

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE QUIMICA

Este exemplar corresponde à redação final da tese, defendida por Maria Cristina Andreazza Costa, e aprovada pela comissão julgadora.

Campinas, 25 de maio de 1993

Yuji Takahata

**ANALISE DE NEOLIGNANAS SINTETICAS COM EFEITO
SOBRE LEISHMANIOSE, ATRAVES DE RELAÇÕES
ENTRE ATIVIDADE BIOLÓGICA E ESTRUTURA QUÍMICA.**

Maria Cristina Andreazza Costa
TESE DE MESTRADO

ORIENTADOR: Prof. Dr. Yuji Takahata /

- 1993 -

AGRADECIMENTOS

- Agradeço ao professor Yuji Takahata pela maneira amistosa como orientou o projeto, sempre disponível para discussões dos resultados e esclarecimentos de dúvidas, incentivando quando possível, a realização de uma nova idéia ou sugerindo outras melhores, sempre respeitando o meu próprio ritmo de trabalho e dessa forma proporcionando um ótimo ambiente de trabalho.
- Ao professor Yoshiyuki Hase, pelo fornecimento dos dados que tornaram possível a utilização do programa MM2 neste trabalho, pelo auxílio no entendimento do mesmo e pelas proveitosas sugestões apresentadas no exame de qualificação.
- Ao professor Roy Edward Bruns, pela gentileza em auxiliar-me na confecção do capítulo IV, através das valiosas discussões sobre os resultados obtidos com os métodos de Reconhecimento de Padrão e pelo demonstrado interesse no meu aprendizado e entendimento dos mesmos, durante todo o curso de mestrado.
- A professora Márcia M. C. Ferreira, pelo excelente curso de Quântica I ministrado, pela amizade e incentivo através das sugestões feitas nos seminários de grupo.
- Aos professores Antonia T. do Amaral e Hugo Kubiniy, pelo curso de QSAR ministrado, o qual veio elucidar-me no rumo que deveria dar ao trabalho.
- Ao professor Lauro E. S. Barata, pela possibilidade de participação no projeto sobre Leishmaniose que julgo muito interessante, sem a qual este trabalho não teria se iniciado.
- A Sra. Angela Paine e ao Dr. Simon Croft da School of Hygiene and Tropical Medicine, University of London, pelos ensaios anti-Leishmaniose realizados.
- Ao professor e amigo Pedro M. Vazquez, pelo auxílio na utilização dos programas gráficos e pelo constante interesse em ajudar.
- Ao amigo Paolo R. Livotto, pelas sugestões no início do trabalho, pelo auxílio na utilização dos programas de otimização de geometria e pela amizade.
- A todos os amigos do Instituto de Química da Unicamp, pelo agradável ambiente de trabalho proporcionado, em especial, ao Jair, Rosana, Anderson, Ester, Luís, Alexandre, Paulo, Marcelo e Nelson.
- Aos funcionários do Instituto de Química, pelo sempre pronto atendimento e dessa forma, pela contribuição ao andamento da tese.
- Aos queridos amigos da "República" com os quais convivi durante todo o tempo em que desenvolvia o projeto da tese, da forma mais agradável possível, dosando trabalho e diversão: Dila, Bete, Jenny e Rodolfo e aos amigos Denise, Sávio, Alexandre, João e Júnior.
- Ao Fernando, pelos desenhos, correções e pela agradável companhia nos momentos mais difíceis.
- Ao CNPq e à CAPES, pelo suporte financeiro.

Ao Fernando
e aos meus pais.

INDICE

RESUMO	I
ABSTRACT	II
I - INTRODUÇÃO	1
I.1 - OBJETIVOS	8
II - ANALISE CONFORMACIONAL	9
II.1 - INTRODUÇÃO	9
II.2 - CALCULOS	15
II.3 - RESULTADOS E DISCUSSAO	25
II.3.1 - COMPARAÇÃO ENTRE OS CALCULOS REALIZADOS NO VACUO E EM MEIO BIOLOGICO	25
II.3.2 - ANALISE DAS MOLECULAS ATIVAS	31
II.4 - CONCLUSOES	37
III - PARAMETROS FISICO-QUIMICOS	38
III.1 - INTRODUÇÃO	38
III.1.1 - ENERGIAS DO HOMO E LUMO	40
III.1.2 - DUREZA E MOLEZA MOLECULAR	43
III.1.3 - MOMENTOS DIPOLARES	45
III.1.4 - ELETRONEGATIVIDADE DE MULLIKEN	46
III.1.5 - VOLUME MOLECULAR	47
III.1.6 - OUTROS PARAMETROS	48
III.2 - CALCULOS	48
III.3 - RESULTADOS E DISCUSSAO	50

IV - RELAÇÕES ENTRE ATIVIDADE BIOLÓGICA E ESTRUTURA	54
IV.1 - INTRODUÇÃO	54
IV.1.1 - INTERAÇÃO FARMACO-RECEPTOR	54
IV.1.2 - MÉTODOS UTILIZADOS PARA CORRELACIONAR ATIVIDADE BIOLÓGICA COM ESTRUTURA QUÍMICA	57
IV.2 - CÁLCULOS	60
IV.3 - RESULTADOS E DISCUSSÃO	61
IV.3.1 - ANÁLISE DOS RESULTADOS, CONSIDERANDO RECEPTOR RÍGIDO	61
IV.3.2 - ANÁLISE DOS RESULTADOS, CONSIDERANDO RECEPTOR FLEXÍVEL	68
IV.4 - CONCLUSÕES	73
V - CONCLUSÕES GERAIS	74
APÊNDICE 1 - O MÉTODO DE MECÂNICA MOLECULAR	76
APÊNDICE 2 - O MÉTODO MINDO-PM3	85
APÊNDICE 3 - MÉTODOS DE RECONHECIMENTO DE PADRÃO	89
APÊNDICE 4 - MÍNIMOS LOCAIS ENCONTRADOS PARA OS COMPOSTOS ATIVOS E INATIVOS	94
REFERÊNCIAS	115

INDICE DE TABELAS

Tabela 1 - Neolignanas sintéticas, sem atividade biológica. Os ensaios biológicos foram feitos <i>in vitro</i>	6
Tabela 2 - Neolignanas sintéticas, ativas contra Leishmaniose. Os ensaios biológicos foram feitos <i>in vitro</i> , mas os valores numéricos não são necessários para a análise realizada neste trabalho.	7
Tabela 3 - Ordem decrescente de estabilidade para as conformações das moléculas inativas, com base nos dados de energia do apêndice 4, no vácuo e meio biológico.	29
Tabela 4 - A conformação D, com os anéis mais distantes é menos estável do que F e G, onde a distância entre eles é menor, devido à variações nas energias de atração ($E_{atr.}$) e repulsão ($E_{rep.}$).	30
Tabela 5 - Cada mínimo local possui conformações com energia variando de E até $E + 2RT$ (onde E é a energia de menor valor do mínimo local) e a função de partição Z para cada mínimo local é mostrada, bem como as probabilidades; com os dados de energia mínima e de probabilidade, obtivemos as ordens de estabilidade e de probabilidade, respectivamente. A tabela continua na página seguinte.	33
Tabela 6 - Métodos utilizados no cálculo dos parâmetros físico-químicos, onde χ é a eletronegatividade de Mulliken, η é a dureza molecular, $1/\eta$ é a moleza, μ é o momento dipolar, T1 e T2 são os ângulos diedros definidos no item II.2 e ΔH_f é o calor de formação.	49
Tabela 7 - Parâmetros físico-químicos calculados para os compostos inativos e ativos, na conformação E e para o conjunto teste (XXI a XXV) também na conformação E.	52
Tabela 8 - Parâmetros físico-químicos calculados para os compostos inativos e ativos, na conformação mais provável e para o conjunto teste (XXI a XXV) também na conformação mais provável.	53
Tabela 9 - Compostos utilizados no conjunto teste.	61
Tabela 10 - Resultados da análise de componentes principais para os compostos na conformação E.	62

Tabela 11 - Resultados obtidos com o método SIMCA para os compostos na conformação E.	66
Tabela 12 - Resultados obtidos com o método SIMCA para o conjunto teste (conformação E).	68
Tabela 13 - Resultados obtidos com a análise de componentes principais, para os compostos na conformação mais provável.	69
Tabela 14 - Resultados obtidos com o método SIMCA, para os compostos na conformação mais provável.	72
Tabela 15 - Resultados obtidos com o método SIMCA, para o conjunto teste na conformação mais provável.	72
Tabela 16 - Valores testados para os ângulos diedros T1 e T2, na referência 50.	84

INDICE DE FIGURAS

Figura 1 - Os compostos a (Surinamensina) e b (Virolina) foram isolados do extrato bruto de folhas de <i>Virola surinamensis</i> e mostraram atividade biológica contra <i>Schistosoma mansoni</i>	2
Figura 2 - Síntese de análogos de neolignanais.	3
Figura 3 - Ciclo dos parasitas causadores da Leishmaniose.	4
Figura 4 - Representação esquemática da variação da energia potencial conformacional com um ângulo de torção variável.	9
Figura 5 a) - Variação da energia potencial conformacional com dois ângulos de torção, T1 e T2, para o composto XII deste trabalho; b) - superfície de potencial para o mesmo composto (XII), mostrando os dois eixos torsionais e o eixo de energia (E); os vales são mostrados com letras maiúsculas e a nomenclatura dos mesmos será explicada no item II.2. c) - Os ângulos de torção T1 e T2 correspondem aos ângulos diedros formados pelos átomos: 14-13-12-6 e 18-17-13-12, respectivamente.	10
Figura 6 - Variação da energia potencial com a conformação, mostrando um mínimo local raso (A) e um profundo (B).	11
Figura 7 - As constantes torsionais para a molécula I, referentes aos ângulos designados pelos átomos 1-6-12-16 e 1-6-12-13 foram parametrizadas.	14
Figura 8 - Os compostos acetofenona e benzaldeído são mostrados; os mesmos foram usados na parametrização de constantes torsionais, usando-se como dados de referência as barreiras de rotação sobre a ligação T1 da acetofenona e do benzaldeído.	14
Figura 9 - O composto I foi girado sobre as ligações T1 e T2, a cada 15 graus e a energia em cada ponto foi determinada.	16
Figura 10 -	17
Figura 11 - Mapas conformacionais para os compostos ativos; as setas indicam os limites dos vales, correspondendo à região mínima (E) e E somado de 2RT.	18
Figura 12 - O radical metóxi foi girado sobre a ligação x, a cada 10 graus, em cada um dos oito mínimos locais encontrados para a molécula I.	22

Figura 13 - Interação do HOMO de uma molécula com o HOMO de outra.	40
Figura 14 - Interação de um orbital ocupado de uma molécula com um desocupado de outra.	41
Figura 15 - Interação do HOMO de uma molécula com o LUMO de outra.	42
Figura 16 - Ligação iônica formada pela interação de orbitais com grande diferença de energia.	44
Figura 17 - Interação entre os substratos ADP e glicose-6-fosfato, e o sítio ativo da enzima.	54
Figura 18 - Representação dos N objetos num espaço M dimensional, onde M é o número de medi- das feitas sobre esses N objetos.	58
Figura 19 - Gráfico traçado usando-se as duas primeiras componentes principais, para os compostos na conformação E.	63
Figura 20 - Dendograma obtido com o método "Agrupamento por Hierarquia", para os compostos na conformação E.	65
Figura 21 - Comparação do composto XXIII do conjunto teste com os compostos XI e XVII.	67
Figura 22 - Comparação do composto XXIV do conjunto teste com os compostos VI e XIII.	67
Figura 23 - Gráfico traçado com as duas componentes principais, para os compostos na conformação mais provável.	70
Figura 24 - Dendograma obtido com o método "Agrupamento por Hierarquia", para os compostos na conformação mais provável.	71
Figura 25 - Curvas mostrando o comportamento harmônico e anarmônico do estiramento de ligação: , função de Morse; , potencial harmônico; , potencial harmônico com o termo cúbico adi- cionado.	78
Figura 26 - Definição dos termos geométricos que aparecem na equação 24.	79
Figura 27 - Estrutura básica dos compostos estudados neste trabalho, com oxigênio na "ponte". Os substituintes estão representados por "R".	81
Figura 28 - As constantes torsionais referentes aos átomos 13-12-6-1 e 14-12-6-1 foram parametriza- das.	82
Figura 29 - As constantes torsionais referentes aos átomos 16-13-12-6 e 14-13-12-6 para a molécula XV, foram parametrizadas e usadas posteriormente na molécula XVI.	83
Figura 30 - Ângulos diedros referentes aos átomos 16-13-12-6 e 17-16-13-12 foram encontrados na	

literatura e utilizados para parametrizar constantes torsionais para a molécula XV deste trabalho.	83
Figura 31 - A otimização de geometria da molécula XV foi feita, inicialmente, para o composto sem os substituintes.	84
Figura 32 - Representação dos dados Y num espaço tridimensional, e do ponto central,	90
Figura 33 - Representação da primeira componente principal, num espaço tridimensional.	91
Figura 34 - Representação das duas primeiras componentes principais num espaço tridimensional.	92
Figura 35 - Representação de modelos com números diferentes de componentes principais, para cada categoria.	93

RESUMO

Neolignanas sintéticas, tendo apresentado atividade biológica anti-Leishmaniose, foram analisadas com o intuito de se investigar os fatores, em termos de parâmetros físico-químicos, responsáveis pela atividade biológica e dessa forma sugerir novos compostos possivelmente mais ativos.

Como primeira etapa do trabalho, foi realizada uma análise conformacional dos compostos, com o método de mecânica molecular, MM2, a fim de se obter a conformação mais estável para cada um deles, uma vez que estes dados não são disponíveis na literatura. Essa análise foi feita para os compostos em meio biológico e também no vácuo, com finalidades comparativas. Uma análise populacional usando-se a distribuição de Boltzmann foi realizada e assim pudemos ter uma idéia das conformações de mínima energia realmente significativas; essa análise auxiliou também as comparações feitas para os compostos no vácuo e em meio biológico.

Na segunda etapa, alguns parâmetros eletrônicos e estruturais foram obtidos para os compostos com as geometrias otimizadas em meio biológico, através do método semi-empírico MNDO-PM3 e do método de mecânica molecular, MM2. Esses parâmetros foram então utilizados na terceira etapa, onde procurou-se correlacionar atividade biológica com estrutura química, usando-se métodos de reconhecimento de padrão.

Algumas sugestões de compostos possivelmente mais ativos foram feitas.



ABSTRACT

Some synthetic neolignans have shown biological activity *in vitro* against Leishmaniasis. We have analysed some of these compounds using computational methods to determine which physical properties may be responsible for activity. In this connection we hope to propose structures of new compounds looking for higher activity. In order to obtain our aim we have studied both active and inactive compounds.

In the first step of this work we have performed conformational analysis using the molecular mechanic programme, MM2, both in a vacuum and biological environment to find out which conformation is favoured for the activity and to compare and establish the importance of biological environment.

Parameters calculations are being performed using the active and inactive compounds supposing their favoured conformations. Finally, biological activity and chemical structures relationship are calculated using statistical methods - Pattern Recognition.

Using these methodologies new compounds which may have biological activities higher than the originals were suggested.

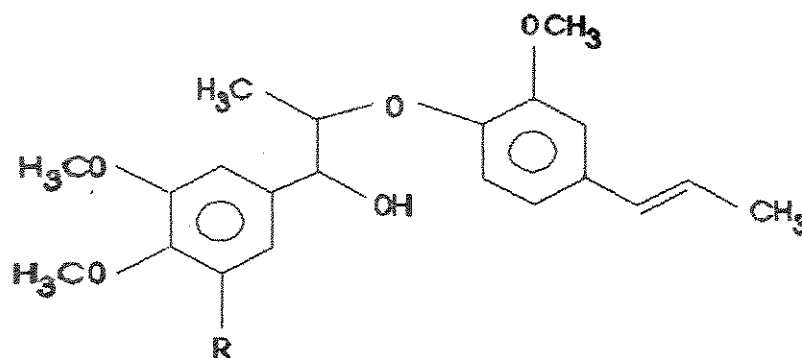
CAPITULO I - INTRODUÇÃO

Calcula-se que o Brasil disponha de aproximadamente 120 mil espécies vegetais (Barata, L.E.S - comunicação pessoal) e provavelmente 40% delas deve conter propriedades terapêuticas. Entretanto, o país chega a importar matéria-prima para 90% dos remédios sintéticos consumidos pela população, gastando com isto, 2 a 3 bilhões de dólares por ano¹. Esses números mostram a necessidade de se explorar a nossa flora, no sentido de se investigar as espécies com propriedades medicinais e seus princípios-ativos, com a finalidade de produzir-se novos medicamentos, mais baratos, talvez mais eficientes do que aqueles usados atualmente e com a possibilidade de se encontrar espécies ativas para doenças, às quais a cura não é ainda conhecida.

Ultimamente, a química e atividade biológica de um certo grupo de compostos, denominados Neolignanais, têm despertado grande interesse.

Neolignanais são dímeros formados pelo acoplamento oxidativo de alil e propenil fenóis; esses compostos são encontrados em plantas da família *Myristicaceae*, sendo o gênero *Viola* o mais representativo nas Américas².

Em 1970, foram realizados os primeiros estudos pelo grupo do professor L.E.S. Barata (IQ - UNICAMP), das folhas de *Viola surinamensis*, uma árvore da família *Myristicaceae*, abundante na Amazônia. O extrato bruto dessas folhas apresentou atividade nos ensaios de proteção contra a penetração de cercárias do *Schistosoma mansoni* em camundongos. O fracionamento desse extrato levou ao isolamento de uma mistura de duas neolignanais denominadas Surinamensina (a) e Virolina (b), mostradas na figura 1².



a) R = -OCH₃ (Surinamensina)

b) R = H (Virolina)

Figura 1 - Os compostos a (Surinamensina) e b (Virolina) foram isolados do extrato bruto de folhas de *Virola surinamensis* e mostraram atividade biológica contra *Schistosoma mansoni*.

Com a finalidade de investigar a atividade biológica de neolignanas, o mesmo grupo citado acima sintetizou (figura 2) uma série de compostos análogos àqueles naturais da figura 1, e os submeteu à avaliações nos ensaios antifúngico, antibacteriano, anti-Leishmaniose, anti-Esquistossomose, anti-PAF (Fator de Agregação Plaquetária), anti-Câncer, tendo mostrado especial interesse em se investigar os fatores responsáveis pela atividade anti-Leishmaniose. A escolha da série de compostos com atividade anti-Leishmaniose foi feita apenas com critérios empíricos, usando-se "intuição química"; por esse motivo, teve início o presente trabalho, no qual pretendemos, através de métodos computacionais, definir tais fatores e sugerir novos compostos para síntese.

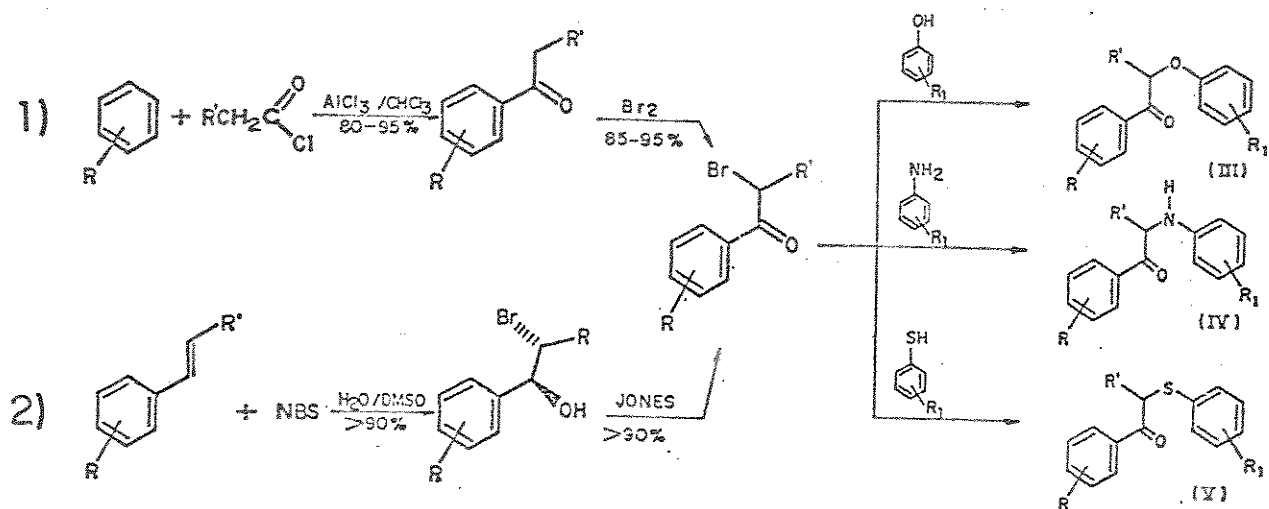


Figura 2 - Síntese de análogos de neolignanas³.

A Leishmaniose é causada por um protozoário que recebeu o nome genérico de *Leishmania*, em homenagem ao bacteriologista escocês W. B. Leishman (1865 - 1929) que o identificou, em 1903. Essa doença atinge cerca de 400 mil pessoas por ano, não só nos países mais pobres, mas também nos desenvolvidos e apresenta-se sob duas formas principais: a visceral e a cutânea. A primeira é quase sempre fatal quando não tratada, ocorrendo principalmente no baço, fígado, medula óssea, pulmões e intestinos. A forma cutânea, conhecida também por úlcera de Bauru, pode atingir as mucosas, destruindo progressivamente os tecidos da cavidade nasal, faríngea e da boca⁴.

Os protozoários causadores das diversas formas de Leishmaniose, hospedam-se primeiro em insetos da subfamília *Phlebotominae* conhecidos por mosquitos-palha no sul do Brasil e birigüi no Norte e Nordeste, depois em várias espécies de mamíferos e no homem. Nos primeiros, os parasitas se desenvolvem sob a forma de promastigotas, fixando-se no aparelho digestivo do inseto-vetor; nos mamíferos, tornam-se amastigotas (Fig.3).

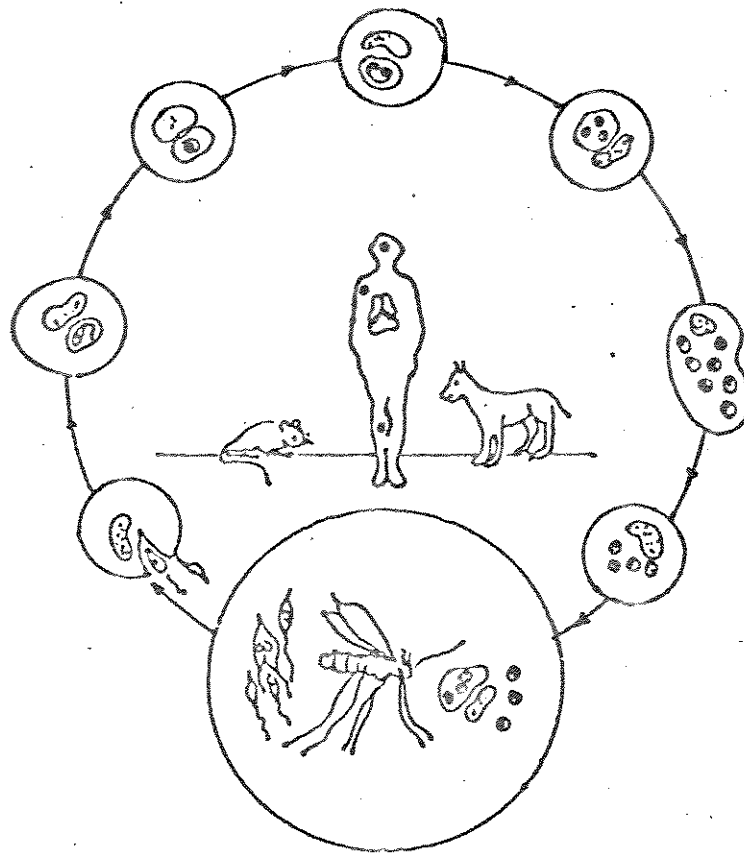


Figura 3 - Ciclo dos parasitas causadores da Leishmaniose⁴.

A transmissão da doença, ocorre da seguinte forma: o inseto transmissor (somente a fêmea se alimenta de sangue), ao sugar o mamífero infectado, ingere as formas amastigotas, que no seu tubo digestivo se transformam em promastigotas. Quando o inseto pica o homem, transmite-lhe as formas promastigotas do protozoário que são fagocitadas pelas células da pele que possuem essa função. Nelas, as promastigotas retornam à forma amastigota, reproduzem-se, rompem a célula e invadem novas células do organismo humano (Fig.3)⁴.

Existem várias espécies de *Leishmania* e a forma visceral, abordada neste trabalho, é causada por *Leishmania donovani*. A Leishmaniose visceral atinge, normalmente, populações rurais e grupos etários mais jovens (0-14 anos), podendo atingir também adultos. As manifestações clínicas são, geralmente, febre irregular, aumento de

volume do baço e do fígado. O comprometimento dos pulmões e intestinos acarreta pneumonia, bronquite, tosse, vômitos, anorexia e enterite. No Brasil, a Leishmaniose visceral tem sido detectada em todos os estados costeiros, do Pará ao Paraná, e no interior de Minas Gerais, Goiás e Mato Grosso do Sul, mas o maior número está no Nordeste, sobretudo no Ceará⁴. A prevenção da doença depende em grande parte de evitar-se o contato com o vetor. Os medicamentos, nem sempre são efetivos e requerem repetidos e extensos períodos de tratamento, produzindo freqüentes efeitos colaterais, como diarreia, vômito, anorexia, dores de cabeça e às vezes, lesão renal⁴. É necessário, portanto, desenvolver novos medicamentos mais eficientes e com menores efeitos colaterais.

Os compostos estudados neste trabalho são mostrados nas tabelas 1 e 2 e foram ensaiados contra *Leishmania donovani* na London School of Hygiene and Tropical Medicine - University of London, sob coordenação do Dr. Ralph Neal. Os ensaios *in vivo* foram feitos apenas para alguns dos compostos listados nas tabelas 1 e 2, enquanto que os testes *in vitro* foram realizados para todos; portanto, utilizaremos os resultados dos testes *in vitro*. Nos ensaios *in vitro* foram utilizados macrófagos peritoniais infectados com amastigotas obtidas de hamsters; as drogas foram adicionadas ao meio de cultura e após 7 dias de exposição à droga, os macrófagos foram fixados e tingidos com Giensa. A proporção dos macrófagos livre de amastigotas foi determinada microscopicamente. As concentrações iniciais de cada composto testado foram de 300 e 100 μ M, e onde a eficácia foi observada, a concentração foi reduzida (3 diluições) nos experimentos subseqüentes, até que uma baixa ou nenhuma atividade foi mostrada. Os resultados foram analisados utilizando análise de regressão linear e teste T para a significância estatística⁵. Neste trabalho, serão estudadas moléculas ativas (tab.2) e inativas (tab.1), uma vez que o método escolhido para tal análise (cap.IV) requer os dois tipos de compostos; além disso, o número de compostos ativos é muito pequeno para uma análise estatística qualquer. Sendo a Leishmaniose, a segunda mais importante doença causada por protozoários (perdendo só para a malária) e, tendo-se em mãos os resultados dos testes biológicos feitos com os análogos às neolignanas naturais da figura 1, achamos que a aplicação dos métodos teóricos a estes resultados, para correlacionar atividade biológica com estrutura química, seria uma boa maneira de

se estudar esses métodos, no sentido de verificar suas limitações e aplicabilidade a este tipo de problema.

MOLECULAS INATIVAS	ESTRUTURA
I	
II	
III	
IV	
V	
VI	
VII	
VIII	
IX	
X	
XI	

Tab.1 Neolignanas sintéticas, inativas contra Leishmaniose. Os ensaios biológicos foram feitos *in vitro*.

MOLÉCULAS ATIVAS	ESTRUTURA
XII	
XIII	
XIV	
XV	
XVI	
XVII	
XVIII	
XIX	
XX	

Tabela 2: Neolignanas sintéticas, ativas contra Leishmaniose. Os ensaios biológicos foram feitos *in vitro*, mas os valores numéricos não são necessários para a análise realizada neste trabalho.

I.1 - OBJETIVOS

Pretendemos, neste trabalho, definir os fatores responsáveis pela atividade biológica, e através desses fatores, correlacionar a estrutura química com atividade, usando apenas química teórica e computacional. Com esses resultados, procuraremos estabelecer uma metodologia à procura de novos novos compostos para síntese.

Pretendemos ainda, verificar a necessidade de realizar os cálculos em meio biológico, uma vez que estamos trabalhando com fármacos e que os mesmos vão atuar em organismos vivos; os fármacos são considerados moléculas flexíveis que podem alterar a sua conformação no momento de interação com o receptor, como veremos nos próximos capítulos. Portanto, desejamos investigar as possíveis modificações nos resultados, quando este fato é levado em consideração.

A análise desses compostos, usando-se apenas química teórica, nos leva também a um último objetivo: o de estudar alguns métodos teóricos, possibilitando-nos verificar suas limitações e aplicabilidade em QSAR (Relações Quantitativas entre Estrutura Química e Atividade Biológica).

CAPITULO II - ANALISE CONFORMACIONAL

II.1) INTRODUÇÃO

A primeira etapa num trabalho onde se pretende planejar novos compostos, com base nos testes biológicos feitos para uma série de análogos, consiste em averiguar a conformação de cada um desses análogos no momento da interação com o receptor. Duas abordagens diferentes são feitas a esse respeito.

A primeira considera a possibilidade da droga alterar a sua conformação preferida (de mínima energia) para ajustar-se ao receptor, mas uma outra idéia seria a de que a droga ligar-se-ia imediatamente ao receptor, mantendo a sua conformação preferida. Entretanto, a primeira possibilidade parece mais provável, de acordo com estudos mais recentes⁶. Sendo assim, é necessário realizar-se uma análise conformacional a fim de se obter a conformação com a maior probabilidade de ser aquela que se ajusta ao receptor. Para tal análise, são feitos cálculos para uma série de posições de uma parte da molécula em relação à outra parte e a energia de cada posição é comparada.

Se há apenas uma ligação na molécula, sobre a qual pode ocorrer rotação, os resultados podem ser apresentados na forma de uma curva de energia *versus* ângulo (Fig.4).

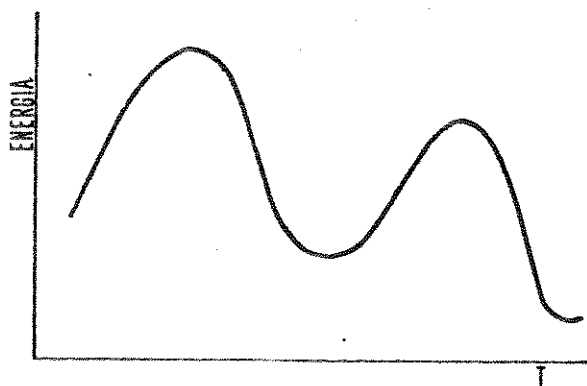


Figura 4 - Representação esquemática da variação da energia potencial conformacional com um ângulo de torção variável.

agrama tridimensional é obtido (energia, ângulos de torção T1 e T2) conforme as figuras 5.a e 5.b que mostram o mapa conformacional do composto XII estudado neste trabalho. A figura 5.c mostra os eixos T1 e T2 do composto XII, sobre os quais é feita a rotação.

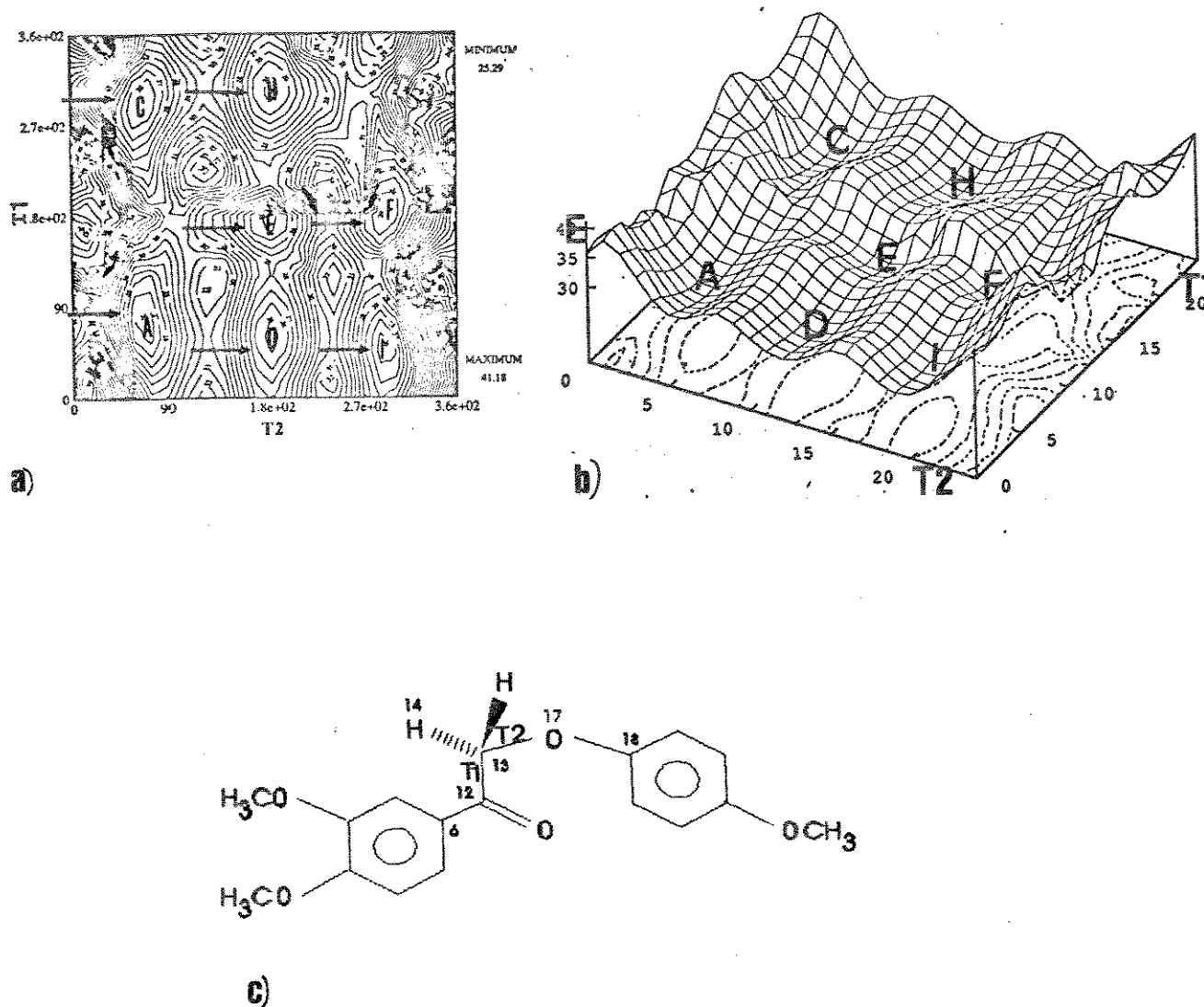


Fig.5 a) - Variação da energia potencial conformacional com dois ângulos de torção, T1 e T2, para o composto XII deste trabalho; b) - superfície de potencial para o mesmo composto (XII), mostrando os dois eixos torsionais e o eixo de energia (E); os vales são mostrados com letras maiúsculas e a nomenclatura dos mesmos será explicada no item II.2. c) - Os ângulos de torção T1 e T2 correspondem aos ângulos diedros formados pelos átomos: 14-13-12-6 e 18-17-13-12, respectivamente.*

* A numeração dos átomos em todas as moléculas do texto segue àquela obtida da saída do programa de otimização de geometria utilizado.

Algumas considerações devem ser feitas na análise da conformação de fármacos e a primeira delas é a de que os cálculos devem ser feitos em meio biológico. Deve-se observar também a flexibilidade das moléculas, isto é, a facilidade com que a molécula passa da conformação de um mínimo local para uma outra e ainda deve-se lembrar que as populações relativas de dois mínimos locais na superfície de potencial dependerá da largura do poço nessa superfície. Por exemplo, na figura 6 , apesar da conformação B ser de energia mais baixa do que de A, a população em A pode ser maior, já que um maior número de conformações está contido nesse poço, em comparação com B⁶.

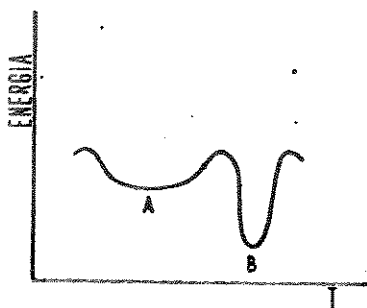


Figura 6 - Variação da energia potencial com a conformação, mostrando um mínimo local raso (A) e um profundo (B)⁶.

A população relativa entre os mínimos locais de energia pode ser calculada usando-se a função de partição molecular (Z), a qual indica como as moléculas estão distribuídas pela superfície de energia. Essa função é descrita da seguinte maneira:

$$Z = \sum_i e^{-E(i)/kT}, \quad \text{eq.1}$$

onde E é a energia calculada para determinada conformação, k é a constante de Boltzmann, T é a temperatura e t é o ângulo que define a conformação da molécula.

Para verificarmos a probabilidade relativa entre os mínimos locais encontrados na molécula, devemos associar cada mínimo com a probabilidade:

$$Z(t) = e^{-E(t) / kT} \quad \text{eq.2}$$

E feito, então, um somatório das funções de probabilidade sobre todos os mínimos, de forma a se obter a função Z (eq.1). Agora, cada função de probabilidade (associada a cada mínimo) pode ser normalizada:

$$Z'(t) = Z(t) / Z \quad \text{eq.3}$$

Como já foi citado anteriormente, a população em cada poço de potencial depende da sua largura. Portanto, deve-se antes definir os limites dos poços, isto é, onde termina um e onde começa o outro. A porção de energia considerada normalmente como pertencente a um determinado poço, é o valor da energia do ponto de mínimo local, somado de $2kT$ ⁶. Assim, na figura 6, cada mínimo local não é considerado apenas como sendo o fundo do poço onde a energia é mínima, mas também toda a superfície em torno desse ponto, definida pela energia mínima mais $2kT$. Portanto, se é necessário grande rotação nas ligações da molécula para atingir o limite do poço ($E_{\text{min.}} + 2kT$), verifica-se que o poço é largo como o poço A da figura 6 e a sua probabilidade alta. Ao contrário, se uma pequena rotação já leva ao limite do poço, este será estreito como o poço B da figura 6, significando que um número menor de conformações está dentro deste poço (baixa probabilidade).

A superfície de potencial completa para uma molécula, isto é, com os ângulos torsionais tendo sido variados de 0 a 360 graus, é chamada de mapa conformacional e toda a informação necessária para esse tipo de análise está contida no mapa. Assim, observando os mapas conformacionais de uma série de análogos ativos, pode-se procurar alguma indicação das conformações favoráveis à atividade. Comparando-se ainda os mapas conformacionais das moléculas ativas e inativas, pode-se verificar se dentro das conformações favoráveis à atividade, existe alguma proibida às inativas (não correspondente a um mínimo local). Se existir, esta será justamente aquela "essencial" à atividade, pois, além de favorecer a atividade, esta não é permitida para as inativas⁶.

Entre os métodos teóricos usados para obter os pontos de mínima energia para

moléculas grandes (no caso deste trabalho, temos moléculas com número de átomos entre 28 e 52), estão os métodos semi-empíricos e o método de mecânica molecular. Preferimos o método de mecânica molecular, ao invés dos métodos semi-empíricos, principalmente pela redução no tempo computacional que o primeiro proporciona, pois um cálculo feito com qualquer método semi-empírico para a molécula I (tabela 1), a qual contém 28 átomos, requer um tempo computacional de aproximadamente uma hora, enquanto que o método de mecânica molecular requer apenas alguns minutos (aproximadamente dez).

Assim, os 576 cálculos citados no item II.2, necessários para a análise conformacional da molécula I puderam ser concluídos em 96 horas com o método MM2, mas com um método semi-empírico, a mesma análise necessitaria de um tempo computacional de 576 horas. Como não seria possível dispender-se de tanto tempo numa só molécula, teríamos que reduzir o número de cálculos, isto é, ao invés de girarmos os eixos T1 e T2 a cada 15 graus, como foi feito, deveríamos fazê-lo a cada 30 ou 45 graus, o que reduziria muito a eficiência desta análise.

Portanto, considerando a necessidade do grande número de cálculos a ser realizado e do tempo computacional de cada método, concluímos que o método de mecânica molecular seria mais eficiente para este conjunto de moléculas.

O método de mecânica molecular emprega um conjunto de equações que são funções potenciais, no cálculo de propriedades moleculares. Esse conjunto de equações é chamado "Campo de Força" e contém parâmetros ajustáveis que são otimizados para obter o melhor conjunto de propriedades para as moléculas, como geometrias, energias conformacionais, calores de formação e outros. Em mecânica molecular, esses parâmetros uma vez ajustados podem ser transferidos de uma molécula para outra; assim, pode-se ajustar um conjunto de parâmetros para moléculas simples e os mesmos podem ser usados nos cálculos para uma molécula maior⁷.

Neste trabalho, por exemplo, foi necessário ajustar alguns parâmetros torsionais para a molécula I, pois precisávamos das constantes torsionais referentes aos átomos 1-6-12-16 e 1-6-12-13 (fig.7).

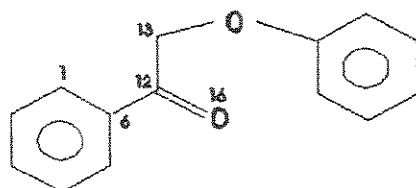
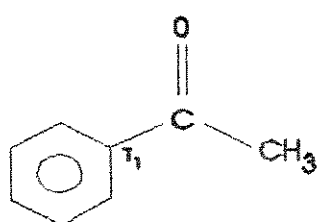
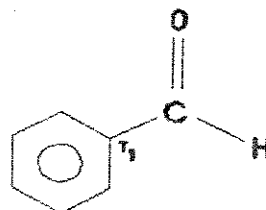


Figura 7 - As constantes torsionais para a molécula I, referentes aos ângulos designados pelos átomos 1-6-12-16 e 1-6-12-13 foram parametrizadas.

O ajuste dos parâmetros foi feito então para as moléculas similares e mais simples, acetofenona e benzaldeído, mostradas na figura 8.



Acetofenona



Benzaldeído

Fig.8 Os compostos acetofenona e benzaldeído são mostrados; os mesmos foram usados na parametrização de constantes torsionais, usando-se como dados de referência as barreiras de rotação sobre a ligação T1 da acetofenona e do benzaldeído.

Para essas duas moléculas, foi possível encontrar dados experimentais na literatura como barreiras de rotação, sobre os eixos T1, os quais foram usados como referência ao ajuste dos parâmetros torsionais (apêndice 1). Finalmente, estes parâmetros assim obtidos puderam ser aproveitados para todas as moléculas deste trabalho.

Um campo de força simples em mecânica molecular inclui estiramento de ligação, deformação angular, torsão e interações de van der Waals:

$$V = \sum V_{est.} + \sum V_{def.ang.} + \sum V_{tor.} + \sum V_{vdW}, \quad eq.4$$

onde "est." é o estiramento de ligação, "def.ang." é deformação angular "tor." é a torsão de ligação e "vdW" é interação de van der Waals.

Assim, na equação 4 as somas são feitas sobre todas as ligações, ângulos de ligação, ângulos de torsão e sobre os átomos não ligados. A soma destes termos é chamada energia estérica da molécula. Para que o campo de força possa reproduzir dados experimentais, os parâmetros devem ser ajustados ou até a inclusão de novos parâmetros pode ser considerada⁷.

O apêndice 1 mostra uma explicação mais detalhada sobre este método.

II.2 CALCULOS

Os cálculos de otimização de geometria foram feitos com o método de mecânica molecular MM2 (campo de força de 1987), acrescido de alguns parâmetros obtidos com uma nova versão do MM2, denominada MM2'⁸, que está brevemente explicada no apêndice 1. Algumas parametrizações necessárias foram realizadas e outras cedidas pelo professor Hase (IQ - UNICAMP).

Iniciamos a otimização de geometria pela molécula mostrada na figura 9 (composto I da tabela 1), uma das menores deste trabalho, contendo 28 átomos e possuindo duas ligações com rotação livre que foram investigadas: a ligação 12-13 e 13-17; vamos chamá-las de ligações T1 e T2 respectivamente, daqui em diante. Consideraremos a molécula I como a estrutura básica de outras moléculas análogas que possuem apenas alguns átomos de hidrogênio substituídos por outros radicais.

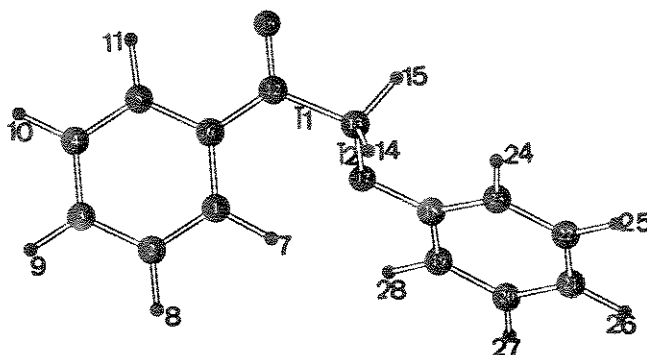


Fig.9 O composto I foi girado sobre as ligações T1 e T2, a cada 15 graus e a energia em cada ponto foi determinada.

Assim, a parte da molécula que compreende os átomos de número 1 a 16 (parte 1) foi mantida numa posição fixa, enquanto a outra parte que compreende os átomos de 17 a 28 (parte 2) foi girada sobre a ligação T2 (através do ângulo diedro determinado pelos átomos 18-17-13-12 e chamado de ângulo diedro T2), de 0 a 360, em intervalos de 15 graus e a energia foi calculada para cada posição da molécula. A seguir, a parte 1 da molécula foi girada de 15 graus na ligação T1, através do ângulo diedro T1, determinado pelos átomos 14-13-12-6 e fixada novamente, enquanto a segunda parte da molécula foi girada, em intervalos de 15 graus, de 0 a 360 sobre a ligação T2, como anteriormente. O processo foi repetido até que o ângulo diedro T1 atingisse 360 graus e assim, 576 posições foram obtidas e portanto, 576 cálculos foram feitos para esta molécula.

Foram encontrados 8 mínimos locais para a molécula I, os quais são mostrados na figura 10.

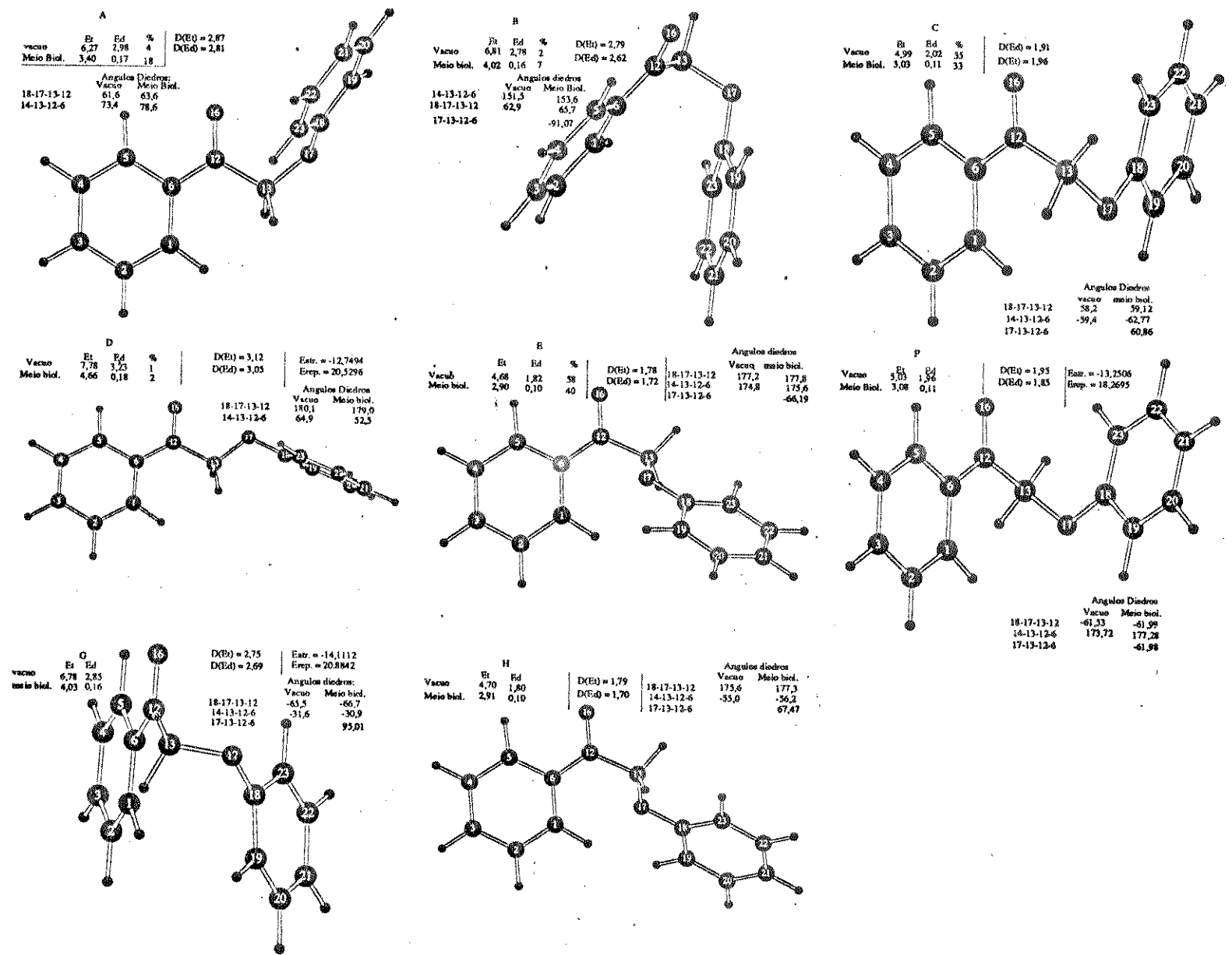


Fig.10 Mínimos locais encontrados para o composto I. Et e Ed significam energia total e energia de interação de dipolos, respectivamente. $D(Et) = Et(\text{vácuo}) - Et(\text{meio biológico})$ e $D(Ed) = Ed(\text{vácuo}) - Ed(\text{meio biológico})$. A porcentagem (%) foi calculada usando a distribuição de Boltzman. Eatr. e Erep. são as energias de atração e de repulsão, respectivamente.

A nomenclatura dada às conformações da molécula I (fig.10), A, B, C, D, E, F, G e H, refere-se aos ângulos diedros T1 e T2 e foi mantida para as outras moléculas, de forma que a conformação A da molécula I, por exemplo, possui os ângulos diedros T1 e T2 próximos aos mesmos ângulos da conformação A das outras moléculas. Os ângulos diedros da conformação A das outras moléculas podem ser vistos no apêndice 4. Para todos os casos, os ângulos T1 e T2 da conformação A estão próximos de 60 graus.

Observando também os mapas conformacionais das moléculas ativas (figuras 11.a a 11.h), veremos que para todas as moléculas, as conformações de mesmo nome estão na mesma região dos mapas. Compare, por exemplo, os mapas das moléculas XII e XIII.

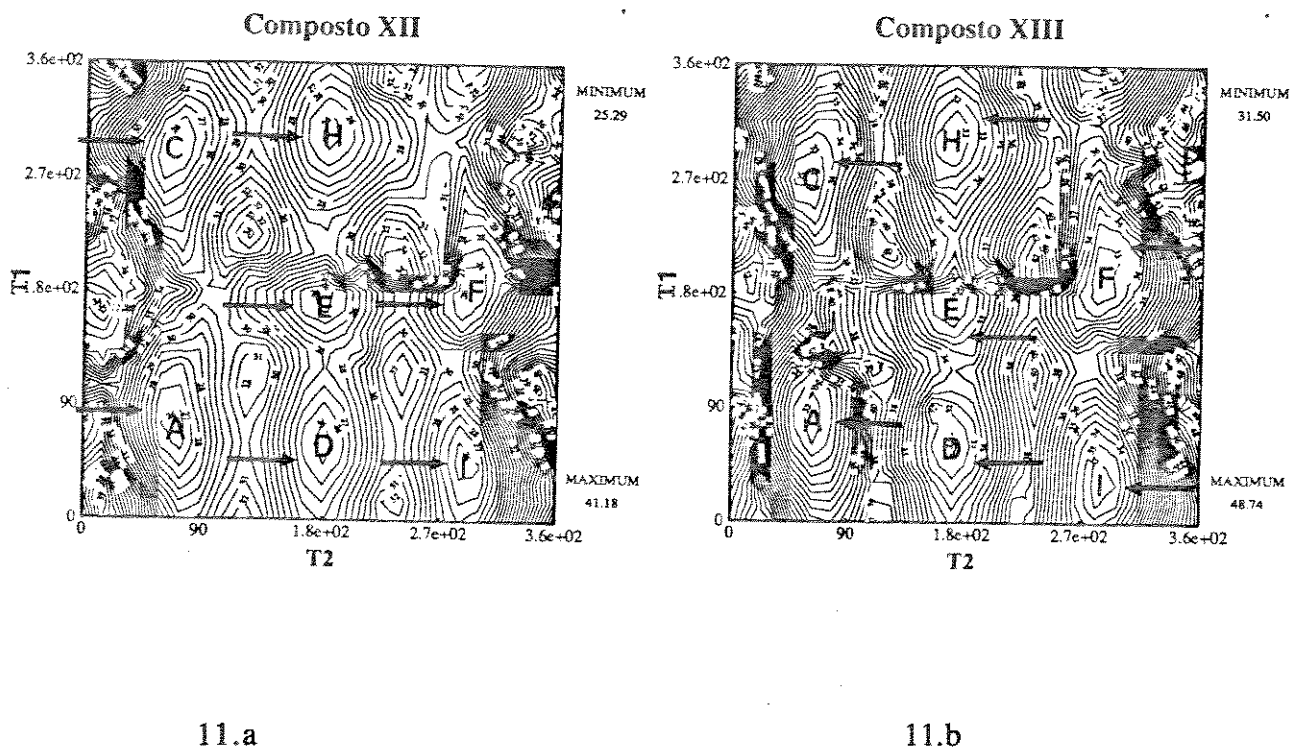
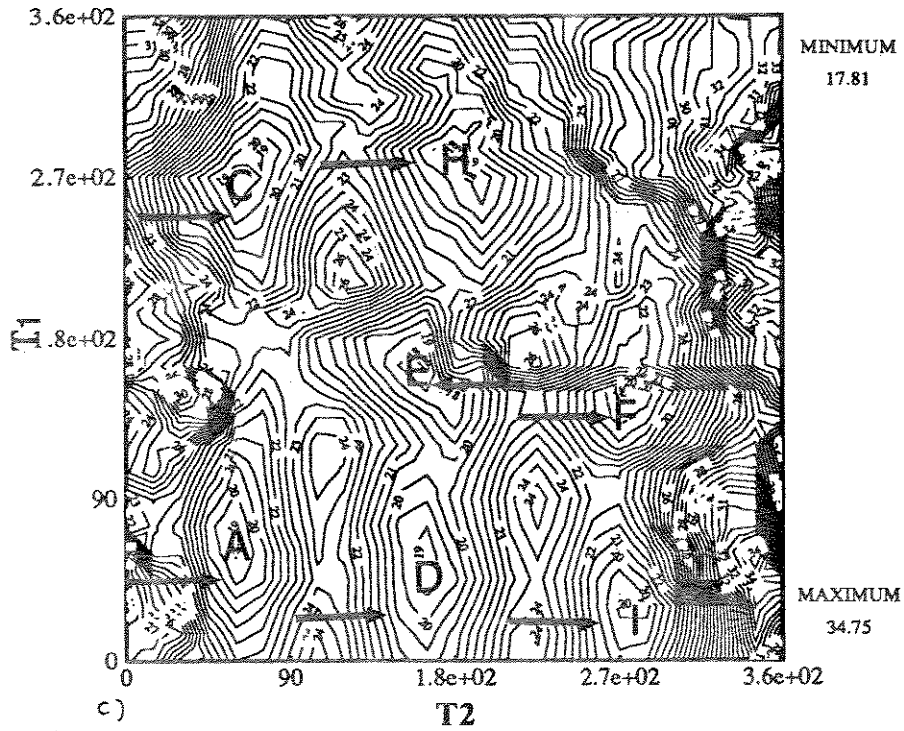


Fig.11 Mapas conformacionais para os compostos ativos; as setas indicam os limites dos vales, correspondendo à região mínima (E) e E somado de 2RT.

Composto XIV



Composto XV

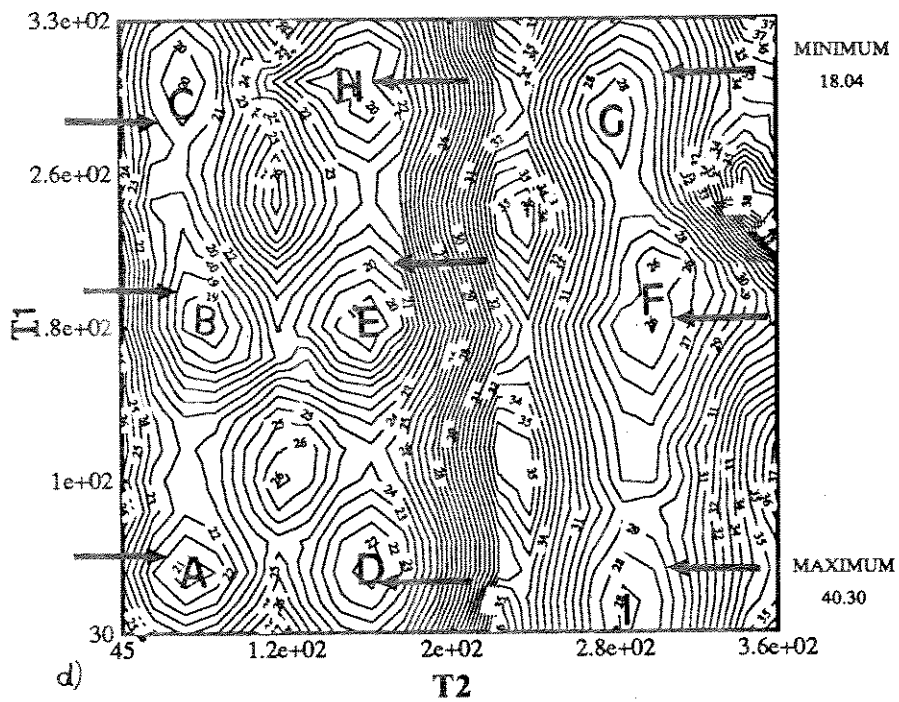
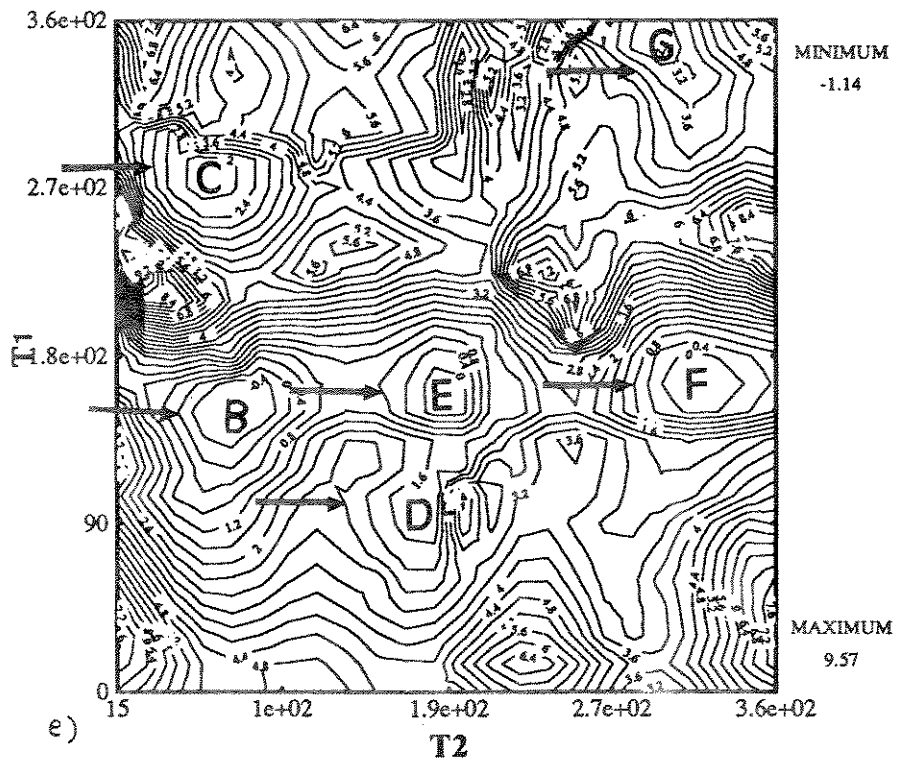


Fig.11 continuação;

Composto XVII



Composto XVIII

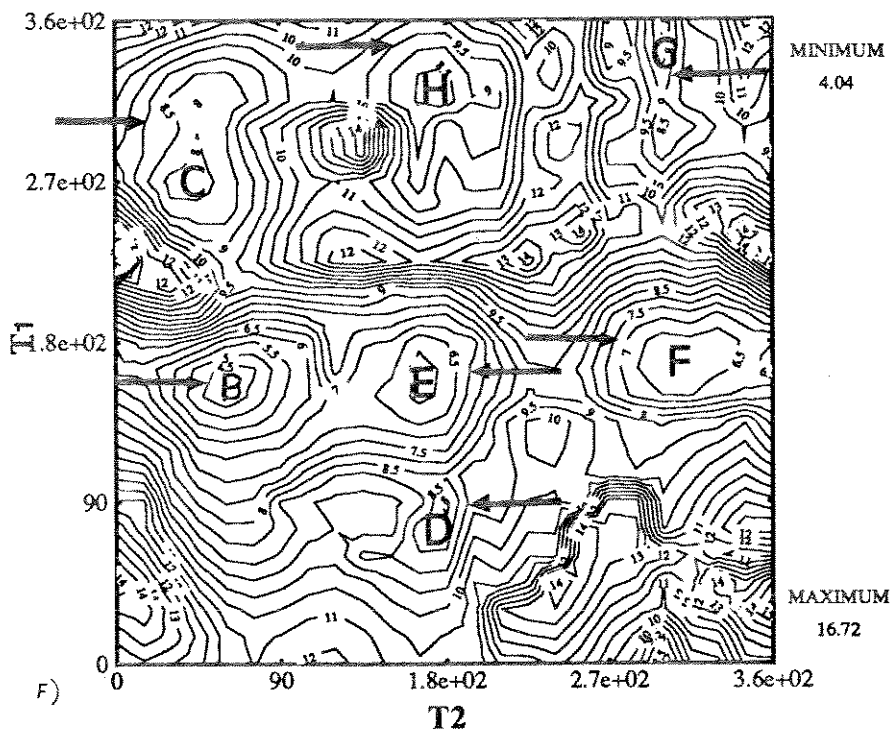
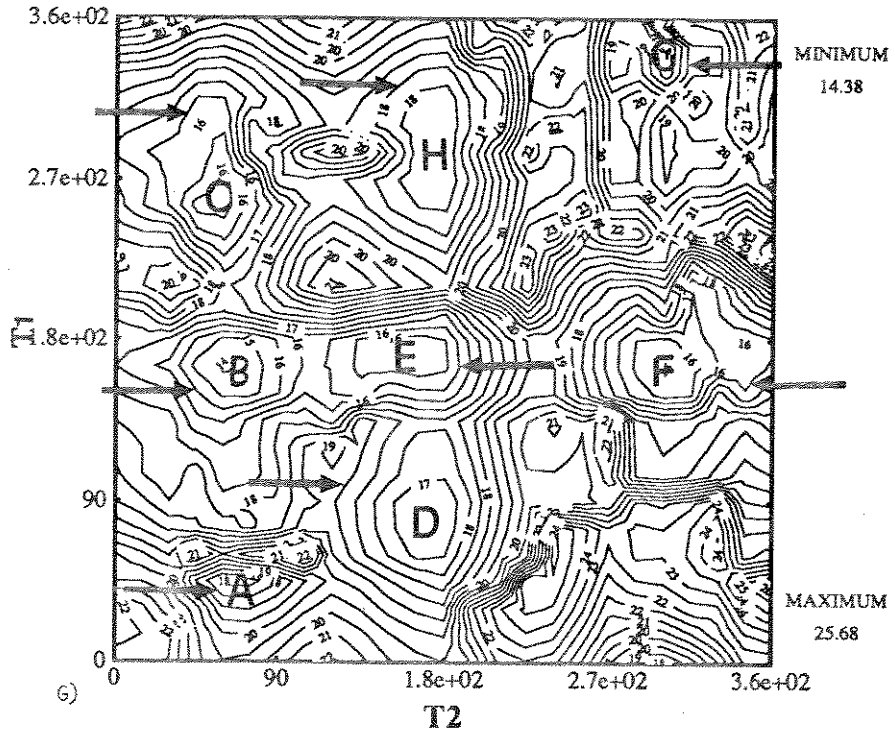


Fig.11 continuação;

Composto XIX



Composto XX

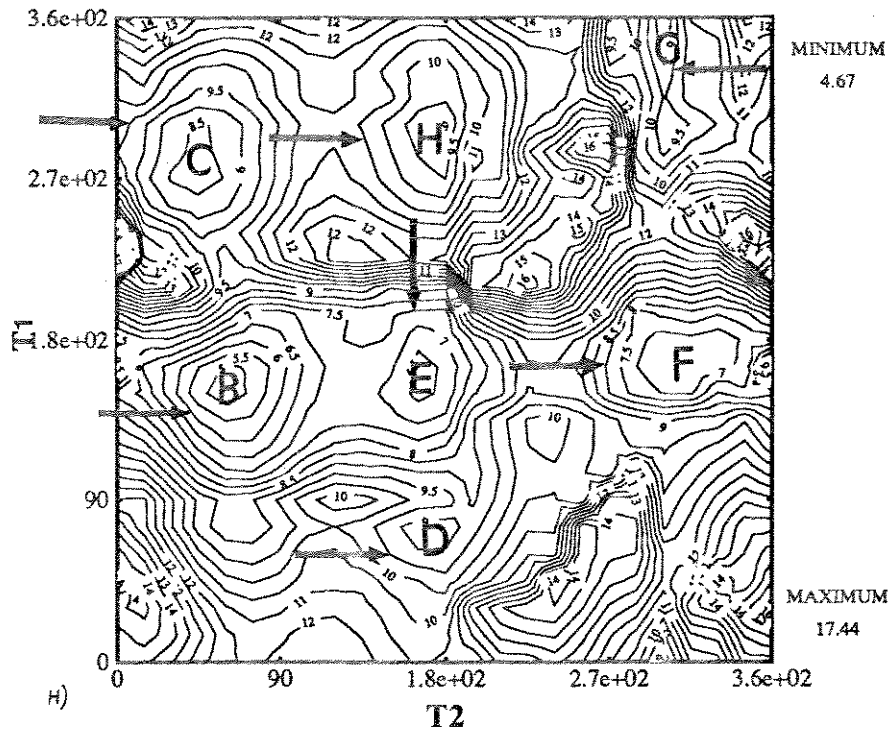


Fig.11 continuação;

Os cálculos dos mínimos locais para as outras 17 moléculas com a mesma estrutura básica da molécula I, foram feitos da seguinte maneira: Para cada um dos oito mínimos

locais encontrados (no caso da molécula I), os grupos substituintes foram sendo colocados um a um nos anéis e para cada um foi feita uma rotação de 0 a 360 graus, com intervalos de 10 graus e a energia para cada ponto foi calculada. Por exemplo, para a molécula IV (figura 12), o radical metóxi foi introduzido em substituição ao hidrogênio 28 na molécula I, para cada um dos oito mínimos locais de I. O radical foi girado em torno da ligação "x", em intervalos de 10 graus para o mínimo local A e a energia para cada ponto foi anotada. Verificou-se a posição do $-OCH_3$ que resultou na menor energia e esta foi considerada a posição com que o radical deve ser colocado no mínimo local A da molécula I para gerar o mínimo local A da molécula IV. O mesmo foi feito para os outros sete mínimos locais de I e assim obtivemos os oito mínimos locais para a molécula IV.

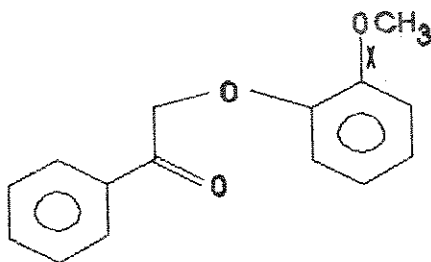


Fig.12 O radical metóxi foi girado sobre a ligação x, a cada 10 graus, em cada um dos oito mínimos locais encontrados para a molécula I.

Apenas para o caso das duas moléculas (XV e XVI da tabela 2) que não possuem a mesma estrutura básica da molécula I, foi necessário investigar novamente os mínimos locais, variando-se as ligações T1 e T2 como para a molécula I. Para estas últimas, nove mínimos locais foram encontrados.

Os cálculos para otimização de geometria foram feitos no vácuo e em meio biológico,

através da alteração da constante dielétrica do meio, no arquivo de dados de entrada do programa MM2⁹. A constante dielétrica do meio biológico foi determinada por Schellman, usando métodos de eletrostática clássica ao problema explicado a seguir.

A energia eletrostática clássica de interação entre duas cargas "q" separadas pela distância "r" em um meio de constante dielétrica "D", é dada por:

$$E = q^2 / D.r \quad \text{eq.5}$$

Assim, para uma dada distância de separação, e com cargas constantes, a magnitude da interação é dependente da constante dielétrica. Quando consideramos interações eletrostáticas entre moléculas pequenas, a distorção da constante dielétrica do meio entre as cargas, pela interferência da estrutura molecular, é pequena e como primeira aproximação, o valor da constante dielétrica da fase principal (80 para água) pode ser usada. Uma proteína globular (uma esfera fortemente empacotada) pode ser tratada como uma região de baixa constante dielétrica, envolvida por um meio de alta constante dielétrica, e para cargas interagindo dentro da proteína, um valor de 2-5 para constante dielétrica tem usualmente sido assumido. Considerando a associação entre um íon em solução e um grupo carregado da proteína, a presença desta região de baixa constante dielétrica tem um efeito marcante na constante dielétrica do meio entre os íons. Schellman, então, aplicou os métodos de eletrostática clássica a este problema de interações iônicas próximas à interface entre regiões de alta e baixa constante dielétrica, encontrando o valor de 28 para a constante dielétrica do meio¹⁰.

Após a obtenção dos mínimos locais para todas as moléculas, foi possível obter os mapas conformacionais para as moléculas ativas; para isso, para cada molécula, partimos da geometria de um mínimo local qualquer e giramos as ligações T1 e T2 em intervalos de 15 graus, como foi feito para a molécula I, de 0 a 360 graus. Os pontos foram anotados e assim o mapa pôde ser obtido. Os mapas são mostrados nas figuras 11.a a 11.h. Nesses mapas, vemos regiões marcadas com letras de A a I. Essas regiões são justamente os vales (ou poços de potencial) que representam os mínimos locais. As curvas em torno desses vales possuem valores de energia crescente, significando que o ponto central de cada vale possui o menor valor de energia (mínimo local). Há também, regiões de máxima energia entre os vales, dando uma idéia da barreira de rotação da molécula, para passar de uma conformação para outra. As setas indicando determinado contorno para cada vale, mostram o limite dos mínimos locais, com o valor de energia igual à energia mínima do poço, somada de $2RT^*$.

* Usamos R ao invés de k, pois estamos trabalhando com mol de moléculas; para macro-sistemas usa-se a constante R.

A probabilidade de cada conformação para as moléculas ativas, foi determinada a partir de informações obtidas com os mapas conformacionais, da seguinte maneira: Para a molécula XII, quatro mínimos locais diferentes foram encontrados, conforme explicaremos no item II.3.2. Esses mínimos possuem os seguintes valores de energia, em meio biológico (em kcal): A = 26,16, C = 25,90, D = 25,74 e E = 25,27. A ordem de estabilidade entre eles é E>D>C>A, segundo os dados acima de energia; entretanto, não podemos dizer que a conformação E, mais estável, é também a mais provável sem antes analisarmos o mapa conformacional da molécula, na figura 11.a. Nessa figura, vemos os quatro poços de energia correspondentes aos mínimos locais diferentes. Quando consideramos que cada mínimo local é formado por todas as conformações com valores de energia entre a energia mínima do poço, E, e E + 2RT, conforme explicamos no item II.1, encontramos conformações com os seguintes valores de energia (em kcal) para os mínimos locais A, C, D e E: A = 26,16, 26,29, 26,31, 26,50, 26,60, 26,64, 26,77, 26,82, 26,99, 27,11 e 27,32. C = 25,95, 26,07, 26,11, 26,30, 26,34, 26,40, 26,74, 27,00, 27,03, e 27,03. D = 25,74, 25,93, 25,94, 26,36, 26,36, 26,42, 26,56, 26,60, 26,83, 26,96, 26,96 e 26,98. E = 25,27, 25,42, 25,66, 25,72, 26,23, e 26,28.

Devemos associar cada conformação com a probabilidade da equação 2 (item II.1) e então, obter a função de partição Z para cada mínimo local através do somatório da equação 1 (item II.1).

$$Z_A = 1,64 \cdot 10^{-18}, Z_C = 2,99 \cdot 10^{-18}, Z_D = 3,30 \cdot 10^{-18}, Z_E = 4,85 \cdot 10^{-18}$$

Agora, fazemos um outro somatório sobre os quatro mínimos locais encontrados e obtemos a função de partição molecular que indica como as moléculas estão distribuídas pela superfície de energia:

$$Z = 1,28 \cdot 10^{-17}$$

Para obtermos a probabilidade relativa entre os mínimos locais, basta dividirmos Z de cada conformação, pelo Z total:

$$P_A = Z_A / Z = 1,64 \cdot 10^{-18} / 1,28 \cdot 10^{-17} = 13\%$$

onde P é a probabilidade. Da mesma forma,

$$P_C = 23\%, P_D = 26\% \text{ e } P_E = 38\%.$$

Vemos que neste caso, a ordem de probabilidade, $E>D>C>A$, coincide com a ordem de estabilidade. Os mesmos cálculos foram feitos para as outras moléculas ativas e os resultados das probabilidades são mostrados na tabela 5.

II.3 RESULTADOS E DISCUSSAO

Na análise conformacional dos compostos envolvidos neste trabalho, procuramos investigar os seguintes pontos que achamos de fundamental importância para o prosseguimento do mesmo:

- 1- Verificar qual é o melhor método para a otimização de geometria.
- 2- Definir a importância de se efetuar os cálculos em meio biológico.
- 3- Analisar os mapas conformacionais das moléculas ativas, a fim de se verificar a possível existência de uma conformação favorável à atividade.

O item 1 que se refere à escolha do método para otimização de geometria, já foi explicado no tópico II.1; os outros serão discutidos aqui.

II.3.1 - COMPARAÇÃO ENTRE OS CALCULOS REALIZADOS EM MEIO BIOLÓGICO E NO VACUO.

Inicialmente, foram encontrados oito mínimos locais para a molécula I, quando as ligações T1 e T2 foram giradas de 0 a 360 graus, conforme descrito no item II.2. Observando a figura 10, onde são mostradas as oito conformações da molécula I, vemos que os

pares de conformações B e G, C e F, E e H possuem valores de energia muito próximos e são imagens especulares não superponíveis entre si (enantiômeros); observe a energia estérica total em kcal (Et) para esses pares, em meio biológico: B (Et = 4,02), G (Et = 4,03), C (Et = 3,03), F (Et = 3,08), E (Et = 2,90), H (Et = 2,91). Observando também as figuras, verificamos a semelhança existente entre esses pares que pode ser comprovada pelos valores dos ângulos diedros T2 e T3 (designado pelos átomos 17-13-12-6), com valores de 65,7 graus e -91,07 graus, respectivamente, para a conformação B, em meio biológico. Para G ("enantiômero" de B) esses valores são -66,7 graus e 95,01 graus, também em meio biológico. Escolhemos esses dois ângulos para analisar, porque eles representam as rotações realizadas sobre os eixos T1 e T2 mostrados na figura 9. Como essa molécula não possui nenhum grupo substituinte, essas conformações podem ser consideradas iguais do ponto de vista físico-químico, isto é, as propriedades físico-químicas calculadas para esses enantiômeros serão idênticas, de forma que a molécula I apresenta apenas cinco mínimos locais diferentes: A, B, C, D e E. A mesma conclusão é obtida se observarmos os valores no vácuo. As figuras das conformações não diferem para a molécula no vácuo ou em meio biológico, pois os valores dos ângulos diedros estudados (T1 e T2) variou muito pouco, como veremos mais à frente.

O apêndice 4 mostra as conformações encontradas para todas as outras moléculas; podemos verificar que para as moléculas II e III, onde o substituinte está em "para" e para a molécula XI que não possui substituintes, apenas as cinco conformações, A, B, C, D e E foram encontradas. Mas para a molécula IV, com o grupo substituinte -OCH₃ em "orto", todas as oito conformações encontradas para I mostraram-se diferentes em energia. As moléculas V, VI e X que possuem o substituinte -CH₃ no carbono de número 13 também apresentaram oito mínimos locais diferentes, pois qualquer substituinte nesse carbono impediria a formação daqueles enantiômeros. Isso pode ser facilmente verificado na figura 10, observando-se que a posição dos hidrogênios ligados a esse carbono está invertida naqueles pares de conformações. Observe, por exemplo, as conformações C e F na figura 10. A molécula IX, com o radical -CH₃ em meta possui apenas 5 conformações diferentes (A, C, D, E e H). Pode-se verificar que os ângulos diedros T2 e T3 são muito próximos para os seguintes pares: A e I, C e F, E e H, bem como as energias desses pares, mas E não é enantiômero de H, pois em E o radical -CH₃ está projetado para frente e os hidrogênios

ligados ao carbono número 13, para trás; mas em H, ambos estão para frente.

No caso da molécula VII, também com o substituinte $-\text{CH}_3$ no carbono 13, apenas seis mínimos locais foram encontrados apesar de não haver nenhum par de conformações. Para obter os mínimos locais de VII, partiu-se dos oito mínimos locais da molécula VI e acrescentou-se o novo grupo substituinte, $-\text{HC}=\text{CH}-\text{CH}_3$ no carbono 21. Verifique as duas moléculas na tabela 1. A otimização foi feita conforme está explicado no item II.2. Entretanto, ocorreram alguns casos em que o novo grupo substituinte foi colocado numa determinada conformação da molécula VI e a otimização resultou em outra conformação para a molécula VII. Por exemplo, para obtermos o mínimo local "B" para a molécula VII, acrescentamos o radical $-\text{HC}=\text{CH}-\text{CH}_3$ no carbono N^o 21 da conformação B da molécula VI e fizemos os cálculos com MM2. Ao verificarmos os resultados, vimos que os ângulos diedros T1 e T2 correspondiam àqueles da conformação C, (com os ângulos diedros T1 \cong 200 graus e T2 \cong 50 graus). Dessa forma, encontramos apenas seis mínimos locais: A, C, D, E, F e H, pois a conformação G de VI também resultou em outra para VII.

O mesmo ocorreu com a molécula VIII, contendo o radical $-\text{NO}_2$ no carbono 19; apesar de não aparecer nenhum par de conformação, apenas cinco mínimos locais diferentes foram encontrados: A, C, E, F e H. A tabela 3 mostra os mínimos locais encontrados para as moléculas inativas (I a XI) e a ordem de estabilidade das conformações.

Os cálculos de otimização de geometria foram feitos no vácuo e em meio biológico, para verificarmos possíveis modificações nas conformações das moléculas, mas observamos que tal alteração não ocorreu, isto é, os ângulos diedros T1 e T2 permaneceram muito próximos. Apenas no caso da molécula X, ocorreu alteração nos ângulos diedros da conformação A, pois quando esta foi calculada no vácuo, obteve-se como resultado a conformação B. A energia estérica, porém, apresentou valores bem menores em meio biológico para a maioria das moléculas. No caso da conformação A da molécula I (figura 10), por exemplo, a energia estérica total (Et) no vácuo é de 6,27 kcal e em meio biológico é de 3,40 kcal, mas os ângulos diedros T1 e T2 quase não sofreram alteração, pois no vácuo, T1 = 73,4 graus e T2 = 61,6 graus e em meio biológico, T1 = 78,6 graus e T2 = 63,6 graus. Todas as outras conformações apresentam essa semelhança quando calculadas no vácuo ou em meio biológico.

A redução no valor da energia estérica total em meio biológico deve-se ao fato de que

neste meio, a constante dielétrica é maior (28) do que no vácuo (1), diminuindo assim, as interações eletrostáticas, segundo a equação 5.

Esta conclusão pode ser comprovada, verificando-se na figura 10 que o decréscimo na energia estérica total $D(Et)$ é bem próximo do decréscimo na energia de interação de dipolos, $D(Ed)$, único termo no campo de força usado em MM2 que envolve constante dielétrica (apêndice 1). Observe por exemplo, a conformação A da molécula I novamente; para esta conformação, $D(Et) = Et(\text{vácuo}) - Et(\text{meio biológico}) = 6,27 - 3,40 = 2,87$ e $D(Ed) = Ed(\text{vácuo}) - Ed(\text{meio biológico}) = 2,98 - 0,17 = 2,81$.

Observou-se também que o decréscimo de energia $D(Et)$ é maior para as conformações menos estáveis (A, B, D), onde o valor das interações de dipolo (Ed) é maior. Comparar, por exemplo, as conformações A (pouco estável) e C (mais estável); para A, $D(Et) = 2,87$ e $Ed(\text{vácuo}) = 2,98$. Para C, $D(Et) = 1,96$ e $Ed(\text{vácuo}) = 2,02$.

Assim, como as conformações de maior energia sofrem um decréscimo também maior de energia, em meio biológico, a diferença de energia entre uma conformação e outra é menor do que no vácuo, de forma que as porcentagens relativas entre as conformações, calculadas através da distribuição de Boltzmann sobre os mínimos locais, são mais uniformes em meio biológico do que no vácuo.

Observe na figura 10, as porcentagens relativas entre as cinco conformações diferentes (A, B, C, D e E) do composto I.

Podemos verificar que aquelas conformações com porcentagens muito baixas no vácuo, tendem a aumentar suas porcentagens em meio biológico e aquelas com porcentagens muito altas no vácuo tendem a diminuir-las em meio biológico. Por exemplo, a conformação A da molécula I (fig.10) possui um aumento de 4% para 18% em meio biológico e a conformação E possui uma queda de 58% para 40%. Esse fato pode ser observado também para outras moléculas (apêndice 4), através da comparação das energias no vácuo em meio biológico.

Apenas o caso da molécula XV é diferente. Para esta, a energia estérica total mostra-se maior em meio biológico do que no vácuo, pois o somatório das energias de interação de dipolos possui valor negativo (energia de atração). Assim, a diminuição nesse valor resultará num valor maior da energia estérica total, da qual será subtraído um valor menor, enquanto que para as outras moléculas, onde a energia de interação de dipolos é de repulsão (+), a diminuição resulta na soma de um valor menor, diminuindo a energia estérica.

Como todas as conformações ficam com as energias mais próximas entre si, em meio biológico, ocorre às vezes, inversão na ordem de estabilidade das conformações.

Para a molécula V, por exemplo (tab.3), a ordem decrescente de estabilidade no vácuo é: E>C>B>H>F=A>D>G e em meio biológico é: B>E>C>A>F>H>D>G. Verificar na tabela 3 esses resultados para as outras moléculas inativas.

Este fato mostra a importância de se realizar os cálculos de otimização de geometria em meio biológico, pois, apesar de não ocorrer alteração nas conformações, ocorre grande alteração nas suas energias, influenciando a conclusão sobre o mínimo global.

Moléculas	Ordem decrescente de estabilidade	
	Vacuo	Meio biologico
I	E>C>A>B>D	E>C>A>B>D
II	E>C>A>B>D	E>C>A>B>D
III	E>C>D>A>B	E>C>D>A>B
IV	F>H>C>B>E>G>A>D	F>H>B=C=E>G=A>D
V	E>C>B>H>F=A>D>G	B>E>C>A>F>H>D>G
VI	F=C>B>E>H>A>D>G	B>C>F>E>A>D>H>G
VII	C>A>E>D>H>F	C>E>A>D>H>F
VIII	H>F>C>A=E	H>F>C>E=A
IX	E=H>D>C>A	E=H>D>C>A
X	B>E>F>A>C>D>H>G	B>E>F>A>D>C>G>H
XI	B>C>E>A>D	B>C>E>A>D

Tabela 3. Ordem decrescente de estabilidade para as conformações das moléculas inativas, com base nos dados de energia do apêndice 4, no vácuo e meio biológico.

Ainda na figura 10, vemos que as conformações mais estáveis não são aquelas em que os anéis estão mais afastados, A e D, como poderíamos pensar à primeira vista, mas aquelas em que a distância é intermediária: C (= F) e E (= H). Tomando como exemplo as conformações D, F e G (= B), vemos que em D, onde os anéis estão mais afastados, a energia de atração ($E_{atr.}$) é menor (em módulo) do que em F e G, enquanto a repulsão ($E_{rep.}$) é alta. À medida que os anéis se aproximam (F), a energia de atração aumenta, enquanto as repulsões diminuem, estabilizando a molécula. Entretanto, quando os anéis se aproximam muito (G), todas as interações são maiores, mas as de repulsão crescem mais do que as interações de atração, desestabilizando o composto, em relação à F, obtendo-se a seguinte ordem decrescente em estabilidade: $F > G > D$. A tabela 4 também mostra esses resultados.

Conformação	$E_{rep.}$ (kcal)	$E_{atr.}$ (kcal)
D	20,53	12,75
F	18,27	13,25
G	20,88	14,11

Tabela 4. A conformação D, com os anéis mais distantes é menos estável do que F e G, onde a distância entre eles é menor, devido à variações nas energias de atração ($E_{atr.}$) e repulsão ($E_{rep.}$).

Se observarmos a ordem de estabilidade dos mínimos locais na tabela 3, veremos que esse raciocínio se aplica às outras moléculas, pois o mínimo global não foi D em nenhuma situação, mas na maioria das moléculas inativas, as conformações mais estáveis em meio biológico foram aquelas em que os anéis apresentavam distâncias intermediárias: E, H, F e C, e em alguns casos, aquela em que os anéis estão mais próximos: B.

Para as moléculas ativas, os cálculos também foram feitos em meio biológico e no vácuo para confirmar estas conclusões. Entretanto, os mapas conformacionais foram feitos apenas em meio biológico, visto que nessa etapa já havíamos concluído a necessidade de se prosseguir os cálculos em meio biológico.

A análise para esse conjunto de moléculas é feita no tópico II.3.2.

II.3.2 - ANALISE DAS MOLECULAS ATIVAS

Os mapas conformacionais para as moléculas ativas foram obtidos conforme está explicado no item II.2, a fim de investigarmos a conformação favorável à atividade e são mostrados nas figuras 11.a a 11.h.

O mapa da molécula XII mostra sete mínimos locais, mas se observarmos a figura correspondente no apêndice 4, veremos que os enantiômeros verificados para algumas moléculas inativas, também ocorrem aqui: A e I, C e F, E e H. Os ângulos diedros T2 e T3 para esses enantiômeros são bem próximos, assim como as energias. Assim, a probabilidade será calculada considerando-se apenas as quatro conformações diferentes: A, C, D e E.

A molécula XIII também apresenta os enantiômeros A e I, E e H, C e F, de forma que novamente temos apenas quatro conformações diferentes: A, C, D e E.

Já no caso da molécula XIV, C e F não formam pares, porque os ângulos T2 e T3 para C não são muito próximos aos mesmos ângulos da conformação F, de forma que os anéis aromáticos em C estão mais próximos do que em F.

As conformações E e H da mesma molécula também não são enantiômeros, pois apesar de possuírem T2 e T3 próximos, o carbono de número 42 em E está em "cis" com o carbono número 22, mas em H, os mesmos carbonos estão em "trans". Entretanto, coincidentemente, são equivalentes em energia. Assim, a molécula XIV apresenta seis conformações diferentes: A, C, D, E, F e H.

As outras moléculas ativas possuem o radical $-CH_3$ no carbono 13 e portanto, como já foi explicado para as moléculas inativas, não apresentam enantiômeros e todos os mínimos locais que aparecem nos respectivos mapas, são diferentes. A questão a ser resolvida agora é: "Qual deles corresponde à conformação com a qual a molécula vai se ligar ao receptor?" Não sabemos nada sobre a conformação do receptor. Se soubéssemos, o problema estaria resolvido, isto é, poderíamos pensar que a droga ligar-se-ia com a conformação que melhor se ajustasse ao receptor.

Contudo, sabemos que existe uma probabilidade para cada mínimo local e que quando a molécula for se ligar ao receptor, ela poderá adotar qualquer uma dessas

conformações de mínima energia e ainda, escolhido o mínimo local supostamente mais próximo da conformação do receptor, a molécula poderá alterar um pouco a sua conformação de forma a adotar uma conformação pertencente ao mínimo local escolhido, mas que não seja aquela com menor energia, isto é, a molécula talvez não esteja no fundo de um poço de potencial, mas em qualquer conformação dentro do limite considerado e indicado pela seta.

Assim, vamos inicialmente considerar que o receptor é uma molécula rígida, que não altera a sua conformação ao ligar-se à droga e que todos os compostos em estudo adotarão uma mesma conformação (de alta probabilidade) para ajustar-se a ele. Devemos então, verificar quais são os mínimos locais comuns a todas as moléculas ativas e escolher entre esses, aquele de alta probabilidade em todos os compostos como sendo a conformação ativa. A probabilidade de cada conformação para todas as moléculas ativas é mostrada na tabela 5 e foi obtida conforme está explicado no item II.2, com exceção da molécula XVI, para a qual não foi possível obter o mapa conformacional, devido ao número extenso de átomos necessário para fazer a rotação sobre o eixo T2. O programa MM2 faz rotação de uma parte da molécula em relação à outra, mas a parte que será girada pode conter no máximo 32 átomos e neste caso, o número era maior.

Molécula	Mínimo local	E (kcal)	Z	P (%)	Estabilidade	Probabilidade
XII	A	26,16	1,64 E -18	13	E>D>C>A	E>D>C>A
	C	25,90	2,99 E -18	23		
	D	25,74	3,30 E -18	26		
	E	25,27	4,8 E -18	38		
XIII	A	32,19	8,55 E -23	16	E>D>A>C	E>D>A>C
	C	32,36	7,00 E -23	13		
	D	31,92	1,17 E -22	21		
	E	31,45	2,78 E -22	50		
XIV	A	18,87	2,46 E -13	7	H>E>F>C>D>A	H>C=E>F>D>A
	C	18,43	6,36 E -13	19		
	D	18,64	3,08 E -13	9		
	E	17,83	6,25 E -13	19		
	F	18,28	4,01 E -13	12		
	H	17,78	11,25 E -13	34		
XV	A	20,55	4,26 E -15	1	B>H>E>C>D>A>F>G>I	B>>E>H>C>A=D>F=G=I
	B	17,91	6,47 E -13	78		
	C	19,80	3,42 E -14	4		
	D	20,28	3,06 E -15	1		
	E	19,60	8,39 E -14	10		
	F	25,11	4,28 E -18	0		
	G	27,02	3,51 E -19	0		
	H	19,41	5,32 E -14	6		
	I	27,40	1,69 E -19	0		

Tabela 5. Cada mínimo local possui conformações com energia variando de E até E + 2RT (onde E é a energia de menor valor do mínimo local) e a função de partição Z para cada mínimo local é mostrada, bem como as probabilidades; com os dados de energia mínima e de probabilidade, obtivemos as ordens de estabilidade e de probabilidade, respectivamente. A tabela continua na página seguinte.

Molécula	Mínimo local	E (kcal)	Z	P (%)	Estabilidade	Probabilidade
XVII	B	-1,14	26,91	53	B>F>E>D>C>G	B>F>E>D>C>G
	C	1,76	0,36	1		
	D	1,32	1,11	2		
	E	-0,26	8,49	17		
	F	-0,40	13,59	27		
	G	2,78	6,44 E -2	0		
XVIII	B	3,29	8,58 E -3	89	B>E>F>D>C>G>H	B>F>E>D=C=G=H
	C	7,81	2,20 E -5	0		
	D	7,76	1,75 E -5	0		
	E	5,93	5,26 E -4	5		
	F	6,06	5,30 E -4	6		
	G	7,81	1,15 E -5	0		
	H	8,73	1,90 E -6	0		
XIX	A	15,38	6,11 E -11	10	B>E>C>A>F> >H>D>G	B>C>E>A>F> >H=D>G
	B	14,38	3,45 E -10	55		
	C	15,24	9,41 E -11	15		
	D	16,97	7,88 E -12	1		
	E	15,20	8,82 E -11	14		
	F	16,15	2,21 E -11	4		
	G	17,41	1,09 E -12	0		
	H	16,75	6,80 E -12	1		
XX	B	4,62	2,09 E -3	82	B>E>F>D>C=G>H	B>E>F>C=D=G=H
	C	8,48	7,15 E -6	0		
	D	8,28	4,57 E -6	0		
	E	6,34	2,53 E -4	10		
	F	6,55	2,11 E -4	8		
	G	8,48	3,16 E -6	0		
	H	9,26	2,04 E -6	0		

Tabela 5 (continuação).

Observando a tabela 5, que mostra a ordem de estabilidade das moléculas ativas (baseando-se nos valores de energia) e a probabilidade calculada baseando-se nos mapas conformacionais, conforme o item II.2, vemos que ocorreu inversão nas probabilidades dos compostos XIV, XV, XVIII, XIX e XX, isto é, na figura 11.c (composto XIV), vemos que o "poço" para a conformação C é nitidamente mais largo do que o poço para as conformações E e F com energia interna mais baixa, resultando na inversão das probabilidades destes mínimos, de forma que C (menos estável que E e F) é tão provável quanto E e mais provável do que F.

Na figura 11.d (composto XV), vemos que o poço E é mais largo do que H, de forma que a conformação E é mais provável do que H, apesar de H ser mais estável do que E. A conformação A também é igualmente provável a D, apesar de D ser mais estável, porque o poço A é mais largo do que o poço D.

Na figura 11.f (composto XVIII), vemos que o poço da conformação F é mais largo do que E, possibilitando a inversão na probabilidade dessas conformações; da mesma forma, C é bem mais largo do que D e apesar de C ser menos estável do que D, os dois são igualmente prováveis. Na figura 11.g (composto XIX), podemos observar que o poço da conformação C é mais largo do que E, invertendo as probabilidades destas conformações. Também observamos que D, com energia mais alta do que H, possui a mesma probabilidade.

O composto XX (fig.11.h) também apresenta inversão nas probabilidades das conformações C e D e de G e H. De fato, podemos ver que a largura do poço da conformação menos estável, C, é maior do que D, possibilitando a inversão de probabilidade. No caso de G e H, H também é bem mais largo do que G e apesar de ser menos estável, possui a mesma probabilidade. Apesar das inversões de probabilidades ocorridas, não houve alteração na probabilidade do mínimo global, de forma que a conformação mais provável é também a mais estável.

As outras moléculas não apresentam inversão nas probabilidades dos mínimos locais.

Observamos (tabela 5) que das nove moléculas ativas, seis (XV, XVI, XVII, XVIII, XIX e XX) apresentam a conformação B, com os anéis fechados, mais provável (ver as figuras no apêndice 4) - com o ângulo T1 entre 150 e 180 graus e T2 entre 50 e 90 graus - enquanto para as outras três (XII, XIII e XIV, as conformações com os anéis abertos H,

($T_1 = -50$ e $T_2 = 180$ graus), no caso da molécula XIV, e E (T_1 e $T_2 \cong 180$ graus) no caso de XII e XIII, são mais prováveis.

Entretanto, se considerarmos que o receptor é uma molécula rígida (a possibilidade da existência de receptores rígidos e flexíveis será discutida no item IV.1), a conformação E, com os ângulos T_1 e T_2 próximos de 180 graus, pode ser a conformação com que as moléculas ativas se ligam, pois esta conformação está presente em todas as moléculas ativas, com alta probabilidade. Observe na tabela 5, a probabilidade de E para todas as moléculas. Entretanto, se o receptor for flexível, poderá alterar a sua conformação ao interagir com o composto ativo; nesse caso, cada um liga-se ao receptor com a sua conformação mais provável.

Proseguiremos o trabalho então, considerando as duas hipóteses.

Se existisse uma conformação provável a todas as ativas e ao mesmo tempo proibida para as inativas, esta seria uma conformação favorável à atividade, mas pelas tabelas 3 e 5, vemos que os mínimos locais que aparecem nas moléculas ativas aparecem também nas inativas. Assumiremos então, as duas primeiras hipóteses já mencionadas, no prosseguimento dos cálculos, onde alguns parâmetros físico-químicos serão calculados.

Quando a primeira hipótese for considerada (receptor rígido), cálculos de parâmetros físico-químicos serão feitos para a conformação E no caso das moléculas ativas e também inativas, pois estamos considerando que essa é a conformação complementar ao receptor e portanto devemos analisar as moléculas, quanto à atividade, nesta conformação, isto é, se a molécula for ativa ela deverá estar nesta conformação e se for inativa, deverá ser inativa apesar de estar nesta conformação, pois além da conformação, outros fatores influenciam a atividade das moléculas e estes serão discutidos nos próximos capítulos.

Quando considerarmos a segunda hipótese, cada molécula (ativa ou inativa) assumirá a sua conformação de maior probabilidade.

II.3 CONCLUSOES

Neste capítulo, três pontos fundamentais ao prosseguimento do trabalho foram mencionados e puderam ser investigados; as seguintes conclusões são mostradas:

1- Para este conjunto de moléculas, o método de mecânica molecular - MM2 mostrou-se mais eficiente e viável na otimização de geometria.

2- Todos os cálculos devem ser efetuados em meio biológico, visto que nesse meio todos os mínimos locais ficam mais próximos em energia do que no vácuo, possibilitando a inversão entre eles, em termos de energia mínima.

3- Concluimos, da análise conformacional para as moléculas ativas, que a conformação E deve ser aquela com a qual a molécula liga-se ao receptor, se este for considerado uma molécula rígida. Concluimos ainda, a necessidade de considerar que o receptor também pode alterar a sua conformação ao interagir com o composto ativo, uma vez que a conformação mais provável não é a mesma para todos os compostos.

CAPITULO III - PARAMETROS FISICO-QUIMICOS

III.1 - INTRODUÇÃO

Um substituinte, numa molécula, produz modificações principalmente nas três seguintes propriedades físicas: eletrônica, estérica e hidrofóbica (ou lipofílica). As relações das duas primeiras propriedades com a atividade biológica foram estudadas, inicialmente, por Hammett¹¹ e Taft¹², respectivamente.

Hidrofobicidade foi o nome dado à propriedade física da molécula, relacionada com a sua partição (P) entre um solvente aquoso e não aquoso, segundo a equação 6:

$$P = \frac{[Droga]_{naq.}}{[Droga]_{aq.}} \quad \text{eq.6}$$

onde P é a constante de equilíbrio da reação $[Droga]_{aq} \rightleftharpoons [Droga]_{naq}$, naq. é o solvente não aquoso e aq. significa que o solvente é a água.

A relação dessa propriedade com a atividade biológica foi estudada inicialmente por Overton (1897)¹³ e Meyer (1899)¹⁴; eles observaram que a atividade narcótica de uma série de compostos é função do seu coeficiente de partição P.

O efeito eletrônico de substituintes no benzeno foi estudado por Hammett (1940 - 1970)¹¹, que relacionou o efeito eletrônico do substituinte σ com a constante de equilíbrio (k) da reação entre o composto não substituído, k_0 , e substituído, k:

$$\log \frac{k}{k_0} = \rho \cdot \sigma \quad \text{eq.7}$$

onde ρ é a constante de proporcionalidade, característica da sensibilidade da reação ao efeito eletrônico.

Por volta de 1950, Taft estendeu a equação de Hammett, incluindo o efeito estérico

dos substituintes em compostos não aromáticos. O efeito estérico do substituinte foi chamado de E_s , e o efeito eletrônico do substituinte de compostos não aromáticos, σ^* .

Na equação de Taft, a constante de equilíbrio está relacionada com σ^* e E_s :

$$\log \frac{k}{k_0} = \rho^* \cdot \sigma^* + \delta \cdot E_s \quad \text{eq.8}$$

Concluído então, o efeito dos substituintes nestas propriedades físicas da molécula, Hansch et al. (1963)¹⁵ postularam o efeito dos substituintes nas propriedades biológicas das moléculas, como resultado da alteração nessas propriedades físicas.

As mesmas propriedades físico-químicas podem ser relacionadas com a atividade biológica, através de outros parâmetros, diferentes desses citados acima. Alguns parâmetros eletrônicos, comumente relacionados com a atividade biológica e calculados com mecânica quântica são as energias do HOMO (orbital molecular mais alto ocupado) e do LUMO (orbital molecular mais baixo desocupado).

A tentativa de buscar a explicação das interações dos fármacos com seus receptores, ao nível eletrônico, é válida, uma vez que as propriedades físico-químicas apresentam origem eletrônica. Algumas revisões¹⁶⁻²⁰ mostram o emprego de modelos de química quântica na correlação da estrutura química com a atividade biológica; esses modelos utilizam resultados de cálculos de orbital molecular, através da resolução da equação de Schroedinger.

Existem alguns métodos em mecânica quântica, que visam a resolução das equações envolvidas com certas aproximações nas integrais; são os métodos semi-empíricos. Estas aproximações consistem em ignorar integrais cujo valor previsto é pequeno, utilizando dados experimentais para aproximar outras, cujo significado físico permite relacioná-los diretamente com observáveis. Esses métodos foram desenvolvidos, visando diminuir o tempo computacional de cálculos para moléculas com muitos átomos e o método semi-empírico utilizado para calcular alguns dos parâmetros físico-químicos deste trabalho, está explicado no apêndice 2.

III.1.1 ENERGIAS DO HOMO E LUMO

Quando duas moléculas se aproximam, numa reação, seus orbitais interagem formando novos orbitais moleculares ligantes e anti-ligantes. A formação dos orbitais ligantes, normalmente é exotérmica, E_1 , enquanto a formação dos orbitais antiligantes é endotérmica, E_2 , conforme a figura 13.

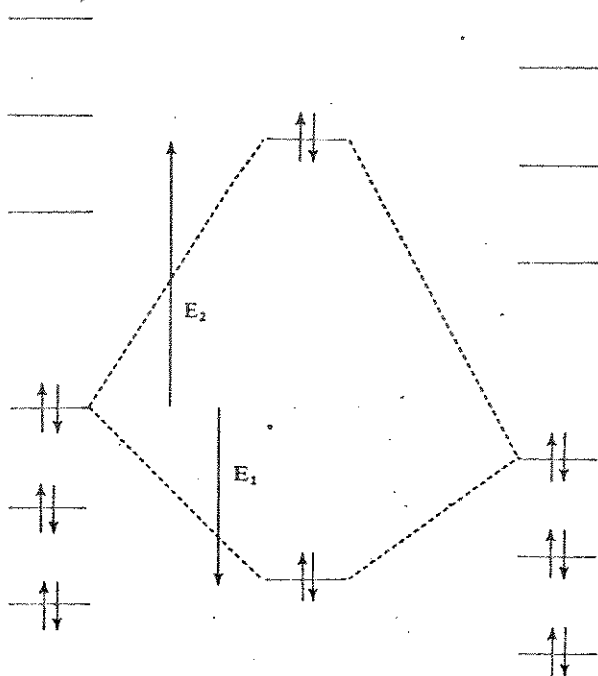


Fig.13 Interação do HOMO de uma molécula com o HOMO de outra.

As interações que produzem um importante efeito de abaixamento de energia são as combinações de orbitais ocupados com desocupados, pois, neste caso não haverá elétrons para ocupar o orbital antiligante, de energia mais alta (fig.14).²¹

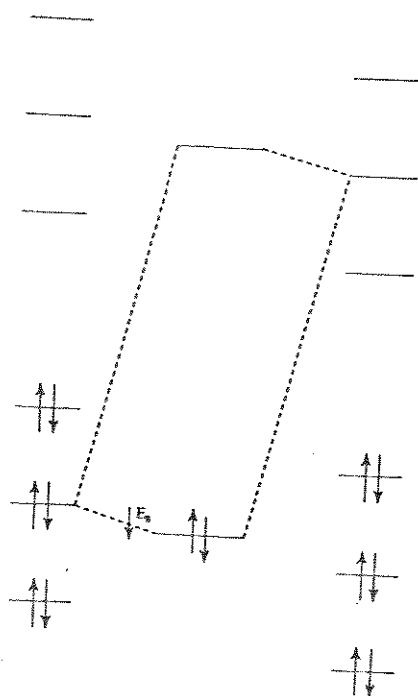


Fig.14 Interação de um orbital ocupado de uma molécula com um desocupado de outra.

Entretanto, é a interação do HOMO (orbital molecular mais alto ocupado) de uma molécula com o LUMO (orbital molecular mais baixo desocupado) de outra que leva ao maior abaixamento de energia, devido à proximidade entre os dois orbitais, pois, o abaixamento de energia é maior para interação entre orbitais com energias similares do que para interação entre orbitais com energias muito diferentes (figura 15).²¹

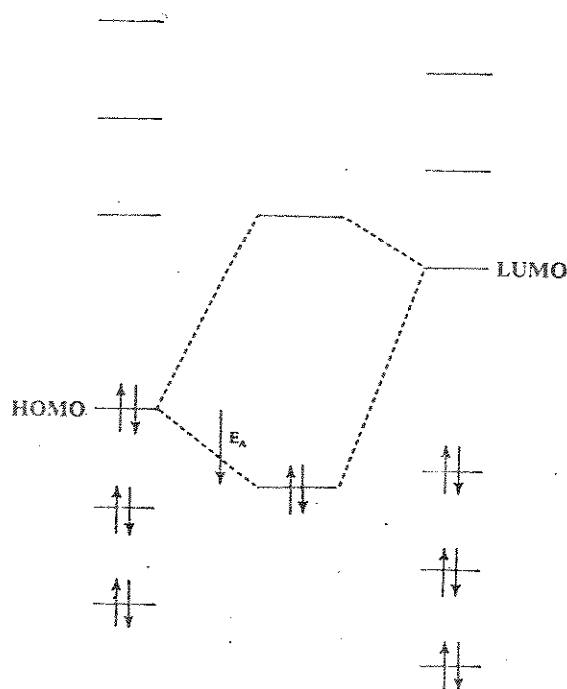


Fig.15 Interação do HOMO de uma molécula com o LUMO de outra.

Isso mostra por que esses orbitais, os orbitais de fronteira como Fukui²² os chamou, são tão importantes.

Klopman²³ e Salem²⁴ propuseram uma equação para o ganho e perda de energia, quando os orbitais de duas moléculas interagem numa reação e mostraram que as energias do HOMO e do LUMO são importantes, pois quando duas moléculas colidem, três forças operam principalmente:

- 1 - Os orbitais ocupados se repelem;
- 2 - Qualquer carga positiva atrai uma negativa e cargas iguais se repelem;
- 3 - Os orbitais ocupados (HOMO, principalmente) interagem com os não ocupados (LUMO), causando atração entre as moléculas.

As energias do HOMO e do LUMO correspondem às energias necessárias para remover um elétron da molécula e para adicionar, respectivamente, isto é, a energia do HOMO é proporcional ao potencial de ionização e a energia do LUMO, à afinidade eletrônica da molécula. Por isso, costuma-se relacionar essas energias com os potenciais de oxidação e

redução.

Em QSAR (Relações Quantitativas entre Estrutura Química e Atividade Biológica), costuma-se usar dados sobre as energias do HOMO e do LUMO para estimar essas propriedades de oxidação-redução; um exemplo deste tipo de aplicação é encontrado no trabalho de Martin et al.²⁵ Vários outros trabalhos foram realizados com a finalidade de correlacionar as energias do HOMO e LUMO com a atividade biológica²⁶⁻²⁹.

III.1.2 DUREZA E MOLEZA MOLECULAR

O conceito de ácidos e bases, duros e moles, foi introduzido por Pearson³⁰; ele classificou ácidos e bases de Lewis num desses dois grupos e verificou que ácidos duros reagem mais rápido e fortemente com bases duras e ácidos moles reagem mais rápido e fortemente com bases moles. Inicialmente, os ácidos e bases foram classificados em duros e moles, de acordo com o tamanho e a carga, de forma que as bases altamente polarizáveis, facilmente oxidáveis e de baixa eletronegatividade foram classificados como bases moles e as bases duras seriam aquelas com propriedades opostas. Os ácidos com baixas cargas positivas, altamente polarizáveis e de grande tamanho foram chamados de moles e aqueles com propriedades opostas, de ácidos duros.

A teoria de orbital molecular classifica como ácidos moles, aqueles que possuem baixa energia do LUMO e bases moles, aqueles que possuem alta energia do HOMO. Os ácidos duros possuem alta energia do LUMO e as bases duras possuem baixa energia do HOMO. Assim, ácidos duros ligam-se fortemente às bases duras, devido à grande diferença entre os orbitais envolvidos, formando ligação iônica, conforme o exemplo da figura 16. Por outro lado, ácidos moles ligam-se fortemente às bases moles devido à proximidade entre as energias dos orbitais envolvidos, formando ligação covalente forte, conforme a figura 15²¹.

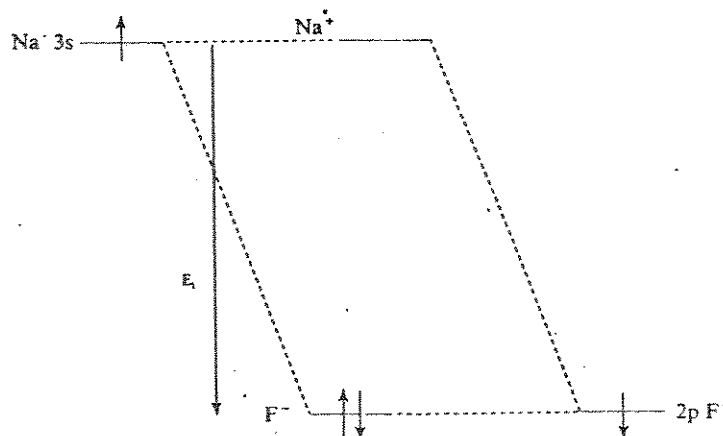


Fig.16 Ligação iônica formada pela interação de orbitais com grande diferença de energia.

Dizemos que uma molécula é mole, quando a separação dos orbitais de fronteira é pequena e a distribuição dos elétrons é facilmente rearranjada por uma perturbação. A molécula dura é aquela que possui grande separação entre os orbitais de fronteira e a distribuição eletrônica resiste ao rearranjo de uma perturbação. A dureza molecular é expressa pela diferença entre a energia de ionização do átomo neutro e seu ânion³¹:

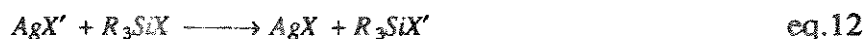
$$\eta = \frac{1}{2}(I - A_e) \quad \text{eq.10}$$

Onde I é o potencial de ionização e A_e é a afinidade eletrônica da molécula. Este valor é a metade da separação entre os dois orbitais de fronteira. A moleza é expressa pelo inverso da dureza:

$$\frac{1}{\eta} = \frac{2}{(I - A_e)} \quad \text{eq.11}$$

Os conceitos de dureza e moleza ajudam o Químico Inorgânico a prever reações; por exemplo, R_3Si^+ é um ácido duro e sais da forma $Ag X'$ convertem R_3SiX a R_3SiX' se X' for mais duro do que X. Isto ocorre, por que a base mais dura X' prefere o átomo Si duro e a

base X, mais mole, prefere o ácido Ag+ também mais mole:



III.1.3 MOMENTOS DIPOLARES

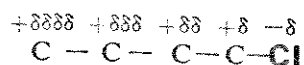
Momento de dipolo, μ , é um vetor, o qual é a medida da magnitude da carga deslocada quando átomos de eletronegatividades diferentes são interligados, e é definido por:

$$\mu = e\vec{d}$$

onde \vec{d} é o vetor distância que separa as cargas positivas e negativas de magnitude e.

A direção do momento de dipolo de uma molécula é baseada nas eletronegatividades relativas dos átomos componentes e o valor é obtido pelo vetor resultante dos momentos de dipolo de cada ligação presente na molécula.

Um dipolo, numa molécula, pode afetar um átomo não ligado diretamente, através de interações eletrostáticas por um efeito indutivo. Por exemplo, devido à diferença de eletronegatividade entre o C e o Cl, o Cl agirá como um retirador de elétrons, induzindo uma carga positiva no carbono adjacente. Este, por sua vez, tentará "acomodar" sua carga positiva induzida sobre todos os átomos a ele ligados, criando um efeito em todas as ligações sucessivas na molécula:



A formação de dipolos numa molécula, produzida por substituintes de diferentes eletronegatividades, altera propriedades moleculares como a acidez e a basicidade do composto, de forma que o momento de dipolo pode dar alguma indicação sobre a reatividade do mesmo³². Momentos dipolares são usados também para correlacionar estrutura química com atividade biológica.³³⁻³⁴

Os momentos dipolares também fornecem informações sobre a geometria, especialmente sobre os ângulos entre as ligações³⁵.

III.1.4 ELETRONEGATIVIDADE DE MULLIKEN

Um dos conceitos mais familiares e úteis em química é a eletronegatividade relativa. Pauling introduziu a eletronegatividade como "o poder de um átomo numa molécula para atrair a si mesmo, elétrons", de forma que a eletronegatividade é uma propriedade de ligação e não de átomos isolados. Assim, Pauling definiu como a diferença de energia entre uma ligação A-B e a média geométrica das energias de ligação A-A e B-B:

$$\Delta' = D(A-B) - [D(A-A)D(B-B)]^{\frac{1}{2}} \quad \text{eq.13}$$

Observou-se que a quantidade Δ' aumentava com a diferença de eletronegatividade entre A e B. Pauling conseguiu ajustar os dados disponíveis a uma expressão empírica:

$$\Delta' = (A-B) = 30(\chi_A - \chi_B)^2 \quad \text{eq.14}$$

onde $\chi_A - \chi_B$ é a diferença de eletronegatividade e $\Delta'(A-B)$ é expresso em quilocalorias. O número 30 é arbitrário, para trazer a escala de eletronegatividade a um intervalo numérico conveniente.³⁵

Uma abordagem diferente e mais física ao conceito de eletronegatividade foi feita por Mulliken.³⁶

Mulliken propôs que, dadas duas espécies "a" e "b", a energia necessária para "a" retirar um elétron de b é $I_b - A_a$, onde I e A são o potencial de ionização e a afinidade eletrônica, respectivamente. Da mesma forma, a energia necessária para "b" retirar um elétron de "a" é $I_a - A_b$. Quando a e b possuem a mesma eletronegatividade, $I_a - A_b = I_b - A_a$; assim, podemos escrever: $I_a + A_a = I_b + A_b$, o que sugere que podemos definir a eletronegatividade da seguinte forma:

$$\chi = \frac{1}{2}(I + A) \quad \text{eq.15}$$

onde $1/2$ é um fator arbitrário³¹.

Esse conceito foi usado por nós, para estimar a eletronegatividade de moléculas, onde I e A são, neste caso, o potencial de ionização e a afinidade eletrônica da molécula, I_m e A_m . O potencial de ionização (I_m) e a afinidade eletrônica (A_m) podem ser calculados ou determinados experimentalmente. A quantidade $\chi_m = \frac{1}{2}(I_m + A_m)$ daria a capacidade de uma molécula M_1 (por exemplo, o fármaco) atrair elétrons de uma outra molécula M_2 (por exemplo, o receptor), quando as duas, M_1 e M_2 se aproximam. O efeito eletrônico produzido por um substituinte numa molécula, pode ser indutivo ou de ressonância; os efeitos indutivos devem-se à migrações eletrônicas ao longo de ligações simples, devido à atração exercida por determinados grupos, em razão da sua eletronegatividade. Portanto, usamos essa propriedade como um parâmetro eletrônico.

III.1.5 VOLUME MOLECULAR

Usamos o volume molecular no nosso conjunto de parâmetros físico-químicos, pois o mesmo está correlacionado com o coeficiente de partição (P). Craig³⁷ observou que o coeficiente de partição está correlacionado com o parâmetro estérico (E_s) e com a refratividade além do volume molecular, mas concluiu ser de pouca valia o uso de mais de um desses parâmetros ao mesmo tempo.

O cálculo de volume molecular foi feito com o programa SURF (Gaudio, A. C., e Takahata, Y.). O algoritmo utilizado para estimar o volume foi proposto por Higo & Go³⁸, e o procedimento do cálculo é explicado abreviadamente, a seguir.

Considera-se, inicialmente, que a molécula está contida numa caixa, com arestas de comprimento múltiplo de um valor "a". A seguir, cada aresta é dividida em unidades de "a", de forma que a caixa transforme-se num conjunto de cubos de aresta com valor "a". Cada cubo é então, classificado em três tipos: interno, externo ou superficial à molécula. Se o cubo for externo, será desprezado. Se for interno, passará a fazer parte do volume molecular. Se for superficial, será subdividido em novos cubos de aresta com valor "a". Os novos cubos serão classificados e processados da mesma forma que os primeiros, e o procedimento

repetir-se-á até o nível estabelecido pelo operador. Finalmente, determinada fração do volume será computada como pertencente à molécula.

O volume foi usado por nós como um parâmetro estérico. Um exemplo da aplicação do volume molecular em QSAR, pode ser visto na referência 34, onde alguns pesquisadores propuseram que diversos parâmetros, entre eles o volume molecular, deveriam ser considerados na relação atividade biológica-estrutura química.

III.1.6 OUTROS PARAMETROS

Alguns ângulos diedros foram utilizados como parâmetros estéricos. No capítulo II, falamos sobre os ângulos diedros T1 e T2, os quais foram alterados a cada 15 graus, de 0 a 360 graus na análise conformacional dos compostos. Os mesmos foram empregados como parâmetros estéricos.

Usamos ainda o calor de formação (ΔH_f), na tentativa de correlacionar com a atividade biológica da nossa série de compostos.

III.2 CALCULOS

Os cálculos de todos os parâmetros físico-químicos foram realizados para os compostos ativos e inativos, na conformação E (conformação com os ângulos T1 e T2 próximos a 180 graus) e posteriormente, na conformação mais provável encontrada para cada composto, tendo sido estas conformações obtidas em meio biológico (a tabela 3 mostra a conformação mais estável para cada composto inativo e a tabela 5 mostra a conformação mais provável para cada composto ativo; para os compostos inativos, a conformação mais estável foi considerada também a mais provável, isto é, aquela com maior probabilidade de se ligar ao receptor).

Alguns parâmetros foram obtidos com o método semi-empírico MNDO-PM3, através

do programa MOPAC 6.0; outros, foram obtidos com o programa MM2 e o volume, com o programa SURF. A tabela 6 mostra os parâmetros calculados e os métodos utilizados:

PARAMETRO	METODO
HOMO	MNDO-PM3
LUMO	MNDO-PM3
χ	MNDO-PM3
η	MNDO-PM3
$1/\eta$	MNDO-PM3
μ	MM2
T1	MM2
T2	MM2
VOLUME	programa SURF
ΔH_f	MNDO-PM3

Tabela 6 - Métodos utilizados no cálculo dos parâmetros físico-químicos, onde χ é a eletronegatividade de Mulliken, η é a dureza molecular, $1/\eta$ é a moleza, μ é o momento dipolar, T1 e T2 são os ângulos diedros definidos no item II.2 e ΔH_f é o calor de formação.

III.3 RESULTADOS E DISCUSSÃO

A tabela 7 mostra os valores das propriedades físico-químicas calculadas para os compostos inativos e ativos, na conformação E e a tabela 8 mostra os valores das mesmas propriedades calculadas para os compostos inativos e ativos, na conformação mais provável.

O método utilizado para calcular os parâmetros eletrônicos já citados, foi o método semi-empírico MNDO-PM3 (modified neglect of diatomic overlap, parametric method 3), isto é, a terceira versão do MNDO (modified neglect of diatomic overlap), sendo que a segunda versão é o método AM1 (primeira reparametrização do MNDO)³⁹

Atualmente, o método MNDO-PM3 é considerado um dos melhores métodos semi-empíricos disponíveis, e por isso foi escolhido por nós.

O método MNDO foi parametrizado para H, C, N, O, F, Al, Si, P, S, Cl, Br e I, sendo que o número de compostos usados na parametrização foi muito menor do que no caso do MNDO-PM3, de forma que as diferenças entre as propriedades calculadas e observadas, usando o MNDO-PM3 foram bem menores do que aquelas usando MNDO ou AM1³⁹.

Os parâmetros do MNDO-PM3 foram otimizados para reproduzir as seguintes propriedades moleculares: calores de formação, momentos dipolares, potenciais de ionização e geometrias moleculares³⁹. A energia do LUMO, usada neste trabalho, não está incluída na lista das propriedades moleculares otimizadas pelo MNDO-PM3 e está relacionada com a afinidade eletrônica (Ae), bem como com a eletronegatividade de Mulliken, dureza e moleza moleculares, os quais são calculados com expressões que incluem o termo afinidade eletrônica.

Um outro possível método para calcular essas propriedades é o método semi-empírico HAM/3 (hydrogenic atoms in molecules) que foi otimizado para calcular a afinidade eletrônica, além das outras propriedades necessárias para o presente trabalho; entretanto, não foi parametrizado para os elementos Cl e S, presentes em alguns compostos da nossa série em estudo. O método HAM/3 foi parametrizado apenas para H, C, N, O e F⁴⁰.

Dessa forma, escolhemos o método MNDO-PM3 entre aqueles que foram parametrizados para todos os elementos por nós requisitados. Além disso, não estamos interessados em saber a energia do LUMO para cada composto independentemente, mas o que importa de

fato é a comparação entre as energias do LUMO (e das outras propriedades relacionadas) dos compostos em estudo.

COMPOSTO	HOMO (ev)	LUMO (ev)	η (ev)	$1/\eta$ (ev)	χ (ev)	μ (D)	T1 (grau)	T2 (grau)	V A ³	ΔH_f kcal
I	-9,261	-0,573	0,230	4,344	4,917	2,813	175,63	177,61	192,2	-11,42
II	-9,141	-0,635	0,235	4,253	4,888	2,211	176,78	181,25	210,9	-17,74
III	-8,807	-0,555	0,242	4,126	4,681	2,933	168,22	185,06	223,0	-46,07
IV	-8,830	-0,482	0,239	4,174	4,656	3,383	174,26	185,751	219,6	-45,30
V	-8,933	-0,452	0,236	4,241	4,692	2,240	177,82	172,21	259,3	-83,92
VI	-8,874	-0,389	0,236	4,242	4,631	2,971	154,93	173,09	283,1	-117,88
VII	-8,876	-0,437	0,237	4,220	4,656	3,671	178,25	161,56	325,7	-101,88
VIII	-9,891	-1,155	0,229	4,368	5,523	5,215	178,58	181,00	216,4	-12,75
IX	-9,362	-0,518	0,226	4,422	4,940	3,056	180,03	178,23	214,9	-18,84
X	-8,911	-0,632	0,242	4,139	4,772	2,536	168,26	172,88	271,3	-43,85
XI	-9,142	-0,685	0,237	4,228	4,913	2,664	158,01	161,80	211,1	27,47
XII	-9,126	-0,557	0,233	4,284	4,841	1,717	177,38	178,97	269,0	-109,15
XIII	-9,062	-0,369	-0,230	4,347	4,716	4,380	175,13	179,7	292,1	-134,56
XIV	-8,879	-0,464	0,238	4,208	4,671	3,616	177,43	167,18	312,0	-99,28
XV	-8,678	-0,132	0,234	4,273	4,405	3,819	179,32	162,34	333,3	-115,75
XVI	-8,603	-0,114	0,236	4,245	4,359	4,580	178,73	165,66	376,6	-134,10
XVII	-9,190	-0,798	0,238	4,196	4,994	2,067	165,70	172,82	240,0	20,10
XVIII	-8,899	-0,608	0,241	4,145	4,753	2,729	167,82	173,23	287,5	-52,83
XIX	-8,981	-0,513	0,236	4,234	4,747	1,260	177,57	172,06	296,7	-90,12
XX	-8,952	-0,796	0,245	4,078	4,874	2,052	168,25	172,80	286,6	-49,83
XXI	-8,640	-0,102	4,269	0,234	4,371	353,1	179,98	162,37	-106,27	3,953
XXII	-9,197	-0,577	4,310	0,232	4,887	211,8	176,21	177,61	-16,30	3,725
XXIII	-9,242	-0,834	4,204	0,238	5,038	224,0	157,88	161,85	21,49	1,859
XXIV	-9,047	-0,343	4,352	0,230	4,695	308,8	178,03	160,71	-135,96	4,436
XXV	-8,912	-0,626	4,144	0,241	4,770	273,0	167,47	174,09	-51,6	2,722

Tabela 7 - Parâmetros físico-químicos calculados para os compostos inativos e ativos, na conformação E e para o conjunto teste (XXI a XXV) também na conformação E.

COMPOSTO	HOMO (ev)	LUMO (ev)	η (ev)	$1/\eta$ (ev)	χ (ev)	μ (D)	T1 (grau)	T2 (grau)	V A ³	ΔH_f kcal
I	-9,261	-0,573	0,230	4,344	4,917	2,813	175,63	177,61	192,2	-11,42
II	-9,141	-0,635	0,235	4,253	4,888	2,211	176,78	181,25	210,9	-17,74
III	-8,807	-0,555	0,242	4,126	4,681	2,933	168,22	185,06	223,0	-46,07
IV	-8,913	-0,453	0,236	4,230	4,683	1,978	188,41	310,451	220,8	-44,13
V	-8,951	-0,403	0,234	4,274	4,677	1,884	161,99	56,381	258,8	-81,32
VI	-8,903	-0,350	0,234	4,277	4,627	2,149	164,08	59,380	283,2	-115,13
VII	-8,809	-0,362	0,237	4,223	4,586	2,305	261,86	60,606	328,5	-102,48
VIII	-9,920	-1,151	0,228	4,384	5,536	5,314	305,24	178,52	217,8	-12,48
IX	-9,362	-0,518	0,226	4,422	4,940	3,056	180,03	178,23	214,9	-18,84
X	-8,990	-0,557	0,237	4,216	4,773	2,854	150,78	57,380	271,0	-40,87
XI	-9,155	-0,686	0,236	4,235	4,921	2,889	146,045	66,74	208,3	29,92
XII	-9,126	-0,585	0,234	4,271	4,855	1,716	315,08	174,77	269,0	-108,42
XIII	-9,053	-0,364	-0,230	4,344	4,709	4,388	300,94	178,3	293,6	-134,37
XIV	-8,865	-0,474	0,238	4,196	4,671	3,609	304,23	180,06	312,9	-99,08
XV	-8,611	-0,093	0,235	4,259	4,352	1,885	186,82	86,170	337,8	-111,27
XVI	-8,533	-0,031	0,233	4,251	4,282	2,024	185,07	86,830	378,0	-129,72
XVII	-9,194	-0,731	0,236	4,231	4,963	2,678	151,70	58,270	239,6	23,35
XVIII	-8,981	-0,538	0,237	4,222	4,760	2,951	154,00	58,630	287,0	-49,21
XIX	-8,984	-0,459	0,235	4,263	4,721	1,984	162,77	56,49	273,9	-87,09
XX	-9,027	-0,708	0,245	4,085	4,868	2,671	152,86	57,71	286,1	-45,98
XXI	-8,556	-0,030	4,263	0,235	4,293	357,4	186,02	86,37	-101,54	2,104
XXII	-9,197	-0,577	4,310	0,232	4,887	211,8	176,21	177,61	-16,30	3,725
XXIII	-9,257	-0,810	4,224	0,237	5,030	223,4	146,28	66,31	24,035	2,545
XXIV	-8,393	-1,406	3,494	0,286	4,900	312,1	99,00	68,49	-28,45	3,129
XXV	-8,999	-0,562	4,219	0,237	4,781	271,8	153,72	59,26	-47,728	2,966

Tabela 8 - Parâmetros físico-químicos calculados para os compostos inativos e ativos, na conformação mais provável e para o conjunto teste (XXI a XXV) também na conformação mais provável.

CAPITULO IV - RELAÇÕES ENTRE ATIVIDADE BIOLÓGICA E ESTRUTURA.

IV.1 - INTRODUÇÃO

IV.1.1 - INTERAÇÃO FARMACO-RECEPTOR

Alguns fármacos manifestam atividade biológica em pequenas concentrações, sendo chamados de estruturalmente específicos; o efeito produzido por eles é atribuído à interação com uma substância receptora específica. Como resultado desta interação, o fármaco forma um complexo com o componente celular que recebe o nome de receptor.

Supõe-se que os fármacos estruturalmente específicos apresentam alto grau de complementaridade para com o sítio em que atuam, sendo que a interação de um fármaco deste tipo com o seu receptor assemelha-se à interação de um substrato com o centro ativo de uma enzima, conforme a figura 17:

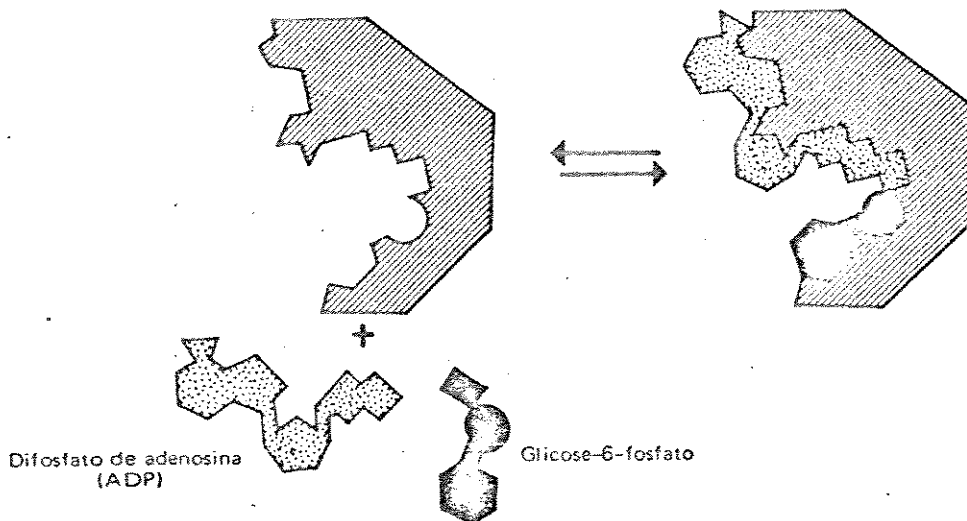


Fig.17 Interação entre os substratos ADP e glicose-6-fosfato, e o sítio ativo da enzima.

Provas experimentais indicam que os receptores estão situados em macromoléculas, a maioria delas com propriedades proteicas e que têm a capacidade específica de interagir pelo menos com os substratos naturais em seu centro ativo.

As forças de interação que ligam os fármacos aos receptores são as mesmas que ocorrem nas interações de moléculas simples. Em geral, formam-se ligações fracas do tipo: iônica, polares, pontes de hidrogênio, transferência de carga, hidrofóbicas, van der Waals. Dessa forma, os efeitos produzidos são reversíveis, isto é, rompem-se as ligações fármaco-receptor e o medicamento deixa de agir assim que diminui a sua concentração, como se deseja na maioria dos casos (que a ação produzida pelo fármaco dure um tempo limitado).

Sobre a topografia de receptores em geral, algumas hipóteses foram feitas.

Inicialmente, admitiu-se que o fármaco deveria apresentar estrutura complementar à do receptor, análoga àquela que há entre chave e fechadura. Entretanto, observações na estrutura e geometria e alguns fármacos sabidamente ativos mostraram que existe a possibilidade do receptor sofrer variações conformacionais para melhor "acomodar" a droga.

Um exemplo desse fato é dos receptores de analgésicos narcóticos; baseando-se num grupo de analgésicos análogos e derivados da morfina que apresentavam uma porção da molécula em comum, foi proposta uma topografia para o receptor, equivalente ou complementar àquela porção. Porém, outros analgésicos sintéticos (narcóticos), com ação análoga à da morfina, têm estrutura muito mais simples, não apresentando aquela porção da molécula em comum.

Assim, a primeira hipótese que admite sítio receptor essencialmente inflexível, no caso de analgésicos, foi modificada e, baseando-se no fato de que as macromoléculas podem sofrer variações conformacionais, Portoghese⁴¹ apresentou a hipótese de que a complexação de analgésicos narcóticos diferentes com os receptores, pode compreender modos diferentes de interação. Portanto, a interação de analgésicos diferentes com uma única espécie de receptores poderia se dar também através de alterações conformacionais do receptor (receptor dinâmico) em alguns casos; a interação poderia ainda, se dar com outras espécies de receptores.

No caso deste trabalho, não temos informação sobre o receptor daqueles compostos que apresentaram atividade. Devemos então, inicialmente, considerar a hipótese do receptor estático e dinâmico. Na análise conformacional obtivemos variações nas conformações mais

prováveis de cada composto ativo, como foi explicado no item II.2, de forma que um grupo de compostos ativos apresentou como conformação mais provável aquela com os anéis mais fechados (conformação B) e outro com os anéis um pouco mais afastados (E e H); observe na tabela 5 a ordem de probabilidade dos compostos ativos e na figura 10, as posições relativas entre os anéis para cada conformação.

Assumindo a hipótese de que o receptor pode alterar a sua conformação para melhor se ajustar à droga, admitimos que cada composto adquiriu a sua conformação mais provável (para os compostos inativos, a conformação mais provável é a mais estável energeticamente) e assumindo a hipótese de receptor estático, sugerimos que todos os compostos ativos apresentariam a mesma conformação e escolhemos a conformação E por ser um mínimo local presente em todos os compostos (ativos e inativos) e com alta probabilidade para os ativos e para a maioria dos inativos. Assumindo que a conformação E é a conformação complementar ao receptor, já estamos levando em consideração também o fato de que a conformação não é o único fator responsável pela atividade, uma vez que alguns compostos inativos apresentam esta conformação como mais estável.

Devemos lembrar ainda que alguns compostos apresentavam conformações com seus respectivos enantiômeros. Este fato deve ser considerado, pois mesmo os dois isômeros apresentando propriedades físico-químicas idênticas, as conformações são diferentes, podendo alterar a atividade, se a conformação do fármaco for fator determinante. Dentro dos parâmetros que usamos para correlacionar a atividade biológica com estrutura, somente os ângulos diedros seriam diferentes para os enantiômeros; o método que usamos para correlacionar atividade biológica com estrutura será explicado no próximo tópico, mas resumidamente, podemos adiantar que esse método classifica um dado composto em estudo como ativo ou inativo, através da semelhança entre os parâmetros físico-químicos do composto em estudo e da série de compostos para a qual foi dada a informação sobre a atividade biológica. Entretanto, num par de enantiômeros não sabemos qual é complementar ao receptor; então devemos escolher qualquer uma das duas conformações, desde que seja a mesma tanto para o conjunto de compostos para o qual a atividade biológica foi medida, quanto para o conjunto que se deseja classificar. Por exemplo, o par de enantiômeros E e H possui as mesmas propriedades físico-químicas, mas ângulos diedros diferentes; então devemos escolher entre essas duas conformações, a mesma para todos os compostos.

IV.1.2 METODOS UTILIZADOS PARA CORRELACIONAR ATIVIDADE BIOLOGICA COM ESTRUTURA QUIMICA.

Para analisar o nosso conjunto de compostos, escolhemos o conjunto de métodos estatísticos chamados de Reconhecimento de Padrão (RP) ⁴², por ser vantajoso em relação aos outros métodos conhecidos, principalmente no seguinte aspecto:

- Em RP, não é fundamental que se conheça os valores das atividades biológicas, bastando saber se o composto apresenta ou não atividade, como é o nosso caso. Métodos como o de Hansch ⁴³ e Free Wilson ⁴⁴ requerem esse tipo de informação.

Os métodos de RP permitem determinar as propriedades físico-químicas capazes de distinguir os análogos ativos dos inativos, reduzir o número de propriedades físico-químicas sem perdas de informações e ainda decidir quais análogos sintetizar.

Em RP, chamamos os compostos de *objetos* e o conjunto de objetos com a mesma propriedade (ativa ou inativa) de *categoria*.

Assim, os compostos ativos são os objetos de uma categoria e os inativos são os objetos de outra categoria.

Os objetos com categorias conhecidas pertencem ao conjunto chamado treinamento, do qual informações são extraídas para tornar possível a classificação de novos objetos (conjunto teste), com base nas medidas feitas para estes objetos. Em algumas aplicações, estas informações são usadas para classificar um conjunto de objetos adicionais, onde se conhecem as medidas, mas as categorias são desconhecidas (fig.18).

		Objetos					
		1	2 . . .	K	N	
Variáveis	1	y_{11}	y_{12}	\cdots	y_{1K}	\cdots	y_{1N}
	2	y_{21}	y_{22}	\cdots	y_{2K}	\cdots	y_{2N}
	3	y_{31}	y_{32}	\cdots	y_{3K}	\cdots	y_{3N}
	...			\vdots		\vdots	
	i	y_{i1}	y_{i2}	\cdots	y_{iK}	\cdots	y_{iN}
M	y_{M1}	y_{M2}	\cdots	y_{MK}	\cdots	y_{MN}	

categoria 1	catego- ria Q	conjunto teste
conjunto de aprendizagem (conjunto de treinamento)		

Fig. 18 - Representação dos N objetos num espaço M dimensional, onde M é o número de medidas feitas sobre esses N objetos.

A representação gráfica mostra uma idéia central dos métodos de RP. Os objetos são considerados pontos no espaço M dimensional, onde M é o número de medidas feitas sobre esses objetos.

A distância euclidiana entre dois pontos i e j de coordenadas Y_{ki} e Y_{kj} é dado por:

$$d_{ij} = \left[\sum_{k=1}^M (Y_{ki} - Y_{kj})^2 \right]^{\frac{1}{2}}$$

onde a soma é feita sobre as M medidas e d_{ij} pode ser relacionado com uma medida de similaridade. Considera-se que quanto menor a distância entre estes, maior a similaridade. Uma medida de similaridade é definida como:

$$s_{ij} = 1 - d_{ij} / (d_{ij})_{\max}$$

onde $(d_{ij})_{\max}$ é a maior distância interponto no espaço M dimensional. Para objetos idênticos, $s_{ij} = 1$, enquanto que $s_{ij} = 0$ descreve dois pontos que são separados pela maior distância interponto para o conjunto sob estudo.

Antes de iniciar a análise com RP, foi necessário escalonar os valores obtidos para as propriedades físico-químicas, pois estas possuem unidades diferentes. O escalonamento dos dados é feito para possibilitar as comparações na mesma escala. Este método é explicado mais detalhadamente no apêndice 3.

Com os dados já escalonados aplicou-se o método de Transformação de Karhunen-Loeve que também é chamado de Análise de Componentes Principais. Este método tem por finalidade reduzir o número de variáveis (propriedades físico-químicas) na análise. O método gera novas variáveis que são combinações lineares das originais, de forma que a primeira nova variável contém a maior variância e assim sucessivamente, até um ponto onde as últimas variáveis têm variâncias que são zero ou próximas a zero. As variáveis com variâncias próximas a zero são truncadas, diminuindo a dimensionalidade do problema.

Gráficos bidimensionais entre as componentes principais foram então gerados com a finalidade de verificar se os compostos podem ser separados nas suas respectivas categorias usando-se apenas duas componentes principais, e dessa forma verificar quais são as propriedades físico-químicas mais importantes na discriminação dos compostos.

Outro método aplicado foi o Agrupamento Hierárquico, que agrupa os compostos similares da seguinte forma: uma matriz de similaridade é obtida e percorrida para o maior valor; os dois pontos produzindo este valor são juntados. Daí em diante, os dois pontos são considerados como único (formando um centro de gravidade) para calcular uma nova e menor matriz de similaridade. Este processo continua até que todos os pontos estejam contidos em um agrupamento.

Finalmente, o método SIMCA (Statistical Isolinear Multiple Component Analysis) foi aplicado. Este método usa análise de componentes principais para construir separadamente um modelo para cada categoria do conjunto de treinamento; cada classe de objetos similares é descrita por um modelo com um pequeno número de componentes principais e uma vez que os modelos para cada classe tenham sido estabelecidos, poderão ser usados para

classificar objetos pertencentes à categorias desconhecidas.

O SIMCA é capaz de indicar quando um objeto não pertence a nenhuma das categorias definidas no conjunto de treinamento. O apêndice 3 mostra como o SIMCA constrói os modelos para cada categoria.

IV.2 CALCULOS

Utilizamos o programa ARTHUR/UNICAMP (Adaptação do Programa ARTHUR/75 de Computadores "Mainframe" para microcomputadores⁴⁵). Os seguintes sub programas do programa ARTHUR, foram utilizados:

SCAL - Usado para escalonar os dados.

DISTAN - Calcula a matriz de distâncias entre todos os objetos.

KARLOV - Faz o cálculo de componentes principais.

SIMCA - Constrói um modelo de componentes principais para cada classe (ou categoria), com a finalidade de classificar objetos do conjunto teste, baseando na similaridade com os modelos que representam as classes.

VARVAR - É usado para fazer gráficos em duas dimensões usando a impressora.

HIER - Produz um dendograma que descreve o agrupamento hierárquico.

Os objetos do conjunto de treinamento foram os vinte compostos das tabelas 1 e 2, e o conjunto teste é mostrado na tabela 9. As propriedades físico-químicas calculadas são mostradas nas tabelas 7 e 8, juntamente com os compostos do conjunto de treinamento.

CONJUNTO TESTE	
COMPOSTO	ESTRUTURA
XXI	
XXII	
XXIII	
XXIV	
XXV	

Tabela 9 - Compostos utilizados no conjunto teste.

IV.3 - RESULTADOS E DISCUSSAO

IV.3.1 - ANALISE DOS RESULTADOS, CONSIDERANDO RECEPTOR RIGIDO

Como já foi explicado, para o caso de receptor rígido, escolhemos a conformação E para os compostos. Os resultados obtidos com a análise de componentes principais são apresentados na tabela 10, onde são mostrados os autovalores, a porcentagem de variância correspondendo à cada autovetor, a porcentagem de variância acumulada e os coeficientes das propriedades físico-químicas calculadas.

Na tabela 10, vemos que os dois primeiros autovetores contêm 71% da variância total.

A soma dos autovetores devem corresponder a 100% da variância total. Os autovetores ou componentes principais são funções lineares das propriedades fisico-químicas e podem ser escritos como segue:

$$\text{C.P. 1} = -0,448 \text{ HOMO} - 0,397 \text{ LUMO} + 0,202 \eta - 0,205 \frac{1}{\eta} + 0,443 \chi - 0,407 \text{ V} - 0,004 \text{ T1} + 0,260 \text{ T2} + 0,359 \Delta H_f - 0,022 \mu$$

$$\text{C.P. 2} = -0,133 \text{ HOMO} + 0,225 \text{ LUMO} + 0,514 \eta - 0,512 \frac{1}{\eta} - 0,025 \chi + 0,143 \text{ V} + 0,423 \text{ T1} - 0,017 \text{ T2} - 0,267 \Delta H_f + 0,367 \mu$$

A primeira componente principal contém 44% da variância total e as propriedades importantes são: HOMO, LUMO, χ , V e ΔH_f ; a segunda componente principal contém 27% da variância total, com η , $\frac{1}{\eta}$, μ e T1 importantes. Praticamente todas as propriedades são importantes para explicar a atividade biológica, pois apenas T2 possui um coeficiente pouco significativo nas duas primeiras componentes principais.

Auto-valor	Vari- ancia	Vari- ancia Acumu- lada	Autovetores									
			HOMO (ev)	LUMO (ev)	η (ev)	$1/\eta$ (ev)	χ (ev)	V A ³	T1 (grau)	T2 (grau)	ΔH_f kcal	μ D
4,397	44,0	44,0	-0,448	-0,397	0,202	-0,205	0,443	-0,407	-0,004	0,260	0,359	-0,022
2,744	27,4	71,4	-0,133	0,225	0,514	-0,512	-0,025	0,143	0,423	-0,017	-0,267	0,367
0,948	9,5	80,9	0,198	0,243	0,000	-0,041	-0,200	-0,283	0,264	0,728	-0,131	-0,404
0,749	7,5	84,4	-0,024	-0,279	-0,328	0,326	0,116	0,168	0,245	0,434	-0,194	0,616
0,669	6,7	95,1	0,102	0,226	0,131	-0,122	-0,246	-0,278	-0,685	0,245	0,069	0,485

Tabela 10 - Resultados da análise de componentes principais para os compostos na conformação E.

Para verificar se as duas primeiras componentes principais são suficientes para separar as duas categorias, traçamos o gráfico dessas duas componentes, mostrado na figura 19.

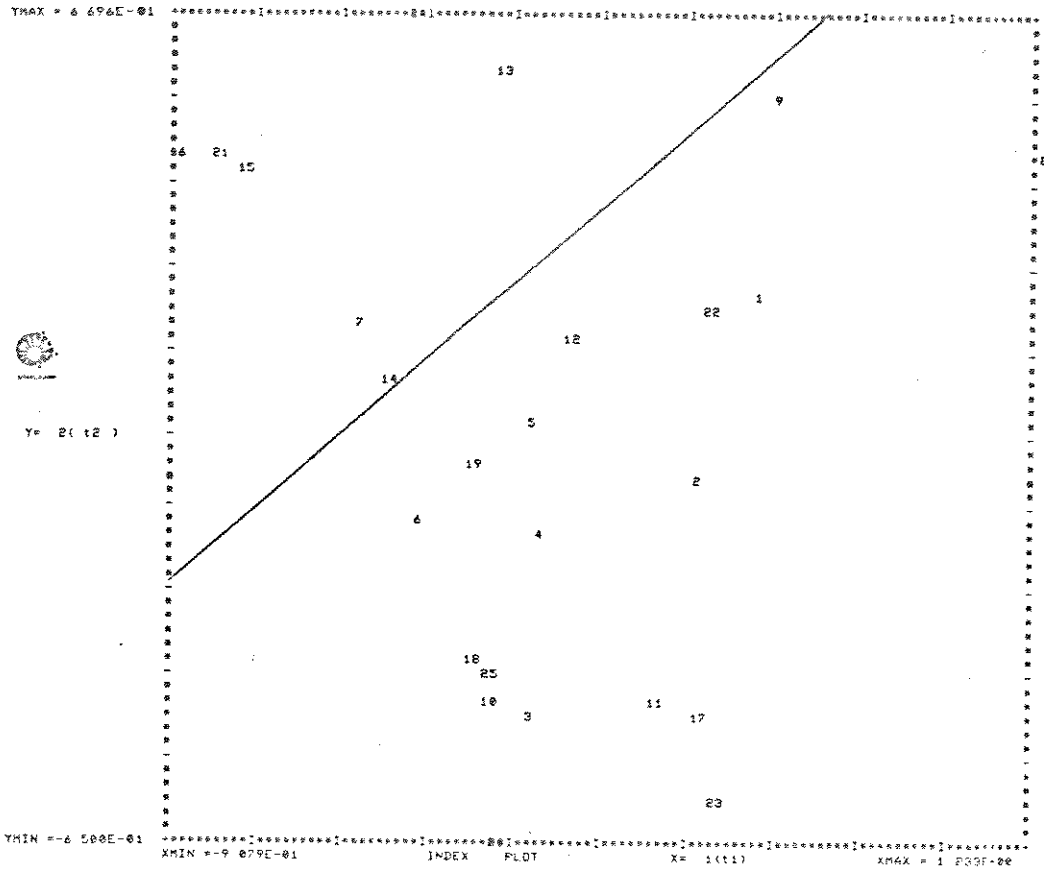


Fig.19 - Gráfico traçado usando-se as duas primeiras componentes principais, para os compostos na conformação E.

Observamos um espalhamento dos dados, o que indica que as duas primeiras componentes principais, apesar de conterem 71% da variância total, não estão explicando bem a atividade biológica. Talvez um número maior de componentes principais seja necessário para discriminar os compostos, o que leva à inclusão da propriedade T2 como fator importante, pois:

C.P. 3 = 0,198 HOMO + 0,243 LUMO + 0,000 η - 0,041 $\frac{1}{\eta}$ - 0,200 χ - 0,283 V + 0,264 T1 + 0,728 T2 - 0,131 ΔH_f - 0,404 μ , onde T2 é realmente um termo significativo.

A seguir, obtivemos o dendograma, mostrado na figura 20. Para o valor de similaridade \cong a 0,6, vemos a separação dos compostos em dois grupos e o composto VIII separado desses grupos. Um dos grupos é formado pelos compostos VII, XIV, XV, XVI e XIII e chamaremos de A e o outro grupo (B) pelos demais compostos. Se voltarmos ao gráfico da figura 19, veremos que os compostos VII, XIV, XV, XVI e XIII estão realmente formando um agrupamento e o composto VIII encontra-se bem separado de todos os outros compostos. Esses resultados são portanto, concordantes, mas não estão separando os compostos nas duas categorias. Podemos apenas concluir que o composto VIII é bem diferente dos demais e que talvez o grupo A, formado pelos compostos VII, XIV, XV, XVI e XIII contenha os compostos com maior atividade, pois nesse grupo há apenas um composto inativo para 4 ativos, enquanto que no outro grupo, há 9 os compostos inativos restantes, agrupados com apenas cinco ativos, restantes. Porém, não podemos nos certificar disto, pois não temos valores para a atividade biológica de todos os compostos.

Entretanto, é interessante ressaltar que o dendograma baseia-se na matriz de distância euclidiana, a qual é construída considerando um espaço em termos de todas as variáveis originais, o que indica a possibilidade dos parâmetros físico-químicos não estarem explicando suficientemente bem a atividade biológica.

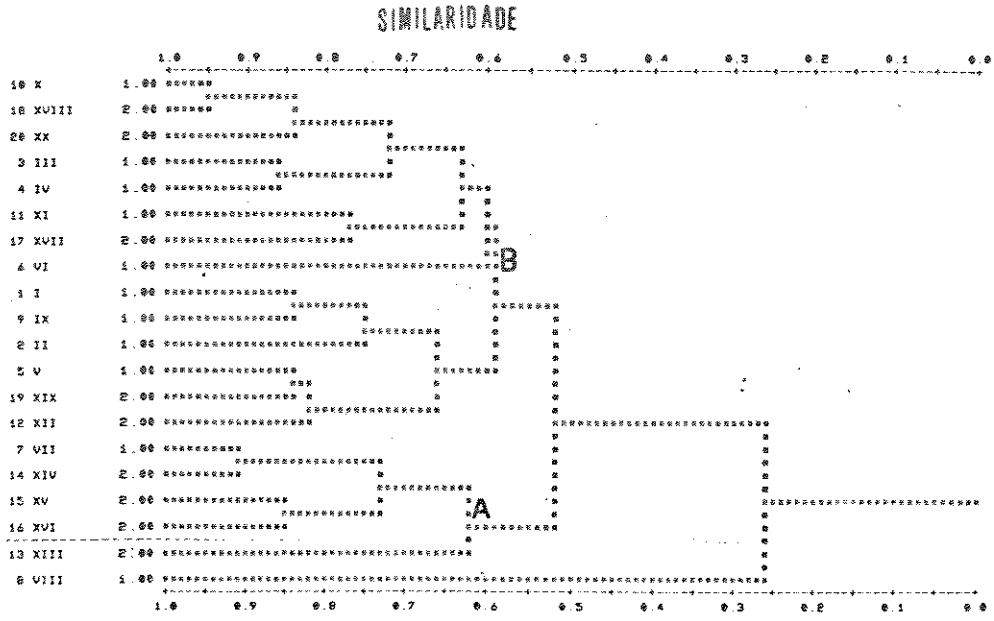


Figura 20 - Dendrograma obtido com o método "Agrupamento por Hierarquia", para os compostos na conformação E.

A aplicação do método SIMCA, entretanto, deu bons resultados. O método conseguiu criar um modelo para cada categoria, com 90% de acerto, usando 4 componentes principais para as duas categorias. O número de componentes principais confirma a necessidade de se incluir todas as propriedades físico-químicas calculadas.

Os resultados são mostrados na tabela 11:

CATEGORIA	NUMERO TOTAL DE COMPOSTOS	COMPOSTOS CLASSIFICADOS INCORRETAMENTE
ATIVA	9	-
INATIVA	11	VII e X

Tabela 11 - Resultados obtidos com o método SIMCA para os compostos na conformação E.

Apenas os dois compostos inativos, VII e X foram classificados como ativos, mas todos os compostos ativos foram classificados corretamente. Esse resultado nos animou a propor um conjunto teste; o composto XXI da tabela 9 foi escolhido, por ser bem parecido estruturalmente, com os compostos ativos XV e XVI e por não haver nenhum composto inativo, parecido com o mesmo. O composto XXV também é estruturalmente parecido com o composto ativo XVIII. Os compostos XXIII e XXIV são parecidos com compostos ativos e inativos (XXIII é parecido com XVII e com XI e XXIV é parecido com XIII e com VI), mas queríamos ter uma idéia, nesses casos, de qual sítio seria mais importante. Por exemplo, o composto XXIII difere do composto XVII, porque em XXIII, o carbono número 13 não é substituído, mas em XVII há um radical $-CH_3$; por outro lado, o composto XI não tem o carbono de número 13 substituído, nem o carbono número 21. A figura 21 mostra essas informações:

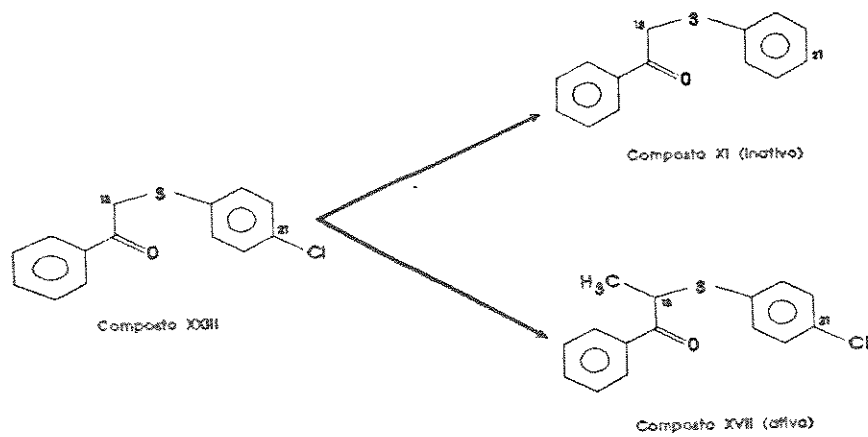


Fig.21 Comparação do composto XXIII do conjunto teste com os compostos XI e XVII.

Como o composto XXIII foi classificado como inativo, é mais parecido com XI do que com XVII. Então, concluímos que a substituição no carbono número 21 não é tão importante, quanto no carbono número 13. Observe agora, na figura 22, um esquema semelhante para o composto XXIV:

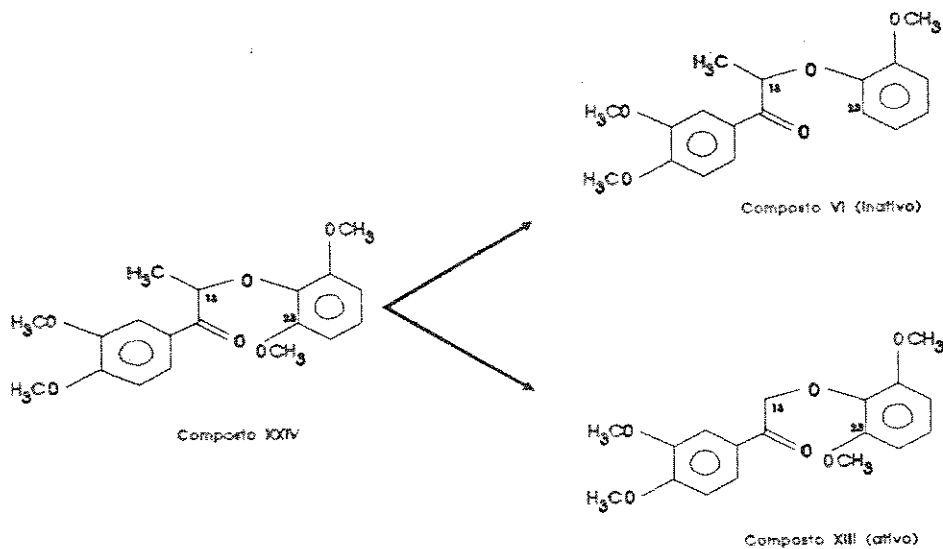


Fig. 22 Comparação do composto XXIV do conjunto teste com os compostos VI e XIII.

Novamente, o composto foi classificado como inativo e concluímos que a substituição no carbono 23 também não é tão importante quanto no carbono 13.

O composto XXII é parecido com os compostos inativos, I, II, III, IV e VIII mas desejávamos verificar uma nova substituição no carbono 19. Esse composto também foi classificado como inativo.

A tabela 12 mostra a classificação do conjunto teste, pelo método SIMCA:

COMPOSTO	CLASSIFICACAO
XXI	ATIVO
XXII	INATIVO
XXIII	INATIVO
XXIV	INATIVO
XXV	ATIVO

Tabela 12 - Resultados obtidos com o método SIMCA para o conjunto teste (conformação E).

O método de escolha do conjunto teste foi intuitivo, sendo que os métodos de RP é que vão nos dizer quais compostos deste conjunto, deve-se sintetizar.

IV.3.2 ANALISE DOS RESULTADOS, CONSIDERANDO RECEPTOR FLEXIVEL

Para este caso (receptor flexível), os compostos foram estudados na sua conformação mais provável (ou mais estável energeticamente), como já foi explicado.

Os resultados aqui, foram bem parecidos com aqueles para os compostos na conformação E.

A tabela 13 mostra os resultados obtidos com a análise de componentes principais:

Auto-valor	Vari- ancia	Vari- ancia Acumu- lada	Autovetores									
			HOMO (ev)	LUMO (ev)	η (ev)	$1/\eta$ (ev)	χ (ev)	V A ³	T1 (grau)	T2 (grau)	ΔH_f kcal	μ D
4,842	48,4	48,4	-0,443	-0,394	0,239	-0,222	0,434	-0,354	0,105	0,179	0,305	0,305
2,329	23,3	71,7	-0,012	-0,237	-0,444	0,458	0,102	-0,244	-0,469	-0,214	0,417	-0,165
1,282	12,8	84,5	0,006	0,254	0,437	-0,449	-0,121	-0,237	-0,520	-0,112	0,198	-0,387
0,978	9,8	94,3	0,164	0,102	-0,128	0,102	-0,141	-0,363	0,075	0,846	0,049	-0,247
0,366	3,7	98,0	0,211	0,239	-0,017	-0,039	-0,230	0,049	-0,372	0,167	0,221	0,792

Tabela 13 - Resultados obtidos com a análise de componentes principais, para os compostos na conformação mais provável.

As duas primeiras componentes principais contêm 71,7% da variância total e são descritas como segue:

$$\text{C.P.1} = -0,443 \text{ HOMO} - 0,394 \text{ LUMO} + 0,239 \eta - 0,222 \frac{1}{\eta} + 0,434 \chi - 0,354 V + 0,105 T1 + 0,179 T2 + 0,305 \Delta H_f + 0,305 \mu$$

$$\text{C.P.2} = 0,012 \text{ HOMO} - 0,237 \text{ LUMO} - 0,444 \eta + 0,458 \frac{1}{\eta} + 0,102 \chi - 0,244 V - 0,469 T1 - 0,214 T2 + 0,417 \Delta H_f - 0,165 \mu$$

Na primeira componente principal, as propriedades HOMO, LUMO, χ , V, ΔH_f e μ são importantes e na segunda componente principal, as mais importantes são: η , $1/\eta$, T1, ΔH_f .

Novamente, T2 não está incluída como uma propriedade importante nas duas primeiras

componentes principais.

O gráfico das duas primeiras componentes principais é mostrado na figura 23 e podemos observar que os compostos XII, XIII, XIV, XV, XVI, VI e VII formam um grupo separado dos demais e novamente o composto VIII aparece fora dos dois grupos.

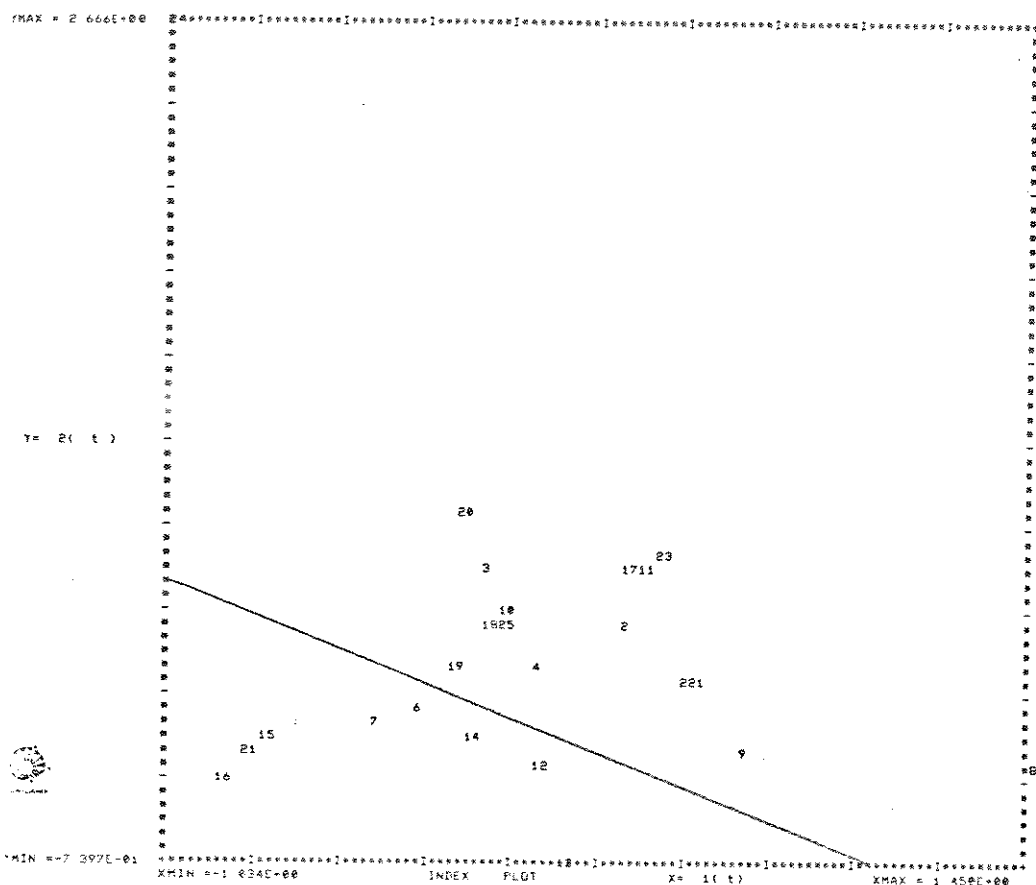


Figura 23 - Gráfico traçado com as duas componentes principais, para os compostos na conformação mais provável.

O dendograma, também mostra essa separação, num valor de similaridade de $\approx 0,6$. A figura 24 mostra esse resultado e chamamos os grupos de A' e B'. O grupo A' neste caso é bem parecido com o grupo A da análise anterior, mas agora mais um composto ativo foi incluído.

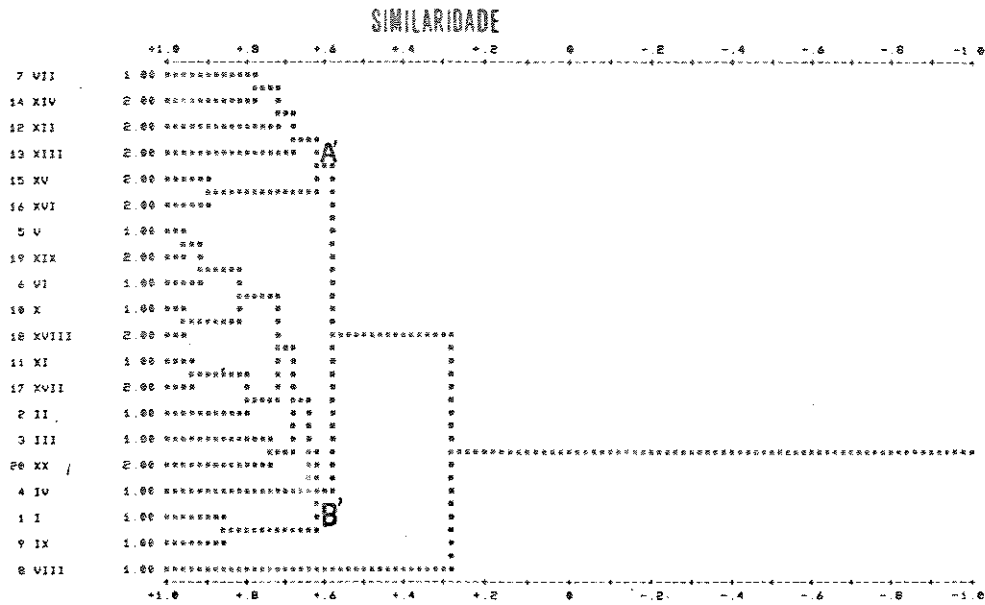


Figura 24 - Dendograma obtido com o método "Agrupamento por Hierarquia", para os compostos na conformação mais provável.

O método SIMCA novamente conseguiu criar um modelo para cada categoria, com 90% de acerto, usando 5 componentes principais para cada categoria. O resultado é mostrado na tabela 14.

CATEGORIA	NUMERO TOTAL DE COMPOSTOS	COMPOSTOS CLASSIFICADOS INCORRETAMENTE
ATIVA	9	XVII e XIX
INATIVA	11	-

Tabela 14 - Resultados obtidos com o método SIMCA, para os compostos na conformação mais provável.

Entretanto, neste caso, dois compostos ativos foram classificados como inativos e nenhum inativo foi classificado incorretamente.

O conjunto teste foi o mesmo escolhido anteriormente (compostos XXI a XXV), mas agora cada um com a sua conformação mais provável. Três compostos foram classificados como ativos, conforme a tabela 15:

COMPOSTO	CLASSIFICACAO
XXI	ATIVO
XXII	INATIVO
XXIII	ATIVO
XXIV	ATIVO
XXV	INATIVO

Tabela 15 - Resultados obtidos com o método SIMCA, para o conjunto teste na conformação mais provável.

Observou-se que apenas o composto XXII não foi classificado como ativo em nenhum caso e este é justamente o composto parecido apenas com os compostos inativos. O

composto XXIII e XXIV agora foram classificados como ativos, mostrando que se o receptor for flexível, as substituições nos carbonos 21 e 23, respectivamente são mais importantes do que no carbono 13. O composto XXV, por sua vez, apresentou-se inativo, mostrando a importância da substituição no carbono 13 neste caso.

IV.4 - CONCLUSOES

Os métodos de RP utilizados, mostraram que todas as propriedades físico-químicas são importantes para classificar os compostos, pois o método SIMCA que deu os melhores resultados nos dois casos (receptor rígido e flexível) usou quatro componentes principais na primeira análise e cinco na segunda.

A tentativa de discriminar os compostos em duas categorias através do gráfico, usando as duas primeiras componentes principais não mostrou bons resultados, confirmando a necessidade de se incluir outras componentes principais; mas, apesar de todas as propriedades físico-químicas serem importantes, talvez não estejam sendo suficientes para explicar a atividade biológica, uma vez que apenas um dos métodos aplicados conseguiu discriminar os compostos.

Notou-se, contudo, que os resultados obtidos com o método "agrupamento por hierarquia" concordam com aqueles obtidos com os gráficos traçados com as duas primeiras componentes principais; resultados numéricos para a atividade biológica poderiam auxiliar a análise neste caso, pois talvez os grupos A e A' nos dendogramas das duas análises sejam formados pelos compostos mais ativos e os grupos B e B' sejam formados pelos compostos inativos e outros com atividade baixa.

Normalmente usa-se esses resultados para escolher quais compostos do conjunto teste não devem ser sintetizados, isto é, a escolha dos "candidatos" para síntese é feita por exclusão (Bruns, R. E. - comunicação pessoal). Portanto, quanto ao conjunto teste, concluiu-se que o composto XXI deve ter a maior prioridade para síntese, pois foi classificado como ativo nos dois casos (receptor flexível e rígido) e a síntese de XXIII, XXIV e XXV deve ter segunda prioridade; o composto XXII não deve ser sintetizado.

CAPITULO V - CONCLUSOES GERAIS

Pretendíamos com este trabalho, propor uma metodologia para estudar as relações estrutura-atividade de uma série de compostos bem diversificada quanto às substituições e onde apenas alguns apresentaram atividade biológica, usando métodos computacionais com o intuito de possibilitar a sugestão de novos compostos para síntese, baseando-nos nas propriedades físico-químicas dos compostos com atividade biológica conhecida.

Concluimos que a análise conformacional desses compostos (com grande número de átomos) é mais eficiente com o método de mecânica molecular do que com os semi-empíricos, devido à rapidez do primeiro que permite fazer uma análise mais detalhada, com maior número de cálculos. Concluimos também a importância de considerar meio biológico na análise conformacional de fármacos, devido à alteração nos mínimos globais que pudemos observar.

No estudo dos compostos com atividade biológica foi importante considerar a flexibilidade das moléculas no momento de interação com o receptor, pois isto levou a uma conclusão mais realista sobre as conformações ativas mais prováveis, isto é, observamos que a ordem das probabilidades relativas entre os mínimos locais, calculadas usando a distribuição de Boltzmann, sofreram alterações quando consideramos que cada mínimo local poderia ser qualquer conformação pertencente ao "poço" de determinado mínimo local, com energia entre a energia mínima do poço (E) e $E + 2RT$.

Outro fator importante considerado aqui foi a possibilidade do receptor ser uma molécula rígida ou flexível, uma vez que não conhecemos nada sobre os receptores de neolignanas ativas contra Leishmaniose. Pudemos notar significativa diferença na sugestão de novos compostos, quando consideramos que todas as moléculas ligar-se-iam ao receptor com uma única conformação ou quando as moléculas adotariam a conformação mais provável ao interagir com o receptor (flexível).

Quanto aos parâmetros físico-químicos calculados, parece que não estão sendo suficientes para explicar a atividade biológica, apesar dos métodos de RP mostrarem que foi necessário a utilização de todos os parâmetros para discriminar os compostos nas duas categorias. Apenas o método SIMCA conseguiu discriminar com 90% de acerto, mas os

resultados obtidos com os outros métodos de RP talvez possam ser melhorados com a inclusão de outras variáveis (parâmetros físico-químicos). Concluimos ainda que existe a possibilidade desses outros métodos com resultados aparentemente ruins estarem separando os compostos mais ativos daqueles com atividade baixa e inativos. Nesse caso, o problema não seria com os parâmetros físico-químicos e valores para a atividade dos compostos auxiliariam essa análise. Um maior número de compostos, também poderia melhorar os resultados obtidos com RP, que baseia-se na comparação entre os objetos de estudo.

APENDICE 1 - O METODO DE MECANICA MOLECULAR. 7

O método de mecânica molecular usa um conjunto de funções potenciais para descrever as interações intramoleculares do sistema em estudo. Esse conjunto de funções é chamado "Campo de força" e contém parâmetros ajustáveis que são otimizados para obter o melhor conjunto de propriedades das moléculas, como geometrias, calores de formação, energias conformacionais ou outras propriedades. Esse método baseia-se na aproximação de Bor-Oppenheimer, comumente usada em mecânica quântica, que estabelece que a equação de Shrödinger para moléculas pode ser separada em duas partes estudadas independentemente: uma parte descrevendo o movimento dos elétrons e outra o movimento do núcleo. Normalmente, a posição do núcleo é fixada e então, a estrutura eletrônica é estudada. Em mecânica molecular, é feito o oposto; os movimentos do núcleo são estudados e os elétrons não são considerados explicitamente, mas como um potencial efetivo. A energia da molécula no estado eletrônico fundamental é função da posição nuclear.

Os cálculos de campo de força têm suas origens na espectroscopia vibracional, sendo necessário entendermos inicialmente o formalismo dos campos de força apropriados para análise vibracional.

Quando uma molécula com "n" átomos, definida por 3n coordenadas "x" é deformada da sua geometria de energia potencial mínima, V_0 , e coordenadas, x_0 a energia potencial pode ser escrita como uma expansão da série de Taylor (equação 16):

$$V_{pot} = V_0 + \sum_{i=1}^{i=3n} \left(\frac{\partial v}{\partial x_i} \right)_0 \Delta x_i + \frac{1}{2} \sum_{i,j=1}^{i=3n} \left(\frac{\partial^2 v}{\partial x_i \partial x_j} \right)_0 \Delta x_i \Delta x_j + \frac{1}{6} \sum_{i,j,k=1}^{i=3n} \left(\frac{\partial^3 v}{\partial x_i \partial x_j \partial x_k} \right)_0 \Delta x_i \Delta x_j \Delta x_k + \dots$$

Na análise vibracional de uma molécula tendo geometria correspondente à energia mínima, o primeiro termo " v_0 " é zero. Da definição de mínimo potencial, segue que nesta geometria o termo da primeira derivada também se anula; de acordo com a aproximação harmônica, os termos maiores do que o quadrático são truncados e a energia potencial depende apenas do terceiro termo.

Chamando a derivada segunda de f_{ij} (constante de força), obtemos a seguinte equação

(eq.17) para um campo de força harmônico:

$$V_{pot} = \frac{1}{2} \sum_{i,j=1}^{i=3n} f_{ij} \Delta x_i \Delta x_j \quad \text{eq.17}$$

Em mecânica molecular, apenas os termos onde $i=j$ são considerados, resultando num campo de força com potencial harmônico, seguindo a lei de Hooke, do tipo mostrado na eq.18:

$$V_{pot} = \frac{1}{2} \sum_i f_{ii} (x_i - x_{0i})^2 \quad \text{eq.18}$$

Os termos (funções potenciais) incluídos num campo de força simples são: estiramento de ligação, deformação angular, torsão e interações entre átomos não ligados, onde cada uma dessas funções potenciais assumem forma da eq.18.

Assim, a função de estiramento de ligação seria:

$$V_b = \frac{1}{2} \sum k_b (r - r_0)^2 \quad \text{eq.19}$$

onde k_b é a constante de força harmônica para o estiramento de uma ligação particular e r_0 é o comprimento de ligação de referência; o somatório é feito sobre todas as ligações. k_b e r_0 são parâmetros ajustáveis no campo de força.

A figura 25 mostra a curva da lei de Hooke, a curva onde os termos de ordem mais alta são adicionados à equação 19 e a curva de Morse para o estiramento de ligação. A lei de Hooke é uma boa aproximação para pequenas distorções da molécula. Para grandes distorções, os termos cúbicos, de quarta ordem, etc. devem ser adicionados à equação 19 e alternativamente, esse campo de força pode ser representado pela função de Morse (equação 20):

$$D (1 - \exp [-a (r - r_e)])^2 \quad \text{eq.20}$$

onde "D" e "a" são parâmetros.

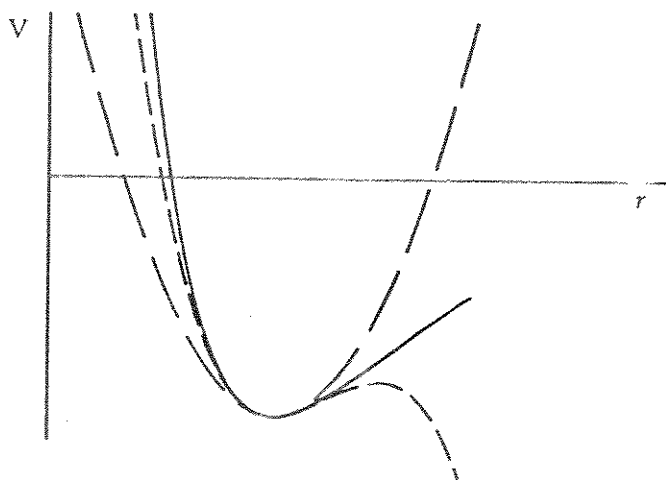


Fig. 25 Curvas mostrando o comportamento harmônico e anarmônico do estiramento de ligação: —, função de Morse; - - -, potencial harmônico; - · - · -, potencial harmônico com o termo cúbico adicionado.

A energia de deformação angular também é representada pela lei de Hooke, numa primeira aproximação:

$$V_{\theta} = \frac{1}{2} \sum k_{\theta} (\theta - \theta_0)^2 \quad \text{eq.21}$$

onde k_{θ} é a constante de força harmônica para a deformação angular, θ é o ângulo de ligação e θ_0 é algum ângulo de referência. Aqui, k_{θ} e θ_0 também são parâmetros ajustáveis.

A energia torsional é aproximada pela equação 22:

$$V_{tor} = k_{\omega} (1 - \cos 3 \omega) \quad \text{eq.22}$$

onde ω é o ângulo torsional.

Em espectroscopia vibracional, é comum escrever a função potencial torsional completa, conforme a equação 23:

$$V_{tor} = \sum_j \frac{1}{2} V_j [1 - \cos(j \omega)] \quad \text{eq.23}$$

onde j determina a periodicidade do movimento de torção e em mecânica molecular, inicialmente encontrou-se que o potencial para $j=3$ era suficiente para descrever todos os tipos de ângulos de torção para um determinado sistema em estudo, mas somente em 1976, foi reconhecido que o melhor nível de precisão era encontrado quando os dois primeiros termos da equação 23 eram também usados.

As interações entre átomos não ligados, comumente incluem as interações de van der Waals e de dipolos. As interações de dipolos são calculadas pela equação 24:

$$V_{\text{dipolos}} = \frac{\mu_i \mu_j}{D \cdot r_{ij}^3} (\cos \chi - 3 \cos \alpha_i \cos \alpha_j) \quad \text{eq.24}$$

onde D é a constante dielétrica, χ é o ângulo entre os dipolos μ_i e μ_j e α são os ângulos que os dipolos formam com o vetor conectando os dois dipolos (fig.26):

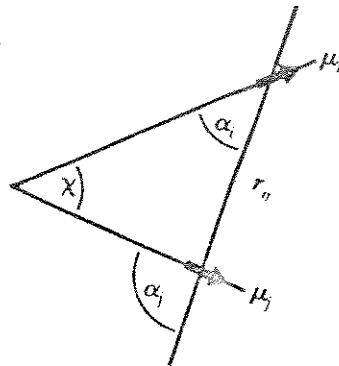


Fig.26 Definição dos termos geométricos que aparecem na equação 24.

As interações de van der Waals são calculadas pela equação 25:

$$VDW = 2,90 \cdot 10^5 \epsilon e^{-12,50 r_0 / r} - 2,25 \epsilon \left(\frac{r_0}{r}\right)^6 \quad \text{eq.25}$$

onde $\epsilon = \sqrt{\epsilon_i \cdot \epsilon_k}$; ϵ_i e ϵ_k são as "durezas" dos átomos (kcal/mol) $r_0 = r_i + r_k$ é raio combinado de van der Waals dos átomos "i" e "k" e "r" é a distância interatômica efetiva.

A soma de todos os termos do campo de força é chamada energia estérica da molécula e para que o campo de força reproduza bem os dados experimentais, é necessário otimizar esses parâmetros ou introduzir outros. Algumas técnicas experimentais usadas para determinar estruturas são: difração de raio-x para cristais e difração de elétrons e espectroscopia de microondas para gases.

Em mecânica molecular, parâmetros e constantes de força podem ser transferidos de uma molécula para outra; assim, é comum otimizarmos os parâmetros para um conjunto de moléculas simples, usando dados experimentais de estrutura para comparação, e posteriormente, usar os mesmos parâmetros para moléculas mais complexas e semelhantes às primeiras.

O primeiro campo de força desenvolvido em mecânica molecular foi o MM1(1973)⁴⁶ e em 1977 surgiu a segunda geração de mecânica molecular - MM2 (mecânica molecular, programa 2)⁴⁷, o qual aperfeiçoou consideravelmente o campo de força de 1973 (MM1), devido principalmente a introdução do termo torsional V2.

O campo de força de 1977 de MM2 ainda necessitava de algum refinamento, pois notou-se falhas na reprodução de barreiras de rotação internas nas ligações simples C-C. MM2 também possui certa tendência em superestimar as distâncias H...H. Alguns parâmetros torsionais aparentemente são muito pequenos e os parâmetros de van der Waals para as interações H...H, muito grandes. Assim, MM2 foi reparametrizado, surgindo um novo campo de força: MM2' (ou MM2 modificado)⁸. Parâmetros torsionais e potenciais de átomos não ligados foram reajustados.

Em 1987, um novo campo de força de MM2 surgiu com novos ajustes no campo de força de 1977⁴⁸. Ligações hidrogênio são consideradas apenas qualitativamente nas 1 versões do MM2, mas não quantitativamente. Experimentalmente, as ligações hidrogênio são mais curtas e fortes do que aquelas calculadas pelas 1 versões do MM2. O novo campo de força de MM2(87), corrige este fato pela redução dos raios de van der Waals dos hidrogênios envolvidos nas ligações hidrogênio e pelo incremento na atração de Van der Waals entre o hidrogênio e os vários átomos eletronegativos.

Para o conjunto de moléculas deste trabalho, usamos o campo de força do MM2 (87), tendo-se incluído alguns parâmetros do campo de força MM2'. Os parâmetros foram obtidos pelo professor Hase (IQ - UNICAMP) através de carta ao autor do método, tendo-nos fornecido posteriormente.

As geometrias dos compostos inativos I a VI e IX a XI (tab.1 - cap.I) e dos compostos ativos XII,XIII e XVII a XX (tab.2) foram otimizadas com o método MM2(87).

Para as moléculas com oxigênio na ponte, as constantes torsionais referentes aos átomos 13-12-6-1 e 16-12-6-1 foram parametrizadas (figura 27):

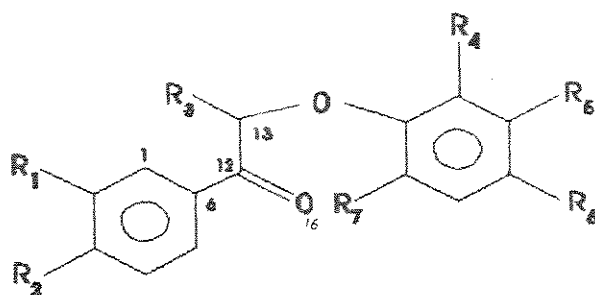


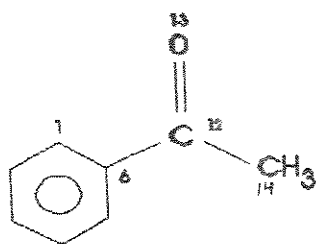
Fig.27 Estrutura básica dos compostos estudados neste trabalho, com oxigênio na "ponte". Os substituintes estão representados por "R".

A parametrização foi feita para as moléculas mais simples, acetofenona e benzaldeído (fig.28) da seguinte maneira:

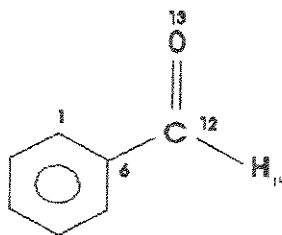
1- Obteve-se da literatura⁴⁹ a barreira de rotação entre as formas planar e ortogonal para os dois compostos, do substituinte em relação ao anel aromático.

2- Realizou-se o cálculo com MM2 para esses compostos, usando valores arbitrários para as constantes torsionais v_1 , v_2 e v_3 referentes aos átomos 13-12-6-1 e 14-12-6-1, para o composto planar e depois ortogonal, até que a barreira de rotação fosse reproduzida.

3- As constantes assim parametrizadas, foram usadas nas moléculas mais complexas citadas. Obteve-se: $v_1 = 0$, $v_2 = 2,0$ e $v_3 = 0$.



a) Acetofenona



b) Benzaldeído

Fig.28 As constantes torsionais referentes aos átomos 13-12-6-1 e 14-12-6-1 foram parametrizadas.

Para as moléculas com "S" na ponte, os parâmetros em falta foram cedidos pelo professor Hase.

Os compostos VI, VII e XIV (tab.1 e 2 do capítulo I) tiveram as geometrias otimizadas com o método MM2' e novamente os parâmetros necessários foram cedidos pelo professor Hase.

Para as moléculas XV e XVI (tab. 2 do cap. I), as constantes torsionais referentes aos átomos 16-13-12-6 e 14-13-12-6 foram parametrizadas (figura 29).

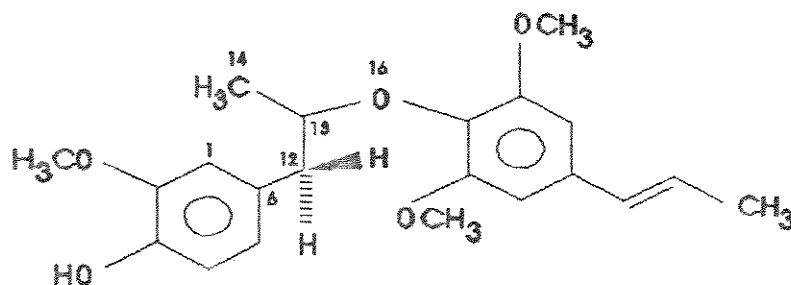


Fig.29 As constantes torsionais referentes aos átomos 16-13-12-6 e 14-13-12-6 para a molécula XV, foram parametrizadas e usadas posteriormente na molécula XVI.

Inicialmente, realizou-se a parametrização das constantes citadas, usando-se dados da literatura sobre ângulos diedros para a molécula mais simples da figura 30. Ângulos referentes aos átomos 16-13-12-6 (T1) e 17-16-13-12 (T2) foram encontrados, de cálculos feitos com o método de mecânica molecular (MMPM1)⁵⁰.

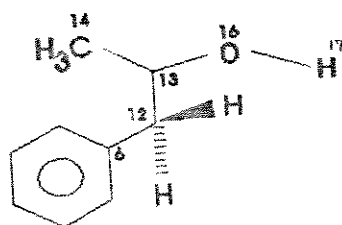


Fig.30 Ângulos diedros referentes aos átomos 16-13-12-6 e 17-16-13-12 foram encontrados na literatura e utilizados para parametrizar constantes torsionais para a molécula XV deste trabalho.

Na referência citada, os seguintes valores foram testados para os ângulos diedros T1 e T2:

Conformacao	Angulos diedros (graus)	
	T1	T2
a	180	180
b	180	60
c	-60	60
d	-60	180
e	-60	-60

Tabela 16 - Valores testados para os ângulos diedros T1 e T2, na referência 50.

Inicialmente, atribuímos valores iguais a zero para as constantes torsionais v_1 , v_2 e v_3 referentes aos átomos 16-13-12-6 e 17-16-13-12 da molécula da figura 30 e calculamos as energias para as conformações a, b, c, d, e. A seguinte ordem de estabilidade foi obtida: $a > d > b > c > e$, a qual não variou muito daquela encontrada na literatura: $a > d > c > b > e$. Para os dois casos, o mínimo mais estável foi o mesmo.

Assim, usando os valores testados das constantes torsionais, obteve-se os mínimos locais para a molécula da figura 31, tendo-se variado os ângulos diedros referentes aos átomos 14-13-12-6 e 19-18-13-12 em intervalos de 15 graus. Após essa etapa, a otimização da geometria da molécula com os substituintes (molécula XV) foi feita da mesma forma que para as outras moléculas, como está explicado no item II.2.

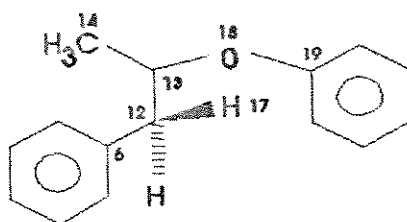


Fig.31 A otimização de geometria da molécula XV foi feita, inicialmente, para o composto sem os substituintes.

APENDICE 2 - O METODO MNDO-PM3

Iniciaremos pelo método MNDO⁵¹ (Modified Neglect of Diatomic Overlap), do qual deriva o método MNDO-PM3^{39,52} (Modified Neglect of Diatomic Overlap - Parametric Method 3). O método AM1⁵³ (Austin Model 1) foi a primeira reparametrização do MNDO e também será explicado aqui, uma vez que os resultados obtidos com ambos os métodos (MNDO e AM1) foram usados para comparação no desenvolvimento do MNDO-PM3.

As aproximações feitas pelo método semi-empírico MNDO à equação de Hartree-Fock-Roothan referem-se à moléculas de camada fechada, considerando-se apenas os elétrons de valência e com primeira aproximação, considera-se os elétrons das camadas internas como parte do núcleo (aproximação do caroço).

No método Hartree-Fock-Roothan⁵⁴, os orbitais moleculares da camada de valência (ψ_i) são representados por combinações lineares de um conjunto de funções de base referentes aos orbitais atômicos (ϕ_ν)

$$\psi_i = \sum_{\nu} C_{\nu i} \phi_{\nu} \quad \text{eq.26}$$

onde os coeficientes $C_{\nu i}$ podem ser obtidos resolvendo a equação de Hartree-Fock-Roothan:

$$\sum_{\nu} C_{\nu i} [F_{\mu\nu} - \epsilon_i S_{\mu\nu}] = 0 \quad \text{eq.27}$$

onde ϵ_i é o autovalor do orbital molecular ψ_i , $F_{\mu\nu}$ é a matriz de Fock e $S_{\mu\nu}$ são as integrais de sobreposição, quando as funções de base ϕ_ν não são ortogonais:

$$S_{\mu\nu} = \int \phi_{\mu}^* \phi_{\nu} dV \quad \text{eq.28}$$

O método MNDO parte da equação de Hartree-Fock-Roothan, já aproximada pelo método semi-empírico NDDO⁵⁵, da seguinte maneira:

$$\sum_{\nu} C_{\nu i} (F_{\mu\nu} - \epsilon_i \delta_{\mu\nu}) = 0 \quad \text{eq.29}$$

onde $\delta_{\mu\nu}$ é o delta de Kronecker.

Os elementos $F_{\mu\nu}$ da matriz de Fock são a soma do Hamiltoniano de um elétron, $H_{\mu\nu}$ (aproximação do caroço) e da parte de dois elétrons, $G_{\mu\nu}$; a energia eletrônica total, E_{el} , é dada por:

$$E_{el} = 1/2 \sum_{\mu} \sum_{\nu} P_{\mu\nu} (H_{\mu\nu} + F_{\mu\nu}) \quad \text{eq.30}$$

onde $P_{\mu\nu}$ é um elemento da matriz de ordem de ligação e $\sum_{\mu} \sum_{\nu} P_{\mu\nu}$ fornece o número total de elétrons sobre determinado átomo.

Assumindo que os orbitais atômicos ϕ_{μ} e ϕ_{ν} estão centrados no átomo A e ϕ_{λ} e ϕ_{σ} no átomo B ($A \neq B$), os elementos da matriz de Fock NDDO são:

$$F_{\mu\mu} = U_{\mu\mu} + \sum_B V_{\mu\mu,B} + \sum_{\nu}^A P_{\nu\nu} [(\mu\mu, \nu\nu) - 1/2(\mu\nu, \mu\nu)] + \sum_B \sum_{\lambda, \sigma}^B P_{\lambda\sigma} (\mu\mu, \lambda\sigma) \quad \text{eq.31}$$

$$F_{\mu\nu} = \sum_B V_{\mu\nu,B} + 1/2 P_{\mu\nu} [3(\mu\nu, \mu\nu) - (\mu\mu, \nu\nu)] + \sum_B \sum_{\lambda, \sigma}^B P_{\lambda\sigma} (\mu\nu, \lambda\sigma) \quad \text{eq.32}$$

$$F_{\mu\lambda} = \beta_{\mu\lambda} - 1/2 \sum_{\nu}^A \sum_{\sigma}^B P_{\nu\sigma} (\mu\nu, \lambda\sigma) \quad \text{eq.33}$$

onde:

- $U_{\mu\mu}$ é a energia de um elétron (um centro) e representa a soma da energia cinética de um elétron no orbital atômico ϕ_{μ} no átomo A e sua energia potencial devido à atração pelo caroço do átomo A.

- $(\mu\mu, \nu\nu) = g_{\mu\nu}$ são as integrais de repulsão de dois elétrons em um centro.
- $(\mu\nu, \mu\nu) = h_{\mu\nu}$ são as integrais de troca
- $\beta_{\mu\lambda}$ são as integrais de ressonância do caroço (dois centros e um elétron).
- $(\mu\nu, \lambda\sigma)$ são as integrais de dois centros de repulsão entre dois elétrons.

A energia total, E_{tot}^{mol} , da molécula é a soma da energia eletrônica E_{el} e das repulsões E_{AB}^{caroco} entre os caroços dos átomos A e B:

$$E_{tot}^{mol} = E_{el} + \sum_A \sum_B E_{AB}^{caroco} \quad \text{eq.34}$$

Os termos de um centro, $g_{\mu\nu}$ e $h_{\mu\nu}$ são obtidos através do método de Oleari⁵⁶ e $U_{\mu\mu}$ é tratado como um termo variável em MNDO. Os valores de $U_{\mu\mu}$ otimizados em MNDO são dados na referência. No método de Oleari, as energias teóricas dos estados de valência dos átomos e seus íons devem corresponder à valores espectroscópicos.

As integrais de dois centros $(\mu\nu, \lambda\sigma)$ representam a energia de interação entre as distribuições de carga $e\phi_{\mu}\phi_{\nu}$ no átomo A e $e\phi_{\lambda}\phi_{\sigma}$ no átomo B, onde e é a carga elementar.

Classicamente, essas integrais são iguais à soma sobre todas as interações entre os momentos de multipolo M_{lm} das duas distribuições de carga, sendo que os subscriptos l e m especificam a ordem e orientação dos multipolos. Baseando neste conceito clássico, as integrais de repulsão de dois centros são expandidas em termos de interações entre multipolos $[M_{l_1 m_1}^A, M_{l_2 m_2}^B]$

$$(\mu\nu, \lambda\sigma) = \sum_{l_1} \sum_{l_2} \sum_m [M_{l_1 m_1}^A, M_{l_2 m_2}^B] \quad \text{eq.35}$$

As interações semi-empíricas entre multipolos mostram o comportamento nos limites $R_{AB} \rightarrow \infty$ e $R_{AB} \rightarrow 0$ (distância interatômica R_{AB}). Para $R_{AB} \rightarrow \infty$, as interações devem convergir para valores clássicos de interação e para $R_{AB} \rightarrow 0$, devem reproduzir valores semi-empíricos para as correspondentes integrais de repulsão de um centro.

Cada multipolo M_{lm} é representado por uma configuração apropriada $[M_{lm}]$ de 2^l cargas pontuais de magnitude $e/2^l$ com separação de cargas D_l . A interação entre dois multipolos $[M_{l_1 m_1}^A, M_{l_2 m_2}^B]$ é então, calculada pela aplicação de uma fórmula semi-empírica apropriada para cada interação entre as cargas nas duas configurações e pelo somatório sobre todas essas interações. Assim, chamando a distância entre as cargas i e j nos átomos A e B, de R_{ij} , obtemos:

$$[M_{l_1 m_1}^A, M_{l_2 m_2}^B] = \frac{e^2}{2^{l_1+l_2}} \sum_{i=1}^{2^{l_1}} \sum_{j=1}^{2^{l_2}} f_1(R_{ij}) \quad \text{eq.36}$$

onde $f_1(R_{ij})$ é substituído por qualquer expressão semi-empírica com o comportamento apropriado nos limites $R_{AB} \rightarrow \infty$ e $R_{AB} \rightarrow 0$. Para $f_1(R_{ij})$ foram usadas as fórmulas baseadas nas aproximações de Dewar-Sabelli-Klopman (DSK)⁵⁷⁻⁵⁸:

$$f_1(R_{ij}) = [R_{ij}^2 + (\rho_{l_1^A} + \rho_{l_2^B})^2]^{-1/2} \quad \text{eq.37}$$

Os valores para ρ_1, ρ_2 são calculados por métodos semi-empíricos e são dados na referência 51.

As atrações caroço-elétron, $V_{\mu\nu, B}$ e repulsões E_{AB}^{caroco} são dadas pelas seguintes equações, em MNDO:

$$V_{\mu\nu, B} = -Z_B (\mu^A \nu^A, S^B S^B) + f_2(R_{AB}) \quad \text{eq.38}$$

$$E_{AB}^{caroco} = Z_A Z_B (S^A S^A, S^B S^B) + f_3(R_{AB}) \quad \text{eq.39}$$

onde $f_2(R_{AB})$ e $f_3(R_{AB})$ incluem um termo exponencial de forma $\exp(-\alpha R_{AB})$ com α ajustável, tal

que a repulsão entre átomos neutros se anule no limite $R_{AB} \rightarrow \infty$

O último termo aproximado em MNDO é a integral de ressonância $\beta_{\mu\lambda}$ que fornece a principal contribuição á energia de ligação da molécula. Como nos métodos semi-empíricos mais antigos, assume-se em MNDO, que $\beta_{\mu\lambda}$ é proporcional às integrais de recobrimento $S_{\mu\lambda}$

$$\beta_{\mu\lambda} = f_4(R_{AB})S_{\mu\lambda} \quad \text{eq.40}$$

As integrais de recobrimento são determinadas analiticamente, sendo que os expoentes das funções orbitais são parâmetros ajustáveis e são dados na referência 51.

As funções $f_4(R_{AB})$ foram investigadas de forma que as expressões referentes contêm um parâmetro ajustável β

O procedimento de parametrização das funções f_1-f_4 é dado na referência 51, bem como os valores para todos os parâmetros otimizados pelo método MNDO.

O método MNDO foi parametrizado inicialmente para H, C, N e O e apresentou alguns erros como a tendência em superestimar repulsões. Dewar et al. modificaram então, a função de repulsão do caroço, E_{AB}^{caroco} , na tentativa de resolver os erros apresentados no método MNDO. Assim surgiu o método AM1, onde E_{AB}^{caroco} é dado por:

$$E_{AB}^{caroco} = Z_A Z_B (S^A S^A, S^B S^B) + f_3(R_{AB}) + (Z_i Z_j / R_{ij}) \left\{ \sum_i K_{Ai} \exp[L_{Ai} (R_{AB} - M_{Ai})^2] + \sum_j K_{Bj} \exp[L_{Bj} (R_{AB} - M_{Bj})^2] \right\}$$

onde apenas o terceiro termo refere-se à modificação feita pelo AM1, sendo que o primeiro e segundo termos são do MNDO.

Os parâmetros M, K e L são dados na referência 53, sendo que K e M são otimizáveis.

A forma das equações em AM1 é a mesma do método MNDO, só diferindo no termo de repulsão, E_{AB}^{caroco} . Com esse novo método, foi feita uma reparametrização do primeiro (MNDO) para os seguintes elementos: H, C, N, O, F, I, Si, Cl e Br., sendo que o MNDO já havia sido parametrizado para: H, C, N, O, F, P, S, I, Al, Si, Cl e Br.

O método MNDO-PM3 consiste da terceira parametrização, usando o mesmo formalismo, isto é, as mesmas equações do método AM1.

Resultados obtidos com o novo conjunto de parâmetros foram comparados com aqueles obtidos com os métodos anteriores, MNDO e AM1, e observou-se que as diferenças entre as propriedades calculadas pelo novo método e observadas eram menores do que para os métodos MNDO e AM1³⁹. O método desta nova parametrização é dado na referência 52.

APENDICE 3 - METODOS DE RECONHECIMENTO DE PADRAO

ESCALONAMENTO ⁴²

O escalonamento consiste numa transformação dos dados (propriedades fisico-químicas), para que eles possam ser comparados numa mesma escala, apesar de possuírem unidades diferentes.

Neste trabalho, as propriedades fisico-químicas (variáveis) foram transformados tal que a média (eq.42) tornou-se zero e a variância, um (eq.43) ; esse tipo de transformação é chamado de "auto-escalonamento".

$$\bar{Y}_i = \sum_{k=1}^N Y_{ik} / N \quad \text{eq.42}$$

onde Y_i é a média da variável i calculada para os N objetos e Y_{ik} é a medida da variável i para o objeto k . A variância para a i -ésima variável é:

$$S_i^2 = \sum_{k=1}^N (Y_{ik} - \bar{Y}_i)^2 / (N-1) \quad \text{eq.43}$$

As propriedades calculadas para os objetos, Y_{ik} , foram transformados em Y'_{ik} , pela seguinte equação

$$Y'_{ik} = (Y_{ik} - \bar{Y}_i) / \bar{S}_i \quad \text{eq.44}$$

tal que:

$$\bar{Y}_i = \sum_{k=1}^N Y'_{ik} / N = 0 \quad \text{eq.45}$$

e

$$S_i^2 = \sum_{k=1}^N (Y'_{ik} - Y_i) / (N - 1) = 1 \quad \text{eq.46}$$

ANALISE DE COMPONENTES PRINCIPAIS ⁵⁹

Inicialmente, o conjunto de pontos representando os objetos (ou compostos), localizados no espaço pelos seus valores das propriedades fisico-químicas calculadas, é representado por seu ponto central, \bar{Y} definido por:

$$\bar{Y} = (\bar{Y}_1 \bar{Y}_2 \cdots \bar{Y}_k \cdots \bar{Y}_N)$$

onde \bar{Y}_k é a média da propriedade fisico-química k para o conjunto.

A matriz de dados Y , da figura 18 é então representada pela equação 47:

$$Y = J\bar{Y} + E^{(0)} \quad \text{eq.47}$$

onde E é a matriz contendo os desvios das médias.

A figura mostra a representação dos dados Y num espaço tridimensional e a representação do ponto central, \bar{Y}

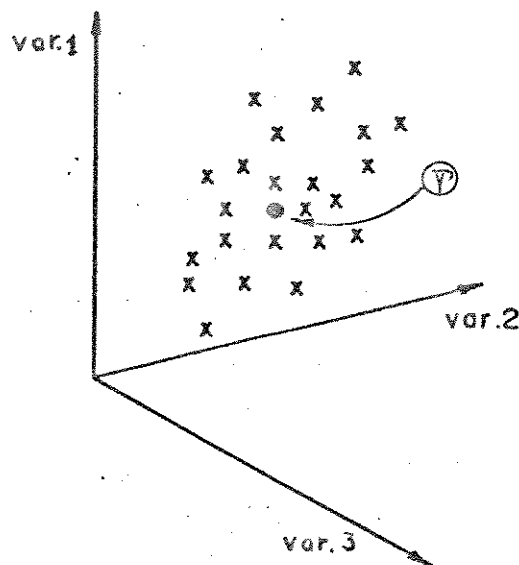


Figura 32 - Representação dos dados Y num espaço tridimensional, e do ponto central, \bar{Y}

Agora, uma reta é ajustada aos M pontos no espaço N, de modo que os desvios são os menores possíveis. Os coeficientes de direção desta reta são chamados "Loadings", sendo um para cada variável k, representados pelo vetor P_1

Quando cada ponto é projetado nesta linha, obtemos os "Scores" t_{i1} , isto é, a

coordenada do ponto i ao longo do eixo P_1

O valor $t_i P_1$ é subtraído de $\underline{E}^{(0)}$ para obter os novos desvios $\underline{E}^{(1)}$:

$$\underline{E}^{(0)} - t_i P_1 = \underline{Y} - \underline{1}\bar{Y} - t_i P_1 = \underline{E}^{(1)} \longrightarrow \underline{Y} = \underline{1}\bar{Y} + t_i P_1 + \underline{E}^{(1)} \quad \text{eq.48}$$

onde $|\underline{E}^{(1)}| < |\underline{E}^{(0)}|$

A equação 48 descreve o modelo de 1 componente principal que é representada por uma linha no espaço N dimensional e uma equação linear com uma variável, a componente principal (figura 33).

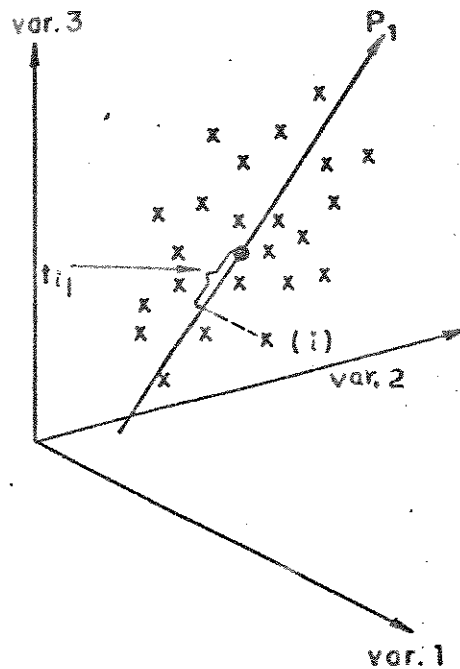


Fig. 33 - Representação da primeira componente principal, num espaço tridimensional.

Essa linha contém o máximo de variância dos dados, ou seja, o máximo de informação estatística.

Em seguida, subtrai-se o modelo de uma componente principal de \underline{Y} , o que implica na remoção da direção P_1 :

$$\underline{Y} - \underline{1}\bar{Y} - t_i P_1 = \underline{E}^{(1)} \quad \text{eq.49}$$

Então, ajusta-se uma outra reta aos pontos (perpendicular à primeira componente principal), de modo que os novos desvios são os menores possíveis. Esta reta representa a segunda componente principal:

$$\underline{y} = \underline{1} \cdot \underline{Y} + \underline{t}_1 \underline{P}_1 + \underline{t}_2 \underline{P}_2 + \underline{E}^{(2)} \quad \text{eq.50}$$

O modelo de duas componentes principais é representado por um plano no espaço N dimensional e uma equação linear com duas variáveis (as duas componentes principais). A figura 34 mostra a representação de duas componentes principais.

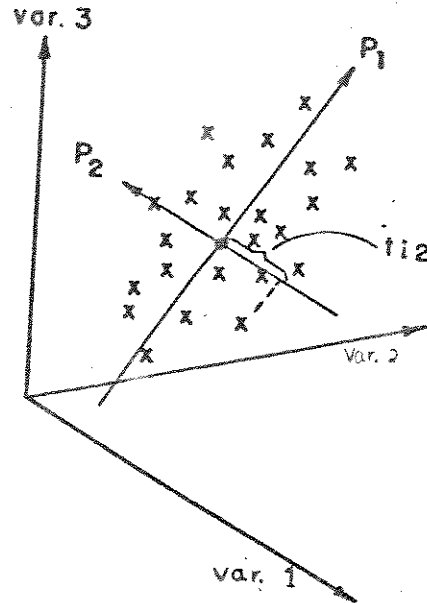


Fig.34 - Representação das duas primeiras componentes principais num espaço tridimensional.

Entre todas as retas perpendiculares à linha da primeira componente principal, a linha da segunda componente principal contém a maior variância possível.

Este procedimento é repetido N vezes e as últimas componentes principais terão variâncias próximas ou iguais a zero, podendo ser ignoradas.

SIMCA (Statistical Isolinear Multiple Component Analysis) ⁶⁶

O método SIMCA constrói um modelo de componentes principais para cada categoria do conjunto de treinamento; esses modelos podem ter números diferentes de componentes principais, conforme a figura 35.

O número de componentes principais pode ser determinado por "Cross Validation".

Uma vez definido o modelo que descreve cada categoria do conjunto de treinamento, o SIMCA pode classificar objetos de categorias desconhecidas (conjunto teste), permitindo ainda, a identificação de amostras que não pertençam às categorias definidas. A classificação do conjunto teste é feita pelo ajuste desses objetos, por regressão múltipla, aos parâmetros dos modelos calculados usando o conjunto de treinamento.

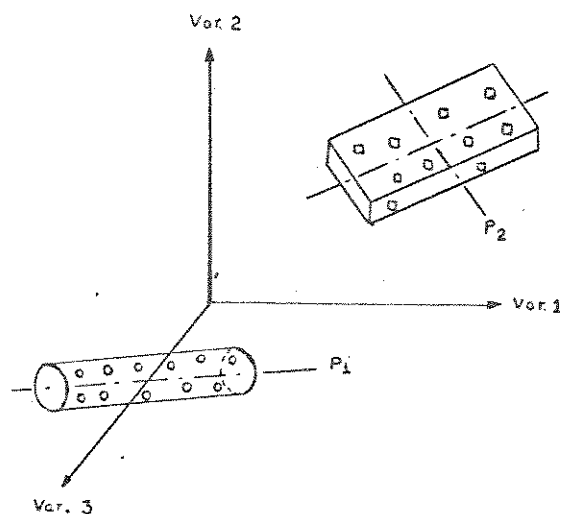
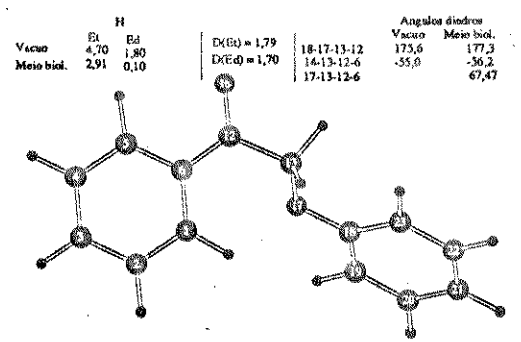
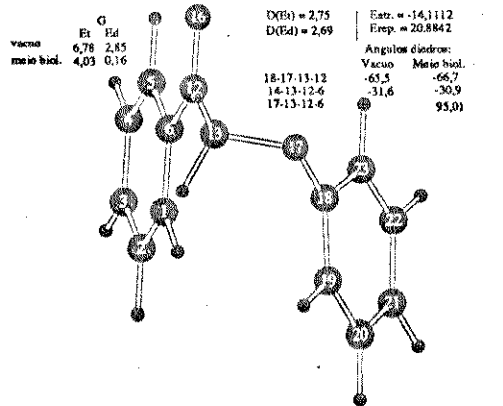
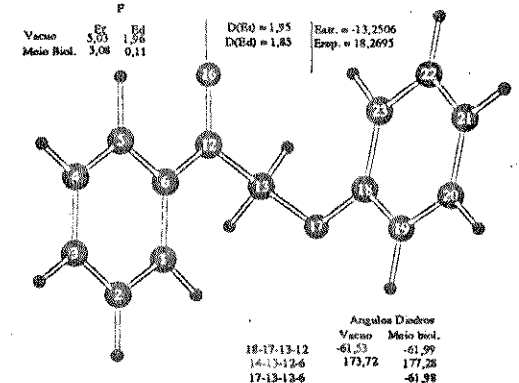
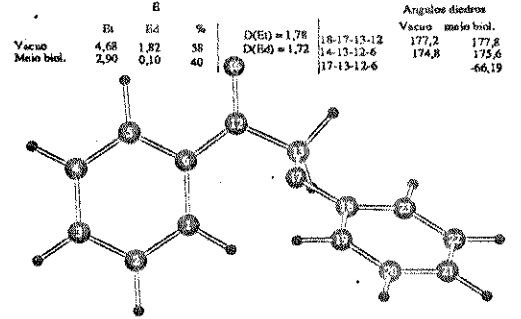
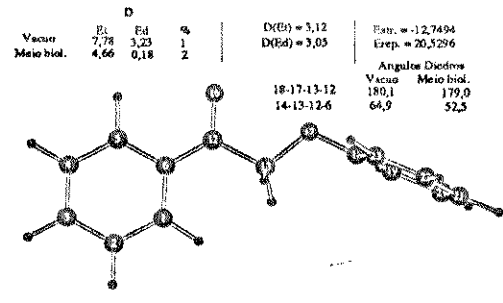
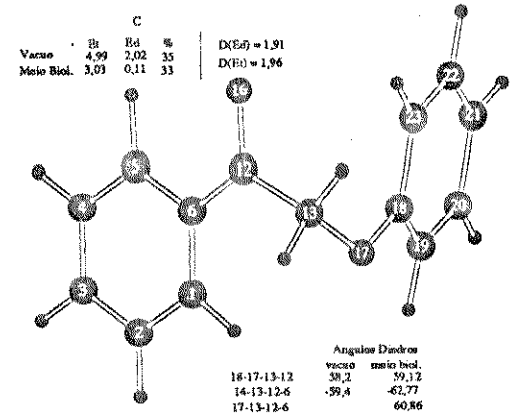
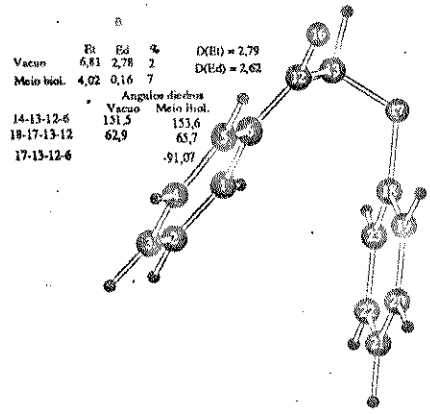
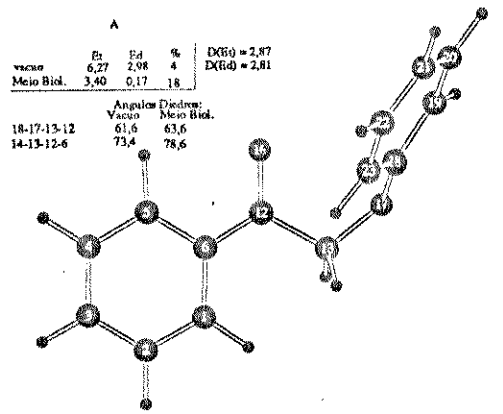


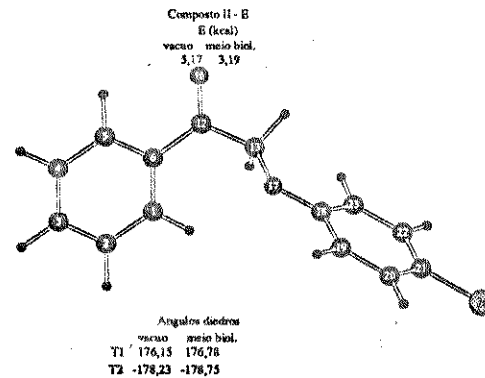
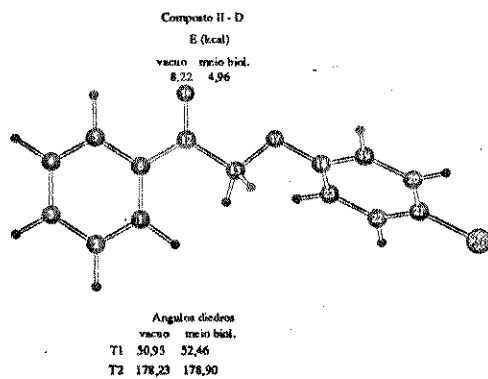
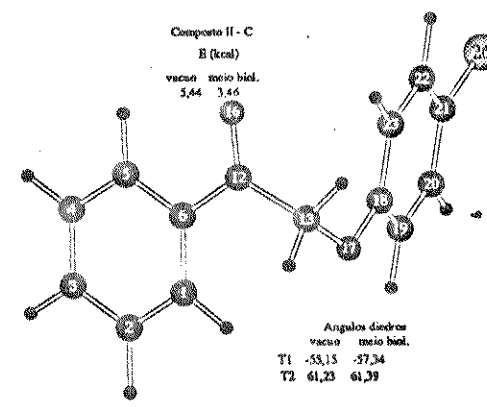
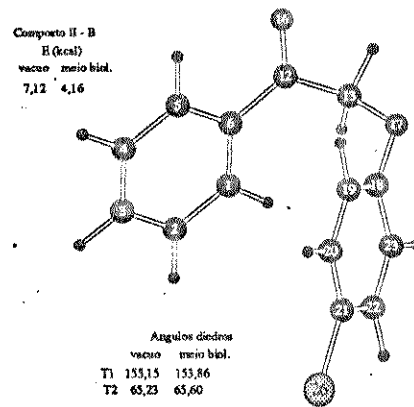
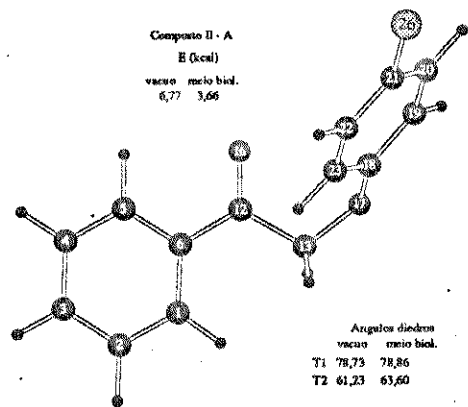
Fig.35 - Representação de modelos com números diferentes de componentes principais, para cada categoria.

AFENDICE 4

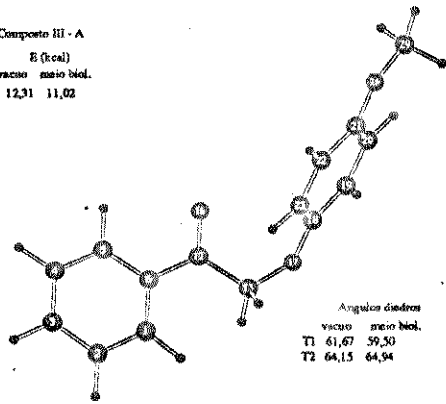
MINIMOS LOCAIS ENCONTRADOS PARA OS COMPOSTOS ATIVOS E INATIVOS



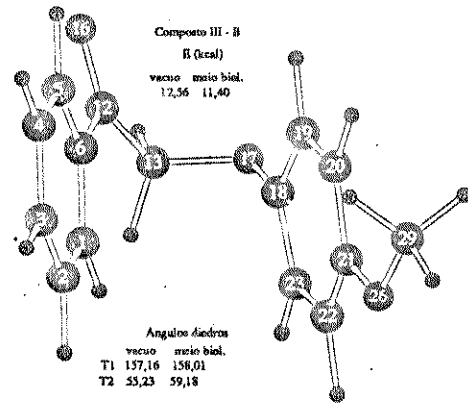
composto I



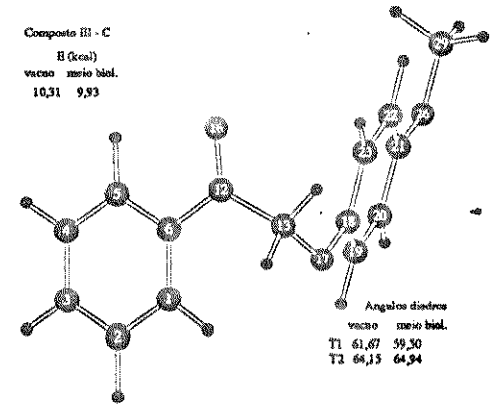
Composto III - A
 E (kcal)
 vacuo meio biol.
 12,31 11,02



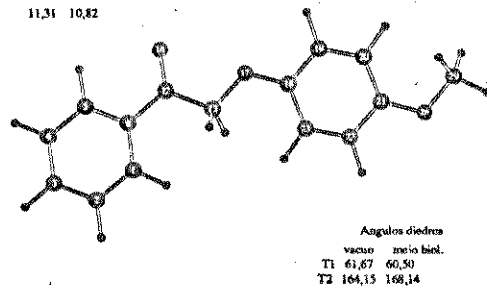
Composto III - B
 E (kcal)
 vacuo meio biol.
 12,56 11,40



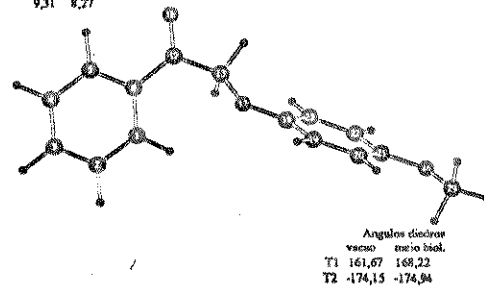
Composto III - C
 E (kcal)
 vacuo meio biol.
 10,31 9,93

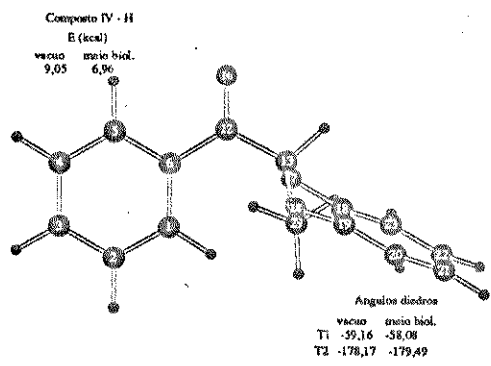
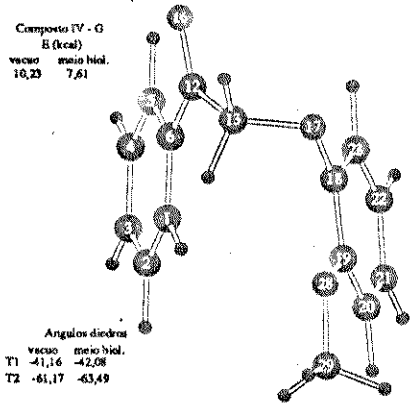
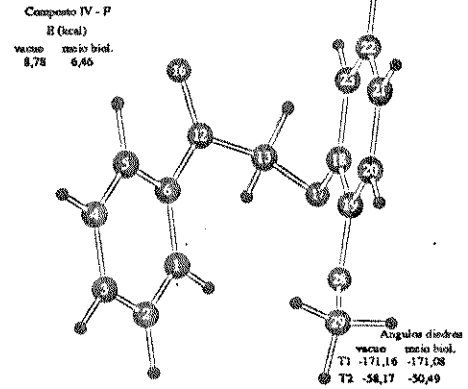
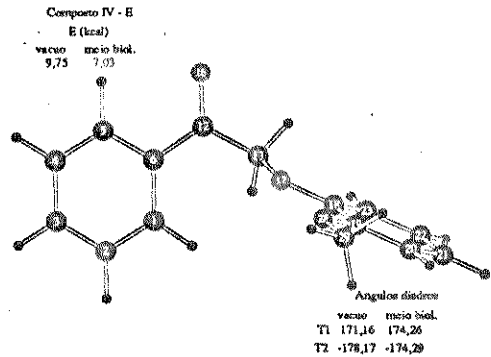
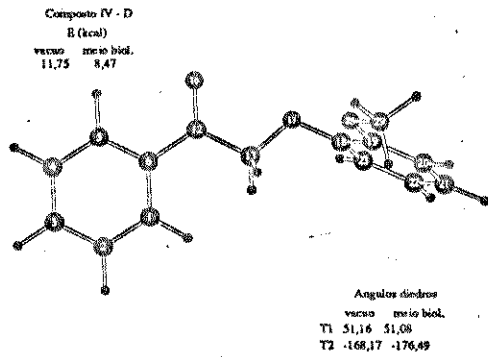
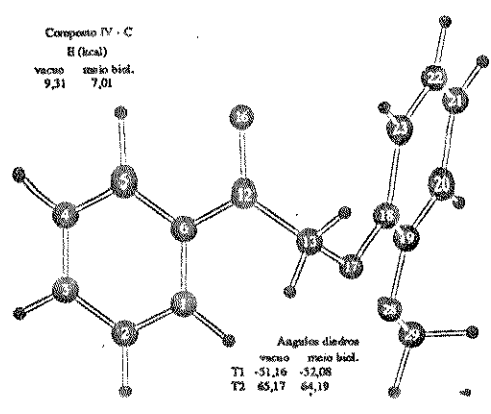
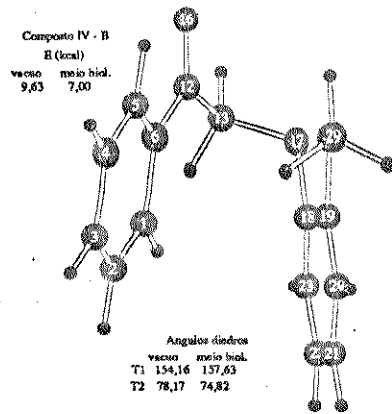
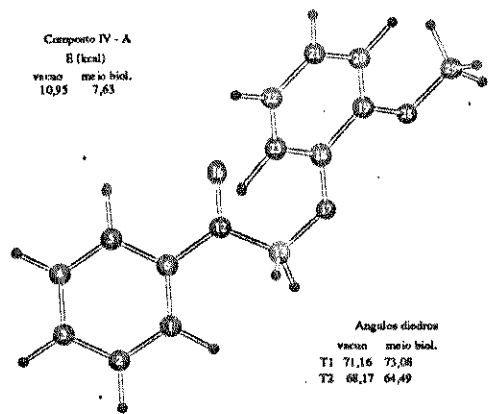


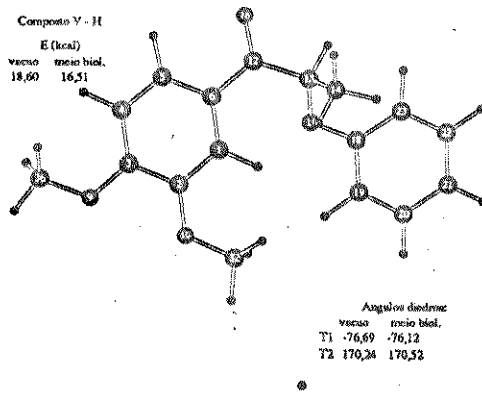
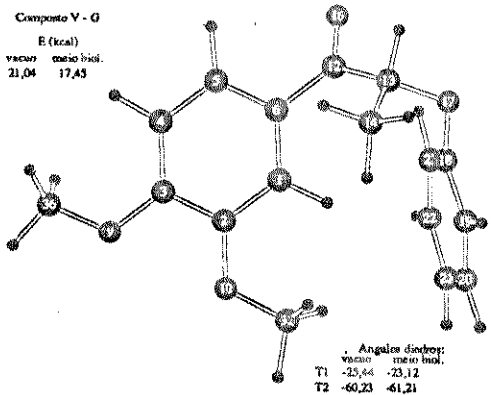
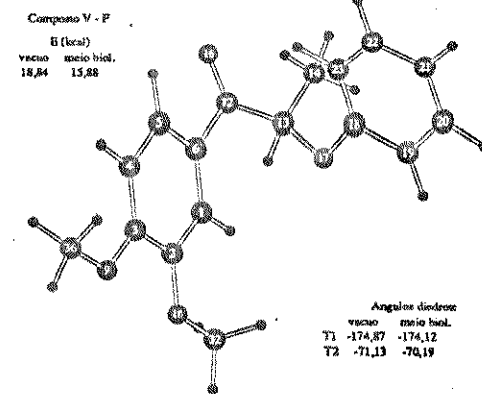
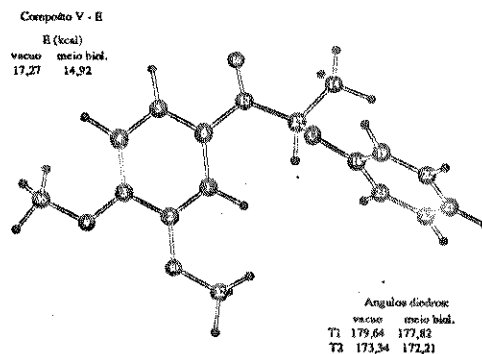
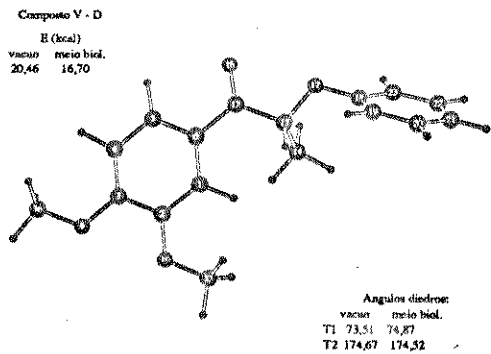
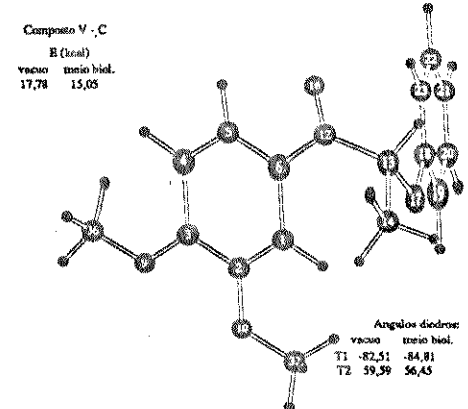
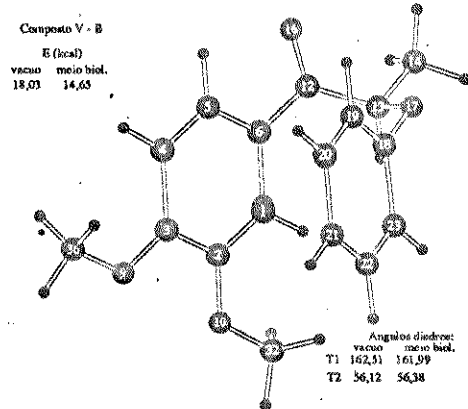
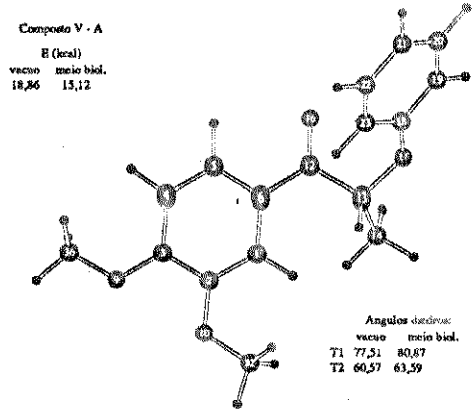
Composto III - D
 E (kcal)
 vacuo meio biol.
 11,31 10,82



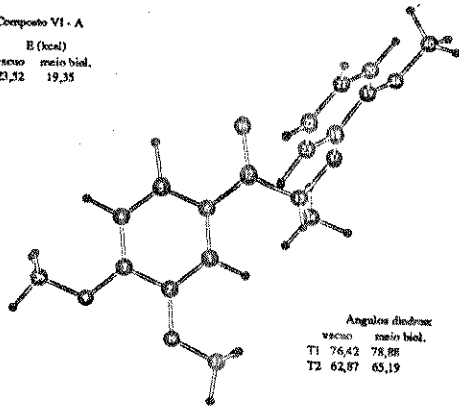
Composto III - E
 E (kcal)
 vacuo meio biol.
 9,31 8,27



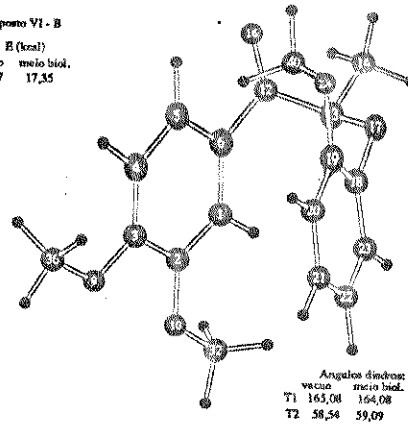




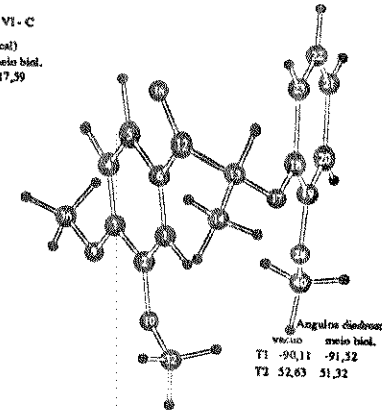
Composto VI - A
 E (kcal)
 vacuo meio biol.
 23,32 19,35



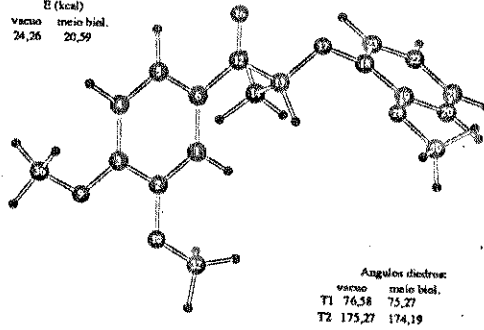
Composto VI - B
 E (kcal)
 vacuo meio biol.
 20,87 17,35



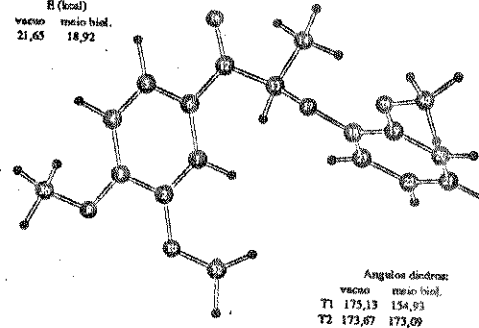
Composto VI - C
 E (kcal)
 vacuo meio biol.
 20,32 17,59



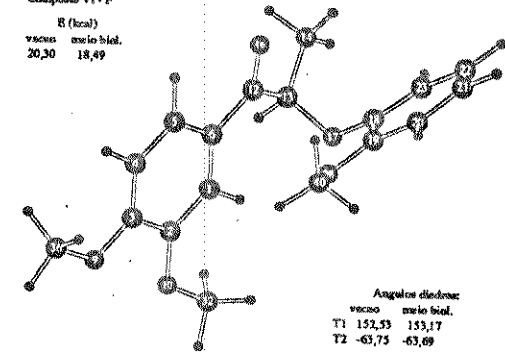
Composto VI - D
 E (kcal)
 vacuo meio biol.
 24,26 20,59



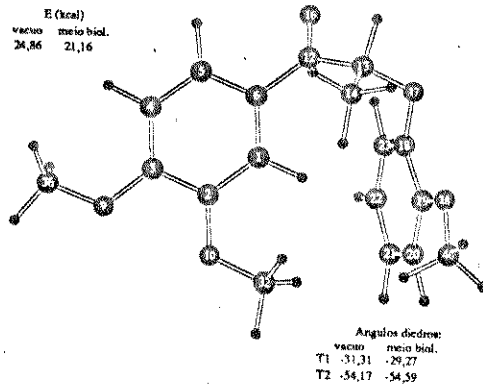
Composto VI - E
 E (kcal)
 vacuo meio biol.
 21,65 18,92



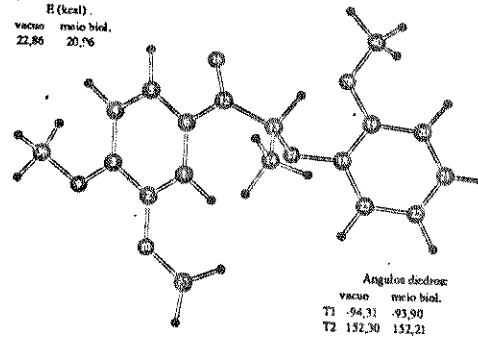
Composto VI - F
 E (kcal)
 vacuo meio biol.
 20,30 18,49



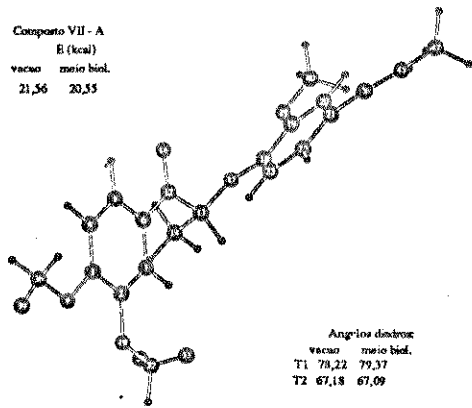
Composto VI - G
 E (kcal)
 vacuo meio biol.
 24,86 21,16



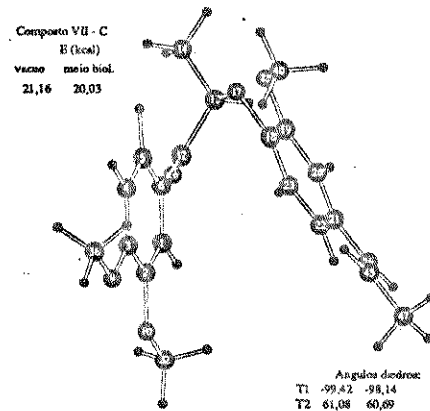
Composto VI - H
 E (kcal)
 vacuo meio biol.
 22,86 20,06



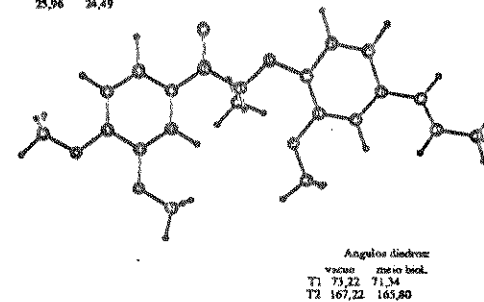
Composto VII - A
 E (kcal)
 vacuo meio biol.
 21,56 20,55



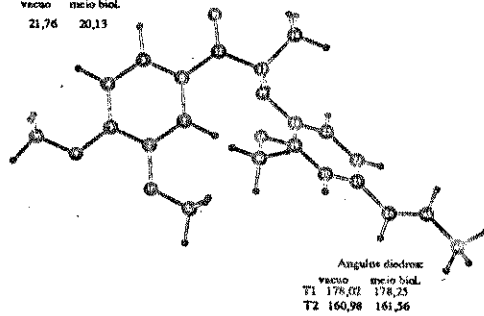
Composto VII - C
 E (kcal)
 vacuo meio biol.
 21,16 20,03



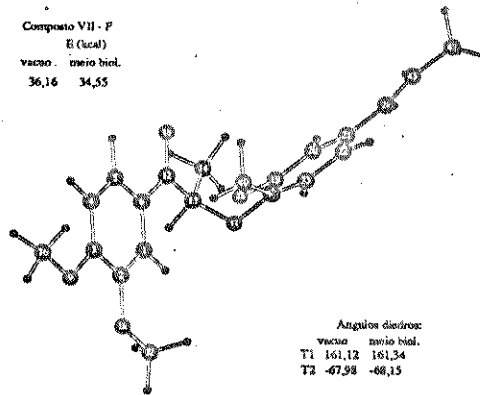
Composto VII - D
 E (kcal)
 vacuo meio biol.
 25,96 24,49



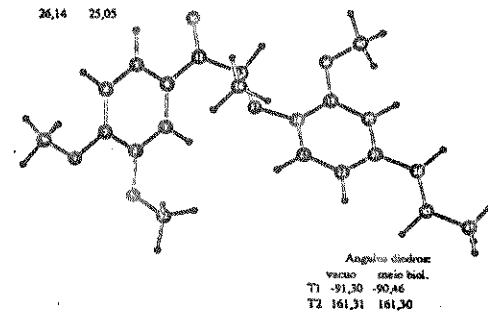
Composto VII - E
 E (kcal)
 vacuo meio biol.
 21,76 20,13



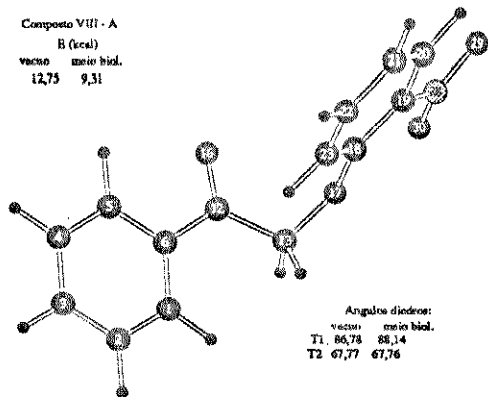
Composto VII - F
 E (kcal)
 vacuo meio biol.
 36,16 34,53



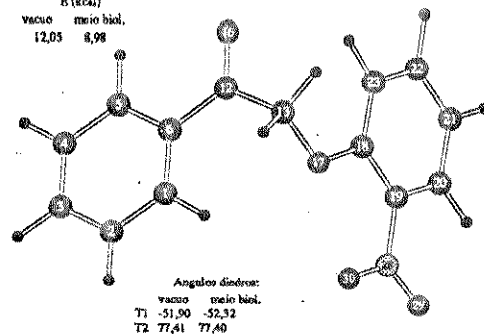
Composto VII - H
 E (kcal)
 vacuo meio biol.
 26,14 25,05



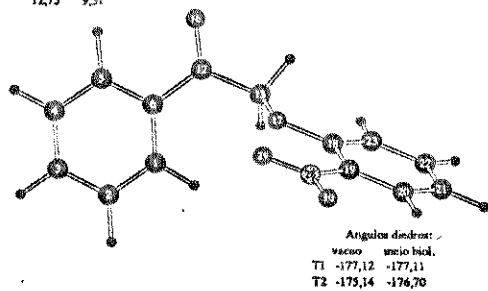
Composto VIII - A
E (kcal)
vacuo meio biol.
12,75 9,31



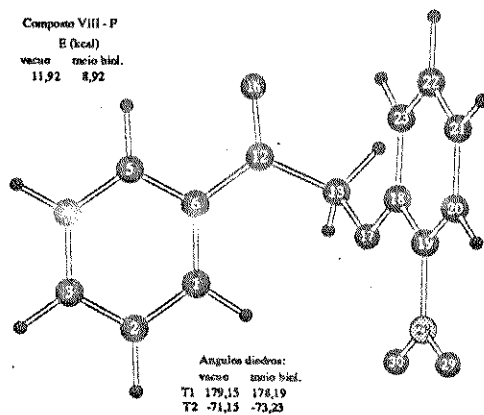
Composto VIII - C
E (kcal)
vacuo meio biol.
12,03 8,98



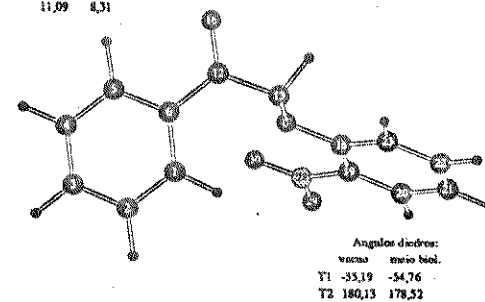
Composto VIII - E
E (kcal)
vacuo meio biol.
12,75 9,31



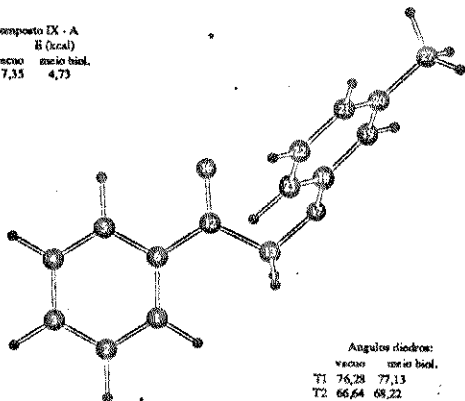
Composto VIII - P
E (kcal)
vacuo meio biol.
11,92 8,92



Composto VIII - H
E (kcal)
vacuo meio biol.
11,09 8,31

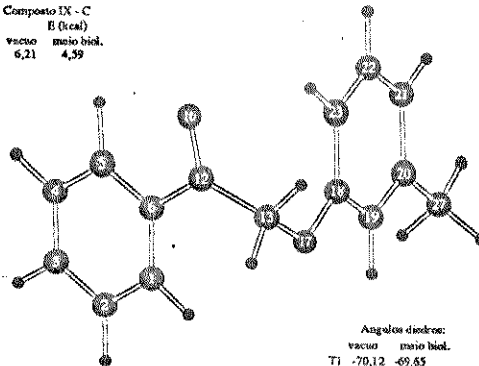


Composto IX - A
 H (kcal)
 vacuo meio biol.
 7,35 4,73



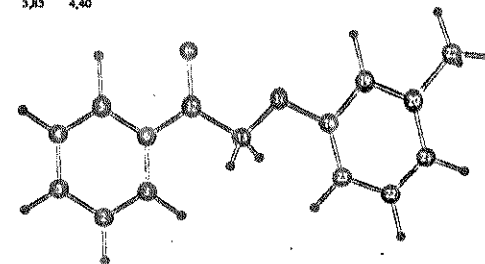
Ângulos diedros:
 vacuo meio biol.
 T1 76,28 77,13
 T2 66,64 68,22
 T3 -164,69

Composto IX - C
 H (kcal)
 vacuo meio biol.
 6,21 4,59



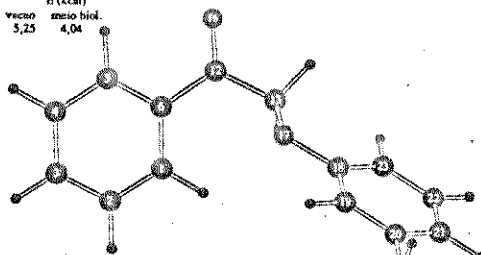
Ângulos diedros:
 vacuo meio biol.
 T1 -70,12 -69,65
 T2 68,54 67,25
 T3 49,09

Composto IX - D
 H (kcal)
 vacuo meio biol.
 5,83 4,40



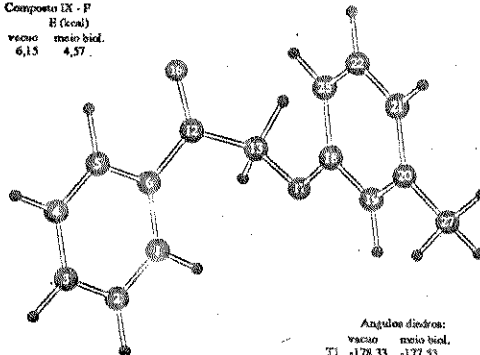
Ângulos diedros:
 vacuo meio biol.
 T1 61,28 60,79
 T2 178,19 179,66

Composto IX - E
 H (kcal)
 vacuo meio biol.
 5,25 4,04



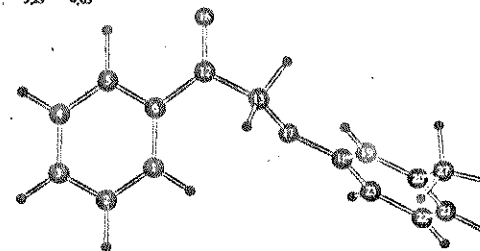
Ângulos diedros:
 vacuo meio biol.
 T1 177,44 -179,97
 T2 178,64 178,23
 T3 -60,40

Composto IX - F
 H (kcal)
 vacuo meio biol.
 6,15 4,57



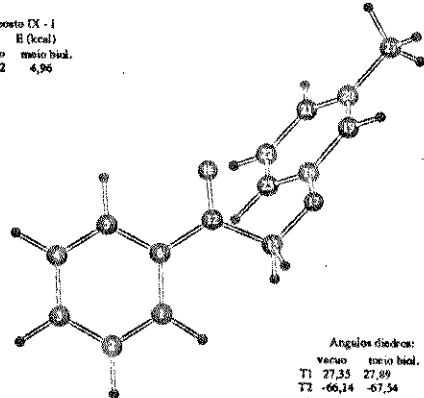
Ângulos diedros:
 vacuo meio biol.
 T1 -178,33 -177,53
 T2 -69,64 -70,57
 T3 -53,39

Composto IX - H
 H (kcal)
 vacuo meio biol.
 5,29 4,05



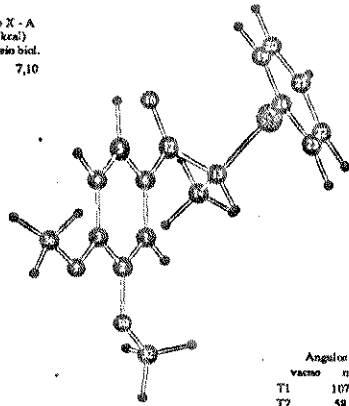
Ângulos diedros:
 vacuo meio biol.
 T1 -52,32 -53,49
 T2 177,12 176,98
 T3 68,59

Composto IX - I
 H (kcal)
 vacuo meio biol.
 7,52 4,96



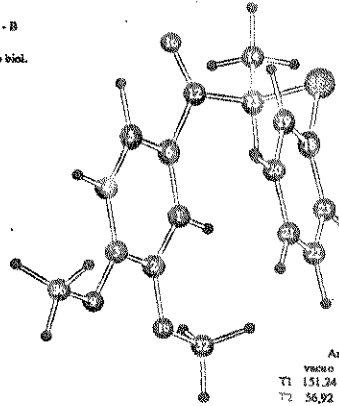
Ângulos diedros:
 vacuo meio biol.
 T1 27,35 27,89
 T2 -66,14 -67,54
 T3 150,56

Composto X - A
E (kcal)
vacuo meio biol.
7,10



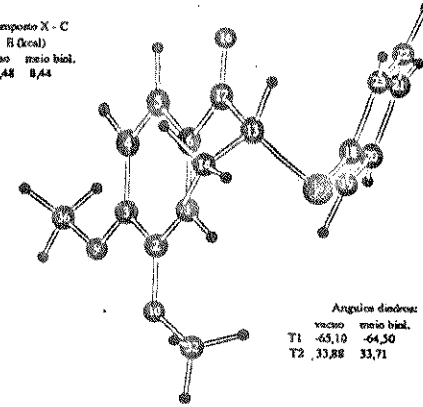
Ângulos diedros:
vacuo meio biol.
T1 107,23
T2 58,78

Composto X - B
E (kcal)
vacuo meio biol.
7,47 3,17



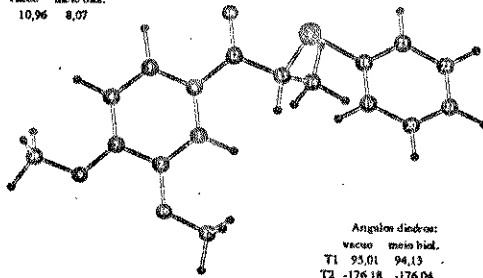
Ângulos diedros:
vacuo meio biol.
T1 151,24 149,74
T2 56,92 37,30

Composto X - C
E (kcal)
vacuo meio biol.
10,48 8,44



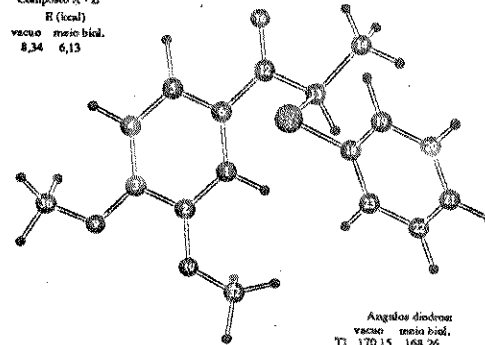
Ângulos diedros:
vacuo meio biol.
T1 -65,10 -64,50
T2 33,88 33,71

Composto X - D
E (kcal)
vacuo meio biol.
10,96 8,07



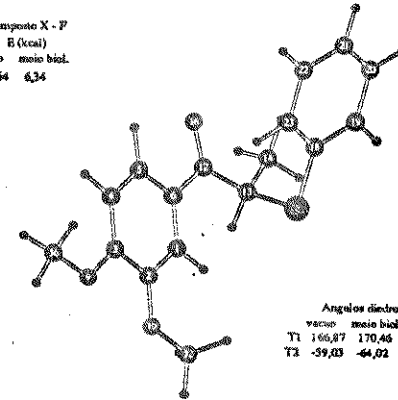
Ângulos diedros:
vacuo meio biol.
T1 95,01 94,13
T2 -176,18 -176,04

Composto X - E
E (kcal)
vacuo meio biol.
8,34 6,13



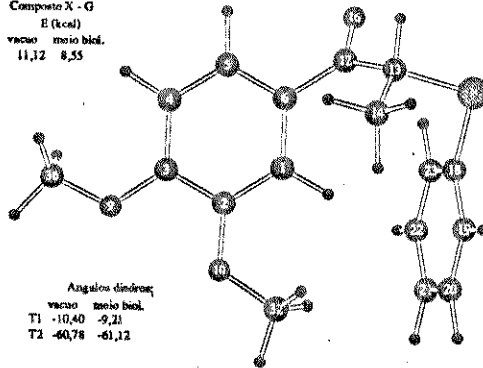
Ângulos diedros:
vacuo meio biol.
T1 170,15 168,26
T2 173,92 172,88

Composto X - F
E (kcal)
vacuo meio biol.
6,54 6,34



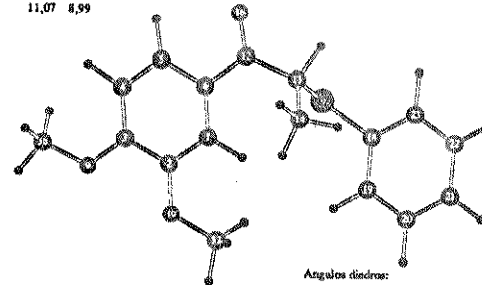
Ângulos diedros:
vacuo meio biol.
T1 166,87 170,46
T2 -59,03 -64,02

Composto X - G
E (kcal)
vacuo meio biol.
11,12 8,55

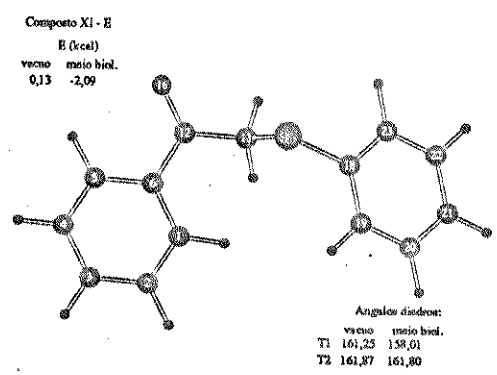
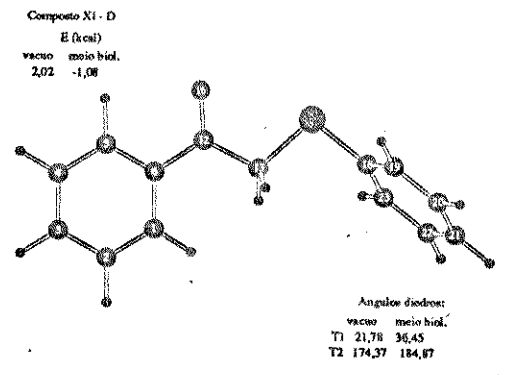
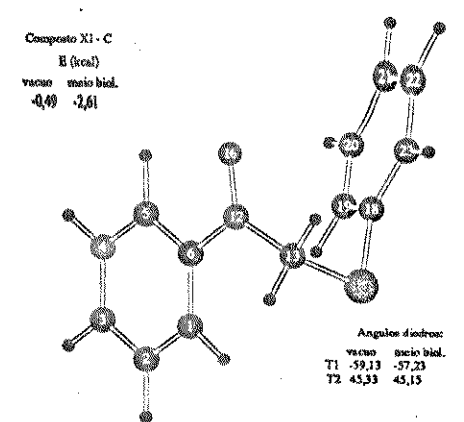
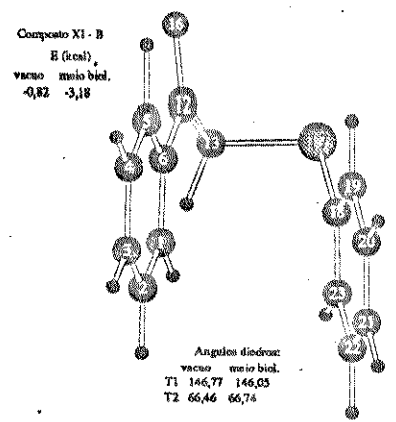
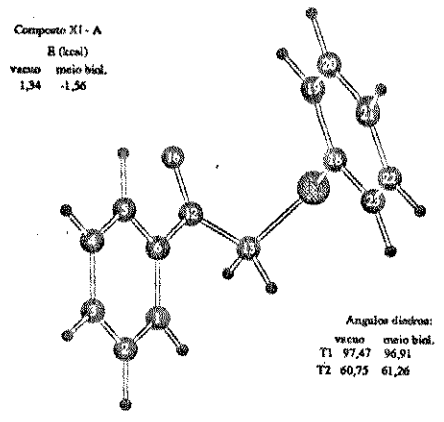


Ângulos diedros:
vacuo meio biol.
T1 -10,60 -9,21
T2 -60,78 -61,12

Composto X - H
E (kcal)
vacuo meio biol.
11,07 8,99

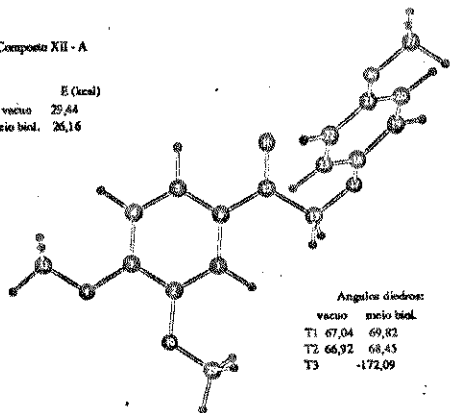


Ângulos diedros:
vacuo meio biol.
T1 -75,79 -72,28
T2 -165,34 -166,02



Composto XII - A

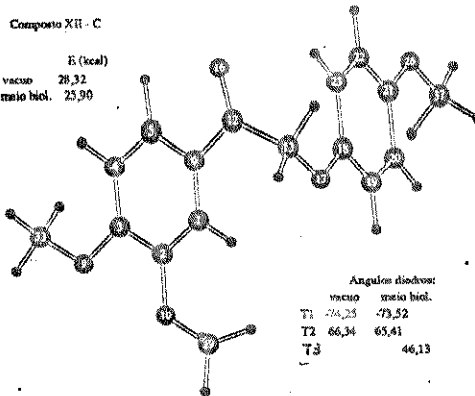
E (kcal)
 vacuo 29,44
 meio biol. 25,16



Ângulos diedros:
 vacuo meio biol.
 T1 67,04 69,82
 T2 66,92 68,45
 T3 -172,09

Composto XII - C

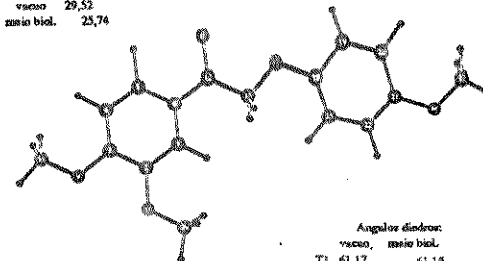
E (kcal)
 vacuo 28,32
 meio biol. 25,90



Ângulos diedros:
 vacuo meio biol.
 T1 -24,25 -79,52
 T2 66,34 65,41
 T3 46,13

Composto XII - D

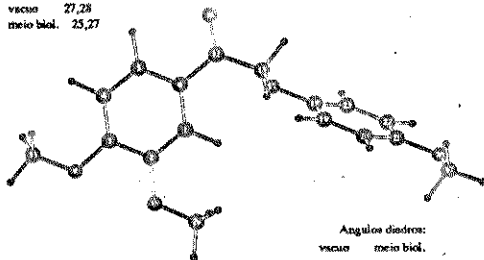
E (kcal)
 vacuo 29,52
 meio biol. 23,74



Ângulos diedros:
 vacuo meio biol.
 T1 61,17 61,15
 T2 178,76 178,76

Composto XII - B

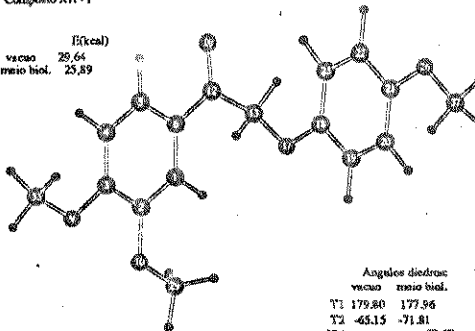
E (kcal)
 vacuo 27,28
 meio biol. 25,27



Ângulos diedros:
 vacuo meio biol.
 T1 -175,71 177,38
 T2 -179,30 178,97
 T3 -63,34

Composto XII - P

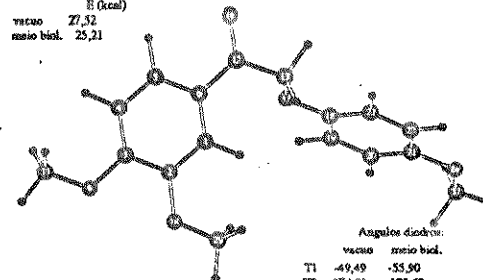
E (kcal)
 vacuo 29,64
 meio biol. 25,89



Ângulos diedros:
 vacuo meio biol.
 T1 179,80 177,96
 T2 -65,15 -71,81
 T3 -37,67

Composto XII - H

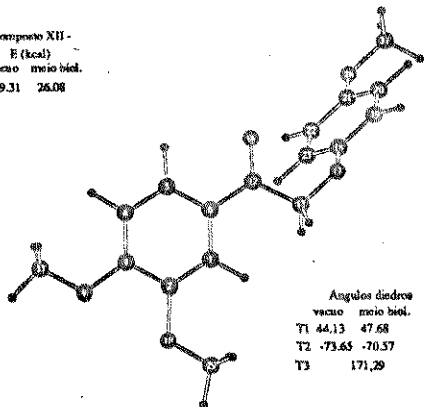
E (kcal)
 vacuo 27,32
 meio biol. 25,21



Ângulos diedros:
 vacuo meio biol.
 T1 -49,49 -55,90
 T2 174,81 -179,63
 T3 65,54

Composto XII -

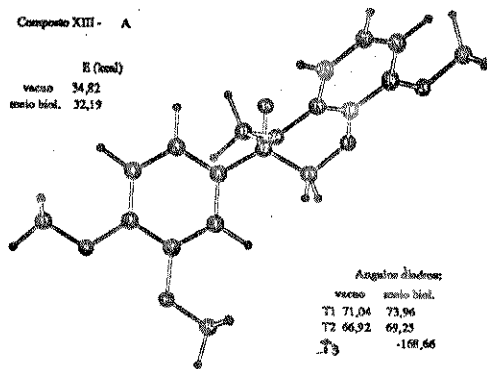
E (kcal)
 vacuo 29,31
 meio biol. 26,08



Ângulos diedros:
 vacuo meio biol.
 T1 44,13 47,68
 T2 -73,65 -70,57
 T3 171,29

Composto XIII - A

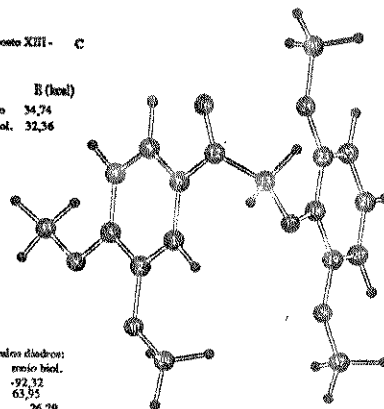
H (kcal)
 vacuo 34,82
 meio biol. 32,19



Ângulos diedros:
 vacuo meio biol.
 T1 71,04 73,96
 T2 66,92 69,23
 T3 -168,66

Composto XIII - C

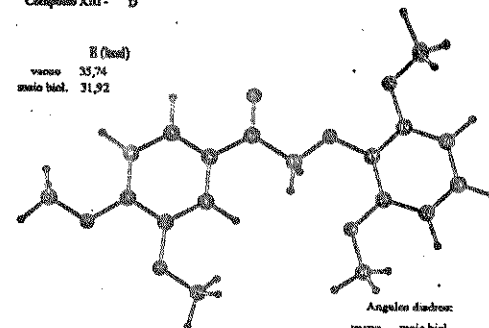
H (kcal)
 vacuo 34,74
 meio biol. 32,36



Ângulos diedros:
 vacuo meio biol.
 T1 -95,34 -92,32
 T2 66,76 63,93
 T3 26,29

Composto XIII - D

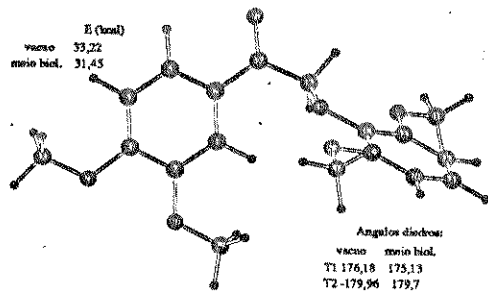
H (kcal)
 vacuo 35,74
 meio biol. 31,92



Ângulos diedros:
 vacuo meio biol.
 T1 73,02 72,43
 T2 178,14 178,81

Composto XIII - E

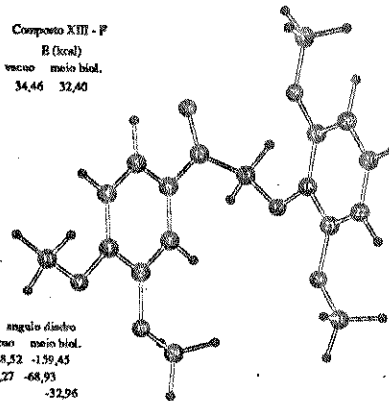
H (kcal)
 vacuo 33,22
 meio biol. 31,45



Ângulos diedros:
 vacuo meio biol.
 T1 176,18 175,13
 T2 -179,96 179,7
 T3 -64,78

Composto XIII - F

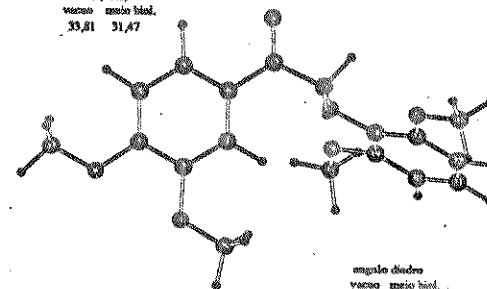
H (kcal)
 vacuo 34,46
 meio biol. 32,40



ângulo diedro
 vacuo meio biol.
 T1 -168,52 -159,45
 T2 -75,27 -68,93
 T3 -32,96

Composto XIII - H

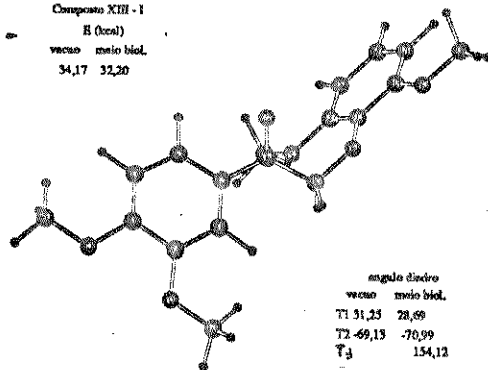
H (kcal)
 vacuo 33,81
 meio biol. 31,47



ângulo diedro
 vacuo meio biol.
 T1 -61,52 -59,06
 T2 181,24 178,28
 T3 62,08

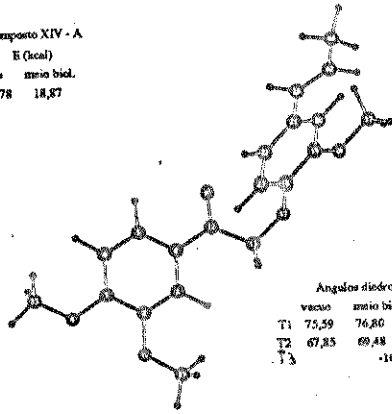
Composto XIII - I

H (kcal)
 vacuo 34,17
 meio biol. 32,20



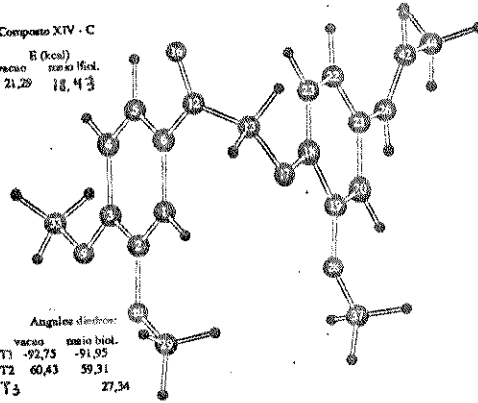
ângulo diedro
 vacuo meio biol.
 T1 31,25 28,69
 T2 -69,13 -70,99
 T3 154,12

Composto XIV - A
E (kcal)
vacuo meio biol.
22,78 18,87



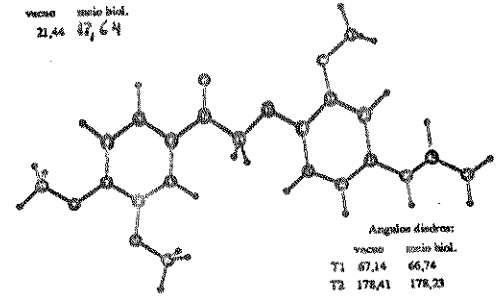
Ângulos diedros:
vacuo meio biol.
T1 75,59 76,80
T2 67,85 69,48
T3 -164,82

Composto XIV - C
E (kcal)
vacuo meio biol.
21,29 18,43



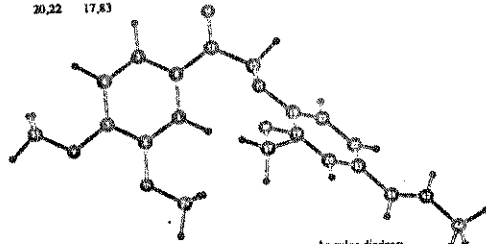
Ângulos diedros:
vacuo meio biol.
T1 -92,75 -91,95
T2 60,43 59,31
T3 27,34

Composto XIV - D
E (kcal)
vacuo meio biol.
21,44 17,64



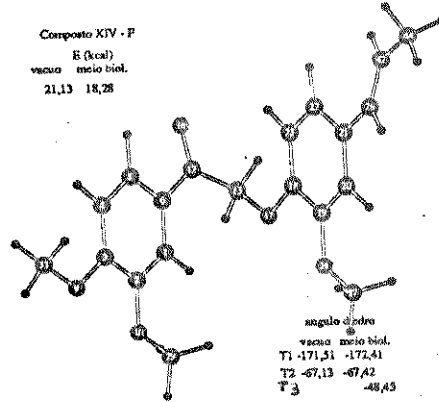
Ângulos diedros:
vacuo meio biol.
T1 67,14 66,74
T2 178,41 178,23

Composto XIV - E
E (kcal)
vacuo meio biol.
20,22 17,83



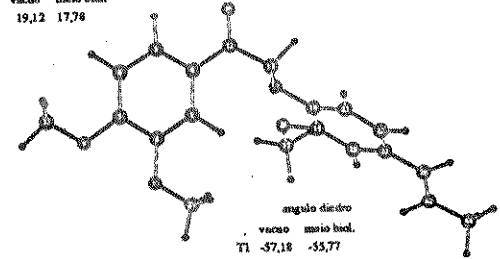
Ângulos diedros:
vacuo meio biol.
T1 -179,65 177,43
T2 166,85 167,18
T3 -62,35

Composto XIV - F
E (kcal)
vacuo meio biol.
21,13 18,28



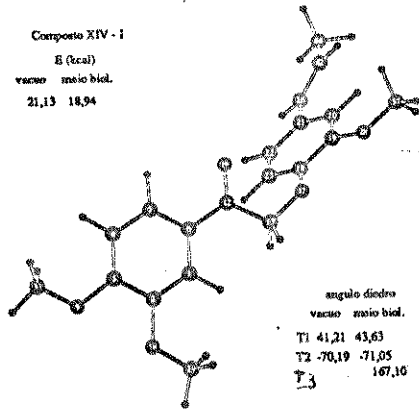
ângulo diedro
vacuo meio biol.
T1 -171,51 -172,41
T2 -67,13 -67,42
T3 -48,45

Composto XIV - H
E (kcal)
vacuo meio biol.
19,12 17,78



ângulo diedro
vacuo meio biol.
T1 -57,18 -55,77
T2 -180,10 -179,94
T3 65,63

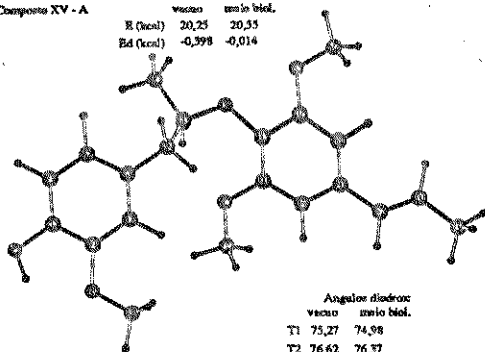
Composto XIV - I
E (kcal)
vacuo meio biol.
21,13 18,94



ângulo diedro
vacuo meio biol.
T1 41,21 43,63
T2 -70,19 -71,05
T3 167,10

Compound XV - A

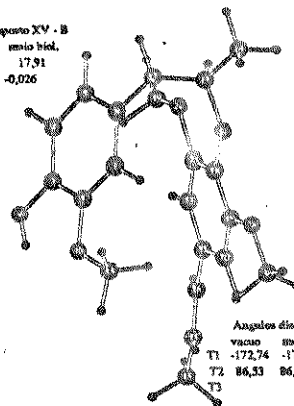
	vacuo	meio biol.
H (kcal)	20,25	20,55
Ed (kcal)	-0,998	-0,014



Angulos dihedros:
vacuo meio biol.
T1 75,27 74,98
T2 76,62 76,37
T3 -165,92

Compound XV - B

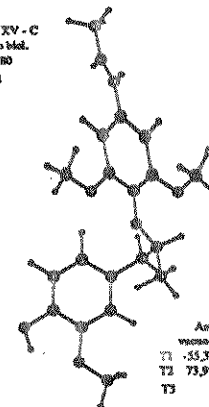
	vacuo	meio biol.
H (kcal)	17,41	17,91
Ed (kcal)	-0,630	-0,026



Angulos dihedros:
vacuo meio biol.
T1 -172,74 -173,18
T2 86,33 86,17
T3 -94,10

Compound XV - C

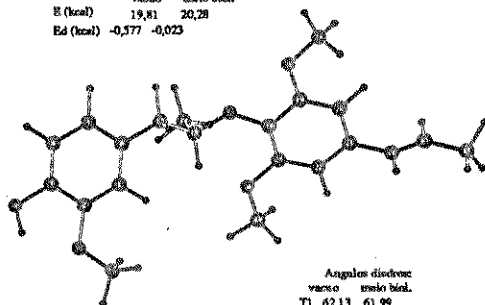
	vacuo	meio biol.
H (kcal)	19,52	19,80
Ed (kcal)	-0,361	-0,012



Angulos dihedros:
vacuo meio biol.
T1 -55,37 -55,55
T2 73,98 73,14
T3 64,99

Compound XV - D

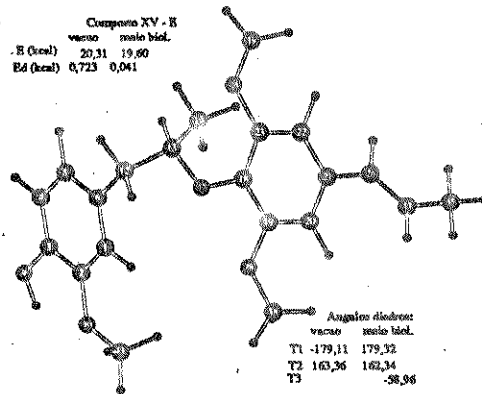
	vacuo	meio biol.
H (kcal)	18,81	20,28
Ed (kcal)	-0,577	-0,023



Angulos dihedros:
vacuo meio biol.
T1 62,13 61,99
T2 161,48 161,43

Compound XV - E

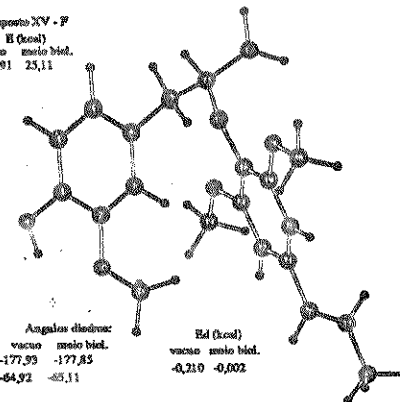
	vacuo	meio biol.
H (kcal)	20,31	19,60
Ed (kcal)	0,723	0,041



Angulos dihedros:
vacuo meio biol.
T1 -179,11 179,52
T2 163,36 162,34
T3 -58,96

Compound XV - F

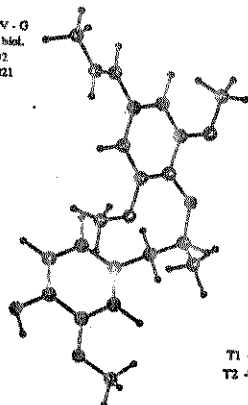
	vacuo	meio biol.
H (kcal)	24,01	23,11



Angulos dihedros:
vacuo meio biol.
T1 -177,99 -177,85
T2 -64,92 -65,11
Ed (kcal):
vacuo meio biol.
-0,210 -0,002

Compound XV - G

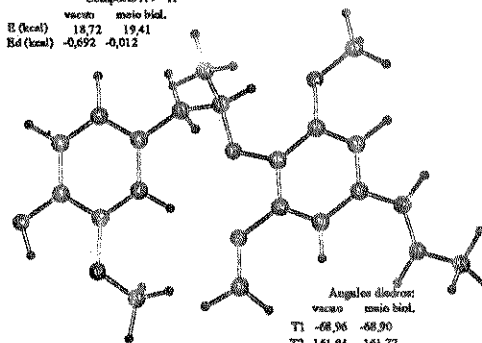
	vacuo	meio biol.
H (kcal)	26,75	27,02
Ed (kcal)	-0,291	-0,021



Angulos dihedros:
vacuo meio biol.
T1 -36,12 -36,37
T2 -73,16 -73,71

Compound XV - H

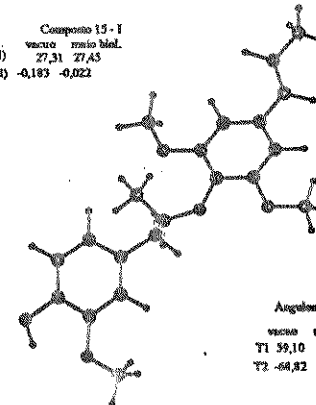
	vacuo	meio biol.
H (kcal)	18,72	19,41
Ed (kcal)	-0,692	-0,012



Angulos dihedros:
vacuo meio biol.
T1 -68,96 -68,90
T2 161,84 161,72
T3 52,15

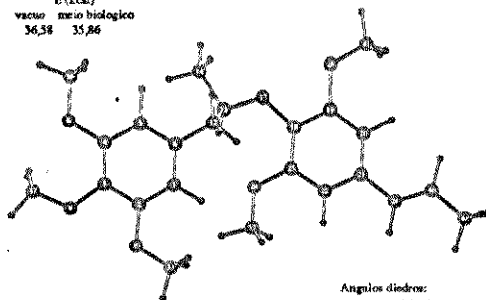
Compound XV - I

	vacuo	meio biol.
H (kcal)	27,31	27,45
Ed (kcal)	-0,183	-0,022



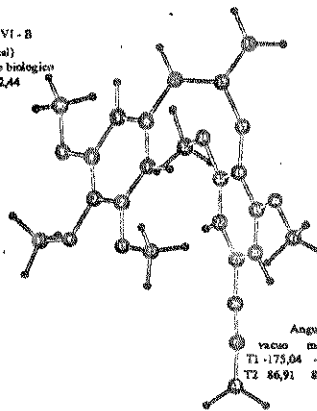
Angulos dihedros:
vacuo meio biol.
T1 59,10 59,45
T2 -68,82 -69,55

Composto XVI - A
 E (kcal)
 vacuo meio biologico
 36,38 35,86



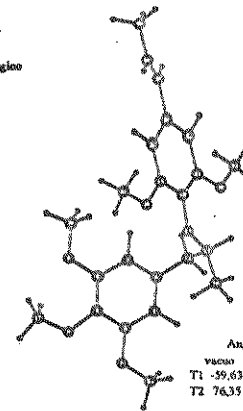
Ângulos diedros:
 vacuo meio biologico
 T1 74,72 74,65
 T2 75,17 75,43

Composto XVI - B
 E (kcal)
 vacuo meio biologico
 33,00 32,44



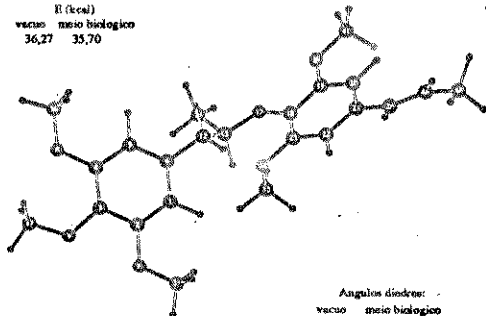
Ângulos diedros:
 vacuo meio biologico
 T1 -175,04 -174,90
 T2 86,91 86,90

Composto XVI - C
 E (kcal)
 vacuo meio biologico
 34,75 35,05



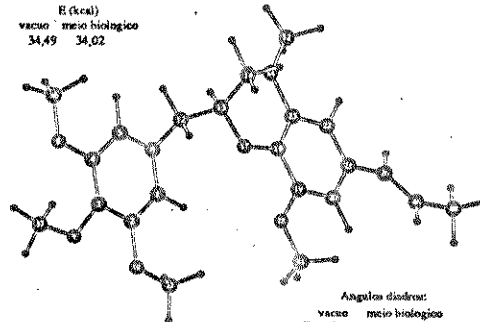
Ângulos diedros:
 vacuo meio biologico
 T1 -59,63 -59,67
 T2 76,35 76,33

Composto XVI - D
 E (kcal)
 vacuo meio biologico
 36,27 35,70



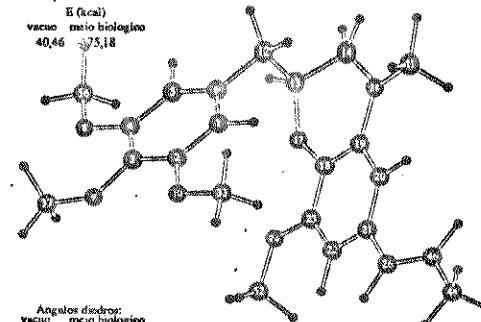
Ângulos diedros:
 vacuo meio biologico
 T1 62,06 62,01
 T2 162,75 162,67

Composto XVI - E
 E (kcal)
 vacuo meio biologico
 34,49 34,02



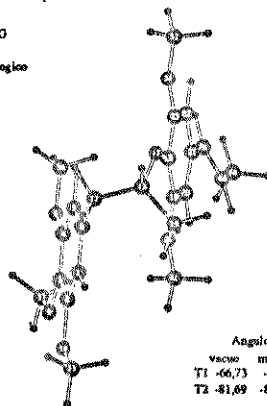
Ângulos diedros:
 vacuo meio biologico
 T1 178,64 178,43
 T2 165,43 165,66

Composto XVI - F
 E (kcal)
 vacuo meio biologico
 40,46 35,18



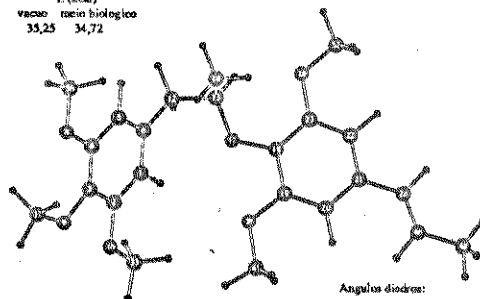
Ângulos diedros:
 vacuo meio biologico
 T1 -175,16 -64,91
 T2 -64,96 75,43

Composto XVI - G
 E (kcal)
 vacuo meio biologico
 41,74 41,13



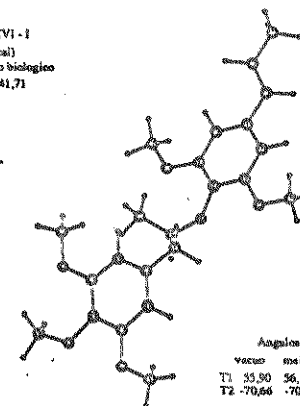
Ângulos diedros:
 vacuo meio biologico
 T1 -66,73 -64,53
 T2 -81,69 -81,25

Composto XVI - H
 E (kcal)
 vacuo meio biologico
 35,25 34,72



Ângulos diedros:
 vacuo meio biologico
 T1 -67,11 -67,22
 T2 161,99 161,98

Composto XVI - I
 E (kcal)
 vacuo meio biologico
 42,46 41,71

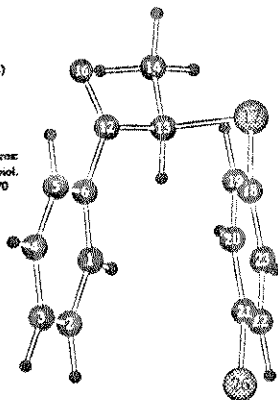


Ângulos diedros:
 vacuo meio biologico
 T1 55,90 56,13
 T2 -70,66 -70,93

Composto XVII - B

	Ei (kcal)
vacuo	1,31
meio biol.	-1,14

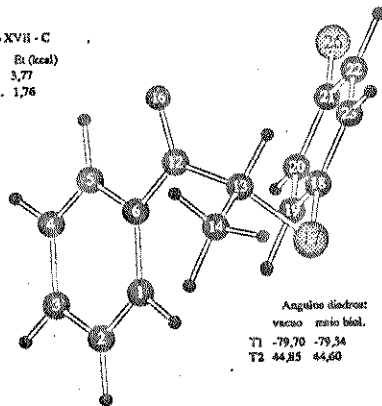
Ângulos diedros:
vacuo meio biol.
T1 152,37 151,70
T2 57,95 58,27



Composto XVII - C

	Ei (kcal)
vacuo	3,77
meio biol.	1,76

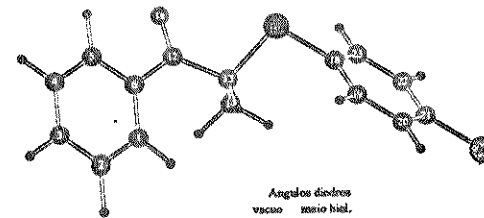
Ângulos diedros:
vacuo meio biol.
T1 -79,70 -79,54
T2 44,85 44,60



Composto XVII - D

	Ei (kcal)
vacuo	4,19
meio biol.	1,32

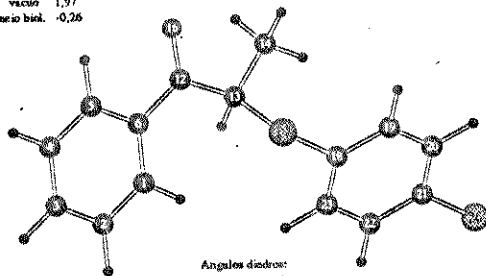
Ângulos diedros:
vacuo meio biol.
T1 102,51 98,35
T2 -177,42 -176,54



Composto XVII - E

	Ei (kcal)
vacuo	1,97
meio biol.	-0,26

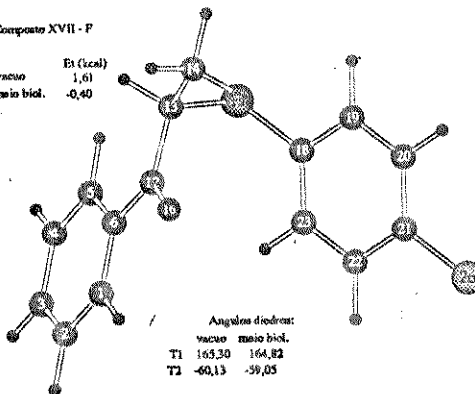
Ângulos diedros:
vacuo meio biol.
T1 166,50 163,70
T2 172,61 172,82



Composto XVII - F

	Ei (kcal)
vacuo	1,61
meio biol.	-0,40

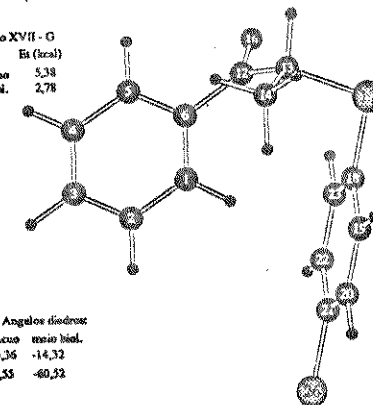
Ângulos diedros:
vacuo meio biol.
T1 165,30 164,82
T2 -60,13 -59,05



Composto XVII - G

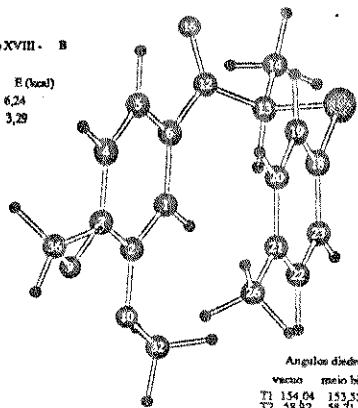
	Ei (kcal)
vacuo	5,38
meio biol.	2,78

Ângulos diedros:
vacuo meio biol.
T1 -15,36 -14,32
T2 -60,55 -60,52



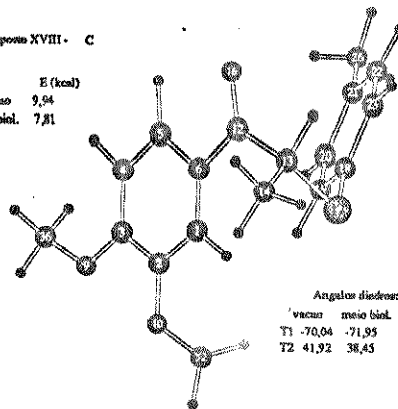
Composto XVIII - B

E (kcal)
vacuo 6,24
meio biol. 3,29



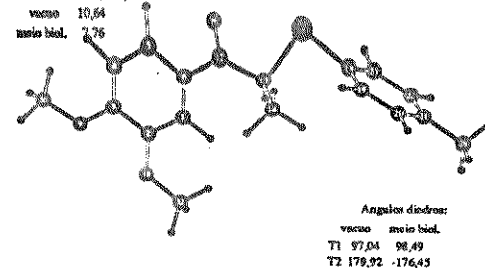
Composto XVIII - C

E (kcal)
vacuo 9,04
meio biol. 7,81



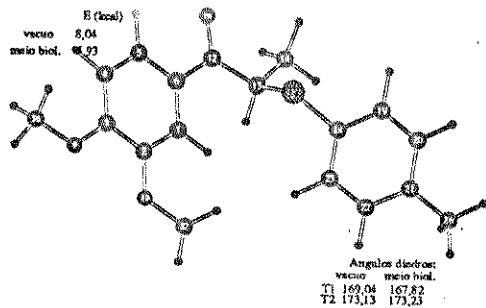
Composto XVIII - D

E (kcal)
vacuo 10,64
meio biol. 7,76



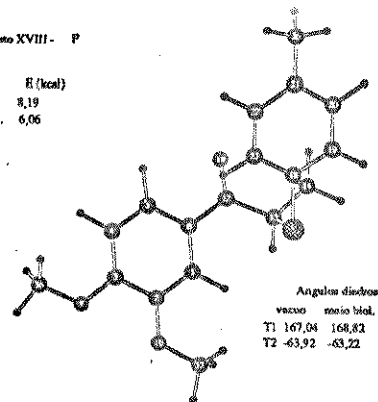
Composto XVIII - E

E (kcal)
vacuo 8,04
meio biol. 4,93



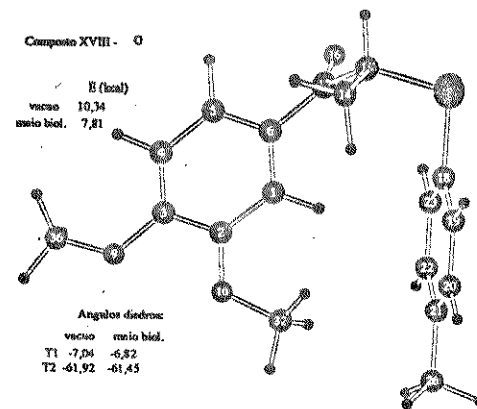
Composto XVIII - F

E (kcal)
vacuo 8,19
meio biol. 6,06



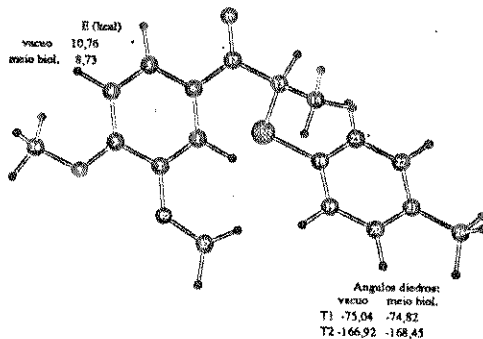
Composto XVIII - G

E (kcal)
vacuo 10,34
meio biol. 7,81

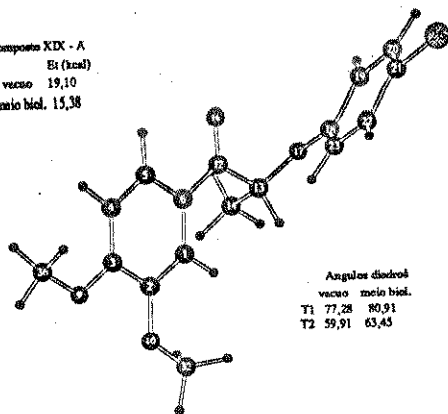


Composto XVIII - H

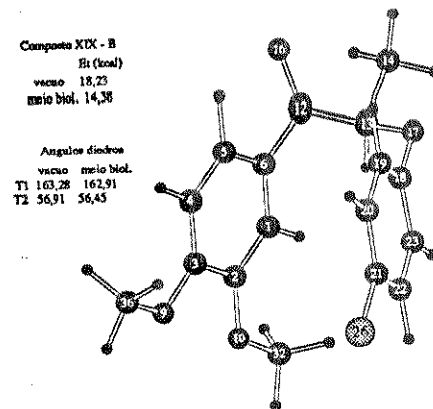
E (kcal)
vacuo 10,76
meio biol. 8,73



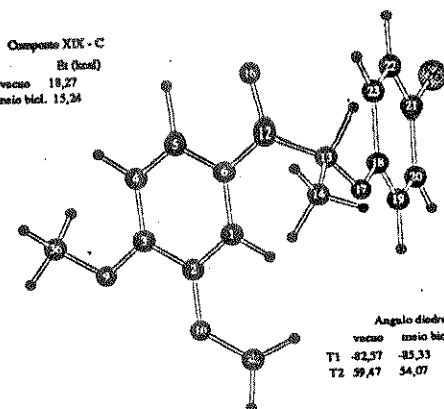
Composto XIX - A
Et (kcal)
vacuo 19,10
meio biol. 15,38



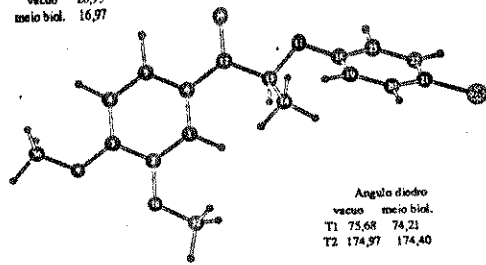
Composto XIX - B
Et (kcal)
vacuo 18,23
meio biol. 14,38



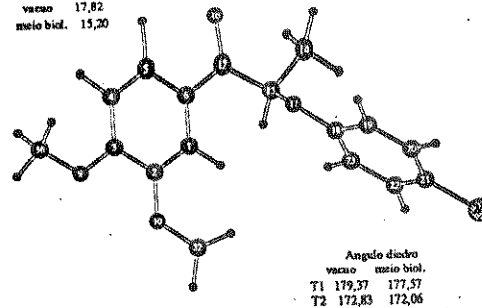
Composto XIX - C
Et (kcal)
vacuo 18,27
meio biol. 15,24



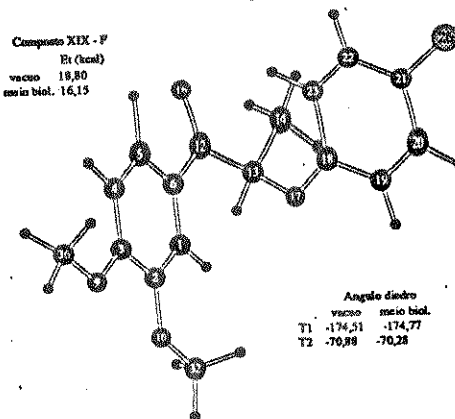
Composto XIX - D
Et (kcal)
vacuo 20,93
meio biol. 16,97



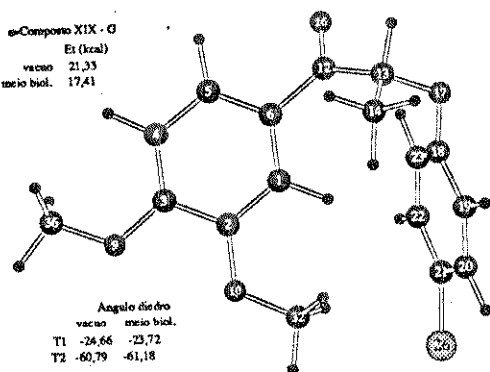
Composto XIX - E
Et (kcal)
vacuo 17,82
meio biol. 15,20



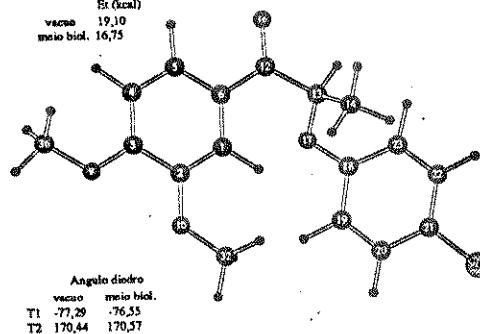
Composto XIX - F
Et (kcal)
vacuo 18,80
meio biol. 16,15

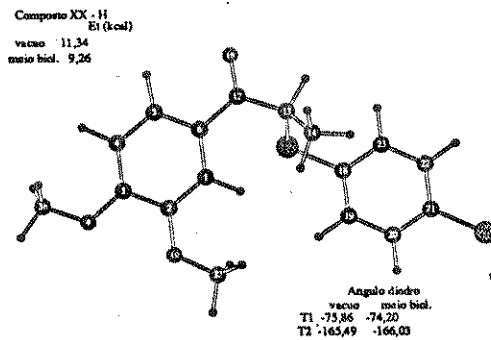
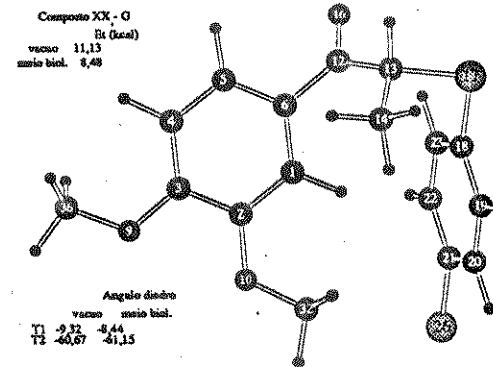
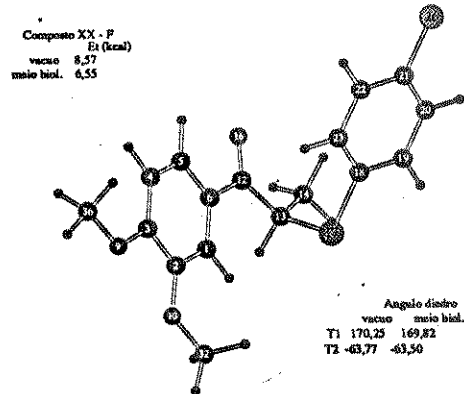
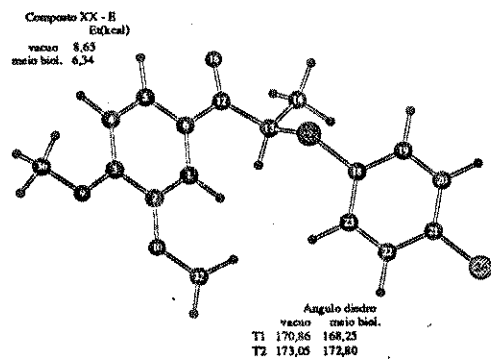
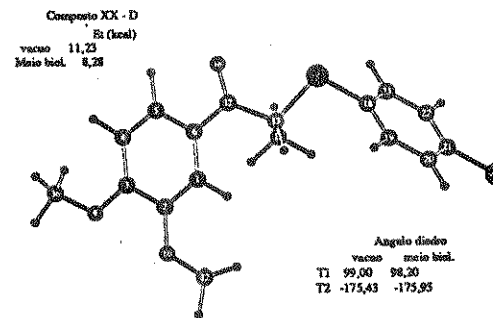
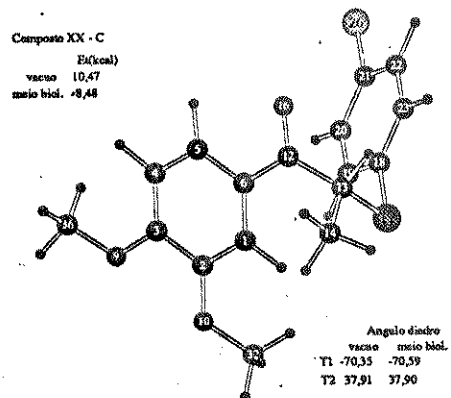
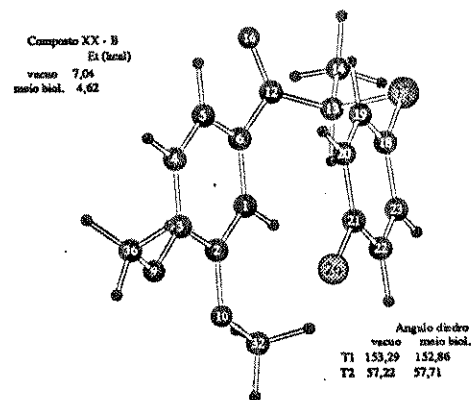


Composto XIX - G
Et (kcal)
vacuo 21,33
meio biol. 17,41



Composto XIX - H
Et (kcal)
vacuo 19,10
meio biol. 16,75





REFERENCIAS

1. - , *Globo Ciência*, **13** , 32-39 (1992)
2. Barata, L. E. S., 2. *Brasilianisch-Deutsches Symposium für Naturstoffchemie* (1991)
3. Barata, L. E. S. e Santos, L. S., *Work-Shop Internacional de Plantas Medicinais dos Países do tratado de cooperação amazônica* (1988)
4. Mayrink, W., *Ciência Hoje*, **11** , 63, 32 (1990)
5. Paine, A. et al., manuscrito não publicado.
6. Richards, W. G. "Quantum Pharmacology", Butter & Tanner Ltd., Great Britain (1977), pp. 119-122
7. Burket, U. and Allinger, N. L. "Molecular Mechanics", ACS Monograph, Washington D.C. (1982), cap. 1 e 2
8. Jaime, C. and Osawa, E., *Tetrahedron*, **39** , 17, 2769 (1983)
9. Allinger, N. L., and Yuh, Y. H., "Operating Instructions For MM2 and MMP2 Programs, 1977 Force Fields" (1980)
10. Gill, E. W., *Progr. Med. Chem.*, **4** , 39 (1965)
11. Hammett, L. P., "Physical Organic Chemistry", McGraw-Hill, New York, (1970).
12. Taft, R. W., Separation of Polar, Steric and Resonance Effects in Reactivity in "Steric Effects in Organic Chemistry", Melvin S. Newman, Ed., Wiley, New York (1956), pp. 556-675
13. Overton, E., *Zeitschrift für Physikalisch Chemie*, **22**, 189 (1897)
14. Meyer, H., *Archiv für Experimentell Pathologie und Pharmakologie*, **42** , 109 (1899)
15. Hansch, C. et al. *J. Am. Chem. Soc.*, **85**, 2817 (1963)
16. Danielli, J. F., et al. , *Fundamental Concepts of Drug-Receptor Interactions*, Academic, New York, (1970)
17. Kier, L. B. , *Molecular Orbital Studies in Chemical Pharmacology*, Springer, Berlin, (1970)
18. Kier, L. B., *Molecular Orbital Theory in Drug Research*, Academic, New York, (1971)
19. Purcell, W. P. et al., "Quantitative Structure-Activity Relationships and Molecular Orbitals in Medicinal Chemistry", in Burger, A., Ed., *Medicinal Chemistry*, 3 ed. Wiley-Interscience, New York (1970), pp. 164-192.
20. Wohl, A. J. "A Molecular Orbital Approach to Quantitative Drug Design", in Ariëns, E. J., Ed., *Drug Design*, Vol. I, Academic, New York (1971), pp. 381-404
21. Fleming, I., "Frontier Orbitals and Organic Chemical Reactions", London, John Wiley & Sons, Ltd. (1978), pp. 5-32
22. Fukui, K. et al., *J. Chem. Phys.*, **20**, 722 (1952)
23. Klopman, G., *J. Am. Chem. Soc.*, **90**, 223 (1968)
24. Salem, L., *J. Am. Chem. Soc.*, **90**, 543 e 553 (1968)
25. Martin, Y. C. et al., *J. Med. Chem.*, **16**, 1089 (1973)
26. Ariëns, E. J., Ed., *Drug Design*, Vol. 1, Academic, New York (1971)
27. Kier, L. B., Ed., *Molecular Orbital Studies in Chemical Pharmacology*, Springer, Berlin (1970)
28. Kier, L. B., *Molecular Orbital Theory in Drug Research*, Academic, New York (1971)
29. Cammarata, A. & Martin, A. N., "Physical Properties and Biological Activity", in Burger, A., Ed., *Medicinal Chemistry*, 3 ed., Wiley-Interscience, New York, (1970), pp. 118-163

- 30 Pearson, R. G., *J. Am. Chem. Soc.*, **85**, 3533 (1963)
- 31 Shriver, D. F., et al. "Inorganic Chemistry", Oxford University Press (1990), pp. 31-35
- 32 Hirsch, J. A., "Concepts in Theoretical Organic Chemistry", Allynand Bacon, Inc., Boston, USA (1975), pp. 88-89.
- 33 Lien, E. J. et al. *J. Pharm. Sic.*, **71** 6, 641 (1982)
- 34 Garrett, E. R., et al., *J. Med. Chem.*, **9** 203 (1966)
- 35 Moore, W. J., "Físico-Química", Vol. 2, Ed. Edgard Blücher Ltda. (1976), pp. 622, 623, 629.
- 36 Mulliken, R. S., *J. Chem. Phys.* **2**, 782 (1934)
- 37 Craig, P. N., *J. Med. Chem.*, **14** 680 (1971)
- 38 Higo, J. and Go, N., *J. Comp. Chem*, **10** 3, 376 (1989)
- 39 Stewart, J. J. P., *J. Comput. Chem.*, **10** 2, 221 (1989)
- 40 Åsbrink, L., et al., *Chem. Phys. Let.*, **52** ,1, 63 (1977)
- 41 Portoghese, P. S. et al., *J. Med. Chem.*, **14**, 144 (1971)
- 42 Kowalski, B. R. and Bender, C. F. , *J. Am. Chem. Soc.*, **9** , 5632 (1972)
- 43 Hansch, C., *Accounts of Chemical Research* **2**, 232 (1969)
- 44 Free, S. M. and Wilson, J. W., *J. Med. Chem.*, **7**, 395 (1964)
- 45 Scarminio, I. S. e Bruns, R. E., *Trends Anal. Chem.* **8**, 326 (1989)
- 46 Allinger, N. L., *Adv. Phys. Org. Chem.* **13**, 1 (1976)
- 47 Allinger, N. A., *J. Am. Chem. Soc.*, **99**, 25, 8127 (1977)
- 48 Allinger, N. L., et al., *J. Comp. Chem.*, **9**, 6, 591 (1988)
- 49 Schaefer, T., et al., *J. Mol. Struct.*, **89**, 93 (1982)
- 50 Bakker, J. M. and Chadwick, D. J., *Acta Chem. Scand. B* **42**, 223 (1988)
- 51 Dewar, M. J. S., and Thiel, W., *J. Am. Chem. Soc.*, **99** , 15, 4899 (1977)
- 52 Stewart, J. J. P., *J. Comput. Chem.*, **10** , 2, 209 (1989)
- 53 Dewar, M. J. S., et al., *J. Am. Chem. Soc.*, **107**, 3902 (1985)
- 54 Blinder, S. M., *Am. J. Phys.*, **33** 6, 431 (1965)
- 55 Pople, J. A. et al., *J. Chem. Phys.* **47**, 2026 (1967)
- 56 Oleari, L., et al., *Mol. Phys.*, **10** , 97 (1966)
- 57 Dewar, M. J. S., and Hojvat, N. L. (Sabelli), *J. Chem. Phys.* **34** , 1232 (1961)
- 58 Klopman, G., *J. Am. Chem. Soc.*, **86**, 4550 (1964)
- 59 Fukunaga, K. e Koontz, W. L. G., *IEEE Trans. Comp.*, **C-19**, 311 (1970)
- 60 Wold, S. *Pattern Recognition*, **8**, 127 (1976)