



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

**Um Estudo Teórico das Relações Quantitativas Estrutura-
Atividade em Derivados de 2-Fenilindan-1,3-dionas.**

Aluizio Galdino da Silva

Dissertação de Mestrado

Recife

2010



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

**Um Estudo Teórico das Relações Quantitativas Estrutura-
Atividade em Derivados de 2-Fenilindan-1,3-dionas.**

Aluizio Galdino da Silva*

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

Orientador: Prof. Dr. Mozart Neves Ramos

Co-orientador: Prof. Dr. João Bosco Paraíso da Silva

* Bolsista CAPES

Recife – PE

Julho / 2010

Silva, Aluizio Galdino da.

Um estudo teórico das relações quantitativas estrutura-atividade em derivados de 2-Fenilindan-1,3-dionas / Aluizio Galdino da Silva. - Recife: O Autor, 2010.

72 p.: il. fig.

Dissertação (Mestrado) - Universidade Federal de Pernambuco. CCEN. Química, 2010.

Inclui bibliografia.

1.Físico-química. 2.Colesterol no sangue. 3.Análise de regressão. I.Título.

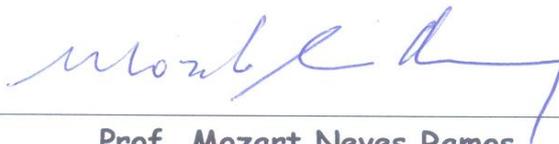
541.3

(22.ed.)

FQ 2010-050

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

Aprovada:



Prof. Mozart Neves Ramos
(Orientador)
Departamento de Química Fundamental
Universidade Federal de Pernambuco



Prof. Sebastião José de Melo
Departamento de Antibióticos
Universidade Federal de Pernambuco



Prof. Alfredo Arnóbio de Souza da Gama
Departamento de Química Fundamental
Universidade Federal de Pernambuco

**"Um Estudo Teórico das Relações Quantitativas Estrutura-Atividade
em Derivados de 2-fenilindan-1,3-dionas"**

Por

Aluizio Galdino da Silva

**Departamento de Química Fundamental
Centro de Ciências Exatas e da Natureza
Universidade Federal de Pernambuco
Recife-PE-Brasil
26 de Julho de 2010**

À minha mãe **Josefa Maria** e
ao meu pai **José Galdino** (*In*
Memorian).

Agradecimentos

Agradeço primeiramente a Deus por ter me dado a vida e me ajudado a superar todos os obstáculos, e aos meus pais pelos sacrifícios feitos em prol da educação dos seus filhos, tanto escolar como doméstica.

Aos meus orientadores, professor Mozart Neves Ramos e João Bosco Paraíso da Silva, pelas discussões, paciência, amizade, sugestões e pela grande ajuda na realização deste trabalho.

Aos professores do DQF Benício de Barros Neto e Ricardo Luiz Longo, pela paciência e pelas milhares de sugestões dadas.

A professora Maria Cristina Falcão Raposo do departamento de estatística da UFPE, pela atenção e sugestões ao longo do trabalho.

Ao meu amigo irmão Alexsandro Cavalcanti, pelo apoio.

Ao pessoal do LQTC, em especial a Eduardo Castro e Victor Rusu, pelas sugestões e discussões.

E a todos do Departamento de Química Fundamental, que ajudaram de forma direta ou indireta, na elaboração desse trabalho.

A CAPES pela bolsa concedida.

Resumo

Neste trabalho apresentamos um estudo teórico das relações quantitativas estrutura-atividade (QSAR) com treze derivados de 2-fenilindan-1,3-dionas possuindo atividade hipolipidêmica. Para estabelecer essas relações, nós usamos descritores eletrônicos, obtidos de cálculos de orbitais moleculares AM1 e B3LYP/6-31G(d,p), com ou sem efeito solvente, além dos descritores empíricos $\log P$ e π . Nesses procedimentos nós realizamos uma completa otimização das geometrias moleculares para esses treze derivados, e para alguns deles, uma análise conformacional foi também realizada.

Na análise estatística multivariada nós empregamos o modelo da regressão linear e a análise de componentes principais.

Nossos resultados indicaram que a atividade na redução dos níveis de colesterol apresenta uma interessante correlação linear com a carga atômica do hidrogênio ao carbono sp^3 do anel indan. Menor a carga positiva desse hidrogênio, maior será a capacidade do derivado em reduzir os níveis de colesterol, após 16 dias de aplicação em ratos machos.

Quanto a redução dos níveis de triglicerídeos, nossos resultados revelaram que menor o momento dipolar da molécula e mais positiva for a constante de lipofilicidade π do substituinte, maior será essa atividade.

A partir das equações matemáticas encontradas, nós temos planejado e proposto dois novos derivados (*orto*-hidroxi-2-fenilindan-1,3-diona e *orto*-flúor-2-fenilindan-1,3-diona) que a priori devem ser potencialmente mais ativos do que os compostos usados no conjunto de treinamento.

Palavras-chaves: QSAR, atividade hipolipidêmica, análise multivariada.

Abstract

In this work we present a theoretical study of quantitative structure-activity relationships (QSAR) with thirteen derivatives of 2-phenylindan-1,3-diones possessing hypolipidemic activity. In order to establish these relationships we have used electronic descriptors obtained from AM1 and B3LYP/6-31G(d,p) molecular orbital calculations, with or without solvent effect, beyond the Log P and π empirical descriptors. In these procedures we have performed a full optimization of molecular geometries for the thirteen derivatives, and for some, a conformational analysis was also performed.

In multivariate analysis we have employed both the model of linear regression and the principal component analysis.

Our results have indicated that the activity in reducing cholesterol levels presents an interesting linear correlation with the atomic charge of hydrogen bonded to the sp^3 carbon of indan ring. Lower the positive charge of this hydrogen, the greater the ability of the derivative to reduce the cholesterol levels after 16 days of administration in CF₁ male mice.

As for reducing triglyceride levels, our results have revealed that the smaller the molecular dipole moment and more positive the π lipophilicity constant of the substituent, the greater it will be this activity.

From these mathematical equations, we have designed and proposed two new derivatives (*ortho*- hydroxy-2-phenylindan-1,3-dione and *ortho*- fluorine-2-phenylindan-1,3-dione), which are predict to be more potent than those used in the training set.

Keywords: QSAR, hypolipidemic activity, multivariate analysis

Sumário

1. Introdução.....	1
1.1. Planejamento de Fármacos e Química Medicinal.....	1
1.2. As Ftalimidas.....	7
1.2.1. Atividade Hipolipidêmica das Ftalimidas.....	8
2. Objetivos.....	16
2.1 Objetivos gerais.....	16
2.2 Objetivos Específicos.....	16
3. Fundamentação Teórica.....	17
3.1. Técnicas Computacionais.....	17
3.2. O Método de Hartree-Fock.....	18
3.3. Métodos ab initio.....	19
3.4. Métodos semi-empíricos.....	20
3.5. Teoria do Funcional da Densidade.....	21
3.6. Efeito do Solvente.....	22
3.6.1 Modelos Teóricos.....	22
3.6.2 Modelos Contínuos.....	22
3.7. Quimiometria Aplicada ao Estudo de QSAR.....	24
4. Cálculos.....	29
4.1 Otimização da Geometria Molecular da Molécula-Mãe e seus Derivados.....	29
4.2 Análise Conformacional da Molécula-Mãe.....	29
4.3 Obtenção das Propriedades Eletrônicas e Empíricas.....	30
4.4 Efeito Solvente.....	31
4.5 Estudo Quimiométrico.....	32
5. Resultados e Discussões.....	33
5.1 Estudo da Molécula Mãe: 2-Fenilindan-1,3-diona.....	33
5.1.1 Análise Conformacional.....	33
5.2 Seleção dos Descritores Eletrônicos e Clássicos.....	40
5.2.1 Método AM1 – Modelo da Molécula Isolada (AM1-MMI).....	41
5.2.1.1 Análise RLM - pC ₁₆	42
5.2.1.2 Análise RLM - pT ₁₆	47
5.2.1.3 Análise de Componentes Principais (PCA).....	50
5.2.2 Método B3LYP/6-31G (d,p) - Modelo da Molécula Isolada (B3LYP-MMI)....	54
5.2.2.1 Usando o Modelo de Regressão Linear (RLM).....	55
6. Efeito Solvente.....	57

6.1	Descritores AM1 no Efeito Solvente.....	57
6.2	Descritores B3LYP no Efeito Solvente.....	59
7.	Novos Agentes Hipolipidêmicos	61
8.	Conclusões.....	66
9.	Referências	68

Índice de Tabelas

Tabela 1	Ângulos diédricos da otimização total de energia mínima dos 13 derivados de 2-fenilindan-1,3-dionas	37
Tabela 2	Atividades hipolipidêmicas, descritores eletrônicos AM1, na molécula isolada, e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.	41
Tabela 3	Matriz de correlação das variáveis (AM1-Molécula Isolada) e das atividades.	42
Tabela 4	Matriz de correlação das variáveis e da atividade após o ajuste dos dois derivados 10 e 11 .	45
Tabela 5	Redução nos níveis de colesterol obtidos dos dados experimentais e previstos usando a equação 5.2, após 16 dias de administração da droga.	57
Tabela 6	Matriz de correlação das variáveis e da atividade após o ajuste dos dois derivados.	49
Tabela 7	Redução nos níveis de colesterol obtidos dos dados experimentais e previstos usando a equação 5.5, após 16 dias de administração da droga.	53
Tabela 8	Atividades hipolipidêmicas, descritores eletrônicos B3LYP, na molécula isolada, e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.	54
Tabela 9	Matriz de correlação para os valores de R mais significativos entre as variáveis previamente selecionadas B3LYP e as atividades hipolipidêmicas.	55
Tabela 10	Atividades hipolipidêmicas, descritores eletrônicos AM1/COSMOS e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.	57
Tabela 11	Matriz de correlação para os valores de R mais significativos entre as variáveis previamente selecionadas (descritores eletrônicos obtidos dos cálculos AM1/COSMOS e descritores empíricos) e as atividades hipolipidêmicas.	58
Tabela 12	Atividades hipolipidêmicas, descritores eletrônicos B3LYP/6-31G(d,p)/PCM e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.	59

Tabela 13	Matriz de correlação para os valores de R mais significativos entre as variáveis previamente selecionadas (descritores eletrônicos obtidos dos cálculos B3LYP/6-31G(d,p)/PCM e descritores empíricos) e as atividades hipolipidêmicas.	60
Tabela 14	Comparação na redução dos níveis de colesterol entre a molécula-mãe (2-fenilindan-1,3-diona) e o novo derivado projetado teoricamente (<i>orto</i> -hidróxi-2-fenilindan-1,3-diona), segundo cálculos AM1 no modelo da molécula isolada.	63
Tabela 15	Comparação na redução dos níveis de triglicerídeos entre a molécula-mãe (2-fenilindan-1,3-diona) e o novo derivado projetado teoricamente (<i>orto</i> -hidróxi-2-fenilindan-1,3-diona), segundo cálculos AM1 no modelo da molécula isolada.	63
Tabela 16	Comparação na redução dos níveis de triglicerídeos entre a molécula-mãe (2-fenilindan-1,3-diona) e o novo derivado projetado teoricamente (<i>orto</i> -flúor-2-fenilindan-1,3-diona), segundo cálculos B3LYP no modelo da molécula isolada.	65

Índice de Figuras

Figura 1	Alguns exemplos de imidas cíclicas: (a) Succinimida; (b) maleimida; (c) ftalimida	7
Figura 2	Ftalimidas e seus derivados	9
Figura 3	N-fenilftalimida e seus derivados	10
Figura 4	Oito derivados de ftalimidas	11
Figura 5	(a) Ftalimida e (b) Indan-1,3-diona.	12
Figura 6	(a) Indan-1,3-diona e (b) 2-fenilindan-1,3-dionas.	12
Figura 7	13 derivados de 2-fenilindan-1,3-dionas.	13
Figura 8	Gráfico mostrando como varia a redução dos níveis de colesterol e triglicerídeos de três derivados da série homóloga em estudo, em função da posição do anel.	14
Figura 9	Gráficos mostrando como varia a redução dos níveis de colesterol e triglicerídeos mudando a natureza dos substituintes na posição <i>orto</i> .	14
Figura 10	Molécula de um soluto solvatada por dipolos	23
Figura 11	Molécula de um soluto imersa em uma cavidade de esferas sobrepostas. Solvatação pelo contínuo dielétrico	24
Figura 12	Molécula-mãe: 2-fenilindan-1,3-diona.	29
Figura 13	Esquema da rotação do grupo fenil para análise da barreira de rotação.	30

Figura 14	Gráfico da barreira de rotação em torno da ligação C ₁₃ – C ₁₇ para a molécula-mãe 2-fenilindan-1,3-diona utilizando o método AM1.	33
Figura 15	Ângulo diédrico C ₈ C ₁₃ C ₁₇ C ₁₈ (7,5°)	34
Figura 16	Ângulo diédrico C ₈ C ₁₃ C ₁₇ C ₁₈ (52,5°)	34
Figura 17	Gráfico da barreira de rotação em torno da ligação C ₁₃ - C ₁₇ para a molécula-mãe 2-fenilindan-1,3-diona utilizando o método B3LYP/6-31G (d, p).	35
Figura 18	Estrutura “Gauche” da molécula-mãe para o ângulo diédrico C ₈ C ₁₃ C ₁₇ C ₁₈ de 7,1°	36
Figura 19	Estrutura “Perpendicular” da molécula-mãe para o ângulo diédrico C ₈ C ₁₃ C ₁₇ C ₁₈ de 57,1°.	36
Figura 20	Estruturas otimizadas para os 13 derivados de 2-fenilindan-1,3-dionas utilizando o método AM1.	38
Figura 21	Estruturas otimizadas para os 13 derivados de 2-fenilindan-1,3-dionas utilizando o método B3LYP.	39
Figura 22	Relação dos compostos da série homóloga estudada com os respectivos valores da redução dos níveis de colesterol (pC16) e triglicerídeos (pT16), após 16 dias de administração da droga.	40
Figura 23	Valores previstos de pC ₁₆ usando equação (5.1) versus seus correspondentes valores experimentais.	43
Figura 24	Gráficos das barreiras de rotação para os derivados (a) <i>orto</i> -metóxi e (b) <i>para</i> -carbóxi utilizando o método AM1 na molécula isolada, onde a energia relativa é diferença do valor da energia total para um dado ângulo diedro e a energia total da conformação mais estável.	44

Figura 25	Valores previstos de pC_{16} usando equação (5.2) versus seus correspondentes valores experimentais.	46
Figura 26	Gráficos dos Resíduos	46
Figura 27	Gráficos das barreiras de rotação para os derivados (a) <i>orto</i> -metóxi e (b) <i>orto</i> -cloro utilizando o método AM1 na molécula isolada, onde a energia relativa é diferença do valor da energia total para um dado ângulo diedro e a energia total da conformação mais estável.	48
Figura 28	Gráfico dos <i>loadings</i> para os descritores: a carga do hidrogênio 14 (qH_{14}), a carga do carbono 17 (qC_{17}), o coeficiente de partição ($\log P$) e o parâmetro lipofílico π	51
Figura 29	Gráficos dos <i>scores</i>	52
Figura 30	Atividade hipolipidêmica prevista em função da atividade prevista explicando 61% da variância do modelo.	53
Figura 31	Estrutura otimizada do <i>orto</i> -hidróxi-2-fenilindan-1,3-diona.	62
Figura 32	Estrutura otimizada do <i>orto</i> -flúor-2-fenilindan-1,3-diona.	64
Figura 33	(a) Conformação “gauche” padrão, correspondente ao mínimo global de energia, para substituintes eletrônicos em posição <i>meta</i> ou <i>para</i> no anel fenil e (b) conformação “quasi-perpendicular” padrão, correspondente ao mínimo global de energia, para substituintes eletrônicos em posição <i>orto</i> no anel fenil.	66

Siglas e Notações

QSAR	<i>Quantitative Structure-Activity Relationship</i>
HOMO	<i>highest unoccupied molecular orbital</i>
LUMO	<i>lowest occupied molecular orbital</i>
AM1	<i>Austin Model 1</i>
DFT	<i>Density Functional Theory</i>
RLM	Regressão Linear Múltipla
PCA	Análise de Componentes Principais
BOA	<i>Born-Oppenheimer Approximation</i>
HF	Hartree-Fock
PCM	<i>Polarizable Continuum Model</i>
COSMOS	<i>Conductor-like Screening Model</i>

1. Introdução

1.1. Planejamento de Fármacos e Química Medicinal

A necessidade do desenvolvimento de novos fármacos, que sejam efetivos contra algumas patologias ainda sem tratamento adequado, e que possam substituir os existentes, porém a custos menores e dotados de menores efeitos adversos, tem impulsionado a comunidade científica a novas e incessantes pesquisas nesta área. A síntese orgânica tem contribuído significativamente neste aspecto, sendo responsável por cerca de 80% dos fármacos existentes no mercado farmacêutico [1]. Cabe ressaltar, porém, que muitos destes fármacos são oriundos de protótipos advindos de produtos naturais, especialmente de plantas, que têm, ao longo dos anos, possibilitado a descoberta de inúmeras moléculas bioativas [1].

Muitas classes de compostos orgânicos têm demonstrado promissores efeitos biológicos e a literatura científica relata um crescimento significativo de novas moléculas com potência similar ou superior àquela de um fármaco, sendo que muitos deles encontram-se em estudos pré-clínicos e clínicos avançados e pormenorizados [1].

Hoje sabemos que a importância medicinal de um fármaco é considerada de acordo com seu sítio ativo ou seu modo de ação. Em muitos casos, a atribuição do modo de ação pode ser apenas uma tentativa, podendo até muitas drogas serem classificadas por mais de um caminho. Por exemplo: algumas drogas produzem seus efeitos através de processos físico-químicos, outras produzem através da reatividade química, e nestes casos, a explicação para seu modo de ação ocorre nos receptores biológicos, ou seja, a área específica, de uma proteína que iniciará uma mudança física ou química na célula quando esta absorve uma molécula do fármaco com forma e polaridade apropriadas conforme sua geometria [2].

A resposta biológica provocada por um composto bioativo ocorre por meio de interações ligante-biomacromolécula alvo, e depende de sua estrutura química e de propriedades físico-químicas intrínsecas. Estas propriedades podem ser de caráter molecular ou atômico e expressam características de efeito eletrônico, reatividade, forma, lipofilicidade, estabilidade e polaridade da molécula, dentre outras.

Neste aspecto, a realização de modificações na estrutura de um fármaco resulta em alterações em suas propriedades físico-químicas e, assim, em sua absorção,

biodisponibilidade, interação com o receptor biológico e nas suas correspondentes bioatividade [3,4,5].

As propriedades físico-químicas, principalmente no âmbito do desenvolvimento de novos agentes terapêuticos, podem ser utilizadas para classificar e quantificar as contribuições da estrutura química sobre a atividade biológica. Para melhor entendimento, as propriedades físico-químicas de compostos bioativos, como a hidrofobicidade, os efeitos eletrônicos e efeito estérico, são geralmente expressas por meio de parâmetros ou descritores estruturais. Neste contexto, o conhecimento de valores numéricos de cada descritor estrutural permite estimar a contribuição, de maior ou menor intensidade, de tais propriedades sobre a atividade biológica [3,6,7].

A avaliação da magnitude da influência das propriedades moleculares, sobre determinado efeito bioativo é realizada por meio de estudos das relações quantitativas estrutura-atividade, QSAR (do inglês, *Quantitative Structure-Activity Relationship*). Já a avaliação da interação entre as propriedades físico-químicas é realizada por meio da análise das relações quantitativas estrutura-propriedade, QSPR (do inglês, *Quantitative Structure-Property Relationship*).

As análises de QSAR e QSPR ocorrem por meio da obtenção de modelos matemáticos que envolvem descritores de uma série de compostos com estrutura química análoga. Estas análises são aplicadas para a investigação da ocorrência de correlação entre determinado descritor com outras propriedades físico-químicas ou ainda com a atividade biológica estudada [3,7]. Desse modo, a partir do conhecimento das relações entre as características moleculares e a atividade biológica, propõe-se o emprego destas informações no planejamento e busca-se a identificação de fármacos que apresentam melhor perfil farmacológico.

A idéia de que a atividade biológica depende da estrutura química de um determinado composto surgiu através dos estudos de Crum-Brown & Fraser, em 1868 [8]. Estes autores registraram que a paralisia desempenhada por uma série de derivados da estriquinina ocorre em função da estrutura molecular destes compostos, e pode ser expressa através da relação abaixo:

$$\text{Atividade Biológica} = f(\text{Constituição da molécula})$$

Há cem anos atrás, Meyer, em 1899, e Overton, em 1901, concluíram, independentemente, que o efeito narcótico apresentado por diversos compostos não ionizados era dependente do coeficiente de partição clorofórmio-água e que a atividade narcótica observada possuía correlação linear com o aumento do coeficiente de partição.

Hammet em 1937 propôs uma constante σ no sentido de avaliar efeitos eletrônicos de diferentes substituintes em reações de ionização em água de ácidos benzóicos *meta* e *para* substituídos. Hammett obteve o parâmetro σ avaliando Relações Lineares de Energia Livre (RLEL), isto é, relações entre logaritmos de constantes de velocidade ou equilíbrio de reações. A equação 1.1 apresenta a relação proposta por Hammett ao introduzir a constante σ .

$$\log K = \log K^o + \sigma \rho \quad (1.1)$$

onde K e K^o representam as constantes de equilíbrio das reações de ionização do ácido benzóico substituído e não substituído, respectivamente, σ é uma constante que depende somente do substituinte e ρ é uma constante da própria reação, que depende do tipo de reação, do meio e da temperatura envolvida.

Para a determinação dos valores de σ , Hammett estabeleceu que o valor da constante de reação ρ era igual à unidade para a reação de ionização do ácido benzóico. Isto lhe possibilitou determinar vários valores de σ para substituintes usados no ácido benzóico nesta reação, e a partir dessas constantes, obter a constante ρ para outras reações em que os substituintes eram os mesmos empregados para calcular σ na série inicial.

A mais importante proposta da equação de Hammett é a separação entre parâmetros relativos ao substituinte e à reação. Esta separação torna possível o uso dos mesmos valores de σ encontrados por Hammett para as reações de ionização do ácido benzóico em uma variedade de séries reacionais com boa precisão, e por isso essa constante é dita ser uma constante de grupo, visto que define o comportamento de um substituinte independentemente da reação em que se encontra [9]. A validade da separação dos parâmetros é possível através de análises gráficas da RLEL entre as constantes de equilíbrio e de velocidade, conforme apontado pelo próprio Hammett em 1935 [10].

Para substituintes iguais nas posições *meta* e *para*, os valores de σ são diferentes. Esta característica pode ser explicada pela variação da intensidade das contribuições

indutiva e de ressonância de um dado substituinte em função da posição do substituinte na molécula. Por exemplo, um grupo nitro ligado na posição *para* de um ácido benzóico terá um efeito de ressonância (R) mais intenso que na posição *meta*. Assim, R irá se somar à contribuição indutiva do substituinte, a qual é aproximadamente a mesma para ambas as posições. Deste modo, o valor de σ_p será necessariamente maior que o de σ_m neste e em casos semelhantes. A aplicação da equação de Hammett a sistemas em que a introdução do substituinte se faz em ponto distante do centro de reação, como em sistemas aromáticos *meta*- ou *para*-substituídos, considera apenas a atuação dos efeitos eletrônicos e é possível detalhar-se bem as relações entre estrutura química e reatividade. Quando, porém, o grupo substituinte é introduzido próximo ao centro de reação, como em sistemas aromáticos *orto*-substituídos, podem operar outros efeitos além dos eletrônicos usuais, como efeito estérico, efeitos de campo, ligações de hidrogênio e outros. Foram propostos valores das constantes sigma *orto*, simbolizadas por σ_o , para substituintes nessa posição do anel benzênico. A determinação desses valores baseou-se em diferentes séries de reações, porém os resultados não se mostraram satisfatórios uma vez que os valores de σ_o para dado substituinte varia enormemente com a reação estudada [9].

Hansch & Fujita em 1964, estudando uma série de derivados de ácidos fenoxiacéticos concluíram que modelos matemáticos lineares obtidos a partir de estudos de QSAR, que combinavam parâmetros físico-químicos múltiplos relativos aos efeitos eletrônicos, hidrofobicidade e estereoquímica, eram fundamentais para descrever a atividade biológica. Neste estudo, os autores apresentaram a formulação do descritor estrutural da hidrofobicidade (π) para grupos substituintes, e o seu emprego em equações multiparamétricas que descrevem a resposta biológica em função de suas propriedades físico-químicas. Este estudo, combinado com a aplicação de regressão múltipla, permitiu identificar as relações entre atividade biológica e estrutura química de compostos bioativos, expressas por propriedades físico-químicas de compostos análogos [3, 11,12].

A hidrofobicidade é a propriedade físico-química que desperta o maior interesse nos estudos QSAR. Este fato ocorre porque a mesma regula tanto a solubilidade dos fármacos, com os processos de absorção, distribuição nos diversos compartimentos biológicos, tempo de meia-vida e excreção [4]. Em conseqüência, sugere-se que a hidrofobicidade interfere na habilidade do fármaco de ação específica em atingir e ligar-se ao sítio receptor, interação que é responsável pela resposta farmacológica e influencia a intensidade da atividade biológica [4,13].

A compreensão desses processos envolve o conhecimento da estrutura, função e propriedade das membranas biológicas e engloba, também, as propriedades físico-químicas como o grau de ionização, em função do pH e lipofilicidade relativa das formas neutra e ionizada das moléculas estudadas [3,5].

Os fármacos são em sua maioria, ácidos ou bases fracas, e quando presentes em solução se apresentam em situação de equilíbrio químico entre a forma não dissociada (neutra) e sua forma ionizada. A forma neutra é geralmente mais lipossolúvel e se difunde através das membranas biológicas por processo passivo, o qual ocorre em função do gradiente de concentração do sistema, do coeficiente de partição e do volume molecular. A forma ionizada, no entanto, somente é absorvida através de transporte ativo que é grandemente influenciada pelo pKa do composto e pH do meio biológico [3,5,7].

Entre os descritores estruturais que expressam a hidrofobicidade destacam-se também o coeficiente de partição ($\log P$), a constante π de Hansch e a constante f de Rekker [12,14].

Nesta mesma linha de trabalho, em estudos posteriores de QSAR, envolvendo os fenômenos de transporte e distribuição de fármacos em sistemas biológicos, Hansch & Clayton, em 1973, descreveram que a hidrofobicidade apresenta relação não linear com a atividade biológica. Hansch atribuiu esse comportamento não linear ao processo de transporte e distribuição dos compostos através das múltiplas biofases [12,15], ou seja, compostos muito hidrofóbicos (por exemplo, compostos com valores de $\log P$ maiores do que 3) tenderiam a permanecer nas membranas biológicas, enquanto que compostos pouco hidrofóbicos (por exemplo compostos com valores de $\log P$ menores que zero) tenderiam a permanecer no meio aquoso, não sendo capazes de cruzar as membranas no processo de transporte. Esta análise fundamenta-se na idéia de que um fármaco deve apresentar valor de hidrofobicidade “ótimo” para conseguir transpor as membranas biológicas e interagir com o sítio receptor (biomacromolécula alvo) resultando na resposta farmacológica.

Na década de 80, com o advento de programas e computadores cada vez mais rápidos, a química teórica passou a auxiliar de forma mais sistemática a química medicinal, através da modelagem molecular, que consiste em um conjunto de ferramentas computacionais que são utilizadas na compreensão de fatores estruturais de uma ou conjunto de moléculas, do mecanismo de ação e, também, para a obtenção de dados e informações necessárias para o estudo das relações entre estrutura química e a atividade biológica [16,17].

Entre os elementos perceptíveis nos estudos de modelagem molecular, alguns podem ser ressaltados, como a flexibilidade e o arranjo conformacional, propriedades estereo-eletrônicas e valores de energia livre (ΔG) e cargas atômicas que influenciam a interação fármaco-receptor [5,16].

A maioria das moléculas bioativas apresenta algum grau de liberdade conformacional, sendo que o arranjo conformacional estável destas moléculas pode mudar de acordo com fatores relacionados ao ambiente onde estas se encontram [13]. Entre as características estruturais estão a torsão angular, o volume e as distâncias intramoleculares, além da conformação de menor energia. As características eletrônicas se concentram no efeito indutivo e mesomérico, cargas dos átomos, valores de energia livre e de orbitais de fronteira HOMO (*highest unoccupied molecular orbital*) e LUMO (*lowest occupied molecular orbital*), da molécula [16,17].

Segundo Silva, em 2003, a conformação molecular é determinada através de técnicas de modelagem molecular e analisam a flexibilidade e o arranjo espacial não idêntico em uma dada molécula e o mesmo é resultante de rotações em torno de uma ou mais ligações simples.

A análise conformacional envolve métodos computacionais de mecânica quântica e mecânica molecular, e é dependente de fatores como ângulo das ligações, energia torsional, distâncias de ligações e barreira de energia de rotação interna da molécula [16,18]. Cada conformação calculada e obtida apresenta valores de energia termodinamicamente singulares, onde a estrutura que apresenta menor valor de energia é considerada aquela de maior estabilidade no ambiente estudado. Desse modo, algumas características estruturais consideradas importantes para a atividade biológica como a torsão angular, volume e distância intramoleculares, são normalmente determinadas a partir da conformação de menor energia obtida pela avaliação conformacional realizada [19].

Outra aplicação determinante da modelagem molecular envolve o estudo de características eletrônicas de compostos bioativos potenciais. Nesse sentido ressalta-se que os valores de energia das propriedades eletrônicas são importantes para a compreensão dos processos que envolvem e regem a interação fármaco-receptor e a ação farmacológica. Como já citadas anteriormente, as propriedades eletrônicas abrangem a densidade eletrônica e efeito mesomérico, cargas dos átomos, valores termodinâmicos de energia livre, valores de energia dos orbitais de fronteira HOMO e LUMO da molécula.

A análise dos valores de energia dos orbitais HOMO e LUMO possibilitam a avaliação e até a quantificação de propriedades eletrônicas importantes para o entendimento da reatividade da molécula e, do mesmo modo, desempenho de sua bioatividade.

Diante do exposto nota-se que com o avanço computacional, muitas ferramentas foram geradas e passaram a ser amplamente utilizadas para construir representações e modelos de estrutura molecular, bem como o cálculo de diversas propriedades moleculares [20, 21, 22]. O comportamento e as características de uma determinada molécula ou conjunto das mesmas podem ser simulados através de ferramentas quânticas e clássicas [23,24], tendo a química quântica um caráter especial na obtenção de descritores que podem a princípio expressar as propriedades energéticas, geométricas e eletrônicas de uma molécula, e suas interações com o ambiente. Estas simulações fornecem varias propriedades que podem ser úteis na química medicinal, pois podem ajudar a explicar conexões ou relações entre a estrutura e a atividade biológica [24]. Desta forma, tais metodologias se apresentam como ferramenta relevante no processo de planejamento e desenvolvimento de novos agentes terapêuticos.

1.2. As Ftalimidas

As imidas cíclicas (figura 1) podem ser divididos em sub-classes, incluindo as maleimidas, succinimidas, glutaramidas, ftalimidas, naftalimidas, etc., e seus respectivos derivados.

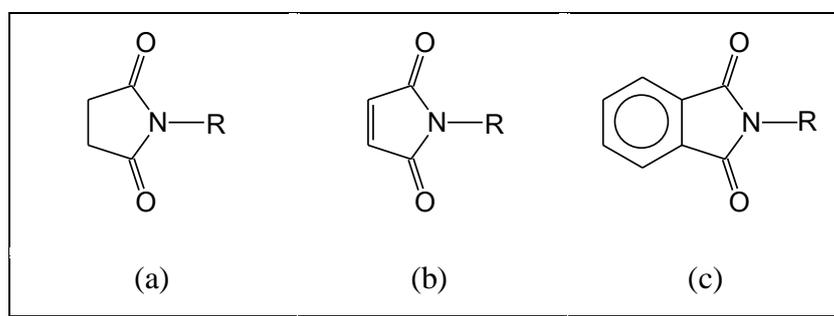


Figura 1: Alguns exemplos de imidas cíclicas: (a) Succinimida; (b) maleimida; (c) ftalimida

As ftalimidas, em particular constituem uma classe importante de compostos heteroaromáticos por apresentarem diferentes atividades biológicas, tais como: Atividade hipolipidêmica [25,26], hipotensiva [27], antitumoral [28], antiviral [29], antimicrobiana [30], anticonvulsiva dentre outras.

Como exemplo, podemos citar o caso da talidomida, que apesar dos significativos efeitos adversos do passado, ocasionando praticamente 100% de teratogenicidade, mesmo em doses clínicas modestas, sendo sua indicação específica para uso na gravidez [1]. Os recentes estudos têm evidenciado um possível uso desta substância para o tratamento de várias patologias incluindo o câncer.

1.2.1. Atividade Hipolipidêmica das Ftalimidas

Hiperlipidemias, ou seja, aumento dos níveis de colesterol (hipercolesterolemia) e triglicerídeos (hipertrigliceridemia) circulantes no plasma sanguíneo estão entre os fatores que mais ocasionam riscos no desenvolvimento de doenças das artérias coronárias (CAD), levando a morbidade e mortalidade em milhares de indivíduos em todo o mundo [31]. A aterosclerose é a principal causadora de enfarte do miocárdio, a forma mais freqüente de doença cardíaca.

Diversos fatores são apontados como causadores da aterosclerose, como hipertensão e hábito de fumar. Outros fatores causadores são: hereditariedade, baixa capacidade vital, dieta alimentar, falta de atividade física, obesidade, diabetes, dentre outros. Entretanto, a aterosclerose é geralmente associada a lipídios e lipoproteínas e, especialmente ao colesterol. Possuindo uma estrutura molecular que as células do organismo são incapazes de quebrar, o colesterol deve ser eliminado, sob pena de se acumular na parede das artérias. Para evitar esse processo, só há duas saídas: interromper a síntese de colesterol pelo organismo ou aumentar sua eliminação.

De modo geral, os fármacos atualmente em uso reduzem o colesterol circulante das seguintes maneiras: favorecendo a retenção, pelo fígado, das lipoproteínas que o transportam, ou bloqueando a produção de lipoproteínas ou ainda combinando essas ações. Certos fármacos atuam de maneira curiosa: incrementam a síntese de lipoproteínas, mas, por outro lado retiram-nas na circulação, e com tal intensidade que o efeito final é que elas se reduzem, provocando assim a redução do colesterol circulante. A resina colestiramina é um exemplo típico de fármaco que exerce esse tipo de ação: fixa os ácidos biliares no

intestino, fazendo aumentar a excreção fecal [32,33]. Como esses ácidos são produzidos pelo metabolismo do colesterol, reduzem-se assim os níveis sanguíneos deste.

O efeito hipolipidêmico da ftalimida e ftalimidas N-substituídas (figura 2) foi primeiramente estudada em 1979 por Chapman *et al.* [34], onde relataram que esta nova classe de compostos era efetiva na redução dos níveis de colesterol e triglicérides em roedores, em doses relativamente baixas de 20 mg/Kg/dia após 16 dias de administração da droga.

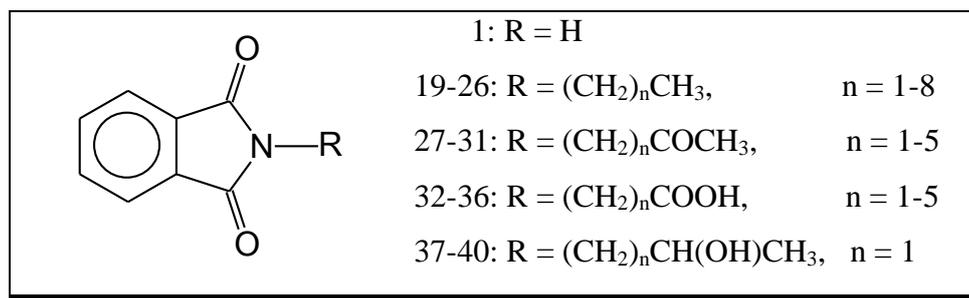


Figura 2: Ftalimidas e seus derivados

Nessa série avaliada, os compostos que apresentaram uma maior redução nos níveis de colesterol e triglicérides nos camudongos foram: *N*-butilftalimida, com 46 e 18%; 1-*N*-ftalimidobutan-3-ona, com 37 e 42% e o ácido 3-*N*-ftalimidopropiônico com 45 e 42% respectivamente.

Em 1983, Hall *et al.* [35,36] examinando uma série de derivados da *N*-fenilftalimidas (figura 3), verificou que a substituição no anel fenílico afeta drasticamente a atividade hipolipidêmica da ftalimida. Por exemplo, quando um substituinte COCH₃ é introduzido na posição *orto* do anel fenila, os níveis de colesterol e triglicérides são reduzidos em 57% e 44% após 16 e 14 dias de administração da droga.

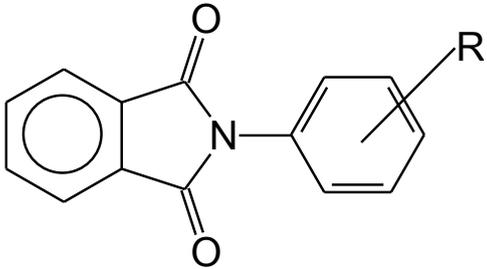
	Substituinte (R)	Redução (%)	
		%C ₁₆	%T ₁₄
 <p>%C₁₆ – Colesterol 16 dias %T₁₄ – Triglicerídeos 14 dias</p>	1: H	43	39
	2: <i>o</i> -COCH ₃	57	44
	3: <i>m</i> -COCH ₃	16	26
	4: <i>p</i> -COCH ₃	45	34
	5: <i>o</i> -C ₂ H ₅	3	8
	6: <i>m</i> -C ₂ H ₅	18	5
	7: <i>p</i> -C ₂ H ₅	11	4
	8: <i>o</i> -COOH	23	21
	8: <i>m</i> -COOH	31	29
	10: <i>p</i> -COOH	47	42

Figura 3: N-fenilftalimida e seus derivados

Os estudos referidos por Hall *et al.* [37], levaram Ramos e Neto [38] à realizarem um primeiro estudo das relações quantitativas estrutura-atividade (QSAR) em oito derivados de ftalimidas (figura 4), e foi observado que pequenas alterações estruturais na ftalimida, levavam a mudanças drásticas na atividade do composto, como por exemplo, a troca de uma das carbonilas da ftalimida por um grupo NH faz com que a redução do nível de colesterol, em 16 dias, caia de 43% para 20%. O mesmo foi observado para a redução do nível de triglicerídeos, em 14 dias, caindo de 56% para 25%. Neste trabalho foi utilizado o método computacional semi-empírico AM1 na descrição de parâmetros moleculares tais como energia dos orbitais HOMO e LUMO, cargas atômicas e momento de dipolo. A partir de um ajuste linear dos dados (utilizando a RLM), Ramos e Neto, encontraram que a redução do nível de colesterol era diretamente dependente da energia do LUMO, enquanto que este último mais à carga total da carbonila descrevia a redução dos níveis de triglicerídeos. Dois novos compostos mais ativos foram propostos, apesar do pequeno número de compostos analisados.

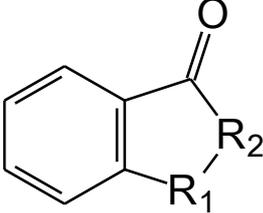
	Substituinte		Redução (%)		
	R1	R2	%C ₁₆	%T ₁₄	
	1. CO	NH	43	56	
	2. CH ₂	NH	13	22	
	3. CNH	NH	44	41	
	4. NH	NH	20	25	
	5. CO	CH(C ₂ H ₅)	40	37	
	%C ₁₆ – Colesterol 16 dias	6. CO	CH ₂	40	21
	%T ₁₄ – Triglicerídeos 14 dias	7. CNH	CH ₂	38	41
		8. CO	CHCH ₃	42	39

Figura 4: Oito derivados de ftalimidas

Motivados por esse estudo anterior, Ramos *et al.* [39] estenderam o estudo de QSAR para as *N*-fenilftalimidas e seus derivados que embora sendo estruturas mais complexas, exigiu uma análise conformacional detalhada, envolvendo barreira de rotação, e obtenção de parâmetros moleculares como: energia de orbitais de fronteiras HOMO e LUMO, momentos de dipolo e cargas atômicas além de parâmetros empíricos. A partir da utilização de uma regressão linear múltipla e regressão de mínimos quadrados parciais (PLS, *do inglês Partial Least Square*), encontraram que a redução do nível de colesterol e triglicerídeos era diretamente dependente do descritor π , e da carga do carbono fenílico que está diretamente ligado a ftalimida.

Em 1985, Hall *et al.* [40] observaram que a substituição do grupo NH da ftalimida por um grupo CH₂ (figura 5), para formar o composto indan-1,3-diona, causa uma mudança significativa na atividade hipolipidêmica, ou seja, a redução no nível de colesterol sofre uma queda de 43% para 40%, enquanto a redução no nível de triglicérides sofre uma redução de 56%, para 21%.

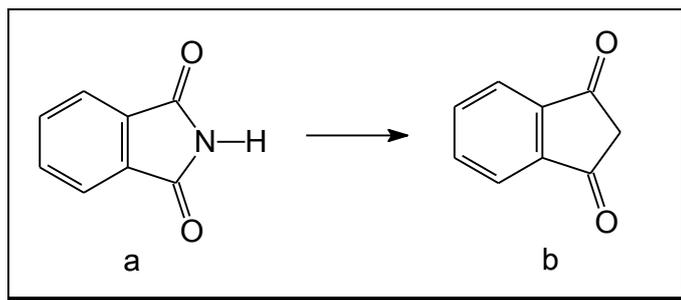


Figura 5: (a) Ftalimida e (b) Indan-1,3-diona.

Foi também observado, que a introdução de um grupo fenila (figura 6), em substituição a um hidrogênio no metileno na indan-1,3-diona leva a outra importante mudança na atividade hipolipidêmica, no que diz respeito aos níveis de triglicerídeos, ou seja, a redução do nível de triglicerídeos sofre um aumento de 21% para 40%, enquanto que a redução dos níveis de colesterol permanece inalterada.

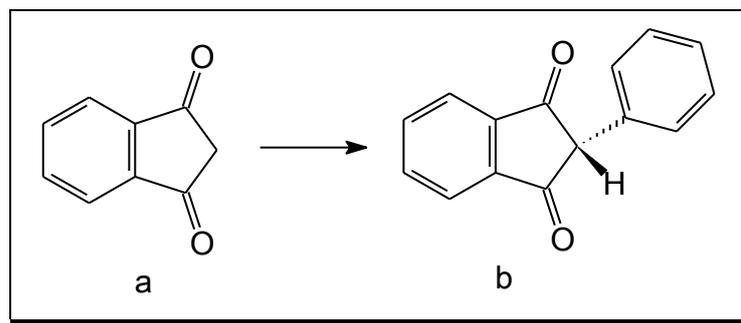


Figura 6: (a) Indan-1,3-diona e (b) 2-fenilindan-1,3-dionas.

Com base no exposto acima, o alvo deste trabalho se encontra nos derivados de 2-fenilindan-1,3-dionas e sua atividade hipolipidêmica (figura 7). Nesse sentido, um estudo desenvolvido ainda em 1985 mostrou a atividade hipolipidêmica de alguns derivados de 2-fenilindan-1,3-dionas. Hall *et al* [40] obtiveram os valores experimentais da redução dos níveis de colesterol e triglicerídeos sanguíneo de roedores.

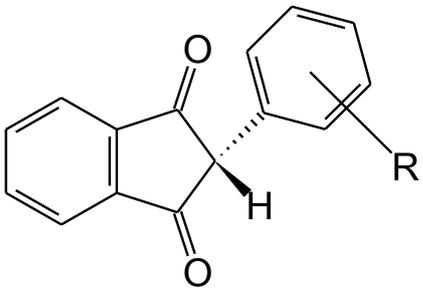
	Substituinte		Redução (%)	
	(R)	%C ₁₆	%T ₁₆	
 <p>%C₁₆ – Colesterol 16 dias %T₁₆ – Triglicerídeos 16 dias</p>	1: H	40	40	
	2: <i>o</i> - OCH ₃	28	52	
	3: <i>m</i> - OCH ₃	33	45	
	4: <i>p</i> - OCH ₃	41	58	
	5: <i>o</i> - CH ₃	31	51	
	6: <i>m</i> - CH ₃	29	49	
	7: <i>p</i> - CH ₃	35	58	
	8: <i>o</i> - Cl	44	45	
	9: <i>m</i> - Cl	40	30	
	10: <i>p</i> - Cl	37	53	
	11: <i>o</i> - COOH	22	48	
	12: <i>p</i> - COOH	22	05	
	13: <i>p</i> - OCH ₂ CH ₃	41	19	

Figura 7: 13 derivados de 2-fenilindan-1,3-dionas.

Também foi observado que a intensidade da atividade hipolipidêmica, por sua vez, depende não só da natureza do substituinte ligado ao anel fenílico, como também de sua posição. Por exemplo, a figura 8 representa respectivamente as reduções dos níveis de colesterol e triglicerídeos dos derivados *orto*-cloro, *meta*-cloro e *para*-cloro. A figura 8 mostra, por exemplo, que quando muda-se a posição do substituinte da posição *orto* para a *meta* e depois para a posição *para* o nível de colesterol sofre duas leves reduções, enquanto que no caso dos triglicerídeos primeiro ocorre uma queda significativa na redução dos níveis de triglicerídeos e logo depois um aumento.

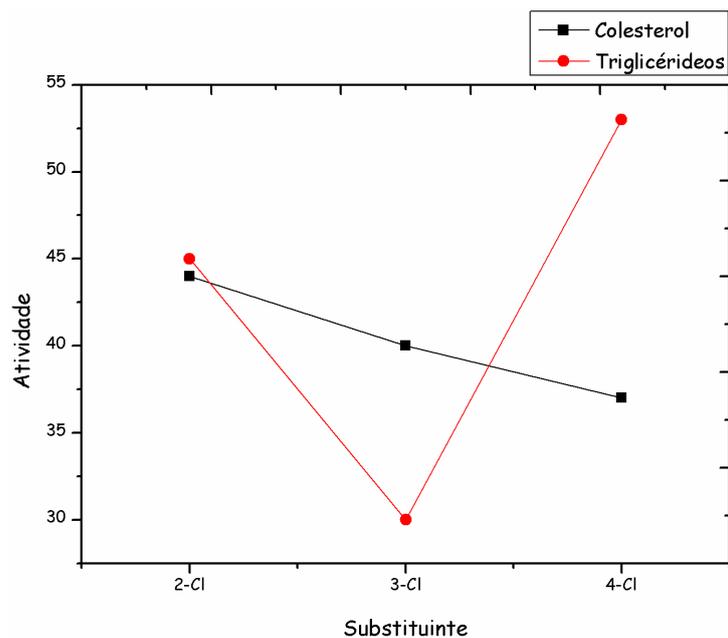


Figura 8: Gráfico mostrando como varia a redução dos níveis de colesterol e triglicerídeos de três derivados da série homóloga em estudo, em função da posição do anel.

A natureza do substituinte, como dito anteriormente, pode causar também mudanças significativas nas atividades hipolipidêmicas das 2-fenilindan-1,3-dionas, como pode ser vista na figura 9:

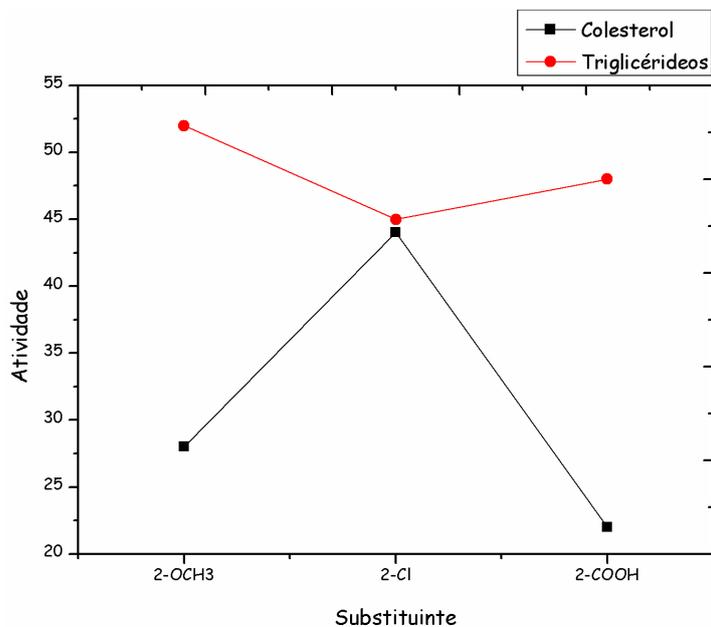


Figura 9: Gráficos mostrando como varia a redução dos níveis de colesterol e triglicerídeos mudando a natureza dos substituintes na posição *orto*.

A substituição do grupo metóxi por cloro na posição *orto*, produz um aumento na redução dos níveis de colesterol e uma diminuição nos níveis de triglicérides, mas a substituição do cloro por um grupo carbóxi causa uma queda na redução dos níveis de colesterol e um leve aumento na redução dos níveis de triglicérides.

Estudos anteriores feitos pelo grupo do professor Mozart N. Ramos, de relação estrutura-atividade em 13 derivados de 2-fenilindan-1,3-dionas, utilizando para isso cálculos de orbitais moléculares semi-empíricos (MNDO, CNDO/2) e de alguns parâmetros clássicos como π e a superdeslocalizabilidade nucleofílica [41], levou a conclusão que a redução do nível de colesterol era diretamente dependente do parâmetro π , e do quadrado do parâmetro π mais a superdeslocalizabilidade nucleofílica de um carbono.

Apesar de já haver um estudo de QSAR envolvendo as 2-fenilindan-1,3-dionas e sua atividade hipolipidêmica, este estudo utilizou metodologias limitadas aos recursos computacionais da época, por essas razões começamos a imaginar um conjunto de descritores moleculares mais vastos, calculados sob a ótica da metodologia semi-empírica AM1 (*Austin Model 1*) e DFT (do inglês, *Density Functional Theory*) e que incluíssem efeitos de solvatação, talvez, fosse obtida uma função resposta com uma interpretação química dos parâmetros mais bem fundamentada.

2. Objetivos

2.1 Objetivos gerais

Este trabalho tem como objetivo principal determinar correlações quantitativas entre os parâmetros eletrônicos e empíricos com as atividades hipolipidêmicas de 13 derivados de 2-fenilindan-1,3-diona. Como medida da atividade “in vivo” foi utilizada as taxas dos níveis de colesterol e triglicerídeos após 16 dias de administração da droga, em doses relativamente baixas de 20mg/Kg/dia intraperitoneamente em roedores.

2.2 Objetivos Específicos

- Estudar e compreender a relação quantitativa estrutura – atividade em 13 derivados de 2-fenilindan-1,3-diona;
- Obter descritores físico-químicos quânticos, oriundos dos cálculos computacionais, utilizando para isso cálculos de orbitais moleculares AM1 e B3LYP/6-31G (d,p), utilizando os modelos da molécula isolada e do efeito solvente.
- Coletar descritores clássicos, extraídos de dados previamente tabelados desses compostos.
- Desenvolver um estudo quimiométrico usando a técnica de regressão linear múltipla (RLM) e análise de componentes principais (PCA); avaliar as regressões obtidas, a fim de obter uma equação apropriada capaz de descrever matematicamente a dependência funcional da atividade com a estrutura molecular da série homóloga estudada.
- Analisar a equação de QSAR obtida através dos métodos quimiométricos citados e avaliar a contribuição das variáveis na atividade para a proposição de um novo composto, cujas características lhe confirmam ao nível teórico um caráter hipolipidêmico maior que as demais moléculas da série;

3. Fundamentação Teórica

3.1. Técnicas Computacionais

Hoje em dia os métodos de química quântica são atualmente utilizados na compreensão e na previsão de propriedades moleculares. Em todas as técnicas utilizadas para a descrição da estrutura eletrônica em química o objetivo principal é resolver, de maneira exata ou aproximada, a equação de Schrödinger [42] (equação 3.1):

$$\hat{H}\Psi = \hat{E}\Psi \quad (3.1)$$

onde \hat{H} é o operador hamiltoniano que contém as contribuições de energia cinética e potencial para os núcleos e elétrons do sistema e Ψ é a função de onda do sistema que depende das coordenadas espaciais e de spin de todas as partículas (elétrons e núcleos) e \hat{E} corresponde a energia total. Esta é uma equação de autovalores e autovetores e sua resolução exata está restrita a sistemas monoelétrônicos. O operador Hamiltoniano é formado pelos operadores de energia cinética eletrônico, de energia cinética nuclear e de energia potencial ou coulombianos que descrevem as atrações eletrostáticas elétron-núcleo e as repulsões eletrostáticas núcleo-núcleo e elétron-elétron, sendo, portanto, necessário realizar algumas aproximações para resolver sistemas multieletrônicos.

Uma das aproximações mais utilizadas é a aproximação de Born-oppenheimer (BOA, *Born-Oppenheimer Approximation*) [43], que consiste em separar os movimentos eletrônico e nuclear do sistema, baseado na idéia de que o movimento dos elétrons é bem mais rápido que o movimento dos núcleos devido a diferença entre as massas dos mesmos. Com a aproximação o problema de determinação da energia molecular fica dividido em dois: um problema eletrônico e outro nuclear.

O problema eletrônico é resolvido para uma configuração fixa dos núcleos, dando origem a superfície de energia potencial que em seguida pode ser utilizada para resolver as equações de movimentos nucleares. A energia total da molécula, para uma dada configuração nuclear, é obtida como a soma da energia da parte eletrônica e da repulsão nuclear já que a contribuição proveniente da energia cinética nuclear é considerada nula, pois está incluída de maneira paramétrica no movimento eletrônico.

A resolução da parte eletrônica da equação de Schrödinger fornece as descrições das vibrações, rotações e translações da molécula.

A BOA é uma aproximação adequada para a grande maioria dos sistemas químicos, no entanto a mesma costuma falhar em sistemas com os estados eletrônicos ou vibracional-eletrônico degenerados.

3.2. O Método de Hartree-Fock

Como foi dito acima a equação de Schrödinger tem solução exata apenas para átomos monoelétrônicos. O motivo para essa limitação na resolução da equação se deve ao termo de repulsão eletrônica, que não permite tornar a equação separável como no caso de sistemas com apenas um elétron.

Para resolver esse problema Hartree propôs um método que aproxima a função de onda atômica como o produto de funções espaciais de um elétron, onde as melhores formas para esses orbitais são encontradas num processo iterativo de cálculo na qual se assume que cada elétron move-se num campo produzido pelo núcleo e por uma nuvem eletrônica hipotética gerada pelos outros elétrons. Esse método denominado modelo de partículas independentes consegue transformar a equação de Schrödinger dependente de n elétrons em n equações de um elétron.

Nesse sentido Hartree propôs que um procedimento variacional fornece os melhores valores para essas funções espaciais de um elétron. Logo em seguida, Fock generalizou as equações de Hartree para um produto de funções de spin orbitais de um elétron antissimetrizadas (determinante de Slater). Esse é o método denominado de método Hartree-Fock (HF) [44].

As equações obtidas são do tipo íntegro-diferenciais cujas soluções exigem alto custo numérico e de difícil programação. Então Roothaan [44] propôs a expansão das soluções das equações de HF num conjunto de funções de bases (orbitais atômicos), tomando as equações íntegro-diferenciais num problema matricial. Agora a função de onda total que resolve a equação de Schrödinger é substituída por orbitais moleculares que são as funções de onda para os elétrons individuais. Esses orbitais moleculares são expandidos numa combinação linear de funções de orbitais atômicos, onde esses orbitais atômicos do sistema são representados por funções ϕ e os orbitais moleculares escritos como uma combinação linear de orbitais atômicos.

Contudo a função de onda HF não é uma representação completa da estrutura eletrônica de uma molécula, devido ao fato da interação, ou correlação efetiva, entre elétrons não acontecer por meio de um potencial médio. Este fato impõe um limite máximo de aprimoramento no valor da energia, mesmo que se utilize a melhor função de onda HF. A diferença entre a energia exata e a energia HF é conhecida como a energia de correlação. Apesar de pequena, a inclusão da mesma é de fundamental importância para descrever adequadamente diversas propriedades moleculares.

A inclusão da correlação eletrônica pode ser feita por vários métodos e consiste, essencialmente, na descrição e resolução do termo de repulsão direta elétron-elétron. Métodos pós-HF são os mais comuns para se fazer isso, em sua maioria são baseados na teoria de perturbação de muitos corpos (MBPT, do inglês, *Many Body Perturbation Theory*), nos quais encontram-se os métodos de Møller-Plesset (MP n , aqui n é a ordem da correção), de *Coupled-Cluster* (CC) e de integração de configurações (CI, do inglês, *Configuration Interaction*). Além disso, também é possível utilizar métodos baseados na teoria do funcional densidade (DFT, do inglês *Density Functional Theory*), os quais vêm sendo largamente utilizados devido aos seus relativos baixos custos computacionais, quando comparados com os métodos pós-HF, precisão química aceitável, e, assim como os métodos MP n , são extensivos.

3.3. Métodos *ab initio*

Os métodos *ab initio* tem por objetivo a determinação das auto-funções associadas ao Hamiltoniano fundamental, em qualquer parametrização [45], sendo por isso considerado teoricamente puro. Os cálculos só usam as quantidades bem estabelecidas, como, por exemplo, constantes fundamentais e números atômicos dos núcleos [45]. A maior vantagem deste método é o alto nível de confiabilidade dos resultados obtidos. Infelizmente, apresenta também a seguinte desvantagem: os cálculos numéricos baseados neste método quase sempre levam muito tempo, que cresce com n^4 , onde n é o número de funções atômicas de base, e consomem muita memória dos computadores.

Na determinação da energia e da função de onda do estado fundamental de um sistema através de cálculos *ab initio*, ou de “primeiros princípios”, resolve-se um grande número de integrais de dois corpos, que trazem um alto custo computacional para sistemas moleculares grandes, ou seja, acima de 20 átomos. Para contornar essas dificuldades,

foram feitas várias aproximações e parametrizações subseqüentes, para a determinação dos elementos da matriz de Fock dando origem dessa forma aos chamados métodos semi-empíricos. A importância desses últimos está relacionada à capacidade de prever as propriedades físicas e químicas de um sistema, respondendo satisfatoriamente aos parâmetros experimentais a serem avaliados.

3.4. Métodos semi-empíricos

Como o próprio nome sugere, métodos semi-empíricos são aqueles que se valem da parametrização a partir de dados empíricos, portanto, diferentemente dos métodos *ab initio*, a exatidão de seus resultados dependem dos dados experimentais fornecidos. Por utilizar dados de origem experimental há um ganho na velocidade dos cálculos, o que os tornam mais vantajosos em aplicações de sistemas mais complexos. O potencial dos métodos semi-empíricos reside não no rigor teórico, mas no fato de que parâmetros ajustáveis são otimizados de forma a reproduzir propriedades químicas importantes.

O objetivo fundamental dos métodos semi-empíricos é o desenvolvimento de um tratamento quantitativo de propriedades moleculares com precisão, confiabilidade e custo computacional suficiente para ser de valor prático em química, em particular em áreas onde há falta de dados experimentais ou onde os procedimentos experimentais não funcionam satisfatoriamente.

Pode-se, então dizer, que os métodos semi-empíricos representam uma abordagem à solução da equação de Schrödinger para o sistema de muitos corpos (átomos, moléculas ou sólidos) usando a simulação numérica e computacional.

Em 1985 foi publicada uma nova versão do método MNDO, o método AM1[46]. Neste método foram parametrizadas as integrais dos seguintes átomos: H, B, C, Si, N, O, S, F, Cl, Br, I, Hg e Zn. Se comparar com o MNDO, o método AM1 calcula melhor a repulsão entre núcleos e proporciona melhores resultados e moléculas envolvidas em processos biológicos, corrigindo a tendência do método MNDO em superestimar as forças de repulsão interatômicas a curtas distâncias.

3.5. Teoria do Funcional da Densidade

Outro modo de obter a estrutura eletrônica de moléculas é através do uso da teoria do funcional da densidade, que surgiu como uma alternativa aos tradicionais métodos *ab initio* e semi-empíricos no estudo de propriedades do estado fundamental de sistemas moleculares. A grande vantagem do método DFT em comparação com os *ab initio* está no ganho de velocidade computacional e espaço em memória. Considerando-se um sistema arbitrário qualquer, modelado com n funções de base, o esforço computacional no estudo desse sistema utilizando-se o DFT aumenta na ordem de n^3 , enquanto que para os métodos Hartree-Fock aumenta em n^4 ou n^5 para técnicas de interação de configuração. Em contraste com métodos semi-empíricos, o hamiltoniano na DFT [47,48] é bem definido e suas características não são obscurecidas por freqüentes e injustificáveis aproximações no procedimento computacional. Por essa razão esse método, é útil no estudo de sistemas moleculares com muitos átomos, e tem sido cada vez mais empregada para sistemas moleculares encontrados em pesquisas farmacológicas, agroquímicas, complexos organometálicos dentre outros, ou seja, em sistemas que estão fora do alcance de métodos de natureza *ab initio*.

A tentativa de se utilizar a densidade eletrônica como variável básica iniciou-se com Drude [49] em 1900, porém foi somente em 1964, com a publicação de dois teoremas por Hohenberg e Kohn, que o uso da densidade eletrônica como variável básica foi legitimado. Nesses teoremas Hohenberg e Kohn [48] forneceram os fundamentos da teoria do funcional de densidade moderna e mostraram que o método deve ser considerado como uma aproximação de uma teoria exata. Estes dois teoremas mostram que a energia total de um sistema eletrônico é um funcional exato da densidade eletrônica, $E[\rho]$, e um princípio variacional exato para este funcional. Então, em 1965, Kohn e Shan propuseram uma forma de fazer cálculo usando a DFT.

Os cálculos DFT têm a vantagem de incluir naturalmente os efeitos de correlação eletrônica que estavam ausentes no método Hartree-Fock, porém não podem ser usados para cálculos de estados excitados das moléculas, para tanto é necessário a aplicação de teorias de perturbação aos sistemas multieletrônicos.

Nesse trabalho, além da aplicação de cálculos de orbitais moleculares semi-empíricos AM1, também foi utilizado o funcional híbrido B3LYP, utilizando o conjunto de funções de base 6-31G(d, p), tanto no método da molécula isolada, como utilizando o

efeito solvente, pois, os mesmos apresentam características importantes, para o trabalho desenvolvido nesta dissertação.

3.6. Efeito do Solvente

3.6.1 Modelos Teóricos

Ao analisarmos a química e a bioquímica, podemos perceber que parte considerável das reações envolvidas ocorre em fase líquida. Há muito se sabe da importância do solvente na condução de reações químicas, o qual afeta a velocidade das reações e determina a formação dos produtos [50,51]. Desta forma, modelos teóricos capazes de incluir o efeito que o meio exerce nas reações são de grande valor.

Os modelos teóricos do solvente podem ser classificados em três categorias: Modelo do contínuo dielétrico, Simulação de líquidos com potenciais obtidos da combinação mecânica quântica/mecânica molecular (QM/MM) [52], e o modelos baseado na Teoria de equações integrais para líquidos [53]. Dentre esses modelos, podemos destacar o modelo em que o solvente é considerado um contínuo dielétrico e circunda o soluto [54]. Este é muito difundido ao se estudar reações em solução devido a sua praticidade, e por considerar várias propriedades importantes na interação soluto-solvente, como:

- ✓ A constante dielétrica;
- ✓ A distribuição de cargas do soluto;
- ✓ Tamanho e forma da cavidade.

3.6.2 Modelos Contínuos

A interação do soluto com o meio ocorre através de forças intermoleculares [54], como:

- ✓ Eletrostáticas;
- ✓ Dispersão
- ✓ Repulsão de Troca

Usualmente, as forças eletrostáticas são as mais importantes, enquanto os termos de dispersão e repulsão tendem a se cancelar. Iremos discutir somente a inclusão de forças eletrostáticas. Ao inserirmos um soluto A com carga negativa em um solvente X, as moléculas do solvente se orientam em torno do soluto. Isto pode ser visto na Figura 10, onde podemos verificar que os dipolos que representam as moléculas do solvente estão orientados eletrostaticamente em torno do soluto.

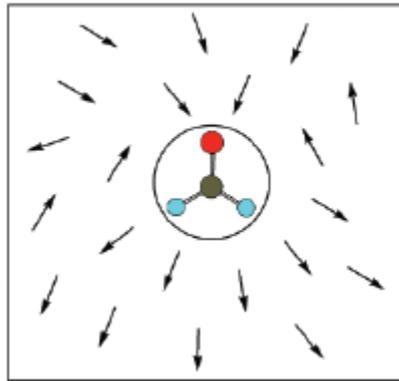


Figura 10: Molécula de um soluto solvatada por dipolos

No modelo contínuo, o solvente é considerado como constituído por um contínuo de dipolos pontuais. Cada ponto no espaço tem um vetor de polarização ($\vec{\rho}$) por unidade de volume, como mostra a equação 3.2:

$$\vec{\rho} = \frac{1}{\Delta V} \sum_i \langle \vec{\mu}_x \rangle \quad (3.2)$$

sendo ΔV o volume da região e μ_x , o momento de dipolo da molécula do solvente. Na aproximação de resposta linear, o vetor de polarização ($\vec{\rho}$) é representado pela equação:

$$\vec{\rho} = \chi \vec{E} \quad (3.3)$$

sendo χ uma constante que depende do meio e \vec{E} , o campo elétrico gerado pelo soluto e pela polarização do solvente.

Além da constante dielétrica do solvente e da distribuição de cargas no soluto, é necessário definir outras propriedades da interação soluto-solvente, como o tamanho e a

forma da cavidade do soluto. O tamanho da cavidade deve ser definido empiricamente e é geralmente baseado no desvio da energia livre de solvatação teórica e experimental. Na figura 11 temos as cavidades obtidas através de esferas centradas nos átomos para a molécula de um soluto qualquer. Nota-se que as esferas se sobrepõem de modo a fornecer as cavidades de forma realistas.

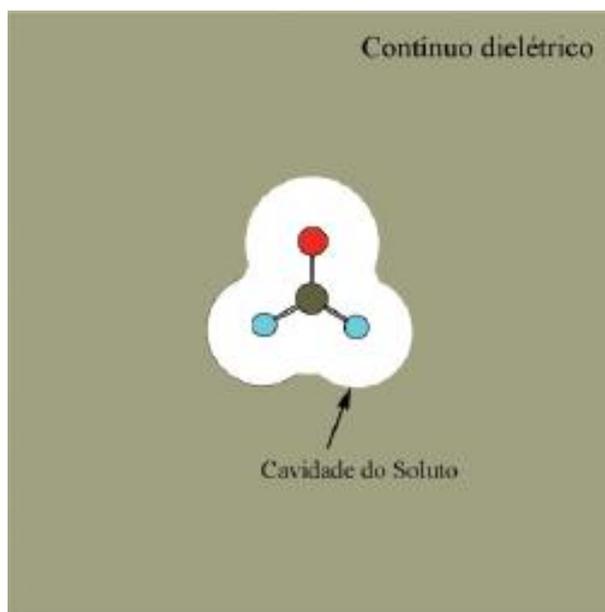


Figura 11: Molécula de um soluto imersa em uma cavidade de esferas sobrepostas.

Solvatação pelo contínuo dielétrico

Nesse trabalho, será empregado cálculos de orbitais moleculares semi-empírico AM1 utilizando o modelo COSMOS, e para o método B3YP/6-31G (d,p) será empregado o modelo PCM, como uma forma de entender o comportamento dessas moléculas num meio onde se tenha solvatação. O solvente utilizado em ambos os modelos foi a água.

3.7. Quimiometria Aplicada ao Estudo de QSAR

Nos diversos ramos da ciência, freqüentemente deseja-se estabelecer relações quantitativas entre um fenômeno observado e algumas variáveis independentes que se acredita ter relevância na explicação do fenômeno. Em outras palavras, deseja-se construir

um modelo matemático que seja capaz de explicar o fenômeno observado e que também seja capaz de proporcionar previsões dentro e, se possível, fora dos limites investigados.

Do ponto de vista estatístico, a RLM é um método de regressão que modela a relação entre a variável dependente Y , associada à resposta biológica dos compostos em teste, com as variáveis independentes X , correspondentes aos diferentes compostos testados e suas características e/ou propriedades.

Para a modelagem dos compostos estudados, assume-se que existe uma reta perfeita capaz de ajustar todos os valores conhecidos da variável dependente, que pode ser representada pela equação 3.4:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (3.4)$$

onde, $\beta_0, \beta_1, \dots, \beta_n$ são os parâmetros do modelo e ε é o erro aleatório relacionado à determinação de y [55, 56]. Um modelo linear de regressão fornecerá valores previstos de y para cada valor correspondente de x_i associado, através da obtenção do conjunto de parâmetros $b = (b_0, b_1, \dots, b_n)$ que mais se aproximam dos coeficientes $\beta_0, \beta_1, \dots, \beta_n$ correspondentes. Observe, no conjunto b , que n é o número total de variáveis. O resíduo deixado – e – corresponde à diferença entre os valores medidos (y) e os previstos (y_p), estes últimos modelados pelos coeficientes b . O objetivo da modelagem é prever y_p da melhor forma, deixando assim o menor valor possível de resíduos. Além disso, espera-se dos resíduos que não apresentem padrão algum quando apresentados em gráfico contra as variáveis x_i (gráfico dos resíduos), indicando que possuem caráter aleatório e, portanto, que o modelo estimado é adequado.

Espera-se ainda da relação entre y e y_p que a mesma seja linear e é importante ressaltar que o modelo é dito linear devido à linearidade da relação estabelecida entre as variáveis y e x_i . Ainda que uma ou mais variáveis seja do tipo x^2 , a linearidade é mantida, pois o modelo é função linear dos parâmetros b , e não das variáveis. Assim, um modelo que envolva y , x e x^2 numa RLM é dito ser linear com relação à x e x^2 , através de seus parâmetros, embora o gráfico de y contra as variáveis não seja linear.

Os parâmetros b são calculados de acordo com a equação matricial 3.5, na qual os termos estão em negrito para indicar que são matriciais:

$$\mathbf{b} = (\mathbf{X}^t \mathbf{X})^{-1} \cdot (\mathbf{X}^t \mathbf{y}) \quad (3.5)$$

onde \mathbf{X}^t é a matriz de dados transposta, $(\mathbf{X}^t\mathbf{X})^{-1}$ representa o inverso do produto da matriz de dados pela sua transposta e y é o vetor que contém o valor das atividades observadas.

De uma forma geral, é sempre possível reconhecer nos dados a seguinte organização: um parâmetro biológico aleatório y está correlacionado através de um conjunto de parâmetros lineares a um conjunto X de variáveis físico-químicas, de modo que $X = (x_1, x_2, \dots, x_n)$, onde n é o número total de variáveis. O objetivo de uma análise de QSAR é justamente encontrar os parâmetros b que possibilitem definir a função expressa pela equação 3.4, através da equação matricial 3.6 apresentada a seguir:

$$y_p = \mathbf{X} \cdot \mathbf{b} \quad (3.6)$$

A equação de QSAR obtida ao final do estudo permitirá, portanto, descobrir quantitativamente como as variáveis estão relacionadas à atividade, e a partir de sua interpretação será possível saber como modificá-las para obter melhores valores de atividade biológica, com a proposição de um ou mais compostos potencialmente ativos.

Após a construção do modelo de RLM, é necessário avaliar a qualidade do mesmo, isto é, verificar o seu poder de previsão bem como os resíduos deixados pelo modelo. Existem formas diferentes de avaliação, a maioria consistindo na quantificação do poder previsivo [57]. Uma delas consiste em retirar um grupo de compostos da modelagem e testar a capacidade de previsão da mesma com os compostos não incluídos, denominados grupo de teste. Isto deve ser efetuado após a obtenção da equação de QSAR com os demais compostos incluídos no modelo, denominados grupo de treinamento. Para que a separação dos compostos em grupo de teste e treinamento não interfira significativamente na qualidade do modelo é interessante que se disponha de um número significativo de compostos. Além disso, convém verificar que no grupo de teste não existam compostos que apresentem características muito diferenciadas dos demais, como consequência não teriam sua resposta adequadamente prevista pelo modelo, por se localizarem em regiões extremas de previsão, o que simultaneamente reduziria a qualidade do modelo, por não abranger um intervalo de dados tão extenso quanto possível.

Um meio de avaliação para um modelo estatístico, também bastante usado, é a análise da variância, denominada ANOVA (do inglês *Analysis of Variance*), baseada na distribuição de Fischer. A ANOVA avalia a soma quadrática dos desvios em torno da média dos valores de y , ou soma quadrática total (SQ_T), separando-se a mesma em soma

quadrática devida à regressão (SQ_R) e soma quadrática residual (SQ_T). Avaliam-se também as médias quadráticas dos desvios devidos à regressão (MQ_R) e aos resíduos (MQ_T). Por fim, obtém-se o coeficiente de variação, ou de determinação (R^2), que corresponde à divisão de SQ_R por SQ_T , a fim de conhecer a porcentagem de variação explicada pelo modelo desenvolvido. Como o valor de R^2 está obrigatoriamente compreendido entre 0 e 1, pois SQ_R é necessariamente menor que SQ_T , por definição, então quanto mais próximo da unidade for o valor de R^2 , melhor será o modelo.

A ANOVA, em geral não abrange o teste de Fischer, que é usualmente realizado em paralelo. Este teste consiste em obter o valor do índice r^2 , dividindo-se MQ_R por MQ_T e comparando-se com o valor tabelado correspondente aos graus de liberdade do problema. Caso o índice calculado seja maior que o tabelado, cerca de três vezes para um intervalo de confiança de 95%, então é possível afirmar que a regressão é significativa. Caso contrário, a significância da regressão não está estatisticamente garantida.

Outro método utilizado é a análise de componentes principais (PCA) que baseia-se na projeção linear do espaço original das variáveis X 's, que possui m dimensões (cada dimensão representando uma variável), num subespaço com k dimensões (cada dimensão desse subespaço representando uma componente principal) que preserve a maior variância possível do conjunto de dados. Em outras palavras, PCA é capaz de transformar dados complexos e apresentá-los numa nova perspectiva em que se espera que as informações mais importantes tornem-se evidentes. Na análise de componentes principais, a atividade biológica não é considerada diretamente, visto que PCA utiliza apenas as variáveis descritivas. PCA está fundamentada na correlação entre as variáveis, agrupando aquelas que são mais correlacionadas numa nova variável chamada de *componente principal*. As novas variáveis representadas pelas componentes principais são mais informativas e em menor número do que os descritores originais, isto é, $k < m$. As componentes principais são obtidas através da combinação linear dos descritores originais e apresentam as propriedades de serem mutuamente ortogonais e definidas em ordem decrescente da quantidade de variância que são capazes de explicar. Em outras palavras, a informação contida numa componente principal não está presente em outra e a variância que elas descrevem é uma medida da quantidade de informação que as mesmas contêm [58]. As relações entre os compostos não são alteradas nessa transformação. Porém, como os novos eixos são ordenados pela sua importância, ou seja, pela ordem de variância citada acima,

pode-se visualizar a estrutura do conjunto de dados em gráficos de baixa dimensionalidade (como por exemplo, PC1, PC2 etc.).

Do ponto de vista matemático a matriz de dados \mathbf{X} é decomposta em duas matrizes, uma de escores (\mathbf{T}) e uma de pesos (*loadings*) \mathbf{L}^T , ou seja, $\mathbf{X} = \mathbf{T} \mathbf{L}^T$. Os escores são as novas coordenadas de cada composto no novo sistema de eixos e a informação de quanto cada descritor original contribui, está contida nos pesos. Os escores \mathbf{T} expressam as relações entre as amostras enquanto que os pesos \mathbf{L}^T mostram as relações entre as variáveis.

No processo de seleção de variáveis, PCA identifica agrupamentos de descritores, proporcionando um entendimento de como eles estão correlacionados e o quanto de informação eles contém. Este é o fundamento de seu uso na seleção das variáveis. No entanto, frequentemente o número de descritores originais selecionados é maior que o desejado. Embora isso não constitua problema do ponto de vista matemático, corre-se o risco de tornar a interpretação físico-química difícil ou mesmo impossível. Portanto, deve haver equilíbrio entre o aspecto matemático e a possibilidade das correlações encontradas poderem ser explicadas através de hipóteses físico-químicas-biológicas. Os resultados da análise de componentes principais costumam ser visualizados em gráficos, facilitando a identificação de agrupamentos. Por exemplo, um gráfico de pesos contém informação sobre as variáveis e é usado para determinar quais delas são mais importantes para descrever a variação dos dados originais. Um gráfico de escores contém informação sobre os compostos tornando visível a similaridade, agrupamentos e diferenças entre os mesmos, com base nas variáveis utilizadas. Portanto, é importante que estes gráficos sejam analisados em conjunto.

O estudo da relação dos descritores físico-químicos com a atividade hipolipidêmica da série 2-fenilindan-1,3-dionas estudada foi realizado usando-se o método de regressão linear múltipla (RLM) e a análise de componentes principais (PCA).

4. Cálculos

4.1 Otimização da Geometria Molecular da Molécula-Mãe e seus Derivados

Conforme relatamos na seção anterior, as geometrias moleculares otimizadas e as propriedades eletrônicas da molécula-mãe e de seus derivados foram obtidas mediante cálculos de orbitais moleculares AM1 e B3LYP/6-31G(d,p) usando o programa GAUSSIAN 03. No processo de otimização seguimos a seguinte rotina: a geometria da molécula-mãe, mostrada a seguir, foi inicialmente otimizada, e em seguida, usamos esse resultado como de partida no processo de otimização das geometrias moleculares de seus derivados, tanto nos cálculos AM1 como aqueles B3LYP/6-31G(d,p).

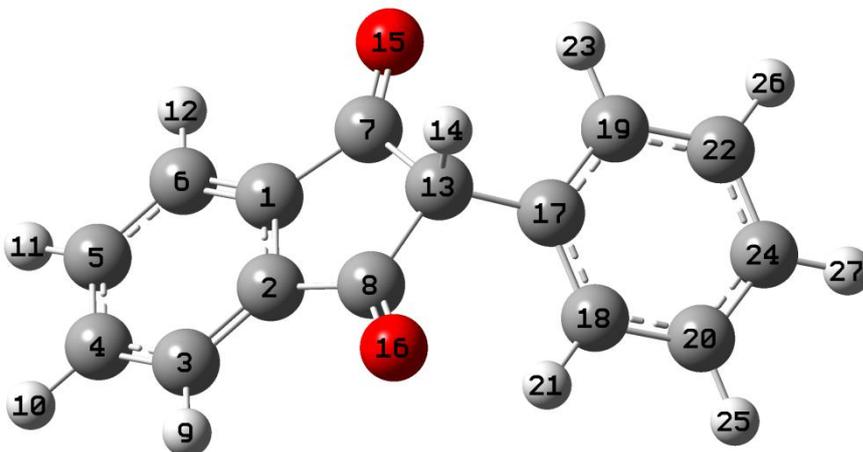


Figura 12: Molécula-mãe: 2-fenilindan-1,3-diona.

4.2 Análise Conformacional da Molécula-Mãe

A partir da geometria otimizada, a molécula-mãe foi submetida a um estudo conformacional para determinar a barreira de rotação em torno da ligação C₁₃-C₁₇, mostrada a seguir (Figura 13), tanto AM1 como B3LYP.

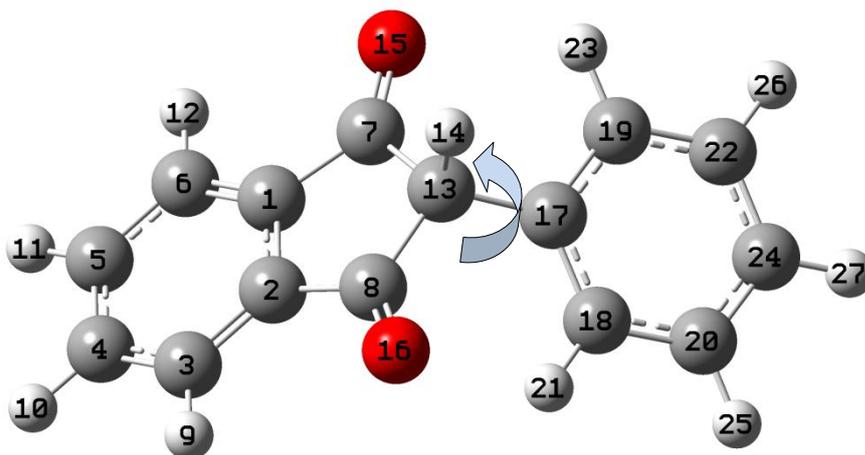


Figura 13: Esquema da rotação do grupo fenil para análise da barreira de rotação.

Nesse procedimento, a barreira de rotação foi avaliada variando-se gradualmente, a cada 10 graus, o ângulo diédrico $C_8C_{13}C_{17}C_{18}$, enquanto as demais distâncias e ângulos de ligação e diédricos eram otimizados completamente. Isso permitiu construir um gráfico, tanto para o cálculo AM1 como para aquele B3LYP/6-31G(d,p), da energia total relativa como função do ângulo diédrico $C_8C_{13}C_{17}C_{18}$, como veremos adiante.

4.3 Obtenção das Propriedades Eletrônicas e Empíricas

A partir das geometrias otimizadas AM1 e B3LYP/6-31G (d,p) da molécula-mãe e de seus derivados foram extraídas as respectivas propriedades eletrônicas empregadas em nossos estudos de relações quantitativas estrutura-atividade das 2-fenilindan-1,3-dionas. Nesse trabalho as propriedades empregadas foram:

- (i) Momento de dipolo molecular,
- (ii) Energias dos orbitais de fronteira HOMO e LUMO, e
- (iii) Cargas atômicas utilizando a partição de Mulliken.

Todos os cálculos foram realizados com o programa GAUSSIAN 03 em microcomputadores de uso comum do Laboratório de Química Teórica e Computacional (LQTC) do Departamento de Química Fundamental da UFPE.

Além desses descritores quânticos, incluímos também em nossas matrizes de correlação estatísticas, dois descritores empíricos:

- (iv) Parâmetro lipofílico de Hansch (π), e
- (v) Coeficiente de partição ($\log P$)

A medida da lipofilicidade ($\log P$) dos compostos da série 2-fenilindan-1,3-dionas foi obtida utilizando o programa online ALOGPS 2.1.

4.4 Efeito Solvente

Em nossos estudos QSAR, os descritores eletrônicos foram obtidos tanto no modelo da molécula isolada, como naquele em que o efeito do solvente é simulado. Para este último caso, dois modelos contínuos diferentes foram empregados, que diferem principalmente no que se refere à definição da cavidade em que se insere o soluto:

- (i) o PCM (*Polarizable Continuum Model*) e
- (ii) o COSMOS (*Conductor-like Screening Model*).

O PCM utiliza um modelo quântico para descrever o soluto; o potencial da superfície é determinado pela diferenciação numérica e sua interação com o solvente é computada de maneira auto-consistente. A cavidade da molécula é definida por meio de esferas atômicas de raio proporcional ao raio de Van der Waals.

O modelo COSMOS também considera cavidades atômicas esféricas para definir a cavidade da molécula (soluto), assim como faz o PCM; entretanto, a sua abordagem para simular as interações eletrostáticas difere significativamente deste último. No modelo COSMOS considera-se o soluto como se estivesse perfeitamente blindado pelo solvente (o solvente tem cargas livres, como um condutor). As cargas são distribuídas na superfície da cavidade e determinadas pela imposição de que os potenciais se cancelam na superfície.

Para realizar esses cálculos foram utilizados os programas GAUSSIAN 03 para o modelo PCM e MOPAC 09 para o modelo COSMOS; nesse procedimento, as estruturas otimizadas pelo método da molécula isolada foram tomadas como de partida, e em seguida, um cálculo *Single Point* foi realizado para a molécula-mãe e seus derivados, incluindo agora o efeito solvente usando esses dois modelos, permitindo assim obter suas propriedades eletrônicas.

4.5 Estudo Quimiométrico

De posse dos parâmetros moleculares e empíricos assim obtidos, bem como das atividades hipolipidêmicas dos derivados de 2-fenilindan-1,3-dionas, estudos quimiométricos com o programa statistica 6.0 foram realizados, utilizando ambos os métodos: RLM e PCA, descritos no Capítulo anterior dessa dissertação.

A partir da matriz de resposta e variáveis, obteve-se a matriz de correlação, que, por sua vez, permitiu determinar, dentre os descritores utilizados, aqueles apresentando uma maior correlação (dentro de nível de 95% de confiança) com a atividade hipolipidêmica. O valor de seleção foi escolhido no sentido de obter um modelo que descrevesse satisfatoriamente os dados, partindo do pressuposto que as variáveis escolhidas estavam suficientemente correlacionadas com a atividade e que eram pouco correlacionadas entre si.

Estudou-se qualitativamente a equação de QSAR no sentido de avaliar quais propriedades contribuíam mais significativamente no valor da atividade e de que modo – positiva ou negativamente. A partir dessa análise e das equações obtidas, dois novos compostos foram projetados, possuindo *a priori* atividades hipolipidêmicas superiores àqueles empregados no conjunto de treinamento.

5. Resultados e Discussões

5.1 Estudo da Molécula Mãe: 2-Fenilindan-1,3-diona

5.1.1 Análise Conformacional

A Figura 14 mostra o resultado AM1 da barreira de energia de rotação do grupo fenil em torno da ligação C₁₃ – C₁₇ (opção *scan* no G03), variando-se a cada 10 graus o ângulo diédrico C₈C₁₃C₁₇C₁₈:

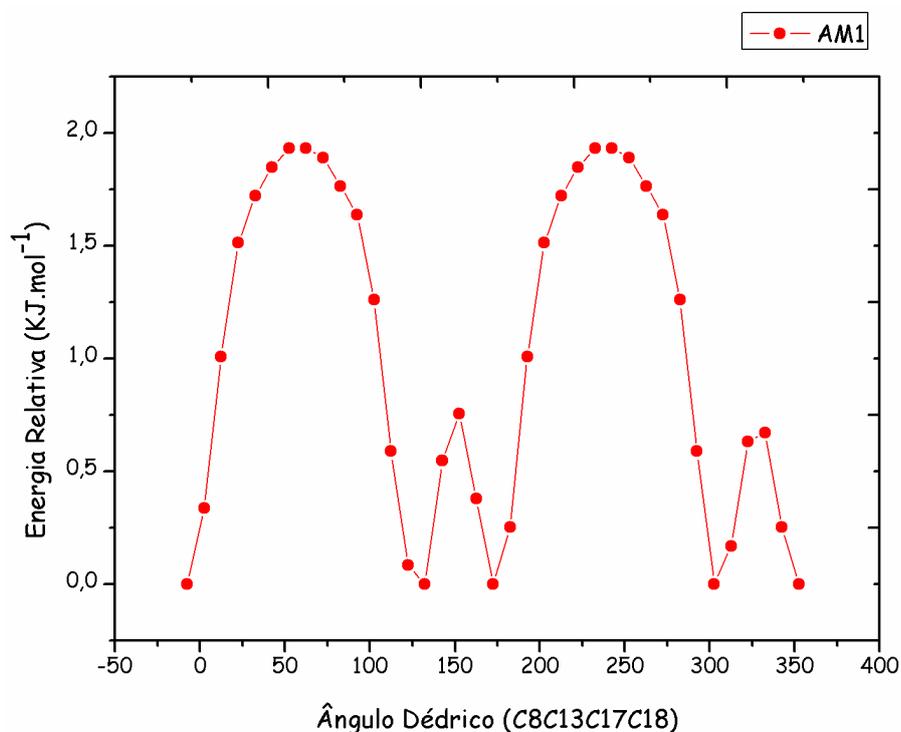


Figura 14: Gráfico da barreira de rotação em torno da ligação C₁₃ – C₁₇ para a molécula-mãe 2-fenilindan-1,3-diona utilizando o método AM1.

Como podemos notar, a metodologia AM1 revela quatro pontos de mínimo de mesma energia, que correspondem aos ângulos diédricos de 7,5°, 132,5°, 172,5° e 302,5°, em função da simetria molecular.

A estrutura correspondente a esses pontos foi chamada de “Gauche”, onde sua estrutura pode ser vista na figura 15:

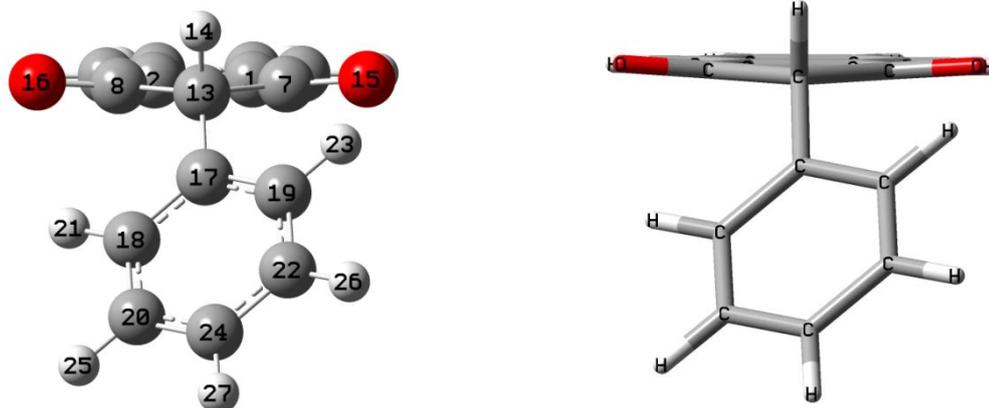


Figura 15: Ângulo diédrico $C_8C_{13}C_{17}C_{18}$ ($7,5^\circ$)

O gráfico da Figura 14 também mostra a existência de dois máximos de mesma energia (globais) em $52,5^\circ$ e $232,5^\circ$. A estrutura correspondente a esses pontos foi chamada de “perpendicular”, onde sua estrutura pode ser vista na figura abaixo:

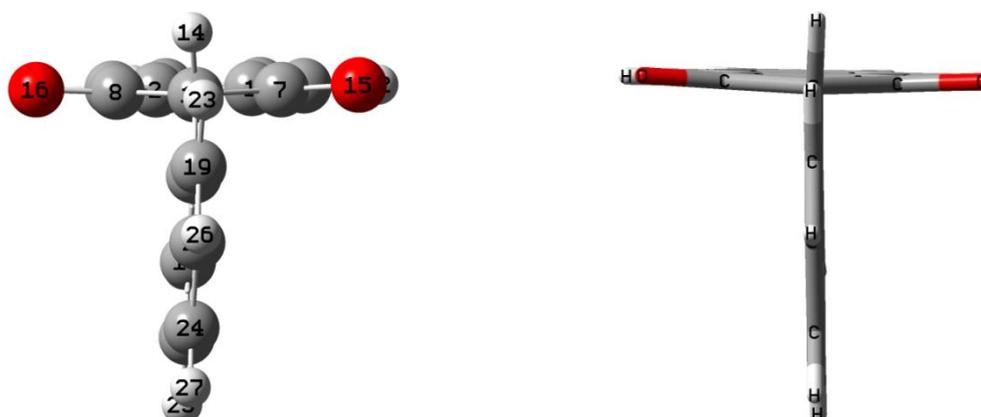


Figura 16: Ângulo diédrico $C_8C_{13}C_{17}C_{18}$ ($52,5^\circ$)

Na figura 14, pode-se também perceber dois outros pontos de máximo de menor energia (máximos locais), relativos aos ângulos de $152,5^\circ$ e $332,5^\circ$ respectivamente.

Ainda no gráfico da Figura 14, é possível extrair o valor da barreira de energia entre o mínimo global e o máximo global para a molécula de 2-fenilindan-1,3-diona, ou seja, de aproximadamente 1,93 kJ/mol. Portanto, o método AM1 prevê uma barreira de energia muito baixa.

Na Figura abaixo mostramos os resultados dessa barreira de rotação para o cálculo B3LYP/6-31G (d,p), cujo perfil da barreira é absolutamente similar àquela AM1.

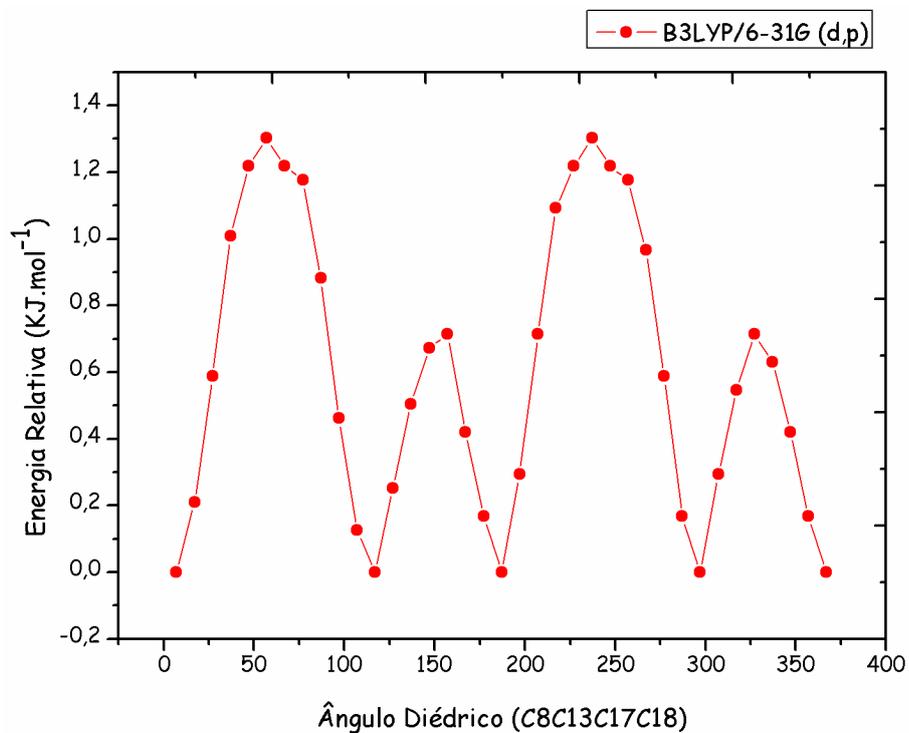


Figura 17: Gráfico da barreira de rotação em torno da ligação C₁₃ – C₁₇ para a molécula-mãe 2-fenilindan-1,3-diona utilizando o método B3LYP/6-31G (d, p).

Os cálculos B3LYP também revelam quatro pontos de mínimo de mesma energia, em ângulos diédricos C₈C₁₃C₁₇C₁₈ também muito similares àqueles AM1, ou seja, em 7,1°, 117,1°, 187,1° e 297,1°. A estrutura (“Gauche”) correspondente ao ângulo diédrico 7,1° é mostrada a seguir:

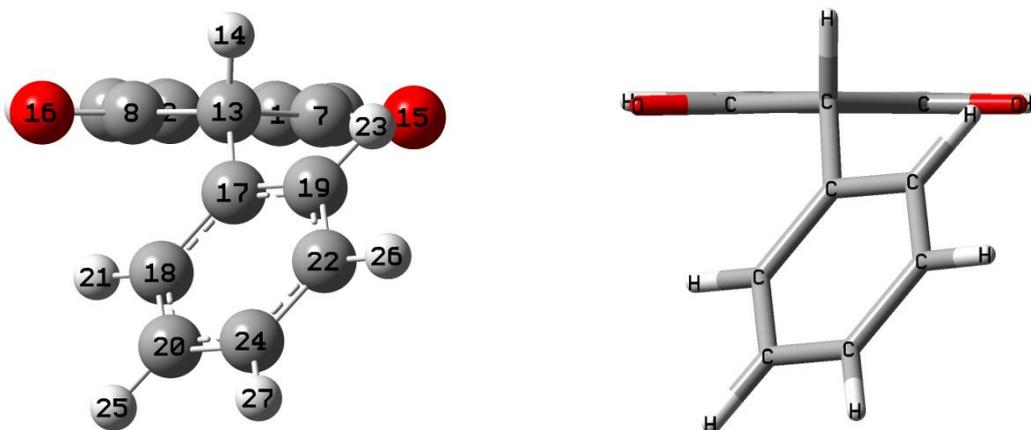


Figura 18: Estrutura “Gauche” da molécula-mãe para o ângulo diédrico $C_8C_{13}C_{17}C_{18}$ de $7,1^\circ$.

Dois pontos de máximos globais foram também observados para os cálculos B3LYP em $57,1^\circ$ e $237,1^\circ$, enquanto àqueles AM1 foram em $52,5^\circ$ e $232,5^\circ$, respectivamente. Portanto, os pontos de máximos globais B3LYP e AM1 ocorrem em ângulos diédricos muito similares. A estrutura “Perpendicular” para o ponto de máximo global B3LYP em $57,1^\circ$ é mostrada a seguir:

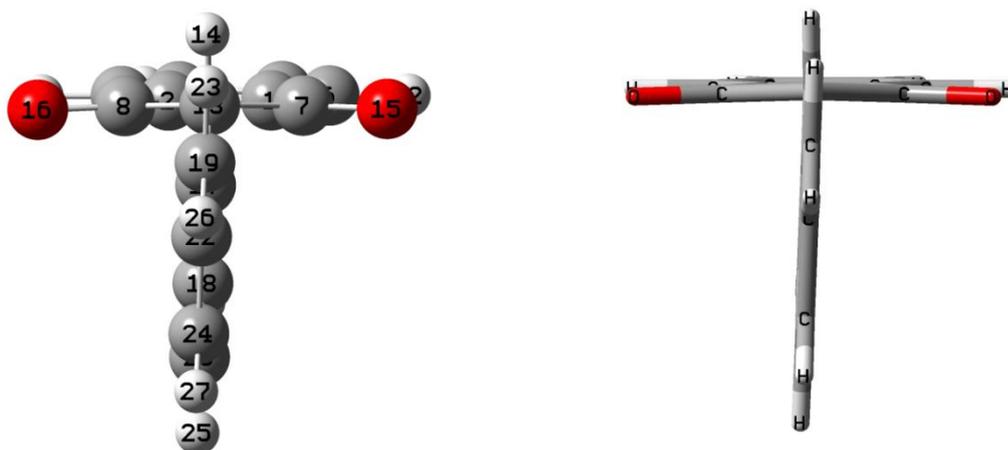


Figura 19: Estrutura “Perpendicular” da molécula-mãe para o ângulo diédrico $C_8C_{13}C_{17}C_{18}$ de $57,1^\circ$.

Como observado também no método AM1, percebeu-se que duas conformações possuem pontos de máximos menos acentuados (máximos locais), onde os ângulos

diédricos são de $147,1^\circ$ e $327,1^\circ$, enquanto àqueles correspondentes AM1 situam-se em $152,5^\circ$ e $332,5^\circ$, respectivamente.

A partir da figura 17 é possível extrair o valor da barreira de energia B3LYP para a molécula de 2-fenilindan-1,3-diona, ou seja, de aproximadamente 1,30 kJ/mol.

Na Tabela 1 mostramos os ângulos diédricos relativos aos mínimos globais da molécula-mãe e de seus 12 derivados para os cálculos AM1 e B3LYP, cujos mínimos não apresentaram frequências imaginárias. Para a obtenção desses ângulos diédricos AM1 e B3LYP dos 12 derivados, partimos do mínimo global da molécula-mãe em $7,5^\circ$ para os cálculos AM1 e $7,1^\circ$ para os cálculos B3LYP, respectivamente.

Tabela 1: Ângulos diédricos da otimização total de energia mínima dos 13 derivados de 2-fenilindan-1,3-dionas.

Compostos	Ângulo Diédrico (graus)	
	$C_8C_{13}C_{17}C_{18}$	$C_8C_{13}C_{17}C_{18}$
	AM1	B3LYP
(a): H	7,5	7,1
(b): <i>o</i> - OCH ₃	26,8	28,7
(c): <i>m</i> - OCH ₃	7,3	7,0
(d): <i>p</i> - OCH ₃	8,0	5,7
(e): <i>o</i> - CH ₃	42,5	36,2
(f): <i>m</i> - CH ₃	7,8	0,1
(g): <i>p</i> - CH ₃	8,7	2,1
(h): <i>o</i> - Cl	33,8	23,0
(i): <i>m</i> - Cl	8,1	4,3
(j): <i>p</i> - Cl	8,3	3,0
(k): <i>o</i> - COOH	59,4	35,1
(l): <i>p</i> - COOH	6,7	14,3
(m): <i>p</i> - OCH ₂ CH ₃	8,2	3,9

As estruturas otimizadas AM1 para a molécula-mãe e de seus 12 derivados são mostradas a seguir:

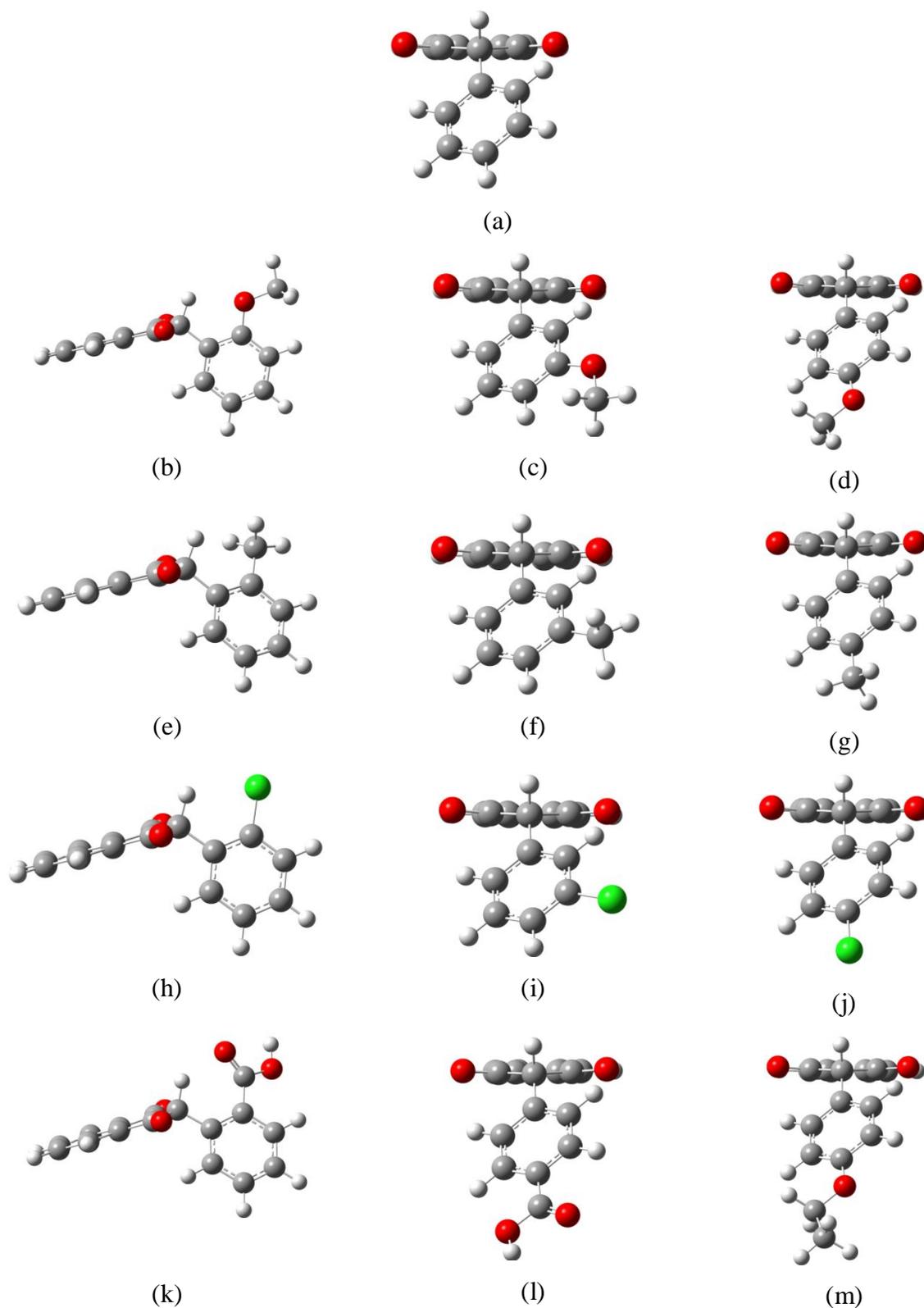


Figura 20: Estruturas otimizadas para os 13 derivados de 2-fenilindan-1,3-dionas utilizando o método AM1.

As estruturas otimizadas para os treze derivados através do método B3LYP/6-31G (d,p) são mostradas abaixo:

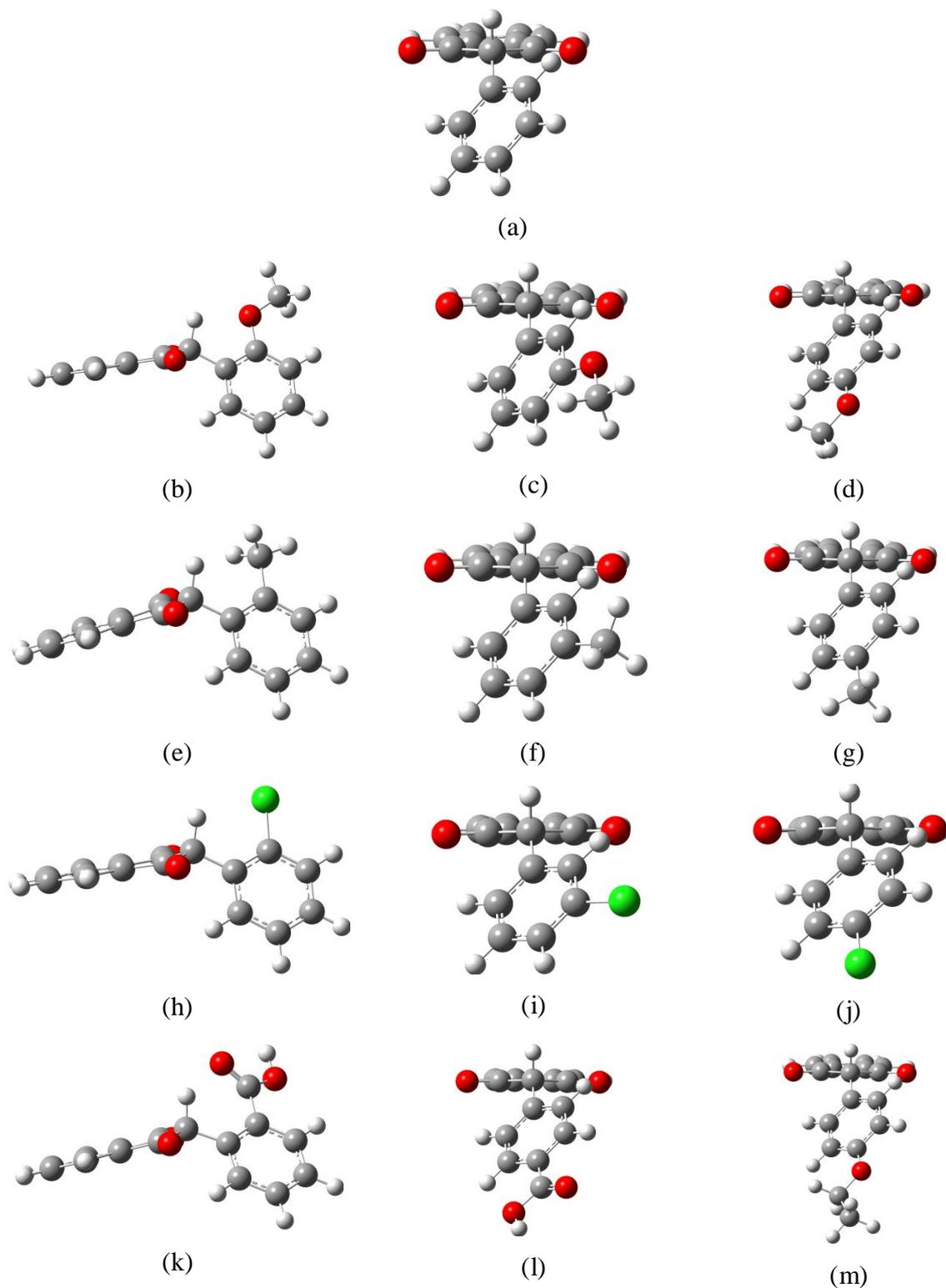
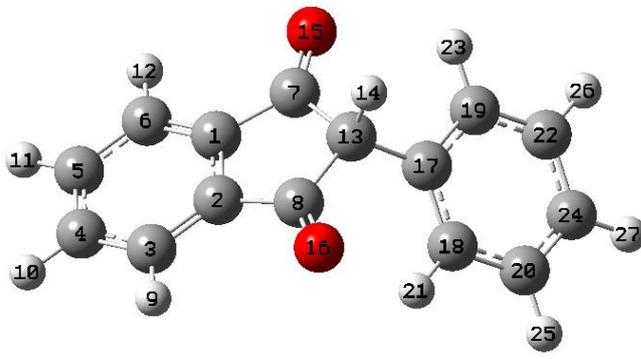


Figura 21: Estruturas otimizadas para os 13 derivados de 2-fenilindan-1,3-dionas utilizando o método B3LYP.

Em geral, podemos notar, de acordo com as figuras 20 e 21, que os derivados (b), (e), (h) e (k) na posição *orto* das 2-fenilindan-1,3-dionas apresentam os maiores ângulos diédricos formando as estruturas chamadas de “Perpendiculares”. Isto se deve a maior repulsão do anel fenílico com o grupo indan-1,3diona, enquanto os menores ângulos pertencem às estruturas chamadas de “Gauche” (que são em grande maioria). Nota-se o mesmo padrão de comportamento tanto para o nível de cálculo AM1, como para o B3LYP.

5.2 Seleção dos Descritores Eletrônicos e Clássicos.

Para a obtenção de um bom modelo de regressão linear, é usualmente desejável que os dados da resposta biológica possuam uma variância apreciável, a fim de conferir ao modelo uma boa previsibilidade. Destacam-se na figura 22, os valores da atividade hipolipidêmica dos compostos da série homóloga estudada.

	Substituinte (R)	Redução (%) ^a	
		pC ₁₆	pT ₁₆
	1: H	1,602	1,602
	2: <i>o</i> - CH ₃	1,491	1,708
	3: <i>m</i> - CH ₃	1,462	1,690
	4: <i>p</i> - CH ₃	1,544	1,763
	5: <i>o</i> - Cl	1,643	1,653
	6: <i>m</i> - Cl	1,602	1,477
	7: <i>p</i> - Cl	1,568	1,724
	8: <i>o</i> - COOH	1,342	1,681
	9: <i>p</i> - OCH ₂ CH ₃	1,613	1,279
	10: <i>p</i> - COOH	1,342	0,699
	11: <i>o</i> - OCH ₃	1,447	1,716
	12: <i>m</i> - OCH ₃	1,519	1,653
	13: <i>p</i> - OCH ₃	1,613	1,763

$$^a pX = \log X$$

Figura 22: Relação dos compostos da série homóloga estudada com os respectivos valores da redução dos níveis de colesterol (pC₁₆) e triglicerídeos (pT₁₆), após 16 dias de administração da droga.

Pode-se observar na figura acima, que os valores das atividades estão expressas em forma de pC_{16} e pT_{16} , onde, esses valores expressam o logaritmo da redução dos níveis de colesterol e triglicerídeos. Quanto maior for o valor de pC e pT , mais ativo será o derivado. Observa-se na figura acima, por exemplo, que o derivado mais ativo em relação a pC é o derivado **5** (*o-cloro*), enquanto os derivados **8** (*o-Carbóxi*) e **10** (*p-Carbóxi*), que possuem o mesmo valor de pC , são os menos ativos.

5.2.1 Método AM1 – Modelo da Molécula Isolada (AM1-MMI)

A partir das conformações mais estáveis, as propriedades eletrônicas especificadas anteriormente, ou seja, energia do orbital molecular mais alto ocupado (ϵ_{HOMO}), energia do orbital molecular mais baixo desocupado (ϵ_{LUMO}), momento dipolar (μ) e cargas líquidas atômicas (q_{ai}), foram calculadas tanto para o método AM1 como para aquele B3LYP. Seus valores AM1 no modelo da molécula isolada são mostrados na Tabela 2, juntamente com os descritores empíricos π e $\log P$, além das atividades hipolipidêmicas da molécula-mãe e de seus 12 derivados.

Tabela 2: Atividades hipolipidêmicas, descritores eletrônicos AM1, na molécula isolada, e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.

Compostos	pC_{16}	pT_{16}	ϵ_{HOMO}	ϵ_{LUMO}	μ	qO_{15}	qO_{16}	qC_{13}	qH_{14}	qCO_1	qCO_2	qC_{17}	π	π^2	$\log P$
1	1,602	1,602	-9,548	-0,981	2,496	-0,263	-0,262	-0,159	0,153	0,018	0,018	-0,072	0,000	0,000	2,970
2	1,491	1,708	-9,339	-0,952	2,302	-0,259	-0,254	-0,159	0,152	0,022	0,022	-0,067	0,560	0,314	3,300
3	1,462	1,690	-9,361	-0,965	2,323	-0,263	-0,262	-0,159	0,153	0,018	0,018	-0,068	0,560	0,314	3,360
4	1,544	1,763	-9,269	-0,962	2,372	-0,263	-0,262	-0,157	0,152	0,018	0,018	-0,078	0,560	0,314	3,360
5	1,643	1,653	-9,562	-0,907	1,327	-0,252	-0,252	-0,163	0,148	0,030	0,029	-0,048	0,710	0,504	3,520
6	1,602	1,477	-9,560	-1,067	3,453	-0,259	-0,261	-0,162	0,155	0,021	0,019	-0,061	0,710	0,504	3,580
7	1,568	1,724	-9,474	-1,077	3,490	-0,262	-0,261	-0,161	0,155	0,019	0,019	-0,070	0,710	0,504	3,580
8	1,342	1,681	-9,687	-1,079	3,310	-0,263	-0,264	-0,162	0,170	0,020	0,016	-0,013	-0,320	0,102	2,480
9	1,613	1,279	-8,933	-0,959	2,877	-0,264	-0,264	-0,151	0,151	0,016	0,016	-0,111	0,380	0,144	3,320
10	1,342	0,699	-9,987	-1,156	4,673	-0,260	-0,259	-0,169	0,158	0,021	0,021	-0,032	-0,320	0,102	2,630
11	1,447	1,716	-9,048	-0,843	1,003	-0,253	-0,253	-0,153	0,169	0,024	0,024	-0,086	-0,020	0,000	2,840
12	1,519	1,653	-9,05	-0,956	2,408	-0,258	-0,261	-0,161	0,155	0,022	0,02	-0,039	-0,020	0,000	2,900
13	1,613	1,763	-8,975	-0,970	2,981	-0,264	-0,264	-0,152	0,151	0,016	0,016	-0,109	-0,020	0,000	2,910

A partir dessa matriz de dados, foi obtida a matriz dos coeficientes de correlação mais significativos (R) entre as variáveis previamente selecionadas e as respectivas atividades hipolipidêmicas, conforme mostrado na Tabela 3 abaixo:

Tabela 3: Matriz de correlação das variáveis (AM1-Molécula Isolada) e das atividades.

	pC ₁₆	pT ₁₆	ε _{LUMO}	μ	qH ₁₄	qC ₁₇	π	logP
pC ₁₆	1,00	0,34	0,38	-0,31	-0,73	-0,60	0,60	0,69
pT ₁₆	0,34	1,00	0,58	-0,64	-0,02	-0,22	0,35	0,27
ε _{LUMO}	0,38	0,58	1,00	-0,95	-0,13	-0,49	0,21	0,19
μ	-0,31	-0,64	-0,95	1,00	0,03	0,27	-0,27	-0,21
qH ₁₄	-0,73	-0,02	-0,13	0,03	1,00	0,40	-0,60	-0,66
qC ₁₇	-0,60	-0,22	-0,49	0,27	0,40	1,00	-0,30	-0,37
π	0,60	0,35	0,21	-0,27	-0,60	-0,30	1,00	0,98
logP	0,69	0,27	0,19	-0,21	-0,66	-0,37	0,98	1,00

O objetivo dessa matriz é o de analisar e selecionar as variáveis que possuem simultaneamente uma maior correlação com a atividade – garantindo que as variações de tais descritores são de fato passíveis de interferência na atividade – e uma menor correlação entre si, diminuindo a redundância da informação e possibilitando uma modelagem de maior qualidade.

5.2.1.1 Análise RLM - pC₁₆

Os valores, em vermelho, da Tabela 3, correspondem aos descritores apresentando os maiores coeficientes de correlação com as atividades hipolipidêmicas com 95% de confiança. Nesse sentido, destacam-se, em relação às atividades de redução nos níveis de colesterol: as cargas atômicas do hidrogênio 14 e do carbono 17, além dos descritores empíricos π e logP. Observando os valores de R, nota-se que alguns dos descritores são razoavelmente correlacionados entre si; como, por exemplo, a carga do hidrogênio 14 apresenta um coeficiente de correlação de 0,66 com o log P, mostrando que são variáveis linearmente dependentes, e, portanto, não devem ser empregados numa mesma equação de QSAR. Da Tabela 3, podemos notar que o parâmetro que possui uma maior correlação

com pC_{16} é a carga atômica do hidrogênio 14 com um coeficiente de correlação (R) de -0,73; quanto menos positiva for a carga do hidrogênio, maior será a atividade na redução nos níveis de colesterol. Além disso, os demais decriptores, além de possuírem uma menor correlação com o pC_{16} , os mesmos possuem uma alta correlação entre si. A partir de um ajuste linear dos dados, foi encontrada a seguinte equação de regressão entre pC_{16} e qH_{14} , como pode ser visto a seguir

$$pC_{16} = 3,24(\pm 0,48) - 11,02(\pm 3,11)qH_{14} \quad (5.1)$$

(n=13; R= -0,73; $R^2 = 0,53$, F=12,57; s=0,072; p = 0,0046)

Os valores entre parênteses correspondem aos desvios-padrão dos coeficientes da equação. Embora o valor de R^2 não seja muito admirável, ainda assim explica 53% da variabilidade da atividade num intervalo de confiança de 95%.

A partir desta equação (5.1) valores de pC_{16} foram previstos e plotados contra seus correspondentes valores experimentais, cujo resultado é mostrado abaixo:

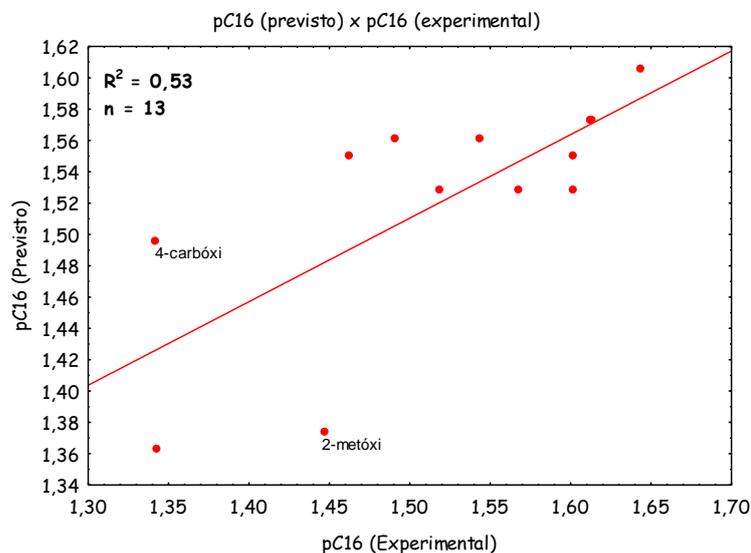


Figura 23: Valores previstos de pC_{16} usando equação (5.1) versus seus correspondentes valores experimentais.

Pode-se observar no gráfico da Figura 23, que os maiores desvios do comportamento linear são verificados para derivados **10** (*p*-Carbóxi) e **11** (*o*-Metóxi). Considerando que a carga do hidrogênio é influenciada pela mudança conformacional, nós

resolvemos realizar um estudo conformacional nesses dois derivados para analisar este efeito sobre as cargas do hidrogênio 14, especialmente porque as respectivas mudanças conformacionais podem ser relativamente baixas em termos de energias relativas, enquanto as mudanças nas cargas podem ser significativas do ponto de vista do ajuste linear.

A Figura 24 mostra as barreiras de rotação para esses dois derivados, ao tempo em que foram analisadas as respectivas cargas atômicas do hidrogênio 14 com a mudança do ângulo diédrico. Inicialmente, podemos notar que as respectivas barreiras são 18,94 kJ/mol para o *orto*-metóxi e 2,39 kJ/mol para o *para*-carbóxi. Como esperado, em decorrência do maior impedimento estérico, a altura da barreira para o derivado *orto* é bem superior àquela do derivado com o substituinte em *para*.

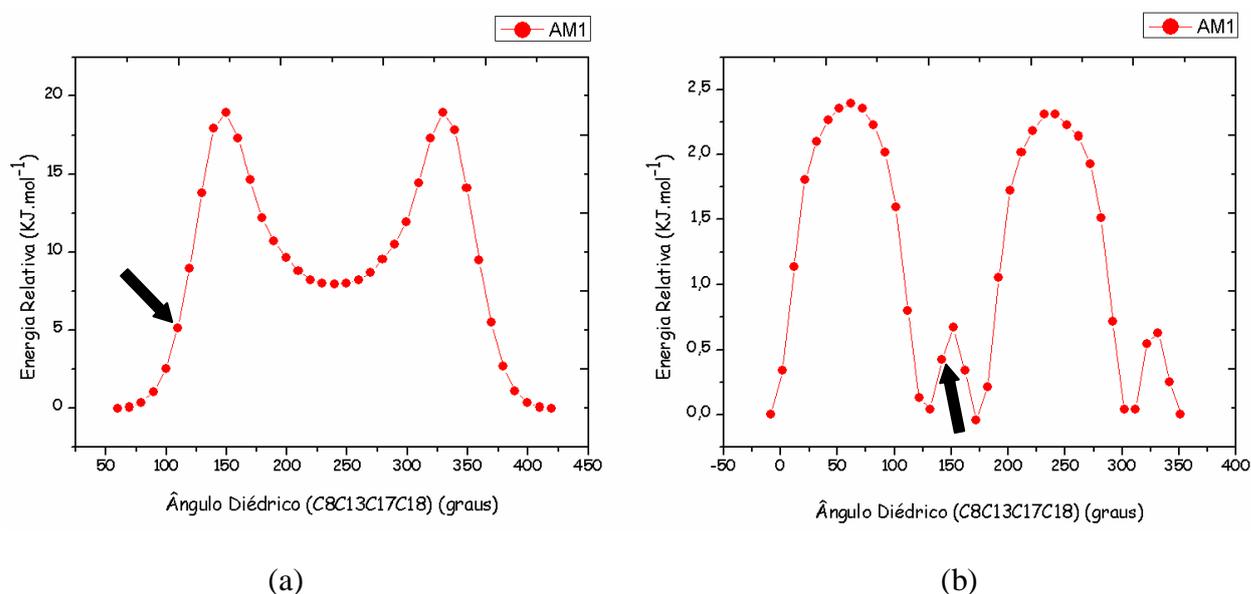


Figura 24: Gráficos das barreiras de rotação para os derivados (a) *orto*-metóxi e (b) *para*-carbóxi utilizando o método AM1 na molécula isolada, onde a energia relativa é diferença do valor da energia total para um dado ângulo diedro e a energia total da conformação mais estável.

Analisando os valores de qH_{14} para cada um dos pontos dessas curvas de potencial, foi possível identificar os valores que otimizam o ajuste linear, conforme equação (5.1), deste descritor com pC_{16} . Os correspondentes valores “ótimos” de qH_{14} para os derivados **10** (*para*-Carbóxi) e **11** (*orto*-Metóxi) são 0,159e e 0,160e, respectivamente, os quais são indicados por setas nas respectivas curvas. As respectivas diferenças de energia destes

pontos com relação aos seus valores na conformação mais estável são 8,95 kJ/mol para o *orto*-metóxi e 0,42 kJ/mol para o *para*-carbóxi.

O uso dessas conformações que não apresentam energia mínima, para a obtenção de propriedades que levam a um melhor ajuste linear na equação de regressão, não deixa de ser importante considerando que esses sistemas estão em um meio biológico e podem naturalmente assumir conformações distintas daquela relativa à conformação mais estável.

De posse dos novos valores de qH_{14} os dois derivados **10** e **11** da série homóloga estudada, foi construída uma nova matriz de dados e por conseguinte, uma nova matriz de correlação foi obtida. As mais significativas correlações são mostradas na Tabela 4 abaixo:

Tabela 4: Matriz de correlação das variáveis e da atividade após o ajuste dos dois derivados **10** e **11**.

	pC ₁₆	pT ₁₆	ε _{LUMO}	μ	qH ₁₄	qC ₁₇	π	logP
pC ₁₆	1,00	0,34	0,35	-0,28	-0,82	-0,56	0,60	0,69
pT ₁₆	0,34	1,00	0,59	-0,70	-0,25	-0,24	0,35	0,27
ε _{LUMO}	0,35	0,59	1,00	-0,87	-0,38	-0,49	0,19	0,16
μ	-0,28	-0,70	-0,87	1,00	0,37	0,12	-0,26	-0,18
qH ₁₄	-0,82	-0,25	-0,38	0,37	1,00	0,46	-0,64	-0,68
qC ₁₇	-0,56	-0,24	-0,49	0,12	0,46	1,00	-0,27	-0,33
π	0,60	0,35	0,19	-0,26	-0,64	-0,27	1,00	0,98
logP	0,69	0,27	0,16	-0,18	-0,68	-0,33	0,98	1,00

O novo modelo de QSAR obtido, a partir da regressão efetuada entre pC₁₆ e qH₁₄, é mostrado pela equação abaixo:

$$pC_{16} = 4,71(\pm 0,67) - 20,66(\pm 4,37)qH_{14} \quad (5.2)$$

(n=13; R = -0,82; R² = 0,67, F=22,38; s=0,060; p = 0,00062)

Comparando-se a equação 5.1 como a 5.2, nota-se que essa última é capaz de explicar agora 67% da variabilidade da atividade para um intervalo de 95% de confiança. Nota-se também o baixo desvio padrão do modelo s=0,060.

O gráfico representado na figura 25 mostra os valores previstos pelo modelo (equação 5.2) *versus* os valores experimentais de pC_{16} . Como esperado, nota-se nesse gráfico, um melhor ajuste linear quando comparado com àquele obtido da equação 5.1.

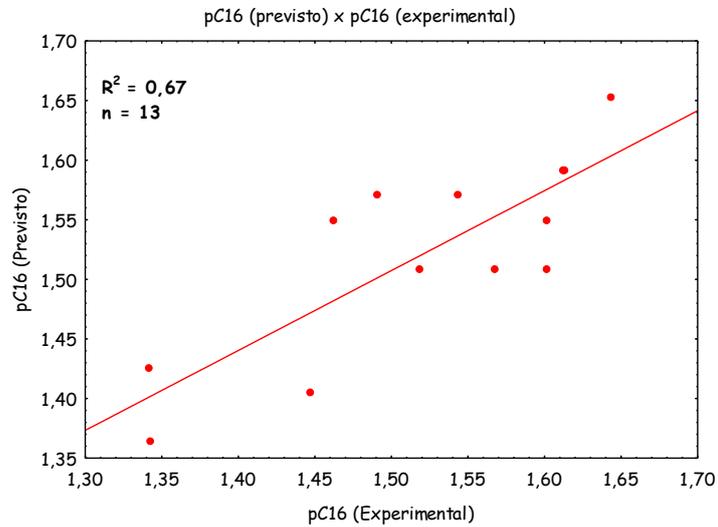


Figura 25: Valores previstos de pC_{16} usando equação (5.2) versus seus correspondentes valores experimentais.

No gráfico da Figura 26, observa-se a distribuição normal dos resíduos em torno de zero, que corresponde à linha horizontal central, ou seja, os resíduos estão aleatoriamente dispersos, e não há nenhum indício de anormalidade, o que garante que o modelo apresenta um bom comportamento linear.

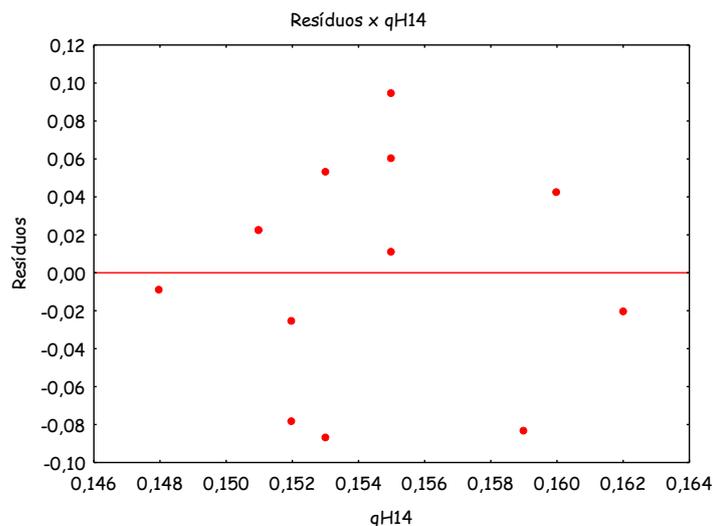


Figura 26: Gráficos dos Resíduos

Os valores numéricos correspondentes ao gráfico da Figura 25 são mostrados na Tabela 5, em função da percentagem do nível de colesterol reduzido (%C₁₆). Como podemos notar há, em geral, uma boa concordância entre valores experimentais e previstos, sendo que a maior diferença é verificada para os derivados **2**, **6** e **10**.

Tabela 5: Redução nos níveis de colesterol obtidos dos dados experimentais e previstos usando a equação 5.2, após 16 dias de administração da droga.

Compostos	%C ₁₆ (Experimental) ^a	%C ₁₆ (Previsto)
1	40±2	35
2	31±3	37
3	29±3	35
4	35±3	37
5	44±4	45
6	40±4	32
7	37±3	32
8	22±2	23
9	41±2	39
10	22±2	27
11	28±2	25
12	33±5	32
13	41±3	39

^aHall *et al.*, 1985.

5.2.1.2 Análise RLM - pT₁₆

Para o pT₁₆ observando a matriz de correlação (Tabela 3) os parâmetros que possuem uma maior correlação com a atividade na redução de níveis de triglicerídeos são: o momento dipolar (μ) com R= -0,64 e a energia do LUMO (ϵ_{LUMO}) com R= 0,58. Infelizmente, ambos não podem ser usados numa equação de regressão, uma vez que são altamente correlacionados com R= -0,95. Assim, a melhor equação de regressão para pT₁₆ é representada por:

$$pT_{16} = 2,10(\pm 0,20) - 0,20(\pm 0,071)\mu \quad (5.3)$$

(n=13; R = -0,64; R² = 0,41, F=7,74; s=0,23; p = 0,018)

A equação 5.3 mostra que quanto menor for o momento de dipolo do derivado, maior será a redução nos níveis de triglicerídeos. Os valores entre parênteses correspondem aos desvios-padrão dos coeficientes da equação. Embora o valor de R² não seja também muito admirável, o mesmo consegue ainda explicar 41% da variabilidade da atividade para um intervalo de confiança de 95%.

Seguindo o mesmo procedimento adotado para pC₁₆, aqui os dois derivados que mais se afastam do ajuste linear, representado pela equação acima, são o *orto*-metóxi e *orto*-cloro, cujas barreiras de rotação para o ângulo diedro C₈C₁₃C₁₇C₁₈ são mostradas na Figura 27.

As barreiras de energia calculadas para o *orto*-metóxi e o *orto*-cloro são 18,94 kJ/mol e 30,16 kJ/mol, respectivamente.

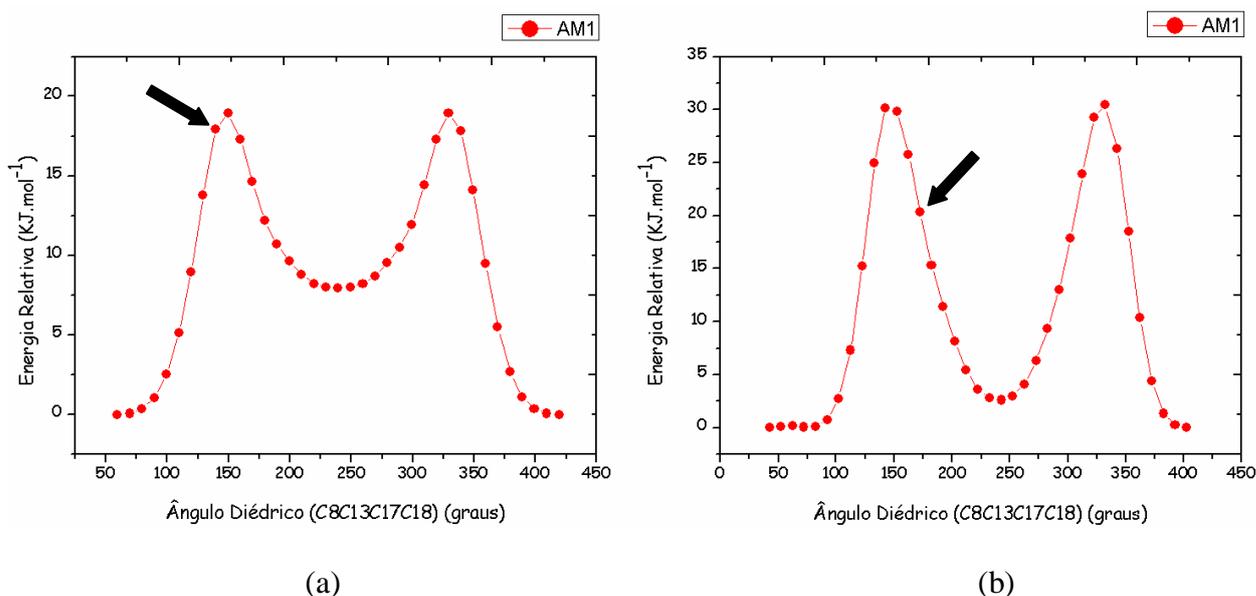


Figura 27: Gráficos das barreiras de rotação para os derivados (a) *orto*-metóxi e (b) *orto*-cloro utilizando o método AM1 na molécula isolada, onde a energia relativa é diferença do valor da energia total para um dado ângulo diedro e a energia total da conformação mais estável.

Analisando os valores de μ para cada um dos pontos dessas curvas de potencial, foi possível identificar os valores que otimizam o ajuste linear, conforme equação (5.3), deste

descritor com pT_{16} . Os correspondentes valores “ótimos” de μ para os derivados *orto-metóxi* e *orto-cloro* são 2,54D e 2,64D, respectivamente, os quais são indicados por setas nas respectivas curvas. Esses valores nas respectivas conformações mais estáveis são 1,00D e 3,21D. Por outro lado, as respectivas diferenças de energia destes pontos com relação aos seus valores na conformação mais estável são 17,9 kJ/mol para o *orto-metóxi* e 20,4 kJ/mol para o *orto-cloro*.

De posse dos novos valores de μ para esses dois derivados, foi construída uma nova matriz de dados e, por conseguinte, uma nova matriz de correlação foi obtida. As mais significativas correlações são mostradas na Tabela 6 abaixo:

Tabela 6: Matriz de correlação das variáveis e da atividade após o ajuste dos dois derivados:

	pC_{16}	pT_{16}	ϵ_{LUMO}	μ	qH_{14}	qC_{17}	π	$\log P$
pC_{16}	1,00	0,34	0,31	-0,38	-0,64	-0,60	0,60	0,69
pT_{16}	0,34	1,00	0,55	-0,76	-0,12	-0,24	0,35	0,27
ϵ_{LUMO}	0,31	0,55	1,00	-0,79	-0,50	-0,55	0,16	0,13
μ	-0,38	-0,76	-0,79	1,00	0,43	0,37	-0,33	-0,28
qH_{14}	-0,64	-0,12	-0,50	0,43	1,00	0,80	-0,47	-0,54
qC_{17}	-0,60	-0,24	-0,55	0,37	0,80	1,00	-0,30	-0,37
π	0,60	0,35	0,16	-0,33	-0,47	-0,30	1,00	0,98
$\log P$	0,69	0,27	0,13	-0,28	-0,54	-0,37	0,98	1,00

Após os dados ajustados o novo modelo de QSAR obtido, a partir da regressão efetuada, para o pT_{16} , é mostrado a seguir:

$$pT_{16} = 2,53(\pm 0,25) - 0,33(\pm 0,085)\mu \quad (5.4)$$

(n=13; R = -0,76; R² = 0,58; F=15,06; s=0,20; p = 0,0026)

Comparando-se as equações 5.3 e 5.4, nota-se que essa última é capaz de explicar agora 58% da variabilidade da atividade, sendo o mesmo significativo no intervalo de 95% de confiança. O coeficiente de correlação, R, que antes era -0,64, agora passa a ser -0,76.

5.2.1.3 Análise de Componentes Principais (PCA)

Na seção anterior, os modelos matemáticos de QSAR, tanto para a redução dos níveis de colesterol (pC_{16}) como para os triglicerídeos (pT_{16}) (equações 5.1/5.2 e 5.3/5.4) respectivamente, foram obtidos utilizando o método da regressão linear (RLM), tendo por objetivo determinar relações quantitativas estrutura-atividade para a série homóloga de 13 derivados de 2-fenilindan-1,3-dionas. Na matriz de correlação mostrada na Tabela 4 foi observado que alguns descritores moleculares e empíricos estavam correlacionados com o pC_{16} , ou seja: a carga do hidrogênio 14 (qH_{14}), a carga do carbono 17 (qC_{17}), o coeficiente de partição ($\log P$) e o parâmetro lipofílico π . Como foi descrito também nos itens 5.2.1.1 e 5.2.1.2, esses quatro descritores não poderiam ser incluídos em um mesmo modelo matemático de QSAR, visto que os mesmos são altamente correlacionados entre si, sendo linearmente dependentes.

Em função dessa dependência linear entre esses descritores, passamos nessa etapa a fazer uma análise de componentes principais (PCA).

A análise de componentes principais realizada foi baseada na matriz de correlação (Tabela 4), utilizando os quatro descritores acima, calculados com o método AM1-MMI. O gráfico dos *loadings* dos dois primeiros componentes (PC1 e PC2), que representa 90,37 % da variância (PC1= 68,84% , PC2= 21,53%), está mostrado pela Figura 28. Podemos notar que pC_{16} é quase coincidente com o lado negativo do eixo PC1.

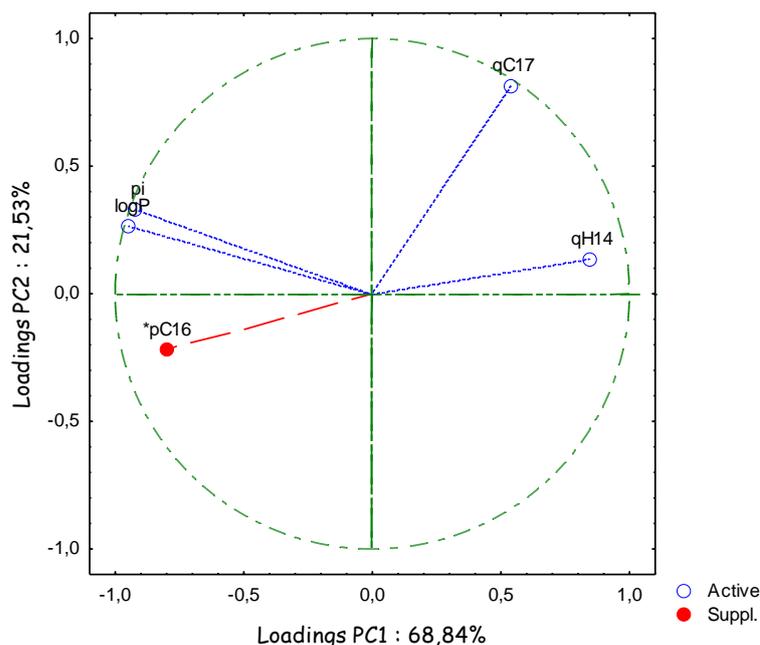


Figura 28: Gráfico dos *loadings* para os descritores: a carga do hidrogênio 14 (qH_{14}), a carga do carbono 17 (qC_{17}), o coeficiente de partição ($\log P$) e o parâmetro lipofílico π

Podemos ainda observar que do mesmo lado da atividade hipolipidêmica, encontram-se também os descritores $\log P$ e π com *loadings* negativos, e os outros dois descritores qC_{17} e qH_{14} apresentam *loadings* positivos todos em PC1, indicando que atividade está associada com os *scores* negativos de PC1. Isto significa que os compostos mais ativos devem apresentar valores mais altos dos descritores $\log P$ e π e valores mais baixos dos descritores qC_{17} e qH_{14} . Este último descritor é o mais importante na descrição do modelo visto que está quase coincidindo com o eixo PC1 positivo, e está diametralmente oposto com a atividade.

O gráfico correspondente aos *scores* encontra-se representado na Figura 29. O gráfico mostra que os derivados mais ativos possuem *scores* negativos no eixo PC1, enquanto os derivados menos ativos, possuem *scores* positivos também no eixo PC1. O derivado mais ativo de acordo com o gráfico corresponde ao composto **5** (2-cloro), enquanto o derivado menos ativo corresponde ao derivado **8** (2-carbóxi). Os *scores* da primeira componente PC1 pode ser interpretada como uma medida teórica da atividade hipolipidêmica dos compostos estudados.

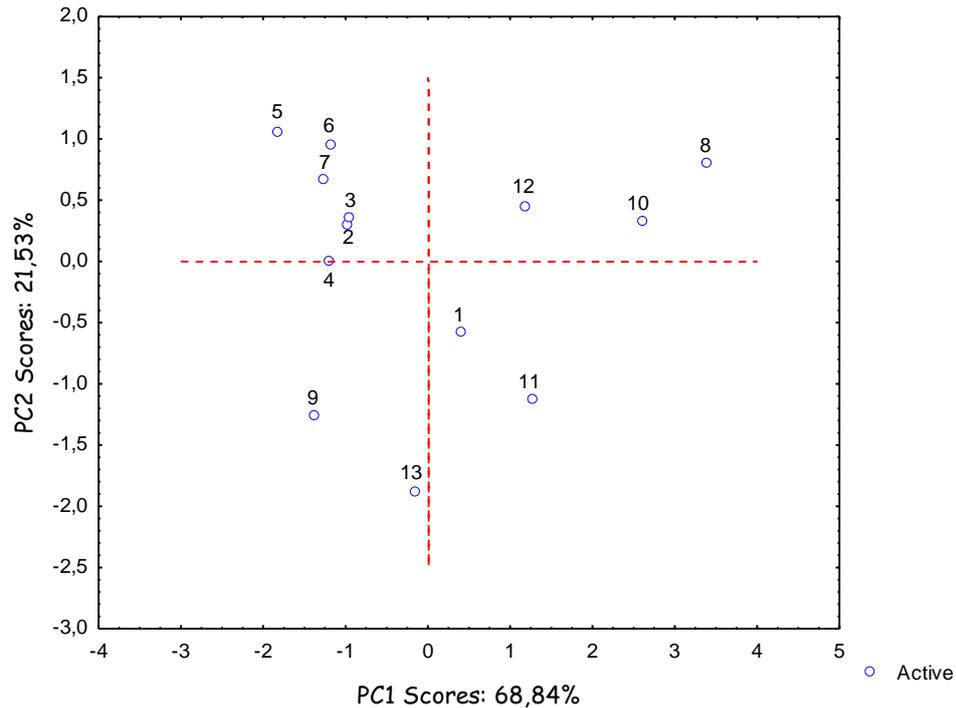


Figura 29: Gráficos dos *scores*

Observando a Tabela de correlação 4 foi realizada uma regressão entre os *scores* de PC1 dos quatro descritores destacados em vermelho, com os valores de pC_{16} . A equação obtida através da análise de componentes principais é mostrada abaixo:

$$pC_{16} = 1,52(\pm 0,018) - 0,04t_1(\pm 0,011) \quad (5.5)$$

($n=13$; $R = -0,78$; $R^2 = 0,61$; $F=17,44$; $s=0,065$; $p = 0,0016$)

onde t_1 representa os *scores* no eixo PC1. Os valores entre parênteses são as estimativas do erro padrão dos coeficientes de regressão. Embora o valor de R^2 não seja muito alto, ambos os coeficientes são estatisticamente significativos em um nível de confiança de 95%.

Para se ter uma idéia do comportamento deste modelo, foi feito um gráfico dos valores previstos *versus* os valores experimentais para pC_{16} , mostrado na Figura 30, e os respectivos valores numéricos encontra-se indicados na Tabela 7 em termos do percentual de redução dos níveis de colesterol.

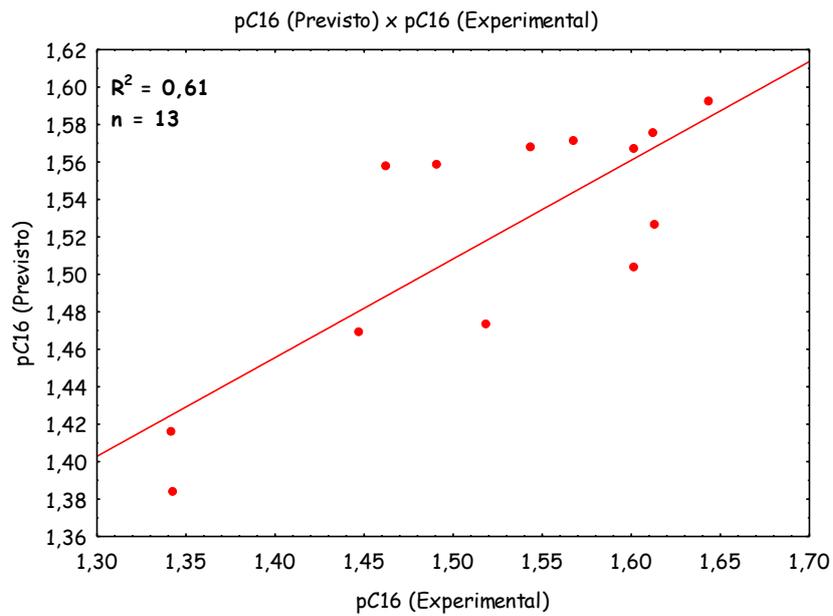


Figura 30: Atividade hipolipidêmica prevista em função da atividade prevista explicando 61% da variância do modelo.

Tabela 7: Redução nos níveis de colesterol obtidos dos dados experimentais e previstos usando a equação 5.5, após 16 dias de administração da droga.

Compostos	%C ₁₆ (Experimental)	%C ₁₆ (Previsto)
1	40±2	32
2	31±3	36
3	29±3	36
4	35±3	37
5	44±4	39
6	40±4	37
7	37±3	37
8	22±2	24
9	41±2	38
10	22±2	26
11	28±2	29
12	33±5	30
13	41±3	34

5.2.2 Método B3LYP/6-31G (d,p) - Modelo da Molécula Isolada (B3LYP-MMI)

Para testar os modelos matemáticos obtidos a partir do estudo AM1, propriedades eletrônicas foram também estimadas usando cálculos B3LYP com um conjunto de base 6-31G(d,p) para os 13 derivados de 2-fenilindan-1,3-dionas. Os descritores eletrônicos B3LYP/6-31G(d,p) foram extraídos das conformações mais estáveis de cada um desses derivados, cujas estruturas são mostradas na Figura 21. A Tabela 8 mostra os valores desses descritores, que são aqueles mesmos empregados nos cálculos AM1, juntamente com as atividades hipolipidêmicas (pC_{16} e pT_{16}) e os descritores empíricos π e $\log P$.

Tabela 8: Atividades hipolipidêmicas, descritores eletrônicos B3LYP, na molécula isolada, e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.

Compostos	pC_{16}	pT_{16}	ϵ_{HOMO}	ϵ_{LUMO}	μ	qO_{15}	qO_{16}	qC_{13}	qH_{14}	qCO_1	qCO_2	qC_{17}	π	π^2	$\log P$
1	1,602	1,602	-6,458	-2,254	3,091	-0,454	-0,464	-0,308	0,168	-0,056	-0,074	0,124	0,000	0,000	2,970
2	1,491	1,708	-6,251	-2,222	2,973	-0,458	-0,455	-0,316	0,158	-0,052	-0,075	0,125	0,560	0,314	3,300
3	1,462	1,690	-6,308	-2,222	2,845	-0,454	-0,464	-0,308	0,167	-0,056	-0,074	0,127	0,560	0,314	3,360
4	1,544	1,763	-6,229	-2,220	2,789	-0,455	-0,465	-0,307	0,167	-0,058	-0,075	0,120	0,560	0,314	3,360
5	1,643	1,653	-6,484	-2,254	3,458	-0,443	-0,452	-0,333	0,183	-0,042	-0,073	0,202	0,710	0,504	3,520
6	1,602	1,477	-6,567	-2,410	4,824	-0,451	-0,464	-0,305	0,175	-0,055	-0,074	0,111	0,710	0,504	3,580
7	1,568	1,724	-6,494	-2,419	5,087	-0,454	-0,465	-0,304	0,174	-0,059	-0,075	0,111	0,710	0,504	3,580
8	1,342	1,681	-6,554	-2,146	2,206	-0,456	-0,465	-0,329	0,212	-0,050	-0,079	0,129	-0,320	0,102	2,480
9	1,613	1,279	-5,730	-2,193	3,015	-0,456	-0,467	-0,306	0,165	-0,059	-0,076	0,119	0,380	0,144	3,320
10	1,342	0,699	-6,773	-2,423	5,088	-0,451	-0,463	-0,309	0,175	-0,055	-0,072	0,128	-0,320	0,102	2,630
11	1,447	1,716	-5,831	-2,051	1,730	-0,451	-0,456	-0,301	0,178	-0,050	-0,076	0,116	-0,020	0,000	2,840
12	1,519	1,653	-5,837	-2,182	2,224	-0,450	-0,462	-0,310	0,166	-0,051	-0,074	0,130	-0,020	0,000	2,900
13	1,613	1,763	-5,760	-2,207	3,170	-0,455	-0,467	-0,306	0,166	-0,058	-0,076	0,119	-0,020	0,000	2,910

A partir dessa matriz de dados, foi obtida a matriz dos coeficientes de correlação; desta última extraímos uma sub-matriz com os valores mais significativos de R, envolvendo as variáveis previamente selecionadas e as respectivas atividades hipolipidêmicas. Esta sub-matriz é mostrada na Tabela 9. Os valores assinalados em vermelho são os mais altos coeficientes de correlação verificados nesta sub-matriz de correlação.

Tabela 9: Matriz de correlação para os valores de R mais significativos entre as variáveis previamente selecionadas B3LYP e as atividades hipolipidêmicas.

	pC ₁₆	pT ₁₆	ε _{LUMO}	μ	qH ₁₄	qC ₁₇	π	logP
pC ₁₆	1,00	0,34	-0,12	0,15	-0,46	0,20	0,60	0,69
pT ₁₆	0,34	1,00	0,49	-0,52	0,01	0,03	0,35	0,27
ε _{LUMO}	-0,12	0,49	1,00	-0,98	0,12	0,06	-0,31	-0,38
μ	0,15	-0,52	-0,98	1,00	-0,10	-0,06	0,30	0,38
qH ₁₄	-0,46	0,01	0,12	-0,10	1,00	0,24	-0,37	-0,43
qC ₁₇	0,20	0,03	0,06	-0,06	0,24	1,00	0,17	0,14
π	0,60	0,35	-0,31	0,30	-0,37	0,17	1,00	0,98
logP	0,69	0,27	-0,38	0,38	-0,43	0,14	0,98	1,00

5.2.2.1 Usando o Modelo de Regressão Linear (RLM)

A partir da Tabela 9, podemos notar que, diferentemente dos cálculos AM1, não se verifica uma correlação significativa entre pC₁₆ e a carga atômica do hidrogênio 14; aqui o valor de R é apenas de -0,46, enquanto nos cálculos AM1 este valor foi de -0,73. A mais importante correlação agora observada nos cálculos B3LYP foi com o coeficiente de partição (log P), ou seja, com um coeficiente de correlação de 0,69. O descritor seguinte mais importante foi à constante de lipofilicidade. Aqui não há como realizar uma equação de regressão linear múltipla, uma vez que os descritores empíricos π e log P, como esperado, são altamente correlacionados (R= 0,98).

Já no caso do pT₁₆, os descritores eletrônicos obtidos dos cálculos B3LYP mais significativos são exatamente os mesmos dos cálculos AM1, ou seja, o momento dipolo da molécula e a energia do orbital LUMO, os quais, como também em AM1, são altamente correlacionados. Porém aqui, para os cálculos B3LYP, os coeficientes de correlação entre pT₁₆ e esses dois descritores (μ e ε_{LUMO}) são relativamente menores do que àqueles AM1. Por exemplo, aqui o valor de R entre pT₁₆ e μ é de -0,52, enquanto nos cálculos AM1 esse valor foi de -0,70.

Aqui, um melhor ajuste linear pode ser obtido fazendo-se uma regressão linear múltipla para pT₁₆ com μ e o descritor empírico π, ao tempo em que há uma perda de

significância estatística pela inclusão de mais uma variável para melhor descrever a mudança nos níveis de triglicérides com a estrutura molecular.

A equação de regressão linear múltipla para pT_{16} e esses dois descritores (μ e π) é mostrada a seguir:

$$pT_{16} = 2,07(\pm 0,20) - 0,19(\pm 0,060)\mu + 0,41(\pm 0,17)\pi \quad (5.6)$$

($n=13$; $R=0,74$; $R^2=0,55$; $F=6,14$; $s=0,22$; $p=0,018$)

Apesar do valor de R^2 não ser expressivo, essa equação é capaz de explicar 55% da variabilidade do modelo num nível de 95% de confiança.

Naturalmente, poderíamos aqui também realizar uma análise conformacional para determinar valores “ótimos” de momento dipolar, de forma a aumentar a correlação linear com pT_{16} . Entretanto, quando assim fizemos nos cálculos AM1, foi com o objetivo de chamar a atenção de que esses descritores podem variar com a conformação, como resultado do próprio ambiente biológico da droga, especialmente se a barreira energética for pequena, como foi o caso dos derivados analisados.

6. Efeito Solvente

Como foi explicado no item 2.1 dos objetivos, foi utilizado o modelo do efeito solvente, a fim de obter outros modelos de QSAR que pudessem explicar melhor a interação fármaco-receptor em um meio biológico. Para isso foi empregado o método COSMOS (*Conductor-like Screening Model*) para o método AM1, enquanto para o funcional B3LYP/6-31G (d,p) foi empregado o método PCM (*Polarizable Continuum Model*). Em ambos os casos a água foi utilizada como solvente.

6.1 Descritores AM1 no Efeito Solvente

A matriz dos descritores eletrônicos AM1/COSMOS e empíricos, juntamente com as atividades hipolipidêmicas, é mostrada abaixo na Tabela 10.

Tabela 10: Atividades hipolipidêmicas, descritores eletrônicos AM1/COSMOS e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.

Compostos	pC ₁₆	pT ₁₆	ε _{HOMO}	ε _{LUMO}	M	qO ₁₅	qO ₁₆	qC ₁₃	qH ₁₄	qCO ₁	qCO ₂	qC ₁₇	π	π ²	logP
1	1,602	1,602	-9,736	-1,037	3,417	-0,319	-0,323	-0,171	0,183	-0,007	-0,009	-0,077	0,000	0,000	2,970
2	1,491	1,708	-9,551	-1,028	3,220	-0,318	-0,323	-0,171	0,181	-0,004	-0,010	-0,083	0,560	0,314	3,300
3	1,462	1,690	-9,584	-1,036	3,225	-0,319	-0,323	-0,171	0,183	-0,007	-0,009	-0,072	0,560	0,314	3,360
4	1,544	1,763	-9,472	-1,035	3,285	-0,320	-0,323	-0,169	0,182	-0,007	-0,009	-0,084	0,560	0,314	3,360
5	1,643	1,653	-9,752	-1,057	4,410	-0,314	-0,316	-0,175	0,182	0,004	-0,004	-0,069	0,710	0,504	3,520
6	1,602	1,477	-9,763	-1,059	4,634	-0,318	-0,321	-0,174	0,187	-0,005	-0,007	-0,062	0,710	0,504	3,580
7	1,568	1,724	-9,653	-1,056	4,648	-0,318	-0,321	-0,173	0,187	-0,005	-0,007	-0,068	0,710	0,504	3,580
8	1,342	1,681	-9,983	-1,075	4,532	-0,314	-0,317	-0,192	0,195	0,001	-0,003	-0,012	-0,320	0,102	2,480
9	1,613	1,279	-9,217	-1,035	4,133	-0,320	-0,323	-0,164	0,181	-0,008	-0,010	-0,112	0,380	0,144	3,320
10	1,342	0,699	-10,07	-1,070	6,079	-0,316	-0,320	-0,182	0,191	-0,002	-0,005	-0,027	-0,320	0,102	2,630
11	1,447	1,716	-9,289	-1,019	1,825	-0,320	-0,320	-0,160	0,185	-0,008	-0,008	-0,113	-0,020	0,000	2,840
12	1,519	1,653	-9,345	-1,044	3,420	-0,318	-0,321	-0,174	0,186	-0,005	-0,007	-0,041	-0,020	0,000	2,900
13	1,613	1,763	-9,223	-1,035	4,187	-0,320	-0,323	-0,165	0,182	-0,008	-0,010	-0,111	-0,020	0,000	2,910

A partir dessa matriz de dados, foi obtida a matriz dos coeficientes de correlação; desta última extraímos uma sub-matriz com os valores mais significativos de R, envolvendo as variáveis previamente selecionadas e as respectivas atividades hipolipidêmicas. Esta sub-matriz é mostrada na Tabela 11. Os valores assinalados em vermelho são os mais altos coeficientes de correlação verificados nesta sub-matriz de correlação.

Tabela 11: Matriz de correlação para os valores de R mais significativos entre as variáveis previamente selecionadas (descritores eletrônicos obtidos dos cálculos AM1/COSMOS e descritores empíricos) e as atividades hipolipidêmicas.

	pC ₁₆	pT ₁₆	ε _{LUMO}	μ	qH ₁₄	qC ₁₇	π	logP
pC ₁₆	1,00	0,34	0,32	-0,11	-0,71	-0,58	0,60	0,69
pT ₁₆	0,34	1,00	0,43	-0,66	-0,34	-0,33	0,35	0,27
ε _{LUMO}	0,32	0,43	1,00	-0,84	-0,78	-0,84	0,19	0,19
μ	-0,11	-0,66	-0,84	1,00	0,49	0,55	-0,11	-0,06
qH ₁₄	-0,71	-0,34	-0,78	0,49	1,00	0,81	-0,55	-0,58
qC ₁₇	-0,58	-0,33	-0,84	0,55	0,81	1,00	-0,34	-0,38
π	0,60	0,35	0,19	-0,11	-0,55	-0,34	1,00	0,98
logP	0,69	0,27	0,19	-0,06	-0,58	-0,38	0,98	1,00

A partir desta matriz, é possível estabelecer duas equações de regressão linear simples, uma para pC₁₆ (equação 5.7) e outra para pT₁₆ (equação 5.8), conforme podem ser vistas a seguir:

$$\mathbf{pC_{16} = 4,68(\pm 0,94) - 17,04(\pm 5,07)qH_{14}} \quad (5.7)$$

(n=13; R = -0,71; R² = 0,51, F=11,28; s=0,074; p = 0,0064)

$$\mathbf{pT_{16} = 2,31(\pm 0,26) - 0,19(\pm 0,064)\mu} \quad (5.8)$$

(n=13; R = -0,66; R² = 0,44; F=8,66; s=0,23; p = 0,013)

Nota-se que as equações acima, utilizando o efeito solvente, são semelhantes àquelas obtidas pelo AM1 no modelo da molécula isolada (equações 5.1 e 5.3,

respectivamente). Confirmando assim que o modelo AM1 é bastante estável, levando-se em conta a inclusão ou não do solvente.

6.2 Descritores B3LYP no Efeito Solvente

Usando o mesmo procedimento anterior para o AM1, nesta seção são estimados os descritores eletrônicos B3LYP na presença do efeito solvente usando agora o modelo PCM. A matriz desses descritores eletrônicos B3LYP/6-31G(d,p)/PCM e empíricos, juntamente com as atividades hipolipidêmicas, é mostrada abaixo na Tabela 12.

Tabela 12: Atividades hipolipidêmicas, descritores eletrônicos B3LYP/6-31G(d,p)/PCM e empíricos da 2-fenilindan-1,3-diona (molécula-mãe) e seus doze derivados.

Compostos	pC ₁₆	pT ₁₆	ε _{HOMO}	ε _{LUMO}	M	qO ₁₅	qO ₁₆	qC ₁₃	qH ₁₄	qCO ₁	qCO ₂	qC ₁₇	π	π ²	logP
1	1,602	1,602	-6,551	-2,192	4,201	-0,502	-0,515	-0,313	0,207	-0,091	-0,108	0,112	0,000	0,000	2,970
2	1,491	1,708	-6,380	-2,189	4,146	-0,503	-0,509	-0,320	0,191	-0,084	-0,111	0,104	0,560	0,314	3,300
3	1,462	1,690	-6,439	-2,186	3,974	-0,502	-0,515	-0,313	0,206	-0,091	-0,108	0,113	0,560	0,314	3,360
4	1,544	1,763	-6,348	-2,184	3,915	-0,503	-0,516	-0,312	0,206	-0,093	-0,108	0,107	0,560	0,314	3,360
5	1,643	1,653	-6,601	-2,201	4,801	-0,497	-0,504	-0,337	0,212	-0,082	-0,104	0,183	0,710	0,504	3,520
6	1,602	1,477	-6,680	-2,241	6,358	-0,500	-0,513	-0,310	0,216	-0,089	-0,105	0,104	0,710	0,504	3,580
7	1,568	1,724	-6,583	-2,237	6,682	-0,501	-0,513	-0,310	0,216	-0,091	-0,106	0,106	0,710	0,504	3,580
8	1,342	1,681	-6,699	-2,168	3,072	-0,502	-0,514	-0,327	0,221	-0,088	-0,108	0,115	-0,320	0,102	2,480
9	1,613	1,279	-5,906	-2,180	4,386	-0,504	-0,517	-0,311	0,205	-0,094	-0,109	0,106	0,380	0,144	3,320
10	1,342	0,699	-6,803	-2,247	6,612	-0,499	-0,513	-0,315	0,217	-0,088	-0,104	0,121	-0,320	0,102	2,630
11	1,447	1,716	-5,986	-2,125	2,590	-0,502	-0,510	-0,304	0,204	-0,089	-0,111	0,093	-0,020	0,000	2,840
12	1,519	1,653	-6,034	-2,182	3,346	-0,501	-0,514	-0,315	0,207	-0,089	-0,108	0,114	-0,020	0,000	2,900
13	1,613	1,763	-5,936	-2,182	4,518	-0,503	-0,516	-0,311	0,206	-0,094	-0,108	0,107	-0,020	0,000	2,910

A partir dessa matriz de dados, foi obtida a matriz dos coeficientes de correlação; desta última extraímos uma sub-matriz com os valores mais significativos de R, envolvendo as variáveis previamente selecionadas e as respectivas atividades hipolipidêmicas. Esta sub-matriz é mostrada na Tabela 13. Os valores assinalados em vermelho são os mais altos coeficientes de correlação verificados nesta sub-matriz de correlação.

Tabela 13: Matriz de correlação para os valores de R mais significativos entre as variáveis previamente selecionadas (descritores eletrônicos obtidos dos cálculos B3LYP/6-31G(d,p)/PCM e descritores empíricos) e as atividades hipolipidêmicas.

	pC ₁₆	pT ₁₆	ε _{LUMO}	μ	qH ₁₄	qC ₁₇	π	logP
pC ₁₆	1,00	0,34	-0,12	0,19	-0,23	0,23	0,60	0,69
pT ₁₆	0,34	1,00	0,51	-0,52	-0,32	-0,07	0,35	0,27
ε _{LUMO}	0,31	0,30	1,00	-0,53	-0,58	-0,35	-0,09	-0,05
μ	0,19	-0,52	-0,95	1,00	0,42	0,14	0,33	0,40
qH ₁₄	-0,23	-0,32	-0,45	0,42	1,00	0,26	-0,23	-0,20
qC ₁₇	0,23	-0,07	-0,21	0,14	0,26	1,00	0,20	0,18
π	0,60	0,35	-0,29	0,33	-0,23	0,20	1,00	0,98
logP	0,69	0,27	-0,36	0,40	-0,20	0,18	0,98	1,00

Assim, como foi verificado para os cálculos B3LYP/6-31G(d,p) no modelo da molécula isolada, aqui também não foi verificado nenhuma correlação significativa entre pC₁₆ e os descritores eletrônicos B3LYP na presença do solvente.

Para pT₁₆ o melhor ajuste linear encontrado é também similar àquele achado para este nível de cálculo sem o efeito solvente, ou seja, B3LYP/MMI, cuja equação de regressão linear múltipla B3LYP/6-31G(d,p)/PCM é mostrada a seguir:

$$\mathbf{pT_{16} = 2,17(\pm 0,22) - 0,16(\pm 0,049)\mu + 0,43(\pm 0,16)\pi} \quad \mathbf{(5.9)}$$

$$(n=13; R = 0,75; R^2 = 0,56; F=6,44; s=0,21; p = 0,016)$$

que é, portanto, muito similar àquela 5.6.

7. Novos Agentes Hipolipidêmicos

Os dados experimentais relativos à redução nos níveis de colesterol e de triglicerídeos apresentados pelos derivados de 2-fenilindan-1,3-dionas, considerando os treze derivados aqui estudados, já revelavam que não há uma correlação linear direta entre pC_{16} e pT_{16} , uma vez que o coeficiente de correlação entre ambos é de apenas 0,34. Isto significa possivelmente que os fatores que governam a redução dos níveis de colesterol devam ser diferentes daqueles governando a redução dos níveis de triglicerídeos. De fato, nossos estudos mostraram, mesmo que de forma limitada, esse fato, pois as equações de regressão encontradas para ambos os níveis de redução são distintas entre si. Isto não implica, por outro lado, que não seja possível projetar um novo agente potencialmente ativo que seja capaz de simultaneamente reduzir os níveis de colesterol e de triglicerídeos no sangue, após 16 dias de aplicação da droga. Para isso, é necessário que o substituinte introduzido no grupo fenil do 2-fenilindan-1,3-diona apresente características eletrônicas que possam satisfazer as condições de otimização de ambos os níveis de redução, em termos de suas atividades hipolipidêmicas.

Os resultados QSAR apresentados e discutidos na seção anterior revelaram que os seguintes aspectos devem ser relevantes para a projeção teórica de um novo agente hipolipidêmico:

1. Quanto menos positiva for a carga atômica do hidrogênio na posição 14, seguindo a numeração adotada nesse trabalho, mais ativo será o derivado no que concerne a redução dos níveis de colesterol, segundo os cálculos AM1, tanto no modelo da molécula isolada, como no modelo do efeito solvente.
2. Quanto menor for o dipolo molecular mais ativo será o derivado no que concerne a redução dos níveis de triglicerídeos, segundo os cálculos AM1 e B3LYP, tanto no modelo da molécula isolada, como no modelo do efeito solvente. Outra condição importante, especificamente para o modelo matemático estabelecido pelos cálculos B3LYP, refere-se a constante de lipofilicidade π do substituinte. Nesse caso, quanto mais positivo for esse descritor mais ativo será o derivado.

Tomando como referência essas características, fomos analisar, na literatura especializada, substituintes que eventualmente pudessem atender a essas condições. Por exemplo, analisamos os valores σ de Hammett de substituintes ligados diretamente ao grupo fenil nas posições *meta* e *para*. Ou ainda, analisando o comportamento eletrônico de substituintes interagindo com o grupo fenil, segundo seus efeitos indutivos e mesoméricos. A partir disso, fizemos algumas simulações com substituintes que não fizeram parte do conjunto de treinamento, e dessas simulações, tomando por base os nossos estudos QSAR, foi possível planejar dois novos agentes potencialmente mais ativos do que àqueles empregados nesse conjunto de treinamento.

O primeiro deles foi o *orto*-hidróxi-2-fenilindan-1,3-diona, cuja estrutura otimizada é mostrada abaixo:

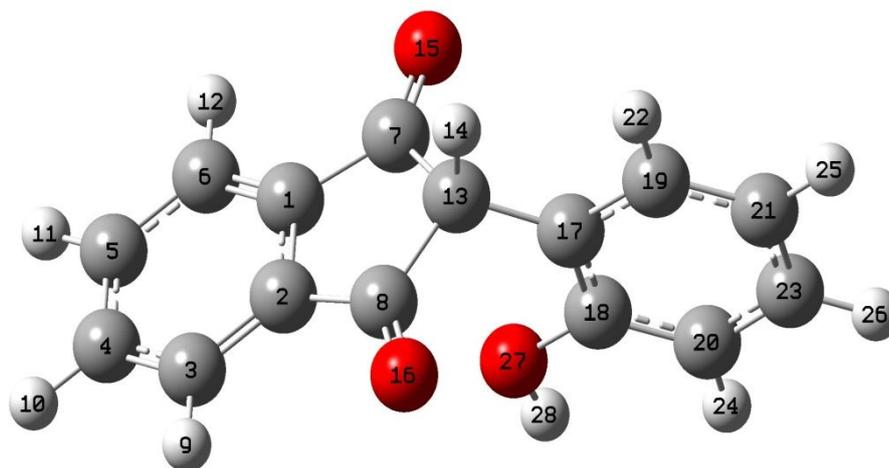


Figura 31: Estrutura otimizada do *orto*-hidróxi-2-fenilindan-1,3-diona.

O grupo $-OH$ tende a doar elétrons e por consequência tende a tornar a carga atômica do hidrogênio, na posição 14, menos positiva. De fato, na molécula-mãe esse valor é $0,153e$, que produz uma redução nos níveis de colesterol em 35%, após 16 dias de aplicação da droga. Mas, substituindo o hidrogênio pelo grupo OH na posição *orto* do grupo fenil, essa carga é reduzida para $0,143e$, e sua atividade prevista pela equação 5.2 é de 57%, tomando por base o modelo matemático estabelecido pelos cálculos AM1, conforme pode ser visto na Tabela 14.

Tabela 14: Comparação na redução dos níveis de colesterol entre a molécula-mãe (2-fenilindan-1,3-diona) e o novo derivado projetado teoricamente (*orto*-hidróxi-2-fenilindan-1,3-diona), segundo cálculos AM1 no modelo da molécula isolada.

Composto	%C ₁₆ [*]	qH ₁₄ (e)
2-fenilindan-1,3-diona	35 (40±2) ^{**}	0,153
<i>orto</i> -hidróxi-2-fenilindan-1,3-diona	57	0,143

*Usando a equação 5.2; **Valor experimental dado entre parênteses;

Conforme indicamos acima, os cálculos AM1 mostraram que a redução dos níveis de triglicerídeos (pT₁₆), depende exclusivamente do potencial do substituinte em diminuir o momento de dipolo da molécula. Isto pode ser visto através da Equação 5.4, que foi usada para determinar pT₁₆. O grupo OH na posição *orto* do grupo fenil produz um momento de 1,828D, bem menor do que aquele da molécula-mãe, que é 2,496D. Por conseguinte, enquanto a redução dos níveis de triglicerídeos projetada teoricamente para a molécula-mãe é de 51%, seu valor correspondente para o *orto*-hidróxi-2-fenilindan-1,3-diona é de 85%, após 16 dias de emprego da droga, conforme podemos ver na Tabela 15.

Tabela 15: Comparação na redução dos níveis de triglicerídeos entre a molécula-mãe (2-fenilindan-1,3-diona) e o novo derivado projetado teoricamente (*orto*-hidróxi-2-fenilindan-1,3-diona), segundo cálculos AM1 no modelo da molécula isolada.

Composto	%T ₁₆ [*]	μ (D)
2-fenilindan-1,3-diona	51 (40±4) ^{**}	2,496
<i>orto</i> -hidróxi-2-fenilindan-1,3-diona	85	1,828

* Usando a Equação 5.4; **Valor experimental dado entre parênteses;

Entretanto, este substituinte possui uma constante de lipofilicidade de $-0,67$, o que não é adequado para ser um potencial agente de redução de níveis de triglicédeos segundo os cálculos B3LYP. De fato, esse percentual cai para apenas 20% usando a equação 5.6.

Isto nos levou a pesquisar outro substituinte, capaz de tornar a carga do hidrogênio menos positiva possível, produzindo um baixo momento dipolar e que possuísse uma constante de lipofilicidade positiva. Isto nos levou ao *orto*-flúor-2-fenilindan-1,3-diona, cuja estrutura otimizada é mostrada abaixo na Figura 32:

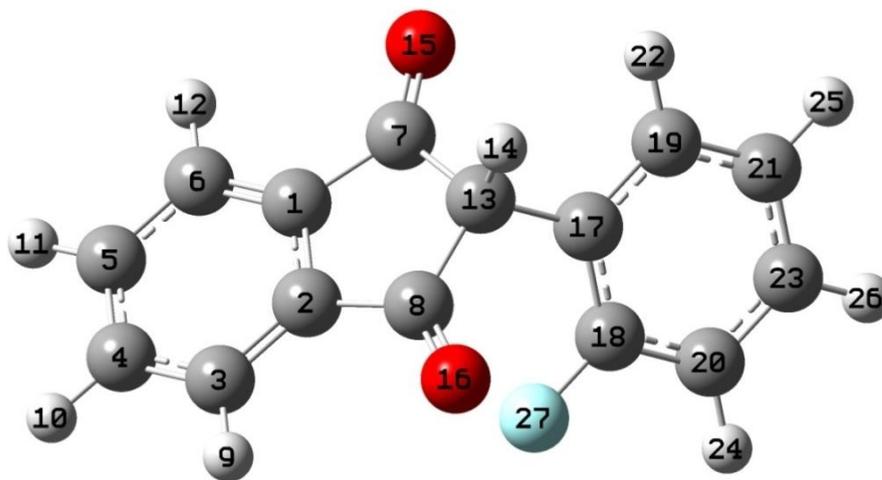


Figura 32: Estrutura otimizada do *orto*-flúor-2-fenilindan-1,3-diona.

Este substituinte produz uma carga no hidrogênio 14 de $0,148e$, mais positiva do que àquela do 2-OH, que como vimos foi de $0,143e$, porém menos positiva do que àquela na molécula-mãe ($0,153e$). Aplicando a equação 5.2, a redução nos níveis de colesterol é de 45%, que é um valor bastante significativo.

Na Tabela 16 mostramos os valores do momento dipolar para o *orto*-flúor-2-fenilindan-1,3-diona. Obtido pelo cálculo B3LYP/6-31G(d,p), e o valor de π deste substituinte. De fato, a introdução de um flúor, na posição *orto* do grupo fenil, produz as características importantes para projetar teoricamente um agente hipolipidêmico potente na redução dos níveis de triglicédeos. Ele reduz esses níveis em 62%, após 16 dias de aplicação da droga. Além disso, o momento dipolar obtido pelos cálculos AM1, na molécula isolada, para o *orto*-flúor-2-fenilindan-1,3-diona é de apenas 1,120D, muito menor que àquela na molécula-mãe (2,496D), prevendo teoricamente um potente agente na redução dos níveis de triglicédeos.

Tabela 16: Comparação na redução dos níveis de triglicerídeos entre a molécula-mãe (2-fenilindan-1,3-diona) e o novo derivado projetado teoricamente (*orto*-flúor-2-fenilindan-1,3-diona), segundo cálculos B3LYP no modelo da molécula isolada.

Composto	% T ₁₆ *	μ (D)	π
2-fenilindan-1,3-diona	30 (40±4)**	3,091	0
<i>orto</i> -flúos-2-fenilindan-1,3-diona	62	1,778	0,140

* Usando a Equação 5.6; ** Valor experimental dado entre parênteses;

8. Conclusões

A partir de nossos estudos de relações quantitativas estrutura-atividade (QSAR), em treze derivados de 2-fenilindan-1,3-dionas, empregando cálculos de orbitais moleculares AM1 e B3LYP/6-31G(d,p), incluindo ou não o efeito do solvente, além de métodos estatísticos multivariados (regressão linear e análise componentes principais), foi possível extrair as seguintes conclusões:

1. Do ponto de vista conformação molecular, nossos resultados revelaram que esses derivados, em relação aos seus mínimos globais de energia, apresentam o que nós poderíamos chamar de uma estrutura “gauche” padrão para os substituintes em posição *meta* ou *para* no anel fenil (conforme Figura 33 (a)), e uma estrutura “quasi-perpendicular” padrão para os substituintes em posição *orto* (Figura 33(b)). Além disso, esses últimos derivados, como esperado, apresentam barreiras de rotação do grupo fenil com respeito à estrutura indan relativamente muito mais alta do que àqueles com substituintes nas posições *meta* e *para*, em decorrência do impedimento estérico.

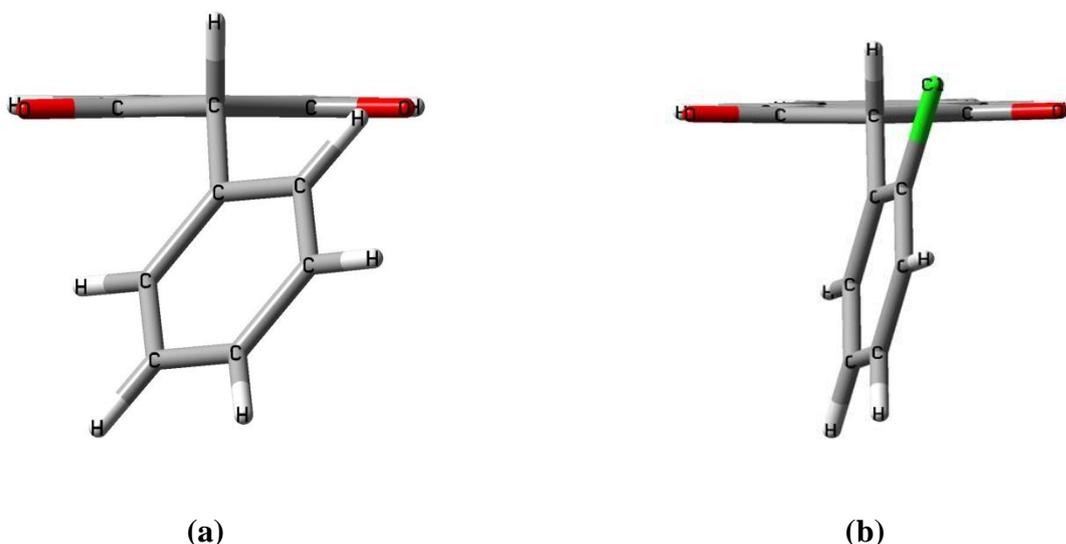


Figura 33: (a) Conformação “gauche” padrão, correspondente ao mínimo global de energia, para substituintes eletrônicos em posição *meta* ou *para* no anel fenil e (b) conformação “quasi-perpendicular” padrão, correspondente ao mínimo global de energia, para substituintes eletrônicos em posição *orto* no anel fenil.

2. Nossos resultados AM1, com ou sem efeito do solvente, revelaram que quanto menos positiva for a carga atômica do hidrogênio (qH_{14} , ligado ao carbono sp^3 do grupo inda, mais ativo será o derivado no que se refere à redução dos níveis de colesterol no sangue, após 16 dias de administração da droga.

Cálculos B3LYP/6-31G(d,p), com ou sem efeito do solvente, apresentaram, por sua vez, uma correlação linear relativamente menor entre essa atividade e a carga atômica do hidrogênio 14, em comparação aos cálculos AM1, levando-se em consideração apenas descritores eletrônicos extraídos dos mínimos globais de energia de cada derivado.

3. Nossos resultados AM1 e B3LYP/6-31G(d,p), com ou sem efeito do solvente, revelaram que o momento dipolar é um importante descritor eletrônico associado à redução dos níveis de triglicerídeos no sangue em derivados de 2-fenilindan-1,3-dionas. Quanto menor for o momento dipolar, mais ativo será o derivado. Em particular, para os cálculos B3LYP um melhor ajuste linear é obtido quando a constante de lipofilicidade de Hansch π é incluída na equação de regressão. Maiores valores de π também contribuem para aumentar a atividade na redução dos níveis de triglicerídeos.

4. A partir dessas constatações foi possível propor dois novos derivados potencialmente mais ativos do que àqueles estudados nesse trabalho: (1) *orto*-hidróxi-2-fenilindan-1,3-diona, na redução de ambos os níveis de colesterol e de triglicerídeos, (2) *orto*-flúor-2-fenilindan-1,3-diona, na redução dos níveis de triglicerídeos, que agora poderão ser alvo de estudos sintéticos e biológicos para comprovar ou não o planejamento molecular aqui proposto.

9. Referências

- [1] V. Cechinel Filho, F. Campos, R. Correia, R. A. Yunes, R. J. Nunes, *Química Nova*, **26**, 230 (2003)
- [2] M. Harfenenist, D .J. Joyner, C. T. Lind, J. F. Batchelor, H. L. White *J. Med. Chem.*, **39**, 1857 (1996).
- [3] L. Costa Tavares, *Química Nova*, **27**, 631 (2004)
- [4] B. G. Katzung., *Farmacologia Básica e Clínica*, Guanabara Koogan, Rio de janeiro, 2001
- [5] E. J. Barreiro, C. A. M. Fraga, *Química Medicinal: As Bases Moleculares da ação de Fármacos*, Artmed Editora, Porto Alegre, 2001.
- [6] M. T. D. Cronin, T. W. Schultz, *J. Mol. Struct (Theo Chem)*, **622**, 39, (2003).
- [7] H. Kubinyi, *QSAR: Hansch Analysis and Related Approaches*, VHC Wheinheim, germany, 1993.
- [8] A. C. Gaudio, *Química Nova*, **19**, 278 (1996)
- [9] M. P. B. Ceneviva, *Química Nova*, **7**, 141 (1984).
- [10] L. P. Hammet, *Chem Rev* **17**, 125, (1935).
- [11] C. Hansch, A. Leo, *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, Washington, USA, 1995
- [12] C. Hansch, T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).
- [13] G. L. Patrick, *An introduction to Medicinal Chemistry*, New York, USA, 2001.

- [14] R. Mannhold, *et al.*, *Quant. Struct.-Act Relat.*, Weinheim, **17**, 578 (1998)
- [15] C. Hansch, J. M. Clayton, *J. Pharm Sci*, **62**, 1 (1973)
- [16] H. D. Höltje, F. Gerd, *Molecular Modeling – Basic Principles and Applications*, Weinheim, 1997
- [17] W. O. Foye, T.L. Lemke, D. A. Williams, *Principles of Medicinal Chemistry*, Baltimore, 1995.
- [18] C. H. T. P. Silva, S. M. Sanches, C.A. Taft, *J. Mol. Graph. Model*, **23**, 89 (2004)
- [19] M. Monastérios, M. Escorche, M. Averdano, *J. Mol. Struct.*, **748**, 49 (2005)
- [20] Y. C. Martin, *Persp. Drug Discov. Design* **3**, 12 (1998)
- [21] G. Cruciani, S. Clementi, M. Pastor, *Persp. Drug Discov. Design* **71**, 12 (1998)
- [22] W. J. Dunn III, A. J. Hopfinger, *Persp. Drug Discov. Design* **167**, 12 (1998)
- [23] B. Beck, T. Clarck, *Persp. Drug Discov. Design* **131**, 9 (1998)
- [24] W. Andreoni, *Persp. Drug Discov. Design* **161**, 9 (1998)
- [25] A. R. K. Murthy, O. T. Wong, D. J Reynolds, I.H. Hall. *Phar. Res.*, **4**, 21 (1987)
- [26] J. M. Chapman, P. J Vordrstd, G. H. Cocolas, I. H. Hall. *J. Med. Chem.*, **26**, 237 (1983)
- [27] J. B. Press, W. B. Wright, P. S. Chan Jr., J. W. Marisco, M.F hang, J. Tauber, A. S. Tomcufcik. *J. Med. Chem.*, **29**, 816 (1986).

- [28] R. Shimazawa, H. Takayama, Y. Fujimoto, M. Komoda, K. Dodo, R. Yamasaki, R. Shirai, Y. Koiso, K. Miyata, F. Kato, Y. Hashimoto. *J. Enzyme Inhibition*, **14**, 259 (1999).
- [29] J. M. Hoffman, J. S. Wai, C. M. Thomas, R. B. Levin, J. A. O'Brien, M. E. Goldman, *J. Med. Chem.*, **35**, 3784 (1992).
- [30] G. Orzeszko, B. Kaminska, A. Orzeszko, B. J. Starosciak. *Farmaco*, **55**, 619 (2000)
- [31] D. R. Sliko, A. D. White. *Trends. Pharmacol. Sci.*, **12**, 194 (1991).
- [32] D. Mathé, F. Chevallier. *Nutrition Reports International*, **15**, 571 (1977).
- [33] C. Lutton, D. Mathé, F. Chevallier, *Biochimica et Biophysica Acta*, **306**, 483 (1973).
- [34] J. M. Chapman, G. H. Cocolas, I. H. Hall, *J. Med. Chem.*, **22**, 1399 (1979).
- [35] I. H. Hall, P. J. Voorstad, M. Chapman, G. G. Cocolas *J. Pharm Sci.*, **72**, 845 (1983).
- [36] J. M. Chapman, G. H. Cocolas, I. H. Hall, *J. Pharm Sci.*, **72**, 1344 (1983).
- [37] J. M. Chapman, D. D. Wyricky, P. J. Voorstad, J. H. Magnire, G. G. Cocolas, I. H. Hall, *J. Pharm., Sci.*, **73**, 1488 (1984).
- [38] M. N. Ramos, B. B. Neto, *J. Comp. Chem.*, **11**, 569 (1990).
- [39] M. D. Neto, W. C. Pinheiro Filho, B. B. Neto, M. N. Ramos, *J. Braz. Chem. Soc.*, **4**, 1399 (1993).
- [40] A. R. Murthy, S. D. Wyrick, I. H. Hall, *J. Med. Chem.*, **28**, 1951 (1985).
- [41] A. M. Simas, Tese de Mestrado, UNICAMP, 1997.
- [42] I. N. Levine, *Quantum Chemistry*, New Jersey, Prentice – Hall International, 1991;

- [43] P. W. Atkins, *Molecular Quantum Mechanics*, Oxford, Oxford Press, 1983;
- [44] C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951)
- [45] P. Gund *et al*, *Three-dimensional modeling and drug design*, *Science* **208**, 1425, 1980
- [46] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **107**, 3902, (1985)
- [47] P. Hohenberg, W. Kohn, *Phys Review*, **136**, B864 (1964)
- [48] W. Kohn, L. J. Sham, *Phys Review*, **140**, A1113 (1965)
- [49] P. Drude, *Annalen der Physik*, **3**, 369, 1900
- [50] C. Reichardt, *Solvents and Effects in Organic Chemistry*, VHC: Weinheim, Germany, 1988.
- [51] A. J. Parker, *Chem. Rev.*, **69**, 1 (1969)
- [52] Y. Mo, J. Gao, *Journal of Computational Chemistry*, **21**, 1558 (2000)
- [53] S. E. Huton, P. J. Rossky, D. A. Zichi, *Journal of the American Chemical Society*, **111**, 5680 (1989)
- [54] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, Wiley, 2002
- [55] P. P. Mager, *Multivariate Chemometrics in QSAR: A Dialogue*. Research Study Press. Grã Bretanha: 1988.

[56] M. F. Pimentel, B. B. Neto, *Quim. Nova*, **19**(3), 268 (1996).

[57] L. Eriksson, E. Johansson, *Chemo and Intell. Lab. Syst.*, **34**, 1 (1996)

[58] M. M. C. Ferreira, A. M. Antunes, M. S. Melo, P. L. O Volpe, *Quim. Nova*, **22**, 724 (1999)