



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

UNIDAD CUAJIMALPA

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

MATERIAL DIDÁCTICO

MODELACIÓN MOLECULAR EN SISTEMAS BIOLÓGICOS

AUTORES:

Dr. Edgar Vázquez Contreras

Dr. Marco Vinicio Vázquez González

Departamento de Ciencias Naturales

Departamento de Matemáticas

Aplicadas y Sistemas

ISBN: 978-607-28-0554-5

Noviembre 2015



Universidad Autónoma Metropolitana
Cuajimalpa

División de Ciencias Naturales e Ingeniería

Modelación Molecular en Sistemas Biológicos

Edgar Vázquez Contreras

Departamento de Ciencias Naturales

Marco Vinicio Vázquez González

Departamento de Matemáticas Aplicadas y Sistemas

Universidad Autónoma Metropolitana

Unidad Cuajimalpa

México, D. F.

MMXIII

ISBN: 978-607-28-0554-5

Resumen

En el presente trabajo se presenta información general sobre las biomoléculas, en particular se desarrolla lo relacionado con las proteínas, iniciando con lo relacionado con los aminoácidos y posteriormente con las fuerzas no covalentes involucradas en mantener la estructura tridimensional de las proteínas; aquí se desarrolla información sobre las interacciones iónicas, los enlaces o puentes de Hidrógeno, así como el efecto hidrofóbico y las fuerzas de van der Waals o de dispersión. Esta información es relevante pues forma parte de los datos que eventualmente habrá que utilizar en los estudios *in silico*. Para esa sección se desarrolla información sobre la modelación molecular, en donde se discute sobre los modelos, su utilidad e importancia, así como los principios extremos y la predicción de equilibrios. Se incluye también una sección sobre los átomos y los modelos de la estructura electrónica, el problema electrónico, los modelos teóricos y sus niveles de teoría. Se presenta información sobre los métodos *Ab initio* y los métodos semiempíricos. Más adelante se describe una sección sobre la selección de los modelos para la aplicación química, el conjunto de bases, la optimización de geometría y los criterios de convergencia; finalmente se comenta sobre los análisis vibracionales de Mulliken. El texto incluye un glosario y la lista de las referencias bibliográficas utilizadas.

Presentación

En la actualidad, los procesos biológicos se pueden estudiar desde sus principios más elementales. Es decir, actualmente se cuenta con una infraestructura experimental, que es suficiente para poder analizar los procesos que mantienen a los eventos biológicos desde el punto de vista atómico y molecular. Esta infraestructura incluye metodologías y técnicas que derivan principalmente de la física y la química, que describieron previamente la naturaleza de los procesos en donde se llevan a cabo transformaciones de la materia que involucran la combinación de varias moléculas, o bien en donde estas transformaciones solo incluyen un tipo de molécula sin la necesidad de combinarse con otras, diferenciando así los procesos químicos de los físicos. Además, la generación de esta información biomolecular se puede estudiar utilizando poderosas herramientas computacionales, que incluyen tanto el desarrollo de los equipos necesarios, como de los algoritmos y ecuaciones para dichos análisis. En este sentido, es posible a partir de diversas teorías, utilizar estos procedimientos para modelar aquellos fenómenos de interés. La mayoría de estos modelajes, utilizan información estructural, energética y funcional de las biomoléculas que previamente se ha conocido a través de las técnicas y metodologías experimentales. Por ejemplo, a la fecha se conoce una cantidad muy elevada de coordenadas espaciales para los átomos constituyentes de un gran número de biomoléculas. Con estos datos se puede generar información básica o aplicada a partir de alimentar una gama de procesos computacionales que generan respuestas en condiciones o tiempos que no se pueden cuantificar o ensayar experimentalmente. El desarrollo de la información relacionada con esta manera de investigación, rinde continuamente frutos. Es precisamente debido a que en la actualidad una buena cantidad de respuestas se pueden obtener a través del uso de estos procedimientos, que se genera este texto con la información mínima indispensable para emprender estas investigaciones.

Edgar Vázquez Contreras,
Marco Vinicio Vázquez González.
México, D. F., noviembre de 2015

Objetivo

Al terminar la revisión de este documento, el lector tendrá un panorama general de la información biomolecular que puede utilizar para responder preguntas de orden molecular, utilizando una gama de herramientas teórico-computacionales. Esto tendrá una repercusión directa en el proceso de enseñanza aprendizaje de los procesos que suceden en la naturaleza a este nivel para alumnos de posgrado.

Índice

I. Las biomoléculas	6
1. Los monómeros de los polipéptidos.....	11
2. Las fuerzas no covalentes involucradas en mantener la estructura tridimensional de las proteínas.....	12
2.1. Las interacciones iónicas (electrostáticas)	12
2.2. El enlace o puente de Hidrógeno	13
2.3. Los puentes de hidrógeno en las biomoléculas.....	13
3. El efecto hidrofóbico.....	15
4. Las fuerzas de van der Waals o de dispersión.....	15
II. Introducción a la modelación molecular	17
1. Modelos, su utilidad e importancia en el estudio de la naturaleza	17
2. Principios extremos y la predicción de equilibrios.....	19
3. El átomo y los modelos de la estructura electrónica.....	19
4. El problema electrónico	24
5. Modelos teóricos y niveles de teoría.....	25
6. Métodos <i>Ab initio</i>	27
7. Métodos semiempíricos	29
8. Selección de modelos y aplicación a la química: ¿qué nivel de teoría usar?.....	30
9. Conjunto de bases	32
10. Optimización de geometría.....	36
11. Criterios de convergencia: ¿dónde se detiene la búsqueda de un mínimo en la superficie de energía potencial?	37
12. Análisis vibracional	38
13. Análisis de Mulliken.....	41
Glosario básico	43
Referencias bibliográficas	46

I. Las biomoléculas

Los organismos vivos mantienen los procesos que les permiten continuar en ese estado, gracias a un fenómeno denominado homeóstasis. La homeostasis es posible gracias a múltiples ajustes dinámicos del equilibrio y los mecanismos de autorregulación celular, lo que en palabras muy simples se podría explicar como el mantenimiento de los procesos vitales o el metabolismo. Para que el metabolismo se pueda llevar a cabo correctamente, las células, que son las unidades fundamentales de los procesos biológicos, deben contener un número determinado de moléculas que les son necesarias para efectuar los procesos vitales [1]. En los humanos por ejemplo, algunas de estas moléculas pueden ser sintetizadas en el interior celular, pero otras deben ser obtenidas a partir de la dieta. A estas moléculas, se les conoce genéricamente con el nombre de biomoléculas, e incluyen macromoléculas y moléculas pequeñas. Dentro de las primeras se encuentran las proteínas, los polisacáridos (azúcares), los lípidos (grasas) y los ácidos nucleicos. Por otra parte, las moléculas pequeñas se conocen como metabolitos y existen de varios tipos. En algunos textos, a las biomoléculas se les denomina como sustancias biogénicas, pero dado que actualmente es posible sintetizar algunas de ellas en el laboratorio, ese término es poco preciso en la realidad. El hecho es que la combinación de dos factores: las concentraciones correctas de las biomoléculas y la sincronía en su utilización, producen un fenotipo correcto. El fenotipo es el conjunto de características observables de un organismo, son sus rasgos, tales como su morfología y su desarrollo; así como sus propiedades bioquímicas y fisiológicas. En el fenotipo también se incluyen el comportamiento y lo que de él se derive (como por ejemplo, los tipos de nidos de las aves). En principio el fenotipo resulta de la expresión de los genes del organismo, pero los factores ambientales tienen una gran influencia en sus características. Entonces desde el punto de vista más general, el fenotipo está directamente relacionado con las biomoléculas y el ambiente en donde éstas se combinan para realizar sus funciones. A partir de esta generalización es que surge la necesidad de estudiar las características de las biomoléculas para poder entender sus funciones, sus papeles en el metabolismo y su resultado directo en el fenotipo.

En general, los procesos biológicos dependen de la expresión de la información genética; hoy en día se sabe que la información almacenada en el genoma, sirve

principalmente para producir proteínas, y son estas macromoléculas las que en general participan en la fabricación y mantenimiento de las otras biomoléculas, es decir, existen proteínas que producen a los ácidos nucleicos, hay unas proteínas que producen a los lípidos, otras que dan origen a los carbohidratos, y otras más que generan a todos los metabolitos; finalmente, también hay proteínas que producen a las otras proteínas; entonces, por detrás de cualquier proceso celular, siempre habrá un tipo particular de proteína asociado; por ejemplo en el genoma humano hay alrededor de 26,000 genes, lo que nos indica que un número de proteínas no muy lejano a ese, se puede producir en una célula humana. Lo que se sabe en la actualidad es que finalmente la relación estructura-función es determinante para que los organismo funcionen y esto empieza a nivel molecular; si la estructura de la proteína es correcta y si las condiciones del medio son las correctas, ésta puede llevar a cabo su función; por el contrario, si las condiciones del medio no son las adecuadas o la estructura de la proteína es inadecuada, la función se verá afectada. De tal forma que la homeóstasis busca preservar en todo momento tanto las concentraciones como el medio en el que reaccionan las biomoléculas.

Desde el punto de vista químico, las biomoléculas están compuestas principalmente por átomos de Carbono (C) que se combinan en diferentes proporciones con átomos de Hidrógeno (H), Oxígeno (O) y Nitrógeno (N). A estos elementos se les conoce como bioelementos primarios; el CHON constituye alrededor del 96 % de la materia viva seca. Por otra parte, existen los bioelementos secundarios, que se dividen en indispensables y variables; en los primeros tenemos al Calcio (Ca), Sodio (Na), Potasio (K), Magnesio (Mg), Cloro (Cl), Hierro (Fe), Yodo (I) y en los variables al Boro (B), Bromo (Br), Cobre (Cu), Flúor (F), Manganeso (Mn) y Silicio (Si), que en conjunto, constituyen el 3 % de la materia viva seca. Finalmente existen los elementos traza, que forman menos del 1% de la materia viva seca y son variables en los diferentes tipos de organismos [2]. La diversidad en la combinación de los bioelementos es amplia en los metabolitos, por lo cual en este texto, no se describirán a detalle; por otra parte, en las macromoléculas se pueden generalizar bastante bien las interacciones químicas que les pertenecen. Por ejemplo, en los lípidos existe una diversidad química muy amplia que va desde moléculas sencillas como los ácidos grasos, hasta moléculas complejas como el colesterol o los fosfolípidos. Pero para los otros tres tipos de macromoléculas, existe un determinado número de monómeros o unidades básicas de formación, como se puede observar en la Tabla 1; esto quiere decir que en el contexto biológico, el término macromolécula se

refiere a moléculas poliméricas y por tanto muchas de sus propiedades están relacionadas directamente con las características de los monómeros que las componen; de hecho un término correcto para referirse a estas macromoléculas, es biopolímeros [3].

macromolécula	Tipo de Monómeros	# de tipos de monómeros posibles	Tipo de enlace covalente entre monómeros
Polisacáridos	Monosacáridos	Decenas	O-Glucosídico (éter)
Proteínas	Aminoácidos	20	Peptídico (amida)
Ácidos nucleicos	Nucleótidos	8*	Nucleotídico (fosfodiéster)

Tabla 1. Información general sobre las biomoléculas. (*) 4 para ADN y 4 para ARN

Los monómeros de las macromoléculas se unen entre sí por medio de enlaces covalentes que son típicos para cada tipo de ellas (ver Tabla 1). Por ejemplo, los monosacáridos se unen mediante un enlace éter, los aminoácidos por medio de una enlace amida y los ácidos nucleicos por uno fosfodiéster. Los monómeros también se pueden ver en algunos tipos de lípidos, como por ejemplo los acilglicéridos que están compuestos de glicerol y dependiendo su grado de esterificación, de uno, dos o tres ácidos grasos. El hecho es que, independientemente de su naturaleza química, los enlaces covalentes que forman a las biomoléculas, mantienen su estabilidad general. Esta propiedad no es poca, en el laboratorio se necesitan condiciones de temperatura y pH drásticas, para lograr la ruptura de estos enlaces covalentes. Aún así, existe otro nivel de estructura que las macromoléculas pueden tener y que está determinado por las interacciones que los monómeros ya unidos en el biopolímero pueden desarrollar. Desde el punto de vista mas general, las interacciones que pueden formarse en los biopolímeros pueden ser de tres tipos, entre los residuos, así se les denomina a los monómeros de la macromolécula que están unidos por un enlace covalente; es decir que los átomos de los residuos que no forman parte del enlace covalente, pueden participar en la generación de otras interacciones químicas, con los átomos de otros residuos del biopolímero. Los átomos de los residuos de las macromoléculas, también pueden interaccionar químicamente con otras moléculas y con el solvente. En general los polisacáridos no tienen otro nivel estructural mas que el definido por los enlaces covalentes, así se forman los disacáridos y demás polisacáridos, pero no debe pasarse por alto que tienen una relación estrecha con el medio que los contiene. Por otra parte, en los ácidos nucleicos y en las proteínas, existen otros niveles estructurales, siendo más diversos en las segundas. La doble hélice

del ADN se forma a través de interacciones no covalentes denominadas puentes de Hidrógeno (Figura 1). Cada una de estas interacciones es muy débil, pero en una hebra de ADN hay miles, por lo que la molécula es muy estable, incluso la doble hélice puede doblarse sobre si misma, pero esa es su máxima variación estructural; la función del ADN es almacenar la información genética y para eso solo necesita esa estructura, la doble hélice.

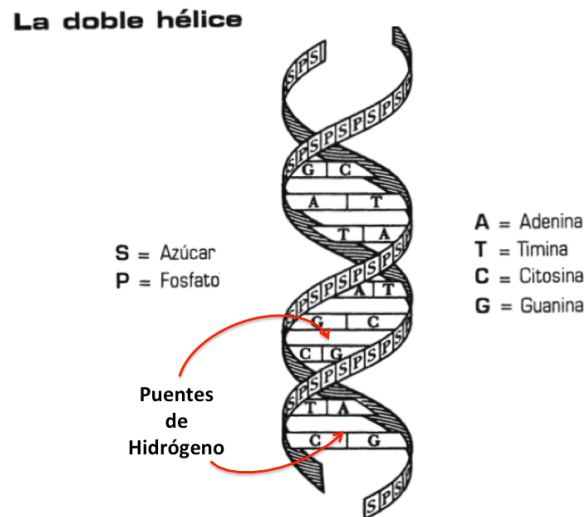


Figura 1. La doble hélice del ADN. Se muestran los componentes principales de esta biomolécula, haciendo énfasis en la posición de los puentes de hidrógeno.

Pero es en las proteínas en donde la diversidad estructural se hace muy evidente y por lo tanto se encuentran muchísimas funciones (Figura 2); búsquese una función de los seres vivos y lo que tiene por detrás es una proteína con una estructura tridimensional precisa que la lleva a cabo; al ADN lo produce un tipo de proteína como se explicó anteriormente. Las interacciones que producen esa estructura funcional son solo de unos cuantos tipos: puentes de Hidrógeno, el efecto hidrofóbicos o polares.

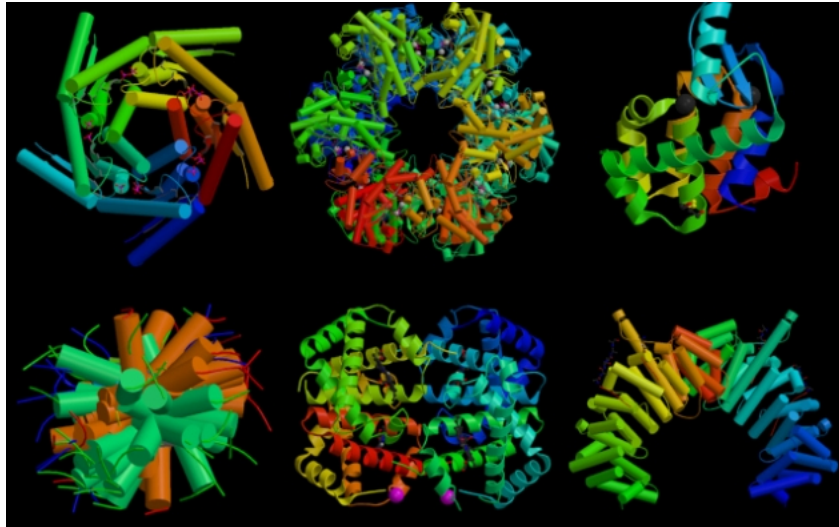


Figura 2. La diversidad estructural de las proteínas. Se muestran ejemplos de proteínas constituidas por solo alfa-hélices. Se representan en dos formas: tubos y listones.

Dado que la diversidad funcional y estructural a nivel molecular se encuentra en las proteínas, en este capítulo se hará énfasis en lo relacionado a estas biomoléculas, pero debe considerarse que esta información puede, con el debido cuidado, ser extrapolada a los otros tipos de biomoléculas.

En la actualidad existen diversas aproximaciones para estudiar las relaciones estructura función de las biomoléculas. Estos procesos van desde experimentos *in vivo*, hasta poderosos cálculos *in silico*. En este texto, se desarrolla la información que está relacionada directamente con el modelaje molecular; para utilizar estas herramientas es necesario conocer sobre las características de las interacciones químicas entre los átomos de las biomoléculas, en la siguiente sección se hace énfasis en las propiedades de las interacciones moleculares en las proteínas.

1. Los monómeros de los polipéptidos.

Polipéptido es un término que se aplica a moléculas que contienen enlaces peptídicos (Figura 4) como las proteínas. En las células, la síntesis de proteínas se lleva a cabo por los ribosomas, que son ribonucleoproteínas que unen de uno en uno los aminoácidos que forman al polipéptido. Esta unión se hace de acuerdo a las instrucciones que contienen el *ARN mensajero*. Hay 20 tipos de aminoácidos proteínicos. Desde el punto de vista fisicoquímico, cada tipo de aminoácido posee ciertas propiedades que lo hacen distintivo de los otros, como la polaridad, la hidrofobicidad, el tamaño y la presencia de cargas positivas o negativas; generalmente los aminoácidos se clasifican en función de estas características.

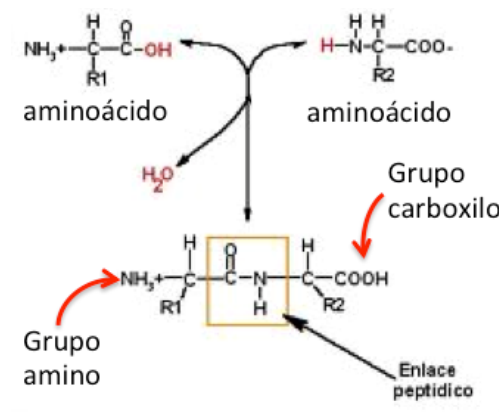


Figura 4. Los aminoácidos y el enlace peptídico. Se muestra la estructura de los aminoácidos, el evento de deshidratación, y la formación del péptido a través del enlace peptídico.

En la naturaleza ocurre un fenómeno denominado plegamiento y es vital para que los polipéptidos formados por la traducción, obtengan la función para la que están diseñados [4]. Es decir que la estructura primaria de una proteína, contiene la información para producir una estructura tridimensional que es relevante para las funciones de los seres vivos. Esta estructura tridimensional, conocida como estructura nativa, se forma por las interacciones que se forman por una parte al ponerse en contacto las cadenas laterales de los residuos de aminoácidos o al alejarse. Esta situación se produce por las afinidades o las repulsiones de los grupos químicos que poseen. La mayoría de estas interacciones son de carácter no covalente [5].

2. Las fuerzas no covalentes involucradas en mantener la estructura tridimensional de las proteínas

La correcta interacción que se desarrolla a través de las interacciones iónicas, los puentes de hidrógeno, el efecto hidrofóbico y las fuerzas de van der Waals producen la estructura tridimensional y funcional de las proteínas [6]. Estas cuatro fuerzas se relacionan de tal forma que en realidad no pueden estar totalmente separadas las unas de las otras, pero su combinación le da la estabilidad necesaria a las proteínas. De hecho, las fuerzas de van der Waals están incluidas en cada uno de los otros tres fenómenos; los puentes de hidrógeno se pueden considerar como un tipo de interacción iónica; a su vez, las interacciones iónicas pueden involucrar a los puentes de hidrógeno, o bien, el efecto hidrofóbico es, mayoritariamente, el reflejo de la suma de los puentes de hidrógeno en el solvente. Finalmente, las interacciones no covalentes son también las responsables del reconocimiento molecular, un evento fundamental para los procesos biológicos y, de dirigir y mantener la estructura tridimensional de las proteínas.

2.1. Las interacciones iónicas (electrostáticas)

Si en una región de una proteína existe un grupo iónico con carga positiva (catión), podría interactuar con otro grupo localizado en la misma proteína que posea carga negativa (anión); la proximidad de estos segmentos, produce atracción entre ellos y con ello un ordenamiento del polipéptidos. Se conoce como par iónico al complejo que se forma cuando un catión y un anión interactúan favorablemente en una solución. Sin embargo, todas las interacciones entre los átomos y moléculas son, en última instancia, electrostáticas, aún aquellas que se llevan a cabo entre moléculas no polares. La energía de estas interacciones, varía de acuerdo con la naturaleza de las cargas involucradas y con la constante dieléctrica (D) del medio. En este sentido existen dos tipos de interacciones: las de largo alcance y las de corto alcance. Cuando se trata de iones con carga neta, si la energía entre ellos decae con respecto a la distancia (r) de acuerdo con la relación $1/Dr$, se trata del primer caso; pero entre dipolos permanentes y orientados al azar, la energía decae a razón de $1/Dr^6$, lo que denota una interacción de corto alcance. Por otra parte, se sabe que la energía entre una carga y un dipolo inducido decae a razón de $1/Dr^4$. Finalmente, se debe recalcar que el valor de D del agua es de 80, mientras que

en una proteína puede tener valores entre 2 y 4.

Debido a la heterogeneidad local de las proteínas y a la interfase proteína-agua que generalmente forman, es difícil de calcular la contribución de las interacciones carga-carga de larga distancia, lo cual produce que el valor de la constante dieléctrica varíe en un rango amplio.

2.2. El enlace o puente de Hidrógeno

Un enlace de hidrógeno es la interacción electromagnética atractiva de un átomo de hidrógeno unido a un grupo polar o molécula química y un átomo electronegativo, tal como nitrógeno, oxígeno o flúor, de otra molécula o grupo químico. Es un enlace químico no covalente. Cuando las atracciones de este tipo se producen entre moléculas, se denomina intermoleculares, pero cuando se produce entre diferentes partes de una misma molécula se le llama intramoleculares [7]. La energía de un enlace de hidrógeno varía entre 5 y 30 kJ mol⁻¹, por lo que es más fuerte que una interacción de van der Waals, pero más débil que cualquier enlace covalente o iónicos. Este tipo de enlace se produce tanto en moléculas inorgánicas como agua y moléculas orgánicas como las biomoléculas.

Los enlaces de hidrógeno intermoleculares que se forman entre los hidrógenos de una molécula de agua y el oxígeno de otra, son responsables de muchas de las características de este líquido. Los enlaces de hidrógeno intramoleculares que aparecen en las proteínas y los ácidos nucleicos son en parte responsable de las estructuras tridimensionales de estas biomoléculas. También juegan un papel importante en la estructura de los polímeros, tanto sintéticos como naturales.

2.3. Los puentes de hidrógeno en las biomoléculas

El enlace de hidrógeno juega un papel importante en la determinación de las estructuras tridimensionales adoptadas por las proteínas y los ácidos nucleicos. En estas macromoléculas, uniones entre las partes de la misma macromolécula causan que se plieguen en una forma específica; esta estructura tridimensional determina la función de la

molécula. La estructura de doble hélice del ADN, por ejemplo, se debe en gran parte a los enlaces de hidrógeno que se producen entre los pares de bases que forman a las hebras complementarias, enlazándolas una con otra y permitiendo su replicación.

En las proteínas, la estructura secundaria, se debe totalmente a la formación de enlaces de hidrógeno entre los átomos de oxígeno y los hidrógenos amida de los residuos de aminoácidos que participan en el enlace peptídico formando así lo que se denomina como la cadena principal, el esqueleto polipeptídico o la secuencia primaria de la proteína. Cuando la separación de los residuos de aminoácidos que participan en un enlace de hidrógeno se produce regularmente entre las posiciones i e $i+4$, se forma una hélice alfa, esta disposición es la más frecuente en las proteínas (Figura 3 A). Cuando los enlaces de hidrógeno se forman entre residuos más próximos, es decir entre las posiciones i e $i+3$, se forma una hélice 3_{10} . Cuando dos segmentos de la secuencia se unen mediante enlaces de hidrógeno que involucran residuos alternantes de la región de la cadena participante, se forma una lámina beta (Figura 3 B).

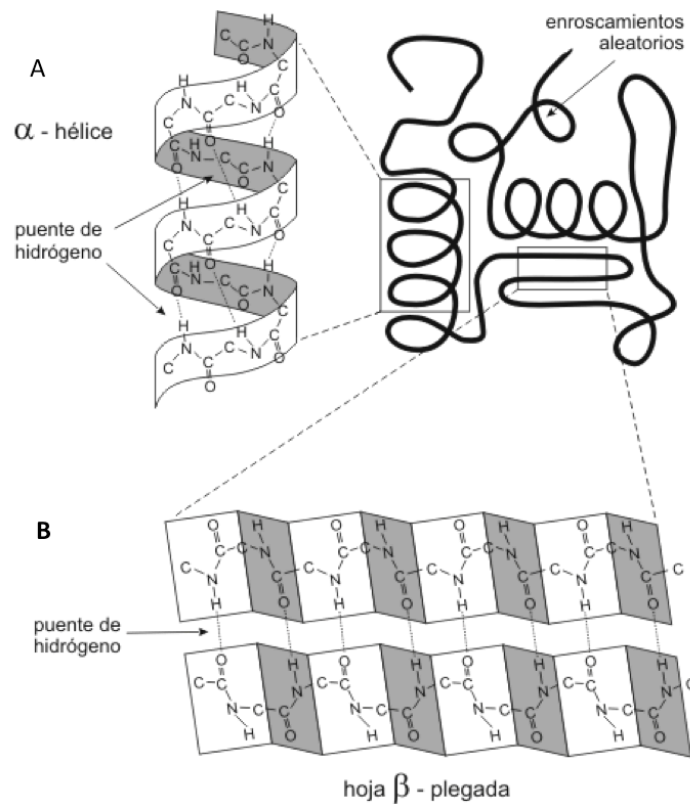


Figura 3. Los arreglos de la estructura secundaria de las proteínas. Se muestra la estructura de las alfa hélices y las beta plegadas. En la parte superior derecha, se muestra la disposición de los elementos de estructura secundaria en la proteína nativa.

Los enlaces de hidrógeno también juegan un papel en la formación de la estructura terciaria de la proteína a través de la interacción de grupos funcionales (-R), que son las cadenas laterales de los residuos de aminoácidos, estas cadenas tienen diferentes naturalezas químicas, pueden ser cargadas, polares y no polares. Son precisamente las polares las que intervienen en la generación de los puentes de hidrógeno.

3. El efecto hidrofóbico

El efecto hidrofóbico se refiere a la tendencia de los compuestos no polares, como los hidrocarburos, a transferirse desde una solución acuosa a una fase orgánica. La teoría del efecto hidrofóbico nos dice que esta segregación entre las moléculas no polares y las del agua, no se debe a que existe una gran energía de interacción entre las moléculas del soluto o a la enorme fuerza de repulsión que se desarrolla entre lo apolar y lo polar; lo que se observa es el resultado de la reorganización de la red de puentes de hidrógeno del agua en presencia de un compuesto hidrofóbico. Un compuesto no polar en el agua produce un arreglo local del entramado de los puentes de hidrógeno. Este arreglo en donde las moléculas de agua se ordenan alrededor del compuesto hidrofóbico, causa una disminución de la entropía del sistema al darle menos grados de libertad. Al extrapolar este concepto a las proteínas, un residuo no polar se dirige a la región hidrofóbica de la proteína, mientras que aquellos polares, se orientan de tal forma que puedan ejercer interacciones con las moléculas del solvente.

4. Las fuerzas de van der Waals o de dispersión

Las fuerzas de dispersión de London siguen la ley del recíproco de la sexta potencia de la distancia: $U = A/r^{12} - B/r^6$, en donde A y B son constantes de proporcionalidad, con valores específicos para cada pareja de donadores-aceptores. Se observa que la energía asociada a la atracción entre dos moléculas es función de $1/r^6$. Nótese que esta relación es característica de la interacción entre dipolos.. Las fuerzas de van der Waals son un tipo de fuerza de dispersión como lo son las de London.

Cualquier molécula, aún siendo no-polar, puede presentar una separación de cargas

como resultado de las fluctuaciones locales de la densidad electrónica en un instante determinado. Debido a que la interacción entre las moléculas depende de la inducción del dipolo, la polarización es un factor importante en la energía de la interacción entre dos átomos cualquiera. La polarizabilidad electrónica (α) se define como el cociente del momento dipolo inducido (p) de un átomo entre el campo eléctrico (E) que induce este momento dipolo: $p = \alpha E$. La polarizabilidad electrónica tiene las unidades del SI de: $C \cdot m^2 \cdot V^{-1} = \text{Å}^2 \cdot s^4 \cdot kg^{-1}$, pero frecuentemente se expresa como volumen de polarizabilidad, con unidades de cm^3 o $\text{Å}^3 = 10^{-24} cm^3$. Es importante aclarar que la polarizabilidad no tiene que ver con la electronegatividad, sino con la deformación de la nube electrónica: mientras más grande y alejada del núcleo sea la nube electrónica y su capacidad de deformarse, mayor será el valor de α .

La polarizabilidad está asociada con la constante B en la ecuación de las fuerzas de dispersión de London, y describe la fuerza de atracción entre dos átomos que están separados a una distancia de 1 Å. Así pues, mientras mayor sea la polarizabilidad, y en consecuencia el tamaño de la nube electrónica de un átomo, mayor será la energía de autointeracción o autoatracción (B). Por ejemplo, el grupo -O- tiene una polarizabilidad de $0.63 mL \cdot 10^{-24}$ y un potencial de autoasociación de $220 kcal \cdot \text{Å}^6/mol$, mientras que para el -CH₂- los valores son 1.80 y 1160, respectivamente. Esto nos indica que el grupo metileno prefiere interactuar con moléculas similares que con el oxígeno.

Aunque todas estas fuerzas de atracción son débiles y las fuerzas de van der Waals son bajas, todas son aditivas y pueden hacer contribuciones significativas a la asociación e interacción cuando se suman dentro de una molécula.

Las consecuencias de la participación de las interacciones iónicas (electrostáticas), de van der Waals, los puentes de hidrógeno y el efecto hidrofóbico son evidentes cuando se aprecia la estructura tridimensional de una proteína.

II. Introducción a la modelación molecular

1. Modelos, su utilidad e importancia en el estudio de la naturaleza

La observación es la primera piedra en los cimientos de la ciencia y su herramienta más poderosa: el método científico. A través de ella tomamos conciencia de la ocurrencia regular de los fenómenos naturales y desencadena un torrente de actividad intelectual: curiosidad, imaginación y la poderosa necesidad de explicar lo que se ha observado. Sin embargo, la sola contemplación de los fenómenos naturales no es suficiente para descubrir todos los principios involucrados. Comúnmente la complejidad de los sistemas en estudio no permite separar adecuadamente los procesos que ocurren en su interior, impidiendo racionalizarlos en la forma de leyes y principios. Solamente en casos excepcionales, estas leyes de la naturaleza, han podido deducirse a partir de la observación directa.



Johannes Kepler (1571-1630)

Un ejemplo notable. La excepción que confirma la regla, es el descubrimiento de las leyes del movimiento de los planetas, hecho por el astrónomo Johannes Kepler (publicado en su obra *Astronomia Nova* en 1609). Comúnmente, el investigador crea condiciones

excepcionales (en un laboratorio), para permitir la separación de los diferentes procesos que ocurren en los sistemas naturales y, entonces, realiza experimentos. Este método (experimental o empírico) permite estudiar el efecto de las condiciones del experimento sobre algunas variables del sistema, escogidas apropiadamente, y con ellos ganar comprensión de los cambios que ocurren en las variables debido a las condiciones experimentales, los cambios en otras variables y de su comportamiento a través del tiempo; esta información suele codificarse en forma de ecuaciones diferenciales o integrales (lo que se llama un *modelo matemático* del sistema), y luego puede llegarse a inferir algo sobre los principios naturales que gobiernan el comportamiento del sistema. Posteriormente, los modelos matemáticos pueden emplearse para hacer predicciones sobre el sistema estudiado (pero también, dentro de ciertos límites, sobre sistemas donde ocurran procesos físicos semejantes), que luego pueden compararse con los experimentos (un procedimiento llamado *validación*), y con ello mejorar aún más el modelo.

La modelación molecular tiene como objeto de estudio un sistema formado por átomos o moléculas (una o muchas) y busca construir una representación de las interacciones entre ellas, con el objeto de predecir ciertas propiedades físicas y químicas del conjunto; útiles para describir el comportamiento químico del sistema (reactividad química), que conducen a la formación de nuevas moléculas, así como caracterizar los cambios en la estructura electrónica de las moléculas de reactivos y productos, conforme avanza una reacción química dada. Estas herramientas pueden utilizarse para emprender el estudio sistemático de problemas de interés en las ciencias biológicas, como las interacciones entre fragmentos moleculares (proteínas, ácidos nucleicos) o moléculas completas, que ocurren en procesos de reconocimiento molecular: en la replicación de ADN, la síntesis de proteínas, la competencia por los sitios de unión en una enzima, o sobre receptores de la superficie de la célula, entre muchos otros.

2. Principios extremos y la predicción de equilibrios

Los átomos y moléculas son sujetos a fuerzas cuyas interacciones determinan el estado de un sistema; tales fuerzas pueden describirse en términos de las dos grandes tendencias observadas en la naturaleza (principios): por un lado, los átomos y moléculas reaccionan entre sí, dando lugar a cambios de conformación, a la disociación de las uniones entre ciertos átomos (enlaces) o a la formación de nuevas uniones, sufriendo también otros procesos físicos y químicos (cambios de fase, solvatación, etcétera) que en conjunto provocan que una propiedad llamada **energía** del sistema alcance su mínimo valor posible o, por otra parte, que otra cantidad llamada **entropía** llegue a su máximo valor posible (ambos extremos corresponden a un tipo de equilibrio). Esto implica que si podemos construir modelos matemáticos de la energía (o de la entropía) de un sistema, podríamos predecir la tendencia al equilibrio de un sistema particular, mediante el cálculo de los mínimos (o máximos) de ciertas funciones matemáticas [8].

3. El átomo y los modelos de la estructura electrónica

La concepción moderna del átomo lo presenta como un núcleo, formado de protones y neutrones, y una envoltura electrónica. La masa de las partículas que forman el núcleo, protones y neutrones, es semejante (1.6725×10^{-27} Kg), distinguiéndose en la carga eléctrica que llevan: el protón tiene una carga positiva de magnitud $+e$, mientras que el neutrón tiene carga 0. e es una constante física, igual al valor absoluto de la carga eléctrica de un electrón: $e = 1.60 \times 10^{-19}$ C. El electrón tiene carga $-e$ y su masa es, aproximadamente, 1/1836 de la masa del protón. Los átomos, inicialmente, son eléctricamente neutros, por lo que se establece la necesidad de que en un átomo dado haya un número de electrones en su envoltura electrónica, igual al número de protones en el núcleo; a este número se le denota por Z , el **número atómico**.

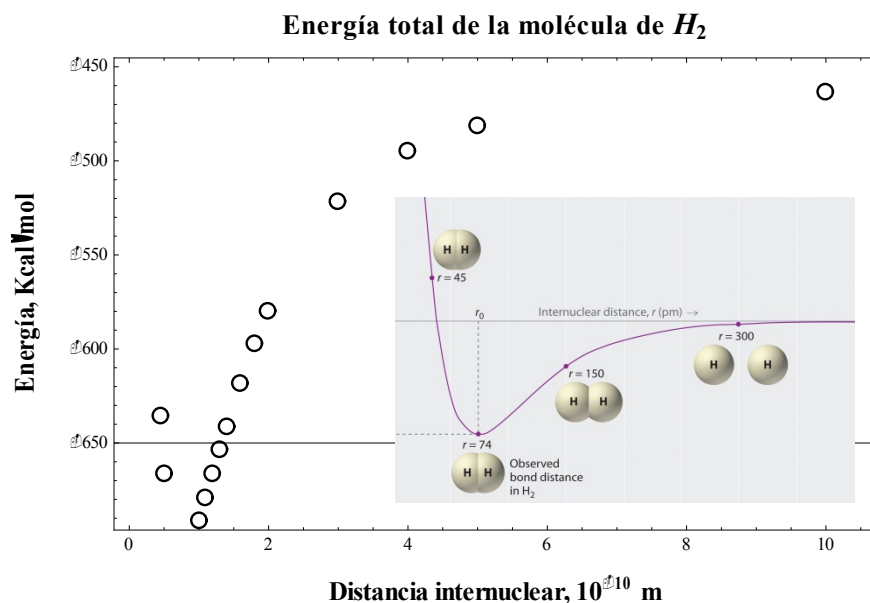


Figura 4. Diagrama de la energía total del dímero $H-H$, como función de la distancia entre los núcleos de hidrógeno, calculado al nivel de teoría HF/6-311+G(2d,p). En el inserto se muestra la interpretación de la interacción química entre los átomos para formar una molécula de H_2 (mínimo en la curva de energía). Observe el incremento sostenido de la repulsión nuclear cuando los átomos se acercan más allá de la distancia de equilibrio (distancia de enlace).

Entre los electrones y los protones existe una fuerza de atracción de naturaleza puramente electrostática, cuantificada por la ley de Coulomb,

$$F = -K \frac{q_1 q_2}{r^2} = -K \frac{e^2}{r^2}$$

dicha fuerza es inversamente proporcional al cuadrado de la distancia entre las partículas. Esta interacción es sumamente poderosa a distancias cortas, la razón por la cuál los electrones no “caen” al núcleo, es parecida a la que mantiene en movimiento los planetas alrededor del sol, contra la fuerza atractiva de la gravedad, y es que —al igual que los planetas— los electrones se mueven a velocidades prodigiosas, en ocasiones cercanas a la de la luz, alrededor del núcleo, girando permanentemente en torno a su propio eje. Este movimiento perpetuo contribuye a disipar un poco de la atracción electrostática para no caer al núcleo, pero no lo suficiente para alejarse de él, cancelando su influencia, a menos que reciban energía extra del exterior. En ese caso, los electrones más alejados del núcleo (en la llamada **capa de valencia**) son susceptibles de ser separados del núcleo, lo cuál puede permitir que el sistema atómico alcance un nuevo y más favorable estado de equilibrio. Un sistema N -electrónico que ha perdido un electrón, quedando con $N-1$

electrones, tiene una carga positiva no equilibrada en su núcleo, es decir, que el átomo neutro se ha convertido en un **ión** con carga formal (en este caso, con una carga positiva, un **catión**). También ocurre lo opuesto, la absorción de un electrón adicional, lo cuál produce un sistema de $N + 1$ electrones, que aloja una carga negativa no equilibrada por los protones del núcleo; a este *ión* se le llama **anión**.

Por las razones discutidas anteriormente los electrones que forman la envoltura electrónica de un átomo dado, se posicionan alrededor del núcleo en diferentes niveles, con energías distintas: una estructura electrónica característica de cada átomo.

Una molécula puede considerarse como un conglomerado de núcleos atómicos, con una envoltura electrónica conformada por los electrones provenientes de los átomos individuales, en una nueva estructura electrónica *molecular* (nuevos niveles, con sus energías distintivas). La química que conocemos trata de las interacciones entre los electrones de las capas, o niveles energéticos, más externos (capas de valencia) de los átomos. La estructura electrónica de una nueva molécula se modela a través de la energía total del sistema; de la misma forma que la ecuación de Newton es la base para modelar la energía de un sistema macroscópico, la **ecuación de Schrödinger** es la ecuación fundamental para describir la energía de los sistemas electrónicos (cuánticos).



Erwin Schrödinger (1887-1961)

En 1926, el físico Erwin Schrödinger, postuló que “toda la información acerca de las propiedades de un sistema cuántico, pueden derivarse de una o varias funciones, ψ , llamada **función de onda** del sistema” [9]. Las funciones de onda se obtienen al resolver la ecuación de Schrödinger para el sistema de interés. Las ecuaciones de Schrödinger para un sistema cuántico dado no son fáciles de plantear, ni de resolver, excepto para casos aislados (los sistemas más simples, de hecho, como el átomo de hidrógeno). También llamadas **ecuaciones de onda**, las ecuaciones de Schrödinger son ecuaciones diferenciales parciales de segundo orden. Para un sistema estacionario (donde ψ no depende del tiempo), la ecuación de Schrödinger puede escribirse de una forma aparentemente sencilla,

$$\hat{H}\psi = E\psi$$

El símbolo \hat{H} es un **operador** de energía total (cinética y potencial), que se describirá con mayor detalle más adelante en esta obra. Una vez conocidas las soluciones de esta ecuación, toda la información sobre las propiedades físicas del sistema (energía cinética, energía potencial, momento lineal, etcétera), se obtienen mediante operaciones sobre las funciones de onda ψ ; estas operaciones se definen mediante operadores mecanocuánticos específicos para cada propiedad.

La función de onda es una función de muchas variables (Las tres coordenadas espaciales de cada uno de los M núcleos y los N electrones de una molécula, es decir, $3M + 3N$ coordenadas) de valores complejos. Sin embargo, el cuadrado de su valor absoluto, $|\psi|^2$, sí es una función de valores reales y se interpreta como una densidad de probabilidad (interpretación debida a Max Born, matemático y físico de origen alemán).



Max Born (1882-1970)

4. El problema electrónico

La evidencia experimental, principalmente en óptica y espectroscopía, revela un carácter dual de la naturaleza del electrón: algunas propiedades se explican considerando a los electrones como partículas sometidas a la acción de fuerzas —principalmente electrostáticas— (enfoque clásico); otras propiedades, sin embargo, sólo se entienden si se considera al electrón como una onda (enfoque no-clásico). En este último caso, las propiedades físicas (observables) del electrón se obtienen a partir de **operadores hermitianos** que se aplican a (operan sobre) la función de onda [10, 11].

La función de onda $\psi(\vec{r})$ en sí misma, no tiene interpretación física (no corresponde a ningún observable físico). Se debe a Max Born la interpretación del cuadrado de la función de onda ($|\psi(r)|^2$) como una densidad de probabilidad. De este modo, la probabilidad de encontrar un electrón dado en todo el espacio debe estar normalizada a la unidad

$$\int_V dV |\psi(\vec{r})|^2 = 1 \quad (1)$$

con los límites apropiados para la integral de volumen (r_i es el vector posición de la i -ésima partícula). Esta función de onda es desconocida y se postula que debe satisfacer una ecuación de valores propios

$$H\psi(\vec{r}) = E\psi(\vec{r}) \quad (2)$$

que tiene la forma general

$$\hat{A}\psi(\vec{r}) = a\psi(\vec{r}) \quad (3)$$

recordemos que el valor propio a es el valor que obtendríamos en una medición de la propiedad física A asociada al operador hermitiano \hat{A} .

De particular importancia, la ecuación (2), donde \hat{H} es el operador Hamiltoniano de energía mecánica total (cinética y potencial), se conoce como **ecuación de Schrödinger**.

El Hamiltoniano para un sistema de M núcleos y N electrones (en unidades atómicas) se escribe explícitamente de la forma siguiente,

$$\hat{H} = -\frac{1}{2} \sum_{\alpha=1}^M \nabla_{\alpha}^2 - \frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{\alpha=1}^M \sum_{i=1}^N \frac{Z_{\alpha}}{r_{i\alpha}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{\alpha=1}^M \sum_{\beta>\alpha}^M \frac{Z_{\alpha} Z_{\beta}}{r_{\alpha\beta}} \quad (4)$$

∇_{α}^2 , es el laplaciano respecto a las coordenadas del α -ésimo núcleo; ∇_i^2 opera respecto a las coordenadas del i -ésimo electrón; $r_{i\alpha}$ es la distancia entre el i -ésimo electrón y el α -ésimo núcleo; r_{ij} es la distancia entre los electrones i -ésimo y j -ésimo; $r_{\alpha\beta}$ es la distancia entre los núcleos α -ésimo y β -ésimo cuyas cargas nucleares son Z_{α} y Z_{β} respectivamente; en los dos últimos términos, las sumas corren sobre todos los pares no repetidos. El operador Hamiltoniano (4), puede escribirse alternativamente en notación de operadores,

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} + \hat{V}_{NN} \quad (5)$$

cuyos términos son: la energía cinética de los M núcleos (\hat{T}_N), la energía cinética de los N electrones (\hat{T}_e), el potencial (atractivo) núcleo-electrón (\hat{V}_{Ne}), y los potenciales (repulsivos) electrón-electrón (\hat{V}_{ee}) y núcleo-núcleo (\hat{V}_{NN}).

La ecuación de Schrödinger, ecuaciones (2) y (4), independiente del tiempo, es en general irresoluble analíticamente, excepto para el átomo de hidrógeno, el más ligero de todos los átomos.

5. Modelos teóricos y niveles de teoría

La energía de un sistema molecular depende tanto de la naturaleza de cada una de las partículas que lo constituyen (masa, carga nuclear, número de electrones), como de sus

coordenadas espaciales (x, y, z) . Bajo la **aproximación de Born-Oppenheimer**, también llamada **aproximación de núcleos fijos**, el modo en que varía la energía del sistema con respecto a dichas coordenadas se describe mediante una función comúnmente conocida como *superficie de energía potencial* (asociada a una *función de onda*). Para cierto estado electrónico, un mínimo en esta superficie es particularmente significativo; las coordenadas de este punto están asociadas a una configuración particular del sistema —con longitudes de enlace y ángulos internos, muy bien definidos—. La energía correspondiente, en los métodos **SCF**, es interpretada *variacionalmente*¹ como la mejor aproximación a la energía del **estado basal** del sistema y la función de onda como la mejor representación de dicho estado. Por ello, una optimización completa de geometría (con el propósito de localizar el conjunto de coordenadas atómicas que minimiza la función de energía del sistema) es el paso inicial que precede a cálculos precisos de energías, frecuencias de vibración, entalpías de formación, entre otros datos de interés, para entender el comportamiento químico del sistema (reactividad química).

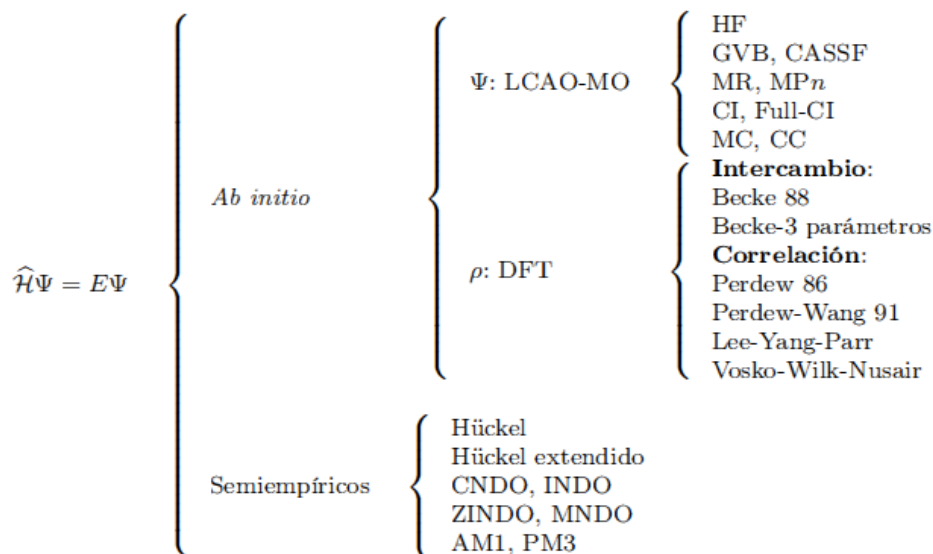


Tabla 2. Un panorama esquemático de algunos de los métodos de resolución de la ecuación de Schrödinger. La diversidad de aproximaciones sirve a propósitos bien definidos, como mejorar sustancialmente la precisión de las predicciones de los modelos o alcanzar un óptimo balance entre la demanda de precisión y la disponibilidad de recursos computacionales. Los métodos conocidos como multirreferenciales, utilizan funciones de onda multideterminantes [12].

Resolver la ecuación de Schrödinger, no es una tarea simple, y la búsqueda de métodos

¹ Una cierta función dará un valor de la energía que es una cota superior para la energía mínima, correspondiente al estado basal de dicho sistema. Bajo este criterio, podemos mejorar la función del sistema, si las energías calculadas en cada iteración son cada vez menores, hasta algún nivel aceptable.

de solución condujo al desarrollo de diversas aproximaciones (llamadas *niveles de teoría*) [13-15]. Éstas se desarrollan en dos vertientes importantes (véase el Tabla 2):

- **Métodos *Ab initio***. Son métodos con un fundamento teórico sólido. Su estructura interna es independiente de mediciones experimentales, al menos hasta la determinación de los valores de algunas constantes universales. Estos se dividen a su vez en dos grandes aproximaciones:
 - **Métodos basados en la función de onda, $\Psi(r)$** . Aquí están los métodos variacionales como: HF restringido (RHF), HF no-restringido (UHF), interacción de configuraciones (CI), interacción completa de configuraciones (full CI) —sólo para sistemas muy reducidos— y los métodos basados en la teoría de perturbaciones: Møller-Plesset de segundo orden (MP2) y superiores (MP3 y MP4), entre otros [13].
 - **Métodos basados en la densidad electrónica, $\rho(r)$** . Esto es, desarrollados a partir de la teoría de los funcionales de la densidad [16-18].
- **Métodos semiempíricos**. Basados en la teoría de Hartree-Fock (HF) y el método de combinación lineal de orbitales atómicos (CLOA), utilizan criterios derivados de determinaciones experimentales para introducir simplificaciones importantes en los cálculos. Ejemplos de éstos son: el método de Hückel, Hückel extendido, AM1, PM3, anulación completa del traslape diferencial (complete neglecting of the differential overlapping, CNDO), anulación intermedia del traslape diferencial (intermediate neglecting of the differential overlapping, INDO) [19].

6. Métodos *Ab initio*

Para estudiar un sistema mediante métodos teóricos tiene que definirse un modelo. De acuerdo con Hehre y colaboradores [20], *“Tratándose de la resolución aproximada de la ecuación de Schrödinger, existen dos tipos de acercamiento conceptual, ampliamente diferentes: el primero, consiste en el examen de cada problema particular empleando el más alto nivel de teoría disponible hasta ese momento para un sistema de ese tamaño; en el segundo, un nivel de teoría se define claramente [13], luego se le aplica uniformemente a sistemas moleculares de todos los tamaños, hasta el límite impuesto por los recursos computacionales disponibles hasta entonces. Una teoría de este tipo, si se prescribe de*

modo único para cualquier configuración de núcleos y para cualquier número de electrones, podrá llamarse modelo teórico, dentro del cual todas las estructuras, energías y otras propiedades físicas pueden ser exploradas...” Dentro de este contexto, la comparación sistemática del modelo teórico empleado contra resultados experimentales conocidos es un mecanismo de validación. Si las comparaciones son favorables, el modelo gana confiabilidad como herramienta de predicción en situaciones en que la información experimental no está disponible directamente.

Aquel que aspira a comprender los fenómenos químicos encuentra que, en principio, la solución a la ecuación de Schrödinger es la llave que le conduciría a la predicción cuantitativa de la mayoría de ellos. Tales predicciones tendrían como único sustento experimental la determinación de los valores de un reducido número de constantes físicas universales (como la carga del electrón, la constante de Boltzmann, la constante de Planck, la permisividad en el vacío). Un procedimiento semejante es lo que se llama una aproximación desde *primeros principios (ab initio)* a la química.

En aplicaciones prácticas, sin embargo, la ecuación de Schrödinger tiene que ser reemplazada por modelos matemáticos aproximados, para los cuales la posibilidad de resolución es razonablemente accesible.

Los métodos se dividen en dos grandes grupos, como se muestra esquemáticamente en el **Tabla 2**. El primero utiliza una función de onda y la teoría de Hartree-Fock [10] (incluyendo los métodos de campo autoconsistente, SCF, y de combinación lineal de orbitales atómicos, CLOA) como eje de sus procedimientos. El segundo gran grupo deriva sus propios métodos a partir de la densidad electrónica como variable fundamental, enfoque validado por la teoría desarrollada a partir del trabajo de Hohenberg y Kohn (1964) [16]. En este último caso, se utiliza el esquema de Kohn-Sham (1965) [17], que a su vez incluye un método iterativo autoconsistente.

7. Métodos semiempíricos

Los métodos semiempíricos de cálculo resuelven la ecuación de Schrödinger utilizando ciertas aproximaciones. Para describir y acortar dichos cálculos, los métodos semiempíricos realizan muchas simplificaciones que tienen como efecto positivo la economización de recursos de cómputo, poniendo al alcance de un sistema, de modesta capacidad, el estudio de sistemas de gran tamaño (como en la investigación de las propiedades farmacológicas de compuestos de origen natural). Recurren, por ejemplo, al cálculo sólo de los electrones de valencia, desestimando las integrales para ciertas interacciones, utilizando funciones base no-optimizadas de orbitales atómicos y utilizando parámetros obtenidos a partir de determinaciones experimentales (calibración de métodos). El uso de parámetros experimentales elimina la necesidad de efectuar el cálculo de ciertas cantidades y de corregir los errores que resultan del uso de aproximaciones [15].

Usados profusamente antes de las computadoras de bajo costo y alto rendimiento, los métodos semiempíricos NDO (Neglect of Differential Overlap, *Anulación del Traslape Diferencial*), consisten en la eliminación de gran parte de las integrales que describen las interacciones interelectrónicas. Un ejemplo de éstos es el método CNDO —el más simple de los métodos SCF— que desestima el cálculo de todas las integrales de traslape diferencial, mientras que otros métodos NDO (por ejemplo: INDO, MNDO/3, MNDO, y AM1) desestiman sólo algunas integrales. El método INDO (Intermediate Neglecting of Differential Overlapping) es una mejora del CNDO que considera ciertas repulsiones entre electrones del mismo átomo. ZINDO/1 y ZINDO/S son versiones modificadas de INDO. ZINDO/S reproduce transiciones espectroscópicas en el UV visible cuando se utiliza con CI (Interacción de Configuraciones) en estados no excitados.



Erich Hückel (1896-1980)

El método de Hückel es el primero de los métodos aproximados de orbitales moleculares, formulado en 1931, es de hecho el primer método ZDO (*Zero Differential Overlapping*, o traslape diferencial nulo) [19]. En un principio, la teoría de Hückel fue concebida para explicar la naturaleza no-aditiva de ciertas propiedades de los compuestos aromáticos y estaba limitada a los sistemas conjugados de electrones π . De poca aplicación en la actualidad como herramienta computacional de investigación, la teoría original de Hückel tuvo extensiones (teoría de Hückel extendida) que hasta hoy, todavía proporcionan aproximaciones cualitativas a la estructura electrónica de moléculas de cierta complejidad, que resultan muy útiles.

8. Selección de modelos y aplicación a la química: ¿qué nivel de teoría usar?

La química computacional representa la implementación práctica de los métodos de cálculo de estructura electrónica. La demanda de recursos de cómputo (tiempo de procesamiento, espacio de memoria y espacio de almacenamiento) de un problema típico de la química, puede crecer considerablemente, convirtiendo la disponibilidad de recursos en un criterio clave para la selección de métodos y modelos.

Siguiendo a Young (2001) [14], quien ilustra la estimación del costo computacional aproximado de un problema típico de capa cerrada en el método de Hartree-Fock, introduciremos el concepto de *complejidad temporal* y contrastaremos la demanda de

recursos de varios modelos teóricos.

En un cálculo de Hartree-Fock con N orbitales, al final deben sumarse las N energías orbitales, con lo que se efectúan N operaciones de suma; las integrales bi-electrónicas que contabilizan los efectos coulómicos y de intercambio requieren N^4 operaciones; un algoritmo común de inversión de matrices realiza N^3 operaciones; además, el sistema debe realizar algunas operaciones rutinarias (inicializar variables, reservar direcciones de memoria, etcétera) que son independientes del problema, lo cual toma unas C operaciones. El tiempo de procesamiento total requerido para realizar un cálculo típico de Hartree-Fock aumenta en proporción a $N^4 + N^3 + N + C$; no obstante, para un valor de N suficientemente grande, el término de cuarto grado es mucho mayor que la suma de los otros tres términos.

Se dice entonces, que el método de HF posee una *complejidad temporal* de orden N^4 , o bien, $O(N^4)$. El orden de complejidad temporal sirve para estimar el consumo de recursos computacionales, así como el tiempo necesario para completar un cálculo, $t_c \propto O[N^k]$, con cierto tamaño de la base, en un nivel de teoría dado.

Método	Orden de complejidad
TFD	N
MM	M^2
DM	M^2 o L^6
Semi empíricos	N^2 a N^3
HF	N^2 a N^4

Tabla 3. Orden de complejidad temporal de diversos métodos y niveles de teoría (adaptado de Young, 2001). M es el número de átomos; L , la arista de una caja donde se incluyen las moléculas en un cálculo con condiciones periódicas a la frontera; y, N es el número de orbitales usados en el cálculo.

La Tabla 3, presenta el orden de complejidad temporal para algunos métodos de uso corriente: TFD, funcionales de la densidad; MM, mecánica molecular; DM, dinámica molecular; HF, Hartree-Fock [14].

Volviendo a los métodos *ab initio*, la interpretación de Born del cuadrado de la función de onda (que en sí misma no es un observable del sistema) como una *densidad de*

probabilidad, perfectamente observable, abre en principio la noción de que el problema electrónico podría describirse en términos de esta propiedad. No obstante, la primera aproximación exitosa a la solución del *problema electrónico* consistió en la búsqueda de las funciones, $\psi_i(\vec{r})$, y las constantes, a_i , asociadas a un operador mecanocuántico \hat{A} mediante una expresión de la forma

$$\hat{A}\psi_i(\vec{r}) = a_i\psi_i(\vec{r})$$

una ecuación de valores propios, cuyas soluciones forman un **conjunto completo** y **ortogonal de funciones**, mismo que puede utilizarse para expandir la función de onda aproximada, de algún estado del sistema

$$\psi(\vec{r}) = \sum_{i=0}^{\infty} c_i \psi_i(\vec{r})$$

La función de onda $\psi(\vec{r})$ contiene, pues, toda la información de nuestro sistema, accesible a nuestra percepción sólo al operar sobre ella con un operador mecanocuántico apropiado, asociado a un número real (en principio accesible a las determinaciones experimentales) mediante una ecuación de valores propios.

Menos evidente, sin embargo, es que todas las propiedades del sistema puedan describirse a partir de las densidades mono y bi electrónicas ($\rho(\vec{r}_1)$ y $\rho(\vec{r}_1, \vec{r}_2)$, respectivamente). Por ello, la teoría de funcionales de la densidad gozó en sus principios de una aceptación gradual, hasta que se publicaron los trabajos pioneros de Hohenberg-Kohn (1964) [16], y Kohn-Sham (1965) [17], donde se ofreció a la comunidad de químicos teóricos una herramienta poderosa y bien fundamentada para atacar el problema electrónico en átomos y moléculas, utilizando la densidad electrónica como variable principal[18].

9. Conjunto de bases

Una función de onda aproximada (función de onda de prueba) para un sistema polielectrónico, usualmente se representa en términos de una base orbital finita, arbitraria. Los métodos basados en la función de onda son sensibles a la calidad de la base

utilizada para construir la aproximación de la función de onda del sistema; así mismo, los métodos de cálculo SCF que dependen de una aproximación inicial de los orbitales moleculares también son afectados por la calidad de la base.

Como la función de onda de prueba puede ser cualquier tipo de función, y no únicamente combinaciones lineales de orbitales atómicos hidrogenoides, John C. Slater propuso utilizar funciones de la forma

$$S_{nlm}(r, \theta, \phi) = N_{nl} r^{n-1} e^{-\zeta r} Y_l^m(\theta, \phi) \quad (6)$$

llamados *orbitales de Slater* [21], donde N_{nl} es una constante de normalización, y Y_l^m representa la parte angular (armónicos esféricos) de la solución a la ecuación de Schrödinger. A diferencia de los orbitales hidrogenoides, el parámetro ζ se escoge arbitrariamente y no se restringe al valor Z/n , lo cual proporciona mayor flexibilidad en los métodos variacionales.

No obstante, los cálculos para sistemas polieletrónicos se complican cuando los orbitales atómicos asociados mediante las integrales bielectrónicas, se centran en átomos distintos. De acuerdo con A. R. Leach (1996) [19], al utilizar **orbitales atómicos hidrogenoides**, las integrales que involucran uno o dos centros, como $(\mu\mu|\mu\mu)$, $(\mu\mu|\nu\nu)$, $(\mu\nu|\nu\nu)$ y $(\mu\nu|\mu\nu)$, se calculan de forma relativamente sencilla. Al emplear orbitales del tipo de Slater (**STO** por sus siglas en inglés, *slater type orbital*), el cálculo de integrales con tres y cuatro centros se vuelve posible, siempre y cuando los orbitales atómicos se localicen sobre el mismo átomo. Sin embargo, este tipo de integrales se vuelven bastante difíciles si los orbitales atómicos se centran en átomos distintos. Por ello, es práctica común en los métodos de cálculo numérico, reemplazar las funciones de Slater por otras basadas en funciones gaussianas.

Una función base de tipo gaussiana (centrada en el A -ésimo núcleo atómico en R_A) tiene la forma general

$$\phi(\alpha, r - R_A) = x^a y^b z^c \exp(-\alpha |r - R_A|^2) \quad (7)$$

en donde a , b y c , son números enteros. El *orden* de la función gaussiana está dado por la suma de los exponentes de x , y y z . El parámetro α determina la *amplitud* de la

curva gaussiana, en una relación recíproca, esto es, valores pequeños de α dan amplitudes grandes de la curva representada por $\phi(\alpha)$. El reemplazo de cada orbital atómico (STO) por una combinación lineal de funciones gaussianas (GTO)

$$\phi_{\mu} = \sum_{i=1}^L d_{i\mu} \phi_i(\alpha_{i\mu}) \quad (8)$$

mejora significativamente la representación, y por lo tanto los resultados, con la concomitante reducción del error [19].

En una **expansión gaussiana**, los coeficientes de la combinación lineal y el mismo exponente de las gaussianas, son parámetros que pueden variarse durante el cálculo en los métodos; en éste tipo de cálculo, las gaussianas utilizadas se llaman *no-contraídas* o *primitivas*. Esta modalidad de uso de las gaussianas es exigente en términos de recursos computacionales. Es más común la utilización de gaussianas *contraídas*, esto es, funciones donde los coeficientes de la contracción y los exponentes están predeterminados y se mantienen constantes durante los cálculos.

Una **base mínima** contiene apenas suficientes funciones para representar los orbitales llenos en cada átomo. Llevado a la práctica, el conjunto base mínimo contiene todos los orbitales atómicos de una capa dada. Los conjuntos base mínimos denotados como STO- n G, utilizan n funciones gaussianas para cada orbital. El uso de bases mínimas en los métodos conlleva un número de deficiencias: los elementos de un nivel energético se describen utilizando el mismo número de funciones base, con lo que se ignora el mayor número de electrones de los elementos hacia el final del periodo; una base mínima no puede representar ciertas características de la distribución de los electrones en las moléculas que, en rigor, no se ajustan a la simetría esférica.

Estos problemas se abordaron tomando más de una función para cada orbital; en este sentido, las bases de tipo *doble zeta* duplican el número de funciones utilizadas en la base mínima, y además con una combinación lineal de una función *contraída* y una función *difusa*. Las **funciones difusas** se añaden a los conjuntos base para describir mejor el comportamiento de la distribución electrónica, en particular de especies cuya nube de electrones tiene tendencia a expandirse (como en el caso de especies aniónicas), deben tener el mismo tipo de momento angular que las funciones presentes

de antemano en la base, pero con exponentes menores a los de cualquiera de ellas, para describir mejor la región de valencia [22]. Las **funciones contraídas** se forman agrupando algunas de las funciones del conjunto base original, por norma a las funciones no contraídas se las llama *primitivas*; a estos conjuntos base contraídas se les exige que contengan un número de funciones significativamente menor que el número de primitivas en el conjunto original, con el objeto de reducir el espacio de almacenamiento de las integrales bielectrónicas (y en menor grado, las matrices que se calculan en el método, como el Hessiano, la matriz de densidad, etcétera), estos conjuntos se obtienen utilizando métodos de contracción que garanticen que los resultados obtenidos con ellas sean al menos tan buenos como los que se obtienen con las bases no contraídas [23-25] Durante el procedimiento SCF se calculan los coeficientes de ambos tipos de funciones, lo cual asegura que se obtiene una representación más contraída o más difusa, según se requiera, para cada tipo de orbital.

Otra vía para solucionar los problemas de la base mínima consiste en duplicar sólo el número de funciones usadas para representar los orbitales de la capa de valencia, pero manteniendo un solo conjunto de funciones (zeta simple) para los orbitales cercanos al núcleo (conjuntos de *valencia dividida*) bajo la suposición de que los orbitales internos influyen pobremente en las propiedades químicas de las moléculas, en tanto que de una molécula a otra su variación es menor.

Incrementar el número de funciones base (triple zeta, cuádruple zeta,...) no elimina totalmente los problemas de las bases mínimas; esta práctica, inclusive, favorece un nuevo tipo de problema relacionado con la distribución, fuertemente anisotrópica, de la carga electrónica de ciertas moléculas. Aunque los conjuntos base de valencia dividida permiten remontar algunos problemas de la anisotropía, no es sino con la introducción de funciones de polarización. Es decir, funciones con un mayor número cuántico angular, que corresponden a funciones p , d , f , o superiores [19] que logra corregirse la deficiencia. Sin embargo, aún permanece la incapacidad conocida de los conjuntos base para representar especies cuya densidad de carga se halla significativamente desplazada lejos del conglomerado de núcleos, como aniones y especies con pares electrónicos libres. La dificultad se origina en el bajo valor de amplitud de las funciones gaussianas hacia los extremos alejados del núcleo [19], para lo cual se añaden funciones adicionales altamente difusas.

Veamos, por ejemplo, la base disponible en la suite de programas *Gaussian*®) [13], denotada como:

6-311+G(2d,p)

ésta utiliza seis gaussianas (conjunto zeta sencillo) para los orbitales internos, mientras que para los orbitales de la capa de valencia se utiliza una gaussiana para la parte contraída y una para la parte difusa (conjunto zeta triple), con funciones difusas adicionales para los átomos pesados (segundo período en adelante) y funciones de polarización para todos los átomos (incluyendo al hidrógeno). Estas últimas utilizan dos conjuntos de funciones tipo d para los átomos del segundo período en adelante y un conjunto de funciones tipo p para el hidrógeno.

10. Optimización de geometría

Para un sistema de N partículas —cada una requiere de 3 coordenadas rectangulares para especificar su posición— la función de energía (la superficie de energía potencial mencionada anteriormente) depende de $3N$ coordenadas cartesianas. La configuración espacial de este sistema que corresponde a un mínimo global de la superficie de energía potencial se identifica, en virtud del principio variacional, con la geometría del sistema en el estado basal (de equilibrio).

No obstante, hallar mínimos en la superficie de energía potencial es un proceso cuya complejidad asciende conforme crece el tamaño del sistema de interés [14]. El problema puede plantearse [19] con cierta formalidad como sigue: "dada una función f de una o más variables independientes x_1, x_2, \dots, x_k , hallar valores de dichas variables donde f tenga un valor mínimo." Dichos puntos satisfacen dos condiciones importantes,

$$\frac{df}{dx_i} = 0 \tag{9}$$

la evaluación de la primera derivada da cero, y la segunda derivada es siempre positiva

$$\frac{d^2 f}{dx_i^2} > 0 \quad (10)$$

Una optimización de geometría consiste en la localización de mínimos en la superficie de energía potencial, proceso también conocido como *minimización*. Los métodos de minimización se dividen en dos clases: aquellos que utilizan las derivadas de la energía respecto de las coordenadas, y aquellos que no las usan. El uso de las derivadas suele mejorar la eficiencia del método, pues las derivadas contienen información acerca de la forma (primera derivada) y la curvatura (segunda derivada) de la superficie [19].

La dirección del negativo del gradiente de la energía indica la dirección del máximo cambio local en sentido decreciente. La localización del mínimo, sin embargo, puede requerir muchas correcciones de las direcciones locales a las que señala el negativo gradiente, para ello se requiere un algoritmo adecuado de minimización.

Idealmente, un buen método de minimización se caracteriza por ofrecer una respuesta con gran rapidez, utilizando de manera eficiente los recursos de cómputo (espacio de memoria, tiempo de procesamiento y espacio de almacenamiento en disco). No obstante, no existe un método único que se desempeñe óptimamente en todos los problemas que se plantean en la modelación molecular.

11. Criterios de convergencia: ¿dónde se detiene la búsqueda de un mínimo en la superficie de energía potencial?

En principio, una optimización completa de geometría se considera finalizada con éxito cuando la energía, obtenida mediante un procedimiento de minimización de la superficie de energía potencial, satisface las condiciones para el mínimo, las ecuaciones (9) y (10). No obstante, el cumplimiento riguroso de las condiciones anteriores, escapa a los métodos numéricos inherentemente sujetos a errores (debido a la representación finita de los números en la memoria de una computadora), por ello se aceptan ciertas cotas de error (criterios de convergencia) sobre la energía así calculada, pretendiendo asegurar que se ha alcanzado un mínimo verdadero [13], dentro de las limitaciones señaladas.

La imposibilidad material para visualizar o explorar completamente una superficie de

energía potencial, hace necesario desarrollar criterios numéricos para investigar la consistencia del cálculo de minimización. Los criterios de convergencia también funcionan para detener un cálculo en un punto razonablemente aceptable, acortando el tiempo de procesamiento. En el conjunto de programas se utilizan, por omisión, los siguientes criterios para probar la convergencia sobre un mínimo de la superficie de energía potencial (el programa permite al usuario configurar valores personalizados) [13],

- El componente de las fuerzas en la dirección del máximo cambio, debe quedar por debajo del valor límite 0.00045 (lo cual se interpreta como cero).
- La raíz cuadrada media (RMS por sus siglas en inglés). La RMS se obtiene sumando los cuadrados de los gradientes de la superficie de energía potencial respecto a las coordenadas, se divide entre el número de coordenadas y se toma la raíz cuadrada. Sea \mathbf{g} el vector gradiente y $3N$ el número de coordenadas, luego, la RMS se define como,

$$RMS = \sqrt{\frac{\mathbf{g}^T \mathbf{g}}{3N}}$$

de las fuerzas debe ser esencialmente cero (debajo de una tolerancia predefinida, tomada como 0.0003).

- El desplazamiento estimado hacia el próximo punto, debe ser menor que el valor límite definido como 0.0018 (también interpretado como cero).
- La RMS del desplazamiento hacia el siguiente punto, deberá estar por debajo del valor límite 0.0012

12. Análisis vibracional

Los métodos de cálculo de estructura electrónica, consideran a las moléculas en un vacío hipotético, a $0K$, donde permanecen virtualmente estáticas. Sin embargo, las mediciones experimentales —en sistemas de interés práctico— se realizan a temperaturas mucho mayores, donde ocurren movimientos moleculares (traslaciones, rotaciones y vibraciones) que contribuyen a la energía del sistema. Esta situación introduce la necesidad de corregir los resultados teóricos para incluir el efecto de este tipo de movimientos.

La corrección se lleva a cabo de acuerdo con las fórmulas de la mecánica estadística,

$$U(T) = U_{\text{trasl}}(T) + U_{\text{rot}}(T) + U_{\text{vib}}(T) + U_{\text{vib}}(T=0) \quad (11)$$

donde los modos traslacional y rotacional se obtienen del *teorema de equipartición*; de tal forma que

$$U_{\text{trasl}}(T) = U_{\text{rot}}(T) = \frac{3}{2} k_B T \quad (12)$$

k_B es la constante de Boltzmann. No obstante, el cálculo de la contribución energética vibracional requiere del conocimiento de las *frecuencias de vibración* [19] del sistema. La energía debida al movimiento de vibración, clásicamente, es igual a la diferencia en la entalpía vibracional a la temperatura T y a $0K$

$$U_{\text{vib}}(T) = \sum_{i=1}^{N_{\text{mn}}} \left(\frac{h\nu_i}{2} + \frac{h\nu_i}{\exp(h\nu_i / k_B T) + 1} \right) \quad (13)$$

Aquí, N_{mn} es el número de *modos normales de vibración* del sistema; cuánticamente, la energía vibracional se calcula como

$$U_{\text{vib}}(T) = \int_0^{\infty} \frac{8\pi hc}{\lambda^5} \left(\frac{1}{\exp\left(\frac{hc}{\lambda k_B \cdot T}\right) - 1} \right) d\lambda \quad (14)$$

empleando la aproximación de Boltzmann [26]. Una molécula con M átomos posee $3M - 6$ modos normales de vibración. Las frecuencias de vibración de los modos normales, así como los desplazamientos individuales de los átomos se calculan a partir de la matriz de segundas derivadas (Hessiano) de la superficie de energía potencial, $V(r)$, del sistema.

Primero, el Hessiano (V'') debe convertirse a la matriz equivalente de constantes de fuerza (F), en coordenadas ponderadas por las masas atómicas

$$F = \mu^{-1/2} V'' \mu^{1/2}$$

la matriz diagonal \mathbf{m} de tamaño $3M \times 3M$, contiene las masas atómicas ($\mu_{1,1} = m_1$,

$\mu_{2,2} = m_1, \mu_{3,3} = m_1, \mu_{2,2} = m_2, \dots, \mu_{3M-2,3M-2} = m_M, \mu_{3M-1,3M-1} = m_M, \mu_{3M,3M} = m_M$). Cada elemento de la diagonal principal de $\mathbf{m}^{-1/2}$ es el recíproco de la raíz cuadrada de la masa del átomo correspondiente. Una fuerza dada tiene un efecto distinto sobre un átomo pesado que sobre uno ligero, por ello deben tomarse en cuenta las masas de los átomos [19]. Luego, se resuelve el determinante secular

$$|F - \lambda I| = 0$$

para obtener los valores propios y los vectores propios de \mathbf{F} . La frecuencia de cada modo normal de vibración se obtiene a partir de los valores propios, mediante la relación

$$\nu_i = \frac{\sqrt{\lambda_i}}{2\pi} \quad (15)$$

En (15) se observa que un valor propio negativo produce una frecuencia imaginaria. Por ello, el cálculo de las frecuencias de los modos normales de vibración suele utilizarse como comprobación adicional de la optimización de la geometría de equilibrio [14], pues un valor propio negativo puede corresponder a un estado de transición o un punto de silla de orden superior [27]. El análisis vibracional es válido únicamente en estados donde las primeras derivadas respecto a los desplazamientos cartesianos son *nulas*, es decir en geometrías de equilibrio, además, por el uso de la matriz de segundas derivadas (Hessiano), un cálculo de frecuencias siempre debe realizarse exactamente al mismo nivel de teoría y con la misma base utilizada en la optimización de geometría.

Ochterski (1999) [27] presenta en detalle la técnica utilizada en el programa para el cálculo de las frecuencias y los modos normales de vibración; también discute algunos casos excepcionales en que la técnica se aplica a estados cuya geometría no es la de equilibrio.

13. Análisis de Mulliken

La pérdida o la adquisición de uno o varios electrones de un átomo individual, son procesos perfectamente plausibles que dan origen a un ión; es decir, una especie química con una carga eléctrica formal sobre el átomo. Sin embargo, un átomo enlazado en un grupo de átomos (molécula) forma parte de una colectividad y por lo tanto los electrones, en su conjunto, son afectados por todos los núcleos de los átomos presentes. La individualidad de los átomos que la conforman es una mera abstracción que da paso a una nueva estructura electrónica molecular. La idea de una carga eléctrica centrada en los núcleos de los átomos que constituyen una molécula no deja de ser un concepto arbitrario, dado que en la mecánica cuántica no se encuentra ningún operador asociado a la carga eléctrica atómica.

Los métodos de análisis de población proponen esquemas teóricos, en los cuales la densidad electrónica de la molécula se reparte entre los núcleos, proporcionando la noción de que cada núcleo está asociado a algún número de electrones; aunque no necesariamente un número entero, dando origen a cargas eléctricas parciales sobre éstos. Diversos procedimientos (como el análisis de Mulliken, el análisis de Löwdin y el análisis de población natural, entre otros) desarrollados para éste propósito están disponibles en la literatura [19, 28] donde se da cuenta de la variabilidad de sus predicciones, y la dependencia de factores como la calidad de la base empleada.

El método de Mulliken [19] para el análisis de población parte de la condición impuesta a la densidad monoeléctrica $\rho(r)$ que debe integrarse exactamente al número de electrones, N , de la molécula

$$N = \int \rho(r) dr$$

Si la densidad monoeléctrica se expresa en términos de una combinación lineal de funciones base, ϕ_μ , normalizadas, puede escribirse el número de electrones en términos de las integrales de traslape

$$S_{\mu\nu} = \int dr \phi_\mu^* \phi_\nu \tag{16}$$

y los elementos de la matriz de densidad,

$$P_{\mu\nu} = 2 \sum_{i=1}^{N/2} c_{\mu i} c_{\nu i} \quad (17)$$

relacionados con los coeficientes de la expansión lineal en términos de las funciones base; esto es

$$N = \sum_{\mu=1}^K P_{\mu\mu} + 2 \sum_{\mu=1}^K \sum_{\nu=\mu+1}^K P_{\mu\nu} S_{\mu\nu} \quad (18)$$

Luego, la carga neta sobre un átomo se obtiene restando el número de electrones sobre A, de la carga nuclear, Z_A ,

$$q_A = Z_A - \sum_{\mu=1:\text{men } A}^K P_{\mu\mu} - \sum_{\mu=1:\text{men } A}^K \sum_{\nu=1:\nu \neq \mu}^K P_{\mu\nu} S_{\mu\nu} \quad (19)$$

En el método de Mulliken se observa una dependencia de la carga atómica sobre el tamaño de la base utilizada (Bachrach). Existen otros métodos que superan los problemas del esquema de Mulliken, a costa de recursos computacionales. Sin embargo, la tendencia general de las cargas obtenidas con el análisis de población de Mulliken y con los otros métodos es similar, como se puede ver en los trabajos reportados previamente [29-34]. En aplicaciones que requieran una mayor precisión, otros métodos (como el método de Bader (1990) [35] basado en la teoría de átomos en moléculas dan mejores resultados que el análisis de Mulliken [28, 36, 37] y deberán ser utilizados.

Glosario básico

Ab initio es una locución latina que significa "desde primeros principios" o "desde las bases", se dice de los métodos de cálculo basados en la **función de onda**, que se emplean para resolver de forma aproximada la ecuación de Schrödinger, teniendo como único sustento experimental la determinación de los valores de un reducido número de constantes físicas universales (como la carga del electrón, la constante de Boltzmann, la constante de Planck, la permisividad en el vacío).

Anión es un **ión** con carga negativa, debido a la absorción de uno o más electrones.

Aproximación de Born-Oppenheimer, también llamada **aproximación de núcleos fijos**, basándose en el régimen de separación entre los movimientos nucleares, habitualmente más lentos que los electrónicos, y los movimientos de los electrones, que ajustan sus posiciones rápidamente en respuesta a los primeros, establece que la ecuación de Schrödinger del sistema puede separarse en una ecuación electrónica y otra nuclear.

ARN mensajero: es el resultado de la transcripción de los genes, es la molécula que transporta la información genética para poder realizar la traducción.

Una **base mínima** contiene apenas suficientes funciones para representar los orbitales llenos en cada átomo. Llevado a la práctica, el conjunto base mínimo contiene todos los orbitales atómicos de una capa dada.

Capa de valencia es el nombre que se da a los electrones más externos de un átomo, usualmente son éstos los que se involucran en la formación del enlace químico.

Catión es un **ión** con carga positiva, debido a la pérdida de uno o más electrones.

Conjunto base. La resolución de la ecuación de onda arroja como resultado un conjunto de soluciones que forman un conjunto completo de funciones. Los métodos Ab initio utilizan las soluciones para el átomo de hidrógeno y con ellas construyen aproximaciones para las funciones de onda de sistemas de mayor tamaño mediante el método denominado OMCLAO (orbitales moleculares como combinaciones lineales de orbitales atómicos).

Conjunto completo de funciones. En matemáticas las funciones forman conjuntos que al cumplir ciertos requisitos se llaman espacios de funciones. Un conjunto completo de funciones posee un número infinito de elementos y cualquiera de ellos puede escribirse como una combinación lineal de los otros elementos del espacio.

Ecuación de onda (v. ecuación de Schrödinger)

Ecuación de Schrödinger es la ecuación fundamental para describir la energía de los sistemas electrónicos (cuánticos). También se la conoce como **ecuación de onda**.

Energía es una propiedad fundamental de la materia; la cantidad de energía en un sistema cerrado se mantiene constante.

Entropía es una propiedad termodinámica de un sistema; depende naturalmente de la temperatura y del volumen. Es una propiedad importante que permite definir la dirección del avance de un proceso, la extensión del avance y el criterio para la detención del mismo.

Estado basal es un término que designa el estado de un sistema con la mínima energía posible. En termodinámica se considera que un sistema tiende a sufrir cambios en dirección a su estado de mínima energía, donde adquiere su mayor estabilidad (a fin de no sufrir más cambios).

Ión es el nombre de un átomo o grupo de átomos que en su conjunto poseen diferente número de protones y electrones, produciendo una especie química con carga eléctrica formal.

Las **funciones contraídas** se forman agrupando algunas de las funciones del conjunto base original, por norma a las funciones no contraídas se las llama *primitivas*

Las **Funciones difusas** se añaden a los conjuntos base para describir mejor el comportamiento de la distribución electrónica, en particular de especies cuya nube de electrones tiene tendencia a expandirse (como en el caso de especies aniónicas), deben tener el mismo tipo de momento angular que las funciones presentes de antemano en la base, pero con exponentes menores a los de cualquiera de ellas, para describir mejor la región de valencia.

Función de onda es la solución que se obtiene al resolver la ecuación de Schrödinger para el sistema de interés. Es una función multidimensional que depende de las posiciones de los M núcleos y de los N electrones del sistema, y a veces, también del tiempo.

Funciones ortogonales. En matemáticas existe una definición más general del producto interno aplicada a funciones, y de manera análoga al producto interno vectorial, las funciones ortogonales son aquellas cuyo producto interno se anula.

Número atómico (Z) es el número de electrones en un átomo eléctricamente neutro, por lo que corresponde también al número de protones en el núcleo.

Polarizabilidad es la tendencia relativa de la nube electrónica de un átomo a ser distorsionada por la presencia de un ión vecino o un dipolo, es decir, por un campo eléctrico externo.

SCF (self consistent field). Métodos basados en el mejoramiento progresivo de la solución de un problema, mediante un procedimiento iterativo autoconsistente.

Referencias bibliográficas

1. Voet, D. and J.G. Voet, *Biochemistry*. 4th ed. 2011, Hoboken, New Jersey: John Wiley & Sons.
2. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Biochemistry*. 7th ed. 2012, New York: W. H. Freeman.
3. McKee, T. and J.R. McKee, *Biochemistry: the molecular basis of life*. 4th ed. 2009, New York: Oxford University Press.
4. Cháñez Cárdenas, M.E. and E. Vázquez Contreras, *El plegamiento de las proteínas*, in *La física biológica en México*, L.S. García-Colín, et al., Editors. 2006, El Colegio Nacional: México, D. F. p. 195-215.
5. Vázquez Contreras, E., et al., *The unfolding of proteins induced by different denaturants*, in *Advances in Protein Physical Chemistry* E. García-Hernández and D.A. Fernández-Velasco, Editors. 2008, Transworld Research Network: India. p. 169-192.
6. Vázquez Contreras, E., *Las interacciones fundamentales en la estructura tridimensional de las proteínas y otras biomoléculas*, in *Introducción a la física biológica*, L.S. García-Colín, et al., Editors. 2010, El Colegio Nacional: México, D. F. p. 237-278.
7. IUPAC, *Hydrogen bond*, M. Nic, J. Jirat, and B. Kosata, Editors. 2006-, IUPAC Compendium of Chemical Terminology.
8. Dill, K.A. and S. Bromberg, *Molecular driving forces: statistical thermodynamics in chemistry & biology*. 1st. ed. 2002, USA: Garland Science.
9. Casabó i Gispert, J., *Estructura atómica y enlace químico*. 1999, España: Editorial Reverté, S. A. 392.
10. Szabo, A. and N.S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*. 1st. revised ed. 1989, New York: McGraw-Hill Publishing Co.
11. Koch, W. and M.C. Holthausen, *A chemist's guide to density-functional theory*. 2nd. ed. 2001, Germany: Wiley-VCH.
12. Simons, J. and J. Nichols, *Quantum Mechanics in Chemistry*. 1997, New York: Oxford University Press.
13. Foresman, J.B. and Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods*. 2 ed. 1996, Pittsburgh, PA: Gaussian, Inc.
14. Young, D., *Computational Chemistry: a practical guide to applying techniques to real world problems*. 1st. ed. 2001, New York: Wiley-Interscience John Wiley and Sons.
15. Hyperchem, *Hyperchem computational chemistry*, 1996, Hyperchem, Inc.

16. Hohenberg, P. and W. Kohn, *Inhomogeneous electron gas*. Physical Reviews, 1964. **136**: p. B864.
17. Kohn, W. and L.J. Sham, *Self-consistent equations including exchange and correlation effects*. Physical Reviews, 1965. **140**: p. A1133.
18. Kohn, W., A.D. Becke, and R.G. Parr, *Density Functional Theory of Electronic Structure*. J. Phys. Chem., 1996. **100**(31): p. 12974.
19. Leach, A.R., *Molecular modelling: principles and applications*. 1st. ed. 1996, England: Addison Wesley Longman Ltd.
20. Hehre, W.J., et al., *Ab initio molecular orbital theory*. 1st. ed. 1986, USA: John Wiley & Sons.
21. McQuarrie, D.A., *Quantum chemistry*. 2nd. ed. 1980, New York: University science Books.
22. Custodio, R. and J.D. Goddard, J. Mol. Struct. (THEOCHEM), 1993. **281**: p. 75-87.
23. Raffenetti, R.C., J. Chem. Phys., 1973. **58**: p. 4452.
24. Almlof, J., T. Helgaker, and P.R. Taylor, J. Chem. Phys., 1988. **32**: p. 3029.
25. Almlof, J. and P.R. Taylor, Adv. in Quantum Chem., 1991. **22**: p. 301.
26. Castellan, G.W., *fisicoquímica*. 2a. ed. 1988, México: Addison-Wesley Iberoamericana.
27. Ochterski, J.W. *Vibrational Analysis in Gaussian*. 1999 Oct. 16, 2003; Available from: http://www.gaussian.com/g_whitepap/vib.htm.
28. Bachrach, S.M., *Population analysis and electron densities from quantum mechanics*, in *Reviews in Computational Chemistry*, K.B. Lipkowitz and D.B. Boyd, Editors. 1994, VCH Publishers, inc.: New York. p. 171-227.
29. Koerner, J.B. and T. Ichiye, J. Phys. Chem. B, 2000. **104**: p. 2424-2431.
30. Lill, M.A., M.C. Hutter, and V. Helms, J. Phys. Chem. A, 2000. **104**: p. 8283-8289.
31. Sakata, K.J., J. Phys. Chem. A, 2000. **105**: p. 10001-10008.
32. Yagi, T., K. Morihashi, and O. Kikuchi, J. Phys. Chem. A, 2001. **105**: p. 8241-8247.
33. Ruzsinszky, A. and G.I. Csonka, J. Phys. Chem. A, 2003. **107**: p. 8687-8695.
34. Braun-Sand, S.B. and O. Wiest, J. Phys. Chem. B, 2003. **107**: p. 9624-9628.
35. Bader, R.F.W., *Atoms in molecules: a quantum theory*. 1990, Oxford, UK: Oxford University Press.
36. de Oliveira, A.E., et al., *A Simple Potential Model Criterion for the Quality of Atomic Charges*. J. Phys. Chem. A, 1999. **103**(25): p. 4918.
37. Ponec, R., et al., J. Phys. Chem. A, 2002. **106**: p. 1019-1025.