

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA

SBI-IFUSP



305M810T2693

MÉTODOS CLÁSSICOS E QUÂNTICOS NO
ESTUDO DE GEOMETRIAS DE MOLÉCULAS
BIOLÓGICAS

Banca examinadora

Prof. Dr. J. A. Fornés	UFGO
Prof. Dr. L. C. Gomide Freitas	UFSCar
Prof. Dr. R. Custódio	UNICAMP
Prof. Dr. Sylvio A. Canuto	IFUSP
Prof. Dr. A. S. Ito - orientador	IFUSP



Tese apresentada ao Instituto de Física da
Universidade de São Paulo, para a obtenção
do título de Doutor em Ciências

Orientador: Prof. Dr. Amando Siuiti Ito
Coorientador: Prof. Dr. Michel Loos

Amando Siuiti Ito

André Tsutomu Ota

São Paulo
1997

Iberê Luiz Caldas

Prof. Iberê Luiz Caldas
Presidente da Comissão de Pós-Graduação

30/09/97

537.5
087 m
D
ex. 1

FICHA CATALOGRÁFICA

**Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo**

Ota, André Tsutomu

**Métodos Clássicos e Quânticos no Estudo de Geometrias
de Moléculas Biológicas. São Paulo, 1997.**

**Tese (Doutoramento) Universidade de São Paulo.
Instituto de Física - Departamento de Física Geral.**

**Área de Concentração: Física da Matéria Condensada
Orientador: Prof. Dr. Amando Siuiti Ito.**

**Unitermos: 1. Energia Livre; 2. Tiosemicarbazona;
3. Triptofano; 4. Ácido Aminobenzóico; 5. Fentanil.**

USP/IF/SBI - 41/97

Resumo

Duas são as principais características deste trabalho. A primeira é a introdução de um método para o cálculo da energia livre absoluta no programa de dinâmica molecular clássica Thor. Este método é geral e pode ser usado no vácuo ou em solução, sob quaisquer condições de pressão e/ou temperatura. Trata-se de um caso especial de integração termodinâmica para cálculos de energia livre, e compara um estado termodinâmico com um sistema modelo para o qual a energia livre é conhecida analiticamente. O estado modelo é conhecido no início, e após uma integração termodinâmica até o estado real final, a energia livre absoluta é calculada como uma soma desta integração mais a energia livre do estado modelo. A energia livre depende do modelo a partir do qual a função de partição é derivada analiticamente. O modelo definido é o sólido de Einstein. Os átomos deste sólido ideal não interagem entre si (como um gás ideal) e estão ligados harmonicamente às suas posições ideais. O acoplamento entre os dois estados é feito diretamente por intermédio da energia. Este método foi calibrado e testado na molécula fentanil, e nossos resultados são comparáveis aos do pacote Discover para ela. Aplicamos este método no aminoácido triptofano que é uma sonda fluorescente para a estrutura de peptídeos e proteínas. A energia livre absoluta é proposta como um parâmetro relevante que determina os estados conformacionais da molécula de triptofano. A segunda característica é o uso do programa quântico semiempírico Geomop. Os cálculos realizados com este programa fornecem-nos a estrutura eletrônica e as geometrias do Abz (ácido antranílico - uma sonda extrínica para o estudo de peptídeos) e com seus compostos com a alanina, a fenilalanina e a prolina. Os espectros resultantes destes cálculos são comparáveis com aqueles obtidos experimentalmente. Outra aplicação do programa é a análise das trajetórias de interconversão da molécula de tiosemicarbazona em duas formas isoméricas, mostrando outra possibilidade do uso de cálculos semiempíricos.

Abstract

This work has two main features. First it introduces a method for absolute free energy calculations in the classical molecular dynamics Thor program. This method is general and could be used in vacuum or in solution, under any conditions of pressure and/or temperature. It is a special case of thermodynamic integration for free energy calculations, and compares one thermodynamic state with a model system for which the free energy is analytically known. The initial model state is known and after an integration to a final real state, the absolute free energy is calculated as a sum of that integration plus the free energy of model state. The absolute free energy depends on the model from which the partition function is analytically derived. The model that we have defined is the Einstein solid. The atoms of this ideal system do not interact (like an ideal gas) and are linked harmonically to their ideal position. The coupling between the two states is done directly through the energy. We calibrated and tested this method of free energy calculations on the fentanyl molecule and our results were comparable to those presented by the Discover package for the same molecule. We applied this method to the aminoacid tryptophan, that is a fluorescent probe for the structure of peptides and proteins. The absolute free energy is proposed as the relevant parameter that determines the rotameric conformational states of tryptophan molecule. The second feature is the use of the semi empirical Geomop quantum program. The calculations made with this program gave us the electronic structure and the geometry's of Abz (anthranilic acid, an extrinsic probe for the study of peptides) and its compounds with alanine, phenilalanine and proline. The resulting spectra are comparable with those obtained experimentally. Another application of the program is the analysis of the paths of the inter-conversion of the thiocarbazono molecule in two isomeric forms, showing another possibility of use of semi empirical calculations.

Agradecimentos:

É muito difícil nesta hora lembrar-me de todos os que me ajudaram nesta jornada. É natural que me lembre apenas os que estiveram mais próximos e demonstraram aflição com o trabalho. Assim, agradeço

ao P.Bisch que juntamente com o PeGePa e o Cláudio S. Shida me “introduziram” nesta área;

a Beth e a Márcia que me levaram a estudar um pouco de Química;

aos que “apenas” discutiram, o Piquini, o Sylvio Canuto;

aos que resolveram alguns “probleminhas” principalmente o Wilk, a Ana Maria, ao Roberto Morato;

aos que “torceram” como a Carla, a Tereza, a Regina, a Cecil, ao João, a Aninha, ao Ernane, a Márcia, a Karin, a Rozane, ao Roberto e, ao Eduardo, ao Pedro Rodrigues, ao Amaury, ao Paulo, ao Reynaldo;

aos que tiveram a paciência de ler as primeiras versões e fazer sugestões que mudaram a “cara” do trabalho, a Vera Henriques, ao Marcos e a Marta;

ao Michel Loos, o co-autor de várias partes;

ao Amando pela orientação geral;

aos que sofreram por ele e por causa dele, a Inês, a Laura, a Nádia, o Iuri e a Iara;

Agradeço à Universidade Estadual de Londrina e a todo o Departamento de Física que me liberaram totalmente de qualquer atividade durante este período, ao IFUSP que colocou ao meu dispor toda a sua estrutura, e a CAPES pelo apoio financeiro.

ÍNDICE

Capítulo 1	
Introdução	1
Capítulo 2	
O Programa Thor	9
2.1 – O Campo de Forças do THOR	11
Figuras do capítulo 2	21
Capítulo 3	
Modelagem Molecular Quântica	25
3.1- A Equação de Schrödinger	26
3.2- O Campo Autoconsistente de Hartree-Fock	28
3.3- O Método CI	38
3.4- Métodos Semiempíricos	40
Figura do capítulo 3	47
Capítulo 4	
Cálculo da Energia Livre	53
4.1- O Cálculo da Energia Livre Usando Dinâmica Molecular	53
4.2- Energia Livre Absoluta	57
4.3- Calibração	62
Figuras do capítulo 4	69
Capítulo 5	
O Triptofano	75
Adendo	80
Figuras do capítulo 5	81
Capítulo 6	
A Tiosemicarbazona	91
Figuras do capítulo 6	96
Capítulo 7	
O Abz	107
Figuras do capítulo 7	111
Capítulo 8	
Iniciação	119
Capítulo 9	
Bibliografia	123

Capítulo 1

INTRODUÇÃO

Uma das técnicas mais utilizadas na análise do comportamento de sistemas biológicos realizadas no Grupo de Biofísica e Física Médica do Instituto de Física é o da espectroscopia de fluorescência. Algumas moléculas biológicas são naturalmente fluorescentes, mas a maioria dos aminoácidos que compõem os peptídeos e proteínas não apresenta atividade fluorescente relevante. Para monitorar o comportamento de peptídeos e proteínas, estejam isolados ou em interação com micelas ou membranas, necessita-se de pelo menos um elemento fluorescente, que pode ser intrínseco, como por exemplo o triptofano, um aminoácido natural, ou extrínseco, como o caso do ácido ortoaminobenzóico, utilizado por nosso grupo. A intensa atividade fluorescente deste ácido permite que ele seja introduzido no conjunto de moléculas biológicas que possuam pouca ou nenhuma atividade fluorescente e serve, dessa maneira, como uma sonda biológica. Muitas vezes modifica-se a molécula fluorescente do triptofano⁽¹³⁾ gerando análogos como o 5hidroxi-triptofano, no qual o grupo hidroxila é colocado na quinta posição do anel aromático, e o 7aza-triptofano, onde um átomo de nitrogênio é colocado na sétima posição do anel aromático, e apresenta, cada um deles, um deslocamento bem definido na fluorescência em relação à do triptofano. A fluorescência do ácido ortoaminobenzóico⁽¹⁷⁾ pode também sofrer modificações quando este está associado a outras substâncias químicas, como por exemplo, a glicina, a prolina e a fenilalanina^(18,19,20). Entender a estrutura eletrônica destas substâncias é de importância vital para a pesquisa que usa a técnica de fluorescência. Este trabalho tem por objetivo auxiliar neste entendimento e descreve os métodos e técnicas de cálculos computacionais que utilizamos para a análise qualitativa e quantitativa das estruturas eletrônicas e geométricas destas sondas

biológicas.

O ponto fulcral do trabalho é a introdução de uma técnica de cálculo da energia livre, quando uma molécula possui duas ou mais conformações isoméricas. As técnicas até então usadas não nos convém devido a dificuldades intrínsecas que elas acarretavam, tanto do ponto de vista computacional *per si* com a introdução de um solvente, que extrapola o limite físico das máquinas, como do ponto de vista conceitual e operacional com a introdução de átomos que se transmütavam em outros, o que se choca com a nossa formação. Este problema foi resolvido com o método do cálculo da energia livre absoluta, que aplica a técnica das perturbações na função hamiltoniana como uma combinação linear de duas funções de energia potencial que descreve os estados limites V_0 e V_H , onde V_0 é a função energia potencial normal (incluindo ligações, ângulos, torções, etc.), e V_H é um potencial localizado de oscilador harmônico definido pela constante de mola do *i*^{ésimo} átomo, a sua coordenada instantânea e a rede de referência dos átomos não interagentes (referido como um sólido de Einstein). Assim, o principal resultado deste trabalho é a introdução deste método de cálculo no programa Thor⁽¹⁾, feita com a ajuda muito importante do professor Michel Loos, e de sua aplicação na triptofano⁽¹³⁾.

Investigamos também como acontece a modificação da estrutura eletrônica do ácido ortoaminobenzóico⁽¹⁷⁾ quando este está ligado a diferentes aminoácidos, quais sejam a glicina, a fenilalanina e a prolina^(18,19,20). Neste último caso, é muito importante notar que a geometria final que obtivemos coincide com os dados espectroscópicos formando dois anéis em planos distintos, o que está em contradição com o campo de forças do programa Thor e de outros programas clássicos que reduzem esta geometria a um único plano. Aliás, o ácido ortoaminobenzóico, além da estrutura tetraédrica do grupo nitrila, possui um deslocamento em relação ao plano do anel benzênico de 7° que pode ser positivo ou negativo em seu estado fundamental, e é planar em seu estado excitado.

Assim, este trabalho resulta do uso conspícuo de métodos clássicos e quânticos de computação, isolados ou tomados em conjunto, de maneira a extrair os melhores resultados de cada um deles. Devemos notar que estas técnicas devem apresentar um caráter de complementaridade que muitas vezes não fica explicitada na descrição de suas partes. Este desacoplamento entre os programas clássicos e quânticos pode ser facilmente notado nos capítulos 2 e 3. Apresentamos no capítulo 2 o programa de mecânica e dinâmica molecular clássicos Thor⁽¹⁾ de maneira a mais sucinta possível, e

no capítulo 3 descrevemos os elementos necessários para um entendimento da linguagem utilizada na computação quântica. Assim, no capítulo 2 descrevemos as forças clássicas de ligação entre os átomos (força de Hooke), as forças torcionais entre dois átomos e as diedrais - necessárias para descrever a posição de três átomos - próprias e impróprias introduzidas quando se adota o modelo de átomo unido. Além das forças de ligação direta átomo a átomo, temos as forças que atuam em átomos não ligados diretamente que são as forças de van der Waals e a de Coulomb. Descrevemos o método *steepest-descent* de minimização, método padrão do programa Thor⁽¹⁾. Finalmente, mostramos como o programa realiza a dinâmica e uma rápida justificativa da correspondência desta com a termodinâmica mostra a coerência dos cálculos que o programa executa.

No capítulo 3, fazemos uma rápida descrição dos métodos e técnicas de computação quânticos. Ele é um pouco mais extenso para dar ao leitor um mínimo de entendimento do vocabulário utilizado e que por vezes transforma-se em um verdadeiro criptograma só decifrável por iniciados. Partimos da equação de Schrödinger, passamos pela aproximação de Born-Oppenheimer e a solução aproximada de Hartree-Fock, e chegamos finalmente nos métodos derivados desse cálculo. Descrevemos a seguir o método de cálculo de interações das configurações. Pretendemos mostrar a origem das bases de cálculo usados em química quântica e fechamos com alguns métodos semiempíricos conhecidos.

O capítulo 4 apresenta o cálculo da energia livre absoluta. Nele apresentamos uma descrição sumária dos métodos de cálculos existentes na literatura e um resumo de seus resultados. A seguir, apresentamos o método que implantamos no programa Thor, o cálculo da energia livre absoluta com a introdução do cristal de Einstein, os resultados da calibração com a molécula de fentanil e comparamos com os resultados apresentados no manual do programa *Discover*⁽⁸⁾. Mostramos que nossos resultados são comparáveis com os desse programa e efetuamos todos os cálculos necessários para a uma boa calibração. Descrevemos quais os tempos necessários para que a *inicialização* e a *termalização* conduzam a resultados coerentes e confiáveis. Como aplicação deste método calculamos e mapeamos a energia livre do triptofano e comparamos com o mapeamento de sua energia interna total.

No capítulo 6, a thiosemicarbazona^(14,21,22,23) é apresentada. Esta molécula, de grande interesse médico, apresenta duas famílias de confôrmeros. As medidas experimentais se mostravam incompatíveis com a teoria química vigente, e, aí mostramos que existe a possibilidade teórica de convalidar os resultados

experimentais, com uma possível rotação da molécula na ligação dupla, que no caso desta molécula pode ter um caráter intermediário entre uma ligação simples e uma ligação dupla. Procuramos também traçar as possíveis trajetórias da interconversão de um isômero a outro, ou melhor, de uma família isomérica em outra, além de adotarmos a hipótese de conversão intrafamiliar. Apresentamos a seguir a seqüência de estudo realizado na molécula do ácido ortoaminobenzóico isolado e quando ele forma compostos com a glicina, a prolina e a fenilalanina.

O texto está dividido em três partes. A parte I composto dos dois primeiros capítulos fornece informações básicas sobre programas clássicos e quânticos. A parte II apresenta o método de cálculo da energia livre absoluta e alguns resultados, e se compõe dos capítulos 4 e 5. Já a parte III mostra um conjunto composto de resultados dos cálculos da thiosemicarbazona e do ácido ortoaminobenzóico (o-Abz), usando cálculos quânticos semipíricos e *ab-initio*, e compõe o capítulo 6 e 7.

As figuras dos capítulos 2 e 3 foram obtidas no endereço eletrônico citado na bibliografia. As figuras de outros capítulos foram geradas pelo programa Rasmol, que é um programa de domínio público. Os gráficos foram gerados no programa Origin 4.0. Todas as figuras sofreram adaptações ou pequenas modificações pelos programas CorelDraw4.0 e Paint Shop Pro3.0. Os programas Thor e Geomop foram executados na HP7632 do grupo de Biofísica do IFUSP.

Parte I
CÃES E GATOS

it's raining cats and dogs

Acho que Schrödinger acreditava que os gatos tivessem sete vidas. Quando ele “superou” (?) esta superstição ajudou a criar a mecânica quântica. Há os que não gostam de gatos. Estes ficam com os ratos ou com os cães. Para a Física, a mecânica clássica é a muleta que tem nos tirado de muitas situações embaraçosas. Como fazem certos cães com seus donos. Quem não tem gato, salva-se com um cão. Como não superamos a briga entre cão e gato, ficamos com os dois. Apresentamos no capítulo dois uma descrição geral do nosso cão, ou melhor, do programa clássico Thor, a sua estrutura geral e de seu campo de forças. No capítulo três, fazemos um apanhado geral dos métodos de cálculo quântico. Assim, não precisamos optar entre ficar com o cão ou com o gato.

Capítulo 2

O PROGRAMA THOR

O programa Thor⁽¹⁾ é um programa de mecânica e dinâmica molecular clássicos, isto é, ele descreve as energias moleculares em termos de um conjunto de potenciais clássicos. As funções potenciais e os bancos de dados formam em conjunto o que denominamos “campo de forças”. O programa está baseado nas seguintes hipóteses:

- Os núcleos são tratados como partículas;
- As partículas atômicas são esféricas (os raios são obtidos a partir de medidas experimentais ou da teoria) e têm uma carga elétrica líquida (obtida da teoria);
- As interações entre as partículas vizinhas estão baseadas em potenciais harmônicos clássicos;
- As interações entre partículas não ligadas diretamente são executadas por meio dos potenciais de van der Waals e eletrostático;

Estas características são comuns a todos os programas de dinâmica molecular. O campo de forças usado pelo programa Thor é o mesmo do programa GROMOS⁽²⁾.

A representação das interações entre os átomos ligados por potenciais harmônicos simplifica os cálculos necessários para a descrição matemática da distribuição espacial das posições das moléculas e de seu comportamento dinâmico. Estes potenciais são o de “estiramento de ligação”, o de “deformação angular”, o de torção, e o “diédrico”. Além deles temos a interação

entre os átomos não ligados quimicamente de maneira direta que é descrita através dos potenciais de van der Waals e do eletrostático.

A *mecânica molecular* permite obter a energia mínima associada a uma molécula. Podemos encontrar todas as conformações energeticamente preferidas de uma molécula (especialmente os rotâmeros); isto equivale a localizar todos os mínimos de sua função energia. O objetivo da minimização na função energia é encontrar uma conformação de energia mínima a partir de uma conformação inicial usando o menor número de cálculos possíveis. Mínimos remotos não são detectados por este método porque isto exige um tempo para o movimento de superação da barreira de potencial. Entretanto, as energias aqui não tem nenhum significado como quantidades absolutas. São significativas somente diferenças de energia entre duas ou mais conformações.

Latu sensu, a *dinâmica molecular* trata do movimento molecular. O movimento é inerente a todo processo químico. Vibrações simples como o estiramento das ligações e vibrações angulares dão origem ao espectro infravermelho. Reações químicas, ligações hormônio-receptor e outros processos estão associados com muitas formas de movimentos intra e intermoleculares, e em princípio podem ser tratados via dinâmica molecular.

A força diretriz para os processos químicos é descrita pela termodinâmica. Os mecanismos de como acontecem os processos químicos são descritos pela cinética. A termodinâmica estabelece as relações energéticas entre os diferentes estados químicos, enquanto a seqüência ou taxa de eventos que acontecem para que as moléculas transitem entre seus vários estados possíveis é descrita pela cinética. Transições conformacionais e vibrações locais são objetos de estudos da dinâmica molecular. A dinâmica molecular atua nos graus de liberdade intramolecular de uma maneira análoga à minimização de energia. Os ingredientes básicos da dinâmica molecular são os cálculos das forças em cada átomo, e, a partir delas, as posições de cada átomo, em um período de tempo especificado (da ordem de picosegundo = 10^{-12} segundos).

A equação da energia no programa THOR tem a seguinte expressão:

$$V(\{r_i\}) = V(r_1, r_2, \dots, r_{N_{at}}) = \sum_{n=1}^{N_b} \frac{1}{2} K_{r_n} (r_n - r_{0n})^2 + \\ + \sum_{n=1}^{N_\theta} \frac{1}{2} K_{\theta_n} (\theta_n - \theta_{0n})^2 +$$

$$\begin{aligned}
& + \sum_{n=1}^{N_{\xi}} \frac{1}{2} K_{\xi_n} (\xi_n - \xi_{0n})^2 + \\
& + \sum_{n=1}^{N_{\tau}} \frac{1}{2} K_{\varphi_n} [1 + \cos(n_n \tau_n - \phi_n)] + \\
& + \sum_{i < j}^{N_{at}} \left[\frac{B_{ij}(i,j)}{r_{ij}^{12}} - \frac{A_{ij}(i,j)}{r_{ij}^6} + \frac{1}{4\pi\epsilon_0\epsilon} \frac{q_i q_j}{r} \right] \quad (2.1)
\end{aligned}$$

Os parâmetros que aparecem nesta equação K_{r_n} , b_{0n} , K_{θ_n} , θ_{0n} , K_{ξ_n} , ξ_{0n} , A , n_n , ϕ_n e, as constantes $B_{ij}(i,j)$ e $A_{ij}(i,j)$ e as cargas atômicas parciais são as utilizadas pelo campo de forças do GROMOS. N_b é o número de ligações entre primeiros vizinhos, N_{θ} o número de ângulos entre duas ligações químicas consecutivas, N_{ξ} o número de ângulos diédricos, N_{τ} é o número de ângulos torcionais e N_{at} é o número de átomos. A força em cada átomo pode ser calculada a partir da mudança na energia em duas posições consecutivas.

A seguir, passamos a descrever os termos da equação (2.1).

2.1 O Campo de Forças do THOR

2.1.1 Os potenciais harmônicos

a. O potencial de estiramento da ligação

Os comprimentos das ligações químicas oscilam próximos ao seu valor de equilíbrio, podendo ser descritos de forma aproximada pela função potencial de Hooke,

$$V_r = \frac{1}{2} \sum_{vizinhos} K_r (r - r_0)^2, \quad (2.2)$$

onde K_r é a constante de Hooke associada à ligação química específica e determina a intensidade da capacidade de restauração da mola e r_0 é a distância de equilíbrio.

Para cada par de ligação entre os átomos necessita-se destes dois parâmetros (por exemplo, C-C, C-H, O-C, etc.) Esta equação estima a energia associada com a vibração dos átomos em torno do comprimento de equilíbrio da ligação.

b. Os potenciais de deformação angular e o diédrico

1.O potencial de deformação angular A descrição destes potenciais também está baseada na lei de Hooke,

$$V_{\theta} = \frac{1}{2} \sum_{\text{ângulos}} K_{\theta} (\theta - \theta_0)^2$$

onde θ é o ângulo entre duas ligações químicas consecutivas e K_{θ} é o parâmetro de restituição angular e θ_0 é o ângulo de equilíbrio. Para cada tripleto de átomos que formam duas ligações consecutivas temos parâmetros com valores específicos (por exemplo, C-C-C, C-O-C, C-C-N, etc.). Esta equação estima a energia associada com a *vibração* dos átomos em torno do ângulo de equilíbrio das ligações.

2.O potencial diédrico No Thor define-se um terceiro potencial harmônico para manter a estrutura tridimensional de quatro átomos. Neste potencial, o átomo central (i) é ligado a três outros átomos (j,k,l) segundo a equação

$$V_{\xi} = \frac{1}{2} K_{\xi} (\xi - \xi_0)^2 \quad (2.3)$$

onde ξ é o ângulo entre o plano formado pelos átomos i, j, k e o plano formado pelos átomos j, k, l.

Novamente, a constante de Hooke está ligada a capacidade de restituição ao ângulo de equilíbrio entre estes planos. A introdução deste termo faz-se necessário quando utiliza-se a aproximação de átomo-unido (ver adiante, comentário sobre potencial eletrostático)

3.O potencial de torção A energia de torção é modelada por uma função periódica simples e é referida como Potencial Dieldral Próprio. Ele é usado para corrigir o valor da energia calculada e não representa um processo físico propriamente dito. Ver nas figuras 2.2 e 2.3 a representação esquemática deste potencial. A energia torcional representa a diferença de energia que deve ser adicionada ou subtraída dos outros termos energéticos para que a energia total concorde com a experiência ou com cálculos mais precisos da mecânica quântica. Toma-se normalmente o termo mais baixo da expansão:

$$V_{\phi} = A [1 + \cos(n\tau - \phi)] \quad (2.4)$$

onde A é a constante que define a altura da barreira de rotação, n é o número de mínimos para a torção de uma ligação específica, τ é o ângulo diedral para a ligação central em uma sequência de quatro átomos e ϕ é a defasagem no ângulo diedral. Para cada quarteto de átomos baseado em seu tipo temos um único conjunto de parâmetros (por exemplo, C-C-C-C, C-O-C-N, etc.). A ligação $\text{CH}_3\text{-CH}_3$, repete sua energia a cada 120° . A conformação *cis* de um ângulo diedral é por convenção tomado como zero.

Todas as constantes dos potenciais harmônicos aqui tratados são obtidos a partir de cálculos quânticos sobre pequenas moléculas acopladas a informações estruturais experimentais.

Os potenciais não-ligados

A energia dos átomos não ligados representa a soma da energia dos pares de todos os possíveis átomos não ligados, e é representada pela soma de dois termos, o termo de van der Waals e o termo de interação eletrostática.

a. O potencial de van der Waals O termo atrativo de van der Waals ocorre para curtas distâncias e se anula rapidamente quando os átomos interagentes se afastam de alguns angstroms. A repulsão ocorre quando a distância entre os átomos interagentes se torna levemente menor do que a soma de seus raios de contacto e cresce rapidamente para distâncias menores. Este potencial é descrito pela equação

$$V_{ij} = \sum_{\text{não ligados } ij} \left[\frac{B_{ij}}{r^{12}} - \frac{A_{ij}}{r^6} \right] \quad (2.5)$$

onde r é a distância de separação entre os átomos não ligados, conforme mostra a figura 2.4. O parâmetro A_{ij} pode ser obtido das medidas de polarizabilidade atômica ou de cálculos de mecânica quântica, enquanto que o parâmetro B_{ij} é obtido de dados cristalográficos de maneira a reproduzir a distância média de contato entre diferentes tipos de átomos em cristais de várias moléculas e determina o grau de "dureza" do átomo.

O potencial de van der Waals é responsável pela coesão nas fases condensadas e é importante na estabilização conformacional de macromoléculas.

b. O potencial eletrostático. A energia potencial eletrostática é modelada pela lei de Coulomb e é função das cargas dos átomos não ligados,

de sua distância interatômica e da constante dielétrica que leva em conta a atenuação da interação eletrostática pelo meio.

$$V_c = \frac{1}{4\pi\epsilon_0\epsilon} \sum_{\text{não ligados}}^{ij} \frac{q_i q_j}{r} \quad (2.6)$$

As cargas são originárias de duas fontes, a primeira, a existência de grupos ionizáveis que podem contribuir com cargas formais que aparecem distribuídas também por vários átomos vizinhos. Outra fonte de cargas elétricas advém da ligação química entre dois átomos diferentes que pode levar a uma deslocalização da nuvem eletrônica devido a eletronegatividades diferentes deles. Isto gera na maioria das ligações covalentes dipolos elétricos dependentes desta diferença de eletronegatividade entre os átomos componentes da ligação.

As cargas atômicas parciais podem ser calculadas para pequenas moléculas usando-se técnicas de cálculo quântico (*ab initio* ou semiempírico) e refinadas por dados experimentais. Em muitos campos de forças, o potencial torcional é calibrado para um método particular de cálculo de cargas, em geral desconhecido do usuário. A consistência do campo de forças pode ser invalidada pelo uso de um método diferente neste cálculo.

A deslocalização eletrônica nas ligações C-H é estimada em 0,05 elétron a 0,10 elétron no vácuo, tendo como resultado a não polaridade deste grupo. O GROMOS optou então pela aproximação de átomo-unido, somando as suas massas e reajustando o termo de van der Waals, introduzindo assim os "átomos" CH, CH₂, CH₃, com massas atômicas 13, 14 e 15 respectivamente e carga residual zero. Os átomos de hidrogênio representados explicitamente são os polares, ligados ao oxigênio ou ao nitrogênio, pois estes são os que apresentam maior eletronegatividade entre os componentes das biomoléculas.

2.1.2 Minimização da Energia Molecular

A energia é uma função dos graus de liberdade em uma molécula (ligações, ângulos, diedros). A procura das conformações energeticamente preferidas de uma molécula equivale matematicamente a localizar os mínimos de sua função energia potencial (equação 1.1). Uma superfície de energia apresenta picos (barreiras de energia), vales (energia mínima) e passagens (pontos de sela). As aproximações energéticas da mecânica molecular têm maior validade em torno dos mínimos. O menor mínimo tende a ser mais populado.

Se considerarmos uma molécula com 6 ligações significativas ao estudo, na qual cada uma possa girar de 30^0 até completar uma rotação de 360^0 teremos 2985984 conformações. Assim, para macromoléculas a dimensionalidade do sistema impossibilita a localização destes mínimos de energia com o uso deste método.

Normalmente todos os graus de liberdade em uma molécula são incluídos nos cálculos. A otimização da energia pode localizar as conformações de energia mínima, mas não pode ser usada para explorar todo o espaço configuracional. Um mínimo de energia pode ser caracterizado por uma pequena mudança na energia entre passos consecutivos ou por um zero no gradiente da função energia. Em geral o gradiente indica a convergência. Uma pequena mudança no gradiente é muito mais significativa do que uma diminuição na função energia. A escolha do algoritmo de minimização pode nestas condições ajudar a encontrar o mínimo de energia.

O THOR utiliza-se do método *steepest-descent*. Neste método, o incremento nas coordenadas de um átomo i , é dado na direção e sentido da força resultante sobre ele. Temos, então

$$\mathbf{F}_i = -\nabla(V\{\mathbf{r}_i\}) \quad (2.7)$$

e

$$\Delta\mathbf{r}_{i,n} = k_n \left(\frac{\mathbf{F}_{i,n}}{|\mathbf{F}_{i,n}|} \right) \quad (2.8)$$

onde k_n é um parâmetro de ajuste no tamanho do passo e $\mathbf{F}_{i,n}/|\mathbf{F}_{i,n}|$ é o vetor unitário na direção e sentido da força resultante sobre o átomo i no passo n . Assim no passo $n+1$

$$\mathbf{r}_{i,n+1} = \mathbf{r}_{i,n} + \Delta\mathbf{r}_{i,n} \quad (i = 1, 2, \dots, N_{at}) \quad (2.9)$$

onde n é o número de interações e $\Delta\mathbf{r}_{i,n}$ é o incremento nas coordenadas do átomo i , se transforma em:

$$\mathbf{r}_{i,n+1} = \mathbf{r}_{i,n} + k_n \left(\frac{\mathbf{F}_{i,n}}{|\mathbf{F}_{i,n}|} \right) \quad (2.10)$$

onde o ajuste de k_n acelera o processo de minimização na energia. Partindo de $k_0=0.1\text{\AA}$, o valor de cada passo será definido pela respectiva componente do vetor unitário em cada direção do espaço. Se o potencial total no passo n

$V_n(\{\mathbf{r}_i\})$ diminui em relação ao passo anterior, no passo seguinte k é aumentado. Se $V_n(\{\mathbf{r}_i\})$ aumenta, o sistema pode estar se afastando do mínimo, e k é diminuído fazendo-se, por exemplo, $k_{n+1} = k_n \times 1.2$ quando $V_n(\{\mathbf{r}_i\})$ aumenta, e $k_{n+1} = k_n \times 0.5$ quando $V_n(\{\mathbf{r}_i\})$ diminui.

Verifica-se em cada passo a diferença entre os valores de energia do passo com o anterior. Se esta diferença for menor que o fator de convergência estipulado, o processo é interrompido.

Outras técnicas de minimização de energia estão implementadas no programa. Uma delas é o *annealing* que é uma técnica de termalização e que permite o sistema transpor barreiras de potencial que em situações normais o sistema não ultrapassaria. Nesta técnica aumenta-se a “temperatura” do sistema até 900 K, por exemplo. Nesta situação, espera-se que as barreiras de energia potencial tenham sido superadas e, assim, o sistema pode alcançar um outro mínimo local, por meio de uma minimização de energia. Para maiores detalhes ver a tese de Pedro Geraldo Pascutti⁽¹⁾.

2.1.3 Dinâmica Molecular

Para a descrição mecânica do sistema microscópico precisamos especificar as posições e momentos das suas partículas. O hamiltoniano \mathbf{H} de um sistema pode ser escrito como a soma das energias cinética \mathbf{T} e potencial \mathbf{V} , em função das coordenadas generalizadas \mathbf{q}_i e dos momentos generalizados \mathbf{p}_i de todos os átomos do sistema N_{at} :

$$\mathbf{H}(\{\mathbf{q}_i, \mathbf{p}_i\}) = T(\{\mathbf{p}_i\}) + V(\{\mathbf{q}_i\}) \quad (2.11)$$

onde $q_i = \mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{N_{at}}$ e $\mathbf{p}_i = \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{N_{at}}$.

A energia potencial $V(\{\mathbf{q}_i\})$ contém os termos de interação inter e intramoleculares, de curto e longo alcance, e pode ser substituída pela função $V(\{\mathbf{r}_i\})$ da equação (1.1). A energia cinética será:

$$T(\{\mathbf{p}_i\}) = \sum_{i=1}^{N_{at}} \frac{\mathbf{p}_i^2}{2m_i} \quad (2.12)$$

onde m_i é a massa do átomo i .

Como a energia potencial independe das velocidades e do tempo, \mathbf{H} é igual a energia total do sistema e as equações de movimento de Hamilton conduzem às equações de Newton do movimento

$$\dot{\mathbf{P}}_i = m \ddot{\mathbf{r}}_i = - \frac{\partial V(\{\mathbf{r}_i\})}{\partial \mathbf{r}_i} = \mathbf{F}_i \quad (2.13)$$

e

$$\dot{\mathbf{r}}_i = \frac{\mathbf{P}_i}{m_i} = \mathbf{v}_i. \quad (2.14)$$

Aqui $\dot{\mathbf{r}}_i$ é a velocidade e a sua derivada é a aceleração e \mathbf{F}_i é a força sobre o átomo i . Na prática as trajetórias são obtidas computando-se as acelerações atômicas a partir das forças e massas. A seguir, são calculadas as velocidades a partir das acelerações. Finalmente, as posições são calculadas a partir destas velocidades. Isto acontece porque cada átomo recebe uma velocidade inicial que se ajusta à energia total do sistema, que por seu turno, é ditada pela temperatura de simulação desejada. A seguir, processa-se um lento "aquecimento" do sistema (inicialmente a temperatura de zero absoluto) e neste aquecimento espera-se a energia equilibrar-se entre os átomos constituintes. A partir da mudança de posição de cada átomo durante um período especificado de tempo, da ordem de picosegundos, obtém-se a trajetória desejada. O tempo deixa de ser contínuo e é discretizado em passos menores que o período de vibração dos átomos de hidrogênio. Isto é permitido porque uma observável macroscopicamente mensurável é na realidade a média temporal da observável (por exemplo, a energia potencial) no espaço de fase tomada sobre um longo intervalo de tempo:

$$A_{obs} = \langle A \rangle_{tempo} = \langle A(\Gamma(t)) \rangle_{tempo} = \lim_{t_{obs} \rightarrow \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A(\Gamma(t)) dt. \quad (2.15)$$

As equações que governam esta evolução temporal, as equações de movimento de Newton, não podem ser estendidas a um tempo infinito de computação, mas devem satisfazer a média sobre um longo tempo finito t_{obs} . Isto é exatamente o que faz a dinâmica molecular. As equações de movimento são resolvidas passo a passo, isto é, um grande número finito φ_{obs} de passos temporais, de comprimento $\delta t = t_{obs}/\varphi_{obs}$. Neste caso, podemos reescrever a equação acima como

$$A_{obs} = \langle A \rangle_{tempo} = \frac{1}{\varphi_{obs}} \sum_{\varphi=1}^{\varphi_{obs}} A(\Gamma(t)). \quad (2.16)$$

Na somatória, φ simplesmente estabelece um índice que percorre a sucessão de passos temporais.

Desta maneira evita-se saltos abruptos e garante-se uma continuidade da energia potencial em função das posições das partículas e estas variam suavemente com o tempo. Portanto, pela integração numérica das equações de Newton obtém-se uma série de subestados entre dois estados, um inicial e um final, separados por um pequeno intervalo de tempo (por exemplo, 1 femtosegundo).

$$\begin{array}{ccccccc} \text{estado inicial} & & & & & & \text{estado final} \\ \mathbf{r}_i^1 & \dashrightarrow & \mathbf{r}_i^{1a} & \dashrightarrow & \rightarrow & \rightarrow & \mathbf{r}_i^2 \\ t & & t+\delta t & & & & \end{array} \quad (2.17)$$

As posições atômicas em t são usadas para prever as posições atômicas em $t + \delta t$. As posições em $t + \delta t$ são utilizadas para se obterem as posições em $t + 2\delta t$ e assim sucessivamente.

O algoritmo usado para cálculos das trajetórias baseadas nas leis de Newton é o de *leapfrog*. O nome é derivado do fato de que as informações da velocidade e da posição são derivadas de maneira alternada sucessivamente a intervalos de meio(1/2) passo no intervalo de tempo.

$$\begin{aligned} m\ddot{\mathbf{r}}_i(t) &= -\frac{\partial V(\{\mathbf{r}_i\})}{\partial \mathbf{r}_i} \\ \dot{\mathbf{r}}_i(t + \delta t/2) &= \dot{\mathbf{r}}_i(t - \delta t/2) + \ddot{\mathbf{r}}_i(t) \delta t \\ \mathbf{r}_i(t + \delta t) &= \mathbf{r}_i(t) + \dot{\mathbf{r}}_i(t + \delta t) \delta t \end{aligned}$$

O processo visualizado no quadro acima é executado continuamente de forma cíclica a partir do cálculo da primeira linha até a terceira linha e desta se retorna a primeira linha reiniciando o processo de cálculo. A vantagem deste método está no fato das posições dependerem das velocidades, o que torna possível acoplar o sistema a um banho térmico por meio das correções das velocidades. Controlar as velocidades significa, em última instância, controlar a energia total do sistema, pois além da energia potencial ser função das posições, agora depende também da velocidade.

A temperatura do sistema é a energia cinética média sobre o número de passos:

$$\frac{3}{2} N_{at} k_b T = \left\langle \sum_{i=1}^{N_{at}} \frac{1}{2} m_i \dot{\mathbf{r}}_i^2(t) \right\rangle_N \quad (2.18)$$

N_{at} é o número de átomos do sistema. O número de passos varia de 20 a 50 passos, dependendo do número de átomos do sistema. O acoplamento a um banho térmico se dá a partir da comparação da temperatura T a cada ciclo de N passos, com uma temperatura de referência T_0 . Se a diferença entre T e T_0 estiver fora do desvio estipulado, as novas velocidades são corrigidas pelo fator:

$$\dot{\mathbf{r}}_i(t + \delta t/2) = \dot{\mathbf{r}}_i(t - \delta t/2) \left[(T_0/T)^{1/2} \right]. \quad (2.19)$$

Para termos um *ensemble* estatisticamente válido, precisamos ter um sistema em que ou N , V e E sejam constantes, o ensemble microcanônico, ou N , V e T constantes, o ensemble canônico. O programa Thor permite trabalhar nestes dois *ensembles*. Na sua configuração básica pode utilizar a temperatura constante e permitir que a energia flutue, ou inversamente, podemos fixar a energia e permitir que a temperatura flutue.

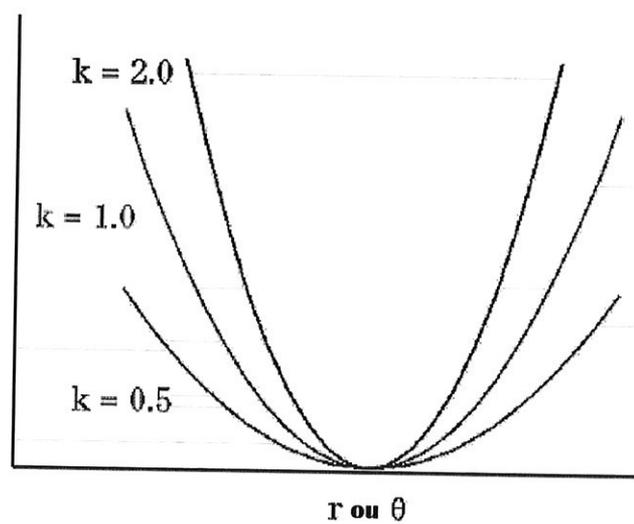


Figura 2.1: O potencial de Hooke.

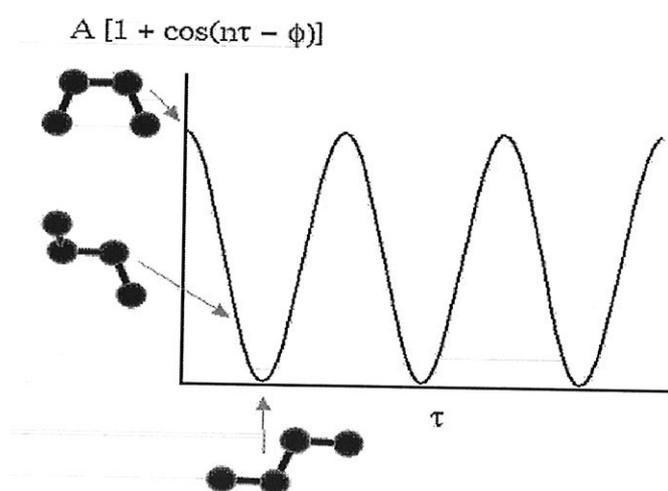


Figura 2.2: O potencial de torção.

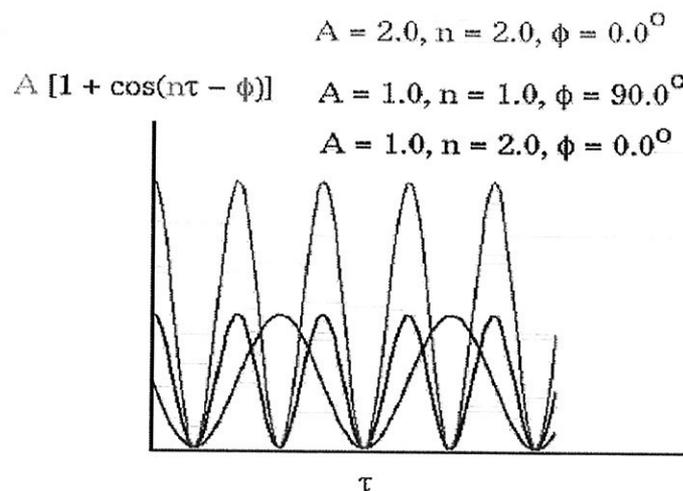


Figura 2.3: Variação do potencial de torção com os seus parâmetros.

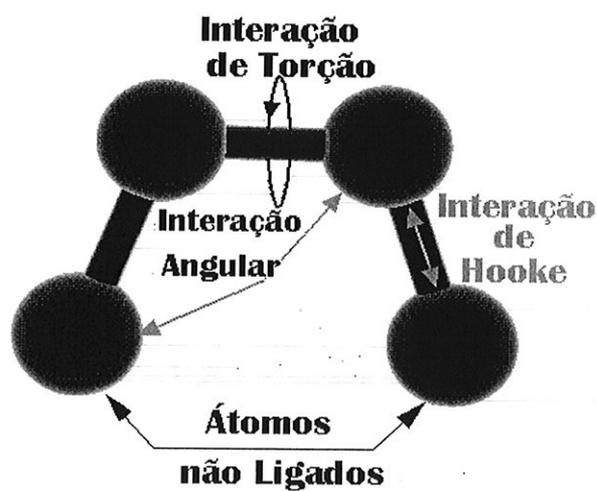


Figura 2.4: Átomos e interações.

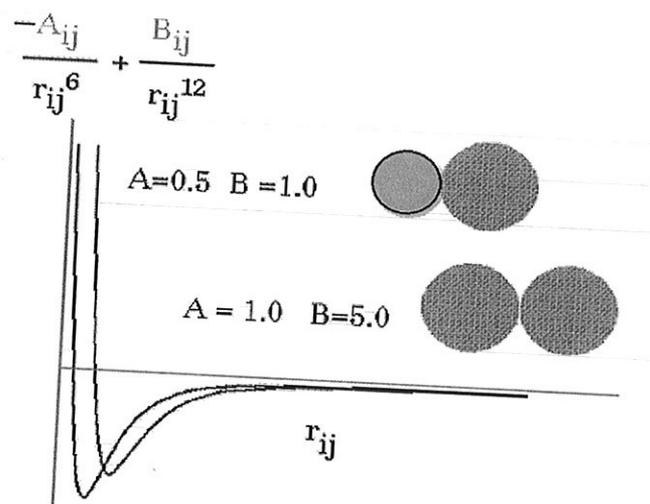


Figura 2.5: Variação do potencial de van der Waals com os parâmetros.

Capítulo 3

Modelagem Molecular Quântica

Muitos aspectos da estrutura molecular e da dinâmica podem ser modelados classicamente. Os campos de forças clássicos estão baseados em resultados experimentais. As médias destes resultados são tomadas em relação a um grande número de moléculas. Devido a esta extensividade na média, os resultados podem ser bons para os sistemas que se aproximam daqueles que foram utilizados como padrão. Se queremos conhecer algo além da estrutura ou propriedades derivadas da superfície da energia potencial, principalmente aquelas que dependam diretamente da distribuição da densidade eletrônica, necessitamos da aproximação quântica.

Trataremos aqui apenas dos aspectos importantes para a modelagem molecular. Basicamente, um sistema químico pode ser descrito por uma equação de onda encontrada a partir da resolução da equação de Schrödinger. Esta equação relaciona os estados estacionários do sistema e suas energias associadas com o operador hamiltoniano. Ou melhor, este pode ser visto como o instrumento para a obtenção das energias associadas às funções de onda que descrevem as posições dos núcleos e dos elétrons do sistema. Como a solução exata das equações de Schrödinger é possível somente para átomos pequenos, recorre-se a aproximações. A aproximação é denominada "*ab-initio*" quando as informações usadas são as constantes fundamentais da natureza, como a constante de Planck, a massa do elétron, etc., necessárias para se chegar a resultados numéricos. Como os métodos *ab-initio* exigem uso de uma grande capacidade computacional, introduzem-se parâmetros empíricos nos métodos quânticos. Estes métodos são denominados "*semiempíricos*" e as suas funções de onda eletrônicas são razoáveis quando comparadas com os métodos *ab-initio*.

3.1 A Equação de Schrödinger

A solução da equação de Schrödinger

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

nos fornece as funções de onda e as energias dos estados estacionários. O operador hamiltoniano \mathbf{H} é composto de três partes: a energia cinética dos núcleos, a energia cinética dos elétrons e a energia potencial dos núcleos e dos elétrons. Ψ , uma das soluções da equação de autovalores, é uma função de onda que depende das coordenadas dos elétrons e dos núcleos. Portanto a equação 3.1 pode ser reescrita como:

$$\mathbf{H}\Psi_{e,n} = E\Psi_{e,n} \quad (3.2)$$

e

$$\mathbf{H} = \mathbf{T}_n + \mathbf{T}_e + \mathbf{V}_{e,n} \quad (3.3)$$

onde e significa elétrons e n núcleos.

As seguintes aproximações simplificam os cálculos:

- desprezam-se todos os efeitos relativísticos;
- usa-se a aproximação de Born-Oppenheimer, isto é, a separação do movimento dos núcleos e dos elétrons;
- usa-se a aproximação orbital, ou seja, confina-se os elétrons em certas regiões do espaço.

A aproximação de Born-Oppenheimer supõe que os elétrons sendo muito mais leves do que os núcleos podem acompanhar o movimento nuclear facilmente. Assim podemos separar o movimento nuclear do eletrônico de maneira que a função de onda total é um produto das duas partes

$$\Psi_{e,n} = \Phi_n \Psi_e \quad (3.4)$$

A parte *eletrônica* da equação é resolvida para um conjunto discreto de posições nucleares fixas pois assumimos que a energia potencial entre os núcleos é quase uma constante e a energia cinética dos núcleos é praticamente nula em comparação ao movimento eletrônico. Devemos notar que

ao desprezarmos a variação da massa com a velocidade desprezamos efeitos relativísticos que se tornam importantes no estudo da estrutura dos átomos pesados. Anulamos também outro efeito relativístico, o acoplamento spin-órbita. Neste caso, estamos desprezando a interação do mesmo nome que é importante em uma análise rotacional/vibracional. Com estas aproximações, a parte eletrônica da equação de Schrödinger é

$$\mathbf{H}_e(\mathbf{R}_n) \Psi(\mathbf{r}_e) = E_e(\mathbf{R}_n) \Psi(\mathbf{r}_e). \quad (3.5)$$

A função de onda eletrônica depende somente parametricamente da configuração nuclear. Se calcularmos a equação acima para n valores das posições \mathbf{R}_n dos núcleos obtemos a curva da energia potencial ao longo do movimento dos núcleos. Para não carregarmos termos desnecessários, introduzimos as unidades atômicas:

e	=	1 carga do elétron	
m	=	1 massa do elétron	
\hbar	=	constante de Planck/ 2π	
1 bohr	=	a_0	= 0.529 Å
1 hartree	=	$\frac{e^2}{a_0}$	= $4.3598 \cdot 10^{18} J = 627.51 \text{ kcal/mol}$.

A função hamiltoniana eletrônica se reduz a¹

$$\mathbf{H}_e = -\frac{1}{2} \sum_{i=1}^n \Delta_i - \sum_{i=1}^n \sum_{A=N}^N \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_{i < j}^n \frac{1}{r_{ij}} \quad (3.6)$$

onde i descreve o elétron, A representa o tipo do núcleo e Δ é o operador de Laplace, Z_A é o número atômico, R é a distância ao núcleo e os r referem-se a distâncias dos elétrons. A função hamiltoniana contém dois termos que agem em um elétron, a energia cinética e a atração elétron-núcleo, e um termo que descreve a repulsão aos pares dos elétrons. A energia total no modelo de Born-Oppenheimer é obtida adicionando a energia de repulsão nuclear à energia eletrônica:

$$E_{tot} = E_n + E_e \quad (3.7)$$

onde

¹Como trataremos apenas da parte eletrônica, todos os índices que aparecem nesta e nas equações subsequentes deste capítulo referem-se unicamente aos elétrons.

$$E_n = \sum_{A < B}^N \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}. \quad (3.8)$$

A energia total define uma hypersuperfície de energia potencial que pode ser usada subsequentemente para resolver a equação de Schrödinger do movimento nuclear, dada por:

$$[\mathbf{T}_n + E(\mathbf{R}_n)] \Phi(\mathbf{R}_n) = \varepsilon_i \Phi(\mathbf{R}_n). \quad (3.9)$$

3.2 O Campo Auto-Consistente de Hartree-Fock

A repulsão de dois elétrons depende das coordenadas de ambos ao mesmo tempo. É o gargalo computacional prático e pode ser resolvido somente para sistemas muito pequenos.

Para contornar este problema introduz-se a aproximação de partícula independente: a interação de cada elétron com todas as outras é tratada de uma forma média, isto é:

$$\mathbf{H}^2 = \sum_{i < j}^n \mathbf{H}_{ij}^2 = \sum_i^n \mathbf{V}_i^{medio}. \quad (3.10)$$

Com esta aproximação, a equação de Schrödinger, que dependia inicialmente das coordenadas espaciais e de spin \mathbf{x} de todos os elétrons, pode ser reduzida ao conjunto de equações:

$$\sum_{i=1}^n (\mathbf{H}_i^1 + \mathbf{V}_i^{medio}) \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = E \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n). \quad (3.11)$$

Vamos trabalhar em orbitais de spin de um elétron como usa-se denominar $\Phi_i(\mathbf{x}_i)$.

O problema é que o potencial de cada elétron devido a todos os outros tem que ser conhecido mas, inicialmente não se conhece nenhum deles. Na prática orbitais tentativos são usados e são modificados iterativamente até uma solução auto-consistente SCF ("Self-Consistent Field") ser obtida, que é expressa como uma solução das equações de Hartree-Fock:

$$\left(\mathbf{H}_i^1 + \mathbf{V}_i^{\text{medio}} \right) \Phi_i = \varepsilon_i \Phi_i \quad (3.12)$$

O objetivo é encontrar a “melhor” função de onda de muitos elétrons .

Muitas técnicas têm sido desenvolvidas para acelerar a convergência, e resolver o problema, mesmo em casos difíceis. As dificuldades costumam acontecer com sistemas com uma estrutura incomum. As energias são interpretadas como energias orbitais. As energias orbitais têm uma interpretação física simples: elas fornecem o acréscimo de energia necessário para levar o elétron fora do orbital molecular, que corresponde à negativa do potencial de ionização experimentalmente observável .

Sendo uma solução da equação eletrônica de Schrödinger, a função de onda deve ser normalizada.

O princípio de Pauli estabelece que as funções de onda eletrônicas devem trocar o sinal quando duas coordenadas eletrônicas intercambiam:

$$\Psi_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_k, \dots, \mathbf{x}_m) = -\Psi_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, \dots, \mathbf{x}_i, \dots, \mathbf{x}_m). \quad (3.13)$$

Para um sistema de dois elétrons, desprezando-se as interações entre eles, a probabilidade de encontrar um dos elétrons em uma dada posição, isto é, Ψ^2 não depende da posição do outro. Este fato pode ser representado matematicamente usando o produto das funções de onda de cada elétron. Obtemos funções de onda simétrica e anti-simétrica. Estas funções dependem das posições espaciais dos elétrons. A função de onda completa é composta de uma função espacial e uma função de spin, também simétrica ou anti-simétrica. Às funções de onda assim contruídas denominaremos spin-orbital. Somente as funções anti-simétricas destas funções spin-orbital têm significado físico e resultam da combinação de funções do espaço e de spin simetricamente opostos. As energias obtidas das funções simétricas não têm correspondência espectroscópica e não obedecem ao princípio de exclusão de Pauli. Sejam as funções $\varphi_1(x_1, y_1, z_1, \sigma_1)$ e $\varphi_2(x_2, y_2, z_2, \sigma_2)$ (onde o sigma representa tanto o estado alfa como o estado beta do spin) de dois elétrons. O produto delas pode ser representada por

$$\Psi_n = \frac{1}{\sqrt{2}} \{ \varphi_1(x_1, \dots, \sigma_1) \varphi_2(x_2, \dots, \sigma_2) - \varphi_2(x_1, \dots, \sigma_1) \varphi_1(x_2, \dots, \sigma_2) \} \quad (3.14)$$

ou

$$\Psi_n = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_1(x_1, \dots, \sigma_1) & \varphi_1(x_2, \dots, \sigma_2) \\ \varphi_2(x_1, \dots, \sigma_1) & \varphi_2(x_2, \dots, \sigma_2) \end{vmatrix} \quad (3.15)$$

Esta equação pode ser generalizada para dar o determinante de Slater de n elétrons

$$\Psi_n = n^{-1/2} |\varphi_1 \varphi_2 \dots \varphi_n|. \quad (3.16)$$

Vamos assumir uma função de onda da forma do determinante de Slater e encontrar uma expressão para a energia. Uma propriedade importante do método **SCF** é que suas soluções satisfaçam o Princípio Variacional, que estabelece que o valor esperado da energia obtida a partir de uma função de onda inexata é sempre maior do que a energia exata, isto é,

$$E_\Psi = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_{\Psi_{\text{exata}}}. \quad (3.17)$$

Como consequência deste princípio, a energia mais baixa está associada com a melhor função de onda aproximada e a minimização da energia é equivalente à otimização de função de onda do estado fundamental. A forma matemática será a de um simples determinante. Os melhores orbitais são as autofunções dos autovalores da equação de um elétron (equação 3.12).

Para um cálculo Hartree-Fock, as energias dos determinantes de Slater são facilmente expressas em integrais de um e dois elétrons. Para o estado fundamental, para camada fechada, temos:

$$E = \sum_i^{\text{ocupado}} H_{ii}^1 + \frac{1}{2} \sum_{i,j}^{\text{ocupado}} [\langle ii|jj \rangle - \langle ij|ij \rangle] \quad (3.18)$$

para os spin-orbitais, e onde

$$H_{ii}^1 = \langle i | \mathbf{H}_i^1 | i \rangle = \int \varphi_i^*(x_i) \mathbf{H}_i^1 \varphi_i(x_i) dx_i \quad (3.19)$$

são os elementos diagonais da matriz \mathbf{H} , e

$$\langle ij|kl \rangle = \int \int \varphi_i^*(x_1) \varphi_j(x_1) \frac{1}{r_{12}} \varphi_k^*(x_2) \varphi_l(x_2) dx_1 dx_2. \quad (3.20)$$

As somatórias percorrem todos os orbitais-spin ocupados. O segundo termo é a integral de 2 elétrons anti-simetrizada. A integral $\langle ij|ij \rangle$ descreve a repulsão entre dois elétrons localizados cada um em um orbital e é denominado integral

de Coulomb, e a integral $(ii|jj)$ é a integral de troca e é um termo puramente quântico, isto é, não possui análogo clássico.

Em muitos casos é vantajoso obrigar o par de elétrons com spins opostos a ocupar o mesmo orbital espacial. Isto conduz ao método Hartree-Fock restrito (**RHF**) em oposição ao método irrestrito (**UHF**). A vantagem deste método é que os momentos magnéticos associados com o spin do elétron se cancelam exatamente para o par de elétrons no mesmo orbital espacial, de maneira que a função de onda **SCF** é uma autofunção dos operadores de spin. Na prática, **RHF** é mais usado para sistemas de camadas fechadas, e **UHF** para sistemas de camadas abertas^(67,68).

Definindo

$$\mathbf{h}(i) = -\frac{1}{2}\nabla_i^2 + \sum_A^M \frac{1}{r_{iA}}$$

como a função hamiltoniana de um elétron, **H** o operador soma de operadores de um e dois elétrons pode ser escrito como

$$\mathbf{H} = \sum_i \mathbf{h}(i) + \sum_{i>j} \frac{1}{r_{ij}}$$

As integrais de um elétron com os spins alfa e beta resultam em $2h_{ii}$ e para a camada fechada teremos $2h_{ii} + h_{jj}$. Introduzindo dois novos operadores, o "operador de coulomb" J e K "operador de troca", isto é as integrais de dois elétrons contribuem J para cada par de elétrons, e -K para cada par de spin paralelo, a energia será reescrita como uma soma para orbitais:

$$E = 2 \sum_i^{\text{ocupado}} f_i h_{ii} + \sum_{i,j}^{\text{ocupado}} \{ \alpha_{ij} (ij|ij) + \beta_{ij} (ii|jj) \} \quad (3.21)$$

$$= 2 \sum_i^{\text{ocupado}} f_i h_{ii} + \sum_{i,j}^{\text{ocupado}} \{ \alpha_{ij} J_{ij} + \beta_{ij} K_{ij} \} \quad (3.22)$$

com

$$f_i = \begin{cases} 1 & \text{se } i \text{ está duplamente ocupado} \\ 1/2 & \text{se } i \text{ está com ocupação simples} \end{cases}$$

$$\alpha_{ij} = \begin{cases} 2 & \text{se } i \text{ e } j \text{ estão duplamente ocupados} \\ 1 & \text{se } i \text{ ou } j \text{ estiver duplamente ocupado e} \\ & \text{o outro com ocupação simples} \\ 1/2 & \text{se } i \text{ e } j \text{ estiverem com ocupação simples} \end{cases}$$

$$\beta_{ij} = \begin{array}{l} -1 \text{ se } i \text{ e } j \text{ forem duplamente ocupados} \\ -1/2 \text{ se } i \text{ ou } j \text{ tiver ocupação simples e o outro ocupado duplamente} \\ -1/2 \text{ se } i \text{ e } j \text{ tiverem ocupação simples, com spins paralelos} \\ 1/2 \text{ se } i \text{ e } j \text{ tiverem ocupação simples, com spins opostos.} \end{array}$$

Estes resultados são obtidos do cálculo dos valores esperados quânticos, onde as funções de onda de uma molécula são representadas como o determinante das funções de onda dos elétrons individuais e o resultado do número ímpar de permutações resulta no sinal negativo da equação (3.23). Dependendo de como os estados eletrônicos individuais são preenchidos, tendo spins paralelos ou antiparalelos, esta equação assume formas explícitas diferentes para camadas abertas e fechadas resumidas na equação (3.24). Assim, na camada fechada temos $N/2$ spins para cima e $N/2$ spins para baixo (spins “up” e “down” ou α e β que usaremos indiferentemente) e a somatória dupla $\sum_{i,j}^{\text{ocupado}}$ em (3.23) exige um fator 4 no termo coulombiano e um fator 2 para o termo de troca, já que o número quântico de spin pertencentes a i e j são os mesmos, sendo uma vez para cima e outra para baixo. Para o caso da última camada estar aberta, temos as interações na camada fechada, na camada aberta e a interação entre os elétrons das camadas aberta e fechada.

3.2.1 Orbitais Moleculares-A forma funcional da função de onda

Na aproximação acima, os orbitais moleculares (MO) são construídos como combinação linear de orbitais de base a partir de uma base finita de orbitais atômicos comumente centrados nos vários núcleos. Eles têm um único centro e têm características *atômicas*. A construção de orbitais moleculares na aproximação combinação linear de orbitais atômicos (linear combination of atomic orbitals) -LCAO é o método mais característico das técnicas de cálculo em uso. Esta aproximação é um exemplo típico do uso do método de bases finitas.

Seja ϕ_i o orbital molecular de N funções de base:

$$\phi_i = \sum_{\mu=1}^N c_{i\mu} \varphi_{\mu} \quad (3.23)$$

onde φ_{μ} é um orbital atômico molecular e $c_{i\mu}$ são os coeficientes de contração que transformam uma base em outra. Temos que compor as energias

eletrônicas nas bases dos orbitais atômicos. Iniciando o cálculo da energia Hartree-Fock camada fechada em termos de orbitais moleculares, temos

$$\begin{aligned} E &= 2 \sum_i^{\text{ocupado}} h_{ii} + \sum_{i,j}^{\text{ocupado}} \{2J_{ij} - K_{ij}\} \\ &= 2 \sum_i^{\text{ocupado}} h_{ii} + \sum_{i,j}^{\text{ocupado}} \{2(ij|ij) - (ii|jj)\}. \end{aligned} \quad (3.24)$$

Como $\langle \phi_i | h | \phi_i \rangle = \int \phi_i^*(\mathbf{r}_1) h(1) \phi_i(\mathbf{r}_1) d\mathbf{r}_1$ e $\phi_i = \sum_{\mu=1}^N c_{i\mu} \varphi_{\mu}$

$$\begin{aligned} \int \int \phi_i^*(\mathbf{r}_1) h(1) \phi_i(\mathbf{r}_1) d\mathbf{r}_1 &= \int \sum_{\mu=1}^N c_{i\mu}^* \varphi_{\mu}^*(\mathbf{r}_1) h(1) \sum_{\nu=1}^N c_{i\nu} \varphi_{\nu}(\mathbf{r}_1) d\mathbf{r}_1 \\ &= \sum_{\mu=1}^N \sum_{\nu=1}^N c_{i\mu}^* c_{i\nu} \int \varphi_{\mu}^*(\mathbf{r}_1) h(1) \varphi_{\nu}(\mathbf{r}_1) d\mathbf{r}_1 \\ &= \sum_{\mu=1}^N \sum_{\nu=1}^N c_{i\mu}^* c_{i\nu} (\phi_{\mu} | h | \phi_{\nu}) \\ &= \sum_{\mu=1}^N \sum_{\nu=1}^N c_{i\mu}^* c_{i\nu} h_{\mu\nu} \end{aligned} \quad (3.25)$$

onde $(\phi_{\mu} | h | \phi_{\nu})$ é uma integral de um elétron sobre orbitais atômicos. Seguindo um procedimento análogo ao anterior, Roothann, C.C.J.^(67,68) expandiu as integrais de dois elétrons em bases dos orbitais moleculares:

$$(ij|kl) = \sum_{\mu=1}^N \sum_{\nu=1}^N \sum_{\lambda=1}^N \sum_{\sigma=1}^N c_{i\mu}^* c_{j\nu}^* c_{k\lambda} c_{l\sigma} (\mu\nu|\lambda\sigma) \quad (3.26)$$

Todas estas integrais dos orbitais atômicos podem ser calculadas, guardadas e usadas quando necessitarmos para o cálculo da energia eletrônica. Neste caso, onde cada orbital molecular está ocupado por dois elétrons, um com spin para cima e outro para baixo, a energia fica

$$\begin{aligned} E &= 2 \sum_i \sum_{\mu=1}^N \sum_{\nu=1}^N c_{i\mu} c_{i\nu} h_{\mu\nu} \\ &+ \sum_i \sum_j \sum_{\mu=1}^N \sum_{\nu=1}^N \sum_{\lambda=1}^N \sum_{\sigma=1}^N c_{i\mu}^* c_{j\nu}^* c_{i\lambda} c_{j\sigma} [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)]. \end{aligned} \quad (3.27)$$

$$(3.28)$$

Definindo

$$\mathbf{S}_{\mu\nu} = \langle \phi_{\mu} | \phi_{\nu} \rangle = \langle \mu | \nu \rangle \quad (3.29)$$

a integral de recobrimento dos orbitais ϕ_{μ} e ϕ_{ν} . A variação de cada orbital molecular ϕ_i se traduz pela variação de cada coeficiente $c_{i\mu}$ de $\delta_{i\mu}$. A condição de ortogonalidade obriga a modificação de todos os outros orbitais moleculares. Obtemos então

$$\mathbf{F}_{\mu\nu} = \langle \mu | \mathbf{F} | \nu \rangle = \langle \mu | \mathbf{H} | \nu \rangle + \sum_i \sum_{\lambda=1}^N \sum_{\sigma=1}^N c_{i\lambda} c_{i\sigma} [2(\mu\nu | \lambda\sigma) - (\mu\sigma | \lambda\nu)] \quad (3.30)$$

que podem ser rescritas como

$$\mathbf{F}_{\mu\nu} = \langle \mu | \mathbf{F} | \nu \rangle = \langle \mu | \mathbf{H} | \nu \rangle + \sum_{\lambda,\sigma} \mathbf{P}_{\lambda\sigma} \left[(\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\lambda | \nu\sigma) \right] \quad (3.31)$$

onde \mathbf{P} é a matriz densidade

$$\mathbf{P}_{\lambda\sigma} = 2 \sum_i^{\text{ocupados}} c_{i\lambda} c_{i\sigma} \quad (3.32)$$

A variação dos coeficientes $c_{i\mu}$, com as condições de ortonormalização dos orbitais

$$\langle \phi_i | \phi_j \rangle = \sum_{\mu=1}^N \sum_{\nu=1}^N c_{i\mu} c_{j\nu} \mathbf{S}_{\mu\nu} = N_{ij} \quad (3.33)$$

deve se anular

$$\delta N_{ij} = 2 \sum_{\mu=1}^N \sum_{\nu=1}^N \delta c_{i\mu} c_{j\nu} \mathbf{S}_{\mu\nu} = 0. \quad (3.34)$$

O teorema variacional nos diz que os valores dos coeficientes de ϕ_i são ótimos quando conduzem a uma variação $\delta E = 0$ se usarmos a variação δN_{ij} . Teremos n elementos diagonais e os elementos não diagonais serão $\frac{n^2+n}{2}$, ou seja,

$$\delta E - \sum_{i=1}^N \sum_{j=1}^N \lambda_{ij} (2\delta N_{ij}) = 0 \quad (3.35)$$

onde λ_{ij} são os multiplicadores de Lagrange. Portanto, fazendo as devidas substituições, obtemos

$$\sum_{i=1}^N \sum_{\mu=1}^N \delta c_{i\mu} \sum_{\sigma=1}^N \left[c_{\mu\sigma} \mathbf{F}_{\mu\sigma} - \sum_{j=1}^N \sum_{\lambda=1}^N \lambda_{ij} c_{j\lambda} \mathbf{S}_{\mu\nu} \right] = 0 \quad (3.36)$$

Se para cada valor de índice i impormos que os multiplicadores devem verificar a relação

$$\sum_{\sigma} c_{i\sigma} \mathbf{F}_{\zeta\sigma} - \sum_{j=1}^N \sum_{\lambda=1}^N \lambda_{ij} c_{j\lambda} \mathbf{S}_{\zeta\nu} = 0 \quad (3.37)$$

para um valor qualquer do índice ζ . A relação anterior deve ser independente da variação dos coeficientes $c_{i\zeta}$ para todo orbital Φ_i , o que equivale a

$$\sum_{\sigma} c_{i\sigma} \mathbf{F}_{\mu\nu} - \sum_{j=1}^N \sum_{\lambda=1}^N \lambda_{ij} c_{j\lambda} \mathbf{S}_{\zeta\nu} = 0 \quad (3.38)$$

para todo $\mu \neq \zeta$. Temos nN equações homogêneas com $\frac{n(n+1)}{2}$ multiplicadores definidos por um fator determinado pelas n condições de normalização dos orbitais moleculares. São $nN + n$ incógnitas (os multiplicadores de Lagrange) ligadas por $nN + n$ equações. Podemos escolher arbitrariamente $\frac{n(n+1)}{2} - n = \frac{n(n-1)}{2}$ incógnitas do sistema (que é o número de multiplicadores λ_{ij} para $i \neq j$). Podemos, portanto, fazer:

$$\lambda_{ij} = \varepsilon_i \delta_{ij} \quad (3.39)$$

e o sistema de equações se escreve reagrupando a somatória σ e λ em

$$\sum_{\nu=1}^N c_{i\nu} [\mathbf{F}_{\mu\nu} - \varepsilon_i \mathbf{S}_{\mu\nu}] = 0. \quad (3.40)$$

Este sistema contém N equações lineares e homogêneas em $c_{i\nu}$ e são conhecidas como equações de Roothaan. O caráter linear e homogêneo permite soluções não nulas onde os multiplicadores de Lagrange ε_i são escolhidos entre as raízes da equação

$$\det [\mathbf{F}_{\mu\nu} - \varepsilon_i \mathbf{S}_{\mu\nu}] = 0. \quad (3.41)$$

Conhecidos todos os coeficientes $c_{i\mu}$, podemos calcular todos os elementos $\mathbf{F}_{\mu\nu}$ da matriz quadrada \mathbf{F} construída com estes elementos; \mathbf{S} é a matriz quadrada construída a partir do reagrupamento das integrais de recobrimento e onde os elementos da diagonal principal são todos iguais a 1; a matriz coluna \mathbf{C} é o arranjo dos $c_{i\mu}$ correspondentes aos diversos orbitais moleculares e finalmente a matriz diagonal ε com os autovalores ε_i correspondentes. Ou, sinteticamente,

$$\mathbf{FC} = \mathbf{SC}\varepsilon. \quad (3.42)$$

A matriz \mathbf{S} introduz uma complicação devido a não ortogonalidade da base dos orbitais escolhidos. A ortogonalização pode ser feita utilizando-se do fato de \mathbf{F} e \mathbf{S} serem matrizes hermitianas. Podemos escrever

$$\mathbf{F}^t = \mathbf{S}^{-1/2} \mathbf{F} \mathbf{S}^{-1/2} \quad (3.43)$$

e obtermos

$$\mathbf{F}^t (\mathbf{S}^{-1/2} \mathbf{C}) = (\mathbf{S}^{-1/2} \mathbf{C}) \varepsilon. \quad (3.44)$$

Como ε é diagonal, a matriz transformada de Fock diagonalizada fornece um conjunto de coeficientes moleculares orbitais transformados, $\mathbf{C}^\dagger = \mathbf{S}^{-1/2} \mathbf{C}$. A matriz $\mathbf{S}^{-1/2} \mathbf{C}$ pode ser transformada de volta para dar o verdadeiro coeficiente matricial do orbital molecular \mathbf{C} . A matriz densidade para estes coeficientes é formada pelo produto $\mathbf{D} = \mathbf{C}^\dagger \mathbf{C}$, e pode novamente ser usada para contruir uma nova matriz de Fock. Já que a matriz de sobreposição \mathbf{S} não depende dos coeficientes dos orbitais moleculares, a mesma transformação unitária pode ser aplicada para a nova matriz de Fock fornecer a nova transformada da matriz. Esta pode ser diagonalizada para produzir novos coeficientes dos orbitais moleculares e o processo se repetir até a convergência. Como passo inicial para a matriz de Fock, pode-se usar uma função hamiltoniana que ignore todas as integrais de dois elétrons, obtendo um coeficiente \mathbf{C} inicial e uma matriz incluindo a interação entre dois elétrons pode ser construída a partir da inicial.

A matriz densidade descreve a densidade eletrônica da molécula. Assim, o critério de convergência do procedimento **SCF** pode se referir tanto a ela quanto à energia. As integrais de dois elétrons sobre funções de base atômicas são os maiores problemas na aplicação do método ab-initio pois uma qualidade razoável exige um grande número de funções de base. Para grandes sistemas é praticamente impossível guardá-las em disco.

3.2.2 Funções de Base

Para contornar a dificuldade do cálculo de integrais multicentricas, definem-se as funções de base. As funções de onda usadas têm os termos radiais de $\varphi(\mathbf{r})$ simplificadas. As mais comuns destas funções de onda são os orbitais do tipo Slater (**STO** - Slater-type orbitals) que possuem a forma de decaimento exponencial

$$\varphi(\mathbf{r}) = e^{-\alpha r}. \quad (3.45)$$

Na prática, o cálculo das propriedades destas funções são muito demoradas. Para permitir a computação das integrais multicêntricas, realiza-se a aproximação a estes orbitais por uma combinação linear de funções de base gaussianas (**GTO** - Gaussian-type orbitals)²(na figura 3.1 pode-se verificar o comportamento esquemático das funções STO e GTO):

$$\varphi^{STO} \approx \sum_{\nu=1}^n k_{\nu} \varphi_{\nu}^{GTO}. \quad (3.46)$$

A forma funcional de um orbital **GTO** é diferente do tipo **STO** especialmente na vizinhança do núcleo. Os expoentes α e os coeficientes de contração k podem ser determinados de diferentes modos, como por exemplo, pela otimização da energia em cálculos rigorosos por métodos auto-consistentes para pequenos átomos e moléculas. Determinados estes valores define-se um conjunto de base padrão que não se modifica em cálculos das outras moléculas. Em resumo, temos um conjunto de base mínimo onde somente os orbitais ocupados de cada átomo isolado são usados para compor os orbitais moleculares. Orbitais não ocupados são denominados virtuais.

3.2.3 Bases Padrão

Muitas bases são utilizadas. As mais populares são as descritas por Pople e co-autores. Não significa que elas formem o melhor conjunto, mas o uso extensivo e a experiência em cálculos práticos transformam-nas nas bases mais utilizadas. Os elementos da segunda linha da classificação periódica têm por funções de base os orbitais atômicos 1s, 2s, 2px, 2py e 2pz. As **STOs** são substituídas por **nGTO**. A menor base é **STO-3G**, isto é, os orbitais **STO** foram aproximados por três gaussianas.

Para aumentar a flexibilidade das funções de onda **SCF**, pode-se aumentar o número de funções de base por átomo.

Na base **ZZ** (double zeta) definem-se duas funções de base para cada orbital atômico da base mínima, um dos quais está mais próximo do núcleo, e o outro permitindo a densidade eletrônica se mover afastada do núcleo. Os elementos da segunda linha são funções de base que dão 1s, 1s', 2s, 2s',

²Não se usa orbitais GTO nos cálculos semiempíricos porque neste tipo de cálculo não se computam integrais multicentricas.

$2p_x, 2p_x', 2p_y, 2p_y', 2p_z, 2p_z'$. Quando esta duplicação de bases mínima é feita somente para os orbitais de valência temos o que se denomina *Split Valence Basis Set*. Temos para a segunda linha, as bases atômicas orbitais $1s, 2s, 2s', 2p_x, 2p_x', 2p_y, 2p_y', 2p_z, 2p_z'$. As mais conhecidas são as bases **3-21G** e **6-31G**.

Na base **3-21G** o orbital atômico $1s$ do elemento da segunda linha é representado por uma combinação fixa de 3 **GTOs**, os elementos $2s$ ($2p_x$, etc.) são aproximados por uma combinação fixa de 2 **GTOs** e os orbitais de valência extra $2s'$ ($2p_x''$, etc.) são ajustados com um **GTO**. Esta base oferece uma qualidade razoável nos resultados. As seqüências maiores tais como **6-31G** ($1s$: 6 **GTOs**; $2s$ ($2p_x$, etc.): 3 **GTOs**; $2s'$ ($2p_x'$, etc.): 1 **GTO**) fornecem uma energia muito menor mas a quantidade de integrais do tipo **GTO** a ser calculada é muito maior.

As funções de base que contém polarização, isto é, os orbitais d para os elementos da segunda linha não podem ser calculadas em átomos porque as funções de polarização não são ocupados. As bases consideradas boas para a geometria da maioria das moléculas são a **6-31G*** ou **6-31G(d)** (**6-31G** + funções d para os átomos da segunda linha), e a **6-31G**** ou **6-31G(d,p)** (**6-31G** + funções d para os átomos da segunda linha, + p para o hidrogênio).

A notação usualmente utilizada é, por exemplo, **6-31G**//3-21G** que representa cálculos com a base **6-31G**** e uma otimização de geometria com a base **3-21G**.

3.3 O método CI

A teoria **SCF Hartree-Fock** restrita não descreve adequadamente a separação de átomos de hidrogênio usando-se orbitais duplamente ocupados. O H_2 tende a dissociar-se em H^+ e H^- que pode ser descrito com um orbital duplamente ocupado em H^- . Este problema não ocorre com o método **UHF**, mas este método tem a desvantagem de não fornecer estados de spin puros. Devido ao uso da aproximação de partículas independentes, a correlação dos movimentos instantâneos dos elétrons é desprezada. Mesmo quando a energia **Hartree-Fock** é de aproximadamente 99% da exata, a diferença pode ser importante quimicamente. Discutiremos rapidamente o método da *interações de configurações* (**CI**) que calcula a energia de correlação (diferença de energia entre a energia exata determinada pela função hamiltoniana e a

energia calculada pelo método Hartree-Fock).

Na teoria Hartree-Fock uma função de onda é representada como um determinante de Slater. No método das interações de Configuração (CI), constrói-se uma combinação linear de determinantes de Slater, usando orbitais não ocupados a partir dos cálculos SCF. Isto porque os métodos SCF subestimam os efeitos da repulsão elétron-elétron, diminuindo a acurácia das funções de onda e energias calculadas. Interações repulsivas podem ser minimizadas permitindo que os elétrons possam existir em mais lugares (orbitais).

A função de onda total é escrita como:

$$\Psi = \Phi_{HF} + \sum_{ij^*} c_{ij}^S \phi_i^{j^*} + \sum_{ijk^*l^*} c_{ijk^*l^*}^D \phi_{ij}^{k^*l^*} + \dots$$

ou seja, as diferentes funções de onda são combinadas de maneira análoga à aproximação LCAO. Em princípio, a energia de correlação exata pode ser obtida a partir de um cálculo CI completo desde que todas as configurações sejam levadas em consideração o que é possível unicamente para sistemas muito pequenos.

HF conjunto de bases mínimo		métodos de correlação de elétrons	CI completo
		→ →	
conjunto ↓			↓
extendido ↓			↓
de bases ↓			↓
			equação de Schrödinger (independente do tempo, não relativística)
limite Hartree-Fock no infinito	→ →		exata

Mesmo para sistemas pequenos, o número de configurações excitadas é enorme. Uma maneira de truncar a expansão CI é considerar apenas as configurações simples e duplamente excitadas (CI-SD). A energia calculada como o valor esperado da função hamiltoniana para CISD é:

$$E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = E_{HF} + \sum_{i < k} \sum_{k^* < l^*} c_{ijk^*l^*}^D [(ij|k^*l^*) - (ik^*|jl^*)]. \quad (3.47)$$

Para executar os cálculos necessitam-se de integrais de dois elétrons sobre os orbitais moleculares. O cálculo é muito demorado, mesmo quando as integrais sobre os orbitais atômicos são acessíveis:

$$(ij|kl) = \sum_{\mu\nu\lambda\sigma} c_{i\mu}c_{j\nu}c_{k\lambda}c_{l\sigma} \langle \mu\nu | \lambda\sigma \rangle.$$

Em geral, o método **CI** não é um método prático para o cálculo da energia de correlação porque o método **CI** completo não é possível e a convergência da expansão **CI** é lenta. O método **CI** truncado não é extensivamente consistente. O cálculo de duas espécies a grandes distâncias não fornece a mesma energia da soma dos cálculos nas espécies separadas. Isto acontece porque faz-se uma seleção diferente das configurações excitadas nos dois cálculos.

3.3.1 Comentários

A correlação eletrônica é capaz de modificar sensivelmente as conclusões, mesmo qualitativas, de um estudo quântico, de maneira que se pratica a correção dos cálculos SCF.

A interação de configuração (**CI**) constitui o complemento natural dos cálculos SCF. Para se levar em conta a mais completa correlação, para uma dada base de orbitais atômicos, devemos desenvolver a função de onda sobre a totalidade das configurações permitidas pela simetria e pelo spin. Este número cresce com o número de elétrons. Chega-se rapidamente ao limite físico de cálculo onde a interação de configurações completa se limita a moléculas pequenas e aos cálculos de referência.

É prática corrente procurar limitar o volume dos cálculos sem afetar muito a precisão dos resultados. A mais simples é limitar-se a certas excitações, duplas, simples e duplas, etc.³

3.4 Métodos Semiempíricos

O número de integrais bieletrônicas, bem como a relativa dificuldade de seu cálculo, fazem dos métodos ab-initio uma máquina numérica pesada, de uso limitado a pequenas moléculas que possam servir de sistemas de referência. A

³Existe o método das perturbações que é conveniente para o estudo de interações moleculares.

necessidade de cálculos que permitissem comparações no interior de famílias de compostos, de sorte a permitir conclusões qualitativas de uma certa generalidade, levou a métodos de cálculos mais rápidos e menos precisos, os métodos semi-empíricos.

Princípio dos Métodos Semiempíricos.

Estes métodos tem o objetivo de reduzir em uma grande proporção o número de integrais a se calcular, particularmente, as bieletrônicas. As aproximações são as seguintes:

- As bases dos orbitais utilizadas são constituídas pelos orbitais de Slater da camada de valência;
- As integrais de superposição são desprezadas na resolução das equações SCF;
- Todas as integrais de dois elétrons, com três e quatro centros, são supostas serem nulas. Algumas das integrais de dois centros também são desprezadas, dependendo do método utilizado;
- Os termos não diagonais da matriz da função hamiltoniana da parte interna da molécula são estimadas por meio de relações empíricas que supõe que estas integrais são proporcionais à integral de superposição dos orbitais atômicos;
- A maior parte das integrais de um e dois elétrons a um centro é estimada a partir de dados espectrais eletrônicos dos átomos ou íons e que são os parâmetros de Slater -Condon. Se utilizarmos os orbitais de Slater (mesma função radial para orbitais diferentes de mesmo número quântico principal) nota-se que a contribuição principal nas diferenças de energia entre estados eletrônicos, onde a configuração eletrônica é obtida a partir de orbitais atômicos de mesmo número quântico principal, provém das integrais coulombianas e de troca. O cálculo dos parâmetros destas integrais repousa no desenvolvimento da quantidade $1/r_{12}$ em harmônicos esféricos e sobre a propriedade destas funções que permite substituir o produto de duas delas por uma soma finita. O cálculo se limita às integrais que dependem de r e a forma analítica da parte radial. Costuma-se representar estas integrais pelos símbolos

F^n e G^n onde n é um número inteiro e estas grandezas são extraídas de certas energias de transição observadas em espectroscopia atômica, com exceção de F^0 que é calculada, ou ainda, obtida por meio de alguma relação empírica.

3.4.1 Os métodos NDDO

NDDO (Neglect Diatomic Differential Overlap) significa que desprezamos todas as integrais de sobreposição de dois orbitais se eles estão centrados em átomos diferentes

Lembrando as equações (3.31) a (3.34) os elementos da matriz de Fock expressas em bases atômicas se simplificam drasticamente nesta aproximação (Zero Differential Overlap), ou seja,

$$\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) = 0 \text{ para } \mu \neq \nu. \quad (3.48)$$

Com esta aproximação, para os elementos de matriz para uma base de orbitais s e p , teremos os termos $F_{\mu\mu}$, $F_{\mu\lambda}$ e $F_{\mu\lambda}$ dados por :

$$F_{\mu\mu} = U_{\mu\mu} + \sum_B V_{\mu\mu}^B + \sum_\nu^A P_{\nu\nu} \left[(\mu\mu|\nu\nu) - \frac{1}{2} (\mu\nu|\mu\nu) \right] + \sum_B \sum_{\lambda\sigma}^B P_{\lambda\sigma} [(\mu\mu|\lambda\nu)] \quad (3.49a)$$

$$F_{\mu\lambda} = \sum_B V_{\mu\nu}^B + P_{\mu\nu} \left[\frac{3}{2} (\mu\nu|\mu\nu) - \frac{1}{2} (\mu\nu|\mu\nu) \right] + \sum_B \sum_{\lambda\sigma}^B P_{\lambda\sigma} [(\mu\nu|\lambda\sigma)] \quad (3.49b)$$

$$F_{\mu\lambda} = \mathbf{H}_{\mu\lambda}^c - \frac{1}{2} \sum_\nu^A \sum_\sigma^B P_{\nu\sigma} [(\mu\nu|\lambda\sigma)]. \quad (3.49c)$$

Aqui estamos utilizando a notação \mathbf{H}^c para designar o operador do núcleo mais os átomos internos (core), $U_{\mu\mu}$ o operador hamiltoniano do átomo em relação ao orbital μ e

$$V_{\mu\nu}^B = \left\langle \phi_\mu \left| \frac{z_B}{r_B} \right| \phi_\nu \right\rangle$$

z_B a carga do core do átomo B e \sum_ν^A representa a soma sobre os orbitais ν centrados no átomo A. Na versão original do método, os termos $U_{\mu\mu}$ são calculados das relações de Slater-Condon a partir do potencial de ionização e da afinidade eletrônica.

3.4.2 O método MNDO

Este método obtém todas as integrais monocêntricas dos valores experimentais a partir do método de Oleari⁽⁶⁾. A lei da variação das integrais bieletrônicas bicêntricas é calculada com a ajuda de um desenvolvimento multipolar das distribuições de carga parametrizadas de maneira que quando a distância interatômica tende a zero, ela tenderá ao valor semiempírico da energia monocêntrica correspondente determinada pelo método de Oleari.

Os termos $V_{\mu\nu}^B$ são calculados por meio de uma relação empírica que supõe que a interação de um elétron representado pela distribuição $\phi_\mu\phi_\nu$ centrada sobre o átomo A com uma carga +1 localizada sobre o núcleo de B é igual, a menos do sinal, a interação deste elétron com um elétron representado por orbitais s da camada de valência de B, simbolizada por s^B , ou

$$V_{\mu\nu}^B = -z_B \langle \mu\nu | s^B s^B \rangle. \quad (3.50)$$

Os termos não diagonais da hamiltoniana do *core* são supostas proporcionais às integrais de superposição

$$H_{\mu\lambda}^c = S_{\mu\nu} \frac{(\beta_\mu + \beta_\nu)}{2} \quad (3.51)$$

onde os β são parâmetros empíricos determinados sobre moléculas de referência, de maneira que os cálculos reproduzam corretamente as quantidades procuradas.

Para se obter a energia da molécula, junta-se à energia eletrônica obtida pelo cálculo **SCF**, a energia de repulsão entre os núcleos, com os elétrons das camadas internas (*core*). Esta quantidade é calculada por meio de uma relação empírica

$$E_{AB}^c = z_A z_B \langle s^A s^A | s^B s^B \rangle [1 + f(R_{AB})] \quad (3.52)$$

com

$$f(R_{AB}) = e^{-\alpha_A R_{AB}} + e^{-\alpha_B R_{AB}}$$

α_A e α_B são os parâmetros onde

$$f(R_{AX}) = R_{XH} e^{-\alpha_X R_{RX}} + e^{-\alpha_B R_{AB}}$$

para X=N ou O.

Este método é muito bem adaptado à maior parte das grandezas de interesse como energias de formação, geometrias moleculares, momentos dipolares e multipolares e constantes de força. Um dos problemas reside na ligação hidrogênio, que levou a versões destes métodos **AM1** (Austin Model) e **PM3** (Parametric Method 3), que diferem por valores de certos parâmetros e também da expressão de E_{AB}^c que utiliza uma função $f(R_{AB})$ mais refinada.

3.4.3 Os métodos CNDO e INDO

CNDO (Complete Neglect of Differential Overlap) simplifica ainda mais os cálculos. Da definição da matriz de superposição

$$S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle = \langle \mu | \nu \rangle = \delta_{\mu\nu} \quad (3.53)$$

onde $\delta_{\mu\nu} = 0$ se $\mu \neq \nu$, e $\delta_{\mu\nu} = 1$ se $\mu = \nu$, o que leva ao anulamento de muitas integrais de dois elétrons

$$(\mu\nu | \lambda\sigma) = \delta_{\mu\nu} \delta_{\lambda\sigma} (\mu\mu | \nu\nu).$$

Para evitar que integrais do tipo $(xy|xy)$ conduzam a um valor zero, esta aproximação admite que integrais do tipo $(\mu\mu|\lambda\lambda)$ são independentes da orientação dos eixos. Estas integrais são as do tipo s, e todas elas são substituídas pelo valor que elas teriam se os orbitais atômicos dos orbitais s tivessem o mesmo número quântico principal, ou

$$(\mu\mu | \lambda\lambda) = \gamma_{AB} \quad (3.54)$$

com os orbitais ϕ_μ e ϕ_λ centrados sobre os átomos A e B respectivamente.

O método **CNDO/2** utiliza um conjunto de hipóteses além das normais CNDO como por exemplo a energia de interação dos centros atômicos a partir da aproximação eletrostática

$$E_{AB}^c = \frac{z_A z_B}{R_{AB}}.$$

Para cobrir os passos entre os métodos acima descritos criou-se o método **INDO** onde o I é de Intermediate. Esta aproximação consiste em acrescentar às integrais de dois elétrons do método CNDO as integrais monocêntricas do tipo $\langle \mu\nu | \mu\nu \rangle$. Versões modificadas deste método tem sido propostas, comparáveis aos do tipo MNDO, que aparecem com o nome de **MINDO/1,2 e 3**,

parametrizadas para permitir resultados confiáveis nas energias de formação das moléculas. Zerner realizou uma parametrização que permite uma boa descrição das transições eletrônicas na região espectral do UV/VIS para moléculas orgânicas, e é conhecido por **ZINDO**.

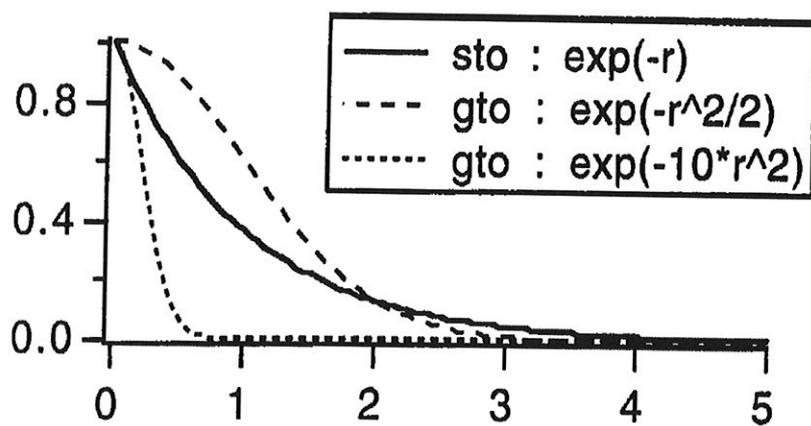


Figura 3.1: Os orbitais do tipo-Slater - STO e os do tipo-Gauss - GTO.

Parte II
CÁLCULOS CLÁSSICOS

We may even judge the degree of perfection to which a science has arrived by the facility with which it may be submitted to calculation. (? ,Quetelet)

Iniciamos o capítulo quatro com um apanhado geral dos métodos de cálculo de energia livre que existem. A seguir, mostramos a técnica de cálculo da energia livre absoluta implantada no programa Thor . Descrevemos a calibração efetuada com a molécula fentanil. O capítulo cinco, apresentamos alguns cálculos realizados com o método que propomos para analisar as possíveis populações dos vários rotâmeros do triptofano *zweiteriônico*.

Capítulo 4

Cálculo da Energia Livre

si non è vero...

Cálculos de diferenças de energias livres configuracionais a partir de simulações computacionais em sistemas complexos seguem, em geral, uma das técnicas, ou o método da integração termodinâmica, ou o método das perturbações.

Para o cálculo da energia livre introduzimos no programa Thor o modelo de sólido de Einstein. Descrevemos a seguir, rapidamente, os resultados existentes nesta área e o modelo teórico que adotamos.

4.1 O Cálculo da Energia Livre Usando Dinâmica Molecular

Introdução

Quando calculamos a função energia total de uma molécula, isto é, a energia potencial mais a energia cinética, estamos, em geral, desprezando a parte que envolve a sua entropia. Assim, o mínimo da energia total não significa que tenhamos necessariamente a sua conformação mais favorável. Em outras palavras, quando uma molécula tem dois isômeros com energias totais muito próximas a conformação favorecida será necessariamente aquela que apresentar o balanço entre energia total e entropia mais baixo, ou seja, aquela que tiver a menor energia livre⁽⁵⁸⁻⁶⁵⁾.

4.1.1 Energia Livre Relativa

O método das perturbações

Define-se a função energia livre por

$$A = -k_b T \ln Q \quad (4.1)$$

onde A é a energia livre, k_b a constante de Boltzmann, T a temperatura absoluta e Q a função de partição do sistema, no *ensemble* considerado (canônico ou das pressões). Dados dois estados 1 e 2, e Q_1 e Q_2 suas funções de partição respectivas, a diferença na energia livre entre estes estados é

$$\Delta A = A_2 - A_1 = -k_b T \ln \frac{Q_2}{Q_1}. \quad (4.2)$$

Definindo $E_i(p, q)$ como a função energia correspondente ao estado i , a razão das funções de partição está relacionada com o valor esperado de $\exp[-(E_2 - E_1)/k_b T]$ por

$$\frac{Q_2}{Q_1} = \int \int \exp[-(E_2 - E_1)/k_b T] P_1 d\vec{r} d\vec{p} \quad (4.3)$$

onde P_1 é a probabilidade de Boltzmann do estado 1 dado por

$$P_1 = \frac{\exp[-E_1/k_b T]}{Q_1}. \quad (4.4)$$

Usando a notação de Dirac a equação (4.3) pode ser escrita

$$\frac{Q_2}{Q_1} = \langle \exp[-\Delta E/k_b T] \rangle_1 \quad (4.5)$$

onde o subscrito 1 indica o estado 1 considerado como estado de referência (a diferença de energia ΔE é computada relativamente a um conjunto de estruturas para o estado 1). Então, a razão das funções de partição pode ser computada a partir de um *ensemble* médio da diferença de energias entre um estado de referência e um estado perturbado. Isto é realizado a partir de uma trajetória definida que conecte os dois sistemas no espaço das configurações. A mudança na energia livre é dada então diretamente por:

$$\Delta A = A_2 - A_1 = k_b T \ln \langle \exp[-\Delta E/k_b T] \rangle_1. \quad (4.6)$$

4.1. O CÁLCULO DA ENERGIA LIVRE USANDO DINÂMICA MOLECULAR 55

A aplicação do método das perturbações envolve essencialmente uma simulação baseada na função energia E_1 (ou E_2), onde E_2 (ou E_1) é também computada a cada passo e a média de $\exp[\pm (E_2 - E_1)/k_bT]$ é formada. Cálculos numéricos sucessivos do conjunto médio desta exponencial via simulação exige que os estados 1 e 2 não sejam díspares. Para evitar tais dificuldades é possível usar um parâmetro de acoplamento que torne viável numericamente uma trajetória entre os estados termodinâmicos 1 e 2, envolvendo estados intermediários que estão suficientemente próximos uns dos outros e computar a diferença de energia livre dada pela equação (4.6).

O método das perturbações é muito utilizado devido a sua simplicidade conceitual. Para pequenas diferenças este método é muito atraente porque basta um único cálculo. A grande vantagem deste método é a facilidade na sua implantação (necessita-se apenas do cálculo da energia, não se necessitando de nenhuma derivada) e conceitualmente o cálculo é direto. Entretanto, a equação (4.6) e a equação reversa (que considera o estado termodinâmico 2) são muito sensíveis a flutuações na diferença de energia, exigindo muitos estados intermediários devido a presença da exponencial. Em outras palavras, o cálculo de energia livre pelo método das perturbações tem bons resultados somente quando a flutuação das diferenças de energia entre os estados inicial e final é pequena. Para casos anômalos pode-se executar uma seqüência de cálculos, cada um computando a mudança na energia livre em um subintervalo⁽⁶⁹⁾.

FDTI

Mezei introduziu um novo método de cálculo da energia livre, que denominou como integração termodinâmica de diferenças finitas (finite difference thermodynamic integration). O traço comum dos vários métodos de cálculos das diferenças de energia livre entre dois sistemas é a definição de um caminho que une os dois sistemas no espaço das configurações. Na maioria das vezes um parâmetro de acoplamento λ é introduzido. Ele varia entre 0 e 1 e leva continuamente o sistema inicial 0 para o sistema final 1. A integração termodinâmica de Kirkwood considera a energia livre como uma função de λ e emprega o teorema fundamental do cálculo

$$A_1 - A_0 = \Delta A = \int_0^1 \frac{\partial A(\lambda)}{\partial \lambda} d\lambda. \quad (4.7)$$

Podemos escrever ainda

$$\Delta A = \int_0^1 \frac{\delta A(\lambda)}{\delta \lambda} d\lambda. \quad (4.8)$$

Assumindo que se pode encontrar um parâmetro de acoplamento razoável, que descreva adequadamente a conversão entre os dois estados continuamente, a equação acima pode ser integrada numericamente. O método FDTI emprega o formalismo das perturbações para computar numericamente as derivadas da função energia livre com relação ao parâmetro de acoplamento. Usando a equação (4.6), é possível calcular as mudanças na energia livre (ΔA_i) para uma perturbação $\delta \lambda$, afastado do ponto i de λ por $\delta \lambda$ como

$$\Delta A = A(\lambda_i) - A(\lambda_i \pm \delta \lambda) = -k_b T \ln \langle \exp[-[E(\lambda_i) - E(\lambda_i \pm \delta \lambda)]/k_b T] \rangle_i. \quad (4.9)$$

Calculando-se ΔA_i para muitos valores diferentes de λ variando no intervalo de 0 a 1, dividindo cada ΔA_i por $\delta \lambda$, e integrando numericamente no intervalo, a mudança na energia livre total pode ser estimada como

$$\Delta A = -k_b T \sum_{i=1}^k \frac{\ln \langle \exp[-[E(\lambda_i) - E(\lambda_i \pm \delta \lambda)]/k_b T] \rangle_i}{\delta \lambda} \Delta \lambda \quad (4.10)$$

onde k é o número de pontos da quadratura na integração numérica. $\frac{\Delta A_i}{\delta \lambda}$ pode ser calculado tanto com a perturbação para a "frente" ($+\delta \lambda$) ou para "trás" ($-\delta \lambda$). Computando ambas ao mesmo tempo, não se necessita um tempo computacional muito maior e é um critério de convergência.

O valor de $d\lambda_i$ na equação (4.8) para um dado i depende do esquema de integração numérica usado. Para uma aproximação trapezoidal tomamos todos os $\Delta \lambda_i$ iguais. Usamos o método de integração das quadraturas de Gaussian-Legendre (ver Numerical Recipes in Fortran) que escolhe os valores necessários, dado o número total de intervalos especificado.

As vantagens deste método são que, ao contrário do método de perturbações, grandes mudanças na energia livre podem ser calculadas em poucos passos, e não necessitamos de derivadas analíticas da hamiltoniana com relação ao parâmetro de acoplamento.

Para maiores detalhes sobre estes métodos ver referências 3,4,5,6 e 7.

4.2 Energia Livre Absoluta

A técnica de cálculo da energia livre absoluta que implantamos no programa Thor com Michel Loos foi apresentado inicialmente no pacote comercial *Discover*⁽⁸⁾ e é muito conveniente computacionalmente, preciso e eficiente, e serve para conformações moleculares estáveis ou instáveis. Ela é geral e pode ser aplicada de maneira clara a sistemas no vácuo ou em solução, sob quaisquer condições de pressão, e ou temperatura. Trata-se de um caso especial da integração termodinâmica para cálculo da energia livre, um método geral da mudança da energia livre quando se passa de um estado termodinâmico para outro. A energia livre absoluta simplesmente vincula um destes estados como o sistema modelo para o qual a energia livre absoluta é conhecida analiticamente. Por integração a partir de um estado conhecido, trabalhado para o estado real final, a energia livre absoluta torna-se a soma dos passos de integração numericamente computada e da energia livre do estado modelo.

Desenvolver o cálculo das energias livres absolutas para conformações particulares é um objetivo importante porque os valores de diferentes estados termodinâmicos da energia livre absoluta podem ser comparados diretamente, não se necessitando acompanhar uma trajetória de transição entre eles, como é feita nos cálculos de energia livre relativa.

Derivação para Sistemas Ideais

O algoritmo da energia livre absoluta depende da definição de um modelo para o qual a função de partição possa ser derivada analiticamente. O modelo que escolhemos é o de um sólido ideal. Os átomos no sistema estão ligados harmonicamente a uma coordenada de referência, analogamente a um sólido, e não devem interagir entre eles, como se fosse um gás ideal. Este sólido, o sólido de Einstein tem a seguinte função hamiltoniana:

$$H_{\text{sólidoideal}} = \frac{1}{h^{3N}} \sum_i^N \frac{1}{2m_i} (p_{x_i}^2 + p_{y_i}^2 + p_{z_i}^2) + \quad (4.11)$$

$$\sum_i^N K_i \left[(x_i - x_i^0)^2 + (y_i - y_i^0)^2 + (z_i - z_i^0)^2 \right] \quad (4.12)$$

onde a primeira somatória é simplesmente a energia cinética, e o segundo termo é uma função harmônica ligando cada átomo a um ponto de referência

correspondente (x_i^0, y_i^0, z_i^0) com uma constante de força K_i . Notemos que não existem termos para interações entre as partículas. Esta simplificação torna possível uma solução analítica. Substituindo este hamiltoniano na função de partição

$$Q = \int \int e^{-H(p,r)/k_B T} dp dr \quad (4.13)$$

resulta em

$$Q = \prod_{i=1}^N \left[\frac{2m_i}{K_i} \left(\frac{\pi k_B T}{h} \right)^2 \right]^{3/2}. \quad (4.14)$$

Retornando à equação inicial da energia livre, esta pode ser escrita para um sólido ideal como

$$A_{\text{sólidoideal}} = \frac{3}{2} k_B T \sum_{i=1}^N \ln \left[\frac{2m_i}{K_i} \left(\frac{\pi k_B T}{h} \right)^2 \right] \quad (4.15)$$

que depende unicamente de m_i e de K_i , não dependendo da escolha dos pontos de referência. Esta propriedade tem importantes implicações práticas no cálculo da energia livre. Os valores da energia do estado de referência são sempre os mesmos para todas as conformações da mesma molécula.

Derivação de Sistemas Reais

A integração da função de partição analiticamente pode ser possível para funções hamiltonianas simples. Entretanto, para sistemas mais realistas que incluam muitas interações entre átomos ligados e não ligados, uma solução analítica é impossível. Usamos a integração termodinâmica para determinar a mudança na energia livre:

$$\Delta A = A_1 - A_0 = \int_0^1 \frac{\delta A(\lambda)}{\delta \lambda} d\lambda. \quad (4.16)$$

Substituindo a equação da energia livre nesta equação, podemos escrever a diferença na energia livre como uma função da função de partição $Q(\lambda)$:

$$A_1 - A_0 = \int_0^1 \frac{\partial [-k_B T \ln Q(\lambda)]}{\partial \lambda} d\lambda = -k_B T \int_0^1 \frac{1}{Q(\lambda)} \frac{\partial Q(\lambda)}{\partial \lambda} d\lambda. \quad (4.17)$$

Sem definir como a função hamiltoniana depende de λ , podemos escrever a diferença nas energias como:

$$A_1 - A_0 = \int_0^1 \left\langle \frac{\delta H(p, r, \lambda)}{\delta \lambda} \right\rangle d\lambda. \quad (4.18)$$

O problema da computação da diferença na energia livre é então simplificada para a computação do valor esperado de uma derivada da hamiltoniana. Como o valor esperado é, pelo postulado de Gibbs, o “*ensemble médio*” da quantidade, é facilmente calculada como a média daquela quantidade sobre um conjunto molecular apropriado instantâneo do sistema. Estes instantâneos podem ser gerados tanto através da dinâmica molecular (nada impede que eles possam ser gerados por simulação de Monte Carlo). Então uma média apropriada resultante de uma trajetória de dinâmica molecular permite que a equação (4.18) possa ser integrada. Entretanto ela está longe de ser mais fácil de ser integrada que a função de partição da equação (4.17).

A equação (4.18) é perfeitamente geral para qualquer sistema clássico e é a equação fundamental de todos os métodos de integração termodinâmica. A hamiltoniana pode ser parametrizada de infinitas maneiras, entre elas, em termos de λ . Como uma função de estado (a energia livre) está sendo integrada, a trajetória entre os pontos terminais é arbitrária. Para o cálculo da energia livre absoluta, o mais adequado é tomar a forma mais simples possível.

Parametrização da função hamiltoniana Definimos a hamiltoniana como uma combinação linear de duas funções de energia potencial que descreve os estados limites:

$$H(\lambda) = K(p) + (1 - \lambda)V_0 + \lambda V_H \quad (4.19)$$

$K(p)$ é a energia cinética (como os dois estados são completamente definidos por suas respectivas funções energia potencial, a energia cinética não deve ser acoplada a λ), V_0 é a função energia potencial normal (incluindo ligações, ângulos, torções, etc.), e V_H é um potencial localizado de oscilador harmônico dado por

$$V_H = \sum_i K_i (r_i - r_i^0)^2 \quad (4.20)$$

onde K_i é a constante de mola do $i^{\text{ésimo}}$ átomo, r_i é a sua coordenada instantânea e r_i^0 a rede de referência dos átomos não interagentes (referido como um sólido de Einstein). Na equação (4.19), λ é maior do que zero e menor do que um, e descreve uma trajetória no espaço das energias entre um sistema descrito por um campo de forças molecular não alterado ($\lambda = 0$) e um que é representado por osciladores harmônicos independentes ($\lambda = 1$).

A equação (4.20) refere-se portanto ao sistema de referência para o qual uma energia livre absoluta pode ser diretamente calculada. O potencial V_H restringe a exploração do espaço de fase a uma região definida por deslocamentos atômicos quadráticos médios relativos ao sólido de Einstein. Conseqüentemente, a escolha de ambos, o estado de referência e as constantes de mola, afetam as energias livres calculadas. Para muitos casos, as coordenadas da energia minimizada de uma estrutura mecanicamente estável são usadas para a referência do sólido de Einstein. Usando a equação (4.16) para a energia do sólido de Einstein A_1 e as equações (4.18) e (4.19), a energia livre absoluta do estado real é¹:

$$A_0 = A_1 - \int_0^1 \left\langle \frac{\partial}{\partial \lambda} [K(p) + (1 - \lambda)V_0 + \lambda V_H] \right\rangle d\lambda = A_1 + \int_0^1 \langle V_0 - V_H \rangle d\lambda \quad (4.21)$$

Cuidados nos Cálculos Computacionais A forma simples da equação (4.21) possibilita um cálculo direto em seu *ensemble* médio. Uma trajetória dinâmica para um dado valor de λ representa um *ensemble* das estruturas para aquele λ . Calculando e tomando-se a média de $(V_0 - V_H)$ para cada estrutura neste *ensemble*, o integrando pode ser numericamente estimado para qualquer valor dado de λ . Fazendo muitos destes cálculos para alguns valores de λ entre 0 e 1, eventualmente a função pode ser numericamente integrada.

Simulações práticas podem ser realizadas a partir da equação (4.21). Quando λ está próximo a zero, o *ensemble* das estruturas calculadas incluem configurações longe do estado de referência. As flutuações de V_H são grandes quando este potencial é desenvolvido neste *ensemble*. Alternativamente, quando λ é próximo de 1, as estruturas geradas são influenciadas muito pouco pelo campo de forças real. Estes *ensembles* podem incluir ligações e ângulos distorcidos e mesmo átomos sobrepostos, levando a grandes

¹ V_H não tem nenhuma interação com o meio externo.

flutuações em V_0 . Grandes flutuações de $(V_0 - V_H)$ diminui a precisão com que a integral pode ser calculada. O resultado final é desestabilizar a integral e, em casos extremos, causar sua divergência.

Estes efeitos indesejados podem ser minimizados com escolhas razoáveis das estruturas de referência, constantes de mola e algoritmo de integração. Usamos para integrar a equação (4.21) o algoritmo de Gaussian-Legendre, que tem a propriedade de não necessitar evoluir a função nos extremos.

A Escolha do Estado de Referência A escolha do estado de referência é o passo mais crítico no cálculo da energia livre absoluta. O estado de referência determina onde a amostra do espaço das configurações está centrado. Escolhendo estados de referência que são somente levemente diferentes (por exemplo, 0,1Å nas distâncias quadráticas médias nas coordenadas) pode mudar a energia livre significativamente se o estado de referência possui alguma tensão residual. Melhores resultados são obtidos quando o estado de referência representa uma conformação de energia mínima na superfície de energia. Se a estrutura de referência é mínima para V_0 e V_H , não existe condições para a conformação sair dela e causar uma divergência. Para configurações que não são minimizadas, como o ponto de sela, o excesso de tensão pode ser removida pela minimização da restrição torcional.

As Constantes de Mola e de λ

Todos os átomos têm o mesmo valor para as constantes de mola. Usamos estas constantes com valores altos dos K_i para obrigar a dinâmica da amostra no espaço de fases a manter-se próxima a estrutura de referência.

Para estruturas mecanicamente estáveis devemos usar constantes de mola renormalizadas, que efetivamente reproduzam flutuações causadas por termos do campo de forças (forças de Coulomb, não ligadas, ligadas, etc.). Para estes estados os deslocamentos quadráticos médios são convergentes, tornando sem significado valores de mola iguais para todos os átomos. Os valores das constantes de mola devem estar no intervalo de 1 a 50 kcalmol⁻¹Å⁻². Estas constantes podem ser estimadas a partir da definição de deslocamento quadrático médio sob a ação de um potencial harmônico usando a expressão:

$$K_i = \frac{k_B T}{\langle \Delta x^2 \rangle_0} \quad (4.22)$$

Para o programa Thor, em nossas simulações com o fentanil, o valor da constante de mola para o sólido de Einstein, que manteve a amostra do espaço configuracional dentro dos limites da região pré-estabelecida, foi o de $50 \text{ kcalmol}^{-1} \text{ \AA}^{-2}$ em relação ao valor de $1 \text{ kcalmol}^{-1} \text{ \AA}^{-2}$. Estes valores foram associados como único valor da constante de mola para todos os átomos. Entretanto, um teste sistemático, para se saber qual o melhor valor da constante de mola, precisa ser efetuado. Devemos monitorar o erro associado com cada escolha explícita, assim como a amostra do espaço configuracional para este cálculo.

Já a escolha dos intervalos de λ depende do grau de precisão e do comportamento da integral e não pode ser antecipado em todos os casos. Para moléculas pequenas e flexíveis, usando uma constante de mola média de $50 \text{ kcalmol}^{-1} \text{ \AA}^{-2}$ e o uso de 10 intervalos conduz a um comportamento razoável da integral com erro estatístico de $\pm 0.5 \text{ kcalmol}^{-1}$. Para cálculos mais cuidadosos, o comportamento da integral deve ser examinado. Mudanças rápidas em $(V_0 - V_H)$ são indicativos de erros sistemáticos. Algumas regiões da integral necessitam então de uma computação longa com mais intervalos para produzir os resultados desejados.

Limitações

A técnica de cálculo da energia livre absoluta deve ser usada para o caso onde uma mesma molécula apresenta diferentes conformações. Este cálculo é muito difícil de realizar pelo método das perturbações, pois a trajetória de uma conformação a outra pode ser difícil de se construir e se modelar. O cálculo da energia livre absoluta é independente da trajetória, pois o que é desenvolvido é a diferença entre funções de estado termodinâmicos diferentes. A integração através de uma trajetória deve ser executada reversivelmente, o que é muito problemático para grandes rearranjos estruturais.

4.3 Calibração

A calibração e teste da rotina foi feita calculando-se a energia livre de estados conformacionais do *fentanil*⁽⁹⁾, o mesmo usado pelo pacote *Discover*, a partir do traçado da superfície de potencial de dois ângulos significativos da molécula, ângulo ϕ_1 da ligação $C_1-C_2-N_3-C_4$, e ângulo ϕ_2 da ligação $C_2-N_3-C_4-C_5$ (ver figura 4.1). Assim temos um padrão de cálculo prévio para com-

parar com nossos resultados. O *fentanil* é uma classe de componentes morfomiméticas, extensivamente estudada por cristalografia de raios-X, NMR e espectroscopia CD², por cálculos quânticos e por mecânica molecular. Apesar disso, muito pouca análise conformacional é encontrada na literatura.

Inicialmente obtivemos uma geometria planar de mínima energia usando o método AM1 do pacote *Geomop*⁽¹⁰⁾ e a seguir examinamos o comportamento conformacional em torno do nitrogênio anilido, forçando os ângulos assinalados a se manterem dentro dos valores escolhidos sem deformarem a estrutura geométrica básica. Desta maneira definimos os pontos de mínima e os de máxima do potencial (ver figura 4.2, onde os pontos A,B são os de mínima e C um ponto de sela).

Obtemos duas classes básicas de energia mínima no mapeamento. Nesta superfície potencial que varre os ângulos de 0° a 360° e de 0° a 180° (há simetria nas curvas de potencial em relação aos ângulos considerados) observamos que os dois poços de mínimo tem uma diferença na energia interna de aproximadamente contorno de 1.5 2.0 kcal/mol/Å. A barreira de energia de A para B é de cerca de 7.5 kcal/mol/Å e de cerca de 9.0 kcal/mol/Å no sentido oposto. O mapeamento em relação ao ângulo ϕ_2 não ultrapassa o valor de 180° porque há simetria neste mapeamento e a energia de transição torna-se muito alta devido a adoção do modelo de átomos unidos do programa Thor. Estes resultados obtidos com o Thor estão em concordância com os do *Discover*, o que mostra que os dois campos de forças em que eles se baseiam apresentam resultados equivalentes, com precisão dos campos de forças semelhantes, independentes da explicitação dos átomos de hidrogênio ou não.

O estado de referência tem dois papéis no cálculo da energia livre. A contribuição trivial é a energia livre como um sólido ideal. Esta contribuição é trivial porque independe da conformação da estrutura de referência. Ela depende unicamente das massas dos átomos e das constantes de mola vinculando cada átomo à rede. O papel não trivial é que ele determina qual parte do espaço configuracional está sendo analisado durante a dinâmica.

Calculamos as energias livres dos estados **A,B** do fentanil. Os pontos escolhidos para nosso cálculo são os mínimos localizados nos poços de potencial. Os estados de referência **A** e **B** correspondem a estruturas completamente minimizadas nestes pontos. O cálculo no ponto **C** representa a energia livre em uma barreira de potencial. Naturalmente será incorreto minimizar esta

²CD - dicroísmo circular

estrutura de referência. Ela é construída neste ponto forçando os ângulos diedrais a adotar os valores do ponto de sela.

É importante escolher a estrutura de energia mais baixa consistente com um dado estado conformacional, de maneira que a energia livre calculada não contenha excesso de entalpia e que se torna muito importante nestes cálculos de energia livre absoluta. O cálculo da energia livre absoluta obriga a amostragem do espaço configuracional a se manter em uma pequena região localizada no entorno da estrutura de referência. Isto limita a distribuição das energias. Movendo a estrutura de referência a uma conformação energética ligeiramente mais alta esta deve apresentar amostras de energias diferentes e mudar a energia livre total. Nossa amostragem não deve sair da região pré-assinalada transpondo a barreira de energia caindo em outro ponto de mínima. Devemos notar que não analisamos todo o espaço configuracional, mas nas diferenças de energia livre entre estados ligeiramente diferentes, que neste caso diferem basicamente pela rotação de dois ângulos diedrais.

4.3.1 Resultados para o Fentanil

A seguir apresentamos os resultados da integração numérica da energia livre do fentanil no estado de mínima energia **A**. Inicialmente temos os gráficos dos valores de $(V_0 - V_H)$ como função do tempo para os valores de λ para os quais geramos o gráfico médio (figura 4.4). Os valores da diferença de energia foram normalizadas para podermos comparar com os valores do programa *Discover* a partir dos resultados no mínimo conformacional em **A**, depois de 10 ps de termalização e 20 ps de medida. Devemos notar que as flutuações nesta grandeza aparecem apenas nos primeiros instantes da termalização e baixos valores de λ . Comparando com o gráfico apresentado no manual do *Discover*, notamos que neste último as flutuações são muito mais pronunciadas. Para valores de λ próximos a 1 esta grandeza apresenta visivelmente uma instabilidade. Em nosso caso, as flutuações são maiores para valores próximos a zero, mas em nenhum momento elas se caracterizam por instabilidades, mas ao contrário, depois de algum tempo elas se estabilizam.

Cada linha corresponde a valores diferentes de λ , partindo de 0.016 até 0.98 no topo do gráfico. O valor médio de $V_0 - V_H$ cresce monotonicamente com λ . Devemos notar que a função cresce rapidamente nas proximidades de $\lambda = 1$ indicando que V_0 é grande para *ensembles* dominados por V_H . Para valores próximos a $\lambda = 0$ a integral flutua inicialmente para depois manter um comportamento constante, mantendo sempre o menor valor. Isto

acontece porque V_0 , que controla a geração do *ensemble* neste ponto, está em um mínimo local. Notamos também que a integral tem um comportamento muito suave para valores intermediários de λ .

Para estados de referência que não correspondam a um mínimo no potencial tal como o estado C espera-se um comportamento semelhante em ambos os extremos de λ .

Na figura 4.5 vemos que não existe diferença significativa entre os valores das integrais de termalizações de 3ps ou de 20ps. A integral da função do gráfico de 0 a 1 em λ é a diferença na energia livre entre o fentanil em um estado de sólido harmônico ideal e o estado livre. Como para amostras com tempos maiores o resultado da integral pode ser influenciada para valores de λ próximos a 1, acreditamos que uma amostragem de 20ps é mais do que suficiente. Mesmo uma amostragem menor, de 3ps ou de 10 ps não modificam significativamente o valor da integral. Se comparamos com a figura correspondente do manual do programa *Discover*, este último tem um comportamento muito estranho com valores muito próximos para $0,2 < \lambda < 0,8$, e apresentando uma divergência pronunciada para valores de λ maiores do que 0,8. Como o manual não apresenta nenhuma tabela de saída, nem arquivo de entrada, não sabemos como este gráfico foi gerado.

tempo-3ps	$\langle V_0 - V_H \rangle$	INT F	LAMBDA
STEP 1	-118,507	-3,951	0,01305
STEP 2	-4,410	-4,280	0,06747
STEP 3	33,476	-0,613	0,16030
STEP 4	52,325	6,432	0,28330
STEP 5	56,790	14,823	0,42556
STEP 6	68,787	24,987	0,57444
STEP 7	77,520	35,424	0,71670
STEP 8	96,014	45,942	0,83970
STEP 9	156,772	57,657	0,93253
STEP 10	277,091	66,894	0,98695

tempo-20ps	$\langle V_0 - V_H \rangle$	INT F	LAMBDA
STEP 1	-131,496	-4,384	0,01305
STEP 2	-5,535	-4,797	0,06747
STEP 3	33,295	-1,150	0,16030
STEP 4	51,929	5,842	0,28330
STEP 5	58,315	14,458	0,42556
STEP 6	63,470	23,837	0,57444
STEP 7	78,940	34,465	0,1670
STEP 8	95,921	44,972	0,83970
STEP 9	160,146	56,939	0,93253
STEP 10	266,518	65,824	0,98695

Estas duas tabelas acima são os resultados de duas simulações do fentanil no estado de energia A para 3 e 20 ps, respectivamente, de termalização. A primeira coluna de valores corresponde a $V_0 - V_H$, a segunda o valor da integral e o terceiro o valor de λ , onde usamos dez divisões para este parâmetro. Os últimos valores da integral correspondem aos valores da energia livre absoluta calculada no ponto A. A tabela abaixo resume este cálculo nos pontos A e B.

estado tempo(ps)	A(kcal/mol)	B(kcal/mol)	$\Delta(B - A)$ (kcal/mol)
3	66,894	67,097	0,203
10	66,566	65,779	-0,787
20	65,673	67,087	1,414
30	65,824	68,132	2,308

Vemos que para cada região pré-definida, o tempo de amostragem influi relativamente pouco no cálculo do valor da energia livre. Aparecem flutuações em seu valor, em alguns casos bastante significativos, diminuindo o seu valor. No caso do fentanil parece que por volta de um tempo de amostragem de 20ps temos uma inversão nesta tendência para as duas regiões. Assim, escolhemos este tempo como um tempo médio razoável para o nosso cálculo, gerando uma diferença de energias livres de 1414 kcal/mol.

A diferença da energia entre as duas regiões amostrais aumenta a medida que se aumenta o tempo total de medida. Como no tempo de 10ps de amostragem, na região B a energia livre diminuiu drasticamente, a diferença

entre estes estados caiu para valores negativos, indicando que este ainda não é um tempo razoável de amostragem. Muito acima de 20 ps, a diferença aumenta muito, indicando que aumenta muito o número de erros aleatórios naturais em cálculos de energia livre.

Temos também os gráficos dos ângulos varridos nos estados conformacionais **A**, **B** (ver figura 4.3) e vemos que eles se comportam de maneira boa, não saindo das proximidades das regiões amostrais pré-definidas inicialmente, não ultrapassando a barreira de potencial e passando ao outro mínimo local. Este gráfico não tem similar no pacote *Discover* e, portanto, não sabemos se os resultados deste último estão de acordo com a boa técnica de cálculo.

Todos os nossos resultados são numericamente semelhantes aos apresentados pelo pacote *Discover* para a mesma molécula, quando este apresenta algum resultado. Alguns dos resultados do programa supra citado não aparecem baseados em números e assim vimo-nos impossibilitados de comparar com os nossos.

Análise de Erros

Nestes cálculos temos fontes de erros tanto sistemáticos quanto aleatórios. A maior fonte de erros sistemáticos nestes cálculos é a falta de convergência, resultado na falha em equilibrar termodinamicamente a cada valor de λ , e uma amostragem insuficiente do espaço configuracional. Para superar em parte estes problemas fazemos uso de uma dinâmica normal de 5ps e a seguir de uma dinâmica termalizada por meio de algoritmo de Berendsen - que mantém o sistema em uma temperatura determinada - também de 5ps. Outras fontes de erros sistemáticos incluem campos de forças sem a acurácia devida que descartamos como premissa de trabalho.

Erros aleatórios são naturais em cálculos de energia livre. A distribuição estatística dos estados acessíveis para uma molécula a uma dada temperatura é precisamente o que define sua entropia. A medida da entropia é um processo estatístico que pode ser quantificado com procedimento de análise de erros.

Para correções de cálculo para valores de λ próximos aos extremos, devemos usar um tempo de equilíbrio e amostragem menor para computar a energia livre entre os valores bem comportados, por exemplo de 0 a 0.2. O segundo cálculo deve computar a energia livre de 0.2 a 1 com tempos de amostragem e medida longos. A energia livre total é justamente a soma das duas, tomando-se o cuidado de somar apenas uma vez a contribuição do

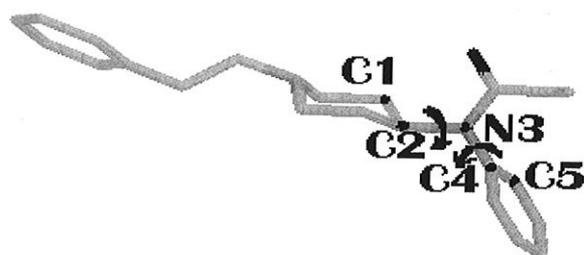
sólido ideal. Matematicamente, isto corresponde a fazer:

$$\int_0^1 d\lambda \langle V_0 - V_H \rangle = \int_0^{0.2} d\lambda \langle V_0 - V_H \rangle + \int_{0.2}^1 d\lambda \langle V_0 - V_H \rangle. \quad (4.23)$$

Não fizemos nossos cálculos com esta análise mais fina, pois estamos interessados em resultados qualitativos comparáveis com o que temos na literatura.

Dados de Saída Para análise da energia livre, utilizamos a saída normal do programa Thor com duas partes que aparecem em seqüência. A primeira parte, a fase de termalização com as colunas correspondendo as saídas normais. Em seguida, a parte do cálculo da energia livre com valores das colunas correspondendo a tempo, temperatura, energia cinética, energia potencial, potencial de referência, valor médio da diferença de energia dos últimos passos, valor médio da diferença das energias desde o início da computação, largura do intervalo de λ instantaneamente calculados, e no final de cada seqüência de cálculo, um resumo com os valores da média das diferenças entre as energias $\langle V_0 - V_H \rangle$, da integral para valor de λ dado, e o valor de λ usado no cálculo no esquema de integração de Gauss-Legendre⁽¹²⁾. Logo a seguir uma nova tabela com o novo λ e assim, sucessivamente, até completar o número de divisões de λ determinado. Para realizar estes cálculos é necessário introduzir em seu cabeçalho (ver tese de doutoramento de Pedro Geraldo Pascutti⁽¹⁾) os comandos FREEEN e FRSTEP=número de divisões de λ , na entrada do programa. Se este número não for colocado o programa calculará λ com 4 divisões (*default* = 4). Além dessas informações é necessário um arquivo que vai indicar o tempo de termalização e o de medida e que tem a terminação .FRE.

Dependendo do tempo de medida e do *ensemble* de amostragem os arquivos de saída podem ser enormes. Para remediar em parte este problema, o programa escreve somente os resultados depois de completados dez ciclos de cálculo.



ÂNGULO 1 = C1-C2-N3-C4

ÂNGULO 2 = C2-N3-C4-C5

Figura 4.1: Fentanil com a definição dos ângulos diedrais significativos no cálculo da energia livre.

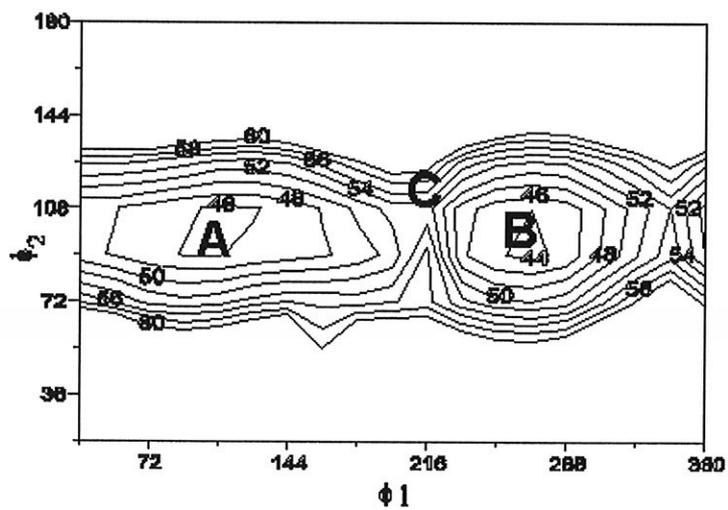


Figura 4.2: Mapa da superfície de energia do fentanil, em kcal/mol. As regiões A e B são de energia mínima e na região C temos um ponto de sela.

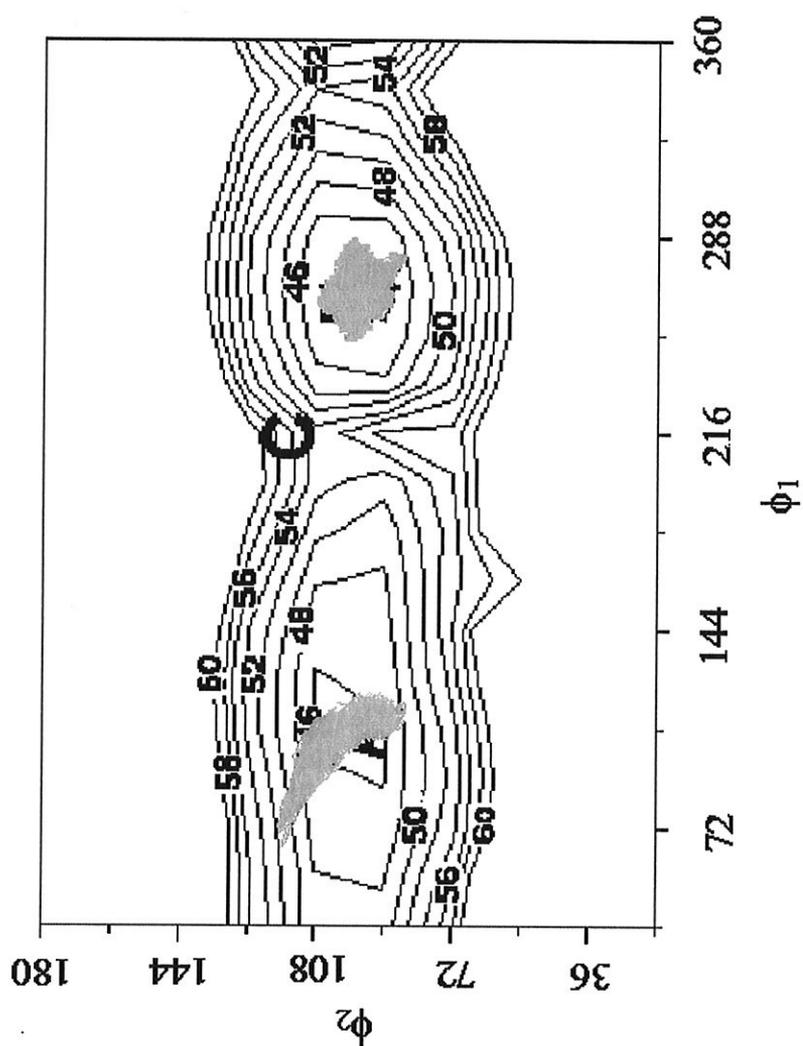


Figura 4.3: Ângulos varridos durante o cálculo da energia livre nas regiões A e B definidos na figura 4.2. Vemos que a molécula não se afasta da região predeterminada passando para outra a outra região.

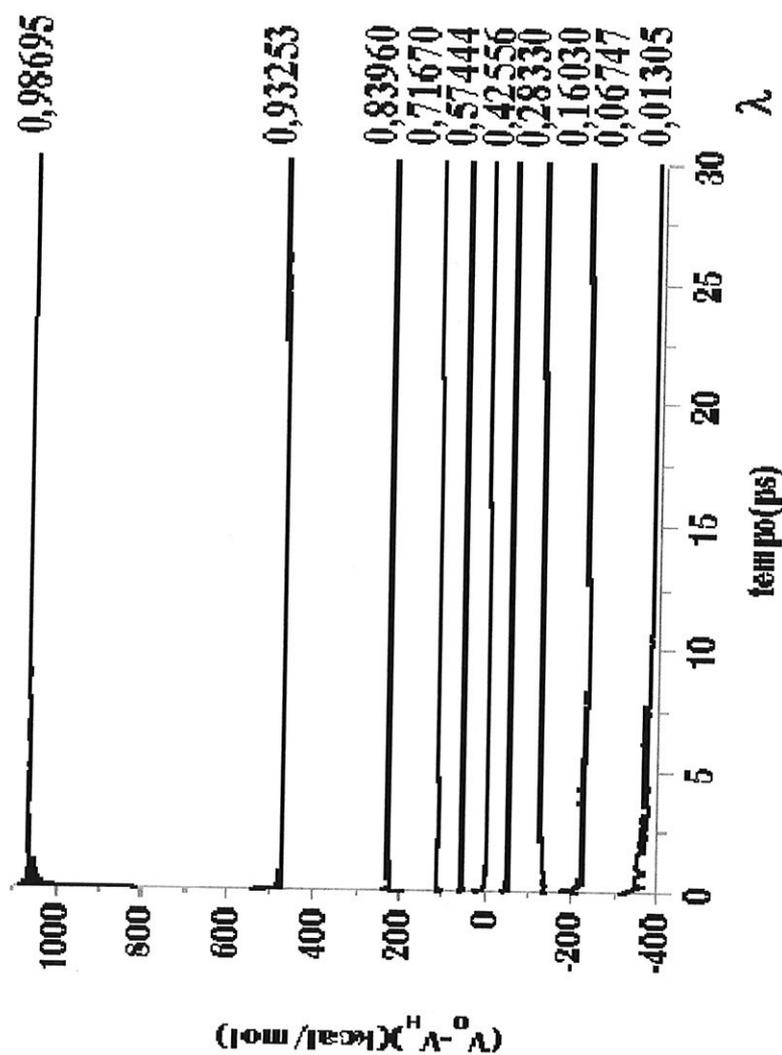


Figura 4.4: Exemplo do comportamento de $V_0 - V_H$. Podemos notar que após os primeiros 10 ps de termalização seu comportamento praticamente se estabiliza. Temos mais 20 ps de medida. Os valores desta diferença de energia foram normalizados para compararmos com os que aparecem no manual do programa Discover.

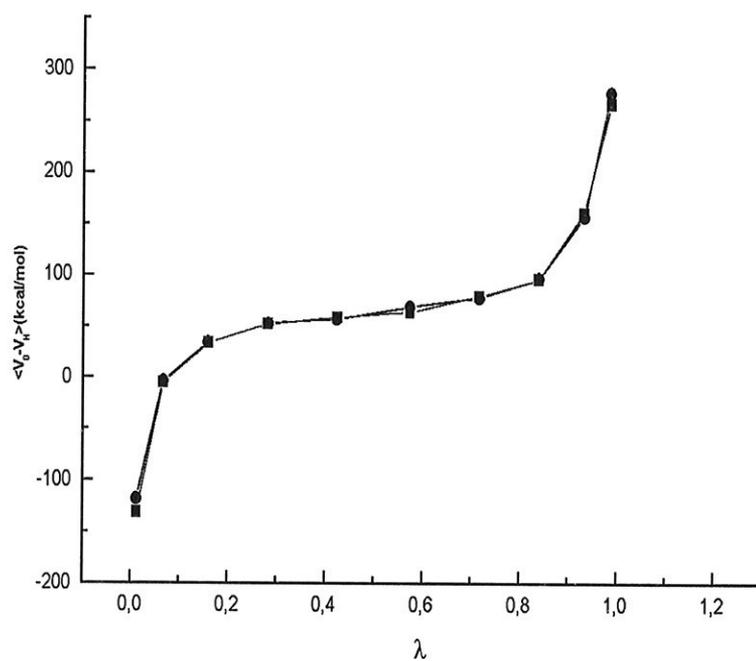


Figura 4.5: Comparação dos valores de $\langle V_0 - V_H \rangle$ para 3ps e 20 ps de medida. Vemos que as diferenças entre as duas simulações são muito pequenas.

Capítulo 5

O Triptofano

Neste capítulo analisamos a possibilidade de transição entre os seis rotâmeros do triptofano⁽¹³⁾. Para isto mapeamos a função energia interna total em função dos valores dos ângulos $\chi_1 = N_\alpha - C_\beta - C_\gamma - C$ e $\chi_2 = C_\alpha - C_\beta - C_\gamma - C_\delta$ da figura 5.1, os dois ângulos significativos do triptofano. Objetivamos obter os mínimos de energia correspondentes a cada rotâmero, bem como a altura das barreiras de energia. Determinadas as configurações nestes extremos, passamos ao cálculo de energia livre para verificarmos a possibilidade de interconversão de uma estrutura a outra. Os cálculos de otimização de geometria foram realizados no programa Thor, com o uso do método do gradiente conjugado com o critério de convergência 10^{-5} kcal/mol. O mapeamento de energia utilizado para definirmos a configuração exata do triptofano em cada extremo, com variação de 10^0 de uma geometria a outra. Para os cálculos de energia livre utilizamos oito divisões de λ , com 10,0ps de termalização e 20,0ps de medida, e variação de 60^0 entre as geometrias. Calculamos também as energias livres do triptofano em cada ponto de mínima energia e nos pontos de sela.

Cálculos

Utilizamos o modelo de átomo-unido para o triptofano *zweiteriônico* e o mapeamento da energia foi construído obrigando os ângulos diedrais a se manterem próximos de seus valores iniciais com a aplicação de uma força de vínculo suficiente grande. Temos uma clara localização dos seis mínimos de energia, bem como os pontos de máximos das barreiras de potencial. Resultados prévios^(13,44-57) indicam que as alturas das barreiras de poten-

cial entre as várias conformações mudam com a representação explícita dos átomos de hidrogênio para representação de átomo-unido, indicando que o mapeamento de energia será ligeiramente diferente em cada caso. Quando o solvente é representado como contínuo, ou quando as suas moléculas são representadas explicitamente, também apresentará outro mapeamento, independente do tipo de modelo utilizado, com localização diferente dos máximos e mínimos⁽¹⁾. Os ângulos diedrais representativos do triptofano $\chi_1 = N_\alpha - C_\beta - C_\gamma - C$ e $\chi_2 = C_\alpha - C_\beta - C_\gamma - C_\delta$ estão definidos na figura 5.1.

Os ângulos que obtivemos para os seis rotâmeros foram os seguintes:

		χ_{1g^+}	χ_{1t}	χ_{1g^-}
ant χ_2	χ_1	55,33	186,93	292,98
	χ_2	269,79	258,91	269,73
perp χ_2	χ_1	62,41	188,06	294,87
	χ_2	103,08	87,32	95,93

com as respectivas energias em kcal/mol:

	χ_{1g^+}	χ_{1t}	χ_{1g^-}
ant χ_2	16,19	15,05	15,90
perp χ_2	15,66	15,67	15,61

Aqui nós devemos considerar que a estrutura ant χ_2t é aquela que apresenta a menor energia total, seguida pelas estruturas perp $\chi_2gauche^-$, $gauche^+$ e $trans$ e finalmente pelas formas ant $\chi_2gauche^-$ e ant $\chi_2gauche^+$. As primeiras estruturas são as candidatas naturais a aparecer mais vezes nas medidas experimentais, se nossa ferramenta de cálculo estiver correta. Resultados de dinâmicas moleculares anteriores realizados por Pedro G. Pascutti⁽¹⁾ para o mesmo meio $\epsilon = 80$, apresentaram a seguinte distribuição de rotâmeros:

		χ_{1g^+}	χ_{1t}	χ_{1g^-}
ant χ_2	%	7,43	20,25	18,20
perp χ_2	%	8,91	30,54	14,67
totais	%	16,34	50,79	32,87

Por estes resultados podemos verificar que se as porcentagens com que os rotâmeros são visitados durante uma dinâmica não coincidem exatamente com a distribuição de energias mínimas; vemos também que eles podem ser classificados na tabela anterior de maneira vertical. Assim, notamos que os rotâmeros $trans$ aparecem com uma frequência maior seguida pelos rotâmeros

gauche⁻ e por fim pelos rotâmeros gauche⁺. Do mapeamento da energia em função dos ângulos diedrais χ_1 e χ_2 (figura 5.2) definimos os pontos de sela.

Definidos os pontos de mínima energia e os pontos de sela, calculamos as energias livres nestas posições. Inicialmente, fizemos este cálculo entre as estruturas ant χ_2 e perp χ_2 , com a constante dielétrica de $\epsilon = 80^1$, e temperatura de 300K. Estes resultados estão tabelados a seguir, em kcal/mol, onde bar representaria a “barreira” de energia:

	bar	χ_{1g^+}	bar	χ_{1t}	bar	χ_{1g^-}	bar
bar		30,442		25,118		25,207	
ant χ_2	23,970	26,518	27,872	33,329	29,074	31,992	23,970
bar		25,037		25,437		24,889	
perp χ_2	23,708	29,261	30,852	28,271	22,685	27,695	23,708
bar		30,442		25,118		25,207	

Verificamos rapidamente que estes valores estão muito diferentes dos valores da energia interna total. Além disso, as energias livres das “barreiras” são, em geral, menores do que às calculadas nos pontos de mínima energia interna total. Por isso, mapeamos os valores de energia livre de 60° em 60° . Supomos que esta variação é razoável, pois os mínimos de energia estão localizados aproximadamente a esta variação do ângulo χ_1 . Na superfície de energia livre obtida desta maneira, temos agora somente três mínimos localizados em 180° , 300° e 360° do ângulo χ_1 e em 120° e 240° do ângulo χ_2 (ver figura 5.3). Notar que nesta figura os eixos vertical e horizontal não terminam em 360° . Isto se deve à dificuldade que tivemos para construir este mapa, pois o número de pontos não corresponde ao número de divisões do programa *Origin*. Para maior clareza, apresentamos a tabela da energia livre em função dos ângulos diedrais em kcal/mol:

$\chi_2 \backslash \chi_1$	60°	120°	180°	240°	300°	360°
60°	25,68	30,79	23,14	26,00	24,66	26,64
120°	25,12	26,70	19,40	23,04	25,90	20,13
180°	24,47	27,60	25,44	27,80	28,31	30,61
240°	24,89	28,35	28,38	26,90	22,20	28,94
300°	28,26	24,86	28,08	25,93	26,82	25,56
360°	23,81	28,80	24,94	31,35	27,12	27,52

¹A constante dielétrica interfere somente nas interações das forças internas (ver equação 2.6)

Notemos que todas as geometrias de energia interna mínima representariam barreiras de energia livre e dificilmente seriam alcançadas. A estrutura $\text{perp}\chi_1\text{t}$ estaria deslocada no ângulo χ_2 de $87,32^\circ$ para 120° . Esta é a única estrutura que está dentro de uma bacia de mínimo de energia interna total. As outras duas estruturas de mínimo de energia livre correspondem a pontos próximos a pontos de sela da energia interna, apesar de uma delas, a de ângulo diedral χ_1 de 360° , a estrutura $\text{perp}\chi_2$ encontrar-se perto da bacia de mínimo de energia interna da estrutura gauche^+ . Procuramos a razão para esta mudança tão radical de um mapeamento a outro, descontadas as diferenças no número de divisões angulares, no fato do programa Thor utilizar o modelo de átomo-unido. Certas configurações difíceis de acontecer na natureza podem aparecer devido à ausência do átomo de hidrogênio. Isto pode ser verificado pelas figuras dos isômeros que nestes cálculos apresentaram a menor energia livre, aparecendo em ordem o $\text{perp}\chi_2(120^\circ)\text{t}$ (figura 5.4), $\text{perp}\chi_2-360^\circ(\text{perp}\chi_2\text{g}^+)$ (figura 5.5), e $\text{ant}240^\circ-300^\circ(\text{ant}\chi_2\text{g}^-)$ (figura 5.6), onde os números definem os ângulos de χ_1 e de χ_2 .

Estes resultados têm importantes implicações para estes modelos rotaméricos do triptofano. Neles são propostos decaimentos fluorescentes multiexponenciais que corresponderiam à presença de um número de estados estacionários rotaméricos, alguns dos quais não se interconverteriam na escala de tempo fluorescente (da ordem de 3ns a 5ns). Assume-se que os rotâmeros individuais tenham decaimento monoexponencial. A relação do decaimento fluorescente com componentes individuais rotaméricos é altamente especulativo. Szabo e Rayner⁽⁵⁵⁾ verificaram experimentalmente um decaimento biexponencial no estudo do triptofano zwitteriônico. Como muitos estudos partiam da hipótese de seis rotâmeros, imaginaram uma rápida interconversão de uma estrutura em outra, dentro da escala de tempo. Outros modelos rotaméricos supõe que os rotâmeros não possuem interconversão durante o tempo de vida do estado excitado. Estudos *via* simulação obtém os dois resultados, dependendo do modelo de campo de forças e parametrização dos programas (ver descrição destes modelos nas referências).

As estruturas do triptofano livre em cristais é predominantemente da estrutura $\text{perp}\chi_2\text{g}^+$. Já em cristais de peptídeos e proteínas o confômero $\text{perp}\chi_2\text{g}^-$ do triptofano é favorecido, o que coincide com resultados encontrados em solução (ver ref.13).

Conclusões

Se nossos cálculos de energia livre estiverem certos, podemos dizer que para o modelo de programa que se utiliza de átomo-unido como o Thor, apesar da possibilidade de seis rotâmeros, apresenta, neste tipo de cálculo, unicamente a probabilidade de se encontrar no máximo três deles, sendo que um deles é um mínimo pouco pronunciado. Nossos resultados estão muito mais próximos de todos os resultados experimentais conhecidos, do que os resultados de uma simples dinâmica onde apenas a energia interna total é considerada. Entretanto, este resultado não pode ser tomado como conclusivo. Para isto acontecer, pretendemos completar esta análise, colocando, no programa, os átomos de hidrogênio explicitamente, e, depois, verificar se os rotâmeros que podem ser visitados num cálculo de energia livre são realmente três, seis ou um outro número qualquer. Além disso, devemos executar estes cálculos com a constante dielétrica $\epsilon = 2$, para simular um meio correspondente ao das proteínas, e verificar se acontece algum deslocamento angular de um dos mínimos. Provavelmente este caminho seja o mais frutífero, ao invés de procurarmos mudar a constante de mola. É muito difícil que tenhamos outros pontos de mínima em uma varredura de ângulos menor como fizemos com a energia interna, mas esta possibilidade não pode ser descartada. Neste último caso, com certeza, teremos uma melhor localização dos mínimos de energia livre, além de conhecermos as alturas das suas barreiras, que possam impedir a passagem de uma forma a outra.

Adendo

A.G.Szabo and D.M.Rayner propuseram em artigo publicado no *J.Am.Chem.Soc.*, 102, 1980, 554 um modelo onde o decaimento biexponencial da fluorescência do triptofano seria devido a existência de várias conformações rotaméricas para a molécula. Não fica claro, todavia, de que modo os seis possíveis rotâmeros resultam apenas dois tempos de vida de decaimento. Um cálculo usando a probabilidade estatística de Boltzmann para os três mínimos mais prováveis dos cálculos de energia livre, conduzem a aproximadamente 76,61%, 22,69% e 0,72%. Outros valores de energia livre têm valores probabilísticos muito menores do que este último. Se tomarmos os fatores pré-exponenciais do decaimento biexponencial, que refletem as populações de cada espécie rotamérica, e compararmos com os nossos resultados, verificamos que eles tem a mesma ordem de grandeza, que gira em torno de 2,8.

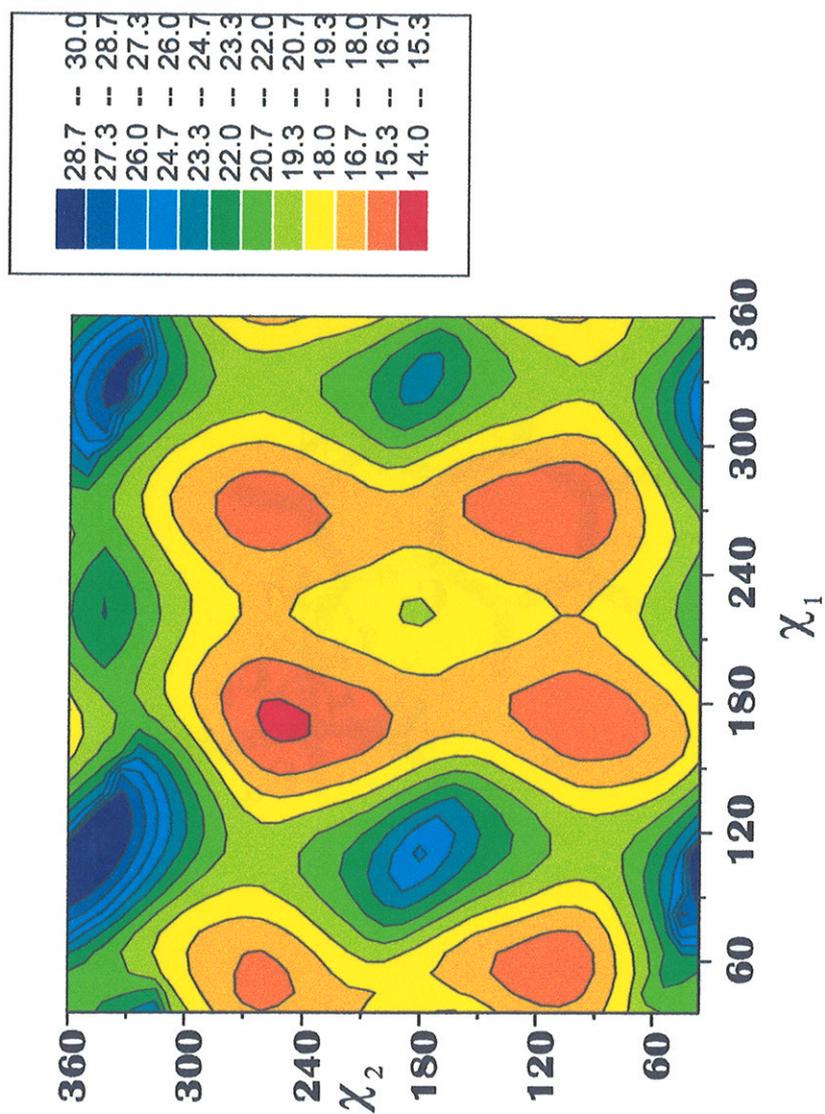


Figura 5.2: Mapa da superfície da energia interna do triptofano (em kcal/mol) em função dos ângulos diedrais definidos na figura 5.1.

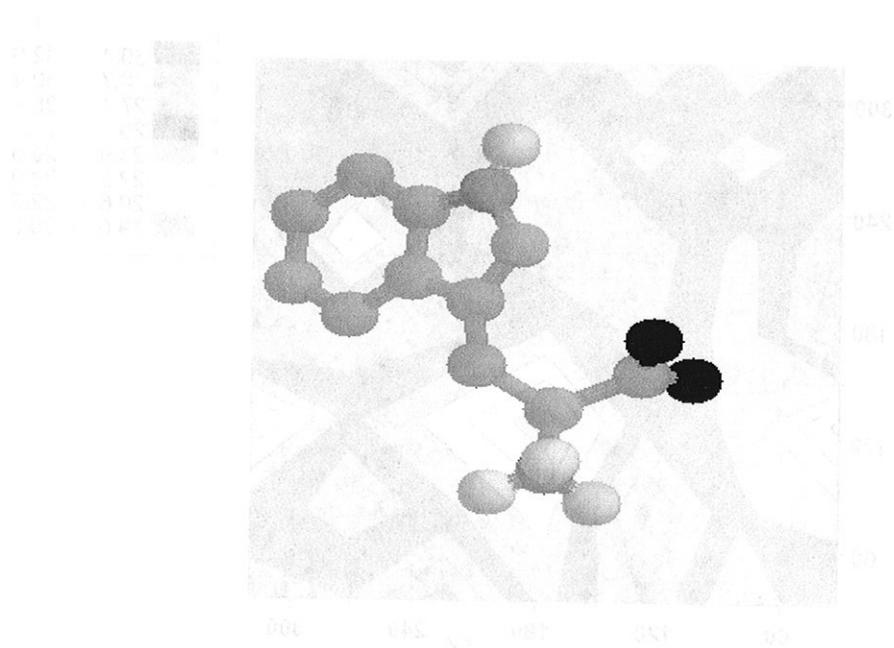


Figura 5.3: Molécula de triptofano com os ângulos diedrais de $\chi_1 = 180^\circ$ e $\chi_2 = 120^\circ$.

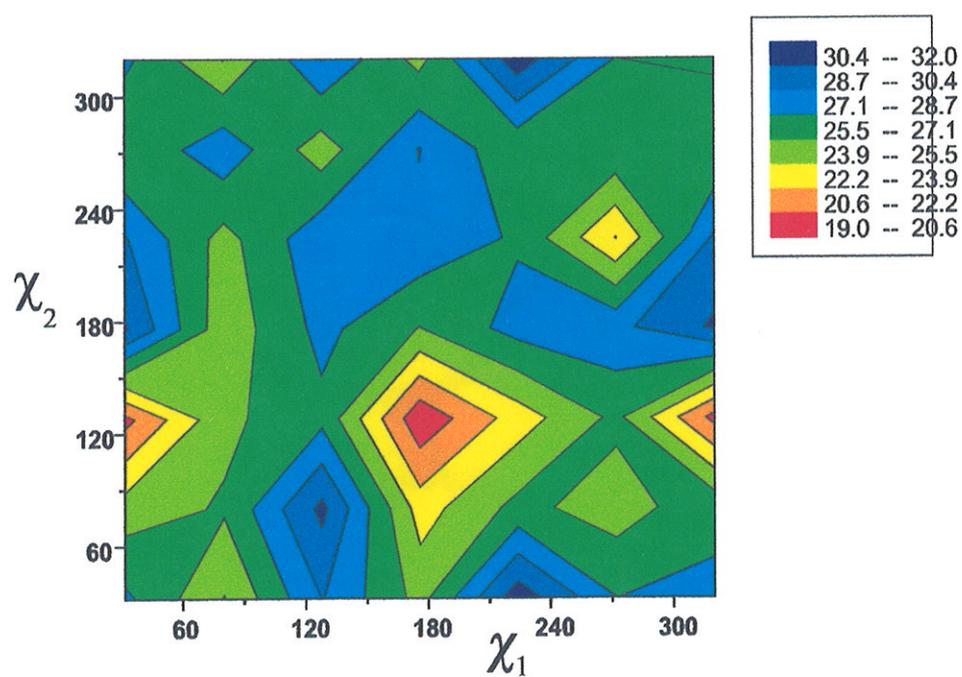


Figura 5.4: Superfície de energia livre (em kcal/mol) do triptofano em função dos ângulos diedrais.

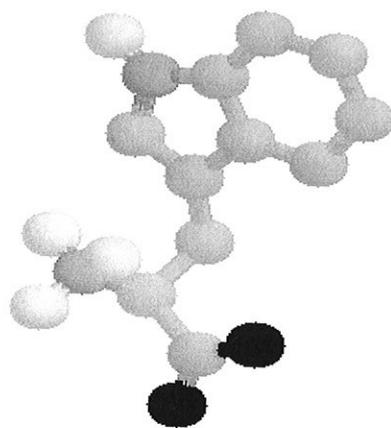


Figura 5.5: Molécula de triptofano com os ângulos diedrais de $\chi_1 = 360^\circ$ e $\chi_2 = 120^\circ$.

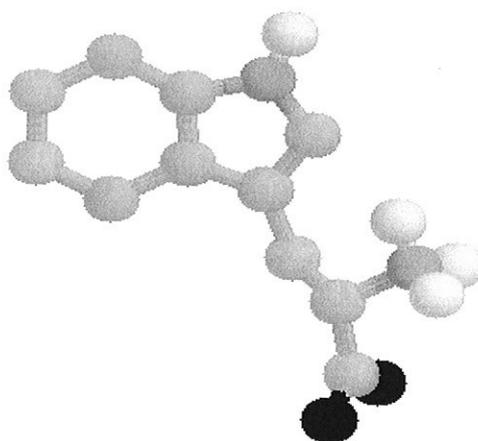


Figura 5.6: Molécula de triptofano com ângulos diedrais de $\chi_1 = 300^\circ$ e $\chi_2 = 240^\circ$.

Parte III
Cálculos Quânticos

If mathematical analysis should ever hold a prominent place in chemistry - an aberration which is happily almost impossible - it would occasion a rapid an widespread degeneration of that science. (? ,A.Comte)

Apresentamos nesta parte resultados de cálculos quânticos. Os cálculos na thiosemicarbazona mostram todas as possibilidades dos métodos quânticos, permitindo especular sobre uma possível trajetória de interconversão isomérica. Já os cálculos com o Abz e seus compostos com a fenilalanina, a prolina e a glicina mostram que os métodos semiempíricos apresentam bons resultados quando referidos a geometria ou a energia de absorção ou de excitação.

Capítulo 6

A Thiosemicarbazona

Muitos derivados da n-pyridinacarbaldéidos thiosemicarbazonas (PATS-n) constituem modificadores das superfícies de eletrodo capazes de tornar ativas eletroquimicamente espécies que em superfícies não modificadas não apresentam resposta eletroquímica, como é o caso de muitas metaloproteínas. Os compostos PATS-n são considerados modificadores de superfície universais pois eles induzem respostas eletroquímicas reversíveis tanto com as proteínas positivamente quanto com sítios de ligação negativamente carregados, tais como a citocroma C e proteínas de cobre azul, respectivamente, no eletrodo⁽²³⁾.

Entre suas funções como modificadoras de superfície, algumas moléculas derivadas da thiosemicarbazona apresentam relevantes atividades biológicas. A família thiosemicarbazona é ativa na terapia do câncer, na tuberculose e nas infecções bacteriais e virais (incluindo a AIDS). A investigação da possível correlação entre traços estruturais e atividade farmacológica constitui portanto tópico relevante na pesquisa nesta área^(24,25,26,27,28,29,30).

Dados de difração de raios-X no PATS-2⁽¹⁴⁾ indicam que sua estrutura cristalográfica é formada por dímeros na configuração E, ostentando um centro de simetria, ligado por ligações de hidrogênio por meio de grupos thioamida⁽²⁹⁾. Moléculas de água também estão envolvidas na ligação. A configuração E é uma das duas formas isoméricas apresentadas por compostos PATS-n. A outra configuração encontrada é denominada forma Z. A isomerização do PATS-2 de Z para E tem sido estudada por espectroscopia vibracional Raman e ressonância nuclear magnética e a evidência da ocorrência espontânea da interconversão $Z \leftrightarrow E$ tem sido verificada⁽³¹⁾. Para simular a transição configuracional entre as duas formas isoméricas, empregamos uma

análise semiempírica, a partir de uma rotação supostamente proibida na ligação dupla (C7 = N1').

6.0.2 Resultados dos cálculos da interconversão Z↔E

Inicialmente modelamos a interconversão Z↔E onde definimos oito estruturas iniciais para cada um dos dois isômeros. Os resultados indicam que a ligação CN do PATS-2 tem um valor significativamente mais baixo que o esperado para uma ligação dupla convencional do mesmo tipo, sugerindo que uma rotação neste ponto na molécula, uma condição para a isomerização observada, é factível. A maior estabilidade do isômero E, observada experimentalmente, deve ser condicionada a considerações energéticas e entrópicas.

A seguir traçamos o caminho de transição forçando a estrutura ficar no estado de transição e liberando para que o sistema fosse naturalmente para as duas estruturas preferenciais dos estados Z e E.

Determinação das conformações de menor energia

As várias conformações possíveis da Thiosemicarbazona foram determinadas a partir das configurações iniciais das suas duas formas isoméricas construídas de parâmetros estruturais da cristalografia de raios-X e constantes de forças obtidas de estudos vibracionais de espécies moleculares semelhantes, com a variação dos ângulos definidos na figura abaixo através do pacote Geomop (Ver figuras 6.1, 6.2 3 6.3).

Usamos cálculos SCF, empregando o método AM1. O critério de convergência usado foi o de 10^{-5} unidades atômicas e as geometrias de energia mínima foram obtidas pelo método de otimização de Rinaldi. Definimos as conformações para a obtenção dos vários conformeros segundo a tabela abaixo:

c ₁	s-cis 1'2'
t ₁	s-trans 1'2'
c ₂	s-cis 2'3'
t ₂	s-trans 2'3'
c ₃	s-cis 2 7
t ₃	s-trans 2 7

Obtivemos os mínimos locais destes conformeros antes de iniciarmos a rotação em torno do dupla ligação C7=N1'.

Impomos rotações à dupla ligação C7=N1' no intervalo de 0° a 180° , com passos de 30° , partindo das várias conformações obtidas com base nos elementos de conformação da tabela acima. Objetivamos com isto traçar uma trajetória possível da conversão da forma isomérica Z para a forma isomérica E e vice-versa, monitorando as energias associadas a cada passo intermediário.

A partir de cada um destes estados de mínima energia assim obtidos, forçamos a molécula a uma conformação geométrica de mais alta energia - um estado de transição - como ponto de partida para gerarmos a trajetória de cada um destes estados de transição para o seu estado de equilíbrio, de acordo com o método de Newton-Raphson. Dessa maneira fomos capazes de traçar os possíveis caminhos que a molécula pode realizar na sua interconversão entre as duas conformações que estão mapeadas a seguir na figura 6.4.

Nas figuras 6.2 e 6.3 podemos notar a existência de uma grande probabilidade da formação de uma ligação intramolecular no isômero Z entre o hidrogênio e o nitrogênio. Esta possibilidade não existe no isômero E, onde o nitrogênio piridinico é muito distante do hidrogênio ligado ao N2' da thiosemicarbazona.

Ao examinar as mudanças na energia de formação ao longo da trajetória a partir da conformação Z até a conformação E passando pelo estado de transição (TS), verificamos que todos os estados de transição exibem uma configuração linear C7-N1'-N2', o que pode ser verificado pelo vetor de transição (figura 6.5) que age em ambas as conformações.

Na figura 6.4, temos a energia de formação em função do tipo de conformação. Vemos ainda que os vetores de transição (figura 6.5) indicam claramente um transição do tipo inversão ao invés de uma rotação. Isto explica a interconversão simultânea s-cis/s-trans 2'-3' e Z/E que observamos aqui. Não é possível caracterizar qualquer estado de transição entre as conformações $t_1t_2c_3E/t_1c_2c_3Z$, como pode ser verificado pela perda da linha unindo estas conformações. Isto pode ser supostamente devido ao fato de que as conformações mencionadas estão entre as de energia mais alta e assim as transições entre famílias de conformações antes da conversão Z/E são menos prováveis neste caso. A energia mínima no estado de transição pode ser encontrada para as conformações $c_1t_2c_3E/c_1c_2c_3Z$. A trajetória para a isomerização Z-E pode envolver transições conformacionais interfamiliares e assim não deve ocorrer necessariamente através de uma única trajetória direta; conformações intermediárias podem muito provavelmente estar envolvidas neste processo.

A figura 6.6 mostra a energia de ativação do estado de transição para as várias famílias conformacionais modeladas. As conformações $c_1t_2c_3E/c_1c_2c_3Z$ e $t_1t_2t_3E/t_1c_2t_3Z$ são as que apresentam a energia de ativação mais baixa. Não conseguimos encontrar a energia de transição para a transformação $t_1t_2c_3E/t_1c_2c_3Z$, dentro dos limites energéticos das outras transições, indicando a probabilidade praticamente nula dela ocorrer. Finalmente $c_1c_2c_3E/c_1t_2c_3Z$ é também uma conformação de energia mínima. Neste caso, é necessário analisar as energias de formação e de ativação quando conformações mais prováveis e trajetórias de interconversão estão sendo estudadas. A conformação $t_1c_2t_3Z$, embora possua uma energia de ativação pequena, apresenta uma alta energia de formação para ser considerada como uma conformação inicial provável para o isômero Z (figura 6.6).

A ordem de ligação calculada para os isômeros Z e E na conformação $c_1t_2c_3$ mostrada nas figuras 6.7 e 6.8 dos dois isômeros apresentam valores consideravelmente inferiores na ligação C7=N1' do que aquelas que se esperam para ligação dupla do mesmo tipo. Este fato aponta para a possibilidade de ocorrência de uma rotação sobre esta ligação na molécula, pois o caráter de dupla ligação diminui significativamente neste caso.

Conformações de energia mínima para os isômeros Z e E estão mostradas nestas figuras 6.7 e 6.8. Ambos os isômeros estão muito próximos da planaridade, mas a planaridade é maior na configuração Z. O isômero E tem uma estrutura que pode ser descrita muito aproximadamente por dois planos, um relacionado ao anel piridínico e o outro com a parte restante da thiosemicarbazona (figuras 6.9 e 6.10). Podemos associar a menor planaridade a um número maior de conformações equivalentes, correspondendo a uma entropia maior. Isto significa que, unicamente com o critério entrópico, o isômero E deve ser a configuração mais provável para o PATS-2. Por outro lado, considerações entálpicas favorecem a ocorrência do isômero Z, embora por uma leve diferença na energia (cerca de 5kcal/mol) relativamente à forma E. Esta pequena diferença de energia, com a contribuição entrópica mais favorável para a forma E, pode explicar a maior estabilidade verificada para este isômero.

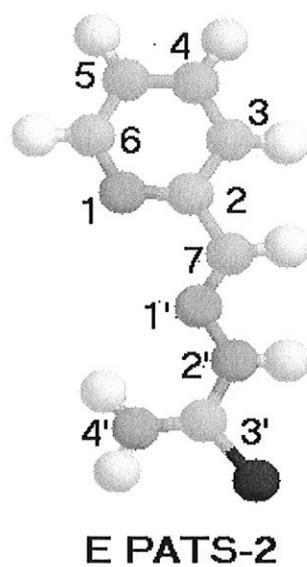
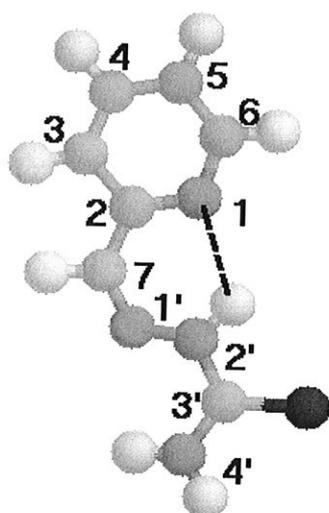


Figura 6.1: O isômero E PATS-2



Z PATS-2

Figura 6.2: O isômero Z PATS-2.

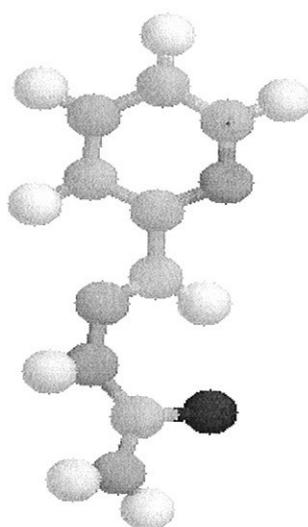


Figura 6.3: Isômero E PATS-2 com energia mínima. Vemos claramente uma rotação na ligação C7-N1'.

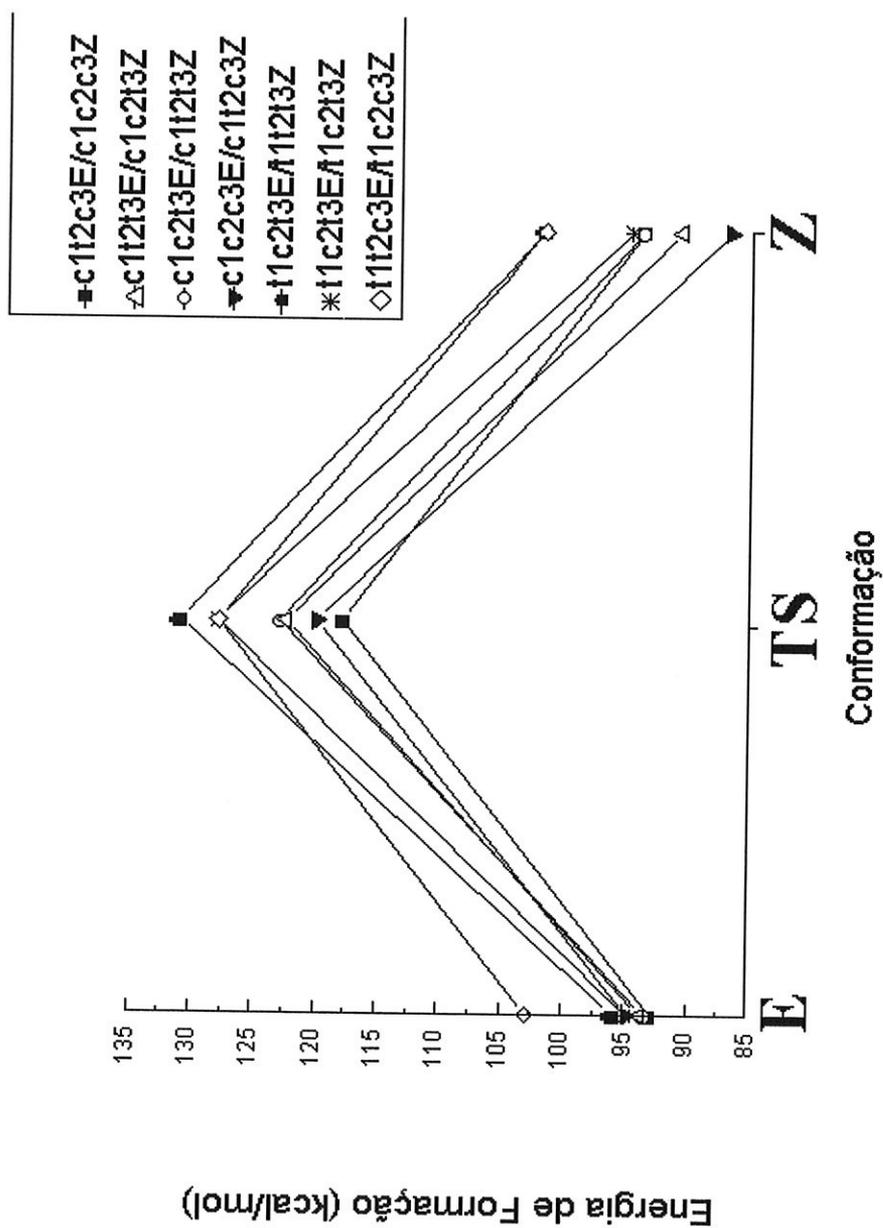
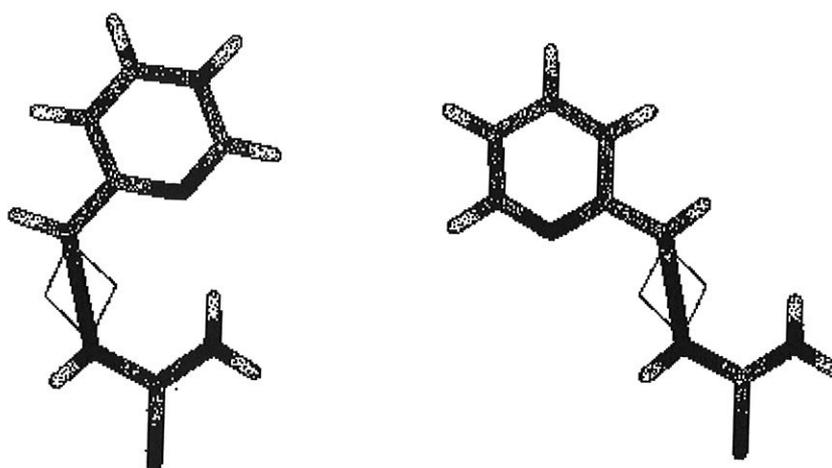


Figura 6.4: Energia de formação associada a várias famílias conformacionais ao longo da interconversão $E \rightarrow Z$ através do estado de transição.



Vetor de Transição

Figura 6.5: Vetores de transição. Vemos que os átomos C7-N1'-N2' estão alinhados.



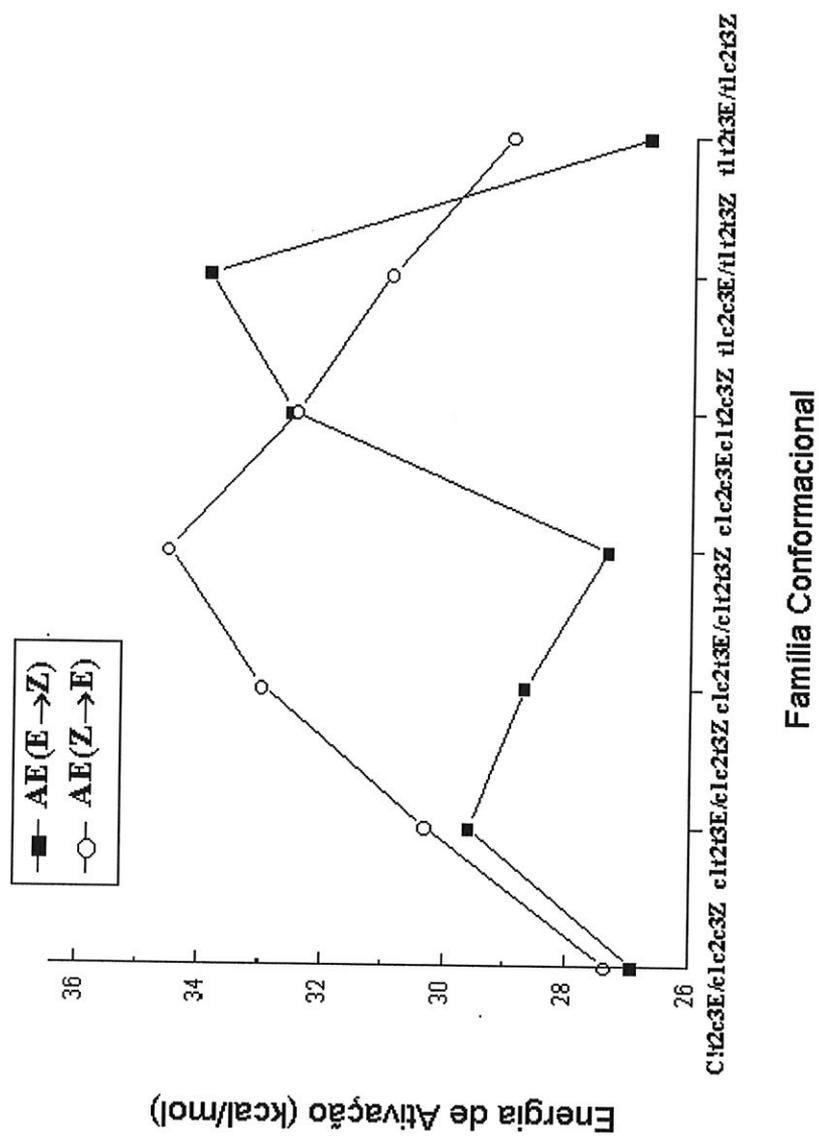


Figura 6.6: Energia de ativação para as várias famílias conformacionais ao longo da interconversão E→Z.

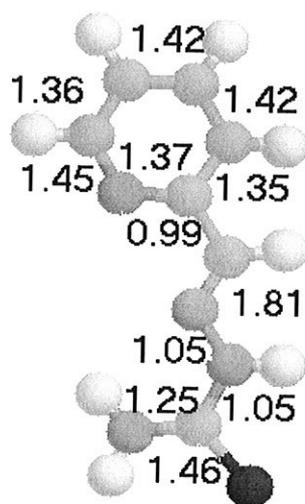


Figura 6.7: Distâncias do isômero E PATS-2 em sua conformação de menor energia.

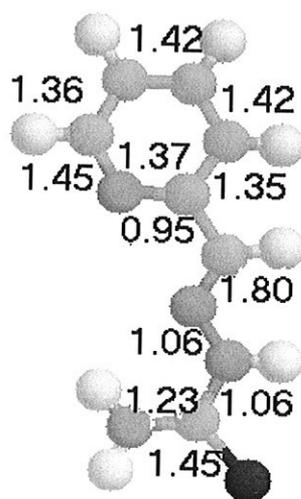


Figura 6.8: Distâncias do isômero Z PATS-2 em sua conformação de menor energia.

**ISÔMERO Z**

Figura 6.9: O isômero Z apresenta total planaridade.



ISÔMERO E

Figura 6.10: Podemos notar que o isômero E apresenta dois planos de simetria que conduzem a uma entropia maior que a do isômero Z.

Capítulo 7

O Abz

tu quoque!

O ácido ortoaminobenzóico⁽¹⁵⁾, também conhecido como ácido antranílico, possui uma intensa atividade fluorescente^(37,38,39,40). Esta propriedade permite que se possa usá-lo como uma sonda externa nos estudos das biomoléculas. Sua baixa solubilidade ($\sim 10^{-5}$ Molar) em água^(41,42,43) permite concluir que a simulação de uma única molécula isolada não deve diferir muito de seu comportamento real. Os resultados que temos de sua geometria foram obtidos pelo uso do pacote Gaussian92 instalado na Universidade de Nancy- Fr. A seguir calculamos seus espectros de absorção e de emissão e a de seus compostos com a glicina, a fenilalanina e a prolina por métodos semiempíricos. Estes cálculos são comparados com os resultados experimentais obtidos no grupo de Biofísica e Física Médica do IFUSP.

Resultados da Simulação

O resultado mais surpreendente que temos de nossas simulações desta molécula é a sua planaridade no estado excitado, em oposição ao estado fundamental que apresenta uma estrutura não planar e com duas conformações perfeitamente equivalentes de energia mínima. Em outras palavras, a molécula do ácido ortoaminobenzóico apresenta dois mínimos de energia com a mesma profundidade no poço de energia potencial. A figura 7.1 representa o ácido ortoaminobenzóico em sua conformação do estado fundamental gerado pelo pacote Gaussian92⁽¹⁶⁾. Nela podemos verificar que o grupo amina NH_2 possui uma estrutura piramidal, o que está de acordo com cálculos anteriores existentes na literatura. Vemos também que o grupo carboxila COO^- está

em um plano diferente do plano do anel benzênico e o primeiro plano pode girar em torno do carbono do mesmo ângulo de 7° no sentido oposto sem modificação em sua energia.

Quando calculamos os espectroeletrônicos com o programa Geomop, usando a opção ZINDO a partir da geometria obtida do programa Gaussian e comparamos com os resultados experimentais, verificamos que o espectro de absorção está bem próximo do espectro medido experimentalmente, enquanto que o espectro de emissão apresenta uma grande discrepância em relação ao valor medido¹.

	Absorção(nm)	Emissão(nm)
Simulação(Gaussian)	313	330
Experierias	310	396

O fato da geometria do estado fundamental não ser planar enquanto que a geometria do estado excitado é planar implica que a análise simplista não é adequada, porque não leva em consideração os níveis vibracionais da molécula (ver figura 7.5).

Na tabela seguinte apresentamos os resultados dos cálculos referentes às energias calculadas nos estados excitado (SCF) e relaxado (CIS) em Hartrees, nas várias geometrias, realizadas com o programa Gaussian:

estado\geometria	não-planar	fundamental-planar	excitada-planar
excitado(CIS)	-472,5712	-472,5754	-472,5844
relaxado(SCF)	-472,7869	-472,7841	-472,7740

O deslocamento de Stokes foi determinado diretamente com os resultados do Gaussian. Temos o deslocamento para o vermelho devido ao estado fundamental de 0,35 eV e 0,34 eV devido ao estado excitado, com um resultado total de 0,69 eV. Este resultado é comparável ao resultado experimental (0,86 eV). Não sabemos se o erro com as computações ZINDO advém da geometria errada do estado excitado ou do método que seria inadequado para geometrias não usuais.

Outros resultados

Temos os seguintes resultados para os espectros de absorção para os derivados do Abz em nanômetros

¹Utilizamos os métodos CIS3 21G* para obtermos a geometria do estado excitado e HF3 21G* para obtermos as geometrias dos estados fundamentais.

	orto-Abz	meta-Abz	para-Abz
experimental	320±10	320±10	320±10
computacional	324.4	326.6	322.7

	Abz-glicina	Abz-fenilalanina	Abz-prolina
experimental	312±10	310±10	290±10
computacional	310.9	299.8	289.3

Os resultados prévios em meta e para-Abz^(18,19,20) permitiu-nos assumir que o grupo NH₂ tivesse conformação piramidal. A partir desta informação calculamos as geometrias e os valores de transição óptica.

A geometria computada do Abz-prolina mostra uma concordância muito boa com os dados cristalográficos, qual seja, a sua estrutura cíclica. Siemion and Wieland⁽³³⁾ mostram que a estrutura desta molécula é totalmente rígida, com as ligações amida girando em relação ao plano do anel aromático: o grupo Ar-NH-CO- forma uma hélice horária e o grupo Ar-CO-N- uma hélice anti-horária, conforme figura 7.7. Na tabela a seguir, apresentamos os resultados mais detalhados do oabzprolina.

exp.(nm) ⁽³³⁾	$\epsilon^2(\text{M.cm})^{-1}$	comp.(nm)	força do osc.
290	4600	289,3	0,014
250	10100	253,7	0,123
230	20200	230,9	0,233

Os estudos de espectroscopia CD³ apresenta máximos em 290nm e 230nm, com fraca absorção óptica a 290nm. Colocamos o valor da força do oscilador para mostrar que o espectro de absorção calculado nesta região é muito reduzido. Em 250nm a força do oscilador duplica em relação ao espectro de 290nm. Já em relação a faixa de 230nm, as duas grandezas dobram os seus valores em relação à faixa de 250nm. Esta redução na intensidade da fluorescência pode ser devida à sua assimetria como tentam mostrar os autores acima citados⁴. Estes resultados mostram que o pacote Geomop esta bem calibrado para cálculos de geometria deste tipo de molécula.

²Coeficiente de extinção molar

³Dicroísmo circular

⁴Os resultados computacionais são muito melhores do que estes se considerarmos que a curva experimental apresenta uma largura que não levamos em consideração.

Em relação aos outros compostos não temos nenhum dado experimental para suas geometrias. Entretanto, a boa concordância entre os dados experimentais de espectroscopia fluorescente e os valores calculados permitem que tenhamos como critério que as geometrias correspondentes podem ser tomadas como representativas das estruturas das moléculas. A discrepância em relação à molécula de Abzfenilalanina sugere que a geometria desta molécula pode ter uma outra conformação distinta da que apresentamos neste trabalho.

Conclusões

Podemos concluir este capítulo afirmando que o programa semiempírico Geomop apresenta um excelente resultado quando queremos encontrar a melhor geometria de uma dada molécula, como comprovamos nos cálculos com o composto Abzprolina. Nos cálculos espectrais, quando realizados com cuidado, utilizando a opção ZINDO, os resultados também são muito bons, como comprovamos com o Abz e os compostos Abzglicina e Abzfenilalanina. Mostramos toda a potencialidade deste tipo de programa com os cálculos efetuados na Tiosemicarbazona.

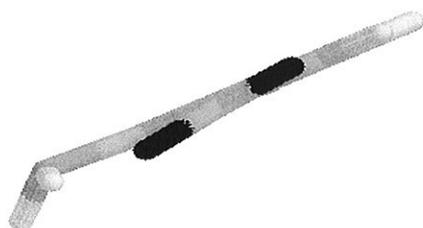


Figura 7.1: O ácido ortoaminobenzóico em seu estado fundamental - vemos que a sua estrutura não é planar , e o grupo carboxila forma um ângulo com o anel benzênico.

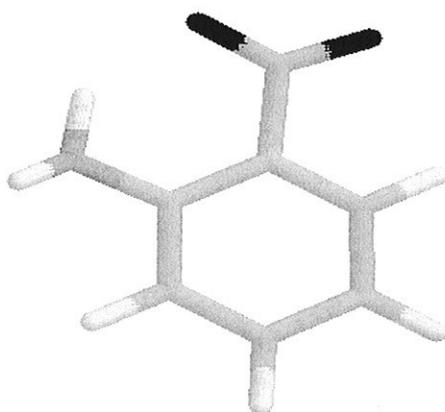


Figura 7.2: O ácido ortoaminobenzóico visto de outro ângulo em relação ao da Figura 7.1.

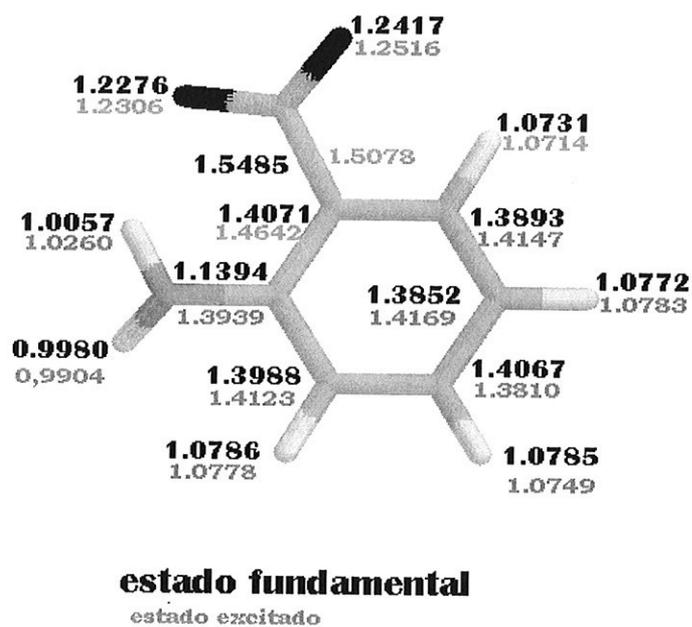


Figura 7.3: Geometrias dos estados fundamental e excitado - os números mais claros referem-se à geometria do estado excitado.

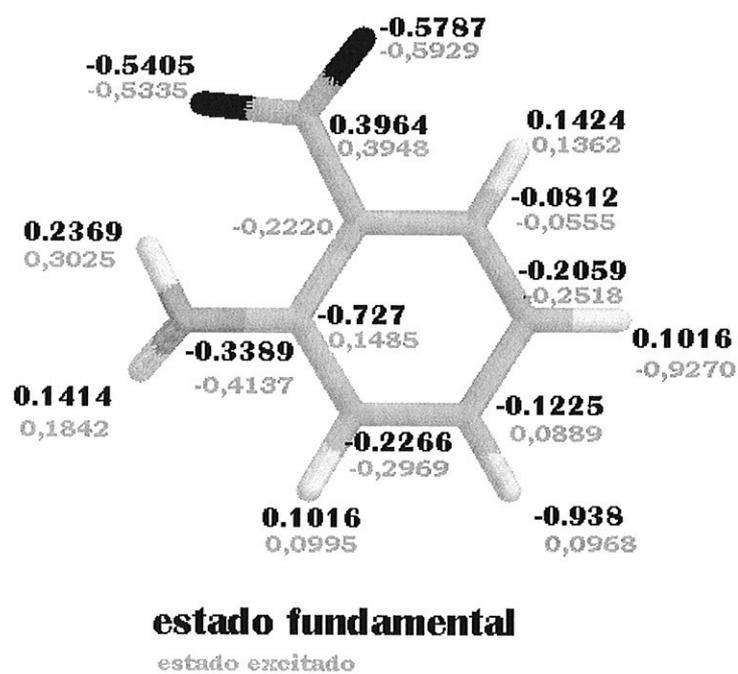


Figura 7.4: Cargas dos estados fundamental e excitado. Os números mais claros referem-se às cargas do estado excitado.

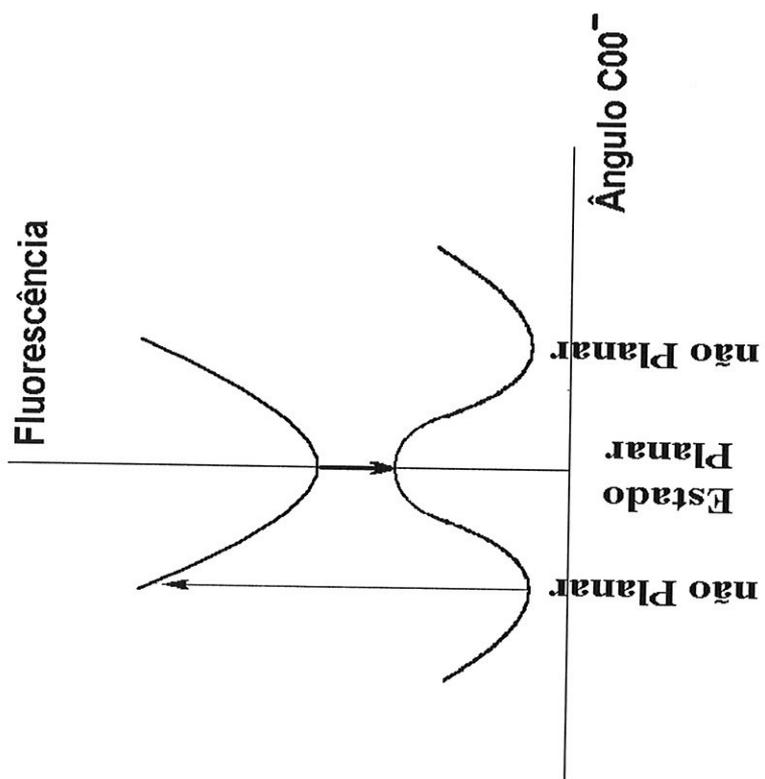


Figura 7.5: Representação esquemática da fluorescência do Abz. Os estados vibracionais não estão representados.

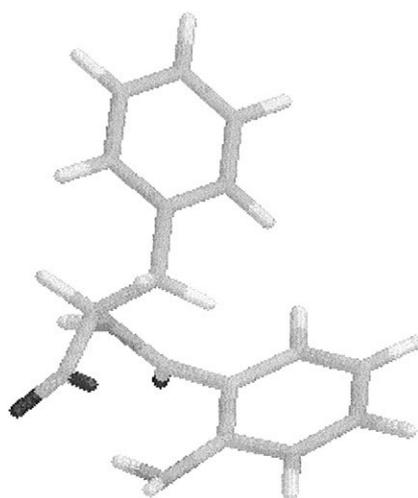


Figura 7.6: O ácido ortoaminobenzóico associado à fenilalanina em uma de suas estruturas de energia mínima.

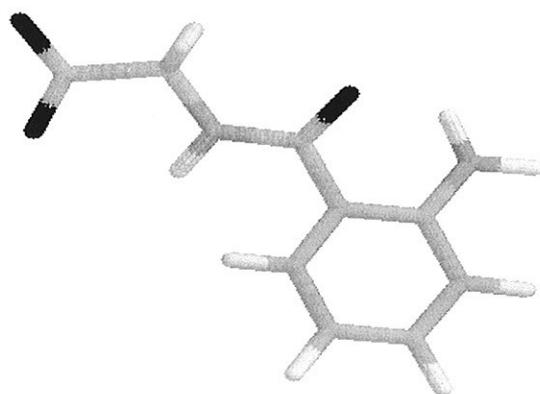


Figura 7.7: O ácido ortoaminobenzóico associado à glicina.

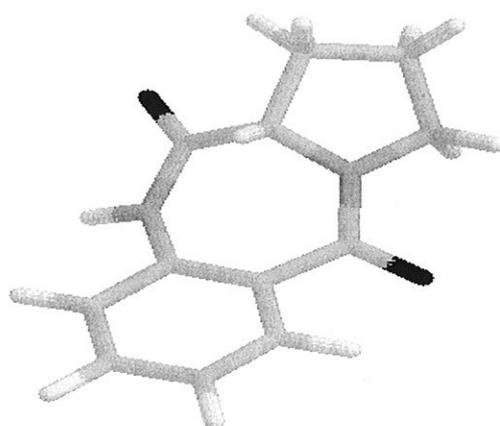


Figura 7.8: O ácido ortoaminobenzóico associado à prolina. Notamos que eles constituem uma estrutura com três ciclos fechados.

Capítulo 8

INICIAÇÃO

Todo ritual de iniciação tem seus estilos e técnicas, que se refletem nos “troles” até que o noviço assimile todas as regras e maneirismos. Das cavernas úmidas e mal-cheirosas às salas refrigeradas e assépticas, esta foi a trajetória do *homo sapiens* ao *homo computus*. Acredito que este trabalho tenha demonstrado que estou em condições de ser admitido no grupo dos noviços dos *homo computus*. Basta verificar que em todo o desenvolvimento do assunto uso o termo *inicialização*, e não, o termo correto *iniciação*.

Terminando assim, creio demonstrar que aqui não trato do *terminus* de um trabalho, mas ao contrário, indico o *principius* de uma nova fase, ainda que tardia.

Capítulo 9

Bibliografía

Bibliografia

- [1] Tese de Doutorado de Pascutti, P.G., apresentado no IFUSP em 1996
- [2] Gröningen Molecular Simulation System (GROMOS), Library Manual, Biomos b.v., Nijenborgh 14, 9747 AG. Gröningen, The Netherlands
- [3] Mezei, M., The finite difference thermodynamic integration, tested on calculating the hydration free energy - difference between acetone and dimethylamine in water, *J. Chem. Phys.*, **86** (2), 1987, 7084.
- [4] Kong, X. and Brooks III, C.L., λ -dynamics: A new approach to free energy calculations, *J. Chem. Phys.*, **105** (6), 1996, 2412.
- [5] Neria, E., Fischer, S., and Karplus, M., Simulation of activation free energies in molecular systems, *J. Chem. Phys.*, **105** (5), 1996.
- [6] Diadorim - programa escrito em linguagem Fortran por Luiz Carlos Gomide Freitas - Departamento de Química - UFSCar, 1992.
- [7] Fleischman, S.H., Brooks III, C.L., Thermodynamics of aqueous solvation: Solution properties of alcohols and alkanes, *J. Chem. Phys.*, **87** (5), 1987, 3029.
- [8] Discover 2.9.5. Manual Program - Chapter 6. Free Energy. Byosim Palo Alto - California
- [9] Tollenaere, J.P., Moereels, H., and Van Loon, M., On conformation analysis, molecular graphics and its derivatives, *J. Med. Chem.*, **29**, 1986, 92.
- [10] GEOMOP - programa escrito em linguagem Fortran por D. Rinaldi, P.E. Hoggan, A. Cartier, K. Baka, G. Monard, M. Loos, A. Mokrane, V. Dillet e V. They

- [11] Berendsen, H.J.C., Postma, J.P.M., van Gusteren, W.F. and Hermans, J., in *Intermolecular Forces: Proceedings of the 14th. Jerusalem Symposium on Quantum Chemistry and Biochemistry*, edited by B.Pullman, Reidl, Dordrecht, 1981, 331
- [12] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., *Numerical Recipes in Fortran*, second edition, Cambridge University Press, 1994.
- [13] Gordon, H.L., Jarrell, H.C., Szabo, A.G., Willis, K.J., and Somorjai, R.L., *Molecular Dynamics Simulations of the Conformational Dynamics of Tryptophan*, *J. Phys. Chem.*, **99** (4), 1992.
- [14] Ota, A.T., Temperini, M.L.A., Arêas, e.P.G., Loos, M., *Modeling of the interconversion between Z and E isomeric forms of pyridine-2-formyl thiosemicarbazone*, to be published in *Theochem*, 1997.
- [15] Frisch, M.J., Trucks, G.W., Head-Gordon, M., Gill, P.M.W., Wong, M.W., Foresman, B.G., Johnson, B.G., Shelegel, H.B., Robb, M.A., Replogle, E.S., Gomperts, R., Arnders, J.L., Raghavaachari, K., Binkley, J.S., Gonzalez, C., Martin, R.L., Fox, D.J., DeFrees, D.J., Baker, J., Stewart, J.J.P., and Pople, J.A., *GAUSSIAN 92* (Gaussain, Pittsburgh, 1992).
- [16] Souten, P.F.W., van Eijck, B.P., *Molecular dynamics of some small organic molecules: problems and results of free energy calculations for conformational transitions of ribose acid and tartaric acids*, *Mol. Simulation*, **4**, 1989,193.
- [17] Ota, A.L., Loos, M., Ito, A.S., *Propriedades eletrônicas do triptofano e do ácido aminobenzóico*, in *XIX Encontro Nacional de Física da Matéria Condensada*, Lindóia, 1996.
- [18] Böhm, S. and Kuthan, J., *An ab initio LCAO-MO Study of the Substituent Effectin Benzenoid Systems: Meta- and para-Substituted Benzoic Acids*, *Int.J. Q.Chem.*,**26**, 1984, 21.
- [19] Palmer, M.H., Moyes, W., Spiers, M., Ridyard, J.N.A., *The Electronic Structure of Substituted Benzene: Ab initio Calculations and Photoelectron Spectra for Nitrobenzene, the Nitrotoluenes, Dinitrobenzenes and Fluoronitrobenzenes*, *J. Mol. Structure*, **55**, 1979, 243.

- [20] Mezey, P.G., and Reynolds, W.F., Ab initio on 4-substituted benzoic acids; a further theoretical investigation into the nature of substituent effects in aromatic derivatives, *Can.J.Chem.*, **55**, 1977, 1567.
- [21] Temperini, M.L.A., dos Santos, M.R., Paoli Monteiro, V.R., *Spectrochimica Acta(A)*, **51**, 1995, 1517.
- [22] Hill, H.A.O., Page, D.J., and Walton, N.J., *J.Electroanal.Chem.*, **217**, 1987, 129.
- [23] Agrawal, K.C., Booth, b.A., DeNuzzo, S.N., and Sartorelli, a.C., *J.Med. Chem.*, **18**, 1975, 368
- [24] Protivinsky, R., *Antbiot.Cheomther.*, **17**, 1971,101.
- [25] Longan, J.C., Fox, M.P., Morgan, J.H., Mokohon, A.M., and Pfau, C.J., *J.Gen. Virology*, **28**, 1975, 271.
- [26] Malatesta, P., Accinelli, G.P., and Quaglia, G., *Ann.Chim.*, **49**, 1959, 397.
- [27] Easmo, J., Heinisch, G., Holzer, W., and rosenwirth, B., *J. Med. Chem.* **35**,1992, 3288.
- [28] Palenik, G.J., Rendle, D.F., and Carter, W.S., *Acta Crystallogr., Sect.B*, **30**, 1974, 2390.
- [29] Biyshkin, V.N., Chumakov, Yu.M., Samus', N.M., and Baka, L.D., *J.Struct.Chem.* **28**, 1987, 119.
- [30] French, F.A., Blantz Jr., E.J., Shanddix, S.C., and Broockman, R.W., *J.Med.Chem.*, **17**, 1974, 172.
- [31] Sathyanarayana, D.N., Volka, K., and Geetharani, K., *Spectrochimica Acta (A)*, **33**, 1977, 517.
- [32] Draeger, J.A. *Spectrochimica Acta (A)*, **39**, 1984, 809.
- [33] Siemion, I.Z., and Wieland, Th., The amide-aromatic-ring system - an inherently dissymmetric chromophore, *Tetrahedron*, **33**, 1977, 155.

- [34] Neidle, S.I., Webster, G.D., Jones, G.B., and Thurston, D.E., Structures of two DNA-minor-groove binders, based on pyrrolo [2,1-c][1,4]-benzodiazepine, *Acta Cryst.*, C **47**, 1991, 2678.
- [35] Siemion, I.Z., Wieland, Th., Pook, Karl-Heinz, Einfluß der Distanz des prolinacarbonyls vom β - und γ - Kohlenstoff auf die Lager der C-NMT-Signale, *Angew. Chem.*, **87**, 1975, 19.
- [36] Feigel, M., Lugert, G., Manero, J., Bremer, M., Conformation of Small Anthranilic-Acid-Proline-Peptides in Crystal, in Solution and a Semiempirical (AM1) Description of the Proline-Ramachandran Maps, *Z. Naturforsch.*, **45b**, 1990, 258.
- [37] Hehre, W.J., Taagepera, M., Taft, R.W., and Topsom, R.D., The relationship between Substituent-Induce Energy and Charge Effects in Proton-Transfer Equilibria, *J. Am.Chem. Soc.*, **103** (6), 1981, 1344.
- [38] Birmachu, W., Reed, J.K., The fluorescence properties of ortho aminobenzoate anesthetics in defined and phospholipid vesicles. *Photochem. an photobiology*, **47** (5), 1988, 675.
- [39] van Gunsteren, W.F., Berendsen, H.J.C., Computer Simulation of Molecular Dynamics: Methodology, Applications and Perspectives in Chemistry, *Angew. Chem. Int.Ed. Engl.*, **29**, 1990, 992.
- [40] Bagno, A., Comuzzi, C., and Scorrano, G., The site of Ionization fo Hydroxamic Acids Probed by Heteronuclear NMR Relaxation Rate and NOE Measurements. An Experimental and Theoretical Study., *J. Am. Chem. Soc.*, 1994, **116**, 916.
- [41] Koppel, I.A., Mishima, M., Sock, L.M., Taft, R.W., and Topsom, R. D., Acidities of 4-substituted benzoic, bicyclo [2.2.2]oct-1-yl and bicyclo[2.2.2]oct-2-enyl carboxylic acids, *J. Phys. Org. Chem.*, **6**, 1993, 685.
- [42] Haebelein, M., Murray, J.S., Brinck, T., Politzer, P., Calculated electrostatic potentials and local surface ionization energies of para-substituted anilines as measures of substituent effects, *Can. J. Chem.*, **70**, 1992, 2209.

- [43] Murray, J.S., Lane, P., Brinck, T., and Politzer, P., Relationships between computed molecular properties and solute-solvent interactions in supercritical solutions, *J. Phys. Chem.* **97**, 1993, 5144.
- [44] Colucci, W.J., Tistra, L., Sattler, M.C., Fronczek, F.R., and Barkley, M.D., Conformational studies of a constrained tryptophan derivative: implications for the fluorescence quenching mechanism., *J. Am. Chem. Soc.*, **112**, 1990, 9182.
- [45] Debuze, B., Dobson, C. M., and Teague, C.E., Conformational analysis of tryptophan in solution using nuclear magnetic resonance methods, *J. Chem. Soc., Perkin Trans.* **2**, 1981, 537.
- [46] Engh, R.A., Chen, L.X.Q., and Fleming, G.R., Conformational dynamics of tryptophan: a proposal for the origin of the non-exponential fluorescence decay, *Chem. Phys. Letters*, **126**(3,4), 1986, 365.
- [47] Goldman, C., Pascutti, P.G., Piquini, P., and Ito, A.S., On the contribution of electron transfer reactions to the quenching of tryptophan fluorescence, *J. Chem. Phys.*, **103**, (24), 1995, 10614.
- [48] Ito, A.S., Castrucci, A.M.L., Hruby, V.J., Hadley, M.E., Krajcarski, D.T., and Szabo, A.G., Structure-activity correlations of melanotropin peptides in model lipids by tryptophan fluorescence studies, *Biochemistry*, **32** (45), 1993, 12264.
- [49] Jäckle, H., and Luisi, P.L., Conformational study of histidine- and tryptophan - containing peptides in solution based upon the combination of nuclear magnetic resonance, circular dichroism, and fluorescence measurements, *J. Phys. Chem.*, **88**, 1984, 1767.
- [50] Mcknight, C.J., Rafalsk, M., and Gierasch, L.M., Fluorescence analysis of tryptophan-containing variants of LamB signal sequence upon insertion into a lipid bilayer, *Biochem*, **30**, 1992, 6421.
- [51] McMahon, L.P., Colucci, W.J., McLaughlin, M.L., and Barkley, Deuterium isotope effects in constrained tryptophan derivatives: implications for tryptophan photophysics, *J. Am. Chem. Soc.*, **114**, 1992, 8442.

- [52] Petrich, J.W., Chang, M.C., McDonald, D.B., and Fleming, G.R., On the origin of non-exponential fluorescence decay in tryptophan and its derivatives, *J. Am. Chem. Soc.*, **105**, 1983, 3824.
- [53] Ross, J.B.A., Wyssbroard, H.R., Poter, R.A., Schwartx, G.P., Michaels, C.A., and Laws, W.R., Correlations of tryptophan fluorescence intensity decay parameters with ^1H NMR - determined rotamers conformations: [tryptophan²]oxitocin, *Biochem*, **31**, 1992, 1585.
- [54] Swaminathan, R., Krishnamoorthy, G., and Periasamy, N., Similarity of fluorescence lifetime distributions for single tryptophan proteins in the random coil state, *Biophys.J.*, **67**, 1994, 2013.
- [55] Szabo, A.G., Rayner, D.M., Fluorescence decay of tryptophan conformers in aqueous solution, *J. Am.Chem.Soc*, **98**, 1980, 9270.
- [56] Slater, L.Sl, and Callis, P.R., Molecular orbital theory of the $^1L_\alpha$ and $^1L_\beta$ states of indole. An ab initio study, *J. Phys.Chem.*, **99**, 1995, 8572.
- [57] Callis, P.R., Vivian, J. T., Slater, L.S., Ab initio calculations of vibronic spectra for indole, *Chem. Phys. Letters*, **244**, 1995, 53.
- [58] Berendsen, H.J.C., Postma, J.P.M., van Gusteren, W.F., DiNola, A., and J. R. Haak, Molecular dynamics with coupling to an external bath, *J.Chem Phis.*, **81**(8), 1984, 3684.
- [59] Jørgensen, E. L., Ravimohan, C., Monte Carlo simulation of differences in free energies of hydration, *J.Chem. Phys.*, **83**(6), 1985, 3050.
- [60] Hummer, G., Szabo, A., Calculation of frre-energy difference from computer simulations of initial and final states, *J.Chem Phis*, **105**(5), 1996, 2004.
- [61] Hoover, W.G., and Ree, F.H. Use of computer experiments to locate the melting transition and calculate the entropy in the solid phase, *J. Chem. Phys.*, **47** (12), 1968, 4873.
- [62] Berens, P.H., Mackay, D.H.J., White, G.M., and Wilson, K.R., Thermodynamics and quantum corrections for molecular dynamics for liquid water, *J. Chem. Phys.*, **79**(5), 1983,2375.

- [63] Boresch, S., and Karplus, M., Jacobian factor in free energy simulations, *J. Chem. Phys.*, **105**(12), 1996, 5145.
- [64] Wallqvist, A., and Covell, D.G., On the origins of the hydrophobic effect: observations for simulations of n-Dodecane in model solvents, *Biophys. J.*, **71**, 1996, 600.
- [65] Jerala, R., Almeida, P.F.F., and Biltonen, R.L., Simulation of the gel-fluid transition in a membrane composed of lipids with two connected Acyl chains: application of a dimer-move step, *Biophys. J.*, **71**, 1996, 609.
- [66] RasMol v2.6- A Molecular Visualisation Program. Roger Sayle, Glaxo-Wellcome Research and Development, Greenfor, Middlesex, U.K.
- [67] Rootham, C.C.J., *Rev. Mod. Phys.*, **23**, 1951, 69.
- [68] Rootham, C.C.J., *Rev. Mod. Phys.*, **32**, 1960, 179.
- [69] Zwanzig, R.W., *J. Chem. Phys.*, **22**, 1954, 1420.

Referências Gerais

Methods in Enzymology, vol 202, Molecular Design and Modeling Concepts and Applications, ed. John J. Langone, Academic Press.

Allen, M.P., and Tildesley, D.J., Computer Simulation of Liquids, Oxford Science Publications, 1994

Engelstaff, P.A., an Introduction to the Liquid State, Oxford Science Publications, 1994.

Mc Weeny, Methods of Molecular Quantum Mechanics, Academic Press, 1992.

Rivail, Jean-Louis, Éléments de Chimie Quantique à l'usage des chimistes, InterÉditions/CNRS Éditions, 1994.

Endereços Eletrônicos

Aqui constam somente os endereços mais importantes sobre modelagem molecular. a partir deles é possível acessar todos os outros endereços eletrônicos.

<http://www.ch.ic.ac.uk/chemime/>

Esta página tem praticamente todos os endereços interessantes para este tipo de trabalho.

<http://cmm.info.nih.gov/modeling/>

Neste endereço foram obtidas as figuras do capítulo 3.

<http://gserv1.dl.ac.uk/CCP/main.html>

CCP significa The Collaborative Computational Projects e cada número CCP focaliza uma área específica. A estrutura eletrônica é o CCP1, o de Simulação das Fases Condensadas é o CCP5, e o de bioseqüência é o CCP11.