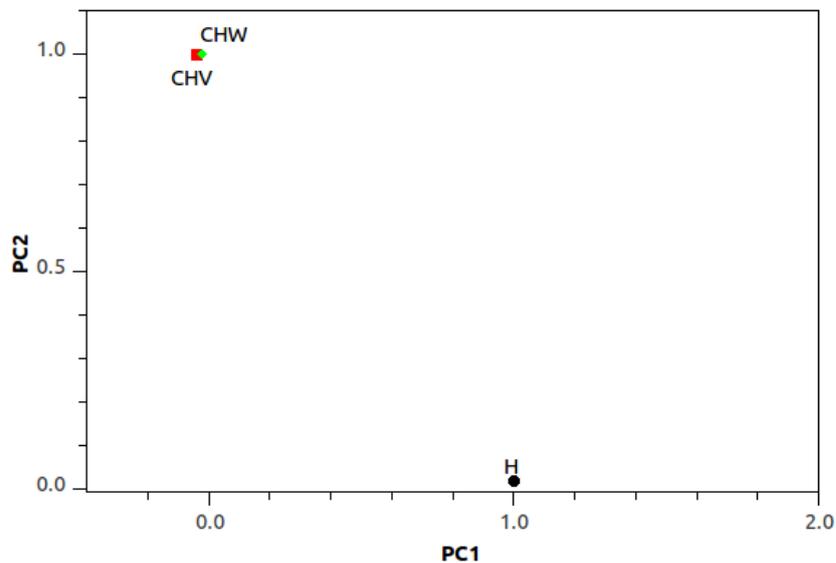


Figura 3.37 – Gráfico loadings da PCA, para enzima DYRK2, com o método AM1, mostrando a distribuição dos parâmetros no plano PC1 X PC2.

No gráfico scores 3.38 e nas equações 3.5 3.6, verificamos que as variáveis H, CHV e CHW, loadings em 3.37 foram suficientes para separar os compostos ativos. Podemos, a partir do resultado da PCA, obter uma relação parecida com a que foi construída para análise binária. A PCA, pode ser relacionada a partir de uma regra do tipo e :

Regra: Se $PC1 > -8.39$ e $PC2 > -0.012$, as moléculas são ativas.

ou

Regra: Se $PC1 < -8.39$ e $PC2 < -0.012$, as moléculas são inativas.

Observamos através do gráfico 3.38 que, com essa regra, os compostos são classificados corretamente sem exceções. Obtemos, então com a PCA 100% de acerto na classificação dos compostos, com método AM1. Notamos que, para a enzima DYRK2, todos os parâmetros (H, CHV e CHW) envolvidos na análise PCA, foram importantes na análise binária com a metodologia de índices eletrônicos. A MIE mostra-se mais uma vez, uma metodologia eficiente e confiável para classificar os compostos quanto a atividade biológica.

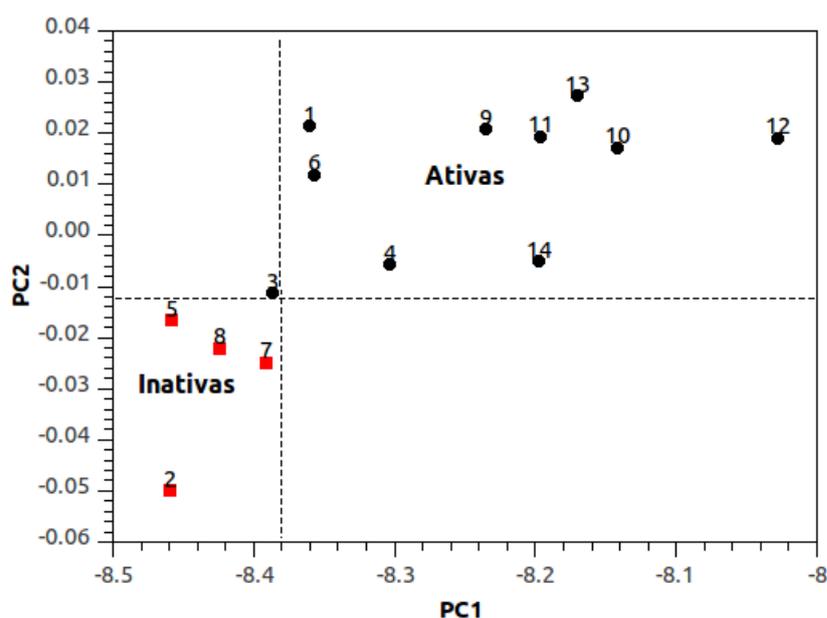


Figura 3.38 – Gráfico scores da PCA, para enzima DYRK2, com o método AM1, mostrando a distribuição das moléculas no plano PC1 X PC2.

Os resultados obtidos para a separação das moléculas ativas e inativas para a enzima DYRK2, utilizando o método PM3, envolve as regiões, A, F e W; e o orbital HOMO–1. Os valores para a variância para este caso, são apresentados na tabela 3.28.

As equações para as variáveis (loadings) PCA-PM3, para a enzima DYRK2, são escritas

Tabela 3.28 – Tabela com valores da variância para as componentes PC1, PC2 e PC3, pelo método PM3, para a enzima DYRK2.

PM3 – PCA			
EIG	Variância	Percentual	Acumulativo
PC1	1131.46	99.9984	99.9984
PC2	0.00892	0.00079	99.9992
PC3	0.00585	0.00052	99.9997

para uma regra do tipo e , como segue:

$$PC1 = 0.99356(H1) - 0.0602(CHA) - 0.0821(CLF) - 0.0203(CHW) - 0.0453(CLW) \quad (3.7)$$

e

$$PC2 = 0.04175(H1) + 0.44802(CHA) + 0.03262(CLF) - 0.6880(CHW) + 0.56843(CLW) \quad (3.8)$$

No gráfico 3.39 podemos verificar o padrão fornecido pelo PCA, para a contribuição de cada variável nas componentes principais. As duas primeiras componentes, possuem juntas, 99,99% da variância total do dados, como mostrado na tabela 3.28.

No gráfico scores 3.39 e nas equações 3.7 3.8, verificamos que as variáveis H1, CHA, CLF, CHW e CLW, loadings em 3.39 foram suficientes para separar os compostos ativos. Podemos, a partir do resultado da PCA, obter uma relação parecida com a que foi construída para análise binária. A PCA-PM3, para a enzima DYRK2, pode ser relacionada a partir de uma regra do tipo e :

Regra: Se $PC1 > -9.1$ e $PC2 > -0.02$, as moléculas são ativas.

ou

Regra: Se $PC1 < -9.1$ e $PC2 < -0.02$, as moléculas são inativas.

Observamos através do gráfico 3.40 que, com essa regra, os compostos são classifica-

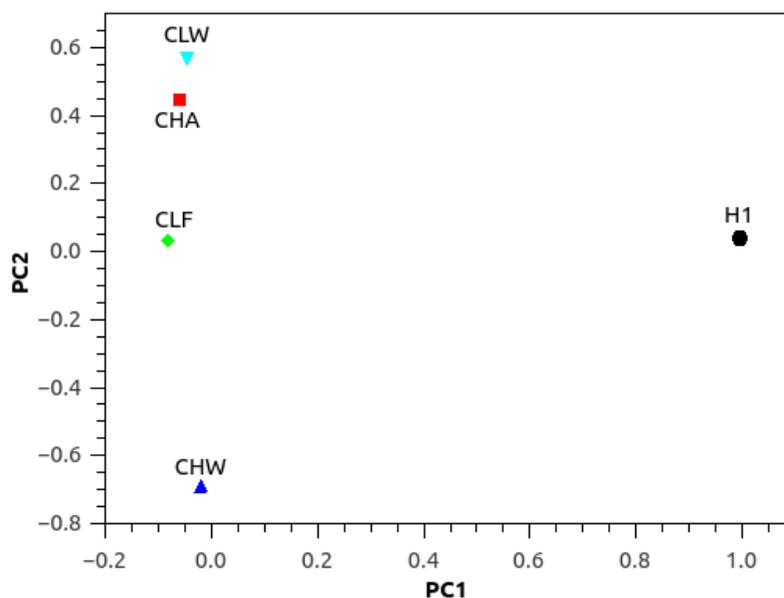


Figura 3.39 – Gráfico loadings da PCA, para enzima DYRK2, com o método PM3, mostrando a distribuição dos parâmetros no plano PC1 X PC2.

dos corretamente sem exceções. Obtemos, então com a PCA 100% de acerto na classificação dos compostos, com método PM3. Notamos que, para a enzima DYRK2, somente dois dos parâmetros (H1, CHA, CLF, CHW e CLW) envolvidos na análise PCA, foram importantes na análise binária com a metodologia de índices eletrônicos, mais especificamente os parâmetros CLW e o orbital HOMO-1. Outros resultados para a PCA foram encontrados com os parâmetros (CHB, CHV e CHW outro com CLP, η LP e η LW ou ainda H, CHV e CHW), que foram utilizados na MIE, eles apresentaram também uma boa separação das moléculas ativas na análise PCA. No entanto, as variáveis selecionadas para exposição do resultado PCA para este caso, foram as que melhor aproximaram as moléculas do grupo de ativas, como pode ser visto no gráfico scores 3.40.

Pelo estudo do HCA, temos

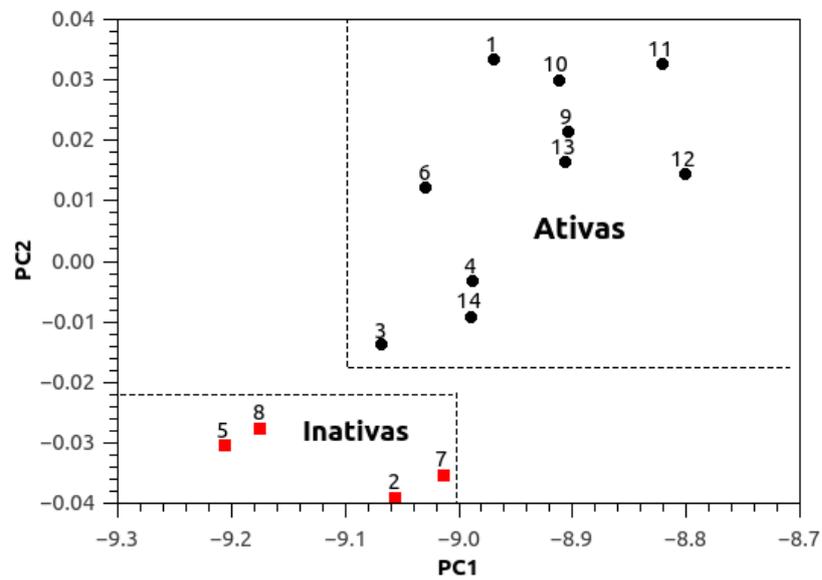


Figura 3.40 – Gráfico scores da PCA, para enzima DYRK2, com o método PM3, mostrando a distribuição das moléculas no plano PC1 X PC2.

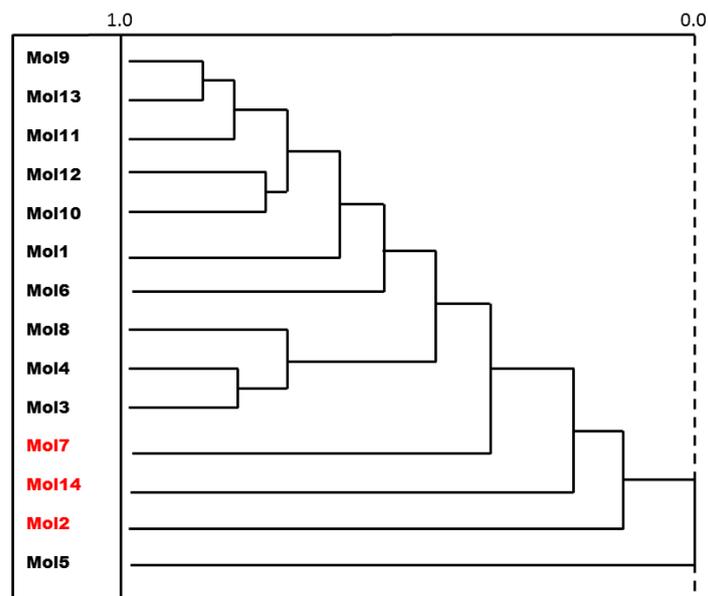


Figura 3.41 – Gráfico com o resultado para HCA da enzima Haspin, pelo método AM1, especificando as moléculas inativas pela cor vermelha e as ativas, preto.

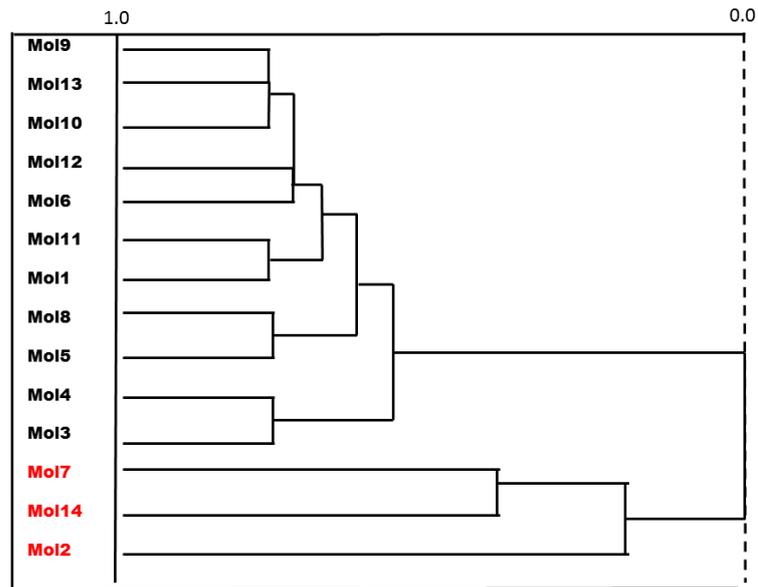


Figura 3.42 – Gráfico com o resultado para HCA da enzima Haspin, pelo método PM3, especificando as moléculas inativas pela cor vermelha e as ativas, preto.

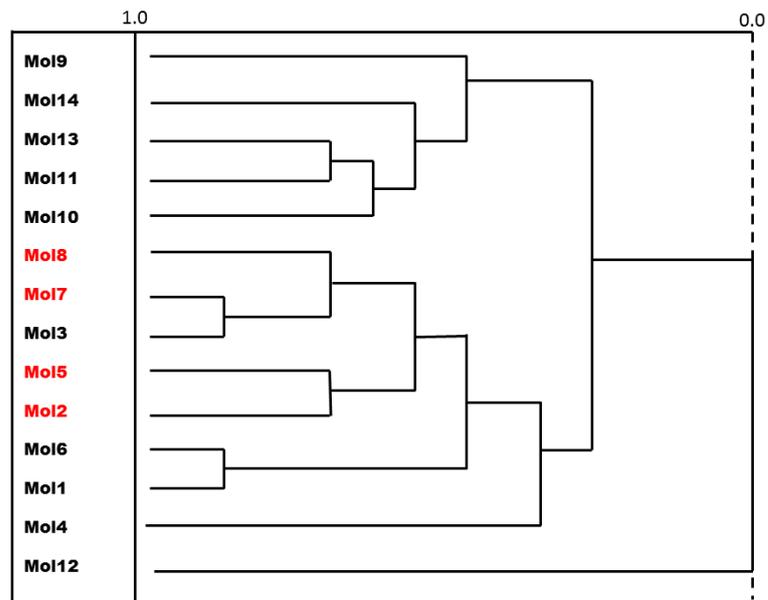


Figura 3.43 – Gráfico com o resultado para HCA da enzima DYRK2, pelo método AM1, especificando as moléculas inativas pela cor vermelha e as ativas, preto.

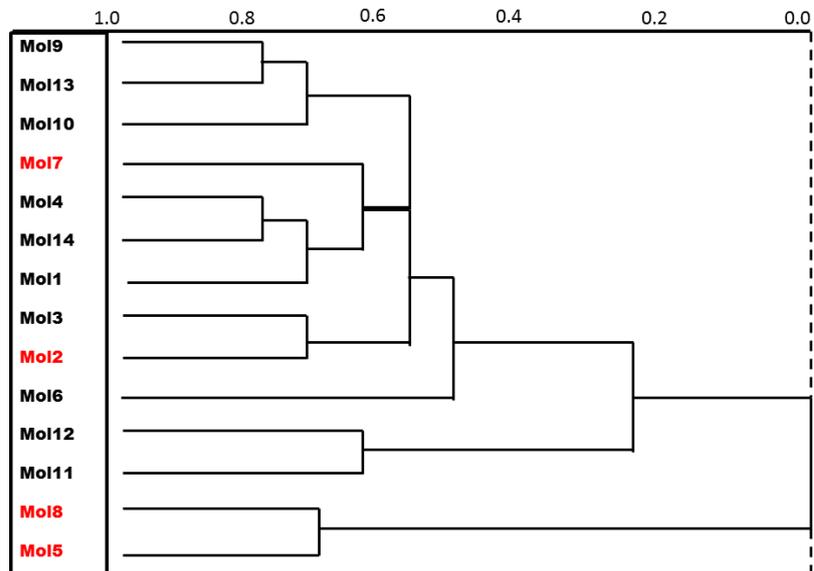


Figura 3.44 – Gráfico com o resultado para HCA da enzima DYRK2, pelo método PM3, especificando as moléculas inativas pela cor vermelha e as ativas, preto.

4 CONCLUSÃO

No presente trabalho, realizamos uma busca sistemática pela geometria molecular de quatorze compostos moleculares, derivados da acridina, que investigados experimentalmente, apresentavam ação biológica inibidora de enzimas, responsáveis pela conclusão do processo de multiplicação celular de células cancerígenas. Percebemos que os valores para as energias eram menores para o método PM3 que para o AM1, as geometrias moleculares desses compostos eram mais estáveis para este método. Percebemos que as estruturas fornecidas pelos métodos PM3 e DFT, após otimização de geometria molecular, bem como os valores para as energias dos orbitais de fronteira e suas respectivas diferenças, foram idênticas. Continuamos o estudo a partir daqui somente com os métodos AM1 e PM3.

Na análise dos parâmetros eletrônicos evidenciamos a eficácia da MIE na classificação dos compostos análogos da acridina, quanto a sua atividade biológica, utilizando apenas dois descritores quânticos moleculares que consideram a importância das propriedades eletrônicas locais. Na análise feita para uma variável a enzima Haspin, apenas a análise para uma variável pelo método AM1, apresentou 92,86% de acerto. As regiões P e W demonstraram-se mais importantes que as demais para essa análise, apresentando poucos erros na predição da atividade. Nos estudos restantes para essa enzima, com duas variáveis para o método AM1, para o qual as regiões de maior importância foram C, V e W e os orbitais HOMO e ΔL , no entanto as contribuições para os orbitais HOMO e LUMO+1 para a região W apresentou o melhor resultado na separação das moléculas ativas e inativas. O estudo com uma variável para o método PM3, mostrou que a contribuição para o orbital HOMO na região P foi suficiente para discriminar as moléculas com acerto de 100% na predição. O estudo com duas variáveis para o método

PM3, identificou as regiões W, K, V e R, e os orbitais LUMO, HOMO e LUMO+1 como as mais eficientes para separar os compostos, a MIE identificou a contribuição CHK e o orbital LUMO+1 como os parâmetros que separa com precisão de 100% moléculas ativas e inativas.

Para a enzima DYRK2, a análise pelo método AM1 com uma variável, as regiões de maior importância na separação do grupo de moléculas ativas e inativas foram as regiões C, K, P, S e V, e os orbitais atômicos LUMO, HOMO e HOMO-1, o parâmetro escolhido, HOMO, foi o que demonstrou o melhor resultado entre os parâmetros encontrados, 92,86%. Para a análise binária, este mesmo método, apresenta as regiões P e K como as mais eficientes na discriminação dos grupos, assim como o orbital LUMO e a diferença ΔH . Escolhemos a contribuição HOMO na região K e o orbital atômico L, como os melhores parâmetros para essa separação, o acerto na predição pela MIE, é de 92,86%. Já para o método PM3, o acerto em ambas as análises, para essa enzima, foi de 100%. Para uma variável as regiões importantes foram A, C, P, V e W, e os orbitais HOMO e HOMO-1, observamos que a contribuição para o orbital LUMO na região W, foi a que exibiu melhor resultado. Para duas variáveis, identificamos as regiões C, D, F, K, P e V, e os orbitais HOMO, HOMO-1, LUMO, e a diferença ΔH , como os melhores parâmetros na identificação de padrão dos compostos. Escolhemos a contribuição para o orbital HOMO na região P e o orbital HOMO-1, para estabelecermos uma regra, já que estes apresentaram o maior acerto.

Nas metodologias de reconhecimento de padrões, PCA e HCA, os melhores resultados implicaram na utilização dos parâmetros da MIE, mostrando mais uma vez a importância da metodologia de índices eletrônicos no estabelecimento de padrões de atividade biológica.

A partir dos parâmetros MIE, que direcionaram a análise de componentes principais, os resultados apresentados pela PCA exibiram 100% de acerto, ao utilizarmos regras do tipo *e* para todas as análises. Os resultados apresentados pela MIE apontam duas as regiões W, P e V, como indispensáveis para classificação dos compostos, as contribuições para essas regiões são as mais presentes nas regras que relacionam os compostos em classes de ativos e inativos. A análise de componentes principais para a enzima Haspin, pelo método AM1, envolveu as contribuições

para o orbital HOMO nas regiões B,F, P e W. Pelo método PM3, as contribuições para o orbital LUMO nas regiões B e V, e para o orbital HOMO na região P, foram os que melhor diferenciou as moléculas ativas e inativas. Para a enzima DYRK2 os resultados para a PCA pelo método AM1 envolveu o orbital HOMO, e as contribuições para este orbital nas regiões V e W. E para o método PM3 os parâmetros envolvidos no padrão foram o orbital HOMO-1, as contribuições para o orbital HOMO nas regiões A e W, e as contribuições para o orbital LUMO nas regiões F e W.

As análises com as metodologias, PCA e HCA, mostraram a eficiência dos parâmetros MIE na separação dos compostos. A HCA, apesar de não ser um método de classificação, nos fornece uma análise dos resultados MIE, obtidos para os compostos estudados, uma vez que nos fornece uma similaridade global. Os agrupamentos hierárquicos obtidos foram regidos pelas componentes principais PC1s, para todas as análises.

Levando em consideração o tempo destinado para encontrar a combinação de descritores que melhor classifica os compostos em ativos e inativos, na metodologia PCA, a MIE possui ainda mais uma vantagem, a rapidez com que os padrões para atividade dos compostos, são encontrados. Para o estudo realizado com os análogos da acridina, a MIE descreveu com alta precisão a atividade biológica dos compostos, mostrando-se uma metodologia simples, contudo primorosa na discriminação dos compostos ativos e inativos.

4.1 PERSPECTIVAS FUTURAS

Pretendemos dar continuidade a investigação de compostos análogos da acridina, que apresentem radicais maiores e diferentes ligados a estrutura principal, ou compostos que tenham características, assim como os compostos estudados, de inibir quinases indispensáveis na conclusão da mitose de células cancerígenas, e o comportamento desses compostos ao intercalarem o DNA. Investigar parâmetros eletrônicos, utilizando a MIE, de compostos análogos da acridina, que sejam utilizados no desenvolvimento de fármacos destinados ao tratamento de neoplasias.

APÊNDICE A – O MÉTODO HARTREE-FOCK

O objetivo principal da química quântica é encontrar soluções para a equação de Schrödinger para descrever as propriedades de sistemas atômicos e moleculares. Uma vez que a solução exata da equação de Schrödinger não é factível, é imprescindível o uso de métodos aproximados, dentre os quais o mais popular é o método de Hartree-Fock (HF). Este método fornece uma boa solução aproximada para o problema de muitos elétrons e além disso, serve como base para métodos semi-empíricos, nos quais outras aproximações são incluídas com a finalidade de redução de tempo computacional, e também para métodos em que acrescenta-se correções à aproximação HF para aprimorá-la e aproximá-la ainda mais da solução exata [31].

Considerando a aproximação de Born-Oppenheimer e desprezando os efeitos relativísticos, o problema resume-se em resolver a equação de Schrödinger eletrônica:

$$H\Phi = E\Phi, \quad (\text{A.1})$$

onde a parte eletrônica do operador Hamiltoniano molecular é dada por:

$$H = -\sum_{i=1}^N \frac{\hbar^2}{2m} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A e^2}{4\pi\epsilon_0 r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{e^2}{4\pi\epsilon_0 r_{ij}} \quad (\text{A.2})$$

onde

- m é a massa do elétron,

- Z_A é o número atômico do núcleo A ,

- r_{ij} é a distância entre os elétrons i e j ,

$-r_{iA}$ é a distância entre o elétron i e o núcleo A ,

$-N$ e M representam, respectivamente, os números de elétrons e núcleos do sistema.

O Hamiltoniano eletrônico pode ser reescrito, usando o sistema de unidades atômicas:

$$H = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^n \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^n \sum_{j>i}^N \frac{1}{r_{ij}}, \quad (\text{A.3})$$

nessa representação, a distância é dada em Bohr e a energia em Hartree.

A.1 A Teoria do Orbital Molecular

A teoria do orbital molecular foi desenvolvida depois do advento da mecânica quântica. A formulação matemática para o problema de muitos corpos, a partir teoria dos orbitais, foi exposta por Hartree, Slater, Roothaan (que na época era um aluno de Mulliken), Löwdin e outros colaboradores, levando ao prêmio Nobel em 1966, concedido a Robert S. Mulliken [32].

Esta abordagem consiste em descrever a função de onda eletrônica em termos das funções de onda para elétrons independentes, os orbitais moleculares (O.M.) $\varphi(r^\mu) = \varphi_i(x^\mu, y^\mu, z^\mu)$ onde x^μ, y^μ, z^μ são as coordenadas do elétron μ e i diz respeito ao estado eletrônico. Se o spin for considerado, usa-se o spin-orbital molecular (S.O.M.):

$$\psi_i(q^\mu) = \varphi_i(x^\mu, y^\mu, z^\mu) \xi(S^\mu) = \varphi_i(\mu) \xi(\mu) = \varphi_i^\mu \xi^\mu, \quad (\text{A.4})$$

onde

$$\xi(S^\mu) = \begin{cases} \alpha(\mu) \uparrow \\ \beta(\mu) \downarrow. \end{cases} \quad (\text{A.5})$$

A função de estado total para o sistema de N elétrons satisfaz ao princípio da exclusão

de Pauli e pode ser escrita como um produto anti-simétrico dos spin-orbitais:

$$\Phi = (N!)^{\frac{1}{2}} \hat{A} \{ \psi_1^{(1)} \psi_2^{(2)} \dots \psi_N^{(N)} \}, \quad (\text{A.6})$$

onde Φ é a função de estado total para N elétrons e \hat{A} é o operador antisimetrizador. Tomou-se $\Phi_{a,b,\dots,k}(q^1, q^2, \dots, q^N)$ por Φ e cada estado $\psi_j(q^\mu)$ por ψ_j^μ . A relação A.6 pode ser escrita na forma de determinante, conhecido como o determinante de Slater:

$$\Phi = (N!)^{-\frac{1}{2}} \begin{bmatrix} \psi_1^{(1)} & \psi_2^{(1)} & \dots & \psi_N^{(1)} \\ \psi_1^{(2)} & \psi_2^{(2)} & \dots & \psi_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{(N)} & \psi_2^{(N)} & \dots & \psi_N^{(N)} \end{bmatrix}. \quad (\text{A.7})$$

em que

- ψ 's são funções das coordenadas espaciais e spin de um único elétron, os S.O.M.

que têm como característica,

- (a) Os S.O.M são linearmente independentes,
- (b) O princípio da exclusão de Pauli é satisfeito automaticamente.

Admitindo que os ψ 's sejam ortonormais, o fator $(N!)^{-\frac{1}{2}}$ é uma constante de normalização para Φ_0 . O determinante único faz uma simplificação no sentido de substituir o problema de encontrar uma função de onda que depende das coordenadas de N elétrons, pelo problema de encontrar N funções de onda de um elétron (S.O.M.). Trocar coordenadas de dois elétrons significa trocar duas linhas do determinante, o que implica numa inversão de sinais. A função de onda Hartree-Fock (HF), definida em termos de um único determinante de Slater, é antissimétrica diante da troca das coordenadas de dois elétrons. Como um determinante com dois S.O.M iguais possui duas colunas iguais, e portanto, é nulo, satisfazendo o princípio da exclusão de Pauli, que afirma que dois elétrons não podem ocupar o mesmo estado quântico (mesmo S.O.) [31].

Os S.O.M. ψ_k são linearmente independentes, isto é:

$$\int \psi_k^*(\mu) \psi_\ell(\mu) d\tau_\mu = \delta_{k\ell}. \quad (\text{A.8})$$

Conseqüentemente, teremos

$$\int \Phi^* \Phi d\tau = 1. \quad (\text{A.9})$$

Então a equação A.9 torna-se:

$$\int \Phi^* \Phi d\tau = \int \psi_1^{(1)\dagger} \psi_1^{(1)} d\tau_1 \int \dots \int \psi_N^{(N)\dagger} \psi_N^{(N)} d\tau_N = 1. \quad (\text{A.10})$$

A energia do sistema é dada por:

$$E = \int \Phi^* \hat{H} \Phi d\tau \equiv E[\Phi], \quad (\text{A.11})$$

onde o operador Hamiltoniano em unidades atômicas é definido por:

$$\hat{H} = \sum_{\mu} \hat{h}(\mu) + \frac{1}{2} \sum_{\mu \neq \nu} \frac{1}{r^{\mu\nu}}. \quad (\text{A.12})$$

Em A.12, $\hat{h}(\mu)$ é o operador hamiltoniano de uma partícula para o μ -ésimo elétron movendo-se no campo dos núcleos; $\hat{h}(\mu)$ é um operador hermitiano e linear expresso por

$$\hat{h}(\mu) = -\frac{\nabla_{\mu}^2}{2} - \sum_A \frac{1}{r^{\mu A}} \quad (\text{A.13})$$

com $\frac{\nabla_{\mu}^2}{2}$ sendo a energia cinética do elétron μ , e $\sum_A \frac{1}{r^{\mu A}}$ a energia coulombiana de interação elétron-núcleo.

Os determinantes de Slater são a essência do método Hartree-Fock [31]. Se nenhuma restrição for feita aos orbitais moleculares, o método é denominado Hartree-Fock não Restrito

(UHF), sendo vantajoso para sistemas de camada aberta. Sistemas de camada aberta podem ser descritos também por uma função de onda do tipo restrita, onde as partes espaciais dos spins-orbitais, duplamente ocupados, sejam obrigatoriamente as mesmas. Nestas condições, tem-se o método Hartree-Fock Restrito de Camada Aberta (ROHF). Para sistemas em que o estado eletrônico for de camada fechada, a restrição na qual cada orbital espacial seja associado a dois elétrons, um com spin α e outro com spin β , é chamado de Hartree-Fock Restrito [31, 52].

A função de onda HF escrita na forma A.7 é normalizada. Supomos então que os spins-orbitais sejam ortonormais, como em A.8, e usando a notação de Dirac, tem-se:

$$\langle \Phi_0 | \Phi_0 \rangle = \frac{1}{N!} \sum_{i=1}^{N!} \sum_{j=1}^{N!} (-1)^{p_i} (-1)^{p_j} \times \langle \mathcal{P}_i \psi_1(1) \psi_2(2) \dots \psi_N(N) | \mathcal{P}_j \psi_1(1) \psi_2(2) \dots \psi_N(N) \rangle \quad (\text{A.14})$$

onde $\psi_1(1) \equiv \psi_1^{(1)}$. As integrais podem ser separadas em integrais do tipo $\langle \psi_a(k) | \psi_b(k) \rangle$. Estas são nulas quando $a \neq b$. Assim, restam somente os termos $i = j$ no somatório duplo, e todas as integrais são iguais a um, resultando em:

$$\langle \Phi_0 | \Phi_0 \rangle = \frac{1}{N!} \sum_{i=1}^{N!} 1 = 1 \quad (\text{A.15})$$

O hamiltoniano eletrônico do sistema (H) pode ser separado em uma parte que só contém operadores de um elétron e outra, com os operadores de dois elétrons:

$$H = O_1 + O_2 \quad (\text{A.16})$$

onde

$$O_1 = \sum_{i=1}^N h(i), \quad (\text{A.17})$$

com

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{1}{r_{iA}} \quad (\text{A.18})$$

e

$$O_2 = \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}. \quad (\text{A.19})$$

Consideramos agora o cálculo de

$$\langle \Phi_0 | O_1 | \Phi_0 \rangle = \langle \Phi_0 | h(1) + \dots + h(N) | \Phi_0 \rangle. \quad (\text{A.20})$$

Como os elétrons são indistinguíveis, temos:

$$\langle \Phi_0 | h(1) | \Phi_0 \rangle = \langle \Phi_0 | h(2) | \Phi_0 \rangle = \dots = \langle \Phi_0 | h(N) | \Phi_0 \rangle \quad (\text{A.21})$$

de forma que

$$\begin{aligned} \langle \Phi_0 | O_1 | \Phi_0 \rangle &= N \langle \Phi_0 | h(1) | \Phi_0 \rangle \\ &= N \times \frac{1}{N!} \sum_{i=1}^{N!} \sum_{j=1}^{N!} (-1)^{p_i} (-1)^{p_j} \\ &\quad \times \langle \mathcal{P}_i \psi_1(1) \psi_2(2) \dots \psi_N(N) | h(1) | \mathcal{P}_j \psi_1(1) \psi_2(2) \dots \psi_N(N) \rangle. \end{aligned} \quad (\text{A.22})$$

A integração sobre as coordenadas dos elétrons 2,3...N só será diferente de zero se cada um desses elétrons ocupar o mesmo spin orbital nas permutações \mathcal{P}_i e \mathcal{P}_j . Mas se os elétrons 2,3,...N ocuparem os mesmos spin-orbitais em ambas as permutações, o elétron 1 também terá que ocupar o mesmo spin-orbital. Assim as permutações \mathcal{P}_i e \mathcal{P}_j serão sempre iguais, e portanto,

$$\begin{aligned} \langle \Phi_0 | O_1 | \Phi_0 \rangle &= \frac{1}{(N-1)!} \sum_{i=1}^{N!} \\ &\quad \times \langle \mathcal{P}_i \psi_1(1) \psi_2(2) \dots \psi_N(N) | h(1) | \mathcal{P}_i \psi_1(1) \psi_2(2) \dots \psi_N(N) \rangle. \end{aligned} \quad (\text{A.23})$$

Considerando todas as permutações, o elétron 1 ocupa cada spin-orbital a , $[(N-1)!]$ vezes, com as integrais sobre os elétrons 2,3,...N iguais a 1. Cada spin orbital a vai de $a = 1, 2, \dots, N$

Assim

$$\langle \Phi_0 | O_1 | \Phi_0 \rangle = (N-1)! \times \frac{1}{(N-1)!} \sum_{a=1}^N N \langle \psi_a(1) | h(1) | \psi_a(1) \rangle, \quad (\text{A.24})$$

que pode ser escrita como

$$\langle \Phi_0 | O_1 | \Phi_0 \rangle = \sum_{a=1}^N \langle a | h | a \rangle, \quad (\text{A.25})$$

usando a notação

$$\langle a | h | a \rangle = \langle \psi_a | h | \psi_a \rangle = \langle \psi_a(1) | h(1) | \psi_a(1) \rangle. \quad (\text{A.26})$$

A integral de dois elétrons é escrita como

$$\langle \Phi_0 | O_2 | \Phi_0 \rangle = \left\langle \Phi_0 \left| \frac{1}{r_{12}} + \frac{1}{r_{13}} + \dots \right| \Phi_0 \right\rangle. \quad (\text{A.27})$$

Como os elétrons são indistinguíveis,

$$\left\langle \Phi_0 \left| \frac{1}{r_{12}} \right| \Phi_0 \right\rangle = \left\langle \Phi_0 \left| \frac{1}{r_{13}} \right| \Phi_0 \right\rangle, \quad (\text{A.28})$$

de forma que

$$\langle \Phi_0 | O_2 | \Phi_0 \rangle = \frac{N(N-1)}{2} \left\langle \Phi_0 \left| \frac{1}{r_{12}} \right| \Phi_0 \right\rangle \quad (\text{A.29})$$

onde $\frac{N(N-2)}{2}$ é o número de pares de elétrons que aparecem em A.27. Logo,

$$\begin{aligned} \langle \Phi_0 | O_2 | \Phi_0 \rangle &= \frac{N(N-1)}{2} \\ &\times \frac{1}{N!} \sum_{i=1}^{N!} \sum_{j=1}^{N!} (-1)^{pi} (-1)^{pj} \\ &\times \langle \mathcal{P}_i \psi_1(1) \psi_2(2) \dots \psi_N(N) | \frac{1}{r_{12}} | \mathcal{P}_j \psi_1(1) \psi_2(2) \dots \psi_N(N) \rangle. \end{aligned} \quad (\text{A.30})$$

Fazendo o mesmo procedimento que foi feito para o caso de um elétron, temos:

$$\langle \Phi_0 | O_2 | \Phi_0 \rangle = \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N (\langle ab | ab \rangle - \langle ab | ba \rangle). \quad (\text{A.31})$$

A restrição de que $a \neq b$ não é necessária, pois o termo $a = b$ no duplo somatório é nulo.

Definindo

$$\langle ab || ab \rangle = \langle ab | ab \rangle - \langle ab | ba \rangle \quad (\text{A.32})$$

podemos escrever

$$\langle \Phi_0 | O_2 | \Phi_0 \rangle = \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N \langle ab || ab \rangle. \quad (\text{A.33})$$

Assim podemos escrever o valor esperado do operador H como

$$E = \langle \Phi_0 | H | \Phi_0 \rangle = \sum_a \langle a | h | a \rangle + \frac{1}{2} \sum_{a,b} \langle ab || ab \rangle. \quad (\text{A.34})$$

A.2 Equação de Hartree-Fock

A ideia do método de Hartree-Fock é utilizar o princípio variacional, supondo que o determinante de Slater é a função de onda que descreve o sistema molecular. De acordo com

o princípio variacional, dada uma função de onda que satisfaz as condições de contorno do problema em questão, o valor esperado do Hamiltoniano é um limite superior para a energia exata do estado fundamental, ou

$$\langle \Phi | H | \Phi \rangle \geq E_{exata} \quad (\text{A.35})$$

Se Φ_{exata} for a solução exata para o Hamiltoniano H, teremos

$$\langle \Phi_{exata} | H | \Phi_{exata} \rangle = E_{exata} \quad (\text{A.36})$$

Como a função de onda que mais se aproxima da solução exata é a que conduz a um mínimo, o valor médio do operador Hamiltoniano, que neste caso é um funcional dos S.O.M., pode ser minimizado,

$$E(\psi) = \sum_a \langle \psi_a | h | \psi_a \rangle + \frac{1}{2} \sum_{a,b} (\langle \psi_a \psi_b | \psi_a \psi_b \rangle - \langle \psi_a \psi_b | \psi_b \psi_a \rangle), \quad (\text{A.37})$$

Impondo a restrição de que os S.O.M. sejam ortonormais,

$$\langle \psi_a | \psi_b \rangle - \delta_{ab} = 0. \quad (\text{A.38})$$

Usando a técnica de multiplicadores de Lagrange para a condição de vínculo,

$$L[\psi] = E[\psi] - \sum_{a,b} \epsilon_{ba} (\langle \psi_a | \psi_b \rangle - \delta_{ab}) \quad (\text{A.39})$$

onde os ϵ_{ab} são os multiplicadores de Lagrange. Impomos as condições:

-L é real,

-E é real,

- a e b assumem os mesmos valores podendo ser trocados no duplo somatório.

Se cada S.O.M. sofrer uma variação $\delta\psi$, então L também variará. Podemos escrever esta

variação como:

$$\begin{aligned} \delta L = \sum_a \langle \delta \psi_a | h | \psi_a \rangle + \sum_{a,b} \langle \delta \psi_a \psi_b | \psi_a \psi_b - \delta \psi_a \psi_b | \psi_b \psi_a \rangle \\ - \sum_{a,b} \varepsilon_{ba} \langle \delta \psi_a | \psi_b \rangle + \text{complexo conjugado.} \end{aligned} \quad (\text{A.40})$$

Definindo os operadores de Coulomb, \mathcal{J}_b , e de Troca, \mathcal{K}_b através das expressões

$$\mathcal{J}_b(1) \psi_a(1) = \left\langle \psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_b(2) \right\rangle \psi_a(1), \quad (\text{A.41})$$

$$\mathcal{K}_b(1) \psi_a(1) = \left\langle \psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_b(2) \right\rangle \psi_b(1). \quad (\text{A.42})$$

Para que L seja um mínimo devemos ter $\delta L=0$. Como $\delta \psi_a$ é arbitrário, tem-se

$$\{h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)]\} \psi_a(1) = \sum_b \varepsilon_{ba} \psi_b(1). \quad (\text{A.43})$$

A expressão entre chaves

$$\mathcal{F}(1) = h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)] \quad (\text{A.44})$$

que é chamada de operador de Fock. Este operador depende das soluções da equação S.O.M.. Existe uma equação para cada orbital atômico, que por sua vez depende dos outros orbitais através do operador de Fock. Portanto, as equações devem ser resolvidas de maneira acoplada, por meio de aproximações sucessivas. Devido a isso o método HF é um método autoconsistente, de modo que, no final do processo os ψ s que são soluções da equação de Hartree-Fock devem ser os mesmos ψ s a partir dos quais o operador de Fock foi obtido.

Em termos deste operador, temos

$$\mathcal{F}(1) \psi_a(1) = \sum_b \varepsilon_{ba} \psi_b(1), \quad (\text{A.45})$$

usando as definições dos operadores de Coulomb e de troca, temos

$$\varepsilon_a = \langle a|h|a\rangle + \sum_b \langle ab||ab\rangle. \quad (\text{A.46})$$

Da equação A.34, temos

$$E = \sum_a \varepsilon_a - \frac{1}{2} \sum_{a,b} \langle ab||ab\rangle. \quad (\text{A.47})$$

A energia eletrônica total, portanto, não é igual à soma das energias dos spins-orbitais. As energias dos orbitais são autovalores do operador de Fock, que inclui a energia de repulsão entre um determinado elétron e os demais elétrons. A energia de repulsão elétron-elétron deve ser descontada, já que é computada duas vezes na soma das energias dos orbitais [31, 52, 53].

A.3 A equação de Hartree-Fock-Roothaan

O método de Hartree-Fock é bem resolvido numericamente para sistemas de átomos ou moléculas com poucos elétrons, tornando-se inviável para sistemas grandes (moléculas poliatômicas). Uma solução para esse problema foi introduzida por Slater, mas formalmente apresentada por Roothaan (1951), por meio da expansão da parte espacial dos spins-orbitais moleculares em termos de um conjunto de funções conhecidas, chamadas de funções base, transformando a solução das equações integro-diferenciais acopladas de HF em um problema de álgebra matricial. A formulação matricial do método de Hartree-Fock é a equação de Hartree-Fock-Roothaan (HFR) [31, 52, 53, 54].

O processo consiste em escrever os orbitais moleculares como

$$\phi_p(r) = \sum_{v=1}^k C_{vp} g_v(r), \quad (\text{A.48})$$

onde

$-g_v(r)$ são as funções base conhecidas,

- k é o número de funções do conjunto,

- C_{vp} são os coeficientes a serem determinados.

As funções de base do tipo gaussianas são as mais utilizadas nos cálculos moleculares, por causa da facilidade de se calcular integrais usando este tipo de função. Uma função gaussiana escrita em coordenadas cartesianas é determinada como

$$g(x, y, z) = Nx^l y^m z^n e^{-\alpha r^2} \quad (\text{A.49})$$

onde l, m e n são números inteiros, $r = \sqrt{x^2 + y^2 + z^2}$ e N é uma constante de normalização.

Por convenção, funções gaussianas cartesianas podem ser classificadas como

$l + m + n = 0 \rightarrow$ função tipo s

$l + m + n = 1 \rightarrow$ função tipo p

$l + m + n = 2 \rightarrow$ função tipo d

$l + m + n = 3 \rightarrow$ função tipo f

Substituindo a expressão A.48 na equação de Hartree-Fock, temos

$$f(\mathbf{r}_1) \sum_{\nu} C_{\nu p} g_{\nu}(\mathbf{r}_1) = \epsilon_p \sum_{\nu} C_{\nu p} g_{\nu}(\mathbf{r}_1) \quad (\text{A.50})$$

Multiplicando por $g_{\mu}^*(\mathbf{r}_1)$ e integrando, tem-se

$$\sum_{\nu} C_{\nu p} \langle g_{\nu}(1) | f(1) | g_{\nu} \rangle = \epsilon_p \sum_{\nu} C_{\nu p} \langle g_{\mu}(1) | g_{\nu} \rangle \quad (\text{A.51})$$

onde $g_{\nu}(1) \equiv g_{\nu}(\mathbf{r}_1)$.

Definindo os elementos da matriz de superposição como

$$S_{\mu\nu} = \langle g_{\mu}(1) | g_{\nu} \rangle \quad (\text{A.52})$$

e da matriz de Fock como

$$F_{\mu\nu} = \langle g_{\mu}(1)|f(1)|g_{\nu}\rangle \quad (\text{A.53})$$

podemos escrever

$$\sum_{\nu} F_{\mu\nu} C_{\nu p} = \varepsilon_p \sum_{\nu} S_{\mu\nu} C_{\nu p}. \quad (\text{A.54})$$

A forma matricial de A.54 é

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}, \quad (\text{A.55})$$

onde $\boldsymbol{\varepsilon}$ é a matriz diagonal que contém as energias orbitais $\varepsilon_{pq} = \varepsilon_p \delta_{pq}$. Essa equação matricial é conhecida como equação de Hartree-Fock-Roothaan[55].

APÊNDICE B – FORMALISMO DOS MÉTODOS SEMI-EMPÍRICOS

Os métodos denominados semi-empíricos procuram resolver as equações de Hartree-Fock Roothaan de forma aproximada. Essa aproximação consiste na utilização de parâmetros, resultados experimentais e ajustes numéricos, para simplificar a resolução das equações. A maior parte dos métodos semi-empíricos fazem uso da teoria de orbitais moleculares para construção da função de onda molecular, a partir de uma combinação linear de orbitais atômicos, com um conjunto de base mínima de valência formado por funções do tipo Slater (STOs- *Slater type orbitals*). A abordagem teórica pode definir os métodos semi-empíricos quanto à aproximação integral adotada, às expressões utilizadas nos cálculos das diversas parcelas das integrais necessárias e ao procedimento de parametrização

Inicialmente considera-se somente os elétrons de valência de forma explícita, e utiliza-se um conjunto mínimo de funções do tipo s e p no conjunto de funções base STO.

A aproximação ZDO (*Zero Differential Overlap*) é a base dos cálculos semi-empíricos. Esta fundamenta-se em desconsiderar todos os produtos entre funções de base que dependam das coordenadas do mesmo elétron localizadas em sítios atômicos diferentes (o produto de dois orbitais atômicos diferentes é nulo), o que simplifica muito as equações de Hartree-Fock Roothaan. Assim, a consequência dessa aproximação é fazer com que a matriz de *overlap* (sobreposição) **S** reduza-se a matriz identidade ($S_{\mu\nu} = \delta_{\mu\nu}$) fazendo com que a equação de autovalores de Hartree-Fock Roothaan seja reduzida a uma forma mais simplificada e solucionada por meio da diagonalização da matriz de Fock.

B.1 Equações MNDO

O método MNDO (*Modified Neglect of Diatomic Overlap*) utiliza uma base de função do tipo STO formada por um orbital s e três p para cada átomo, com exceção do hidrogênio [31].

Os elementos de matriz de Fock para as moléculas de camada fechada e para átomos que possuam esse conjunto de base são expressas como:

$$F_{\mu\mu} = H_{\mu\mu} + \sum_{\nu}^A P_{\nu\nu}[(\mu\mu|\nu\nu) - \frac{1}{2}(\mu\nu|\mu\nu)] + \sum_B \sum_{\lambda\sigma}^B P_{\lambda\sigma}(\mu\mu|\lambda\sigma) \quad (\text{B.1})$$

$$F_{\mu\nu} = H_{\mu\nu} + \frac{1}{2}P_{\mu\nu}[3(\mu\nu|\mu\nu) - (\mu\mu|\nu\nu)] + \sum_B \sum_{\lambda\sigma}^B P_{\lambda\sigma}(\mu\mu|\lambda\sigma) \quad (\text{B.2})$$

$$F_{\mu\nu} = H_{\mu\nu} - \frac{1}{2} \sum_{\nu}^A \sum_{\sigma}^B P_{\nu\sigma}(\mu\mu|\lambda\sigma) \quad (\text{B.3})$$

onde

$$P_{\lambda\sigma} = \sum_i^{occ} c_{\lambda i}^* c_{\sigma i}. \quad (\text{B.4})$$

A equação B.1 é aplicada para os elementos da diagonal da matriz de Fock ($\mu = \nu$). Se μ e ν são diferentes, mas estão localizados no mesmo centro, aplicamos, para estes elementos fora da diagonal, a equação B.2. Por último, se μ e λ estão em centros diferentes, os elementos da matriz de Fock correspondentes são calculados por meio da equação B.3.

No método MNDO, as integrais monoatômicas de repulsão eletrônica são obtidas utilizando dados experimentais. Para cada átomo, que usa uma base mínima formada de orbitais s e p , existe um total máximo de cinco destas integrais, que são obtidas com base no procedimento proposto por Oleari e colaboradores [31].

As integrais diatômicas de repulsão eletrônica ocorrem num total de 22 formas distintas para cada par de átomos pesados. Isto é válido dentro do formalismo NDDO (Neglect of Diatomic Differential Overlap), considerando apenas moléculas que possuam átomos contendo orbitais do tipo s e p .

A forma geral para as integrais de dois elétrons, considerando a expansão de multipolos, é dada por:

$$(\mu\nu|\lambda\sigma) = \sum_{l_1 m_1 l_2 m_2} [M_{l_1 m_1}^A, M_{l_2 m_2}^B] \quad (\text{B.5})$$

onde

$$[M_{l_1 m_1}^A, M_{l_2 m_2}^B] = \frac{e^2}{2^{l_1+l_2}} \sum_{i=1}^{2l_1} \sum_{j=1}^{2l_2} f_1(R_{ij}) \quad (\text{B.6})$$

e R_{ij} representa a distância entre as distribuições de carga i e j , que pode ser obtida a partir da distância interatômica entre os átomos A e B.

A integral de um elétron e um centro, ou hamiltoniano do caroço, $H_{\mu\mu}$, é aproximada como a soma de uma parcela que representa a energia cinética de um elétron, $U_{\mu\mu}$, que é obtida, mais uma vez, parametricamente, somada a uma parcela que está associada à interação elétron-núcleo, calculada a partir de integrais de repulsão eletrônica. A integral $H_{\mu\mu}$ é a energia que um elétron (num orbital atômico ϕ_μ), teria se todos os demais elétrons fossem removidos, e é escrita da seguinte forma:

$$H_{\mu\mu} = U_{\mu\mu} - \sum_{B \neq A} (\mu\mu|s_B s_B). \quad (\text{B.7})$$

A aproximação para a integral de ressonância (um elétron e dois centros) é feita usando a integral de *overlap* $S_{\mu\lambda}$:

$$H_{\mu\lambda} = \frac{1}{2} S_{\mu\lambda} (\beta_{\mu} + \beta_{\lambda}) \quad (\text{B.8})$$

O último conjunto de equações para este método, são as integrais de repulsão caroço-caroço, que dependem do par de elétrons envolvidos.

Embora o método MNDO realize cálculos de propriedades termodinâmicas e mecanismos de reação com eficiência, no entanto, contém algumas falhas, como a descrição das ligações de hidrogênio, entre outros problemas. A correção para estas falhas foi proposta com a criação de um novo método a partir do MNDO, o denominado método AM1 (*Austin Model 1*) introduzido em 1985, se diferenciando do MNDO pela inclusão de funções gaussianas esféricas nas integrais de repulsão caroço-caroço (os quais eram superestimadas pelo MNDO):

$$E_N(A, B) = E_N^{MNDO}(A, B) + \frac{Z_A Z_B}{R_{AB}} \left(\sum_k a_{k_A} e^{-b_{k_A} (R_{AB} - c_{k_A})^2} \sum_l a_{l_B} e^{-b_{l_B} (R_{AB} - c_{l_B})^2} \right) \quad (\text{B.9})$$

onde

a, b e c são os coeficientes ajustáveis que definem, respectivamente, a intensidade, largura e posição destas funções Gaussianas.

Com a introdução das funções Gaussianas foi possível resolver com o AM1 algumas falhas do MNDO, como as ligações de hidrogênio. Contudo, surgiram erros de parametrização. Em sequência, o método PM3 (*Parametric Method 3*) foi publicado por Stewart em 1989, com o aprimoramento da técnica de parametrização, o que reduziu os erros, principalmente para entalpia de formação.

Atualmente estes dois métodos (AM1e PM3) estão entre os mais utilizados métodos semi-empíricos para descrever propriedades físico-químico-estruturais de uma gama de compostos em sistemas biológicos[31, 32].

Referências

- [1] SANTOS, H. F. et al. **Escola Brasileira de Estrutura Eletrônica**, 1ª edição. São Paulo: Livraria da Física, 2002.
- [2] Ministério da Saúde, INCA-Instituto Nacional de Câncer José Alencar Gomes da Silva. **ABC do Câncer: Arbudagens básicas para o controle do câncer**. Rio de Janeiro–RJ, 2011.
- [3] ROSSIT, A.; FROES, N. Suscetibilidade genética, biometabolismo e câncer. *Sociedade Brasileira de Cancerrologia*, n. 10, p. 26–31, 2001.
- [4] ABBAS, A.; KUMAR, V.; FAUSTO, N. **Robbins & Cotran: Patologia-Bases Patológicas das Doenças**, 7ª Edição. Rio de Janeiro: Elsevier, 2005.
- [5] FRIAS, J.; SILVA, C. **Saúde do Trabalhador no Maranhão: uma visão atual e proposta de atuação**, 137f. Dissertação (Mestrado em Saúde Pública), Programa de Pós-Graduação em Saúde Pública, Fundação Instituto Oswaldo Cruz. Rio de Janeiro, 1999.
- [6] KARP, G. **Biologia Celular e Molecular**, 3ª edição. Rio de Janeiro: Editora Manole, 2005.
- [7] OLIVEIRA, A. **Determinação de Hidrocarbonetos Aromáticos Policíclicos Presentes no Material Particulado da Atmosfera Uberlandense**. Dissertação (Mestrado em Engenharia Química), Faculdade de Engenharia Química, Universidade Federal de Uberlândia. Uberlândia–MG, 2008.
- [8] DUARTE, R.; PACHOAL, M. Marcadores Moleculares no Câncer de Pulmão: papel prognóstico e sua relação com o tabagismo. *Jornal Brasileiro de Pneumologia*, v. 32, p. 56–65, 2005.
- [9] TONINI, R.; REZENDE, C.; GRATIVO, A. Degradação e biorremediação de compostos de petróleo por bactérias: Revisão. *Oecologia Australis*, v. 14, n. 4, p. 1025–1035, 2010.
- [10] COSTA, A. **Avaliação da Contaminação Humana por Hidrocarbonetos Aromáticos Policíclicos (PAHs): 1-Hidroxipireno Urinário**, 98f. Dissertação (mestrado em Saúde Pública), Centro de Estudos da Saúde do Trabalhador e Ecologia Humana, Fundação Oswaldo Cruz. Rio de Janeiro, 2001.
- [11] NETTO, P. et al. Avaliação da Contaminação Humana por Hidrocarbonetos Policíclicos Aromáticos (hpas) e seus derivados nitrados (nhpas): uma revisão metodológica. *Química Nova*, v. 23, 2000.
- [12] MEIRE, R.; TORRES, J. Aspectos Ecotoxicológicos de Hidrocarbonetos Policíclicos Aromáticos. *Oecol. Bras.*, v. 11, n. 2, p. 188–201, 2007.

- [13] ALMEIDA, V. et al. Câncer e Agentes Antineoplásicos Ciclo-celular específicos e Ciclo-celular não específicos que Interagem com o DNA: uma introdução. *Química Nova*, v. 28, p. 118–129, 2005.
- [14] HUSSEIN, M. S.; RABELLO, P. Combate ao Câncer Ganha Novas Técnicas: a Física Nuclear e o Tratamento de Tumores. *Revista USP*, n. 66, p. 80–95, 2005.
- [15] SILVA, W. A. A importância dos estudos genéticos sobre o câncer de pulmão. *Jornal Brasileiro de Pneumologia*, v. 35, n. 8, p. 721–722, 2009.
- [16] SILVA, P.; RIUL, S. Câncer de Mama: Fatores de Risco e Detecção Precoce. *Revista Brasileira de Enfermagem*, v. 64, n. 6, p. 1016–1021, 2012.
- [17] OLIVEIRA, R.; ALVES, R. Agentes Antineoplásicos Biorredutíveis: uma nova alternativa para o tratamento de tumores sólidos. *Química Nova*, v. 25, p. 976–984, 2002.
- [18] FERDINANDI, D.; FERREIRA, A. Agentes alquilantes: Reações adversas e complicações hematológicas. *Científica*, 2005.
- [19] AZEVEDO, W. Ácidos nucléicos com alvos para drogas. *Química Medicinal*, 2012.
- [20] PAULIKOVA, H. et al. DNA binding acridine–thiazolidinone agents affecting intracellular glutathione. *Bioorganic & Medicinal Chemistry*, v. 20, p. 7139–7148, 2012.
- [21] WAINWRIGHT, M. Acridine: a neglected antibacterial chromophore. *Journal of Antimicrobial Chemotherapy*, v. 47, p. 1–13, 2001.
- [22] FINLAY, G. J.; RIOU, J. F.; BAGULEY, B. C. From amsacrine to daca (n-[2-(dimethylamino) ethyl]acridine-4-carboxamide): Selectivity for topoisomerases i and ii Among Acridine Derivatives. *European Journal of Cancer*, v. 32A, p. 708–714, 1996.
- [23] SEDLACEK, O. et al. Polymer conjugates of acridine-type anticancer drugs with ph-controlled activation. *Bioorganic e Medicinal Chemistry*, v. 20, p. 4056–4073, 2012.
- [24] JOSEPH, T. et al. Suscetibilidade genética, biometabolismo e câncer. *Bioconjugate Chem.*, n. 15, p. 1230–1235, 2004.
- [25] CUNY, G. et al. Structure–activity relationship study of acridine analogs as haspin and dyrk2 kinase inhibitors. *Bioorganic e Medicinal Chemistry Letters*, v. 20, p. 3491–3494, 2010.
- [26] REZENDE, L. **Estrutura Eletrônica de Hidrocarbonetos Aromáticos Heterocíclicos**. Dissertação (mestrado em física), Programa de Pós-graduação em Física, Universidade Federal de Juiz de Fora. Juiz de Fora, 2002.
- [27] SATO, F. **Estudo da Estrutura-Atividade das Esterilquinolinas e Tetraciclínas Através de Descritores Quânticos**, 120f. Dissertação (mestrado em Física), Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas. Campinas, 2003.
- [28] LLANES, S. **Investigação da Atividade Biológica de Taxóides e Benzo[c]quinolizina-3-onas Através de Descritores Teóricos**, 187f. Tese (Doutorado) — Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, Campinas, 2003.

- [29] SANTO, L. *de Métodos Semiempíricos ao Estudo da Estrutura Eletrônica de Compostos Bioativos*. Tese (Doutorado) — Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, Campinas, 2001.
- [30] PASSAMANI, F. **Modelagem molecular e Avaliação da Relação Estrutura-Atividade Acoplados a estudos Físico-químicos e Toxicológicos *In Silico* de derivados Heterocíclicos com atividade Antiviral**, 150f. Dissertação (mestrado em Ciências Farmacêuticas), Faculdade de Farmácia, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2009.
- [31] MORGON, N.; COUTINHO, K. **Métodos de Química Teórica e Modelagem Molecular**. São Paulo, Livraria da Física, 2007.
- [32] VIANNA, J.; FAZZIO, A.; CANUTO, S. **Teoria Quântica de Moléculas e Sólidos**. São Paulo, Livraria da Física, 2004.
- [33] BURKE, K. Perspective on density functional theory. *The Journal of Chemical Physics*, v. 136, p. 1509011–9, 2012.
- [34] HOHENBERG, P.; KOHN, W. Inhomogeneous Electron Gas. *Physical Review*, v. 136, n. 3b.
- [35] KOHN W. AND SHAM, J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, v. 140, n. 4A.
- [36] BECKE, A. D. Density-functional thermochemistry. III. The role of exact exchange. *Journal Chemistry Physics*, v. 98, n. 7.
- [37] LEE, C.; YANG, W.; PARR, R. Development of the Colle-Solvetti correlation-energy formula into a functional of the electron density. *Physical Review*, v. 37, n. 2, p. 785–789, 1988.
- [38] DEWAR, M. et al. Am1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem.*, n. 107, p. 3902–3909, 1985.
- [39] QUIRANTE, J. Study of the Bimolecular Pyrolysis of Acetic Acid by the Austin Model 1 Semi-empirical Method. *Journal of Analytical and Applied Pyrolysis*, n. 31, p. 169–175, 1995.
- [40] BRAGA, R.; BARONE, P.; GALVÃO, D. Identifying Carcinogenic Activity of Methylated and non-methylated Polycyclic Aromatic Hydrocarbons (pahs) Through Electronic and Topological Indices. *Brazilian Journal of Physics*, v. 30, p. 560–568, 2000.
- [41] VENDRAME, R. *Estudo das Relações entre Estrutura-Atividade de Progestogênios Contraceptivos, Através da Utilização de Parâmetros Físico-Químicos Calculados e Métodos de Reconhecimento Padrão*, 402f. Tese (Doutorado) — Instituto de Química, Universidade Estadual de Campinas, Campinas, 1998.
- [42] ARROIO, A.; HONÓRIO, K.; SILVA, A. Propriedades Químico-Quânticas Empregadas em Estudos das Relações Estrutura-Atividade. *Quim.Nova*, n. 3, p. 694–699, 2010.
- [43] LYRA, W. et al. Classificação Periódica: Um exemplo Didático para Ensinar Análise de Componentes Principais. *Quim.Nova*, n. 7, p. 1594–1597, 2010.

- [44] FERREIRA, M. et al. Quimiometria I: Calibração Multivariada, um Tutorial. *Quim.Nova*, n. 5, p. 724–731, 1999.
- [45] FERREIRA, E. et al. Análise Exploratória dos Teores de Constituintes Inorgânicos em Sucos e Refrigerantes de Uva. *Eclét.Quím.*, p. 77–90, 2002.
- [46] KIAMETIS, A. *Modelagem Molecular de Proteínas Candidatas a Inibidores da Acetilcolinesterase*, 90f. Tese (Doutorado) — Instituto de Física, Universidade de Brasília, Brasília, 2012.
- [47] PAULA, A. *Estudo de Estrutura Eletrônica e Análise Multivariada de Candidatos a inibidores da AChE*, 104f. Tese (Doutorado) — Instituto de Física, Universidade de Brasília, Brasília, 2009.
- [48] CORREIA, P.; FERREIRA, M. Reconhecimento de Padrões por Métodos não Supervisionados: explorando procedimentos quimiométricos para tratamento de dados analíticos. *Quim.Nova*, n. 2, p. 481–487, 2007.
- [49] ALMEIDA, J. et al. Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. *ScienceDirect*, v. 87, p. 208–217, 2007.
- [50] NETO, J.; MOITA, G. Uma Introdução à Análise Exploratória de Dados Multivariados. *Química Nova*, v. 21, n. 4, p. 467–469, 1998.
- [51] SIQUEIRA, T.; CAMARGO, L. Estudos Teóricos da Relação Estrutura Atividade Contra o Câncer. *Revista Processos Químicos*, p. 116–121, 2009.
- [52] SZABO, A.; OSTLUND, N. **Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory**, 2ª edição. New York, Dover Publications, 1996.
- [53] LEVINE, I. **Quantum Chemistry**, 6ª edição. New York, Prentice Hall, 1996.
- [54] ARRUDA, P. *Algumas Considerações sobre Conjuntos de Bases para Cálculos de Propriedades Elétricas*. Centro de Ciências Exatas, Universidade Federal do Espírito Santo. Dissertação (Mestrado), Vitória, 2009.
- [55] CAPELLE, K. A Bird's-Eye View of Density-Functional Theory. *Brazilian Journal of Physics*, v. 36, n. 4A, p. 1318–1343.