

COMUNICAÇÕES

1 RESUMO DE TESE

*Autor: Marcelo Albano Moret Simões Gonçalves
Prof. Auxiliar do Departamento de Ciências Exatas*

Curso: Mestrado em Física

Instituição: Universidade Federal da Bahia — UFBA

Área de Concentração: Física da Matéria Condensada

Subárea: Modelagem Biofísica

*Título da Dissertação: Modelagem Molecular Clássica usando
o Método Estocástico: Simulação por Termalização - GSA*

(Generalized Simulated Annealing)

Orientador: Prof. Dr. Kleber Carlos Mundim

Defesa: 1996

RESUMO — *Propomos, neste trabalho, uma estratégia estocástica para a determinação dos estados de equilíbrio, de sistemas moleculares semiclássicos, com base no método de Simulação por Termalização Generalizada (GSA). As geometrias conformacionais de equilíbrio são obtidas, fazendo um acoplamento entre um campo clássico de força (programa THOR) e o método estocástico GSA. A aplicabilidade do método foi testada em diferentes sistemas moleculares, cujo número de átomos, nas moléculas, variou significativamente. Os resultados obtidos, utilizando a termostática de Tsallis, mostram que o procedimento computacional torna-se mais rápido e eficiente do que aqueles baseados nas máquinas de Boltzmann e Cauchy.*

ABSTRACT— *We propose a stochastic optimization technique based on GSA - Generalized Simulated Annealing method to mapping minima points of a semi-classical energy surface. The conformational geometries of equilibrium are obtained making a coupling between a semi-classical force field (THOR-package) and GSA procedure. Its utility in molecular mechanics is here illustrated, applying it to different molecular systems. The results obtained, using the Tsallis thermostatics, show that data procedures are faster and more efficient than the results based upon Boltzmann and Cauchy's machines.*

1 INTRODUÇÃO

É bem conhecido que sistemas moleculares, em geral, podem existir em diferentes conformações moleculares estáveis. O número dessas organizações espaciais dos átomos, de um sistema molecular,

crece à medida que o número de graus de liberdade do sistema aumenta. Em particular, biomoléculas apresentam milhares de conformações moleculares estáveis, i.e., mínimos locais do sistema molecular. A maior dificuldade, do ponto de vista computacional, está na determinação do mínimo global do sistema molecular. Por esse motivo têm surgido diferentes métodos computacionais em física molecular.

Os modelos teóricos usuais aplicados tanto à física clássica quanto à quântica são baseados em métodos gradientes, os quais garantem a localização de apenas um dos vários mínimos que possam existir no sistema analisado. A estratégia usada para a obtenção do estado fundamental de sistemas moleculares tem sido o mapeamento pela força bruta.

De forma a minimizar esse tipo de dificuldade, é proposto, neste trabalho o acoplamento entre o código computacional THOR, o qual, nesse caso, serve para calcular a energia conformacional, com um algoritmo baseado em Simulação por Termalização, ou seja, o algoritmo GSA_ND. O procedimento usado, nesse acoplamento, consiste em comparar a energia conformacional semiclássica obtida a partir do código computacional THOR com as geometrias obtidas randomicamente pela rotina GSA_ND.

O acoplamento entre o código computacional THOR e o método de Simulação por Termalização se mostra adequado para a determinação das estruturas estáveis, i.e., os mínimos do sistema molecular analisado. Como este trabalho é baseado nesse acoplamento, será feita uma descrição do código computacional THOR e do algoritmo GSA_ND, para uma melhor compreensão acerca do conteúdo deste texto.

O código computacional THOR é baseado na utilização de um campo clássico de energia para investigar as propriedades físicas de sistemas moleculares. O estudo das estruturas é feito através da Dinâmica Molecular, a qual se mostra eficiente na análise da energia conformacional. O código tem a capacidade de fornecer todos os aspectos de vários sistemas moleculares de interesse, a partir da função energia conformacional, i.e., do campo clássico que determina as forças a cada instante da simulação da dinâmica. Esse código também é capaz de obter uma estrutura conformacional de equilíbrio, usando o método gradiente *Steepest Descent*, o qual busca a estrutura estável mais próxima, minimizando as forças entre os átomos do sistema molecular.

A proposta deste trabalho é obter as conformações estáveis (estruturas de mínimo) de sistemas moleculares, usando o método de Simulação por Termalização, ou melhor, o algoritmo GSA_ND e

não pela simulação da Dinâmica ou por métodos gradientes, como é usualmente feito em outros códigos computacionais de simulação para sistemas moleculares. O algoritmo GSA_ND é baseado em uma estratégia estocástica para a determinação dos estados de equilíbrio. Assim a localização do mínimo global, para qualquer função potencial, é feita usando o método Monte Carlo em um processo de resfriamento do sistema molecular em estudo. A aceitabilidade de uma estrutura, obtida randomicamente pelo algoritmo GSA_ND, durante a simulação, é dada pelo critério de Metrópolis. Esse procedimento permite mapear a hiper-superfície de potencial, além de obter o mínimo global do sistema analisado.

Na seção 2, é introduzida, de forma sucinta, a proposta de generalização do método de Simulação por Termalização, feita recentemente por Stariolo e Tsallis. O algoritmo GSA_ND é baseado nessa generalização. Entretanto o algoritmo GSA_ND difere do proposto por Stariolo e Tsallis, pois considera o processo de minimização como sendo feito pela função distribuição de visitação e não pela de *Lévy-Flight*. Esse algoritmo depende de três parâmetros que definem a probabilidade de aceitabilidade, dada por um critério de Metrópolis generalizado, a probabilidade de visitação, a qual fornece a perturbação nas coordenadas minimizadas e a temperatura inicial para o processo de resfriamento. Esses parâmetros são q_A , q_v e $T(q_v)$, respectivamente. É discutido, ainda, o processo para a obtenção da perturbação nas coordenadas, a qual é dada a partir da função distribuição de visitação. Essa perturbação foi calculada a partir da inversão da função distribuição de distribuição. Ainda nessa seção, é apresentado o método gradiente, onde se está dando ênfase ao método de minimização *Steepest Descent*. É interessante ressaltar que não é objetivo desse trabalho escolher um dos vários métodos de minimização como o “correto”, quando aplicado a sistemas moleculares, e sim utilizá-los em sistemas adequados. Portanto quando se deseja obter o mapeamento da hiper-superfície de potencial e/ou o mínimo global de um sistema molecular, deve-se assumir a minimização por Simulação por Termalização e caso o objetivo seja encontrar apenas uma conformação estável muitas vezes o método gradiente se mostra mais rápido computacionalmente.

A seção 3 é uma análise sucinta do código computacional THOR. Nessa seção são encontrados subsídios suficientes para a compreensão da arquitetura desse código. O código tem a capacidade de obter as estruturas moleculares de interesse, usando a Dinâmica Molecular, pode também minimizar um sistema molecular por métodos gradientes, além de ter a capacidade de simular a penetração de uma molécula em uma membrana. A estrutura do

código THOR foi pensada a mais simples possível, pois este código foi montado para ser manipulado por vários usuários distintos.

As conclusões acerca das aplicações do acoplamento entre o código THOR e o algoritmo GSA_ND são apresentadas na seção 4. Na tese, é mostrado que existem valores dos parâmetros do algoritmo GSA_ND que otimizam a simulação, ou seja, a obtenção do mínimo global de um sistema molecular é dada em um menor tempo de computação. A quantidade de coordenadas a serem minimizadas é um fator preponderante na escolha do parâmetro q_v , pois este se mostra ser dependente da dimensão do problema de minimização. Cada problema de minimização leva a valores diferentes do parâmetro q_v , pois esses sistemas têm diferentes graus de liberdade. Para problemas de minimização que tenham a mesma quantidade de coordenadas minimizadas, o comportamento do algoritmo é similar, pois o valor de q_v que otimiza a simulação, em um menor número de transições, é dado em um intervalo pré-determinado, como é analisado nessa seção. Entretanto as diferenças que ocorrem, no resultado, são devido às diferenças nos sistemas físicos analisados, ou seja, os graus de liberdade do sistema como todo e a forma da função energia conformacional, principalmente.

2 MÉTODO DE SIMULAÇÃO POR TERMALIZAÇÃO

Uma grande variedade de problemas (em física, química, estatística, engenharia, economia, psicologia etc.) tem como ponto central a minimização de uma apropriada função energia $E(x)$ (função custo), definida em um espaço contínuo D -dimensional $x \in \mathbf{R}^D$. Sobre este espaço D -dimensional são definidos os vetores x . Para o processo de minimização, otimização de problemas, existem vários algoritmos, os quais são adequados em alguns casos e, em outros, não. Caso a função energia seja convexa (um único mínimo) qualquer método gradiente facilmente soluciona o problema, entretanto se a função energia for não-convexa (de vários mínimos), a solução requer um método mais sofisticado. Isto ocorre porque os métodos gradientes convergem para o mínimo “mais próximo”, o qual, geralmente, é um mínimo local, pois em uma hiper-superfície de potencial onde se tem n mínimos a probabilidade de se convergir para o mínimo global é, pelo menos, $1/2^n$, pois todos os n mínimos, nesse caso, têm a mesma probabilidade de ocorrência. A sofisticação do método, para a obtenção do mínimo global, consiste na capacidade do algoritmo de passar por regiões da hiper-superfície

de potencial onde o valor da função energia $E(\mathbf{x})$ é maior do que o de onde o sistema se encontra.

Nos últimos anos, o chamado Método de Simulação por Termalização (de Recozimento) ou, como é expresso em inglês, *Simulated Annealing* (SA), tem descrito problemas de minimização global de forma precisa. Esses problemas de minimização global possuem uma função energia $E(\mathbf{x})$, definida em um espaço contínuo de "D" dimensões. Por esse método, encontra-se o mínimo (idealmente global) da função energia $E(\mathbf{x})$ em uma hiper-superfície de potencial, a qual possui vários mínimos locais. O aspecto básico de um algoritmo, baseado em Simulação por Termalização, consiste em uma otimização estocástica, onde uma ou mais temperaturas artificiais (ruídos) são introduzidas e gradativamente resfriadas, em completa analogia com a técnica de "recozimento" (*Annealing*), freqüentemente usada em metalurgia para tornar o metal fundido em uma estrutura cristalina (mínimo absoluto da energia termodinâmica). Essa temperatura artificial (ou conjunto de temperaturas) serve como fonte de estocasticidade, extremamente conveniente, pois existe a possibilidade do sistema estar na bacia atratora de um mínimo local da hiper-superfície de potencial. Espera-se que, perto do fim do processo, o sistema seja atraído para a bacia atratora do mínimo absoluto do sistema ou um dos mínimos absolutos, caso haja degenerescência. A temperatura é praticamente nula e o procedimento assintoticamente se torna equivalente ao método gradiente (p.e. *Steepest Descent*). O método de Simulação por Termalização possui o atrativo de poder otimizar problemas com um grande número de parâmetros minimizados, especialmente quando se deseja a obtenção do mínimo global (absoluto) do sistema.

Kirkpatrick et al., para sistemas clássicos, e Ceperley et al., para sistemas quânticos, mostraram que a probabilidade de se obter o mínimo absoluto do sistema, usando Simulação por Termalização, é igual a 1. Estes trabalharam no formalismo da estatística de Boltzmann-Gibbs do *quasi*-equilíbrio e usaram uma distribuição de visitação Gaussiana. Esse algoritmo é conhecido como máquina de Boltzmann ou, como é expresso em inglês, *Classical Simulated Annealing* (CSA). Geman e Geman mostraram que, para o caso clássico, a condição necessária e suficiente de se ter probabilidade igual a um, na obtenção do mínimo global (absoluto) do sistema é tal que a temperatura decresça logaritmicamente com o tempo. Szu e Hartley propuseram uma distribuição de visitação de Cauchy-Lorentz, no lugar de uma Gaussiana. Esse algoritmo é conhecido como máquina de Cauchy ou, como é expresso em inglês, *Fast Simulated Annealing* (FSA). Szu e Hartley mostraram que o resfriamento

pode ocorrer muito mais rapidamente, portanto, mais eficientemente, se a temperatura decrescer com o inverso do tempo. Tsallis et al. propuseram recentemente uma generalização da Simulação por Termalização, baseada na estatística de Tsallis, a qual será chamada de máquina de Tsallis ou, como é expresso em inglês, *Generalized Simulated Annealing* (GSA). A manipulação de dois parâmetros (q_A , q_V) pode levar o algoritmo para o escopo da CSA (1, 1), para FSA (1, 2) ou para valores onde o resfriamento seja otimizado.

Os métodos CSA, FSA e GSA foram demonstrados no ensemble canônico, de Boltzmann-Gibbs para os dois primeiros e da estatística de Tsallis no terceiro. De acordo com a probabilidade de aceitabilidade de cada algoritmo, os saltos/perturbações do vetor x_t , serão sempre aceitos se o novo vetor x_{t+1} levar o sistema para uma função energia $E(x_{t+1}) < E(x_t)$ e poderão ser aceitos, os saltos onde $E(x_{t+1}) > E(x_t)$. Os saltos devido à distribuição de visitação Gaussiana são sempre locais, os devido a distribuição de visitação de Cauchy-Lorentz são frequentemente locais, contudo podem ser gerados saltos muito longos, saltos estes que ocorrem com uma frequência bem menor que os locais. De fato, esta é como uma distribuição de Lévy-Flight. O já estabelecido método de simulação Metrópolis, na termodinâmica do equilíbrio, difere dos métodos Simulação por Termalização, pois a temperatura, neste, é mantida constante. A generalização dos métodos de Simulação por Termalização contém as máquinas de Boltzmann e Cauchy como casos particulares.

O ponto central da estatística de Tsallis está na proposta de generalização da entropia, inspirada na teoria de multifractais. Seguindo essa proposta de generalização da estatística, podem-se obter as expressões corretas para a probabilidade de aceitabilidade e para a distribuição de visitação deste formalismo generalizado.

Agora será resumido o procedimento para se obter o mínimo global de uma dada função energia (função custo) $E(\mathbf{x})$:

(i)

Fixam-se os parâmetros q_A e q_V . Começa-se o procedimento ($t = 1$) com um valor arbitrário x_1 , um valor alto, o bastante da temperatura $T(q_V)_1$ e calcula-se $E(x_1)$. O valor da temperatura deve ser, pelo menos, da ordem da energia mínima de $E(\mathbf{x})$;

(ii)

Gera-se aleatoriamente o valor Dx_t para se obter x_{t+1} a partir da expressão da distribuição de visitação generalizada:

$$g(Dx_t, q_V) = \text{CONST} [T(q_V)_t]^{1/(3 - q_V)} \{1 - (q_V - 1)[Dx_t / (T(q_V)_t)]^{1/(3 - q_V)}\}^{2 \cdot \frac{1}{1/(q_V - 1) - 1}}$$

onde $CONST = [(q_v - 1)/p]^{1/2} \{ [1/(q_v - 1)] / [(3 - q_v)/(q_v - 1)] \}$
(iii)

Calcula-se $E(\mathbf{x}_{t+1})$:

se $E(\mathbf{x}_{t+1}) < E(\mathbf{x}_t)$, troca-se $\mathbf{x}_t$ por $\mathbf{x}_{t+1}$,
se $E(\mathbf{x}_{t+1}) > E(\mathbf{x}_t)$, gera-se um número aleatório $rand \in [0,1]$.
Caso $rand > P(q_A)$, onde
 $P(q_A) = [1 + (q_A - 1)(E(\mathbf{x}_{t+1}) - E(\mathbf{x}_t))/T(q_v)_t]^{-[1/(q_A - 1)]}$
é a probabilidade de aceitabilidade, mantém-se $\mathbf{x}_t$, caso contrário, troca-se $\mathbf{x}_t$ por $\mathbf{x}_{t+1}$.

(iv)

Calcula-se uma nova temperatura $T(q_v)_t$:

$$T(q_v)_t = T(q_v)_t [(2^{q_v-1} - 1)/t^{q_v-1}]$$

e volta para o passo (ii) até que o mínimo da função energia $E(\mathbf{x})$ seja encontrado, com a precisão desejada.

É interessante notar que, ao se utilizar Simulação por Termalização (GSA), para o processo de minimização de uma função custo qualquer (energia), não é necessário calcular o gradiente desta (força), também não é necessário o cálculo de nenhuma matriz, como a Hessiana utilizada nos Métodos Gradientes. Por esses motivos, é possível se economizar tempo de CPU, quando os valores dos parâmetros do algoritmo GSA_ND forem otimizados.

Entretanto não se está querendo obter comparações entre os métodos de minimização comentados acima, pois cada um deles tem a sua aplicabilidade específica. Os Métodos Gradientes são utilizados para a obtenção do mínimo mais próximo, enquanto Simulação por Termalização mapeia a hiper-superfície de potencial, além do bônus de ser possível obter o mínimo absoluto do sistema analisado.

3 INTRODUÇÃO AO CÓDIGO COMPUTACIONAL THOR

A utilização da função empírica energia conformacional para a investigação de propriedades físicas de moléculas obteve um avanço significativo nas últimas décadas. Estudo das estruturas, intervalos de energias e frequências vibracionais de moléculas através do uso das simulações Monte Carlo e da Dinâmica Molecular em soluções e/ou líquidos puros, para analisar energias conformacionais, mostram-se eficientes. Nesta seção será descrita a arquitetura do código computacional THOR, a qual pode eficientemente fornecer todos os aspectos de vários sistemas de interesse, em especial biológicos e farmacológicos, a partir da função empírica energia conformacional.

A partir do código computacional THOR, é possível: — Modelar a construção de uma estrutura de interesse; — Minimizar a energia dessa estrutura, usando o método gradiente (o método usado é o Stepest Descent) ou usando Simulação por Termalização, o qual em Inglês é expresso como *Generalized Simulated Annealing*; - Simular a dinâmica molecular; — Penetrar a molécula em estudo em uma membrana; — Estudar o conjunto de fosfolipídios, formando uma monocamada; — Obter as constantes de força e as constantes da geometria, usando simulação por Termalização; além do bônus dado pela análise estrutural das propriedades de equilíbrio e da dinâmica molecular, determinadas por este cálculo, obtido numericamente. A atual versão do programa é capaz de analisar tanto o caso de moléculas isoladas, quanto um conjunto complexo de moléculas, por exemplo a simulação de uma membrana.

No desenvolvimento do programa, quatro aspectos foram priorizados para que a sua arquitetura permita um sistema mais eficiente possível: — O programa obtém um conjunto de soluções com o menor tempo possível; — Os dados de entrada/saída são simplificados. Basicamente, é necessário apenas o conjunto das coordenadas de cada átomo e suas ligações, i.e., uma matriz Z; — O programa é facilmente compreendido por qualquer usuário, portanto, torna-se simples para modificá-lo (quando necessário/desejado); — Este tem alternativas para se trabalhar com qualquer sistema operacional (UNIX, DOS entre outras plataformas).

A parte de visualização das conformações moleculares e da simulação da dinâmica molecular é realizada a partir de rotinas gráficas próprias. As interações atômicas no código computacional THOR são simuladas por um campo clássico, parametrizado a partir de dados experimentais e cálculos quânticos. A energia conformacional de uma dada molécula é expressa classicamente como a soma das contribuições dos potenciais de "curto alcance" (quimicamente ligados) e de "longo alcance" (não ligados).

A unidade fundamental usada no código computacional THOR é o átomo. Nessa aproximação, um átomo é considerado ser um ponto massivo e, portanto, sem propriedades direcionais e graus de liberdade internos. As massas dos átomos que constituem a molécula em estudo são expressas em unidade atômica. Cada átomo real do sistema pode corresponder a um único átomo no código computacional THOR, contudo para muitos sistemas, vários (ou todos) hidrogênios são combinados ao átomo vizinho, no qual estão ligados. Esta combinação de átomos é conhecida, na literatura, como "representação do átomo estendido". No código computacional THOR somente os hidrogênios ligados ao oxigênio e ao nitrogênio

são considerados explicitamente, enquanto que a maioria dos grupos CH_2 e CH_3 são assumidos como sendo uma unidade atômica.

Por limitação computacional, a representação de átomo estendido era a única forma de se simular moléculas, há alguns anos. A partir da metade da década passada, com as melhorias na eficiência computacional, tornou-se possível escolher quais hidrogênios deveriam ser considerados e quais não. Os hidrogênios pouco carregados (como os alifáticos, citados acima e os aromáticos) podem ser postos na forma de átomos estendidos, sem perdas, pois estes pouco influenciam nas pontes de hidrogênios. Com isso as maiores desvantagens do átomo estendido são removidas. No código computacional THOR, pode-se escolher se serão listados todos os hidrogênios, todos os hidrogênios que participam efetivamente do potencial da molécula, como os hidrogênios de uma estrutura hélice —, ou somente átomos estendidos.

4 CONCLUSÕES E PERSPECTIVAS

Durante a confecção do algoritmo GSA_ND, optou-se por propor um único vetor N-dimensional como sendo o vetor de minimização. Após alguns testes, notou-se que este vetor, o qual aparentemente é isotrópico, privilegia a perturbação nas primeiras coordenadas em detrimento das últimas, pois as perturbações, nas últimas coordenadas, são multiplicadas por um número maior de cossenos, os quais têm seus valores compreendidos no intervalo [0 ; 1]. Por exemplo, enquanto a primeira coordenada é multiplicada por um seno, a segunda é por dois cossenos, a terceira, por três e assim sucessivamente até a k -ésima coordenada que é multiplicada por k valores compreendidos no intervalo [0 ; 1]. Essa sucessão de multiplicações de valores compreendidos entre [0 ; 1] leva a uma diminuição sistemática na perturbação. Portanto há um achatamento dessas coordenadas em relação às primeiras. Além desse achatamento, existe o problema associado ao tipo de coordenada que se está minimizando, ou seja, para ângulos um acréscimo de 10° pode ser adequado, entretanto um acréscimo de 10 Å em uma coordenada radial é inconcebível, para sistemas moleculares. Por esse motivo, escolheu-se minimizar as coordenadas uma a uma, independentemente, como se o problema fosse o acoplamento de vários problemas unidimensionais. Esta escolha fornece o resultado da simulação em um tempo de CPU menor, além de fornecer a convergência de forma melhor.

Agora será analisado o comportamento dos parâmetros do

algoritmo GSA_ND, quando aplicados a sistemas moleculares. Esses parâmetros irão definir o tempo de CPU para que a convergência será alcançada. Os próximos parágrafos são dedicados aos parâmetros $T(qv)_0$, q_A e q_v , respectivamente.

A evolução do número de ciclos com a temperatura mostra ter um comportamento semelhante, para problemas de minimização de sistemas moleculares de um mesmo tipo. A forma da curva dessa evolução, para a molécula ico-alanina, mostra-se semelhante para um, dois ou oito pares de ângulos minimizados, quando se varia a quantidade de coordenadas minimizadas. No caso da molécula penta-alanina, esse comportamento também é observado. E outros sistemas, como os alifáticos, também mostraram esse comportamento, sendo que esse comportamento semelhante é observado para dois sistemas físicos distintos, por exemplo, entre um propano e um butano. Este comportamento sugere que o processo de minimização é dependente tanto da função energia que se está minimizando quanto da forma desta função no espaço.

Com relação ao parâmetro q_A , que fornece a probabilidade de aceitabilidade por um critério de Metrópolis generalizado, por simplicidade, escolheu-se um valor onde os sistemas analisados têm um comportamento razoável. Este valor não deve ser o melhor, mas também está longe de ser o pior. Penna fez uma análise acerca desse parâmetro, para o problema combinatório do caixeiro viajante, o qual é expresso em inglês por *Traveling Salesman*, e conclui que valores negativos fornecem um comportamento otimizado, i.e., a convergência é obtida em um menor número de transições. É interessante que nessa comunicação é dito que o critério de Metrópolis “foi generalizado nela”, caso o autor dessa comunicação tivesse um conhecimento mais aprofundado, descobriria que a generalização é feita a Stariolo e Tsallis, pois o espaço em que esta generalização é feita é o mesmo em ambos os casos, i.e., no caso do problema de minimização baseado na generalização da termodinâmica o espaço das probabilidades é o mesmo do problema combinatório, ou seja, o problema do cacheiro viajante. A única novidade dessa comunicação é a forma elegante que o autor define o critério de Metrópolis generalizado.

Com relação ao parâmetro q_v , que determina a forma da função distribuição de visitação, foi observado que os seus valores irão determinar a rapidez com que a convergência será alcançada. Portanto existem valores de q_v que são ruins, para os quais a convergência será dada com um número de ciclos infinito e outros que otimizam a simulação. Este parâmetro é a “alma/sentimento” do método/algoritmo, pois a escolha correta desse leva a obtenção

do estado fundamental do sistema analisado em um número de ciclos extremamente pequeno em relação às infinitas possibilidades de conformações distintas que este sistema pode assumir. O intervalo de convergência proposta por Tsallis et al. foi obtido sem que houvesse discrepâncias, ou seja, o algoritmo é compatível com os agiríamos baseados em Simulação por Termalização já existentes. Como já foi observado, esses outros algoritmos, conhecidos, utilizam a função *Lévy-Flight* como função distribuição de visitação, portanto são aproximações do GSA_ND.

Foi observado também que a minimização de sistemas moleculares com muitos graus de liberdade ainda apresenta dificuldades e a convergência é obtida após um número de ciclos muito alto. Para ser feita a minimização de um sistema molecular desse tipo, como uma proteína, será necessária a utilização de supercomputadores mais potentes, pois o número de ciclos irá aumentar bastante, o que limita a utilização de *WorkStations*, tipo a usada nesse trabalho. Um outro fator que está associado à dimensão do problema de minimização é que, se esta for muito grande, o intervalo de convergência será muito estreito e próximo de $q_v = 1$. Com isso as perturbações nas coordenadas minimizadas serão dadas de forma local, com um comportamento próximo ao da máquina de Boltzmann.

Foi encontrada, pela simulação, a estrutura em hélice — no peptídeo ico-alanina. Este resultado da tese é um resultado inédito que a comunidade científica vem perseguindo (desde a década de 50!). A partir desse resultado, houve a motivação de se lançar uma nova teoria para dinâmica molecular. Essa nova teoria para se fazer dinâmica molecular, i. e., *Stochastic Molecular Dynamics*, a qual está sendo submetida à apreciação da comunidade acadêmica pela referência Moret et al., tem a capacidade de obter várias (ou melhor, todas) as características físicas de sistemas biológicos.

Um outro trabalho, inédito, que será desenvolvido é o mapeamento das dupla-hélice-a em DNA. Para a confecção desse trabalho, será necessária a utilização de um supercomputador, como o CRAY da UFRJ.

Deve-se observar que o acoplamento entre o código computacional THOR e o algoritmo GSA_ND, o qual foi chamado de THOR_GSA é apenas uma nova versão do código computacional THOR, na qual é possível minimizar sistemas por Simulação por Termalização. Portanto o leitor deve ter em mente que o THOR, agora, pode obter o mapeamento da hiper-superfície de potencial e a conformação do estado fundamental, de qualquer sistema molecular, usando Simulação por Termalização, i.e., usando a rotina GSA_ND, desenvolvida nesse trabalho.

BIBLIOGRAFIA BÁSICA

- ARÊAS, E.P.G., et al. Molecular Dynamics Simulation Sequence at a Membrane/Water Interface, *J. Phys. Chem.*, n.99, p.14882-14885, 1995.
- BROOKS, B. R., et al. The Karplus, The CHARMM approach, M., *J. Comp. Chem.* n.4, p.187 - 199, 1983.
- CEPERLEY, D., ALDER, B. Classical Simulated Annealing, *Science* N. 231, p.555-559, 1986.
- GEAR, C. W. Numerical Initial Value Problems in Ordinary Differential Equations. Prentice-Hall, Englewood Cliffs, NJ, 1971.
- KIRKPATRICK, S., GELATT, C. D., VECCHI, M. P. Optimization by Simulated Annealing. *Science*, n.220, p.671-690, 1983.
- METROPOLIS, N. et al. Equation of State calculations by fast computing Machines, *J. Chem. Phys.* n.21 p.1087-1092, 1953
- MORET, M. A., et al. Stochastic Molecular Dynamics Using GSA, Preprint, *J. Comp. Chem.*, 1996.
- MORSE, P., FESHBACH, H. Methods of Theoretical Physics, International Student Edition, NY: McGraw-Hill, 1953.
- MUNDIM, K. C., PASCUTTI, P. G., BISCH, P. M., Developed the package THOR-Force Field and Molecular Dynamics Simulations.
- MUNDIM, K. C., TSALLIS, C., Geometry Optimization and Conformational Analysis Through Generalized Simulated Annealing, *Int J. Quant. Chem.*, v. 58 n.4, p.373, 1996.
- PENNA, T. J. P., Travelling salesman problem and Tsallis statistics, *Phys. Rev. E.*, v. 51 n.1, R1, 1995.
- PRESS, W. H., et. al. Numerical Recipes in Fortran. Cambridge University Press, Cambridge, 1992 .
- STARIOLO, D. A., TSALLIS, C., Optimization by Simulated Annealing: Recent Progress, World Scientific Publishing, 1995.
- STRYER, L., BIOCHEMISTRY, W. H. Freeman and CO., NY, 1988.
- SZU, H., HARTLEY, R., Fast Simulated Annealing, *Phys. Lett. A* 122, 157 1987.
- TSALLIS, C., SOUZA, A. M., MAYNARD, R., Lévy-Flight and Related Topics in Physics, Heidelberg: Springer-Verlag, 1995.