

# Refinamiento de la metodología de la simulación de la hidratación de los ácidos nucleicos

E. González,<sup>1</sup> F.I. Cedeño,<sup>2</sup> A.V. Teplukhin,<sup>3</sup> G.G. Malenkov,<sup>4</sup> y V.I. Poltev<sup>1,5</sup>

<sup>1</sup>*Colegio de Física, Facultad de Ciencias Físico Matemáticas, Universidad Autónoma de Puebla  
Apartado postal 1152, 72000 Puebla, Pue., Mexico*

<sup>2</sup>*Escuela de Biología, Universidad Autónoma de Puebla  
72000 Puebla, Pue., Mexico*

<sup>3</sup>*Instituto Matemático de Problemas Biológicos, Academia de Ciencias de Rusia  
142292 Pushchino, Región de Moscú, Rusia*

<sup>4</sup>*Instituto de Físico Química, Academia de Ciencias de Rusia  
117915 Moscú, Rusia*

<sup>5</sup>*Instituto Teórico y Experimental de Biofísica, Academia de Ciencias de Rusia  
142292 Pushchino, Región de Moscú, Rusia*

Recibido el 27 de abril de 2000; aceptado el 29 de junio de 2000

Presentamos un conjunto de funciones de potencial para la simulación de la hidratación de fragmentos de los ácidos nucleicos. Estos potenciales son un refinamiento de aquellos que fueron propuestos por dos de nosotros en 1984. Los nuevos potenciales conservan la misma forma analítica como los valores anteriores de la carga de los átomos de las moléculas del agua y del ADN. Los nuevos potenciales reproducen satisfactoriamente los datos experimentales de la energía de interacción intermolecular y la estructura del agua líquida, así como la entalpía de formación de los complejos "base del ADN-agua" y las entalpías de hidratación de las bases del ADN. Estos potenciales serán utilizados para el estudio de la contribución de la hidratación en el funcionamiento del ADN.

*Descriptor:* ADN; hidratación; Monte Carlo

Potential functions are presented for simulation of hydration of nucleic acids. These potentials are a refinement of those suggested by two of us in 1984. The refined potentials retain the same simple analytical form (1-6-12 and 1-10-12 potentials) and the values for atomic charges of water and DNA molecules. The new potentials reproduce well the experimental data of the intermolecular interaction energy and the structure of liquid water, as well as those of the enthalpies of water-base complex formation and the enthalpies of DNA base hydration. These potentials will be used for our study on the hydration contribution into DNA functioning.

*Keywords:* ADN; hydration; Monte Carlo

PACS: 87.15.K; 35.20.G; 02.50.N

## 1. Introducción

Las interacciones de las moléculas de agua con el ADN son importantes para determinar su estructura espacial (es necesario un determinado número de moléculas de agua para formar la doble hélice [1]), así como para el funcionamiento biológico del ADN como portador de la información genética (las moléculas de agua son capaces de formar algunos pares de bases incorrectos o estorbar la formación de otros, estos detalles se discuten en la Ref. 2). Hoy en día se tienen muchos datos experimentales [3] que testifican sobre el importante papel del agua respecto al funcionamiento biológico. Esto dio motivo para considerar al agua como uno de los componentes estructurales del ADN. Los datos más directos sobre el papel del agua fueron obtenidos de investigaciones de rayos X en cristales de oligonucleótidos biespirales [4, 5], éstos confirman que el agua es una parte integrante e importante del ADN. Esas investigaciones demostraron la dependencia de la estructura del agua en ciertas regiones cercanas al ADN respecto a la secuencia de bases. La más conocida

es la formación de una doble capa de moléculas de agua o "espinas de hidratación" en el surco menor de la doble espiral en la región llamada AT track [4]. Al estudio de esta "espinas de hidratación" se le han dedicado decenas de experimentos e investigaciones computacionales [6]. A pesar de todas las importantes investigaciones de rayos X, no se ha logrado determinar los detalles geométricos de los alrededores acuosos y una descripción de la red de enlaces de hidrógeno (enlaces-H). Esto se debe a que las investigaciones experimentales no permiten localizar a los átomos de hidrógeno y la resolución en la localización de los átomos de oxígeno no es alta. Una detallada descripción de los alrededores acuosos del ADN puede ser construida solamente con los resultados de los experimentos computacionales, y utilizando funciones de potencial clásicas para determinar la energía de interacción entre las componentes del sistema "ADN-agua". En 1984 dos de nosotros propusimos uno de los primeros sistemas de funciones de potencial átomo-átomo para la simulación de la estructura y propiedades de la capa de hidratación del ADN [7]. Estos potenciales en la actualidad son utiliza-

dos para la simulación de la hidratación de fragmentos de ADN, permitiendo hacer conclusiones interesantes (ver, por ejemplo, Refs. 2 y 8). Después de haber realizado una gran cantidad de trabajos de simulación sobre la hidratación del ADN, aparecieron nuevos datos experimentales y aumentó la capacidad y la rapidez de cálculo de las modernas computadoras, dando lugar a la necesidad de refinar los potenciales, sin alterar su forma analítica simple para proporcionar una mayor concordancia con los datos experimentales.

## 2. Metodología de la simulación de la hidratación del ADN

En este trabajo consideramos sistemas simples, conteniendo una base y una molécula de agua (encontramos los diferentes mínimos locales, que corresponden a la formación de enlaces-H agua-base) y sistemas conteniendo una base y 400 moléculas de agua en una celda elemental con condiciones periódicas a la frontera. Para encontrar las características energéticas y estructurales promedio de la capa de hidratación utilizamos el método de Monte Carlo (algoritmo de Metrópolis). Este procedimiento es análogo al utilizado y descrito antes (ensamble NTV, 300 K, ver por ejemplo Ref. 2 y 8). La geometría de la molécula de agua y de las bases del ADN es la misma que en trabajos anteriores [7, 8]. Durante la selección de los potenciales que describen las interacciones entre las moléculas de agua, realizamos los cálculos en una celda elemental cúbica, conteniendo 256 moléculas de agua. Aquí se usa una técnica especial de truncado, la cual toma en cuenta todas las interacciones entre una molécula y la imagen vecina de las otras moléculas, y se utilizan condiciones periódicas a la frontera impuestas sobre la celda elemental [9]. Para los cálculos del sistema conteniendo una base del ADN con 400 moléculas de agua, utilizamos una celda elemental de forma de un paralelepípedo rectangular. En este caso consideramos a las moléculas de agua tal que a una distancia entre los átomos de oxígeno de las moléculas de agua no sobrepasen los 8 Å. Como demostraron los cálculos para agua pura con esta condición, prácticamente no ocurren cambios en la distribución de las distancias interatómicas. Un pequeño cambio sucede (menos de 0.5 kcal/mol) únicamente en la energía promedio de la interacción interatómica. Este cambio no influye en la energía de hidratación y tampoco en la energía y estructura de la capa de hidratación de la base.

## 3. Interacciones entre moléculas de agua

El ajuste de los parámetros de las funciones de potencial (FP) se inicia con la investigación de las interacciones entre moléculas de agua. En la literatura es posible encontrar bastantes funciones de potencial con diferente grado de complejidad que permiten describir diferentes propiedades del agua líquida y cúmulos de agua. Para una óptima aplicación al estudio de los ácidos nucleicos en disoluciones acuosas y debido a la complejidad de los sistemas tratados, es necesario

TABLA I. Parámetros de los potenciales para el cálculo de la interacción agua-agua.

	<i>A</i>	<i>B</i>	<i>R</i> <sub>0</sub>	$\epsilon$
OO	366	882847	4.11	0.038
OH	7193	18356	1.75	4.45
HH	35	5914	2.64	0.0516

utilizar potenciales para agua de los más simples posibles, aunados a los potenciales que describen las interacción agua-ADN y las interacciones intermoleculares de los ácidos nucleicos. No hicimos el intento de utilizar potenciales agua-agua de otros autores (o sus modificaciones), ya que nuestros potenciales [7] son los que mejor describen las interacciones agua-base [8]. El uso de otros potenciales nos obligaría a elaborar de nuevo los sistemas de potenciales para el cálculo de las otras componentes de energía de los ácidos nucleicos en disolución acuosa.

Ajustamos las funciones de potencial Poltev y Malenkov propuestas en 1984 [7]. Estas funciones de potencial átomo-átomo tienen una forma bastante simple como se muestra en las Ecs (1) y (2).

$$U(r_{ij}) = k \frac{q_i q_j}{r} - A_{ij}^{(10)} r_{ij}^{-10} + B_{ij}^{(10)} r_{ij}^{-12}, \quad (1)$$

$$U(r_{ij}) = k \frac{q_i q_j}{r} - A_{ij} r_{ij}^{-6} + B_{ij} r_{ij}^{-12}. \quad (2)$$

Correspondiendo a las funciones tipo 1-10-12 para las interacciones entre átomos que forman enlaces de hidrógeno [Ec. (1)] y del tipo 1-6-12 para las interacciones restantes [Ec. (2)]. Cada función tiene un término electrostático y componentes de van der Waals,  $r_{ij}$  es la distancia interatómica,  $q_i$  y  $q_j$  son las cargas de los átomos,  $A_{ij}$  y  $B_{ij}$  ( $A_{ij}^{(10)}$  y  $B_{ij}^{(10)}$ ) son coeficientes que dependen del tipo de átomos, de sus estados de valencia (en algunos casos) y de los átomos vecinos.

Estos potenciales reproducen satisfactoriamente los datos experimentales para el agua líquida, aunque originan una desviación sistemática respecto a ellos. Por ello proponemos el refinamiento de los potenciales para salvar esta deficiencia.

Las funciones de potencial tienen la misma forma analítica, se diferencian un poco respecto a los valores de los parámetros *A* y *B*. El ajuste se realiza conservando las cargas atómicas anteriores ( $e = 0.68$  partes de la carga del electrón). Después de realizar varias decenas de ensayos con diferentes conjuntos de parámetros (incluyendo parámetros con cambios en las cargas de los átomos) obtuvimos los valores de *A* y *B* que se muestran en la Tabla I. En esta tabla también se dan los valores de la energía mínima ( $\epsilon$ ) y la distancia de este mínimo ( $R_0$ ). Estos valores se modificaron durante la elección de los coeficientes.

Los parámetros anteriores simulan satisfactoriamente la energética y la estructura del agua líquida. En particular la energía promedio de la interacción intermolecular a 300 K que tiene un valor de 10.08 kcal/mol por par molecular, las

TABLA II. Comparación de los valores de la energía de interacción de moléculas de agua, calculados y obtenidos experimentalmente de cúmulos de agua. Todas las energías están dadas en kcal/mol,  $E_n$  es la energía calculada y  $\Delta H$  la entalpía determinada experimentalmente [11] de la unión de la  $n$ -ésima molécula de agua al cúmulo que contiene  $n - 1$  moléculas.  $E_\Sigma$  es la energía total de la interacción intermolecular en el cúmulo.

$n$	2	3	4	5	6
$E_n$ [en este trabajo]	-6.5	-8.7	-12.0	-8.7	-9.6
$E_n$ [7]	-6.36	-9.9	-11.0	-8.2	-7.5
$E_\Sigma$	-6.5	-15.1	-27.1	-35.8	-45.4
$\Delta H$	-4.2	-10.9	-33.7	-46.9	-55.0

funciones de distribución radial (FDR) calculadas nos dan una buena aproximación a las construidas con los datos de difracción de neutrones (ver abajo). Al igual que otros potenciales para modelar agua líquida, los valores que se obtienen de la energía de cúmulos de agua, poco se diferencian de los valores experimentales (y de los valores calculados con los más sofisticados métodos mecánico cuánticos).

En la Tabla II se presentan los valores de la energía de cúmulos de agua, conteniendo de 2 a 6 moléculas de agua. Las energías calculadas son para la configuración del cúmulo con la mínima energía. Para  $n$  de 3 a 5, la configuración es un polígono regular en el plano, para  $n = 6$  se tiene una configuración de tipo "libro abierto", es decir, dos cuadriláteros regulares con un lado común. La configuración plana regular del hexágono tiene energía en valor absoluto un poco menor (43.7 kcal/mol) como se observa en la Tabla II. Los enlaces de hidrógeno en todos los cúmulos (excepto  $n = 3$ ) son prácticamente lineales, la distancia O...O tiene el valor de  $2.81 \pm 0.01 \text{ \AA}$ .

Las funciones radiales de distribución,  $g_{OO}$ ,  $g_{OH}$  y  $g_{HH}$  del agua líquida, se presentan en la Fig. 1, éstas son obtenidas de resultados de la simulación y de los datos de difracción de neutrones reportados por Soper [10]. Durante la simulación nosotros consideramos para cada par de átomos, intervalos de distancia interatómica del ancho de  $0.01 \text{ \AA}$ . Después calculamos estos datos con el mismo formato que tienen los datos de Soper (el ancho del intervalo para la distancia OO fue de  $0.05 \text{ \AA}$ , para OH y HH de  $0.1 \text{ \AA}$  para ambos) consideramos las distancias hasta  $8.375 \text{ \AA}$ ,  $7.35 \text{ \AA}$  y  $7.35 \text{ \AA}$  respectivamente para las distribuciones OO, OH y HH. Para estimar cuantitativamente la desviación en la FRD respecto de los datos de Soper, calculamos una cantidad igual a la suma en valor absoluto de estas desviaciones para todos los intervalos  $\Delta OO$ ,  $\Delta OH$  y  $\Delta HH$ , estas cantidades corresponden a  $10.4$ ,  $4.2$  y  $2.4 \text{ \AA}$ . Para el potencial de 1984 esta cantidad fue igual a  $12.6$ ,  $3.5$  y  $2.3 \text{ \AA}$ , respectivamente. La mayor cantidad corresponde a  $\Delta OO$ , la cual está relacionada con un mayor número de sumandos (intervalos dos veces más cortos y sumando distancias más largas). Como se ve en la Fig. 1 la posición de los dos primeros máximos y los dos primeros mínimos de las

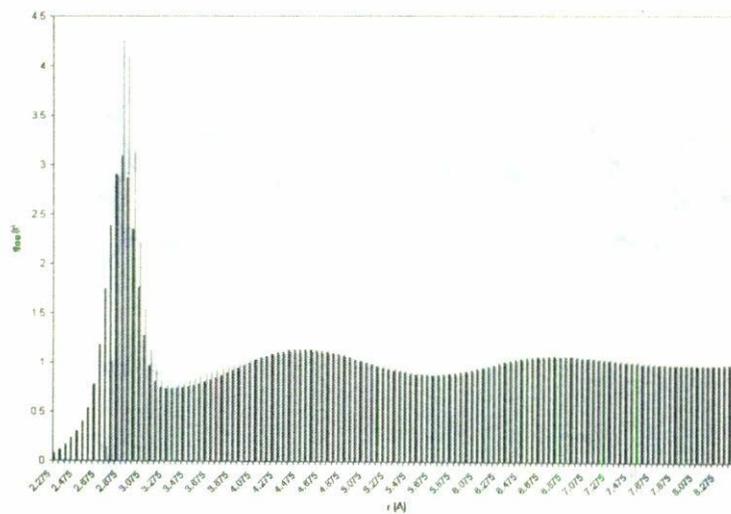
TABLA III. Coeficientes de las funciones de potencial que describen la interacción agua-base.  $H_W$  y  $O_W$  son los átomos H y O de la molécula de agua. C1 y C2 son los átomos de carbono alifáticos y aromáticos respectivamente, N1, N2 y N3 son átomos de nitrógeno del grupo amino, y nitrógenos en el anillo pirrol y piridin. Los valores de los coeficientes no mencionados en la tabla no cambiaron en comparación con los potenciales de 1984 [7].

	$H_W$				$O_W$			
	A	B	$R_0$	$\epsilon$	A	B	$R_0$	$\epsilon$
C1	75	40352	3.2	0.035	195	270672	3.75	0.035
C2	101	19512	2.7	0.13	240	378590	3.83	0.038
N1	-	-	-	-	286	605608	4.02	0.034
N2	-	-	-	-	273	519359	3.95	0.036
N3	9266	27291	1.88	2.8	348	631937	3.92	0.048
O	10667	31084	1.87	3.4	311	433075	3.75	0.056

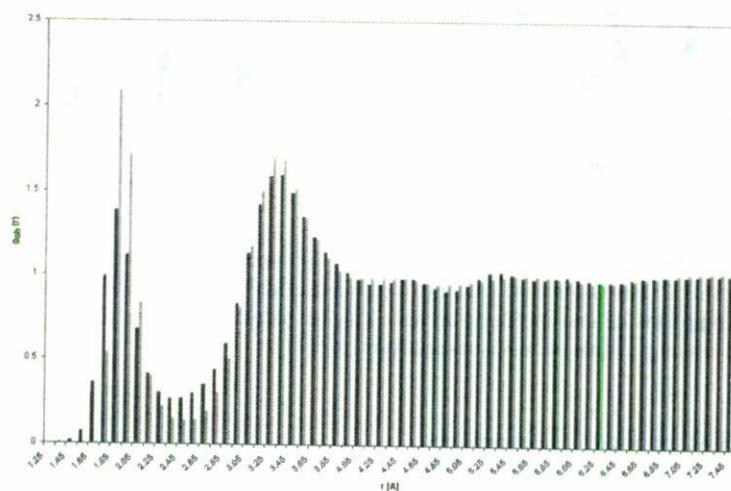
FDR, obtenidos con nuestro potencial, coinciden con aquellos para la FDR de Soper. Una marcada diferencia se ve para  $g_{OO}$  y corresponde a una distancia menor del primer máximo. De los resultados de nuestros cálculos en agua prácticamente no existen distancias OO menores a  $2.7 \text{ \AA}$ . Se observa que de los datos de Soper la distancia OO más grande es menor a  $2.5 \text{ \AA}$ . Una distancia bastante corta en la FDR de Soper nos hace pensar que se trata de un artificio. Es difícil comprender porqué en agua líquida dos átomos pueden encontrarse a distancias más cortas que los contactos interatómicos más cercanos en cristales.

#### 4. Interacción de una molécula de agua con las bases del ADN

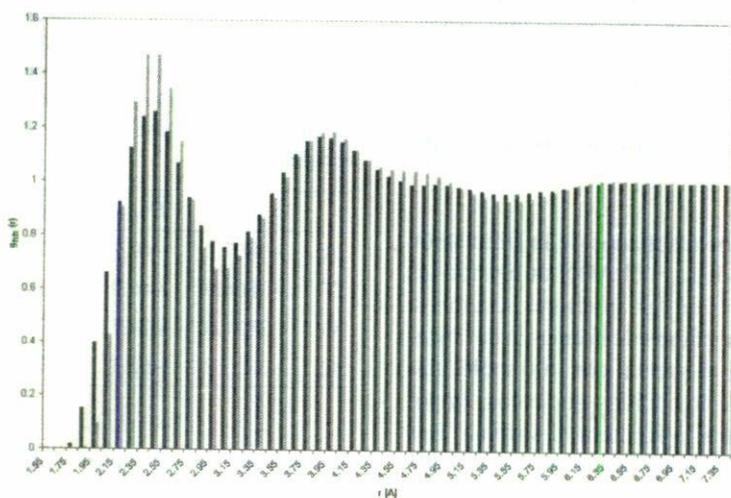
Realizamos los cálculos de la energía de interacción de una molécula de agua con cada una de las bases del ADN en las posiciones de los mínimos locales y cálculos de la energía de hidratación de las bases (ver abajo). Esto se realizó después de que nosotros elegimos un conjunto de coeficientes para las FP, para esta elección se probó con decenas de conjuntos de potenciales, los seleccionados se muestran en la Tabla III. Las posiciones mutuas de las moléculas de agua y bases junto con los valores de la energía en estos mínimos se presentan en la Fig. 2. Los valores de la energía potencial para los más profundos mínimos en cada una de las bases (con exactitud hasta el error experimental) coincide con los datos de espectroscopía de masas [11] y se presentan en los pies de la Fig. 1. Igual que los cálculos con los potenciales de 1984, para cada una de las bases reportan que el más profundo mínimo corresponde cuando la molécula de agua forma dos enlaces-H un poco curvados entre dos átomos hidrofílicos de la base. La longitud de cada uno de los enlaces-H se diferencia significativamente de la longitud encontrada con el potencial de 1984 y está en concordancia con la longitud del enlace-H, agua-ba-



(a)



(b)



(c)

FIGURA 1. Las funciones de distribución radial de las distancias interatómicas en agua líquida; (a) para OO, (b) para OH y (c) para HH, las barras negras corresponden a los datos experimentales [10] y las barras grises a los obtenidos en este trabajo.

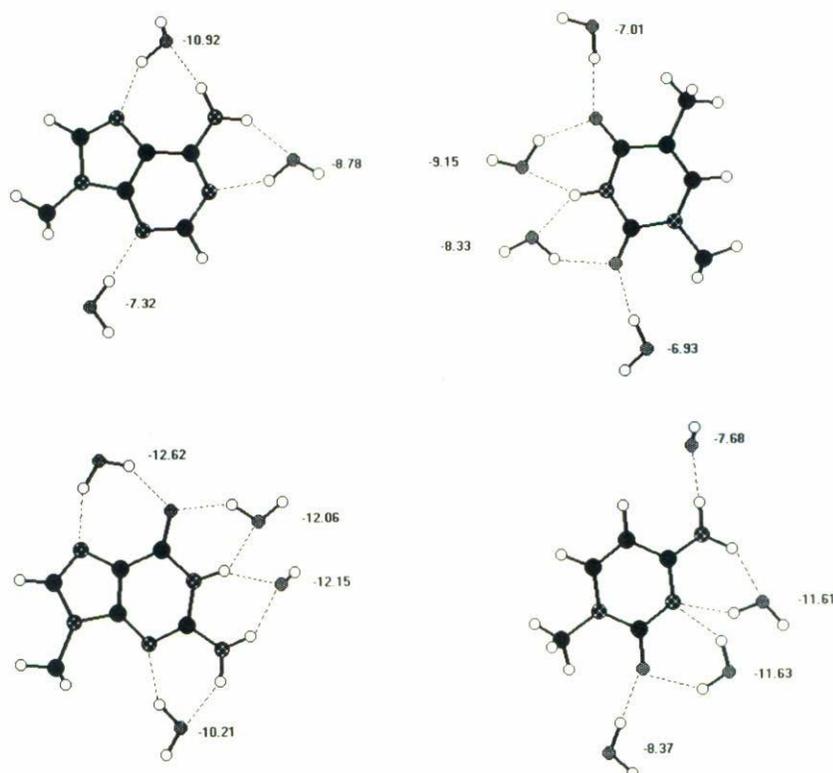


FIGURA 2. Posiciones mutuas de las moléculas de agua y las bases en los mínimos locales de energía de interacción entre moléculas de agua y adenina (arriba izquierda), timina (arriba derecha), guanina (abajo izquierda) y citosina (abajo derecha). Los valores experimentales de entalpía de formación de los complejos de una molécula de agua con las bases son iguales a  $10.6 \pm 1.0$ ,  $10.4 \pm 0.9$ ,  $14.1 \pm 1.0$  y  $11.4 \pm 0.8$  kcal/mol para adenina, timina guanina y citosina respectivamente [11].

se formado en cristales de hidratos reportados en experimentos de rayos X. En cambio en disoluciones acuosas las posiciones mutuas de las moléculas de agua del tipo antes mencionado rara vez se encuentran en bases y otros fragmentos de ADN. Los puentes formados entre átomos hidrofílicos con dos o tres moléculas de agua (ver abajo) son los más probables.

## 5. Simulación de la hidratación de las bases de los ácidos nucleicos

En esta sección presentamos los resultados de la simulación de la hidratación de las cuatro bases del ADN utilizando las funciones de potencial descritas en las anteriores secciones. En la Tabla IV presentamos los valores de la energía del sistema, consistiendo de una celda elemental con 400 moléculas de agua y una de las bases de los ácidos nucleicos, imponiendo condiciones periódicas a la frontera al sistema. Los datos de la Tabla IV demuestran una buena concordancia de los resultados calculados con el potencial refinado y con los datos experimentales. Esto nos da confianza para poder aplicar los potenciales para cálculos en sistemas más complejos, tales como los diferentes pares de bases y otros fragmentos de ADN. Estos datos serán presentados en otro trabajo. Aquí

TABLA IV. Características energéticas y enlaces-H agua-base de las bases de los ácidos nucleicos.  $E_T$  es la energía total del sistema,  $E_{WB}$  y  $E_{WW}$  son las contribuciones a la interacción agua-base y agua-agua,  $\Delta E$  es la energía de hidratación de la base calculada como la diferencia en la energía del sistema con bases y el mismo sistema conteniendo solo 400 moléculas de agua (kcal/mol),  $\Delta H$  es el valor experimental de la entalpía de hidratación de las bases, calculada como la diferencia de entalpía de disolución y de sublimación de la base [12, 13],  $H$  es número total de enlaces-H agua-base.

	$E_T$	$E_{WB}$	$E_{WW}$	$\Delta E$	$\Delta H$	$H$
A	-3807.1	-49.7	-3757.4	-25.9	-23.0	6.59
T	-3806.0	-46.0	-3760.0	-24.8	-23.4	5.54
G	-3821.6	-74.7	-3746.9	-40.4	-36.5	9.31
C	-3815.8	-62.3	-3751.5	-32.6	-29.5	6.57

exponemos algunos datos sobre la estructura de la capa de hidratación de cada una de las bases, en particular, el número promedio de enlaces-H con cada uno de los átomos hidrofílicos y la probabilidad de formación de algunos puentes de agua. Del número total de enlaces-H para cada base presentada en la Tabla IV, la mayor parte son enlaces-H en los

cuales los átomos de las bases participan como aceptores de protones. En particular para la adenina el número promedio de enlaces-H con los átomos N1, N3 y N7 es igual a 1.651, 1.457 y 1.792 respectivamente para la guanina 1.883, 2.505 y 2.073 con los átomos N3, O6 y N7. El número promedio de enlaces-H con todos los grupos de hidrógeno NH y NH<sub>2</sub> es igual a  $0.9 \pm 0.05$ . La posición de las moléculas de agua que forman enlaces-H con dos átomos hidrofílicos de la base (por consiguiente muy próximos a los mínimos de energía de interacción de una molécula de agua con las bases) se encuentran raramente. Por ejemplo, la probabilidad de formar enlaces-H de una molécula de agua entre los átomos N7 guanina o

adenina con átomos O6 o N6H, es de 6% y 19%, respectivamente. La probabilidad de formación de puentes es mucho mayor con dos moléculas de agua, es decir posiciones de las moléculas de agua cuando dos de ellas se enlazan entre sí y forman enlaces-H con diferentes átomos hidrofílicos de la base. Esta probabilidad para los átomos antes mencionados para adenina y guanina es igual a 38% y 37%, respectivamente.

### Agradecimientos

Este trabajo fue realizado dentro del proyecto 25238-E CO-NACYT (México).

1. R.E. Franklin and R.G. Gosling, *Nature* **171** (1953) 740.
2. E. González-Jiménez *et al.*, *J. Mol. Struct. (Theochem)* **493** (1999) 301.
3. F. Vovelle and J.M. Goodfellow, in *Water and Biological Macromolecules*, edited by E. Westhof, (CRC Press, Boca Raton, 1993) p. 244.
4. J. Petruska, L.C. Sowers, and M.F. Goodman, *Proc. Natl. Acad. Sci. USA*. **83** (1986) 1559.
5. H.R. Drew and R.E. Dickerson, *J. Mol. Biol.* **151** (1981) 535.
6. V.P. Chuprina, *Nucleic Acids Res.* **15** (1987) 293.
7. V.I. Poltev, J.I. Grokhlina, and G.G. Malenkov, *J. Biomol. Struct. Dyn.* **2** (1984) 413.
8. V.I. Poltev *et al.*, *J. Biomol. Struct. Dyn.* **13** (1996) 717.
9. M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids*, (Oxford University Press, New York, 1987).
10. A.K. Soper and M.G. Phillips, *Chem. Phys.* **107** (1986) 47.
11. L.F. Sukhodub, *Chem. Rev.* **87** (1987) 589.
12. A. Zielenkiewicz *et al.*, *J. Sol. Chem.* **13** (1984) 757.
13. A.B. Teplisky *et al.*, *Biophys. Chem.* **11** (1980) 17.