

Estudio de la influencia del agua en la citosina mediante métodos DFT

Study of the influence of water on cytosine using DFT methods

Francisco E. Rojas González^a, César Castillo Quevedo^b, Edgar Paredes Sotelo^c, Eduardo Robles Chaparro^d, Ana María Mendoza Wilson^e, Anned de León Flores, Khirbet López Velázquez^f, Edwin Hoil Canul, Adolfo López Sánchez, Jesús Cob Cantu, Luis Maldonado López^g, Gerardo Martínez Guajardo^h y José Luis Cabellos Quiroz^h

Resumen / Abstract

Las interacciones del agua con el ADN ocasionan un incremento en la estabilidad termodinámica de este último es de vital importancia el conocimiento de las estructuras más favorables energéticamente, sus interacciones, así como sus propiedades geométricas, ópticas y electrónicas como función de la cantidad de agua. En este trabajo presentamos un estudio teórico-computacional basado en la teoría funcional de la densidad de los efectos de la solvatación en las propiedades geométricas y vibracionales de la citosina. En los resultados se expone un análisis de las estructuras de más baja energía y de sus sitios de hidratación considerando una y dos moléculas de agua interactuando con la citosina.

*. El Doctor Rojas González desea expresar su agradecimiento a CONAHCYT por la beca otorgada 492050. De igual manera el Doctor Paredes Sotelo también agradece al CONAHCYT por la beca otorgada 1008864.

Conflicto de intereses. Los autores declaran no tener conflicto de interés.

Financiamiento. Esta investigación no recibió fondos.

a. Departamento de Física, Edificio 3F, Universidad de Sonora, 83000, Hermosillo, Sonora, México.

b. Departamento de Fundamentos del Conocimiento, Centro Universitario del Norte, Universidad de Guadalajara, Carretera Federal No. 23, Km. 191, C.P. 46200, Colotlán, Jalisco, México.

c. Departamento de Investigación en Polímeros y Materiales, Edificio 3G. Universidad de Sonora. Hermosillo, Sonora, México.

d. Los Profesores Robles Chaparro y De León Flores laboran como Profesores Investigadores en el Departamento de Ciencias Químico Biológicas, Edificio 5A. Universidad de Sonora. Hermosillo, Sonora, México.

e. Centro de Investigación en Alimentación y Desarrollo, A.C. Coordinación de Tecnología de Alimentos de Origen Vegetal, Carretera Gustavo Enrique Astiazarán Rosas No. 46, 83304 Hermosillo, Sonora, México.

f. Los Profesores e Investigadores López Velázquez, Hoil Canul, López Sánchez, Cob Cantú y Cabellos Quiroz laboran en la Universidad Politécnica de Tapachula, Carretera a Puerto Madero Km. 24+300, Tapachula - Puerto de San Benito, 30830 Puerto de San Benito, Chiapas.

g. Departamento de Física, CINVESTAV-Unidad, Mérida, Yuc, México.

h. Unidad Académica de Ciencias Químicas, Área de Ciencias de la Salud, Universidad Autónoma de Zacatecas, Km. 6 carretera Zacatecas-Guadalajara s/n, Ejido La Escondida C. P. 98160, Zacatecas, Zacatecas. Correo electrónico: jose.cabellos@uptapachula.edu.mx.

Palabras clave: 61.46.-w,65.40.gd,65.,65.80.-g,67.25.bd,71.15.-m,71.15.Mb,74.20.Pq,74.25.Bt,74.25.Gz,74.25.Kc, Teoría Funcional de la Densidad, citosina, agua, solvatación, DFT.

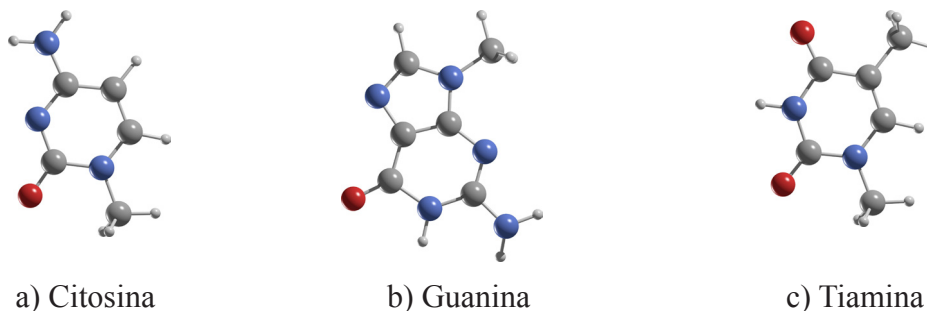
The interactions of water with DNA cause an increase in the thermodynamic stability of the latter; consequently, knowledge of the most energetically favorable structures, their interactions, and their geometric, optical, and electronic properties as a function of the quantity of water is essential in many areas. In this work, we present a theoretical-computational study based on the density functional theory of the effects of solvation on cytosine's geometric and vibrational properties. The results show an analysis of the lowest energy structures and their hydration sites, considering one and two water molecules interacting with cytosine.

Keywords: rainwater, pollutants, solids, electrocoagulation system, sustainability.

1. Introducción

El ADN es considerada la biomolécula más importante, siendo depositario de la información genética, cuyas instrucciones vienen codificadas por el ordenamiento de las bases nitrogenadas que la conforman: (A)denina, (C)itosina, (G)uanina, (T)imina. En la Figura 1, se muestra la citosina, guanina y timina. El ADN es una hélice doble que consta de dos cadenas complementarias de polinucleótidos, donde cada monómero está compuesto por una molécula de azúcar (desoxirribosa), un grupo fosfato y una de las bases mencionadas. La importancia de la hidratación en la estructura de ADN fue reconocida de manera casi inmediata después de su descubrimiento, en el año 1953¹. No obstante, debido a complejidades tales como el fuerte acoplamiento electrostático, tanto en la vecindad como entre la molécula negativamente cargada², y considerando que la estructura del ADN es tridimensional, el estudio de la dinámica del solvente sigue siendo un gran desafío, siendo hasta el día de hoy una pregunta abierta e importante línea de investigación activa³⁻⁷. Sin embargo, existe una vasta literatura acerca de la relación entre la sostenibilidad de la vida y las propiedades anómalas del agua, que debido principalmente a las características de los puentes de hidrógeno, le permiten actuar de diversas maneras en diferentes condiciones⁸.

Figura 1. Bases nitrogenadas



Fuente:

Generalmente, las funciones biológicas suceden en una solución mayoritariamente de agua, acompañada de una regulación de iones.⁹ Específicamente, la presencia del agua juega un rol importante en la conformación e interacción de ácidos. Basados se evidencia experimental, sabemos que la conformación geométrica y la estabilidad de la doble hélice de ADN dependen de la humedad y de la actividad del agua¹⁰. Un buen ejemplo de ello son los cambios conformacionales en su estructura, demostrando gran influencia en las transiciones desde la forma B (de alta hidratación) a la forma A (de baja hidratación)³. Este fenómeno puede explicarse como la consecuencia de la interacción entre las moléculas de agua y moléculas más grandes (ADN o una proteína), formándose una red alrededor de ellas, principalmente por puentes de hidrógeno. Estas interacciones pueden cambiar la estructura de la misma, así como alterar sus modos vibracionales¹¹⁻¹³, y como consecuencia cambiar sus propiedades electrónicas, ópticas y alterar las funciones biológicas, entre ellas la expansión, contracción o torsión de la hélice de ADN^{14,15}. Estudios experimentales de Resonancia Magnética Nuclear (RMN) han confirmado estructuras de moléculas de agua alrededor del DNA^{9,16,17}. Por otro lado, un número considerable de cálculos *Ab Initio* usando Teoría del Funcional de la Densidad (DFT) han sido publicados en estudio de la interacción de moléculas de agua con diferentes bases nitrogenadas, pares de bases o moléculas de interés biológico¹⁸⁻²⁵. También se ha usado como solvente para disociar ácido clorhídrico²⁶, así como estudios teóricos con dinámica molecular²⁷⁻²⁹. La mayoría de estos estudios se enfocan en los efectos del solvente sobre afinidad protónica¹⁸, sitios de localización o efectos del solvente²⁵.

La estructura tridimensional de ADN, está determinada principalmente por las interacciones de las bases nitrogenadas de ácidos nucleicos y por las interacciones con las moléculas de agua. La naturaleza de la estabilización los diferentes tipos de pares que pueden existir es debida a los puentes de hidrógeno, que son de origen electrostático y el apilamiento que se debe a las fuerzas de dispersión de London. Consecuentemente, las fuerzas que estabilizan la doble hélice de ADN juegan un papel importante en los procesos de transcripción y recombinación³⁰. Estudios experimentales recientes han confirmado que la doble hélice ADN es estabilizada principalmente por interacciones hidrofóbicas y dispersión relacionada con el apilamiento (*stacking*)³¹, contrario al pensamiento tradicional, donde los puentes de hidrógeno son los principales responsables de la estabilización de la doble hélice³²⁻³⁵. La motivación principal para incluir las fuerzas de London en los cálculos de la exploración de superficie potencial, es su importante contribución a la determinación de los isómeros (geométrica y energéticamente) que influyen de manera determinante en la espectroscopia y otras propiedades del sistema.

Debido al gran esfuerzo computacional requerido para el número de cálculos en el estudio a primeros principios o *Ab Initio* sobre una macromolécula tal como un gen o una proteína, el cual sobrepasa nuestras capacidades actuales, el presente estudio se limita al estudio en detalle de la solvatación de la base nitrogenada citosina. Entre los principales motivos de la selección está el interés por la comprensión de su estructura no planar, donde la deformabilidad específica en el grupo amino es esencial para la explicación de muchas interacciones base-base observadas en estructuras cristalizadas de ADN³⁶, confirmando consecuencias positivas en las interacciones de apilamiento, incrementando la estabilidad de la hélice³⁷. En particular, los ángulos Buckle y Propeller twisted, que caracterizan la rotación de las bases sobre el plano medio y el eje largo del par

de bases respectivamente, proveen un efecto de bloqueo estereoquímico, confiriendo rigidez a los segmentos escalonados con geometría no planar, siendo esto, junto con la piramidalización del grupo amino, factores clave para la comprensión de la flexibilidad del ADN³⁸. Como observación, una de las bases, la guanina tiene el más bajo potencial de ionización de todas las bases purínicas y pirimidínicas, volviéndola en el sitio más reactivo para un daño oxidativo³⁷. Dependiendo de las condiciones del ambiente, las bases de ácidos nucleicos, así como las pares de bases pueden existir en distintas formas tautoméricas. La posible existencia de una o más de las bases del ADN en forma tautomerizada incrementa la probabilidad de los mispairings o discordancias entre las purinas y pirimidinas, ocasionando mutaciones puntuales.

II. FUNDAMENTOS TEÓRICOS Y METODOLÓGICOS

Ecuación de Schrödinger

El estudio de sistemas químicos mediante la mecánica cuántica se basa en la solución de la ecuación de Schrödinger estacionaria independiente del tiempo dada en la ecuación (1)

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi(x) = E\Psi(x) \quad (1)$$

La ecuación de Schrödinger^{39,40} dependiente del tiempo está dada en la ec. (2)

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = \hat{H}\Psi(x, t) \quad (2)$$

El Hamiltoniano sin considerar efectos relativistas, \hat{H} , es la suma de la energía cinética y potencial $V(x, t)$ del sistema, esto es:

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x)$$

donde \hbar es la constante de Planck, m es la masa de los electrones y Ψ es la función de onda asociada al estado en el que se encuentra el sistema y contiene toda la información de las propiedades de este. La solución exacta de la ec. (2) se conoce para el átomo de hidrógeno, pero para sistemas que tienen muchos electrones no existe solución analítica exacta. Con el objeto de encontrar una solución se recurre a métodos numéricos a diferentes aproximaciones. Para el Hamiltoniano no relativista e independientes del tiempo, donde el potencial $V(r)$ sólo depende de la r se obtiene la ecuación de Schrödinger independiente del tiempo no relativista ec (4)(4)

$$\hat{H}\Psi(x, t) = E\Psi(x) \quad (4)$$

La solución de la ec.(4) es compleja y generalmente requiere el uso de la aproximación de *Born-Oppenheimer*^{41,42} (ABO). La energía total de un sistema molecular es:

$$E_t = E_{CM} + E_{CE} + E_{NN} + E_{EE} + E_{NE} \quad (5)$$

donde

- E_{CM} Energía cinética del núcleo;
- E_{CE} Energía cinética del electrón;
- E_{NN} Energía repulsiva del núcleo-núcleo;
- E_{EE} Energía repulsiva electrón- electrón: y
- E_{NE} Energía atractiva núcleo-electrón.

En la aproximación de Born-Oppenheimer, debido a que el núcleo no se mueve, la energía cinética del núcleo se considera cero y la energía es constante.

A. Teoría del funcional de la densidad (Density Functional Theory DFT)

En los últimos 30 años, la teoría del funcional de la densidad (*Density Functional Theory*, DFT) es uno de los temas más activos en el área de ciencia de materiales, química y física⁴³. Para tener un punto de referencia, el artículo más notable sobre DFT ha sido citado 84719 veces⁴⁴ (11/06/2024), y uno de los científicos más citados es *John P. Perdew*^{43,45}. La DFT aplicada en átomos, moléculas y sólidos, es un procedimiento variacional que resuelve la ecuación de Schrödinger (6) considerando muchos electrones e implementado en métodos numéricos y algoritmos escritos en código de programación, generalmente FORTRAN

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \Psi(x) = E\Psi(x) \quad (6)$$

En la Ec. (6), donde \hbar es la constante de Planck, m es la masa de los electrones, $V(x)$ es el potencial externo, E las energías propias y $\Psi(x,t)$ es la función de onda asociada al estado en el que se encuentra el sistema y contiene toda la información de las propiedades de este. El átomo de hidrógeno, sistema con dos cuerpos electrón-núcleo, es uno de los casos en que la solución de la ecuación de Schrödinger tiene una solución analítica (la partícula libre es otra). Para átomos multi-electrónicos la ecuación de Schrödinger⁴⁷. Consecuente y con el objeto de encontrar una solución, se recurre a DFT y/u otros métodos. La idea fundamental de DFT radica en la minimización de la energía total del sistema con respecto a la densidad electrónica del estado base, en lugar de usar la función de onda. La DFT es una teoría que describe adecuadamente las propiedades del estado base y su ventaja principal con respecto a métodos que usan la función de onda es que la densidad electrónica depende sólo de las coordenadas espaciales x,y,z , contrario a la función de onda, que posee un número grande de parámetros. La primera DFT fue propuesta por *Thomas* y *Fermi*^{48,49} en 1928. Originalmente el modelo fue propuesto para estudiar átomos multi-electrónicos.

En ese trabajo, el cálculo de la energía total del átomo se realizó expresando la energía cinética y las expresiones de las interacciones núcleo-electrón y electrón-electrón como función de la densidad electrónica la densidad electrónica. Los fundamentos que cimientan la *DFT* moderna fueron publicados por *Peter Hohenberg y Walter Kohn* en su artículo de 1964⁵⁰, donde establecieron que la energía es un funcional de la densidad, es decir, la energía es función de la densidad y esta es función de las coordenadas geométricas, una función de funciones. En su artículo⁵⁰, Peter Hohenberg y Walter Kohn mostraron que la energía es un funcional de la densidad electrónica estableciendo la Ec. 7:^{46,47}. Consecuente y con el objeto de encontrar una solución, se recurre a DFT y/u otros métodos. La idea fundamental de DFT radica en la minimización de la energía total del sistema con respecto a la densidad electrónica del estado base, en lugar de usar la función de onda. La DFT es una teoría que describe adecuadamente las propiedades del estado base y su ventaja principal con respecto a métodos que usan la función de onda es que la densidad electrónica depende sólo de las coordenadas espaciales x,y,z , contrario a la función de onda, que posee un número grande de parámetros. La primera DFT fue propuesta por *Thomas y Fermi*^{48,49} en 1928. Originalmente el modelo fue propuesto para estudiar átomos multi-electrónicos. En ese trabajo, el cálculo de la energía total del átomo se realizó expresando la energía cinética y las expresiones de las interacciones núcleo-electrón y electrón-electrón como función de la densidad electrónica la densidad electrónica. Los fundamentos que cimientan la *DFT* moderna fueron publicados por *Peter Hohenberg y Walter Kohn* en su artículo de 1964⁵⁰, donde establecieron que la energía es un funcional de la densidad, es decir, la energía es función de la densidad y esta es función de las coordenadas geométricas, una función de funciones. En su artículo⁵⁰, Peter Hohenberg y Walter Kohn mostraron que la energía es un funcional de la densidad electrónica estableciendo la Ec. 7:

$$E[\rho] = E_{xc}[\rho] + \int \rho(\mathbf{r}) V(\mathbf{r}) d\mathbf{r} \quad (7)$$

donde $E_{xc}[\rho]$ es el funcional que contiene la energía cinética y la energía debida a la interacción electrón-electrón. El problema de *DFT* es que no se conoce la forma exacta del funcional E_{xc} . Además, establecieron que la densidad electrónica que minimiza el funcional de la energía es la densidad del estado fundamental; la energía evaluada en esa densidad es la energía del estado fundamental. En 1965 *Walter Kohn y Lu Jeu Sham* en su artículo⁵¹ propusieron el término correlación e intercambio, que incluye las contribuciones de correlación electrónica dinámica y las desconocidas, que no están consideradas en la DFT. Las ecuaciones resultantes son conocidas como *Ecuaciones de Kohn-Sham* y son las que están implementadas, usando diferentes estrategias en un variedad de códigos de estructura electrónica como *Gaussian*⁵² y *NWChem*⁵³, entre otros. La ecuación de Schrödinger independiente del tiempo en notación (8):

$$\hat{H}|\Psi(\mathbf{x})\rangle = E|\Psi(\mathbf{x})\rangle \quad (8),$$

La cantidad más importante en la DFT es la energía del estado base, E_0 . El principio variacional establece que E_0 se obtiene mediante la minimización de la Ec.(9) con respecto a la densidad electrónica.

O en términos de la Ec. (7):

$$E_0 = \min \left\{ E_{xc}[\rho] + \int \rho(\mathbf{r}) V(\mathbf{r}) d\mathbf{r} \right\} \quad (10)$$

Los teoremas de *Hohenberg* y *Kohn* muestran que la densidad electrónica es la variable fundamental e importante en la descripción de un sistema en estado basal. Los dos teoremas de *Hohenberg* y *Kohn* en esencia dicen así:

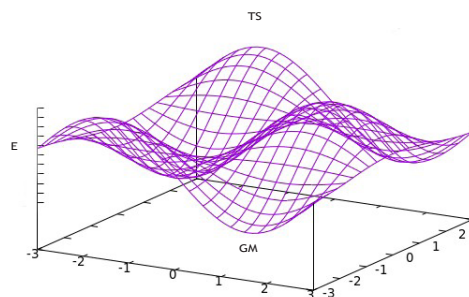
1. Todas las cantidades físicas son un funcional de la densidad electrónica.
2. La densidad electrónica que minimiza el funcional de la energía es la densidad del estado fundamental, Ec.(9) y la Ec. (10) .

La minimización del funcional en la Ec.(10) relativa a la densidad electrónica conduce a las ecuaciones de *Kohn-Sham*, las cuales tienen forma de ecuación de Schrödinger y como se mencionó en el párrafo anterior, son las ecuaciones implementadas en muchos códigos de estructura electrónica.

B. Superficie de energía potencial

La relación entre la geometría de un conjunto de átomos y su energía potencial define la superficie de energía potencial^{54,55} (PES por sus siglas en inglés) y juega un papel preponderante en el estudio del sistema molecular. En la Figura 2, se muestra una representación de una PES, y se muestran sus características más importantes.

Figura. 2. Representación de una superficie de energía potencial. La dimensionalidad de la PES es función de $3N-6$, donde N es el número de átomos. Es posible visualizar la PES multidimensional mediante una figura dimensional usando gráficos de desconexión (disconnectivity graphs).^{56,57}



Fuente:

El concepto de PES surge de la forma lógica y natural al resolver la ec. (4) bajo la aproximación de Born-Oppenheimer, considerando que la molécula es un conjunto de electrones y núcleos interactuando. En 1931 Michael Polanyi y Henry Eyring propusieron, para la reacción $H + H_2$ por primera vez el cálculo semiempírico PES.⁵⁸ Las PES fueron utilizadas para calcular las constantes de velocidad de reacción en la teoría del estado sólido de transición en 1935. La construcción de la superficie de energía potencial, que posee un sólo mínimo global y varios mínimos locales, es complicado y representa un desafío en estudios teóricos y experimentales⁵⁹, principalmente debido a estos factores:

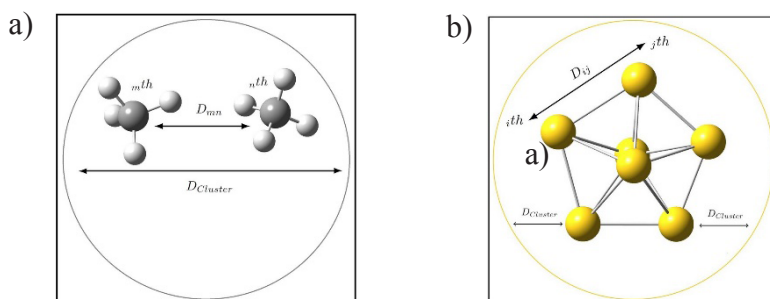
- El número de grados de libertad de la PES es $3N$, donde N es el número de átomos del sistema.
- El gran número de mínimos locales que la molécula posee y que crece exponencialmente como función del número de átomos, por ejemplo, una molécula con 147 átomos posee del orden de 10^{43} mínimos locales⁶⁰.
- El cálculo de una PES aproximada a la demanda real en el uso de métodos mecánico-cuánticos avanzados con capacidad para realizar cálculos con precisión química ($1\text{kcal} \cdot \text{mol}^{-1}$).

Actualmente para sistemas moleculares relativamente pequeños como BrH₂, de tres o cuatro átomos, el cálculo de la PES se ha realizado con métodos Ab-initio multireferenciales (*multireference configuration interaction (MRCI) method*) a muy alto nivel⁶¹⁻⁶³ para más de 10000 geometrías.

C. Exploración de la superficie de energía potencial

Con el objetivo de explorar eficaz y metódicamente la superficie de energía potencial (PES) de los sistemas estudiados en este artículo, se utiliza metodología heurística modificada *de la patada (kick)* implementada en el código escrito en *Python* llamado *Polytechnic University of Tapachula Search Atomic Cluster (PUTSAC)* y disponible en el cluster de la Universidad. Existe una amplia variedad de metodologías heurísticas para la búsqueda de estructuras de mínima energía en áreas de química, física y biología, tales como como Monte Carlo, Dinámica Molecular (MD), recocido simulado, algoritmos genéticos, *cuckoo search algorithm*⁶⁴⁻⁶⁹, entre otros. La eludación de estructuras de energía mínima un problema demasiado complejo y no existe ningún método que garantice encontrar el mínimo de energía usando recursos de cómputo y tiempo razonables. *PUTSAC* explota la paralización de manera eficiente a través del sistema administrador de colas de *Linux* y permite utilizar átomos individuales para formar las estructuras candidatas. La principal ventaja de utilizar átomos es la eficiencia para optimizar las estructuras iniciales debido a que estos fueron construidos con fragmentos preoptimizados. El método implementado en *PUTSAC* y el utilizado en este artículo para la búsqueda de la estructura de mínima energía sobre la PES, comienza con la generación de posiciones aleatorias, una por una, utilizando el algoritmo aleatorio para la generación de números aleatorios twister de Mersenne⁷⁰, en el que se colocan los individuales como el ejemplo mostrado en la Figura. 3 y donde,

Figura. 3. Parámetros importantes en la construcción de cúmulos, a) cúmulo atómico. b) cúmulos molecular



Fuente:

las estructuras ensambladas cumplen con tres condiciones geométricas:

1. Todos los átomos están colocados en el interior de una esfera de diámetro (ver Fig. 3) y es la suma de los radios covalentes de todos los átomos por los que está compuesta la molécula.
2. La distancia entre cualquier par de átomos de ser más larga que la suma de sus radios covalentes especificada en la figura a, b) $3 D_{mn}$.
3. Cada átomo debe estar conectado con otro átomo, como se muestra en la Figura 3.

La esfera está representada por el círculo verde en la Figura 3 y donde se muestra D_{ij} o D_{mn} que son las distancias mínimas a los que se pueden colocar dos átomos ij o moléculas mn . Todas las posiciones atómicas deben de quedar dentro de la esfera de diámetro $D_{cluster}$ y que satisfacen la siguiente Ec.(12).

$$D_{cluster} = \sum_i^{n-mol} D_{motif} \times \sigma \quad (11)$$

donde σ es un factor introducido por el usuario y varía típicamente entre y La distancia mínima de, D_{mn} , debe satisfacer la siguiente ecuación,

$$D_{mn} = \frac{[(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2]^{\frac{1}{2}}}{2\sigma} \quad (12)$$

Donde A y B son coordenadas atómicas de las moléculas o átomos y σ es un factor introducido por el usuario y varía entre y típicamente.

D. Nivel de teoría y detalles computacionales

El paso inicial consiste en la optimización geométrica del cúmulo a un nivel bajo de teoría. Remarcando en este artículo, todas las optimizaciones geométricas y cálculos de propiedades espectroscópicas se realizaron usando la teoría del funcional de la densidad tal y como está implementada en el código de estructura electrónica Gaussian 09⁵². También el algoritmo utilizado para realizar todas las optimizaciones geométricas realizadas en los sistemas de estudio en este artículo fueron con el algoritmo de *Berny*⁷¹ (*Berny algorithm*) tal y como está implementado en *Gaussian*.

Para la exploración de la superficie de energía potencial se usó el código PUTSAC escrito en *Python* y está acoplado a *Gaussian*. Para una primera optimización a la que nombramos Nivel bajo de teoría y donde se utilizó el funcional híbrido (*PBE0*)⁷², basado en el desarrollado del funcional por *Perdew, Burke, y Ernzerhof* (*PBE*)⁷³. *PBE0* considera las energías de intercambio de *PBE* y la energía de intercambio de *Hartree-Fock* y adecuado para modelar sistemas con puentes de hidrógeno⁷⁴, como es el caso de los sistemas estudiados en este artículo de citosina solvatadas con agua. La base usada en esta etapa es LANL2DZ⁷⁵, que es un conjunto de funciones base de los *Los Alamos*

National Laboratory, double- ζ quality basis set, double zeta (DZ), combinada con Effective Core Pseudopotentials (ECP), además de la inclusión de las fuerzas de Van der Waals.

IV. RESULTADOS

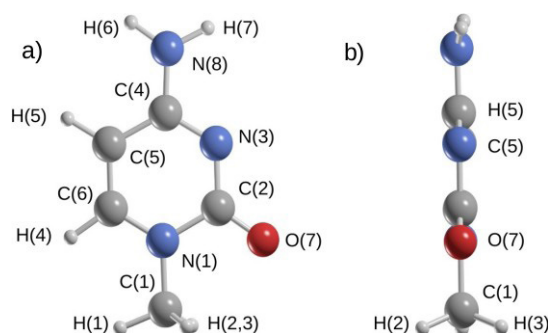
A. Citosina

La citosina es una de las cinco bases nitrogenadas y contiene un anillo aromático heterocíclico, un grupo amino y un grupo cetónico. La citosina es la base más pequeña que es común tanto en DNA como RNA y sus derivados son de gran importancia biológica. La citosina también es la base más alcalina en solución acuosa, esta última característica juega un importante rol en una gran cantidad de procesos bioquímicos⁷⁶. Es sabido que la estructura ceto-amino es la más estable⁷⁷. Estudios experimentales y teórico-computacionales se han realizado con la citosina con el objetivo de investigar sus propiedades geométricas y ópticas tanto en fase gas como en solución acuosa^{77,78}. Sitios de hidratación han sido investigados usando cálculos DFT^{79,80}. La citosina casi siempre se asocia con la guanina y viceversa.

B. Citosina en fase gas

La construcción de la citosina se realizó usando el visualizador molecular *chemcraft*. En la Figura (4) se muestra la citosina en fase gas en vista frontal en el inciso a) y en el inciso b) en vista lateral optimizada sin restricciones al nivel de teoría PBE0/Def2TZVP/GD3 usando el código de estructura electrónica Gaussian⁵². Las coordenadas atómicas se enlistan en el Apéndice A. Nuestro cálculo muestra que la estructura es planar, en concordancia con otros estudios teóricos^{77,79} como se puede observar en la Figura (4) y que por inspección visual se confirma en la columna Z de las coordenadas atómicas del Apéndice 1, donde los valores mayores a cero corresponden a los hidrógenos del grupo metilo,

Figura 4. Representación de citosina en fase de gas. a) vista de frente. b) vista de lado. En color rojo átomos de oxígeno (O), gris átomos de carbono (C), blanco átomos de hidrógeno (H) y en azul átomos de nitrógeno (N)



Fuente:

C. Longitud de enlace en la citosina en fase gas

La longitud del enlace para el ν_{12} , obtenida en nuestro cálculo es de 1.3580481825 \AA a nivel de teoría *PBE0/Def-2TZVP/GD3* (color verde en la Tabla I), en muy buena concordancia con el valor experimental de 1.361 \AA obtenido a través de rayos-X en la literatura⁸¹ y también con el valor obtenido de 1.361 \AA en cálculos Ab-initio⁷⁷. Nuestra longitud de enlace obtenida está ligeramente subestimada, sólo el 0.1% con respecto a la experimental, lo que consideramos en excelente acuerdo con los datos experimentales. Siguiendo el enlace de interés en este estudio de la citosina, es la longitud del enlace calculada para el ν_{15} es de $1.09048710778 \text{ \AA}$. El valor experimental medido es publicado en la referencia⁸¹ sólo ligeramente 1% por abajo del valor experimental, el valor teórico reportado en la misma referencia⁸¹ es de 1.361 \AA .

En la Tabla I se enlistan las distancias de enlaces más importantes que posee la citosina. La distancia más corta, enlace número 1 en la Tabla I, y coloreada en azul, ocurre entre los átomos $N(8)$, y $H(6)$, y es $1.00198186158 \text{ \AA}$. La distancia de enlace más larga ocurre en el enlace número 13 mostrado en la Tabla I y coloreado en color rojo y está entre $N(8)$ y $H(7)$ con valor de $1.00499463771 \text{ \AA}$.

Tabla 1. Longitudes de enlace optimizadas a nivel de teoría *PBE0/Def2TZVP/ GD3* de la citosina

Número de enlace	Átomo (i)	Átomo (j)	r [(i) - (j)] (Å)
1	N (3)	C(2)	1.3580481825
2	N (3)	C (4)	1.31009687535
3	N (1)	C (2)	1.42365965155
4	N (1)	C (6)	1.34307604947
5	N (1)	C (1)	1.44659309422
6	C(2)	O (7)	1.21330966932
7	C (4)	C (5)	1.42678052019
8	C (4)	N (8)	1.35152310711
9	C (5)	C (6)	1.35304815749
10	C (5)	H (5)	1.08042201191
11	C (6)	H (4)	1.08408155317
12	N (8)	H (6)	1.00198186158
13	N (8)	H (7)	1.00499463771
14	C (1)	H (3)	1.09051337582
15	C (1)	H (2)	1.09048710778
16	C (1)	H (1)	1.08986363233

Fuente:

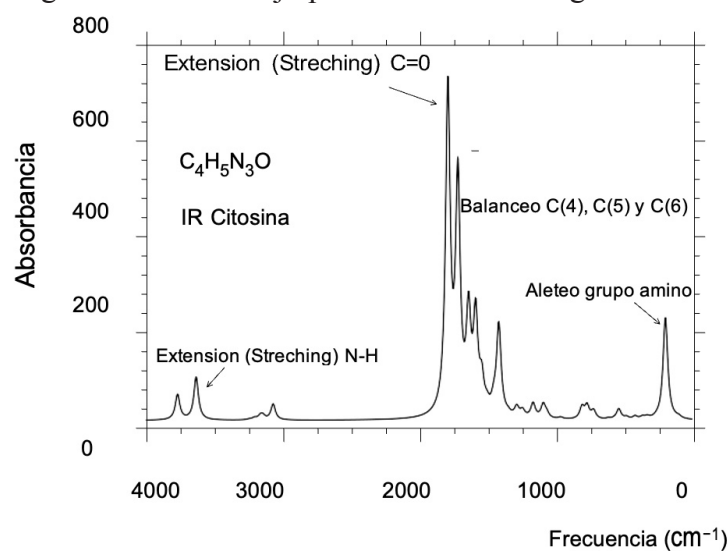
Los enlaces $C(5) - C(6)$ y $C(6) - C(1)$ no toman parte en la estructura resonante, esa es la razón por la que pueden ser predichos con muy buena exactitud. El enlace $C2 = O7$ es muy sensitivo a la presencia del solvente (todavía solvente) y a las correcciones por anarmonicidad, de manera similar la disminución de la longitud en los enlaces $C(2) - N(1)$, $C(2) - N(3)$, $C(4) - N(8)$, son de-

bidos a efectos de resonancia⁷⁷. En general, los enlaces calculados concuerdan muy bien con las mediciones experimentales⁷⁷.

D. Espectro infrarrojo de la citosina

El espectro infrarrojo de la citosina ha sido estudiado mediante DFT en fase gas y en soluciones acuosas⁸²⁻⁸⁶. La referencia⁷⁷ contiene una lista de los valores de las frecuencias experimentales para la citosina. El número de modos vibracionales para la citosina se calcula usando la ecuación $3N-6$, donde N es el número de átomos. La citosina posee átomos, aplicando la ecuación anterior se obtienen modos vibracionales. El modo más bajo calculado es de 100.3 cm^{-1} comparado con el experimental, de 131 cm^{-1} acorde con la referencia⁸⁷. El modo más bajo obtenido está subestimado en , con respecto al valor experimental. Las causas pueden ser varias, entre ellas podemos mencionar el nivel de teoría empleado, así como correcciones al conjunto de bases que no han sido usadas y el número de isómeros que se usaron en el cálculo. El espectro se muestra en la Figura 5, de los modos vibracionales que posee la citosina. Mucha atención ponemos a los de mayor intensidad, que están en el rango de 1500 cm^{-1} a 2000 cm^{-1} . El pico de mayor intensidad, en el espectro mostrado en la Figura 5, corresponde a una extensión o alargamiento en el plano en el enlace formado por $C(2) \equiv O(7)$. El modo también involucra el balanceo o compresión con menor intensidad del heterociclo.

Figura 5. IR infrarrojo para la citosina fase gas calculado



Fuente:

V. CITOSINA CON $[H_2O]_n$ ($n=1$)

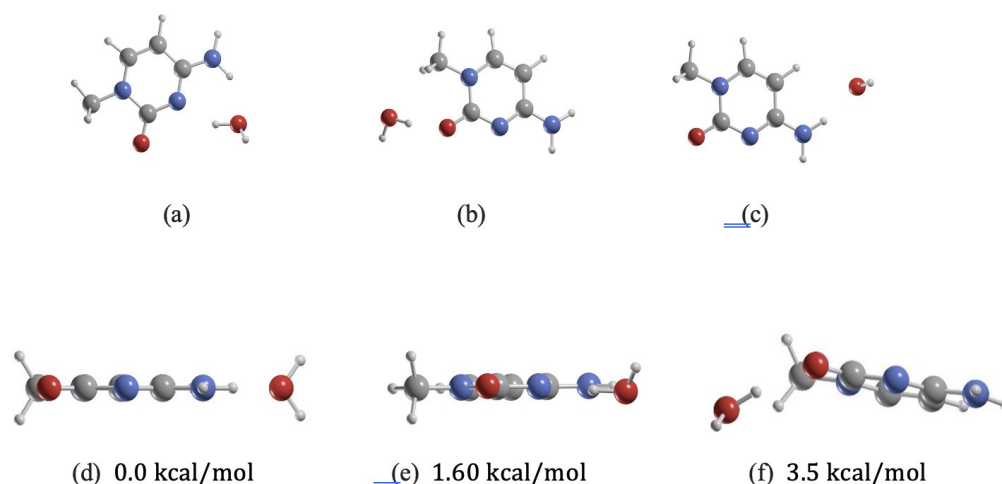
En esta sección se muestra la evolución de las propiedades geométricas de la citosina interactuando con moléculas de agua desde 1 hasta , con énfasis en cómo es la variación de los puen-

tes de hidrogeno que ocurren entre la citosina y las moléculas de agua (cómo y dónde). También, podemos resaltar la identificación de las tres estructuras energéticas más estables. Recalcando, es interesante investigar los cambios en las propiedades electrónicas, ópticas y geométricas como una función del número de moléculas de agua⁷⁹. Estudios *Ab-initio* previos se han realizado con cuatro moléculas sobre la molécula cromona (chromone)⁸⁸, donde se investigaron las energías y geometrías del sistema. En la misma dirección se han realizado estudios *Ab-initio* de la solvatación de timina y citosina⁸⁹ donde se investigó la transferencia de carga como función de la solvatación. También en este artículo identificamos los sitios de hidratación más favorables en todos los sistemas, y al final del capítulo presentamos los resultados acerca de la estabilidad energética.

A. CITOSINA CON [H₂O]₁

En la Figura 6 se muestra la citosina interaccionando con una molécula de agua y las tres estructuras energéticamente más bajas. El mínimo global energético se muestra en inciso a), Figura 6, posee energía relativa y de referencia cero, el segundo isómero, mostrado en inciso b) Figura 6, se localiza a arriba del mínimo global energético. La diferencia en energía entre el isómero de más baja energía y el segundo isómero es mayor a una 1 kcal/mol. El tercer isómero mostrado el en inciso c) Figura 6, se localiza a , muy arriba energía con respecto al mínimo global energético, lo que consideramos una diferencia energética considerable. Los tres isómeros mostrados en la Figura 6 son energéticamente, lo que comprobamos mediante el cálculo de las frecuencias vibracionales, resultando todas para todos los isómeros, como se muestra en la Figura 6. Para el mínimo global mostrado, inciso b) Figura 6, la molécula de agua prefiere asociarse con el nitrógeno,

Figura. 6. Vista de frente y de lado para las tres estructuras energéticamente más bajas con 1 molécula de agua, ([H₂O]₁), energías en paréntesis



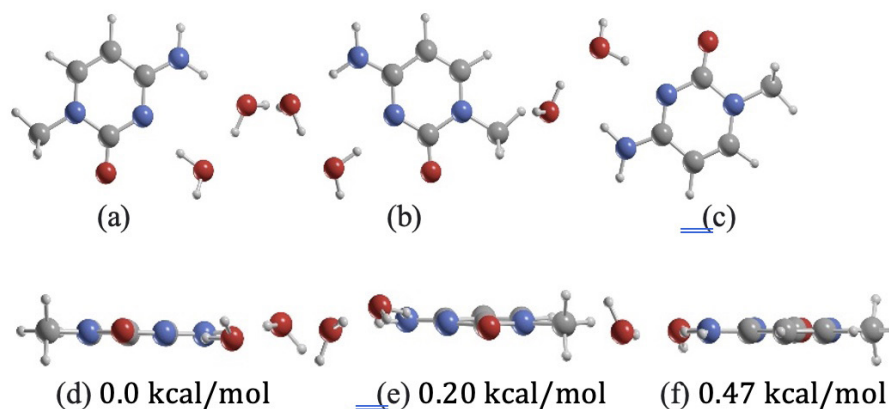
Fuente:

de la citosina a través de un puente de hidrógeno con longitud de 1.9547 Å y con el grupo amino NH_2 también con un puente de hidrógeno de longitud 1.8813 Å. Nuestros resultados, comparados con la longitud del puente de hidrógeno de 2.0910 Å con el grupo amino NH_2 y 2.105 Å reportados en un estudio similar⁸⁹, son sólo subestimadas con respecto a la reportada en trabajos anteriormente publicados⁸⁹. Esto se debe a la inclusión de la dispersión tomada en consideración en este artículo. Basados en la Figura 6, inciso a) el sitio de hidratación de la citosina es entre el grupo amino NH_2 y uno de los nitrógenos, en la Fig. 6 podemos inferir que los otros sitios, donde se coloca la molécula de agua, no son favorables en términos energéticos. Para el mínimo global, Figura.6 inciso a), la molécula de agua está en el plano, igual que para el segundo isómero, mientras que en el tercer isómero la molécula de agua permanece fuera del plano.

B. CITOSINA CON $[H_2O]_2$

Para el caso de dos moléculas de agua, el mínimo global así como la segunda y tercer estructuras energéticamente más bajas se muestran en la Fig 7 (inciso a), b) y c) vista frontal e incisos d), e) y f) vista lateral respectivamente). Mediante análisis vibracional probamos que estas fueran mínimos energéticos verdaderos.

Figura 7. Vista de frente y de lado para las tres estructuras energéticamente más bajas, con 2 moléculas de agua, $[H_2O]_2$, energías en paréntesis



Fuente:

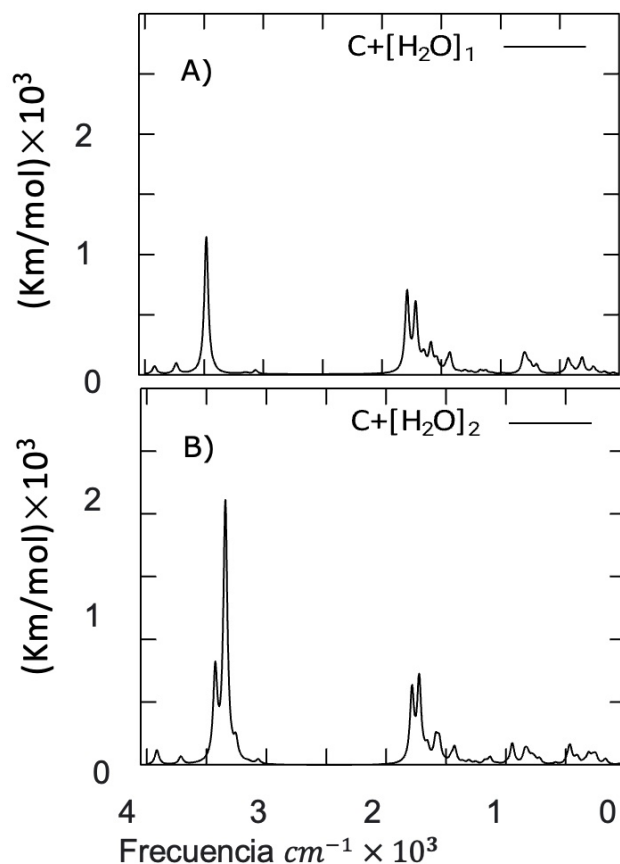
El cálculo arrojó todas las frecuencias vibracionales positivas. Los tres isómeros están localizados en un rango menor de 0.5 kcal/mol. En el mínimo global, Figura 7, inciso a) y d), las dos moléculas de agua se enlazan a la citosina por medio de puentes de hidrógeno, uno formado con el grupo amino y el otro con el nitrógeno localizado a un lado del grupo amino NH_2 . La longitud de enlace del puente de hidrógeno con el grupo amino NH_2 es de 1.8271 Å y la longitud de enlace del puente de hidrógeno con el nitrógeno es de 1.8019 Å. Los sitios de hidratación favorables están en el *filo* de la citosina y con el grupo amino NH_2 y el nitrógeno localizado a un lado del grupo

amino NH_2 , el segundo isomero localizado solo a 0.20 kcal/mol por arriba del mínimo global. La citosina no es planar (el enlace $C=O$ está fuera del plano). En estudios *ab-initio* previos⁸⁹ encontraron que la citosina es no planar en el grupo amino, una diferencia significativa entre nuestro resultados y los publicados. Explicamos estas diferencias a un efecto de las fuerzas de *van der Waals*, tomadas en consideración en nuestro cálculo. El tercer isómero localizado arriba solo 0.47 kcal/mol, las moléculas de agua permanecen en el plano como se puede apreciar en la Figura 7, inciso f).

VI. ESPECTROS INFRARROJOS $[H_2O]_n$ (n=1,2)

El espectro infrarrojo para la citosina (absorbancia) que se muestra en la Figura. 8 incisos A), y B) corresponden al espectro infrarrojo con 1 y 2 moléculas de agua, respectivamente y calculados a nivel de teoría *PBE0/Def2TZVP/GD3*. En la Figura. 5 se muestra el espectro infrarrojo para la citosina en fase gas, donde se observa claramente que el pico con mayor intensidad corresponde al elongamiento del enlace $C=O$. En la Figura. 8, inciso a) se puede observar que el pico de mayor intensidad en el espectro, localizado a 3483.19 cm^{-1} corresponde a una combinación de los hidrógenos del grupo amino NH_2 y una elongación de uno de los hidrógenos de la molécula del agua. El segundo pico en intensidad en el espectro IR de la Figura. 8 y localizado exactamente en 1789.71 cm^{-1} corresponde a la extensión/contracción del enlace $C=O$. El tercer pico en intensidad localizado a 1717.51 cm^{-1} corresponde a la vibración del esqueleto del heterociclo en combinación con un tijeeteo del grupo amino NH_2 .

Figura. 8. Espectro IR para la citosina. Panel, A) con una molécula de agua. Panel B), con dos moléculas de agua



Fuente:

VII. CONCLUSIONES

El agua es crucial y estudiarla es un reto de gran envergadura. Sabemos (*Ab studiis praevis*) que el agua es crucial en los sitios reactivos del ensamble proteína-ligando, y que la remoción o reorganización de partículas de agua del sitio reactivo o de su vecindad ocurre a nivel molecular con la participación de no más de una docena de moléculas de agua. La reorganización de unas cuantas moléculas de agua induce grandes cambios en las energías de interacción ocasionando enlaces más fuertes o débiles en el sistema proteína-ligada. Nuestros resultados y análisis muestran que las geometrías, longitudes de enlace y ángulos de los sistemas tratados a nivel *PBE0/Def2TZVP* y considerando la dispersión están en un muy buen acuerdo con datos experimentales y cálculos teóricos reportados en estudios previos. Cabe mencionar que nuestros datos se compararon con los datos experimentales y teóricos disponibles en la literatura. El efecto de la dispersión es acercar y/o mejorar el valor de las energías calculadas a los valores experimentales. Es interesante que las moléculas de agua prefieren ubicarse sobre una cara y no sobre las dos caras de la citosina, es

decir los sitios hidrófilos se localizan sobre una cara del sistema. Encontramos que sistemas con moléculas de agua localizadas en las dos caras de la citosina no son favorables energéticamente. Respecto a la geometría, la citosina en fase gas es plana mientras con la presencia de agua esta pierde la planariedad. Respecto a los espectros IR, basados en el análisis de resultados, concluimos que se presenta discrepancia entre los valores calculados del espectro fase gas y los reportados experimentalmente. Aunque los espectros presentan discrepancias, podemos capturar los cambios significativos y tendencias. Podemos afirmar que en el rango menor a 1000 cm^{-1} , la influencia a del agua es incrementar la intensidad de los picos IR. En el rango de $2000\text{-}3000\text{ cm}^{-1}$ no hay incremento de las intensidades en los picos del espectro IR, ni desplazamiento en frecuencias de los picos IR. En el rango de $2000\text{-}3000\text{ cm}^{-1}$ corresponde a modos vibracionales relacionados con el puente de hidrógeno, adicionar agua no cambia el modo. En el rango de $3000\text{-}4000\text{ cm}^{-1}$ hay muchos incrementos en intensidades y en ensanchamientos de los picos IR, debido a los OH que se adicionan con el agua, estos cambios si afectan el puente de hidrógeno. Futuro trabajo es incluir un nivel de teoría CCSD(T).

Referencias bibliograficas

- ¹ R. E. Franklin and R. G. Gosling. The structure of sodium thymonucleate fibres. I. The influence of water content. *Acta Crystallographica*, 6(8-9):673–677, Sep 1953.
- ² Him Shweta and Sobhan Sen. Dynamics of water and ions around dna: What is so special about them? *Journal of Biosciences*, 43(3):499–518, Jul 2018.
- ³ Surjit B. Dixit, Mihaly Mezei, and David L. Beveridge. Studies of base pair sequence effects on dna solvation based on all-atom molecular dynamics simulations. *Journal of Biosciences*, 37(3):399–421, Jul 2012.
- ⁴ Olga Kennard Wolfram Saenger, William N. Hunter and. DNA conformation is determined by economics in the hydration of phosphate groups. *Nature*, 324(6095):385–388, Sep 1986.
- ⁵ Berman HM. Schneider B, Patel K. Hydration of the phosphate group in double-helical DNA. *Biophys J.*, 5(75):2422–2434, Jan 1998.
- ⁶ S Leikin, V A Parsegian, D C Rau, and R P Rand. Hydration forces. *Annual Review of Physical Chemistry*, 44(1):369–395, 1993. PMID: 8257560.
- ⁷ Tigran V. Chalikian and Kenneth J. Breslauer. Volumetric properties of nucleic acids. *Biopolymers*, 48(4):264–280, 1998.
- ⁸ Martin Chaplin. Water's hydrogen bond strength, 2007.
- ⁹ Jaydeep Bardhan. Biomolecular electrostatics—i want your solvation (model)*. *Computational Science and Discovery*, 5:013001, 11 2012.
- ¹⁰ A.G.W. Leslie, Struther Arnott, Rengaswami Chan drasekaran, and R.L. Ratliff. Polymorphism of dna double helices. *Journal of Molecular Biology*, 143(1):49–72, 1980.
- ¹¹ F. H. C. Crick J. D. Watosn and. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171(130):737–738, 1953.
- ¹² Martin F. Chaplin. Water: its importance to life. *Biochemistry and Molecular Biology Education*, 29(2):54–59, 2001.

- ¹³ Martin F. Chaplin. Do we underestimate the importance of water in cell biology? *Nature Reviews Molecular Cell Biology*, 7(11):861–866, 2006.
- ¹⁴ G. Albiser, A. Lamiri, and S. Premilat. The a–b transition: temperature and base composition effects on hydration of dna. *International Journal of Biological Macromolecules*, 28(3):199 – 203, 2001.
- ¹⁵ Monika Fuxreiter, Mihaly Mezei, Istv'an Simon, and Roman Osman. Interfacial water as a "hydration fingerprint" in the noncognate complex of bamhi. *Biophysical journal*, 89(2):903—911, August 2005.
- ¹⁶ David G. Gorenstein. Conformation and dynamics of dna and protein-dna complexes by 31p nmr. *Chemical Reviews*, 94(5):1315–1338, 1994.
- ¹⁷ Qiang Chen, Hua-Jin Zhai, Si-Dian Li, and Lai-Sheng Wang. On the structures and bonding in boron-gold al-loy clusters: B6aun-and b6aun (n = 1-3). *J. Chem. Phys.*, 138(8):084306, 2013.
- ¹⁸ G. Maes, J. Smets, L. Adamowicz, W. McCarthy, M.K. Van Bael, L. Houben, and K. Schoone. Correlations between ab initio and experimental data for isolated h-bonded complexes of water with nitrogen bases. *Journal of Molecular Structure*, 410 411:315 – 322, 1997. *Molecular Spectroscopy and Molecular Structure* 1996.
- ¹⁹ Jir'ı Sponer and Pavel Hobza. *DNA Bases and Base Pairs: Ab Initio Calculations*. American Cancer Society, 2002.
- ²⁰ Barbara Marchetti, Tolga N. V. Karsili, Michael N. R. Ashfold, and Wolfgang Domcke. A 'bottom up', ab initio computational approach to understanding fundamental photo-physical processes in nitrogen containing heterocycles, dna bases and base pairs. *Phys. Chem. Chem. Phys.*, 18:20007–20027, 2016.
- ²¹ Pavel Hobza and Jir'ı Sponer. Structure, energetics, and dynamics of the nucleic acid base pairs: Nonempirical ab initio calculations. *Chemical Reviews*, 99(11):3247–3276, 1999. PMID: 11749516.
- ²² Jir'ı Sponer, Imre Berger, Nada Spackov'a, Jerzy Leszczynski, and Pavel Hobza. Aromatic base stacking in dna: From ab initio calculations to molecular dynamics simulations. *Journal of Biomolecular Structure and Dynamics*, 17(sup1):1–24, 2000. PMID: 22607400.
- ²³ C. Espejo and Rey-Gonzales. R. Stabilization and interaction energies of non planar dna base pairs guanine-cytosine and adenine-thymine. a study based on the siesta method. *Revista Mexicana de Física*, 53:212–216, 2007.
- ²⁴ Semen A. Trygubenko, Tetyana V. Bogdan, Manuel Rueda, Modesto Orozco, F. Javier Luque, Jir'ı Sponer, Petr Slav'ıcek, and Pavel Hobza. Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution part 1. cytosine. *Phys. Chem. Chem. Phys.*, 4:4192–4203, 2002.
- ²⁵ Federico Moroni, Antonino Famulari, and Mario Raimondi. Ab initio study of the crystallographic solvation pattern of the cytosine-guanine base pair in dna. *The Journal of Physical Chemistry A*, 105(7):1169–1174, 2001.
- ²⁶ Alba Vargas-Caamal, Cabellos, Jose Luis, Filiberto Ortiz-Chi, Henry S. Rzepa, Albeiro Restrepo, and Gabriel Merino. How many water molecules does it take to dissociate hcl? *Chemistry – A European Journal*, 22(8):2812– 2818, 2016.
- ²⁷ V. Fritsch, G. Ravishanker, D. L Beveridge, and E. Westhof. Molecular dynamics simulations of poly(da)·poly(dt): Comparisons between implicit and explicit solvent representations. *Biopoly-*

- mers, 33(10):1537–1552, 1993.
- ²⁸ Alona Furmanchuk, Olexandr Isayev, Oleg V. Shishkin, Leonid Gorb, and Jerzy Leszczynski. Hydration of nucleic acid bases: a Car-Parrinello molecular dynamics approach. *Phys. Chem. Chem. Phys.*, 12:3363–3375, 2010.
- ²⁹ Santiago Tolosa, Jorge A. Sansón, and Antonio Hidalgo. Theoretical study of adenine to guanine transition assisted by water and formic acid using steered molecular dynamic simulations. *Frontiers in Chemistry*, 7:414, 2019.
- ³⁰ Bobo Feng, Robert P. Sosa, Anna K. F. Martensson, Kai Jiang, Alex Tong, Kevin D. Dorfman, Masayuki Takahashi, Per Lincoln, Carlos J. Bustamante, Fredrik West-erlund, and Bengt Nordén. Hydrophobic catalysis and a potential biological role of dna unstacking induced by environment effects. *Proceedings of the National Academy of Sciences*, 116(35):17169–17174, 2019.
- ³¹ Theodore T. Herskovits. Nonaqueous solutions of dna: Factors determining the stability of the helical configuration in solution. *Archives of Biochemistry and Biophysics*, 97(3):474 – 484, 1962.
- ³² O. Sınanoğlu and S. Abdunur. Hydrophobic stacking of bases and the solvent denaturation of dna*. *Photochemistry and Photobiology*, 3(4):333–342, 1964.
- ³³ Maxim Frank-Kamenetskii. How the double helix breathes. *Nature*, 18(328):17—18, August 1987.
- ³⁴ Peter Yakovchuk, Ekaterina Protozanova, and Maxim D. Frank-Kamenetskii. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Research*, 34(2):564–574, 01 2006.
- ³⁵ S. Walter Englander. A hydrogen exchange method using tritium and sephadex: Its application to ribonuclease*. *Biochemistry*, 2(4):798–807, 1963. PMID: 14075117.
- ³⁶ Jiri Sponer and Pavel Hobza. Nonplanar geometries of dna bases. ab initio second-order moller-plesset study. *The Journal of Physical Chemistry*, 98(12):3161–3164, 1994.
- ³⁷ Manoj K. Shukla and Jerzy Leszczynski. Tautomerism in nucleic acid bases and base pairs: a brief overview. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(6):637–649, 2013.
- ³⁸ Noriyuki Kurita, Victor I. Danilov, and Victor M. Anisimov. The structure of watson–crick dna base pairs obtained by mp2 optimization. *Chemical Physics Letters*, 404(1):164 – 170, 2005.
- ³⁹ E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, Dec 1926.
- ⁴⁰ Vanessa A. Seifert. An alternative approach to unifying chemistry with quantum mechanics. *Foundations of Chemistry*, 19(3):209–222, Oct 2017.
- ⁴¹ M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389(20):457–484, Jan 1927.
- ⁴² Michael Baer and Robert Englman. A modified born- oppenheimer equation: application to conical intersections and other types of singularities. *Chemical Physics Letters*, 265(1):105 – 108, 1997.
- ⁴³ Axel D. Becke. Perspective: Fifty years of density functional theory in chemical physics. *The Journal of Chemical Physics*, 140(18):18A301, 2014.
- ⁴⁴ Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle salvetti correlation energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, 1988.
- ⁴⁵ Gus Scuseria, Mel Levy, and Kieron Burke. Special issue in honor of john p. perdew for his 65th birthday. *Journal of Chemical Theory and Computation*, 5(4):675–678, 2009. PMID: 26609571.

- ⁴⁶ Zhongqi Ma. Quantum three-body problems. *Science in China Series A: Mathematics*, 43(10):1093–1107, Oct 2000.
- ⁴⁷ Book Review: Quantum mechanics. L. I. Schiff, Second edition. McGraw Hill Book Co, New York, Toronto, London. 1955. xii-417 pages. *McGraw Hill*, 1:133–133, 1956.
- ⁴⁸ L. H. Thomas. The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23(5):542–548, 1927.
- ⁴⁹ Elliott H Lieb and Barry Simon. The thomas-fermi theory of atoms, molecules and solids. *Advances in Mathematics*, 23(1):22 – 116, 1977.
- ⁵⁰ P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.
- ⁵¹ W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.
- ⁵² M. J. Frisch, G. W. Trucks, H. B. Schlegel, and G. E. Scuseria et al. Gaussian 09. *Gaussian Inc. Wallingford CT 2009*, jan 2009.
- ⁵³ M. Valiev, E.J. Bylaska, N. Govind, K. Kowalski, T.P. Straatsma, H.J.J. Van Dam, D. Wang, J. Nieplocha, E Apra, T.L. Windus, and W.A. de Jong. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477 – 1489, 2010.
- ⁵⁴ Errol G. Lewars. Computational chemistry, introduction to the theory and applications of molecular and quantum mechanics. *Book*, 1(1):1–300, Oct 2011.
- ⁵⁵ I. N. Levine. Quantum chemistry. *Pearson Prentice Hall*, 1(1):1–400, Oct 209.
- ⁵⁶ David J. Wales and Tetyana V. Bogdan. Potential energy and free energy landscapes. *The Journal of Physical Chemistry B*, 110(42):20765–20776, 2006. PMID: 17048885.
- ⁵⁷ Lewis C. Smeeton, Mark T. Oakley, and Roy L. Johnston. Visualizing energy landscapes with metric disconnectivity graphs. *Journal of Computational Chemistry*, 35(20):1481– 1490, 2014.
- ⁵⁸ Henry Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.
- ⁵⁹ Si-Da Huang, Cheng Shang, Pei-Lin Kang, and Zhi-Pan Liu. Atomic structure of boron resolved using machine learning and global sampling. *Chem. Sci.*, 9:8644–8655, 2018.
- ⁶⁰ C. J. Tsai and K. D. Jordan. Use of an eigenmode method to locate the stationary points on the potential energy surfaces of selected argon and water clusters. *The Journal of Physical Chemistry*, 97(43):11227–11237, 1993.
- ⁶¹ Bin Jiang, Changjian Xie, and Daiqian Xie. New ab initio potential energy surface for brh2 and rate constants for the h + hbr h2 + br abstraction reaction. *The Journal of Chemical Physics*, 134(11):114301, 2011.
- ⁶² Bin Jiang, Changjian Xie, and Daiqian Xie. New ab initio coupled potential energy surfaces for the br(2p^{3/2}, 2p^{1/2}) + h2 reaction. *The Journal of Chemical Physics*, 135(16):164311, 2011.
- ⁶³ Changjian Xie, Bin Jiang, and Daiqian Xie. State-to-state quantum dynamics of the h + hbr reaction: Competition between the abstraction and exchange reactions. *The Journal of Chemical Physics*, 134(18):184303, 2011.
- ⁶⁴ Carlos Emiliano Buelna-Garcia, Eduardo Robles Chaparro, Tristan Parra-Arellano, Jesus Manuel Quiroz Castillo, Teresa del Castillo Castro, Gerardo Martínez-Guajardo, Cesar Castillo

- Quevedo, Aned de León Flores, Gilberto Anzueto Sanchez, Martha Fabiola Martin del Campo Solis, Ana María Mendoza Wilson, Alejandro V´asquez Espinal, and José Luis Cabellos Quiroz. Theoretical prediction of structures, vibrational circular dichroism, and infrared spectra of chiral Be₄B₈ cluster at different temperatures. *Molecules*, 26(13), 2021.
- ⁶⁵ Carlos Emiliano Buelna García, José Luis Cabellos Quiroz, Jesús Manuel Quiroz Castillo, Gerardo Martínez Guajardo, Cesar Castillo Quevedo, Aned de León Flores, Gilberto Anzueto Sánchez, and Martha Fabiola Martin del Campo Solís. Exploration of free energy surface and thermal effects on relative population and infrared spectrum of the Be₆B–11 fluxional cluster. *Materials*, 14(1), 2021.
- ⁶⁶ Abril C. Castro, Edison Osorio, José Luis Cabellos Quiroz, Erick Cerpa, Eduard Matito, Miquel Solá, Marcel Swart, and Gabriel Merino. Exploring the potential energy surface of E2P4 clusters e group 13 element the quest for inverse carbon-free sandwiches. *Chemistry - A European Journal*, 20(16):4583–4590, 2014.
- ⁶⁷ Zhong-hua Cui, Yi-hong Ding, Jose Luis Cabellos Quiroz, Edison Osorio, Rafael Islas, Albeiro Restrepo and Gabriel Merino. Planar tetracoordinate carbons with a double bond in CA13E clusters. *Phys. Chem. Chem. Phys.*, 17:8769–8775, 2015.
- ⁶⁸ Zhong-hua Cui, Valentin Vassilev-Galindo, Jose Luis Cabellos Quiroz, Edison Osorio, Mesías Orozco, Sudip Pan, Yi-hong Ding, and Gabriel Merino. Planar pentacoordinate carbon atoms embedded in a metallocene framework. *Chem. Commun.*, 53:138–141, 2017.
- ⁶⁹ Maryel Contreras, Sudip Pan, Mesías Orozco-Ic, José Luis Cabellos Quiroz, and Gabriel Merino. E₃m₃⁺ (e=c-pb, m=li-cs) clusters the smallest molecular stars. *Chemistry – A European Journal*, 23(47):11430–11436, 2017.
- ⁷⁰ Makoto Matsumoto and Takuji Nishimura. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, 8(1):3–30, January 1998.
- ⁷¹ Xiaosong Li and Michael J. Frisch. Energy-represented direct inversion in the iterative subspace within a hybrid geometry optimization method. *Journal of Chemical Theory and Computation*, 2(3):835–839, 2006. PMID: 26626690.
- ⁷² Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of Chemical Physics*, 110(13):6158–6170, 1999.
- ⁷³ John P. Perdew, Matthias Ernzerhof, and Kieron Burke. Rationale for mixing exact exchange with density functional approximations. *The Journal of Chemical Physics*, 105(22):9982–9985, 1996.
- ⁷⁴ R. E. A. Kelly and L. N. Kantorovich. Planar nucleic acid base super-structures. *J. Mater. Chem.*, 16:1894–1905, 2006.
- ⁷⁵ P. Jeffrey Hay and Willard R. Wadt. Ab initio effective core potentials for molecular calculations. potentials for k to au including the outermost core orbitals. *The Journal of Chemical Physics*, 82(1):299–310, 1985.
- ⁷⁶ Ram Avtar Yadav, Rashmi Singh, and Mayuri Srivastava. Vibrational studies and dft calculations of cytosine, thiocytosine and their cations and anions. 2015.
- ⁷⁷ Roksana Walesa, Teobald Kupka, and Ma-Igorzata A. Broda. Density functional theory (dft) prediction of structural and spectroscopic parameters of cytosine using harmonic and anharmonic approximations. *Structural Chemistry*, 26(4):1083–1093, Aug 2015.

- ⁷⁸ A.K Chandra, M.T Nguyen, and Th Zeegers-Huyskens. Theoretical study of the protonation and deprotonation of cytosine. implications for the interaction of cytosine with water. *Journal of Molecular Structure*, 519(1):1 – 11, 2000.
- ⁷⁹ Carlos Alem'an. Hydration of cytosine using combined discrete/scrf models: influence of the number of discrete sol- vent molecules. *Chemical Physics*, 244(2):151 – 162, 1999.
- ⁸⁰ Tanja van Mourik, David M. Benoit, Sarah L. Price, and David C. Clary. Ab initio and diffusion monte carlo study of uracil-water, thymine-water, cytosine-water, and cytosine-(water)₂. *Phys. Chem. Chem. Phys.*, 2:1281– 1290, 2000.
- ⁸¹ Donald Voet and Alexander Rich. The crystal structures of purines, pyrimidines and their intermolecular complexes. volume 10 of *Progress in Nucleic Acid Research and Molecular Biology*, pages 183 – 265. Academic Press, 1970.
- ⁸² D. Ferro, L. Bencivenni, R. Teghil, and R. Mastromarino. Vapour pressures and sublimation enthalpies of thymine and cytosine. *Thermochimica Acta*, 42(1):75 – 83, 1980.
- ⁸³ Tabish Rasheed and Shabbir Ahmad. Approximate solution of the mode–mode coupling integral: Application to cytosine and its deuterated derivative. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 77(2):446 – 456, 2010.
- ⁸⁴ Philippe Carbonniere, Sandrine Thicoipe, Thibault Very, and Xavier Assfeld. Vibrational analysis beyond the harmonic from ab initio molecular dynamics: Case of cytosine in its anhydrous and aqueous forms. *International Journal of Quantum Chemistry*, 112(9):2221–2230, 2012.
- ⁸⁵ Philippe Carbonniere, Sandrine Thicoipe, Thibault Very, and Xavier Assfeld. Vibrational analysis beyond the harmonic from ab initio molecular dynamics: Case of cytosine in its anhydrous and aqueous forms. *International Journal of Quantum Chemistry*, 112(9):2221–2230, 2012.
- ⁸⁶ A. Aamouche, M. Ghomi, L. Grajcar, M. H. Baron, F. Romain, V. Baumruk, J. Stepanek, C. Coulombeau, H. Jobic, and G. Berthier. Neutron inelastic scattering, optical spectroscopies and scaled quantum mechanical force fields for analyzing the vibrational dynamics of pyrimidine nucleic acid bases 3 cytosine. *The Journal of Physical Chemistry A*, 101(51):10063–10074, 1997.
- ⁸⁷ P.C. Mishra. Vibrational structures and intensity distributions in the electronic absorption spectra of nucleic acid bases: evidence for non-planarity of guanine. *Journal of Molecular Structure*, 144(3):309 – 317, 1986.
- ⁸⁸ Carlos Aleman and Sergio E. Galembeck. Solvation of chromone using combined discrete/scrf models. *Chemical Physics*, 232(1):151 – 159, 1998.
- ⁸⁹ Carlos Alem'an. Solvation of cytosine and thymine using a combined discrete/scrf model. *Chemical Physics Letters*, 302(5):461 – 470, 1999.

Appendix A: Coordenadas XYZ citosina

16

0.000000000 cytosine_phase_gas.out

N	-0.727755000000	-1.065905000000	0.001802000000
N	1.106176000000	0.443283000000	0.000343000000
C	0.619442000000	-0.894587000000	-0.000489000000
C	-1.534724000000	-0.033850000000	-0.002360000000
C	-1.077954000000	1.317839000000	-0.001551000000
C	0.264143000000	1.489626000000	0.000498000000
H	-1.751615000000	2.162494000000	-0.008659000000
H	0.718516000000	2.473891000000	0.001127000000
N	-2.861847000000	-0.288140000000	-0.028739000000
H	-3.530629000000	0.441917000000	0.125258000000
H	-3.151046000000	-1.245402000000	0.071399000000
O	1.435541000000	-1.792419000000	0.000202000000
C	2.543022000000	0.610921000000	0.001968000000
H	2.974786000000	0.134033000000	0.882523000000
H	2.976586000000	0.133236000000	-0.877237000000
H	2.779486000000	1.674823000000	0.001735000000
19			
0.000000000 C+H2O			
H	3.567300000000	-1.910778000000	0.547735000000
O	3.273831000000	-1.289701000000	-0.121054000000
H	2.292345000000	-1.302548000000	-0.085739000000
N	0.657249000000	-0.378286000000	0.021598000000
N	-1.670339000000	0.058738000000	-0.007184000000
C	-0.604991000000	-0.877243000000	0.016361000000
C	0.886461000000	0.921844000000	0.018162000000
C	-0.173061000000	1.877960000000	-0.000783000000
C	-1.430819000000	1.381986000000	-0.014905000000
H	0.008965000000	2.942791000000	-0.005140000000
H	-2.302083000000	2.026798000000	-0.032151000000
N	2.163539000000	1.328044000000	0.037298000000
H	2.390282000000	2.302449000000	-0.009654000000
H	2.895912000000	0.622397000000	0.027797000000
O	-0.878946000000	-2.059634000000	0.029845000000

C	-3.012334000000	-0.482665000000	-0.019826000000
H	-3.150846000000	-1.119952000000	-0.893703000000
H	-3.177742000000	-1.095971000000	0.866533000000
H	-3.727891000000	0.338719000000	-0.042039000000

Contribuciones de cada uno de los coautores al documento.

Francisco E. Rojas González. Investigación, software, escritura inicial manuscrito,
César Castillo Quevedo. Investigación, software, supervisión, financiamiento, escritura inicial del manuscrito, escritura final del manuscrito;
Edgar Paredes Sotelo. Investigación, software, escritura inicial manuscrito;
Eduardo Robles-Chaparro. Investigación, software, escritura inicial manuscrito,
Ana María Mendoza Wilson. Investigación, software, supervisión, financiamiento, escritura inicial manuscrito, escritura final manuscrito;
Anned de León Flores. Investigación, software, supervisión, financiamiento, escritura inicial manuscrito, escritura final manuscrito
Khirbet López Velázquez. Investigación, software, supervisión, financiamiento, escritura inicial manuscrito, escritura final manuscrito,
Edwin Hoil Canul. Investigación, software, escritura inicial manuscrito,
Adolfo López Sánchez. Investigación, software, escritura inicial manuscrito,
Jesús Cob Cantú. Investigación, software, escritura inicial manuscrito,
Luis Maldonado López. Investigación, supervisión, escritura inicial manuscrito,
Gerardo Martínez Guajardo. Investigación, software, escritura inicial manuscrito,
José Luis Cabellos Quiroz. Idea original, escritura final del manuscrito.