



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

FACULTAD DE CIENCIAS

"ISOMEROS DE LA ADENINA"

T E S I S

QUE PARA OBTENER EL TITULO DE:
F I S I C O
P R E S E N T A :

VEGA MUNGUA JOSE



DIRECTOR DE TESIS: DR. KARO MICHAELIAN PAUW

MEXICO, D. F.

279302

2000



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

División de Estudios Profesionales
Facultad de Ciencias, UNAM.

MAT. MARGARITA ELVIRA CHÁVEZ CANO
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias.
PRESENTE

Comunicamos a usted que hemos revisado el trabajo de tesis que lleva el título:

"Isómeros de la alenina"

realizado por: Vega Manguía José

con número de cuenta 3924511-1 , pasante de la carrera de Física

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Dr. Karo Michaelian Pauw
Propietario

K. Michaelian

Propietario Dr. Rubén Santamaría Ortiz

Rubén Santamaría Ortiz

Propietario Dr. Germinal Cocho Gil

G. Cocho

Suplente M.C. Augusto Cabrera Manuel

Augusto Cabrera Manuel

Suplente Dr. Rubén Barrera Pérez

Rubén Barrera Pérez

Consejo Departamental de Física



Margarita Elvira Chávez Cano

FACULTAD DE CIENCIAS
DEPARTAMENTO DE FÍSICA

Agradecimientos

Al gran dador.

“Cuando se es muy joven y se sabe un poco, las montañas son montañas, el agua es agua, y los árboles son árboles. Cuando se ha estudiado y se es leído, las montañas ya no son montañas, el agua ya no es agua, y los árboles ya no son árboles. Cuando se es sabio, nuevamente las montañas son montañas, el agua es agua y los árboles son árboles.”.
Antiguo refrán del Budismo Zen

A mi padres: Sofía Munguía Bravo y Juvenal Vega Martell.

A mis hermanos: Elio, Lila y Nadia.

A la familia Vega Martell, en particular a mis tíos: Lila, Virginia, Hugo y Octavio.

Al Ing. Antonio Cerna González y familia.

A mis amigos, especialmente a Manuel V., compañero de tantas jornadas.

A mi asesor, Dr. Karo Michaelian Pauw.

A la UNAM, máxima casa de estudios.

Introducción

Capítulo 1. Análisis vibracional.

1.1 Introducción.

1.2 Análisis vibracional.

1.3 Solución de la aproximación armónica.

1.4 Coordenadas normales.

1.5 Coordenadas internas.

1.6 La ecuación secular.

1.7 Simplificación de la ecuación secular.

1.8 Consistencia del potencial con valores de la energía.

1.9 Optimización de los parámetros de la energía.

Capítulo 2. El potencial utilizado.

2.1 Desarrollo del potencial.

2.2 El potencial semiempírico.

2.3 tipos de átomos.

2.4 Parámetros de átomos no ligados.

2.5 Obtención de los parámetros de θ y R .

2.6 Parámetros de torsión.

2.7 Obtención de los valores para las cargas q_i .

2.8 R , ϕ y θ como función de las coordenadas.

Capítulo 3. Algoritmos genéticos.

- 3.1 Métodos de optimización.
- 3.2 Algoritmos genéticos.
- 3.3 Creación de la población inicial.
- 3.4 Mecanismos de selección.
- 3.5 Reproducción: (1) entrecruzamiento.
- 3.6 Reproducción: (2) mutación.
- 3.7 Similaridades importantes.
- 3.8 Los *schema* y sus propiedades.
- 3.9 Efecto de la selección.
- 3.10 Efecto del entrecruzamiento.
- 3.11 Efecto de las mutaciones.
- 3.12 Representación Gray.

Capítulo 4. Resultados y conclusiones.

- 4.1 La búsqueda global.
- 4.2 Relajación con cálculos de primeros principios.
- 4.3 Resultados
- 4.4 Referencias.

1 Capítulo 1

1.1 Introducción

Uno de los problemas fundamentales en el estudio de las propiedades estructurales y energéticas de las moléculas, ha sido el de desarrollar una función analítica $V(\mathbf{R})$ la cual pueda describir correctamente la energía del sistema en términos de sus $3n-6$ grados de libertad, ($3n-5$ para el caso de moléculas lineales), donde n es el número de átomos en el sistema (-6 grados por la invariancia de la molécula bajo rotación y traslación). Aunque comunmente se asume que el uso de funciones de energía empíricas y semiempíricas en el análisis conformacional es "clásico", en contradicción a cálculos de mecánica cuántica, en realidad la diferencia básica es que el último es un método deductivo, que busca el predecir fenómenos observables a partir de una base fundamental (la ecuación de Schrödinger y la conservación de la energía), mientras que el primero es un método inductivo, que busca una representación analítica común a un gran conjunto de fenómenos observables. De hecho, no hay nada "clásico" en estas funciones, ya que ellas no son deducidas de la física clásica. Ellas pueden ser justamente representadas como una representación empírica de la aproximación de Born-Oppenheimer, según la cual el estado base de las moléculas es una función continua de las coordenadas atómicas [1]. En el área de proteínas y ácidos nucleicos, el método de optimización por mínimos cuadrados de los parámetros de la energía debido a Lifson y Warshel [2] fué pionero en la obtención de funciones analíticas $V(\mathbf{R})$.

1.2 Análisis vibracional

Para escribir la ecuación de movimiento interno de nuestra molécula, necesitaríamos tener la expresión para la energía cinética T y para la energía potencial V . Empleando como coordenadas los $3n$ desplazamientos de las posiciones de equilibrio de los n átomos: Δx_i , Δy_i y Δz_i , la energía cinética

toma la siguiente forma:

$$2T = \sum_{i=1}^n m_i (\Delta \dot{x}_i^2 + \Delta \dot{y}_i^2 + \Delta \dot{z}_i^2) \quad (1.1)$$

donde m_i es la masa del i -ésimo átomo. Con el objeto de simplificar la ecuación (1.1), podemos indicar por r_α la coordenada de desplazamiento cartesiana genérica respecto a la del equilibrio y reescribir la ecuación (1.1) en la forma más compacta

$$2T = \sum_{\alpha=1}^{3n} m_\alpha \dot{r}_\alpha^2 \quad (1.2)$$

en la cual la suma va de 1 a $3n$ y $m_\alpha = m_1$ para $\alpha = 1, 2$ y 3 ; $m_\alpha = m_2$ para $\alpha = 4, 5$ y 6 y así sucesivamente. En vez de usar las coordenadas de desplazamiento cartesianas r_α , es a menudo conveniente usar un conjunto de coordenadas cartesianas con pesos de masa q_α definidas por la relación:

$$q_\alpha = \sqrt{m_\alpha} r_\alpha \quad (1.3)$$

en términos de la cual la energía cinética toma la más simple forma

$$2T = \sum_{\alpha=1}^{3n} \dot{q}_\alpha^2 \quad (1.4)$$

Para pequeños desplazamientos, considerando la expansión de Taylor de la función de energía potencial intermolecular $V(\mathbf{r})$ de las coordenadas atómicas \mathbf{r} , alrededor de las coordenadas de equilibrio \mathbf{r}_0 ,

$$V(\mathbf{r}) = V(\mathbf{r}_0) + \sum_{\alpha=1}^{3n} \left(\frac{\partial V}{\partial r_\alpha} \right)_{\mathbf{r}_0} r_\alpha + \left(\frac{1}{2!} \right) \sum_{\alpha, \beta=1}^{3n} \left(\frac{\partial^2 V}{\partial r_\alpha \partial r_\beta} \right)_{\mathbf{r}_0} r_\alpha r_\beta + \left(\frac{1}{3!} \right) \sum_{\alpha, \beta, \gamma=1}^{3n} \left(\frac{\partial^3 V}{\partial r_\alpha \partial r_\beta \partial r_\gamma} \right)_{\mathbf{r}_0} r_\alpha r_\beta r_\gamma + \dots \quad (1.5)$$

El primer término representa la energía de conformación de la molécula; ya que no estamos interesados en el valor absoluto de la energía potencial, si no únicamente en su variación con las coordenadas r_i , podemos cambiar el cero de la energía tal que la energía de la configuración de equilibrio V_0 sea

cero. La desaparición del segundo miembro, esto es, la solución del conjunto de ecuaciones

$$\left(\frac{\partial V}{\partial r_\alpha}\right)_{\mathbf{r}_0} = 0 \quad (1.6)$$

representa la condición de que \mathbf{r}_0 es una conformación de equilibrio. Para pequeñas excursiones de los átomos alrededor de sus posiciones de equilibrio, los términos más altos de la ecuación (1.5) pueden ser despreciados y ésta se convierte en:

$$2V = \sum_{\alpha, \beta=1}^{3n} \left(\frac{\partial^2 V}{\partial r_\alpha \partial r_\beta}\right)_{\mathbf{r}_0} r_\alpha r_\beta = \sum_{\alpha, \beta=1}^{3n} f_{\alpha\beta} r_\alpha r_\beta \quad (1.7)$$

1.3 Solución de la aproximación armónica

Un conjunto de ecuaciones lineales para $\delta \mathbf{r}$ se obtiene al expandir el gradiente $\nabla V(\mathbf{r})$ en una serie de Taylor en un punto cerca del equilibrio, como $\nabla V(\mathbf{r})$ es una función de $\mathfrak{R}^{3n} \Rightarrow \mathfrak{R}^{3n}$, debemos utilizar el siguiente teorema:

Teorema: Si $\mathbf{f} = (f_1, f_2, \dots, f_m)$ de $\mathfrak{R}^n \Rightarrow \mathfrak{R}^m$ es diferenciable en el punto \mathbf{r} , entonces:

$\mathbf{f}(\mathbf{r} + \delta \mathbf{r}) = \mathbf{f}(\mathbf{r}) + D\mathbf{f}(\mathbf{r})\delta \mathbf{r}$, donde $D\mathbf{f}(\mathbf{r})$ es la matriz jacobiana de la función \mathbf{f} de \mathfrak{R}^n en \mathfrak{R}^m y viene dada por:

$$D\mathbf{f}(\mathbf{r}) = \begin{pmatrix} \frac{\partial f_1(\mathbf{r})}{\partial r_1} & \frac{\partial f_1(\mathbf{r})}{\partial r_2} & \dots & \frac{\partial f_1(\mathbf{r})}{\partial r_n} \\ \frac{\partial f_2(\mathbf{r})}{\partial r_1} & \frac{\partial f_2(\mathbf{r})}{\partial r_2} & \dots & \frac{\partial f_2(\mathbf{r})}{\partial r_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{r})}{\partial r_1} & \frac{\partial f_m(\mathbf{r})}{\partial r_2} & \dots & \frac{\partial f_m(\mathbf{r})}{\partial r_n} \end{pmatrix} \quad (1.8)$$

Sustituyendo \mathbf{f} por $\nabla V(\mathbf{r})$

$$\nabla V(\mathbf{r} + \delta \mathbf{r}) = \nabla V(\mathbf{r}) + f_{\mathbf{x}}(\mathbf{r})\delta \mathbf{r} \quad (1.9)$$

donde $f_{\mathbf{x}}$ es la matriz de segundas derivadas

$$f_{\alpha\beta} = \frac{\partial^2 V}{\partial r_\alpha \partial r_\beta} \quad \alpha, \beta = 1, 2, \dots, 3n \quad (1.10)$$

si hacemos $\mathbf{r}_0 = \mathbf{r} + \delta \mathbf{r}$ entonces $\nabla V(\mathbf{r} + \delta \mathbf{r}) = \nabla V(\mathbf{r}_0) = \mathbf{0}$ y por lo tanto:

$$\nabla V(\mathbf{r}) + f_{\mathbf{x}}(\mathbf{r})\delta \mathbf{r} = \mathbf{0} \quad (1.11)$$

resolviendo la ecuación (1.11) para δr tenemos:

$$\delta r = -fx^{-1}\nabla V(r) \quad (1.12)$$

La conformación de equilibrio es entonces:

$$r_0 = r + \delta r = r - fx^{-1}\nabla V(r) \quad (1.13)$$

En esta ecuación, los términos $fx_{\alpha\beta}$ son constantes, las cuales son normalmente llamadas "constantes de fuerza", ya que ellas representan los factores de proporcionalidad entre los desplazamientos de los átomos y las fuerzas restauradoras actuando sobre ellos (en la aproximación armónico simple). Debe observarse de la ecuación (1.10) que hay en principio $3n * (3n + 1)/2$ constantes de fuerza ya que, como V es una función continua, el orden de diferenciación no importa y así:

$$fx_{\alpha\beta} = fx_{\beta\alpha} \quad \alpha, \beta = 1, 2, \dots, 3n \quad (1.14)$$

Una vez que hemos obtenido T y V , podemos escribir las ecuaciones de movimiento para vibraciones moleculares. Como T es función solamente de las velocidades y V es función solamente de los desplazamientos, la ecuación de movimiento en forma lagrangiana puede ser escrita:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{r}_i} \right) + \left(\frac{\partial V}{\partial r_i} \right) = 0 \quad i = 1, 2, \dots, 3n \quad (1.15)$$

usando (1.2) y (1.7), queda

$$m_i \ddot{r}_i + \sum_{j=1}^{3n} fx_{ij} r_j = 0 \quad (1.16)$$

Este es un conjunto de $3n$ ecuaciones diferenciables de segundo orden simultáneas, cuyas soluciones son del tipo oscilador armónico de mecánica clásica

$$r_i = A_i \cos(2\pi\nu t + \phi) \quad (1.17)$$

donde ν es la frecuencia de vibración en cm^{-1} , ϕ es un factor de fase y A_i es el desplazamiento máximo del átomo, sustituyendo (1.17) en (1.16) se obtiene

$$\sum_{j=1}^{3n} fx_{ij} A_j - (2\pi\nu)^2 m_i A_i = 0 \quad (1.18)$$

la cual puede ser reescrita como

$$\sum_{j=1}^{3n} (f_{x_{ij}} - \lambda m_j \delta_{ij}) A_j = 0 \quad (1.19)$$

donde $\lambda = 4\pi^2\nu^2c^2$ y δ_{ij} es la delta de Kronecker. En forma matricial la ecuación (1.19) es:

$$\begin{pmatrix} f_{x_{11}} - \lambda m_1 & f_{x_{1,2}} & \dots & f_{x_{1,3n}} \\ f_{x_{2,1}} & f_{x_{2,2}} - \lambda m_2 & \dots & f_{x_{2,3n}} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_{3n,1}} & f_{x_{3n,2}} & \dots & f_{x_{3n,3n}} - \lambda m_{3n} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{3n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (1.20)$$

Para que las A_j sean diferentes de cero para $j = 1, 2, \dots, 3n$ se requiere que el determinante

$$\begin{vmatrix} f_{x_{1,1}} - \lambda m_1 & f_{x_{1,2}} & \dots & f_{x_{1,3n}} \\ f_{x_{2,1}} & f_{x_{2,2}} - \lambda m_2 & \dots & f_{x_{2,3n}} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_{3n,1}} & f_{x_{3n,2}} & \dots & f_{x_{3n,3n}} - \lambda m_{3n} \end{vmatrix} = 0 \quad (1.21)$$

Este es el determinante secular, cuando se expande da un polinomio de orden $3n$, implicando $3n$ valores de λ ; esto es, $3n$ frecuencias ν_k . En realidad, 6 de las λ 's son cero. Cuando una raíz de λ , digamos λ_k , es sustituida en la ecuación (1.18), se podrán encontrar $3n$ soluciones para A_i denotadas más correctamente como A_{ik}

1.4 Coordenadas normales

Las expresiones para T y V pueden ser reescritas en una forma compacta y simple, usando la notación matricial. Si llamamos \mathbf{r} un vector columna, con componentes dadas por las $3n$ coordenadas de desplazamiento r_α de los átomos, la ecuación (1.4) se transforma en

$$2T = \dot{\mathbf{r}}^T M \dot{\mathbf{r}} = \dot{\mathbf{q}}^T \dot{\mathbf{q}} \quad (1.22)$$

donde M es una matriz diagonal de las masas atómicas y $\mathbf{q} = M^{\frac{1}{2}} \mathbf{r}$. La ecuación (1.7) se convierte en

$$2V = \mathbf{r}^T \mathbf{f} \mathbf{r} = \mathbf{q}^T \mathbf{f} \mathbf{q} \quad (1.23)$$

donde $f = M^{-\frac{1}{2}} f_x M^{-\frac{1}{2}}$, la T indica el vector transpuesto y f_x es una matriz simétrica cuadrada de tamaño $3n \times 3n$ cuyos elementos son las $f_{x_{\alpha\beta}}$. La ecuación (1.22) es una suma de los cuadrados de cada una de las coordenadas de las velocidades, la ecuación (1.23) es más complicada y contiene todos los posibles términos de productos cruzados de las coordenadas, ya que en general la matriz f tiene elementos distintos de cero fuera de la diagonal. Este hecho no solo complica el tratamiento clásico si no que hace imposible la solución del problema cuántico en términos de las coordenadas r_i , ya que se sabe como resolver la ecuación de Schrödinger para el oscilador armónico solo si se pueden escribir las energías cinética y potencial como una suma de términos separados, siendo cada término función de una simple coordenada.

Estas consideraciones muestran que las coordenadas cartesianas, aunque convenientes para visualizar el problema, son, desde el punto de vista de encontrar una solución analítica, una mala elección. Es, por lo tanto, legítimo el preguntarnos si podemos encontrar un conjunto mejor de coordenadas, el cual tenga la deseada propiedad de hacer las matrices de las energías potencial y cinética diagonales. A este nuevo conjunto de coordenadas se les llama "coordenadas normales" y definen un nuevo vector columna Q con componentes dadas por las $3n$ coordenadas normales Q_α .

Para obtener las energías potencial y cinética en forma diagonal, la transformación de las coordenadas cartesianas con factor de masa a coordenadas normales debe ser una transformación lineal del tipo

$$Q_\alpha = \sum_{\beta=1}^{3n} l'_{\alpha\beta} q_\beta \quad (1.24)$$

lo cual significa que cada coordenada normal Q_α es una combinación lineal de las $3n$ coordenadas cartesianas con factor de masa con los coeficientes $l'_{\alpha\beta}$ a ser determinados. En notación matricial tal transformación se escribe como

$$Q = \mathcal{L}^{-1} q \quad (1.25)$$

donde por conveniencia se ha usado el símbolo \mathcal{L}^{-1} para indicar la matriz de transformación de coordenadas cartesianas con factor de masa a coordenadas

normales, cuyos elementos son los coeficientes $l'_{\alpha\beta}$ de la ecuación (1.24). La transformación inversa de coordenadas normales a coordenadas cartesianas con factor de masa es por consiguiente

$$\mathbf{q} = \mathcal{L}\mathbf{Q} \quad (1.26)$$

donde \mathcal{L} es la inversa de la matriz \mathcal{L}^{-1} , esto es, $\mathcal{L}\mathcal{L}^{-1} = I$, I siendo la matriz unidad. Como discutimos antes, se requiere que Q_α lleve las energías cinéticas y potencial simultáneamente a forma diagonal, esto es, a expresiones del tipo

$$2T = \dot{\mathbf{Q}}^T \dot{\mathbf{Q}} \quad (1.27)$$

y

$$2V = \mathbf{Q}^T \Lambda \mathbf{Q} \quad (28)$$

donde Λ es una matriz diagonal cuyos elementos son los parámetros de frecuencias normales

$$\lambda_k = 4\pi^2 c^2 \nu_k^2. \quad (1.29)$$

Las ecuaciones (1.27) y (1.28) imponen severas condiciones sobre la matriz de transformación L , las cuales encontraremos por la introducción de la ecuación (1.26) en las correspondientes expresiones de T y V , (1.22) y (1.23), Considerando primero la energía cinética, tenemos

$$2T = \dot{\mathbf{Q}}^T \mathcal{L}^T \mathcal{L} \dot{\mathbf{Q}} \quad (1.30)$$

comparándola con (1.27)

$$\mathcal{L}^T \mathcal{L} = I \quad (1.31)$$

la cual muestra la primera propiedad interesante de la matriz \mathcal{L} , es decir

$$\mathcal{L}^T = \mathcal{L}^{-1} \quad (1.32)$$

del mismo modo obtenemos para V :

$$2V = \mathbf{Q}^T \mathcal{L}^T \Lambda \mathcal{L} \mathbf{Q} \quad (1.33)$$

la cual, comparándola con la ecuación (1.28)

$$\mathcal{L}^T \Lambda \mathcal{L} = \Lambda \quad (1.34)$$

la cual da la segunda propiedad importante de \mathcal{L} . Como \mathcal{L} es ortogonal, la ecuación anterior puede ser reescrita en la forma

$$f\mathcal{L} = \mathcal{L}\Lambda \quad (1.35)$$

la cual es una ecuación de eigenvalores. Formalmente podemos considerar a la matriz f como un operador actuando sobre una columna dada \mathcal{L}_k de la matriz \mathcal{L} para dar como resultado la misma columna \mathcal{L}_k multiplicada por un factor constante λ_k , el cual es el correspondiente k -ésimo elemento diagonal de la matriz Λ . Para cada columna \mathcal{L}_k de \mathcal{L} tenemos

$$f\mathcal{L}_k = \mathcal{L}_k\lambda_k \quad (1.36)$$

la cual puede ser reescrita en la forma

$$(f - I\lambda_k)\mathcal{L}_k = (M^{-1/2}f_xM^{-1/2} - I\lambda_k)\mathcal{L}_k = 0 \quad k = 1, 2, \dots, 3n \quad (1.37)$$

o

$$|f - I\lambda_k| = 0 \quad (1.38)$$

vemos que la ecuación (1.38) sería la análoga a la (1.19), pero debemos observar que cada amplitud desconocida A_{jk} no es numericamente igual a \mathcal{L}_{jk} , de hecho, de la ecuación (1.31) se sigue que para la k -ésima columna de \mathcal{L} : $\sum_j \mathcal{L}_{jk}^2 = 1$, lo cual no es cierto para los elementos A_{jk} correspondientes. Como se dijo anteriormente, podemos determinar las relaciones entre las $3n$ A_{jk} , y podemos construir un conjunto de amplitudes normalizadas de el tipo

$$\frac{A_{jk}}{\sqrt{\sum_{j=1}^{3n} A_{jk}^2}}, \quad (1.39)$$

las cuales son numericamente equivalentes a los elementos \mathcal{L}_{jk} . Estas consideraciones nos muestran más precisamente el significado de las columnas de la matriz \mathcal{L} ; esto es, los renglones de la matriz \mathcal{L}^T . Como $\mathbf{Q} = \mathcal{L}^{-1}\mathbf{q} = \mathcal{L}^T\mathbf{q}$, vemos que para una coordenada normal dada Q_k , el correspondiente renglón de \mathcal{L}^T da el conjunto de coeficientes con los que los desplazamientos cartesianos con factor de masa toman parte para esa coordenada normal.

1.5 Coordenadas internas

Vimos que un nuevo conjunto de coordenadas llamado "coordenadas normales" está asociado con una vibración normal de la molécula, estas coordenadas nos permiten escribir de una forma muy simple, sin términos cruzados, las energías potencial y cinética. Desafortunadamente, las coordenadas normales son desconocidas hasta que la ecuación secular ha sido resuelta y no son usadas en el problema de calcular las frecuencias vibracionales. Estas son coordenadas solo en el sentido que describen el movimiento molecular; pero, por el momento, las coordenadas cartesianas son el único marco de trabajo que tenemos a nuestra disposición. Así, el método según el cual dado un conjunto de constantes de fuerzas es posible obtener el conjunto de frecuencias vibracionales es realmente teórico en la mayoría de los casos. En la práctica, las frecuencias vibracionales pueden ser determinadas experimentalmente del análisis de espectros infrarrojos y Raman; mientras que las constantes de fuerza, esto es, los elementos de la matriz f_x , son desconocidos, es aquí donde surge nuestro interés de que, a partir de las frecuencias de vibración, podamos calcular las constantes de fuerza.

Para esto, debemos de trabajar en un sistema de coordenadas que sea más característico de la estructura molecular que las coordenadas cartesianas de los átomos. Los cambios en la longitud de enlace y ángulos entre enlaces químicos son las más apropiadas coordenadas físicas para la descripción de la energía potencial. Este tipo de coordenadas son conocidas como coordenadas internas, ya que ellas describen justamente el movimiento de la molécula. En el análisis vibracional, los tipos de coordenadas internas que son usualmente usados son:

(1) coordenada de alargamiento de enlace: variación en la longitud de un enlace químico.

(2) coordenada de doblamiento de ángulo de enlace: variación de el ángulo entre dos enlaces químicos teniendo un átomo en común.

(3) coordenada de torsión: variación en el ángulo entre los planos determinados por tres enlaces consecutivos conectando 4 átomos.

Sea s_k una coordenada interna genérica. La relación más general entre s_k y las coordenadas cartesianas puede ser escrita como

$$s_k = \sum_i B_i^k r_i + \frac{1}{2} \sum_{ij} B_{ij}^k r_i r_j + \dots \quad (1.40)$$

donde los coeficientes B_i^k, B_{ij}^k están determinados por la geometría molecular. Si nos limitamos a pequeñas amplitudes de vibración, los términos más altos que el primero pueden ser despreciados y la ecuación (1.40) puede ser escrita como una relación lineal, cuya forma matricial sería

$$s = Br \quad (1.41)$$

donde B es una matriz cuyos elementos son los coeficientes B_i^k de la ecuación (1.40). La matriz B no es en general una matriz cuadrada. Si consideramos una molécula triatómica con $3 * 3 - 6 = 3$ grados de libertad dados por las longitudes de enlace l_1 y l_2 y el ángulo α entre ellos, entonces la matriz B tendrá 3 renglones y 9 columnas (ya que hay 9 coordenadas cartesianas para 3 átomos). Como B no es una matriz cuadrada, no puede ser invertida. Nosotros podemos incluir 6 coordenadas correspondientes a la traslación y rotación en el vector s, de este modo B es cuadrada y matriz no singular y por lo tanto puede ser invertida, conduciendo así a la transformación de coordenadas internas a cartesianas

$$r = B^{-1}s \quad (1.42)$$

1.6 La Ecuación Secular

Las expresiones de la energías cinética y potencial en coordenadas internas son fácilmente obtenidas

$$2T = \dot{q}^T \dot{q} = \dot{r}^T M \dot{r} = \dot{s}^T (B^{-1})^T M B^{-1} \dot{s} = \dot{s}^T G^{-1} \dot{s} \quad (1.43)$$

y también

$$2V = q^T f q = r^T M^{\frac{1}{2}} f M^{\frac{1}{2}} r = r^T f r = s^T (B^{-1})^T f B^{-1} s = s^T F s \quad (1.44)$$

donde $f_x = M^{\frac{1}{2}} f M^{\frac{1}{2}}$ es la matriz de constantes de fuerza en coordenadas cartesianas.

$$F = (B^{-1})^T f_x B^{-1} \quad (1.45)$$

es la matriz de constantes de fuerza en coordenadas internas y

$$G^{-1} = (B^{-1})^T M B^{-1} \quad (1.46)$$

es la matriz de energía cinética en coordenadas internas, la cual ya no es diagonal como en coordenadas cartesianas. Entonces

$$s = Br = BM^{-\frac{1}{2}} q = BM^{-\frac{1}{2}} L Q = L Q \quad (1.47)$$

donde

$$L = BM^{-\frac{1}{2}} L \quad (1.48)$$

Siguiendo el mismo procedimiento de la sección anterior y utilizando $2V = Q^T \Lambda Q$ y $2T = Q^T Q$ tenemos

$$2V = Q^T L^T F L Q \quad (1.49)$$

y

$$2T = \dot{Q}^T L^T G^{-1} L \dot{Q} \quad (1.50)$$

vemos que

$$L^T F L = \Lambda \quad y \quad L^T G^{-1} L = I \quad (1.51)$$

de la ecuación (1.51) obtenemos

$$G^{-1} = (L^T)^{-1} L^{-1} \quad (1.52)$$

y por inversión

$$G = L L^T \quad (1.53)$$

multiplicando la izquierda de la primera ecuación en (1.51) por L tenemos

$$L L^T F L = L \Lambda \quad (1.54)$$

de la cual se sigue, utilizando la (1.53) que

$$G F L = L \Lambda \quad (1.55)$$

la cual es la forma más general de la ecuación secular en la aproximación armónica para el problema vibracional. La solución de la ecuación (1.55) es el problema central en espectroscopía vibracional.

1.7 Simplificación de la ecuación secular a la aproximación armónica

A continuación mostramos la simplificación de la ecuación (1.55) a la forma en que se utilizó en el proceso de ajuste de mínimos cuadrados para los parámetros del potencial:

$$\text{GFL} = [\text{BM}^{-1}\text{B}^T][(\text{B}^T)^{-1}\text{f}_x\text{B}^{-1}][\text{BM}^{-1/2}\mathcal{L}] = \text{L}\Lambda = [\text{BM}^{-1/2}\mathcal{L}]\Lambda \quad (1.56)$$

Del lado izquierdo, expresando a $\text{M}^{-1} = \text{M}^{-1/2}\text{M}^{-1/2}$ y eliminado $\text{BM}^{-1/2}$ de ambos lados queda:

$$\begin{aligned} \text{M}^{-1/2}\text{f}_x\text{M}^{-1/2}\mathcal{L} &= \mathcal{L}\Lambda \\ \text{f}\mathcal{L} &= \mathcal{L}\Lambda \end{aligned} \quad (1.57)$$

Esta ecuación puede considerarse como una ecuación de eigenvalores, al actuar sobre la columna \mathcal{L}_k de la matriz \mathcal{L} , toma la forma:

$$\text{M}^{-1/2}\text{f}_x\text{M}^{-1/2}\mathcal{L}_k = \text{f}\mathcal{L}_k = \mathcal{L}_k\lambda_k \quad (1.58)$$

tomando la traspuesta de ambos lados en (1.57) y tomando en cuenta que $\text{f}^T = \text{f}$ y que las matrices M , y Λ son diagonales queda

$$\mathcal{L}^T\text{f} = \Lambda\mathcal{L}^T \quad (1.59)$$

de (1.32), como $\mathcal{L}^{-1} = \mathcal{L}^T$, entonces

$$\text{f} = \mathcal{L}\Lambda\mathcal{L}^T \quad (1.60)$$

Multiplicando ambos lados por el vector \mathbf{q} y recordado de (1.26) que $\mathcal{L}^T\mathbf{q} = \mathbf{Q}$ obtenemos:

$$\text{f}\mathbf{q} = \mathcal{L}\Lambda\mathbf{Q} \quad (1.61)$$

Ahora, fijémonos en el k -ésimo elemento de esta ecuación vectorial:

$$\sum_j f_{kj} q_k = \lambda_k \sum_j \mathcal{L}_{kj} Q_k \quad (1.62)$$

Pero como $\mathbf{q} = \mathcal{L}\mathbf{Q}$, implica $\sum_j \mathcal{L}_{kj} Q_k = q_k$ Por lo que se obtiene finalmente

$$\sum_j f_{kj} q_k = \lambda_k q_k \quad (1.63)$$

o en forma de ecuación de eigenvalores:

$$\mathbf{f}\mathbf{q} = \lambda\mathbf{q} \quad (1.64)$$

La cual es la aproximación armónica que se presentó en las ecuaciones (1.19)-(1.21).

1.8 Consistencia del potencial con valores experimentales de la energía

La contraparte experimental del exceso de energía de conformación es el exceso de entalpía en el estado gaseoso, como se ha derivado de medidas calorimétricas. El exceso calculado de energía de conformación de una molécula cicloalcana con n átomos de carbono, sobre la energía de conformación del ciclohexano en forma de silla se define como

$$\Delta V^{(n)} = V(\mathbf{r}_0^{(n)}) - \left(\frac{n}{6}\right) V(\mathbf{r}_0^{(6)}) \quad (1.65)$$

donde $\mathbf{r}_0^{(n)}$ es la conformación de equilibrio de un anillo con n átomos de carbono. Para comparar $\Delta V^{(n)}$ con los valores experimentales del exceso de energía,

$$\Delta H^{(n)} = H^{(n)} - \left(\frac{n}{6}\right) H^{(6)} \quad (1.66)$$

es necesario el agregarle a $\Delta V^{(n)}$ los correspondientes excesos de energía de vibración, rotación y traslación: $\Delta H_{vrt}^{(n)}$. Las energías vibracionales no están disponibles experimentalmente para la mayoría de los cicloalcanos debido a problemas en las asignaciones de frecuencias fundamentales por la falta de conocimiento de las constantes de fuerza apropiadas. Entonces, si las frecuencias vibracionales y la energías de conformación son derivadas teóricamente

del mismo conjunto de funciones, la energía vibracional puede ser combinada con la energía potencial de conformación para dar una predicción consistente del exceso de entalpía. Esto es llevado a cabo aquí al insertar las frecuencias calculadas ν_α en la fórmula de mecánica estadística

$$H_{vrt}^{(n)} = kT \sum_{\alpha=1}^{3n-6} \left[\frac{1}{2} + (e^{x_\alpha} - 1)^{-1} \right] + 3kT \quad (1.67)$$

donde $x_\alpha = h\nu_\alpha/kT$, $\Delta H_{vrt}^{(n)} = H_{vrt}^{(n)} - \frac{n}{6} H_{vrt}^{(6)}$ es entonces sumando a ΔV^n
Para dar

$$\Delta H^{(n)} = \Delta V^{(n)} + \Delta H_{vrt}^{(n)} \quad (1.68)$$

1.9 Optimización de los parámetros del potencial

Para alcanzar un mejor ajuste entre los resultados experimentales y los calculados teóricamente, se usó un método de mínimos cuadrados que a continuación se describe para ajustar los parámetros. Estos parámetros de las funciones de energía fueron variados iterativamente para que gradualmente se reduzca la suma de los cuadrados de las diferencias entre las cantidades experimentales y las calculadas, hasta que esta alcance un mínimo. El conjunto de funciones con los parámetros optimizados se le conoce como campo de fuerza consistente obtenido por un análisis vibracional y conformacional. Mientras más se incremente el ajuste del cálculo teórico al experimental, mayor será el entendimiento de la naturaleza de la interacciones moleculares. La descripción del procedimiento de optimización es la siguiente:

Sea $\mathbf{x} = \{x_m\}$ un vector cuyos componentes son los parámetros de las diversas funciones de energía que componen la energía conformacional total. Sea $\delta\mathbf{y} = \mathbf{y}_{calc} - \mathbf{y}_{exptl}$ un vector cuyos componentes son las diferencias entre los valores calculados y los observados de los excesos de entalpía, conformaciones de equilibrio y frecuencias vibracionales de todas las moléculas bajo consideración; esto es, $\Delta\mathbf{y} = \Delta\Delta H^{(n)}, \Delta\Delta s_0^{(n)}, \Delta\Delta\nu_k^{(n)}$, donde el índice (n) denota las diversas moléculas empleadas (alquenos). Las conformaciones de equilibrio son representadas aquí por coordenadas internas s_0 , para asegurar una correspondencia uno a uno entre las coordenadas experimentales y

las calculadas. Sin embargo, todos los cálculos de las diversas propiedades moleculares fueron desarrollados en coordenadas cartesianas por conveniencia y para ahorrar tiempo de CPU. Los resultados finales fueron transformados entonces a coordenadas internas. En general Δy es una función no lineal de x , cuya expansión de Taylor en un punto x es:

$$\Delta y(x + \delta x) = \Delta y(x) + Z\delta x \quad (1.69)$$

(donde Z es una matriz cuyos elementos son $Z_{im} = \partial y_i / \partial x_m$, su cálculo se muestra en el apéndice) establece una relación lineal entre $\Delta y [= \delta y(x + \delta x)]$ y pequeños incrementos δx . El ajuste entre los cálculos y los experimentos es mejorado si δx es tal que reduce el valor de la expresión $(\Delta y^T P)(P \Delta y)$, donde P es una matriz diagonal de factores de peso. Este mejoramiento es obtenido cuando $(\Delta y^T P)(P \Delta y)$ es minimizado, sujeto a la condición de que δx es suficientemente pequeño de tal forma que los términos más altos en la expansión de Taylor son despreciables. La condición es impuesta al requerir que $\delta x^T \delta x$ sea una pequeña constante. Así δx tiene que ser solución del conjunto de ecuaciones:

$$\left(\frac{\partial}{\partial \delta x_m} \right) (\Delta y^T P^2 \Delta y + \xi \delta x^T \delta x) = 0 \quad (1.70)$$

donde ξ es un multiplicador de Lagrange. Las ecuaciones (1.69) y (1.70) dan

$$Z^T P^2 (\Delta y(x) + Z\delta x) + \xi \delta x = 0 \quad (1.71)$$

entonces

$$\delta x = -(Z^T P^2 Z + \xi I)^{-1} Z^T P^2 \Delta y(x) \quad (1.72)$$

da el requerido mejoramiento lineal de x , con tal de que ξ sea suficientemente grande, correspondiendo a valores suficientemente pequeños de $\delta x^T \delta x$. Este procedimiento es iterativo, hasta que la secuencia de valores $x^{(k+1)} = x^{(k)} + \delta x^{(k)}$ converga a los valores óptimos de los parámetros de la energía.

2 Capítulo 2

2.1 Desarrollo del potencial

Durante los últimos 15 años, un gran esfuerzo se ha desarrollado para tratar de obtener formas analíticas para las funciones de energía empíricas las cuales puedan describir correctamente las estructuras, energías y frecuencias vibracionales de sistemas macromoleculares biológicos complejos. La mayor parte del esfuerzo ha sido aplicado a proteínas; es más recientemente que las funciones potenciales han sido desarrolladas para ácidos nucleicos y han sido empleadas en simulaciones dinámicas.

El procedimiento computacional descrito en el capítulo anterior ya ha sido aplicado a diversos conjuntos de funciones de energía. Cada conjunto tenía sus parámetros de energía optimizados para dar un mejor acuerdo con un cuerpo cada vez mayor de datos experimentales. Estos últimos incluyeron frecuencias vibracionales, conformaciones de equilibrio, energías de esfuerzos de conformación, y entalpías (energías) de vibración-traslación-rotación. Este método fué inicialmente realizado por A. Warshel y S. Lifson en un conjunto de funciones de energía de un campo de fuerza para alcanos. [1]

Varias funciones de energía fueron examinadas y las mejores fueron elegidas, para así mejorar sistemáticamente la representación de la superficie de energía de Born-Oppenheimer y consecuentemente obtener un mejor ajuste a los datos experimentales. Al conjunto resultante de funciones de energía se le llamo campo de fuerza consistente (CFC), el término "campo de fuerza" se tomó del análisis de modos normales de vibraciones moleculares, y el término "consistente" enfatiza la característica principal de este método, es decir, el cálculo simultáneo y consistente de diferentes propiedades de familias de moléculas.

Este campo de fuerza de coordenadas internas y distancias interatómicas fué posteriormente extendido por A. Warshel y S. Lifson al estudio de la fase cristalina [2]: este incluyó el cálculo de los parámetros de la celda unitaria cristalina y su dependencia con la temperatura, vibraciones de red, energías

de vibración y calores de sublimación de las moléculas n-hexano y n-octano, todo esto calculado con las funciones del potencial para interacciones interatómicas optimizadas. Así esto permitió el análisis de los parámetros de dichas funciones. Con esta extensión, se obtuvo una mejor representación de la energía potencial interatómica, tanto intra- como intermolecular.

Este campo de fuerza fué posteriormente transferido (y parcialmente re-examinado) por A.T. Hagler, E. Huler y S. Lifson para crear una función de energía para péptidos y proteínas que incluyera el puente de hidrógeno [3]. En este análisis un mejor ajuste se buscó para las estructuras y energías de red de los cristales amida, compuestos únicamente formados de $C = O$, NH y grupos CH_n .

El aspecto particular que requiere un comentario aquí es el modo en que la geometría cristalina fué tratada: Todas las moléculas de la celda unitaria son mantenidas fijas y los parámetros de la energía son cambiados de modo que den un mejor ajuste a las nueve componentes de los tres vectores de la celda unitaria. Sea \mathbf{a} el vector formado por estas nueve componentes, en vez de calcular \mathbf{a}_{calcd} por una tediosa minimización de la energía, se estimó $\Delta\mathbf{a} = \mathbf{a}_{calcd} - \mathbf{a}_{exptl}$ por

$$\Delta\mathbf{a} = -\mathbf{F}(\mathbf{a}_{exptl})^{-1}\nabla_{\mathbf{a}}V(\mathbf{a}_{exptl}) \quad (2.1)$$

basado en la fórmula de Newton-Raphson. Aquí \mathbf{F} es la matriz de segundas derivadas, $F_{\alpha\beta} = \partial^2V/\partial a_{\alpha}\partial a_{\beta}$, y $\nabla_{\mathbf{a}}V$ es un vector cuyas componentes son $\partial V/\partial a_{\alpha}$. Tanto \mathbf{F} y ∇V son calculados en la estructura cristalina experimental (\mathbf{a}_{exptl}). Los parámetros, representados por un vector \mathbf{p} , son cambiados iterativamente de acuerdo a la fórmula

$$\Delta\mathbf{p}^i = -[(Z^T\mathbf{W}^2\mathbf{Z})^{-1}Z^T\mathbf{W}^2\Delta\mathbf{y}]^{i-1}. \quad (2.2)$$

La cual se tomó de la fórmula iterativa de optimación por mínimos cuadrados $\delta\mathbf{x} = -[Z^T\mathbf{P}^2\mathbf{Z} + \epsilon\mathbf{I}]^{-1}Z^T\mathbf{P}^2\Delta\mathbf{y}(\mathbf{x})$ desarrollada por Warshel y Lifson y que se presentó en el capítulo anterior. En esta ocasión el vector $\Delta\mathbf{y}$ incluye todas las componentes del vector $\Delta\mathbf{a}$ dadas por la ecuación (2.1), así como

las diferencias en la energía de red (ΔV_L) y momentos dipolares ($\Delta\mu$) en la $(i - 1)$ iteración, Z es la matriz jacobiana, $Z_{lm} = \partial y_l / \partial p_m$, y W es la matriz diagonal de los factores de peso de los varios observables.

Finalmente, S.J. Weiner *et al* [4] desarrollaron un campo de fuerza empírico para ácidos nucleicos y proteínas dentro del contexto de refinamiento de energía de Lifson y Warshel. Este conjunto de parámetros fué el primer campo de fuerza para ácidos nucleicos en el cual se llevo a cabo una optimización completa de todos los grados de libertad atómicos, y el cual fué razonablemente exitoso en reproducir las estructuras, energías y frecuencias vibracionales de los sistemas modelo. Por razones de eficiencia computacional ese campo de fuerza usó una representación de los grupos CH , CH_2 y CH_3 como si fueran un sólo átomo esférico. Debido a esta aproximación, algunos compromisos se tuvieron que hacer, los cuales condujeron en algunos casos a ajustes no tan óptimos con los experimentos. Calculos recientes sobre interacciones de ácidos nucleicos también han sugerido que cuando uno está examinando pequeñas diferencias de energía, una representación esférica de los grupos CH conduce a un más pobre acuerdo con el experimento que una representación que incluya a todos los átomos.

Esta función de campo de fuerza presentada aquí, fué el resultado de solucionar estos inconvenientes y fué presentada por S.J. Weiner, P.A. Kollman, D.T. Nguyen y D.A. Case [5] e incluye todos los átomos explícitamente de las proteínas y los ácidos nucleicos. Los parámetros de esta función también se usaron para llevar a cabo el primer análisis vibracional general de las cinco bases de los ácidos nucleicos con un potencial de mecánica molecular aproximado.

2.2 El potencial semiempírico

La función de energía semiempírica que usamos tiene la siguiente forma:

$$V = \sum_{\text{enlaces}} K_R(R - R_0)^2 + \sum_{\text{angulos}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{torsiones}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{332.17q_i q_j}{\epsilon R_{ij}} \right] + \sum_{H-\text{enlaces}} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right] \quad (2.3)$$

donde R , K_R and R_0 son la longitud de enlace, constante de fuerza de alargamiento de enlace, y distancia de equilibrio respectivamente para los átomos enlazados covalentemente; las funciones de alargamiento y doblamiento son cuadráticas (armónicas sencillas), las cuáles permiten una descripción adecuada de las estructuras y energías para unas relativamente poco tensionadas moléculas de ácidos nucleicos y proteínas. θ , K_θ y θ_0 son el ángulo de enlace, constante de fuerza para doblamiento de ángulo y ángulo de equilibrio, respectivamente para los ángulos entre los enlaces covalentes a un átomo dado; ϕ es un ángulo diedro entre 4 átomos (si A-B-C-D son cuatro átomos y "-" representa un enlace covalente, entonces los átomos A-B-C formarán un plano 1, y los átomos B-C-D formarán otro plano 2, el ángulo ϕ es entonces el ángulo formado por la intersección de los planos 1 y 2, $0 < \phi < \pi$). Una expansión de Fourier de la energía torsional en términos del ángulo diedro ha sido llevada a cabo, esta serie nos ofrece una simulación bastante precisa de las preferencias conformacionales en moléculas simples y complejas.

$V_n/2$ es la constante de fuerza, n es la multiplicidad en la expansión (más de un valor de n puede ser usado por ángulo diedro) y γ es un valor de fase. Juntos, estos tres términos dan la variación en la energía de enlace covalente de la molécula. La cuarta suma es el término no ligado y representa las energías electrostática y de Van der Waals (Lennard-Jones).

Hagler et al. [3] compararon las funciones no ligadas 6-12 y 6-9 en cálculos de empaquetamiento cristalino y encontraron que ninguna es claramente superior. Mientras que la distancia interatómica no esté debajo de la suma de los radios de Van der Waals, la forma 6-12 será adecuada. Las cantidades R_{ij} , q_i y q_j son las distancias de no-enlace y las cargas respectivamente para el par de átomos i y j . El factor ϵ es un parámetro dieléctrico efectivo. Weiner et al. usaron $\epsilon = r_{ij}$ con la razón de que usar una constante dieléctrica dependiente de la distancia nos ayuda a crear los efectos de polarización en interacciones atractivas, con interacciones mas fuertes multiplicadas por un factor de peso mayor. Además, esto ayuda a compensar la pérdida de solvatación explícita al amortiguar más el rango mayor de las interacciones

de carga y a su vez le da un mayor peso a las interacciones de rango más corto. Hay un soporte computacional y empírico para tal modelo, dado que el solvente (agua) en los experimentos no está explícitamente incluido en el cálculo. Sin embargo, cuando el agua está explícitamente incluida, una ϵ constante es probablemente más apropiada. El último término es la energía de los enlaces de hidrógeno, la cual aparece en la unión de los pares de bases (Adenina-Timina y Citosina-Guanina). Las unidades de V son Kcal/mol con R en angstroms, θ y ϕ en radianes, q_i en unidades electrostáticas (dada la carga Mulliken), y 332.17 es un factor de conversión.

2.3 Tipos de átomos

La base de un campo de fuerza es la elección de los tipos de átomos; esto es, la selección de átomos los cuáles son bastante parecidos, tanto físicamente como químicamente, para ser tratados idénticamente en los refinamientos de mecánica molecular. En el caso de un cálculo de mecánica cuántica, uno necesita únicamente un simple "tipo" de átomo; por que los grados de libertad electrónicos son tomados en cuenta explícitamente. En nuestro cálculo clásico las decisiones sobre los tipos de átomos son compromisos inevitables entre poseer la representación más precisa de muchas moléculas y el tener un número manejable de tipos.

En la figura 2.1 mostramos las 2 versiones de la molécula adenina: la que se presenta en la naturaleza con los átomos verdaderos, y la que está formada por unos tipos de átomos que resultaron del estudio de varios compuestos y que nos permiten una representación a la vez aproximada y manejable de los verdaderos.

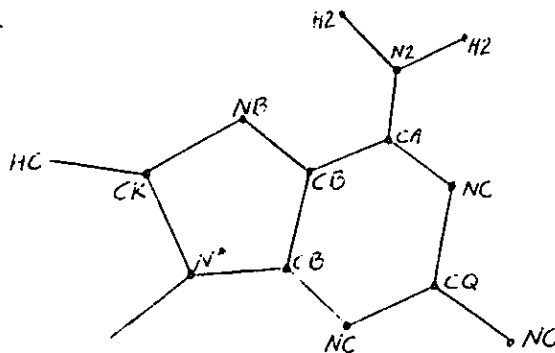


Figura 2.1. Representación de los átomos de la molécula adenina, en la representación natural C es Carbono, N nitrógeno y H hidrógeno, para nuestro modelo el significado de los átomos está dado en la tabla 2.1.

En la siguiente lista se muestran los tipos de átomos usados y sus características. Los átomos del tipo sp^3 son realmente más típicos en la naturaleza, pero se han incluido más del tipo sp^2 para asegurar una mayor precisión geométrica para sistemas de cadena tales como purinas, pirimidinas, indoles e imidasoles.

átomo	tipo
carbonos unidos a hidrógeno(s)	
CQ	sp^2 carbono en un anillo de purina de seis miembros entre dos nitrógenos "NC" y ligado a un hidrógeno
CK	sp^2 carbono aromático en un anillo de 5 miembros entre dos nitrógenos y ligado a un hidrógeno (en purinas)
demás carbonos	
CA	sp^2 carbono aromático en un anillo de seis miembros con un sustituyente
CB	sp^2 carbono aromático situado en la unión entre anillos de 5 y 6 miembros (esto es, C4 y C5 en purinas)
nitrógenos	
NB	sp^2 nitrógeno en un anillo de cinco miembros con pares solos (esto es, N7 en las purinas)
N*	sp^2 nitrógeno que se encuentra en las purinas y las pirimidinas con el grupo alquil ligado (N9 en las purinas y N1 en las pirimidinas)
NC	sp^2 nitrógeno en un anillo de seis miembros con los pares solos (esto es, N3 en la adenina)
N2	sp^2 nitrógeno en los grupos NH_2 base y arginina NH_2
hidrógenos	
H2	hidrógenos amino del NH_2 en las purinas y las pirimidinas
HC	hidrógeno explícito ligado a carbono

Tabla 2.1. Etiquetas de los átomos de la adenina y su correspondiente significado de su origen.

2.4 Parámetros de átomos no ligados

El conjunto más difícil para derivar a priori son los de átomos no ligados. Como punto de partida para los átomos sp^2 se utilizaron los parámetros para interacción 6-12 y 6-9 derivados por Hagler et al. [3]. La diferencia significativa entre los valores para 6-12 y 6-9 de R (distancia donde se encuentra el mínimo de Van der Waals) y ϵ (profundidad del potencial de Van der Waals) originó que los valores no se tomaran directamente, sino que sufrieran un pequeño cambio; así, los valores originales para el campo de fuerza 6-9 sufrieron algunas modificaciones al cambiarlos al de la forma 6-12. Para los grupos CH , CH_2 y CH_3 también se utilizaron parámetros apropiados para la energía 6-12 de Van der Waals, tomados de datos experimentales sobre cristales de hidrocarburos así como simulaciones de líquidos con alcoholes y éters. Estos dos conjuntos de parámetros obtenidos tanto experimentalmente como teóricamente fueron muy similares, sugiriéndonos así unos valores apropiados para los átomos unidos (así se les conoce a los grupos CH , CH_2 y CH_3). En la figura 2.2 se muestra el potencial que existe entre 2 átomos no ligados covalentemente dependiendo de su distancia de separación, los parámetros A_{ij} y B_{ij} que aparecen en la fórmula (2.3) vienen dados en función de los valores correspondientes R_i , R_j , ϵ_i y ϵ_j (mostrados en la tabla 2.2) para el par de átomos i, j . $U(R_{ij}) = \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6}$ con $A_{ij} = \frac{1}{2^{1/2}}(R_i + R_j)^{12} \sqrt{\epsilon_i \epsilon_j}$, $B_{ij} = \frac{1}{2^3}(R_i + R_j)^6 \sqrt{\epsilon_i \epsilon_j}$. Los valores de la distancia de equilibrio y la energía mínima son: $r_m = \left(\frac{2A_{ij}}{B_{ij}}\right)^{1/6}$, y $E_d = -\frac{B_{ij}^2}{4A_{ij}}$.

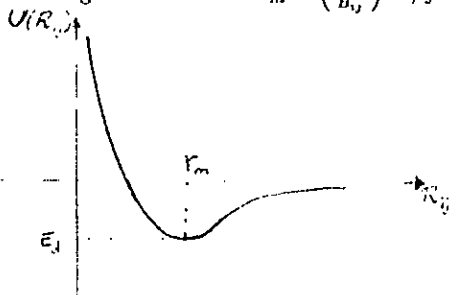


Fig. 2.2. Potencial Van der Waals entre cualquier 2 átomos de la Adenina.

En la tabla 2.2 se muestran los valores de R_i y ϵ_i para los distintos tipos de átomos que conforman la molécula adenina.

átomo	R_i	ϵ_i
CA	1.85	0.120
CB	1.85	0.120
CK	1.85	0.120
CQ	1.85	0.120
NB	1.75	0.160
NC	1.75	0.160
N*	1.75	0.160
N2	1.75	0.160
HC	1.54	0.010
H2	1.00	0.020

Tabla 2.2. R_i y ϵ_i para los distintos átomos, las unidades de R_i son angstroms y las de ϵ_i son kcal/mol.

2.5 obtención de los parámetros de θ y R

Las longitudes de enlace de equilibrio y ángulos fueron obtenidos de datos de microondas, difracción de neutrones y rayos-X en compuestos apropiados. Los valores de K_R fueron obtenidos de cálculos de modos normales en los que la K_R fué variada para mejor ajustar las frecuencias experimentales de varias moléculas. Los valores iniciales de K_θ vinieron de MM2, pero el análisis de modo normal también jugo un papel importante.

ángulo	K_θ	θ_{eq}
N*-CK-NB	70.0	113.9
CK-NB-CB	70.0	103.8
NB-CB-CB	70.0	110.4
CB-CB-N*	70.0	106.2
CK-N*-CB	70.0	105.4
CB-CB-CA	85.0	117.3
CB-CA-NC	70.0	117.3
CA-NC-CQ	70.0	118.6
NC-CQ-NC	70.0	129.1
CB-NC-CQ	70.0	111.0
CB-CB-NC	70.0	127.7
HC-CK-NB	35.0	123.0
NB-CB-CA	70.0	132.4
CB-CA-N2	70.0	123.5
CA-N2-H2	35.0	120.0
H2-N2-H2	35.0	120.0
H2-N2-CA	35.0	120.0
N2-CA-NC	70.0	119.3
NC-CQ-HC	35.0	115.4
NC-CB-N*	70.0	126.2
N*-CK-HC	35.0	123.0

Tabla 2.3. Valores de K_θ y θ_{eq} para la aproximación armónica de la energía de variación de un ángulo formado por 3 átomos unidos covalentemente. Las unidades de K_{theta} son kcal/mol * radianes² y las de θ_{eq} son grados.

En la siguiente tabla están los valores de K_r y r_{eq} para los distintos enlaces covalentes de la molécula adenina, esto en la aproximación de considerar a la energía de variación de enlace como si fuera la de un resorte (armónico simple).

enlace	K_r	r_{eq}
CK-NB	529	1.304
NB-CB	414	1.391
CB-CB	520	1.370
CB-N*	436	1.374
N*-CK	440	1.371
CB-CA	469	1.404
CA-NC	483	1.339
NC-CQ	502	1.324
NC-CB	461	1.354
HC-CK	340	1.080
CA-N2	481	1.340
N2-H2	434	1.010
CQ-HC	340	1.080

Tabla 2.4. Valores de K_r y r_{eq} utilizados en la aproximación armónica de la energía de variación de enlace entre 2 átomos unidos covalentemente. Las unidades de K_r son kcal/mol * angstroms² y las de r_{eq} angstroms.

2.6 Parámetros de torsión

Los parámetros de torsión inicialmente vinieron de datos experimentales sobre conformación de moléculas en equilibrio. Sin embargo, como los términos no-ligados y los de torsión son muy dependientes entre ellos, apropiadas variaciones tuvieron que ser hechas. Es importante recalcar aquí que de los ángulos de torsión presentados en la siguiente tabla, algunos representan a 4 diferentes ángulos de torsión y otros solo representan a 2 diferentes ángulos de torsión. Esto se puede ver de la figura 2.1 que representa la molécula, por ejemplo, el ángulo de torsión X-CA-CB-X engloba los 4 siguientes: N2-CA-CB-NB, N2-CA-CB-CB, NC-CA-CB-NB y NC-CA-CB-CB. Mientras que el ángulo de torsión X-CA-NC-X representa las 2 siguientes: N2-CA-NC-CQ y N2-CA-NC-CQ.

torsión	$V_n/2$	γ	n
X-CA-CB-X	10.2	180.0	2
X-CA-N2-X	6.8	180.0	2
X-CA-NC-X	9.6	180.0	2
X-CB-CB-X	16.3	180.0	2
X-CB-NB-X	5.1	180.0	2
X-CB-N*-X	6.6	180.0	2
X-CK-NB-X	20.0	180.0	2
X-CK-N*-X	6.7	180.0	2
X-CQ-NC-X	13.5	180.0	2
X-CB-NC-X	20.0	180.0	2

Tabla 2.5. Posibles valores de los parámetros de la energía de variación de ángulo de torsión, las unidades de $V_n/2$ son kcal/mol, γ está en grados y n es adimensional.

2.7 Obtención de los valores para las cargas q_i

Los valores de las cargas fueron derivados por medio de un uso de potenciales electrostáticos derivados cuánticamente, los cuales fueron ajustados a un modelo de carga puntual para átomos en determinadas moléculas. Aunque este procedimiento parece ser uno de los mejores para derivar los valores de las cargas, todavía está sujeto a varias discusiones. Primero, mientras que los valores de las cargas pueden depender de la conformación de la molécula usada, es impráctico poseer un conjunto separado de cargas para cada conformación. Segundo, hacer cálculos cuánticos es práctico solo para fragmentos de polímeros, por lo que estos después deben ser "pegados" para formar la molécula total.

Viendo las posibles variaciones de los valores debido a estos inconvenientes, Weiner et al. [4] encontraron que estos pueden variar en un rango del 0 a 10% ,(para esto utilizaron conformaciones típicas y diferentes conjuntos de bases de moléculas). En la tabla 2.6 se muestran los diferentes valores de las cargas para los átomos que forman la adenina.

átomo	$q_\alpha = c_\alpha e$
CK	0.263
NB	-0.543
CB	-0.097
CB	0.546
N*	-0.073
CA	0.769
NC	-0.774
CQ	0.661
NC	-0.728
HC	0.062
H2	0.324
N2	-0.768
H2	0.335
HC	-0.032

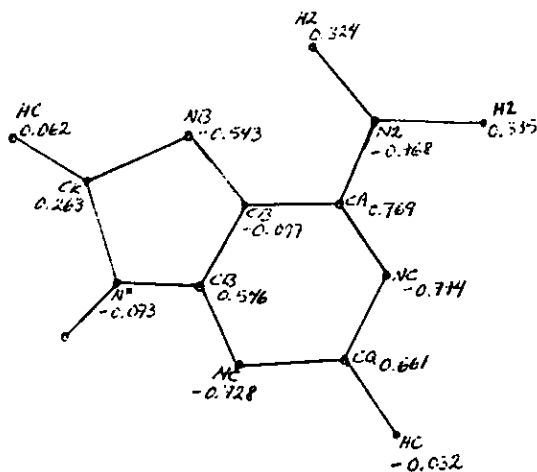
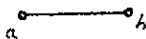


Tabla 2.6. Valores de las cargas para los átomos de la adenina, nótese que dos mismos átomos pueden tener cargas diferentes, para ver la correspondencia correcta a la derecha se muestra la molécula y las cargas para cada átomo.

2.8 R , ϕ y θ como función de las coordenadas

El potencial, a pesar de ser función de R , θ , ϕ y \hat{R}_{ij} , debe ser convertido a una función de las coordenadas cartesianas de los 14 átomos de la molécula, hicimos esto para que el número de variables de manejar por el algoritmo genético sea más pequeño (aproximadamente igual al número de grados de libertad del problema). Para que esto quede claro, enseguida mostraremos como las variables R , θ y ϕ son funciones de las coordenadas de los átomos respectivos.

a) Coordenada de variación de enlace.



Considere dos átomos ligados como en la figura. El vector del átomo a al

átomo b puede ser escrito en términos de cualquier sistema cartesiano como

$$\mathbf{r}_{ab} = (x_b - x_a)\mathbf{i} + (y_b - y_a)\mathbf{j} + (z_b - z_a)\mathbf{k}$$

donde \mathbf{i} , \mathbf{j} y \mathbf{k} son vectores unitarios a lo largo de los ejes X, Y y Z del sistema de referencia. El módulo r_{ab} del vector \mathbf{r}_{ab} está entonces dado por:

$$r_{ab} = \sqrt{\mathbf{r}_{ab} \cdot \mathbf{r}_{ab}} = [(x_b - x_a)^2 + (y_b - y_a)^2 + (z_b - z_a)^2]^{1/2}$$

b) Coordenada de doblamiento de ángulo de valencia.



Considere dos enlaces con un átomo en común, como en la figura. El producto escalar entre los vectores \mathbf{r}_{ab} y \mathbf{r}_{ac} es

$$\mathbf{r}_{ab} \cdot \mathbf{r}_{ac} = r_{ab} r_{ac} \cos \alpha = (x_b - x_a)(x_c - x_a) + (y_b - y_a)(y_c - y_a) + (z_b - z_a)(z_c - z_a)$$

c) Coordenada de torsión.

Considere el ángulo diedro ϕ entre los planos determinados por tres consecutivos enlaces conectando cuatro átomos como esta mostrado en la figura de abajo:



Por convención, limitaremos ϕ al intervalo $-\pi \leq \phi \leq \pi$ y consideraremos ϕ positivo si, viendo los átomos en la dirección de el enlace ac , el ángulo del plano bac a el plano acd con a más cerca del del observador, es trazado en el sentido de las manecillas del reloj. La figura 1a muestra un ejemplo de un ángulo ϕ positivo, mientras que la figura 1b muestra un ejemplo de un ángulo ϕ negativo.

La relación analítica que conecta ϕ con las coordenadas cartesianas de los átomos puede ser encontrada al formar el producto escalar entre un vector perpendicular al plano bac y un vector perpendicular al plano acd . Entonces tenemos en este caso:

$$\begin{aligned}
 (\bar{r}_{ab} \times \bar{r}_{ac})(\bar{r}_{ac} \times \bar{r}_{cd}) &= r_{ab} r_{ac}^2 r_{cd} \sin \beta_a \sin \beta_c \cos \phi \\
 &= [(x_c - x_a)(y_d - y_c) - (y_c - y_a)(x_d - x_c)] [(x_a - x_b)(y_c - y_a) - (y_a - y_b)(x_c - x_a)] \\
 &+ [(z_c - z_a)(x_d - x_c) - (x_c - x_a)(z_d - z_c)] [(z_a - z_b)(x_c - x_a) - (x_a - x_b)(z_c - z_a)] \\
 &+ [(y_c - y_a)(z_d - z_c) - (z_c - z_a)(y_d - y_c)] [(y_a - y_b)(z_c - z_a) - (z_a - z_b)(y_c - y_a)]
 \end{aligned}$$

3 Capítulo 3

3.1 Métodos de optimización

En la actualidad, los métodos de búsqueda y de optimización se pueden agrupar en 3 diferentes tipos: los métodos basados en el Cálculo, los Numerativos y las técnicas inteligentes; A continuación daremos una semblanza breve de estos, empezando por los que históricamente fueron los primeros en desarrollarse: los métodos basados en el Cálculo. Los métodos basados en técnicas de Cálculo se subdividen en 2 principales clases: directos e indirectos. Los métodos indirectos buscan los extremos locales tratando de resolver el conjunto usualmente no lineal de ecuaciones que resultan de igualar el gradiente de la función a cero. Este método es la generalización a varias dimensiones de la noción de puntos extremos con pendientes de cero en todas las direcciones dada en cálculo elemental. Por otra parte, los métodos directos buscan los óptimos locales al situarse en la superficie de la función y luego moverse en la dirección dada por el gradiente.

Aunque ambos métodos han sido grandemente estudiados, extendidos y mejorados, un análisis simple muestra su falta de efectividad. Primero, ambos métodos son locales en su alcance, ya que los puntos óptimos que ellos buscan están en una porción del dominio completo; además, si queremos medir la altura de dos picos con este método para ver cual es más alto, primero debemos subir uno y luego otro para comparar cual es el más alto, ya que este método no nos dice cual pico hay que subir desde el principio. Segundo, los métodos basados en cálculo dependen de la existencia de continuidad y de derivadas, limitando así su aplicación a un número limitado de problemas, ya que en el mundo real no todos los espacios de investigación son continuos y derivables.

Los diversos métodos numerativos están basados en una idea bastante simple: en un espacio de búsqueda finito, o en un espacio de búsqueda infinito discretizado, este algoritmo comienza a mirar los valores de la función objetivo en cada punto del espacio, uno por uno. Aunque la simplicidad de

este tipo de algoritmo es atractiva, ultimadamente tal método debe ser rechazado por una simple razón: falta de eficiencia. Muchos espacios prácticos son simplemente demasiados grandes para buscar punto por punto y todavía tener una posibilidad de usar esta información para un fin práctico. Más aún, el mejor programa enumerativo desarrollado, llamado por su creador Bellman "dynamic programming", falla por falta de eficiencia en problemas de tamaño y complejidad moderados. Entonces debemos concluir que los métodos numerativos no son métodos de optimización eficientes. Las técnicas de búsqueda inteligentes están basadas en el método numerativo, pero usan información obtenida del espacio de búsqueda para guiar la búsqueda. Las dos principales subclases son simulated annealing y algoritmos evolutivos.

3.2 Algoritmos genéticos

La característica principal de los cálculos evolutivos es la aplicación de los conceptos de evolución por selección natural a una población de "individuos"; estos usan movimiento al azar en un espacio de parámetros codificados. Usar movimiento al azar como una herramienta en un proceso de búsqueda parece extraño al principio, pero pronto veremos como su efectividad es muy buena. Las diferencias de los algoritmos genéticos que hacen posible sus buenos resultados comparados con los otros métodos tradicionales son 4: (1) Los algoritmos genéticos trabajan con un conjunto de parámetros codificados, no con los parámetros mismos. (2) Los algoritmos genéticos buscan en una población de puntos de soluciones intermedias, no en un simple punto; escalando así muchos picos simultáneamente, por lo que la probabilidad de quedarnos en el pico no global se reduce. (3) Los algoritmos genéticos usan información de una función objetivo, no derivadas u otro conocimiento auxiliar. (4) Los algoritmos genéticos usan reglas de transición probabilísticas, no reglas deterministas para que a partir de las soluciones intermedias buenas se construyan nuevas soluciones al problema. Los algoritmos genéticos requieren que el conjunto de parámetros de los que dependa el problema sean codificado en una cadena de longitud finita usando algún alfabeto finito.

Como ejemplo, consideremos el siguiente problema: deseamos maximizar el valor de la función $f(x) = x^2$ sobre el intervalo de enteros $[0,31]$. Con algoritmos genéticos el primer paso del proceso de optimización es codificar la variable x en una cadena de longitud finita; aunque hay varios métodos para codificar la variable x , nosotros utilizaremos en este ejemplo la forma más usual y más simple para los operadores, que es la representación binaria (utilizando 0's y 1's) de la variable x . En la última sección mostraremos la representación binaria Gray que es más conveniente y es la que utilizamos nosotros.

3.3 Creación de la población inicial

En general, es aconsejable utilizar una población inicial bastante grande, ya que esto mejorará la variabilidad y entonces la posibilidad a tener algunos valores en común con la solución óptima. Si la población inicial es bastante pequeña, una solución inicial relativamente buena puede solamente tener muchos valores en común con una solución subóptima y las probabilidades serán muy altas de que sea esta solución, y no la óptima, la única que sea encontrada. En general, las soluciones iniciales vienen de elegir totalmente al azar los valores de los parámetros dentro de los rangos definidos. Para nuestro ejemplo, el número de valores de x que utilizaremos como población inicial será de tamaño $n = 4$ (muy pequeño para los algoritmos genéticos estándar) y serán los siguientes 4 números con cada dígito generado totalmente al azar:

```
0 1 1 0 1
1 1 0 0 0
0 1 0 0 0
1 0 0 1 1
```

Un modo de crear esta población inicial consiste en repetir 4 veces 5 lanzamientos de una moneda, donde el águila es igual a 1 y el sol el igual a 0. Después de esto, poblaciones sucesivas son generadas utilizando el algoritmo genético. Los mecanismos de un algoritmos genético simple son sorprendentemente sencillos, involucrando nada más complejo que copiado de cadenas e

intercambio de cadenas parciales, siendo esta misma sencillez de operaciones y su gran efecto dos de los principales atractivos de los algoritmos genéticos. Definiremos enseguida un conjunto de operaciones simples que manipularán esta población inicial y generarán poblaciones sucesivas que (eso esperamos) sean cada vez mejores. Un algoritmo genético simple que produzca buenos resultados está compuesto de tres operadores:

1. Selección.
2. Entrecruzamiento.
3. Mutación.

3.4 Selección

La selección es un proceso en el cual las cadenas individuales son copiadas de acuerdo a los valores de la función objetivo $f(x)$ (conocida también como función de fineza). Intuitivamente, el valor de $f(x)$ es el que se desea optimizar y es una medida de que tan bueno es el valor de la variable x . El copiado de cadenas de acuerdo a su fineza significa que las cadenas con un valor mejor de la función objetivo tienen una más alta probabilidad de contribuir con uno o más descendientes en la siguiente generación. Este operador, por supuesto, es una versión artificial de la selección natural de Darwin, según la cual en la naturaleza los individuos que estén mejor adaptados para la competición de los recursos y su supervivencia de los depredadores son los que sobreviven y entonces tienen una mayor oportunidad de reproducirse. En nuestro proceso artificial, la función objetivo es el árbitro de la vida o muerte de las creaturas-cadenas, y es ésta la que juega el papel de presión del medio ambiente en la evolución darwiniana. La manera más fácil de implementar el operador de reproducción es creando una rueda de ruleta en donde cada cadena de la población tenga una porción de la ruleta en proporción a su fineza, mostrándose esto en la figura 1. Para nuestro ejemplo, calculando los valores f de las cuatro cadenas y sumándolos, obtenemos que la fineza total de las cuatro cadenas es de 1170, el porcentaje de la fineza total para cada cadena también está mostrado en la figura 1 así como la rueda de ruleta.

Para la cadena número 1 el valor de la fineza es de 169, el cual corresponde a 14.4% de la fineza total; como consecuencia, a la cadena 1 se le da 14.4% de la rueda de ruleta. Para reproducir, simplemente giremos nuestra rueda de ruleta así definida cuatro veces; así, en cada giro, la probabilidad de que el dardo arrojado se clave en la zona de la cadena 1 es de 0.144.

número de cadena	cadena	fineza	porción de ruleta
1	01101	169	0.14
2	11000	576	0.49
3	01000	64	0.06
4	10011	361	0.31

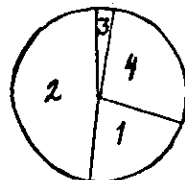


Figura 3.1. Forma de la rueda de ruleta para la población inicial.

Cada vez que necesitemos otro individuo para reproducción, un simple giro de nuestra rueda de ruleta nos lo dará. De este modo, las cadenas más finas tienen un número más grande de descendientes en la siguiente generación. Una vez que una cadena ha sido elegida para reproducción, una réplica exacta de la misma es hecha. Para casos más generales, los pasos anteriormente descritos se pueden generalizar fácilmente; si f_i denota la fineza de la cadena i de un total de n cadenas, la probabilidad de que esta cadena sea elegida para reproducción será de $f_i / \sum_{k=1}^n f_k$ y el número aproximado de copias que esta cadena tendrá en el "club de matrimonio" será de $n * f_i / \sum_{k=1}^n f_k$, este número es igual al número de copias que tendrá esta cadena después de que el proceso de reproducción se ha efectuado n veces.

3.5 Reproducción: (1) entrecruzamiento

Esta cadena es entonces introducida en un grupo de sobrevivientes para una posterior acción del operador genético de entrecruzamiento. Después de la reproducción, el simple entrecruzamiento procede en dos pasos. Primero, a

una cadena del grupo de sobrevivientes se le asignará al azar otra cadena del mismo grupo. Segundo, cada par de cadenas sufrirá entrecruzamiento como sigue: un número entero de posición k a lo largo de la cadena es elegido al azar entre 1 y la longitud de la cadena menos uno $[1, l - 1]$. Dos nuevas cadenas son creadas por el intercambio de todos los caracteres entre las posiciones $k + 1$ y l inclusive. Por ejemplo, consideremos las cadenas A_1 y A_2 de la población inicial de nuestro ejemplo:

$$A_1 = 0\ 1\ 1\ 0\ | \ 1$$

$$A_2 = 1\ 1\ 0\ 0\ | \ 0$$

Supóngase que al escoger un número al azar entre 1 y 4, nosotros obtuvimos $k = 4$ (como está indicado por el símbolo separador $|$). Como resultado del entrecruzamiento se producirán 2 nuevas cadenas A'_1 y A'_2 que serán parte de la nueva generación

$$A'_1 = 0\ 1\ 1\ 0\ 0$$

$$A'_2 = 1\ 1\ 0\ 0\ 1$$

3.6 Reproducción: (2) mutación

El operador de mutación induce alteraciones al azar y esporádicas en las cadenas de bits, tal como se muestra abajo, después de que en la cadena ha habido una mutación en un lugar indicado por la flecha, la nueva cadena que resulta muestra el resultado

$$1\ 1\ 0\ \downarrow\ 1\ 0$$

$$1\ 1\ 0\ 0\ 0$$

La probabilidad de ocurrencia de una mutación en los algoritmos genéticos es muy pequeña (del orden de centésimas o milésimas). Por ejemplo, si $p_{mut} = 0.001$ significa que de cada 1000 bits manejados uno cambiará de valor a su complemento; en general, si la generación actual tiene n cadenas y cada cadena tiene l bits, entonces en ésta generación ocurrirán $n * l * p_{mut}$ mutaciones. En realidad, hay mucha confusión acerca del rol de la mutación

en la genética (tanto natural como artificial). La mutación es necesaria para el intentar prevenir una convergencia prematura a una solución subóptima, ya que aunque la reproducción y el entrecruzamiento efectivamente buscan y combinan diferentes soluciones, ocasionalmente ellos pueden perder un material genético potencialmente útil (1's o 0's en localizaciones particulares). En los algoritmos genéticos, el operador de mutación es una protección contra tal pérdida irreversible.

Regresando al problema de maximizar $f(x) = x^2$ para x entero entre 0 y 31, una simulación real de este proceso usando algoritmos genéticos con los 2 procesos anteriormente descritos (reproducción por mutación y entrecruzamiento) está mostrado en la tabla 3.1. Asumiendo que la probabilidad de mutación en este ejemplo es de 0.001. Con 20 bits transferidos se esperarían que $20 \cdot 0.001 = 0.02$ bits sufran mutación durante una generación dada, por lo que durante esta primera generación no se espera que ocurra una mutación.

Cadena No.	Poblacion inicial generada al azar	Valor de x entero	$f(x)$ x^2	probabilidad de seleccion para reprod. $f_i / \sum f$
1	0 1 1 0 1	13	169	0.14
2	1 1 0 0 0	24	576	0.49
3	0 1 0 0 0	8	64	0.06
4	1 0 0 1 1	19	361	0.31
		Suma	1170	1.00
		Promedio	293	0.25
		Valor máx	576	0.49

Cadena número	número esperado $f_i/\bar{f} = n f_i / \sum f$	valor real después de la rueda de ruleta con factor de peso
1	0.58	1
2	1.97	2
3	0.22	0
4	1.23	1
suma	4.00	4.00
promedio	1.00	1.00
Valor máx	1.97	2.00

Cadena número	gpo de sobrevivientes después de la reprod. mostrando el lugar para entrecruzamiento)	Pareja elegida al azar	Sitio de entrecruzamiento elegido al azar	Población nueva (siguiente generación)
1	0 1 1 0 1	2	4	0 1 1 0 0
2	1 1 0 0 0	1	4	1 1 0 0 1
3	1 1 0 0 0	4	2	1 1 0 1 1
4	1 0 0 1 1	3	2	1 0 0 0 0

	Valor de x (de la pobl. nueva)	$f(x)$ x^2
	12	144
	25	625
	27	729
	16	256
Suma		1754
Promedio		439
Valor máx		729

Tabla 3.1. Ejemplo de como actua un algoritmo genético en una población inicial.

Después de reproducción con entrecruzamiento y mutación, tendremos una nueva población lista para ser examinada. Primero, decodificamos las

nuevas cadenas creadas por el algoritmo genético y después calculamos los valores de la función de fineza para estos valores de x . De la tabla, notemos como ambos, los valores máximo y el promedio, han sido mejorados en la nueva población: la fineza promedio de la población se ha mejorado de 293 a 439 en una generación, y la fineza máxima se ha incrementado de 576 a 729 durante el mismo período.

3.7 Similaridades importantes

Ya hemos visto que a pesar de que los algoritmos genéticos manejan operaciones guiadas al azar verdaderamente muy simples, los resultados que se obtienen son muy buenos. Para ver el porqué de esto, en esta sección introduciremos unos conceptos más generales que nos ayudarán a mostrar el fundamento matemático que hay detrás de los algoritmos genéticos y la razón de su enorme efectividad y potencia. Para esto, consideremos la población inicial del ejemplo pasado:

cadena	fineza
0 1 1 0 1	169
1 1 0 0 0	576
0 1 0 0 0	64
1 0 0 1 1	361

Qué información está contenida en esta población que guía una búsqueda para su mejoramiento? A un primer vistazo, parece que no mucha: 4 cadenas diferentes e independientes con sus respectivos valores de fineza; pero si las observamos con más detalle, veremos que las cadenas que comienzan con un 1 parecen estar entre las mejores (ciertamente con nuestra función $f(x) = x^2$ y nuestra codificación binaria sabemos que ésto es cierto). Aquí, nosotros estamos haciendo dos procesos distintos; primero, estamos buscando similitudes entre la población; segundo, estamos buscando una relación entre estas similitudes y su alta fineza. Al nacerlo así, estamos aceptando el peso de nueva información para ayudar al proceso de búsqueda, para ver

que tanto y que tipo de información aceptaremos, veremos a continuación el concepto importante de schema (en plural schemata) o similarity template, lo que traerá como consecuencia que ya no estemos interesados en las cadenas como objetos aislados.

3.8 Los *schema* y sus propiedades

Un schema describe un subconjunto de cadenas con similaridades en ciertas posiciones de cadena (para esta discusión, vamos a trabajar sin pérdida de generalidad con el alfabeto binario 0, 1). Para que el concepto de schema sea más fácilmente introducido, adicionemos al alfabeto binario el símbolo * o símbolo de "cualquier valor". Con éste alfabeto extendido ahora se podrán crear cadenas (schemata) usando el alfabeto ternario {0, 1, *}. El significado de schema será claro si pensamos a éste como un subconjunto formado por las cadenas que tienen 0's o 1's en las posiciones correspondientes de los 0's y 1's del schema, pero que pueden tener 0 o 1 en las posiciones donde el schema tiene un *. Como ejemplo, el schema *0000 es igual al subconjunto 10000, 00000, asimismo, el schema *1111* describe un subconjunto con 4 miembros 01110, 01111, 11110, 11111.

Algunos schemata son más específicos que otros, por ejemplo, el schema 011*1** nos define unas similaridades mayores que el schema 0*****. Además, el schema 1****1* recorre una porción más grande de la cadena que el schema 1*1****. Esto motiva la introducción de 2 propiedades de los schema: el orden y la longitud definida de los schema. El orden de un schema H, denotado por $o(H)$ es simplemente el número de posiciones fijas no *'s presentes en el schema, por ejemplo, el orden del schema 011*1** es 4, mientras que el orden del schema 0***** es 1. La longitud definida de un schema H, denotado por $\delta(H)$ es la distancia entre la primera y última posiciones de una cadena. Por ejemplo, el schema 011*1** tiene una longitud definida $\delta = 4$ porque la última posición específica es 5 y la primera posición específica es 1, y la distancia entre ellas es $\delta(H) = 5 - 1 = 4$. Para el schema 0*****, la primera y últimas posiciones específicas son las mismas y la longitud definida es $\delta = 0$.

El número de posibles schemata para cadenas de longitud l será de 3^l porque cada una de las posiciones puede ser 0, 1 o *. El número de schemata representados en una cadena individual es 2^l , ya que cada posición puede tomar o su valor actual o el símbolo *, lo cual resulta que una población de tamaño n representa entre 2^l y $n2^l$ schemata, dependiendo de la diversidad. Entonces, de los 2^l a $n2^l$ schemata contenidos en un población. Cuántos son realmente procesados de una manera útil por el algoritmo genético? Para obtener una respuesta, consideraremos el efecto de selección, entrecruzamiento y mutación en el crecimiento o decaimiento de los schemata de generación en generación.

3.9 Efecto de la selección

Empezemos primero con el efecto de selección. Supongámos que a un paso de tiempo dado t hay m ejemplos de un particular schema H contenidos dentro de la población $A(t)$. Esto es, $m = m(H, t)$ (hay posiblemente diferentes cantidades de diferentes schemata H a diferentes tiempos t). Durante la selección, una cadena es copiada de acuerdo a su fineza, o más precisamente una cadena A_i es elegida con probabilidad $p_i = f_i / \sum f_j$. Después de recoger una población de tamaño n con reemplazo de la población $A(t)$, esperaremos tener $m(H, t+1)$ representativos del schema H en la población al tiempo $t+1$ dados por la ecuación $m(H, t+1) = m(H, t) * n * f(H) / \sum f_j$, donde $f(H)$ es la fineza promedio de las cadenas representando al schema H al tiempo t . Si reconocemos que la fineza promedio de la población entera se puede escribir como $\bar{f} = \sum f_j / n$ entonces podremos reescribir la ecuación de crecimiento del schema por reproducción como sigue

$$m(H, t+1) = m(H, t) * \frac{f(H)}{\bar{f}}$$

En palabras, un particular schema crece como la razón de la fineza del schema a la fineza promedio de la población. Puesto de otro modo, los schemata con valores de fineza arriba del promedio de la población recibirán un mayor número de muestras en la siguiente generación, mientras que los schemata

con valores de fineza abajo del promedio de la población recibirán un número menor de muestras. Para que esto se vea mejor, supongámos que un particular schema H permanece arriba del promedio una cantidad $c\bar{f}$, c una constante. Bajo esta suposición, la ecuación para schema se reescribe como

$$m(H, t + 1) = m(H, t) \frac{\bar{f} + c\bar{f}}{\bar{f}} = (1 + c) * m(H, t)$$

Comenzando de $t = 0$ y asumiendo un valor constante c , se obtiene la ecuación

$$m(H, t) = m(H, 0) * (1 + c)^t$$

Esto es, con la selección los schemata con fineza arriba (abajo) del promedio se incrementarán (disminuirán) exponencialmente.

3.10 Efecto del entrecruzamiento

Aunque resulta interesante que el operador de selección pueda incrementar y disminuir los números de individuos de un schemata dado de acuerdo a la fineza del mismo. Por otra parte, la selección sola no promueve la exploración de nuevas regiones del espacio de búsqueda, ya que no son investigados nuevos puntos. Si solamente trabajamos con los mismos individuos, como podremos explorar nuestro espacio para investigar la posible existencia de individuos con valores mejores de la función objetivo ? Es precisamente aquí donde el operador de entrecruzamiento entra.

El entrecruzamiento crea nuevas estructuras con una mínima perturbación a la estrategia de colocación de los mejores schema dictada por la reproducción. Esto resultará en un incremento (disminución) de las proporciones de un schema en una población. Para ver cuales schemata son afectados por entrecruzamiento y cuales no, consideremos una cadena particular de longitud $l = 7$ y dos schemata representativos dentro de esa cadena:

$$A = 0 1 1 1 0 0 0$$

$$H_1 = * 1 * * * * 0$$

$$H_2 = * * * 1 0 * *$$

Claramente, los dos schemata H_1 y H_2 están representados en la cadena A, pero para ver el efecto de entrecruzamiento sobre el schemata, primero recalquemos que el simple entrecruzamiento procede con la selección al azar de una pareja, la selección al azar de un sitio para entrecruzamiento y el consiguiente intercambio de subcadenas. Supongamos que la cadena A ha sido escogida para reproducción y entrecruzamiento. En esta cadena de longitud 7, supongámos que tiramos un simple dado para escoger el sitio de entrecruzamiento (hay 6 sitios en una cadena de longitud 7). Supongámos que el dado nos dió un 3, significando que el sitio de corte tomará lugar entre las posiciones 3 y 4. El efecto de este entrecruzamiento sobre nuestros dos schemata H_1 y H_2 puede ser visto fácilmente en el siguiente ejemplo, donde el sitio para entrecruzamiento ha sido marcado con el símbolo separador |:

$$\begin{aligned}
 A &= 0\ 1\ 1\ | \ 1\ 0\ 0\ 0 \\
 H_1 &= * \ 1\ * \ | \ * \ * \ * \ 0 \\
 H_2 &= * \ * \ * \ | \ 1\ 0\ * \ *
 \end{aligned}$$

A menos que la pareja de la cadena A sea idéntica a A en las posiciones fijas de el schema (una posibilidad que consideraremos sin importancia), el schema H_1 será destruido debido a que el 1 en la posición 2 y el 0 en la posición 7 serán colocados en diferentes descendientes (ellos están en los lados opuestos del símbolo separador marcado por el símbolo de corte). Es igualmente claro que con el mismo punto de corte (entre los bits 3 y 4), el schema H_2 sobrevivirá porque el 1 en la posición 4 y el 0 en la posición 5 serán llevados intactos a un simple descendiente.

Aunque hemos usado un punto de corte específico para ilustración, es claro que el schema H_1 es menos probable que sobreviva al entrecruzamiento que el schema H_2 porque en promedio el punto de corte es más probable que caiga entre las posiciones extremas fijas. Para cuantificar esta observación, notemos que el schema H_1 tiene una longitud definida de 5. Si el sitio de corte es elegido uniformemente al azar entre los $l - 1 = 7 - 1 = 6$ sitios probables, entonces claramente el schema H_1 será destruido con probabilidad $p_d = \delta(H_1)/(l - 1) = 5/6$ (o probabilidad de sobrevivencia $p_s = 1 - p_d =$

1/6). Similarmente, el schema H_2 tiene longitud definida $\delta(H_2) = 1$ y será destruido durante ese único evento en 6 cuando el sitio de corte sea elegido que ocurra entre las posiciones 4 y 5 tal que $p_d = 1/6$ o que la probabilidad de sobrevivencia sea $p_s = 1 - p_d = 5/6$. Más generalmente, vemos que un límite inferior sobre la probabilidad de sobrevivencia al entrecruzamiento p_s puede ser calculada para cualquier schema. Debido a que un schema sobrevivirá cuando el sitio de cruce cae fuera de la longitud definida, la probabilidad de sobrevivencia bajo el entrecruzamiento es $p_s = 1 - \delta(H)/(l - 1)$, ya que el schema será destruido siempre que un sitio dentro de la longitud definida sea el elegido de los $l - 1$ sitios posibles. Si el entrecruzamiento es desarrollado por elección al azar, digamos con probabilidad p_c en una particular unión, la probabilidad de sobrevivencia estará dada por la expresión:

$$p_s \geq 1 - p_c \frac{\delta(H)}{l - 1}$$

la cual se reduce a la anterior expresión cuando $p_c = 1.0$. El efecto combinado de reproducción y entrecruzamiento ahora lo podremos considerar. Como cuando tratamos únicamente la reproducción, estamos interesados en calcular el número de individuos de un schema particular H que se espera en la siguiente generación. Asumiendo la independencia de las operaciones de reproducción y entrecruzamiento, podemos estimar que:

$$m(H, t + 1) \geq m(H, t) \frac{f(H)}{f} \left[1 - p_c \frac{\delta(H)}{l - 1} \right]$$

Comparando esta con la previa expresión para reproducción solamente, el efecto combinado de entrecruzamiento y reproducción es obtenido por multiplicar el número esperado de schemata por reproducción únicamente por la probabilidad de sobrevivencia p_s bajo entrecruzamiento.

Una vez mas el efecto de las operaciones es claro, el schema H aumenta o disminuye dependiendo de un factor multiplicativo. Con ambos, el entrecruzamiento y la reproducción, ese factor depende de dos cosas: si el schema está arriba o abajo del promedio de la población y si el schema tiene una longitud definida relativamente corta o larga. Claramente, aquellos schemata

con ambas cosas: desarrollo observado arriba del promedio y corta longitud definida, van a ser escogido con una relación de incremento exponencial.

3.11 efecto de las mutaciones

El último operador que falta considerar es la mutación. Usando nuestra definición previa, la mutación es una alteración al azar de una simple posición con probabilidad p_m . Para que un schema H sobreviva a la mutación, todas las posiciones específicas deben en sí mismo sobrevivir. Por lo tanto, ya que el valor de una posición simple sobrevive con probabilidad $(1 - p_m)$, y ya que cada una de las mutaciones es estadísticamente independiente, un schema particular sobrevivirá cuando cada una de las $o(H)$ posiciones fijas dentro del schema sobreviva. Multiplicando la probabilidad de sobrevivencia $(1 - p_m)$ por si misma $o(H)$ veces, obtenemos la probabilidad de sobrevivir la mutación, $(1 - p_m)^{o(H)}$. Para pequeños valores de p_m ($p_m \ll 1$), la probabilidad de sobrevivencia del schema puede ser aproximada por la expresión $(1 - o(H)p_m)$. Por lo tanto, concluimos que un particular schema H recibirá un número esperado de copias en la siguiente generación baja la reproducción, entrecruzamiento y mutación dado por la siguiente ecuación (ignorando pequeños términos de productos cruzados):

$$m(H, t + 1) \geq m(H, t) \frac{f(H)}{f} \left[1 - p_c \frac{\delta(H)}{l-1} - o(H)p_m \right]$$

La adición de cambios por mutación cambia nuestras conclusiones previas muy poco: Los schemata arriba del promedio en fineza, bajo orden y longitud corta recibirán una cantidad de individuos que se incrementará exponencialmente en las subsecuentes generaciones. Esta conclusión es tan importante que se le ha dado un particular nombre: el Teorema de Schema o el Teorema Fundamental de los Algoritmos Genéticos.

3.12 Representación Gray

Hay dos formas de representar los enteros como cadenas binarias. La primera usa el código binario estándar, mientras que la segunda usa el código Gray;

cuando variables continuas son usadas, un método común para representarlas es usar su representación entera. Cada variable primero es linealmente mapeada a un entero definido en un rango específico, y el entero es entonces codificado usando un número fijo de bits binarios. La principal desventaja de codificar las variables como cadenas binarias son las grandes distancias Hamming entre los códigos binarios de enteros adyacentes (la distancia Hamming es el número de bits que son necesarios cambiar para que un número sea igual a otro). Así, por ejemplo, 011 y 100 son las representaciones enteras de 3 y 4 respectivamente (tabla 2), y tienen una distancia Hamming de 3. Para que el algoritmo genético mejore el código de 3 a el de 4, este debe alterar todos los bits simultáneamente, tal situación representa un problema para el funcionamiento de los algoritmos genéticos. Para sobrepasar este problema, un código Gray puede ser usado. Los códigos Gray tienen la propiedad de que donde quiera que se incremente o decremente un número por 1 es siempre por un cambio de un bit (tabla 2). Por lo tanto, enteros adyacentes siempre presentan una distancia Hamming de 1.

Enteros	Cód. Binario	Cód. Gray
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111

Tabla 3.2. Comparación de enteros codificados en binario y Gray.

Este mapeo no es único, pero el código Gray, $g(i)$, más a menudo usado en métodos genéticos puede ser obtenido de la representación binaria estándar, $b(i)$, por numerar las posiciones de bit de 1 a k , donde k es la longitud de la cadena de bits, con $b(1)$ siendo el bit de orden más alto y que usa las

siguientes reglas:

```
g(1) := b(1);  
for i = 2 through k  
g(i) := b(i) if b(i-1) = 0;  
g(i) := COMPLEMENT(b(i)) if b(i-1) = 1;
```

Similarmente, un número en código Gray puede ser convertido a su representación estándar por

```
b(1) := g(1);  
for i = 2 through k  
b(i) := g(i) if b(i-1) = 0;  
b(i) := COMPLEMENT(g(i)) if b(i-1) = 1;
```

4 Capítulo 4

4.1 La búsqueda global

En la actualidad, a pesar del avance en las técnicas de optimización y de búsqueda, el tratar de encontrar las estructuras de baja energía estables es todavía un problema difícil. Aún con los sorprendentes márgenes de eficiencia que han presentado la maduración de técnicas de dinámica molecular con cálculos *ab initio* (de primeros principios), las optimizaciones globales llevadas a cabo con este método son muy costosas computacionalmente debido a la gran cantidad de grados de libertad. Las aproximaciones que emplean potenciales clásicos con parámetros ajustados a datos experimentales o resultados *ab initio* seleccionados son una alternativa efectiva. Tales potenciales a menudo son la puerta de entrada a búsquedas más refinadas, ya que los resultados obtenidos con la optimización clásica pueden ser usados como una entrada configuracional hacia un completo cálculo *ab initio* o funcional de la densidad.

En esta investigación empleamos un algoritmo genético parecido al descrito en el capítulo 3 con el potencial del campo de fuerza descrito en el capítulo 2 como técnica de búsqueda global. El algoritmo genético, originalmente propuesto por Holland [6], está basado en la técnica desarrollada por la naturaleza para resolver problemas al hacer evolucionar un código genético microscópico a través de mutación y entrecruzamiento, con la selección basada en la fineza de los correspondientes individuos. Esta es una aproximación inteligente y eficiente para optimizaciones globales multi-variables y ha sido aplicada exitosamente a un gran número de problemas complejos de la física. McGarrah y Judson [7] fueron los primeros en aplicar el algoritmo genético para tratar de determinar la conformación molecular de la molécula cíclica hexaglicina. Judson *et al* [8] han observado la efectividad de los algoritmos genéticos en encontrar el estado base de un polímero plano comparada con las aproximaciones de búsqueda global simulated annealing y al azar. Ellos llegaron a la conclusión general de que mientras más grande

sea la molécula. mayor es la efectividad de los algoritmos genéticos sobre las otras aproximaciones.

El comienzo de la búsqueda global genética se llevo a cabo de la siguiente manera: primero generamos una población al azar de distintas configuraciones de los 14 átomos de la adenina (5 nitrógenos, 5 carbonos y 4 hidrógenos). Estas son generadas en posiciones distribuidas al azar dentro de una esfera de radio $r = 5$ angstroms centrada en el origen de un sistema de coordenadas cartesianas tridimensional. Este radio es un poco mayor al de una esfera que encerraría por completo la conocida estructura "natural" de la adenina. A cada configuración de la población le corresponde una cadena genética, esta última se formó al juntar las 3 coordenadas espaciales de los 14 átomos, siguiendo un orden preestablecido para los átomos. Un código Gray [9] binario de 8 bits para las coordenadas fué escogido.

En una cadena hay $3 \cdot 14 \cdot 8 = 336$ bits que corresponden a las coordenadas espaciales de una configuración dada. Estas cadenas se hacen evolucionar de acuerdo a las técnicas estándar de los algoritmos genéticos empleando mutación y entrecruzamiento, con la selección basada en los resultados de una función de fineza para la molécula [10], la cual por supuesto se tomó como la energía de interacción entre los átomos dada en el capítulo 2, (así, las cadenas escogidas para formar parte de la siguiente generación son las que tengan más baja energía). La evolución de las cadenas prosigió hasta que la energía más baja de de la molécula no hubiese cambiado durante 7 generaciones consecutivas. Es entonces cuando la configuración ha caído en una cuenca de atracción (o mínimo local de energía) de la superficie del potencial.

Para optimizar los detalles de la configuración así encontrada, una optimización utilizando gradiente conjugado local se llevo a cabo. La iteración del procedimiento así descrito nos permitió una precisión de la energía de configuración de $\Delta V/V = 10^{-8}$. Finalmente, para checar que la configuración no estuviera en un punto de silla y que fuera realmente una estructura estable, se obtuvo y diagonalizó el Hessian a una alta precisión numérica y

se descartó aquellas estructuras que presentaron eigenvalores negativos.

4.2 Relajación con cálculos de primeros principios

Tres distintos cálculos cuánticos: PM3, Hartree-Fock y funcionales de la densidad, fueron empleados en una optimización local de las configuraciones interesantes encontradas con el potencial. El método MNDO-PM3 utiliza un semiempírico, orbitales moleculares de campo auto-consistentes (SCF-MO) H Hamiltoniano parametrizado [11] usando los calores estándar de formación de un gran conjunto de moléculas de referencia [12] Está diseñado para reproducir los calores estándar de formación de las energías totales (con la inclusión de calores de atomización precisos) para geometrías moleculares correspondientes al valor mínimo de SCF. Nosotros utilizamos un conjunto base mínimo del tipo de orbitales de Slater (STO) para describir los electrones de valencia en la aproximación de coraza cerrada. El método da resultados confiables para los orbitales moleculares y afinidades electrónicas. Los errores debidos a las simplificaciones relacionadas con la aproximación MNDO son parcialmente compensados por el ajuste a datos experimentales precisos de un gran conjunto de moléculas de referencia en sus estados base. Los cálculos MO y gradiente de energía fueron desarrollados con el programa MOPAC [13].

El método Eigenvector fué el siguiente método escogido de búsqueda para el mínimo [14][15].

En el segundo conjunto de cálculos cuánticos, optimizaciones de orbital molecular ab-initio fueron llevadas a cabo con el campo auto-consistente Hartree-Fock (esquema computacional de combinación lineal de orbitales atómicos HF-LCAO) implementado con el programa Gaussian 94 [16]. La diagonalización del Hamiltoniano fué desarrollada usando como conjunto base un 6-31G*. Para la optimización de las coordenadas nucleares un algoritmo de Berny [17] fué incorporado sin imponer unas restricciones de simetría sobre la geometría. Un criterio de convergencia en el cual a las fuerzas se les impuso que fueran menor que 10^{-4} au fué usado para detener la optimización

de la geometría.

En el tercer conjunto de cálculos cuánticos usamos teoría completa de funcional de la densidad auto-consistente (DFT), los cálculos fueron desarrollados para resolver las ecuaciones auto-consistentes estándar Kohn-Sham en la aproximación de densidad local (LDA) [18]. Estos cálculos fueron desarrollados utilizando el código SIESTA [19][20]. Los electrones de la coraza han sido eliminados y reemplazados por los estándar potenciales conservativos de norma Troullier-Martins en su forma completa no-local (Kleinman-Bylander) [21][22]. Combinaciones lineales flexibles de potenciales atómicos (pseudo) numéricos (PAO) fueron usados como el conjunto base, permitidos para orbitales de polarización dobles. Para limitar el rango de la orbitales base pseudoatómicos, ellos son ligeramente excitados por un cambio de energía común ($\delta E_{PAO} = 0.02$), y truncados en el nodo radial resultante [23]. Las funciones base y la densidad electrónica son proyectadas hacia una red de espacio-real uniforme como medidas para calcular los potenciales de correlación de intercambio, los Hartree y los elementos de matriz. La fineza de la red es controlada por la "energía de corte" ($E_{corte} = 120Ry$) de las ondas planas que pueden ser representadas en esta sin alizamiento. La relajación estructural fué desarrollada utilizando un método de gradiente conjugado no construido con las fuerzas DFT-LDA descritas arriba con un criterio de convergencia de 10^{-4} au.

Para comprobar si efectivamente son estructuras estables, se calculó numéricamente y diagonalizó la matriz Hessiana de segundas derivadas. Solamente aquellas configuraciones que tuvieran todos los eigenvalores positivos (excepto los 6 eigenvalores iguales a cero correspondientes a los grados de libertad de traslación y rotación) fueron considerados interesantes.

4.3 Resultados

En 13000 corridas del algoritmo genético con la función del potencial, comenzando de configuraciones generadas completamente al azar, se encontraron más de 100 estructuras estacionarias (mínimos o puntos de silla). En la tabla

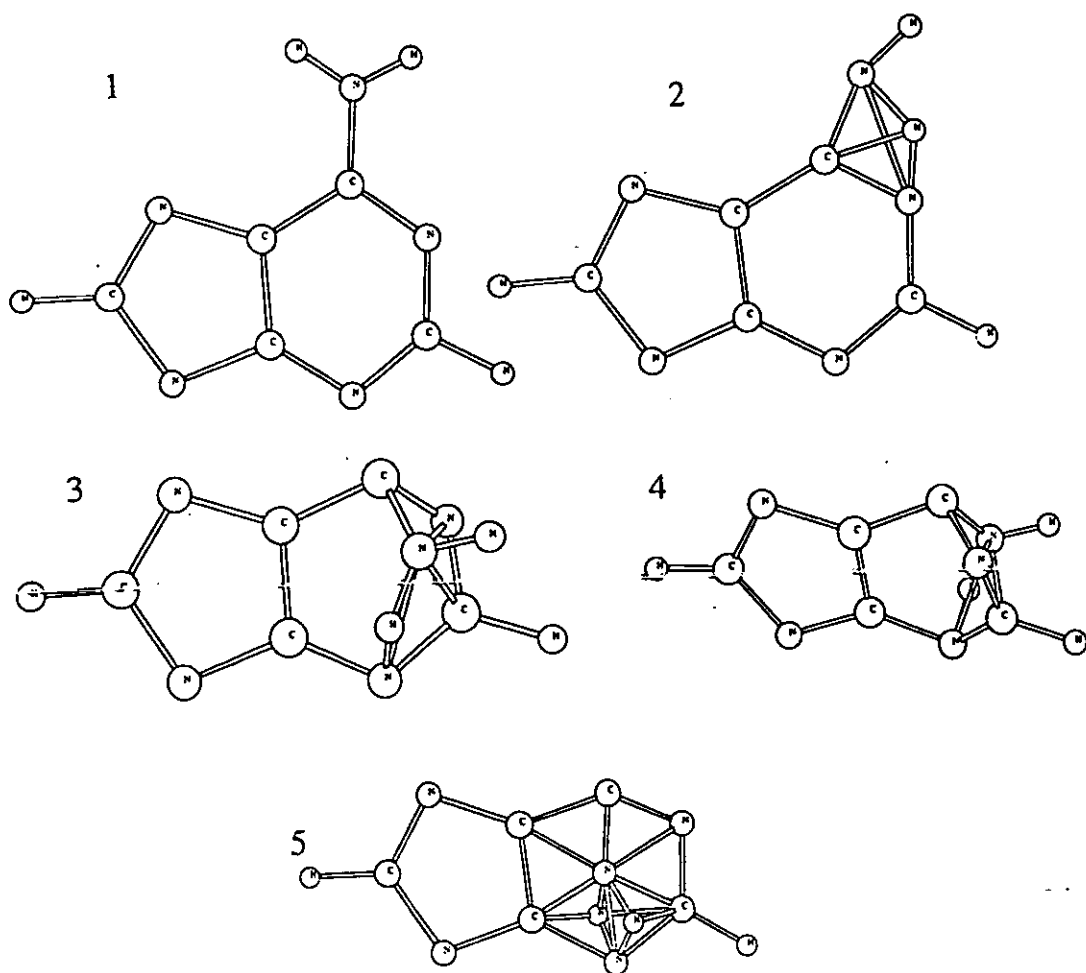
siguiente 4.1 se muestran las energías y el número de veces que se encontraron para las 34 configuraciones de más baja energía. El número de veces que se encontró una estructura mínima local o punto de silla es proporcional, en primera aproximación, al tamaño de su cuenca de atracción.

La configuración con energía mas baja encontrada correspondió a la bien conocida configuración natural, y esta configuración fué la que se encontró un mayor número de veces, lo cual significa que la conformación natural tiene la cuenca de atracción más grande y más profunda en la superficie de potencial.

Energía Kcal/mol	número de veces encontrados
-550.8708	9365
-491.5054	9
-488.2690	76
-487.2949	8
-485.1609	19
-484.5831	18
-481.8493	7
-473.4648	1
-465.1080	1
-461.5263	103
-459.7043	94
-456.3002	71
-454.1580	73
-441.7946	152
-435.6713	705
-433.9844	722
-426.3164	1
-424.7625	1
-418.7526	7
-417.3381	1
-416.4531	3
-415.4279	1
-414.6390	1
-410.0581	1
-408.0559	1
-407.5979	2
-405.8675	2
-402.1395	1
-398.5010	1
-394.7580	1
-392.6488	1
-389.5229	16
-388.6041	30
-384.6376	12

Tabla 4.1. Las energías y número de veces encontrados de los 34 más bajos puntos estacionarios encontrados en 13000 optimizaciones globales de configuraciones iniciales al azar de los átomos de la adenina, utilizando el potencial de campo de fuerza con el algoritmo genético de búsqueda.

Figura 4.1: Las geometrías de las más interesantes (ver texto) configuraciones estables de más baja energía encontradas para los átomos de la adenina en 13,000 optimizaciones globales comenzando de configuraciones al azar. Estos resultados fueron obtenidos con el potencial de campo de fuerza descrito en el capítulo 2.



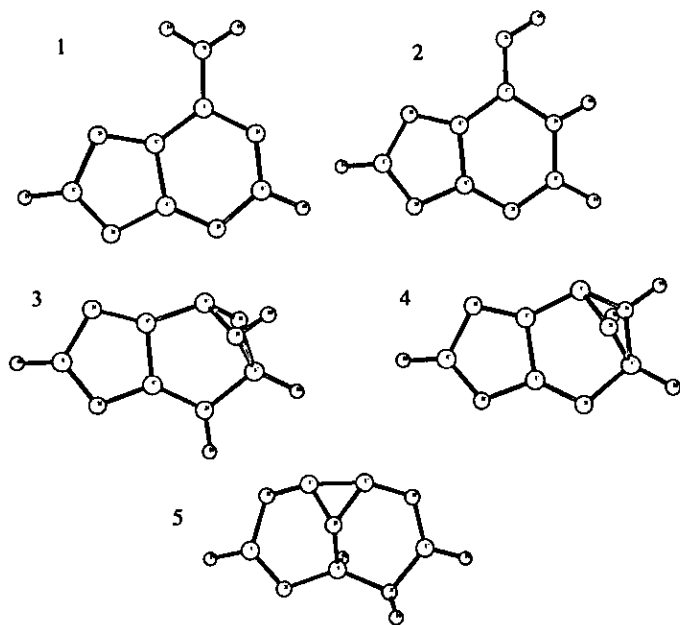


Figura 4.2: Las geometrias de la figura 4.1 despues de optimizacion local utilizando calculos semiempiricos de PM3 .

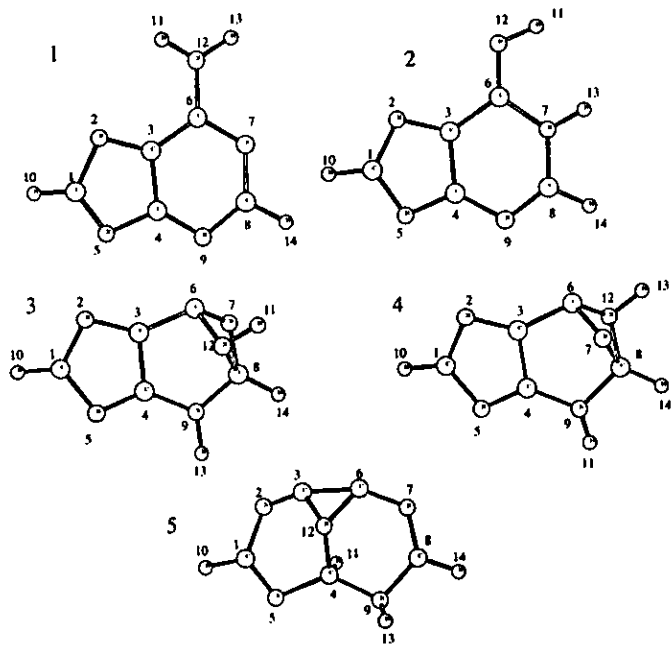


Fig. 4.3 : Las geometrias de la figura 4.1 despues de optimizacion local utilizando calculos Hartree-Fock.

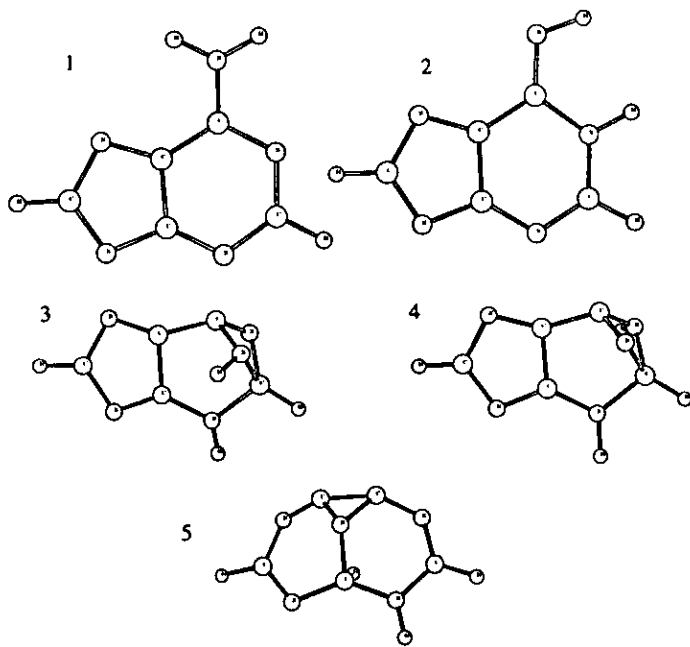


Fig. 4.4: Las geometrias de la figura 4.1 despues de optimizacion local utilizando calculos de funcionales de la densidad.

En la tabla 4.2 damos las brechas de energía de las 10 primeras configuraciones de la tabla 4.1, las unidades de esta tabla son kcal/mol y eV/molécula, estos valores se obtuvieron de tomar en cuenta el siguiente factor de conversión: 1 kcal/mol = 0.0434 eV/molécula.

$E_i - E_j$	kcal/mol	eV/molécula
$E_1 - E_2$	59.3654	2.5767
$E_2 - E_3$	3.2365	0.1404
$E_3 - E_4$	0.9741	0.0422
$E_4 - E_5$	2.1340	0.0926
$E_5 - E_6$	0.5778	0.0251
$E_6 - E_7$	2.7338	0.1186
$E_7 - E_8$	8.3845	0.3628
$E_8 - E_9$	8.3568	0.3627
$E_9 - E_{10}$	3.5817	0.1554

Tabla 4.2. Brechas de energía de las 10 configuraciones con más baja energía.

Es aquí necesario el recalcar que el potencial fué ajustado utilizando datos experimentales obtenidos de las configuraciones "naturales", si bien la diversidad en número y tipos de compuestos fué grande para su formulación, es posible que este potencial sea adecuado para conformaciones naturales o aquellas que se asemejen a estas. Las configuraciones que se alejen mucho de las naturales es probable que no sean bien descritas por el potencial. Esto limita a que nos concentraremos en aquellas conformaciones interesantes que no se alejen demasiado de la natural. En la figura 4.1 mostramos las más interesantes de estas.

Debido a un artificio del potencial modelo, casi la mayoría de las estructuras tenían su gemelo que resulta del intercambio de los hidrógenos en el enlace H-N-H, esto a consecuencia de los distintos valores de las cargas asignadas a los dos hidrógenos y a su no optimización en el proceso de encontrar los mínimos.

Una comparación de las diferencias de energía total entre la configuración "natural" de energía más baja y cada isomero, para aquellos isómeros listados en las figuras 1 - 4 está dada en la tabla 4.3. Note que el ordenamiento de energía de los isómeros es el mismo para los cuatro cálculos y que los tres cálculos cuánticos están en acuerdo cuantitativo razonable para las diferencias de energía de todos los isómeros.

isomero	energía (kcal/mol)	dif. de energía (kcal/mol)	dif. de energía (kcal/mol)	dif. de energía (kcal/mol)	dif. de energía (kcal/mol)
	Pot. mod.	Pot. mod.	PM3	HF	DF
1	-550.871	0.0	0.0	0.0	0.0
2	-485.161	+65.71	+12.94	+6.97	+17.75
3	-435.671	+115.20	+93.73	+98.13	+96.41
4	-433.984	+116.89	+99.43	+98.14	+96.54
5	-389.523	+161.35	+130.32	+142.07	+127.34

tabla 4.3. Comparación de las diferencias en la energía total entre la configuración natural de la adenina (isomero 1) y los isómeros de más alta energía listados en las figs. 1-4 para el potencial modelo, PM3, Hartree-Fock y cálculos de funcional de la densidad.

4.4 Conclusiones

Hemos realizado una búsqueda global de los isómeros estables de la adenina utilizando un algoritmo genético junto con un potencial de campo de fuerza. Con esta búsqueda un gran número de isómeros estables fueron encontrados, tanto planares como no planares. Aquellas configuraciones similares a la configuración natural y estables resultaron interesantes y fueron posteriormente estudiadas con aproximaciones PM3, Hartree-Fock y funcionales de la densidad. El hecho de que los mínimos encontrados con el potencial fueran cercanamente similares a aquellos encontrados con cálculos cuánticos nos da confianza en la validez de este potencial para uso como primera aproximación

para una búsqueda mas global de isomeros.

Debido a que los isomeros encontrados aqui han sido confirmados con cuatro cálculos teóricos independientes y se verificó que son estables tanto a un nivel clásico como cuántico, concluimos que ellos pueden ser observados en la naturaleza. Estos resultados podrían tener implicaciones en la actividad biológica de la adenina y en las características nanoelectrónicas de la adenina atada a nanocúmulos de metal, esto nos parece interesante para posteriores estudios mas detallados.

Los resultados presentados en esta tesis fueron aceptados para publicarse en una revista internacional [24]. Las direcciones para futuras investigaciones incluyen cálculos las energías libres de los isomeros a temperatura ambiente, determinación de la barrera de potencial entre los isomeros, incluyendo el par de base timina y el esqueleto de ADN, aplicar la misma búsqueda global a las otras bases y pares de bases, estudiar el enlace de los isomeros a nanocúmulos de metal e investigar las implicaciones biológicas de los isomeros.

References

- [1] A. Warshel et. al. The Journal of Chemical Physics, Vol. 49, núm. 11, 1968, pág. 5116.
- [2] A. Warshel et. al. The Journal of Chemical Physics, Vol. 53, núm 2, 1970, pág. 582.
- [3] Hagler et. al. Journal of the American Chemical Society, Vol. 96, núm 17, 1974, pág 5319.
- [4] Weiner et. al. Journal of the American Chemical Society, Vol. 106, 1984, pág. 765.
- [5] Nguyen et. al. Journal of Computational Chemistry, Vol 7, núm 2, 1986, pág 230.
- [6] J. Holland, Adaptation in Natural and Artificial Systems (University of Michigan Press, Ann Arbor, 1975).
- [7] D. B. McGarrah and R.S. Judson, J. Comp. Chem. vol. 14, (1993) 1385.
- [8] R. S. Judson, M.E. Colvin, J.C. Meza, A. Huffer, D. Gutiérrez, Int. J. Quan. Chem. 44 (1992) 277.
- [9] G. Chang, M. Lewis, Acta Cryst. D50 (1994) 667.
- [10] K. Michaelian, Am. J. Phys. 66 (1998) 231.
- [11] M.J.S. Dewar, W. Thiel, J. Am. Chem Soc. 99 (1997) 4907.
- [12] J.J.P. Stewart, Comput. Chem. 10 (1989) 209.
- [13] J.J.P. Stewart, MOPAC, v.6. Disponible para PC por L. A. Montero en el Laboratorio de Química Teórica y Computacional, Universidad de La Habana, 1993-1997.
- [14] P. Culot, G. Dive, V.N. Nguyen, J.M. Ghuysen. Theor. Chim. Acta 82 (1992) 189.

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

- [15] J. Baker. *J. Comput. Chem.* 7 (1986) 385.
- [16] GAUSSIAN 94, Gaussian. Inc., Pittsburgh PA, 1995.
- [17] M.J. Frisch, A. Frisch. J.B. Forersman, *Gaussian 94 User's Reference*, Gaussian Inc., Pittsburgh. 1996. p. 111.
- [18] W. Kohn. L.J. Sham, *Phys. Rev.* 145 (1965) 561.
- [19] P. Ordejón, E. Artacho, J.M. Soler, *Phys. Rev. B.* 53 (1996) 10 441.
- [20] D. Sánchez-Portal, et al.. *Int. J. Quantum Chem.* 65 (1997) 453.
- [21] N. Troullier, J.L. Martins, *Phys. Rev. B.* 43 (1991) 1993.
- [22] L. Kleinman, D.M. Bylander, *Phys. Rev. Lett.* 48 (1982) 1425.
- [23] O.F. Sankey, D.J. Niklewski, *Phys. Rev. B.* 40 (1989) 3979.
- [24] J. Vega. K. Michaelian, I.L. Garzón, M. R. Beltrán, L. Hernández, aceptado para publicar en la revista *Journal of Molecular Structure (Theochem)*, julio 1999.

5 Apéndice: cálculo de los elementos de matriz

Z_{lm}

Los renglones de la matriz $Z_{lm} = \partial y_l / \partial x_m$ son de tres clases, $\partial \Delta H^{(n)} / \partial x_m$, $\partial s_0^{(n)} / \partial x_m$, y $\partial \nu_\alpha^{(n)} / \partial x_m$. Ellos son calculados como sigue:

(1) La derivada del exceso de entalpía es, por la ecuación (1.12),

$$\partial \Delta H^{(n)} / \partial x_m = (\partial \Delta V^{(n)} / \partial x_m) + (\partial \Delta H_{vrt}^{(n)} / \partial x_m)$$

El primer término se deriva directamente de la energía potencial como función explícita de \mathbf{x} . El segundo término se ha encontrado que es relativamente insensible a cambios en \mathbf{x} y es por lo tanto despreciable. Notemos, sin embargo, que ΔH_{vrt} sigue siendo parte del exceso de entalpía calculada en los componentes de el vector $\Delta \mathbf{y}$.

(2) Las derivadas de las coordenadas de equilibrio con respecto a x_m son primero obtenidas en coordenadas cartesianas, esto es, $\partial \mathbf{r}_0 / \partial x_m$, y entonces transformadas a coordenadas internas por la matriz de transformación B,

$$\delta \mathbf{x} = B \delta \mathbf{r}, \quad B_{\alpha\beta} = \partial s_\alpha / \partial r_\beta$$

entonces

$$\partial s_0 / \partial x_m = B \partial \mathbf{r}_0 / \partial x_m$$

Las variables de equilibrio \mathbf{r}_0 no están dadas como funciones explícitas de \mathbf{x} , si no que están dadas por la ecuación $\partial V(\mathbf{r} : \mathbf{x}) / \partial r_\alpha = 0$. Para obtener $\partial \mathbf{r}_0 / \partial x_m$ consideremos su definición

$$\frac{\partial \mathbf{r}_0}{\partial x_m} = \lim_{\delta x_m \rightarrow 0} \frac{\mathbf{r}_0(\mathbf{x} + \delta \mathbf{x}_m) - \mathbf{r}_0(\mathbf{x})}{\delta x_m}$$

donde $\delta \mathbf{x}_m$ es un vector $\delta \mathbf{x}$ con δx_m como la única componente distinta de cero, $\mathbf{r}_0(\mathbf{x} + \delta \mathbf{x}_m)$ son las coordenadas de equilibrio pertenecientes a $\mathbf{x} + \delta \mathbf{x}_m$, y $\mathbf{r}_0(\mathbf{x})$ son las coordenadas de equilibrio pertenecientes a \mathbf{x} . Ahora consideremos la solución de $\partial V(\mathbf{r}, \mathbf{x} + \delta \mathbf{x}_m) / \partial r_\alpha = 0$ por la ec. (1.13),

$$\mathbf{r}_0(\mathbf{x} + \delta \mathbf{x}_m) = \mathbf{r}(\mathbf{x} + \delta \mathbf{x}_m) - \mathbf{F}^{-1}(\mathbf{r}, \mathbf{x} + \delta \mathbf{x}_m) \nabla V(\mathbf{r}, \mathbf{x} + \delta \mathbf{x}_m)$$

Si escogemos $\mathbf{r}(\mathbf{x} + \delta\mathbf{x}_m)$ de tal forma que

$$\mathbf{r}(\mathbf{x} + \delta\mathbf{x}_m) = \mathbf{r}_0(\mathbf{x})$$

De las dos ecuaciones anteriores se sigue que:

$$\frac{\partial \mathbf{r}_0}{\partial \mathbf{x}_m} = \lim_{\delta \mathbf{x}_m \rightarrow 0} \frac{-\mathbf{F}^{-1}[\mathbf{r}_0(\mathbf{x}); \mathbf{x} + \delta\mathbf{x}_m] \nabla V[\mathbf{r}_0(\mathbf{x}); \mathbf{x} + \delta\mathbf{x}_m]}{\delta \mathbf{x}_m}$$

este límite es igual a

$$-\mathbf{F}^{-1}(\mathbf{r}_0; \mathbf{x}) \partial \nabla V(\mathbf{r}_0; \mathbf{x}) / \partial \mathbf{x}_m$$

Las expresiones \mathbf{F}^{-1} y \mathbf{F} han sido ya usadas en la derivación de \mathbf{r}_0 y ν_α , y $\partial \nabla V / \partial x_m$ se deriva de las expresiones analíticas de el gradiente ∇V como función explícita de \mathbf{x}

(3) Considerando las frecuencias de modos normales como funciones de las constantes de fuerza y usando el método de mínimos cuadrados para obtener los valores óptimos de las constantes de fuerza. De la ecuación (1.8) se

obtiene que:

$$\frac{\partial \lambda_\alpha}{\partial x_m} = \delta \mathbf{q}_\alpha^T \left(\frac{\partial f}{\partial x_m} \right) \delta \mathbf{q}_\alpha$$

donde f no depende de las coordenadas de equilibrio. En el presente caso, sin embargo, f es una función no lineal tanto de los parámetros de energía \mathbf{x} como de las coordenadas de equilibrio \mathbf{r}_0 , las cuales son funciones no lineales de \mathbf{x} :

$$\frac{\partial \lambda_\alpha}{\partial x_m} = \delta \mathbf{q}_\alpha^T \left[\left(\frac{\partial f}{\partial x_m} \right)_{\mathbf{r}=\mathbf{r}_0} + \sum_i \frac{\partial f}{\partial r_i} \frac{\partial r_i}{\partial x_m} \right] \delta \mathbf{q}_\alpha$$

La evaluación de la tercera derivada $\partial^3 / \partial r_\alpha \partial r_\beta \partial r_\gamma$, en esta expresión ocuparía mucho tiempo de computo. En vez de esta, se uso la siguiente expresión

$$\frac{\partial \lambda_\alpha}{\partial x_m} = \delta \mathbf{q}_\alpha^T \left(\lim_{x_m \rightarrow 0} \frac{(f[\mathbf{x} + \delta\mathbf{x}_m; \mathbf{r}_0(\mathbf{x} + \delta\mathbf{x}_m)] - f(\mathbf{x}, \mathbf{r}_0))}{(\delta x_m)} \right) \delta \mathbf{q}_\alpha,$$

donde $\mathbf{r}_0(\mathbf{x} + \delta\mathbf{x}_m)$ es obtenida de las ecuaciones (1.68)-(1.71), y la ecuación última fué evaluada numericamente usando la expresión analítica para f .