



Universidad Autónoma del Estado de Hidalgo
Instituto de Ciencias Básicas e Ingeniería
Área Académica de Química

**Determinación de sitios activos en la interacción
glucosa-insulina, utilizando métodos de
química cuántica computacional.**

Tesis

Que para obtener el título de
Licenciada en Química

Presenta

Ana María Vera Meneses

Director de Tesis

Dr. Luis Humberto Mendoza Huizar



La presente tesis se realizo bajo la dirección del Dr. Luis Humberto Mendoza Huizar en el Laboratorio de Electroquímica y en la Estación de Trabajo del Centro de Investigaciones Químicas de la Universidad Autónoma del estado de Hidalgo.

Parte de este trabajo ha sido publicado en:

Revista Electrónica Ciencia Latinoamericana, (2009) Vol.1, No.1, X, XX © 2009
CENBIOEL ISSN 0718-6762 (en revisión)





A mi Josué, porque a pesar de ser
pequeño me dio la alegría más grande.





Agradecimientos

A mi Mamá por su apoyo incondicional, porque a pesar de todo siempre confió en mí. Por alentarme y por ser también una amiga.

A mi asesor Dr. Luis Humberto Mendoza Huizar por darme su apoyo en todo momento, por creer en mí y por la oportunidad que me brindó al realizar este proyecto.

A mi tutora Dra. Noemí Andrade López por alentarme y darme los mejores consejos, por brindarme su amistad y escucharme cuando lo necesitaba. Y a cada uno de mis profesores.

A mi mejor amiga Madaí por estar ahí siempre que te necesite, y brindarme tu amistad.

A mis sinodales por los consejos aportados a esta tesis que fueron de gran ayuda

A ti Luis por apoyarme en los momentos difíciles, por ser más que un amigo y porque llegaste en el momento más indicado a mi vida, por dedicarme cada uno de tus alegrías así como yo ahora comparto contigo una más.



Resumen

El propósito de este trabajo de investigación es proveer información de cuáles son los sitios activos que intervienen en la interacción glucosa-insulina. Se realizó un estudio teórico (cuántico) de la afinidad electrónica de los sitios activos de la glucosa e insulina, así como un cálculo tipo *Docking* que ayudó a predecir la zona de interacción. Las moléculas de la glucosa se caracterizaron mediante un análisis a un nivel de teoría HF/6-31G y diferentes métodos de teoría como AM1, MM, funcionales de la densidad, cálculos de frecuencias y determinación de orbitales moleculares, como análisis vibracional, cálculos de la dureza para la identificación de la molécula glucosa más estable. En el estudio de la molécula 6-desoxi-D-glucosa se encontró que los átomos H y O participan en la interacción mostrando puentes de hidrógeno con la insulina 9INS de la cadena β . Se analizó la molécula de la insulina a partir del método ONIOM con dos niveles de teoría la primera MM para el nivel bajo y HF/ 6-31G para el nivel alto.

Los resultados sugieren que la reactividad de la insulina se debe a los grupos funcionales exclusivamente en la cadena β , involucrado a los residuos 33 Val, 34 Glu 35, Ala 36, Leu 37, Tyr y 45 Phe, los rótulos de los átomos nucleofílicos y electrofílicos más reactivos se detallan en el capítulo 4. Del análisis del *Docking* se obtiene que la molécula 6-desoxi-D-glucosa forma puentes de hidrógeno con la insulina debido a la existencia de pares de electrones libres del oxígeno.

Índice General

Lista de siglas y símbolos	i
Índice de Tablas	iii
Índice de Figuras	iv

Capítulo 1. Antecedentes

1.1 Introducción	1
1.1.1 Tipos de diabetes	2
1.1.2 Glucosa e Insulina	3
1.1.3 Mecanismos de regulación metabólica	4
1.2 Introducción a los métodos de cálculo	8
1.2.1 Mecánica molecular	11
1.2.2 Métodos de estructura electrónica	12
1.3 Métodos de Cálculo	28
1.3.1 <i>Semiempíricos</i>	28
1.3.2 <i>Ab initio</i>	28
1.3.3 <i>Teoría de los Funcionales de la Densidad</i>	31
1.3.4 El método ONIOM	34
1.4 <i>Utilización del método ONIOM</i>	34
1.5 Justificación	37
1.6 Objetivos	38
1.6.1 Objetivo general	38
1.6.1.1 Objetivos específicos	38
1.7 Referencias	39

Capítulo 2. Metodología

2.1 Introducción	41
2.2 Método de calibración	43
2.2.1 <i>Mecánica Molecular</i>	46
2.2.2 <i>Semiempíricos</i>	49
2.2.3 <i>Ab initio</i>	49
2.2.4 <i>Funcionales de la densidad</i>	49
2.3 Referencias	51

Capítulo 3. Propiedades electrónicas de la glucosa

3.1 Introducción	52
3.2 Dureza	55
3.3 Distribución de sitios activos	58
3.4 Análisis Vibracional	59
3.5 Conclusiones	61
3.6 Referencias	62

Capítulo 4. Propiedades electrónicas de la insulina

4.1 Estudio a nivel semiempirico PM6	64
4.1.1. Optimización de geometría	66
4.2 Estudio a nivel <i>ab initio</i> utilizando el método <i>ONIOM</i>	66
4.2.1 Distribución de sitios activos	66
4.2.2 Mapa del orbital HOMO.	67
4.2.3 Mapa del orbital LUMO.	68
4.3 Conclusiones	70
4.4 Referencias	72

Capítulo 5. Interacción Glucosa-Insulina utilizando *Docking*

5.1 En qué consiste el <i>Docking</i>	73
5.2 Optimización de geometría después de utilizar el proceso de <i>Docking</i>	77
5.3 Resultados y discusión	78
5.4 Conclusiones	79
5.5 Referencias	80

Capítulo 6. Conclusiones Generales

81

Capítulo 7. Perspectivas

83

Anexos

A.1 Mapeo del HOMO y LUMO. Método <i>ONIOM</i>	84
A.2 Artículo en revisión generado del desarrollo de esta tesis	86

Lista de siglas y símbolos

QSPR	Identificación de patrones, tendencias y correlaciones entre estructuras químicas y sus propiedades
GLUT	Transportador de glucosa a través de la membrana apical del intestino
QM	Mecánica cuántica
MM	Mecánica molecular
MME	Métodos de estructura electrónica
ψ	Función de onda del sistema
R	Vector que representa todas las coordenadas de las diferentes partículas que componen el sistema
E	Energía total del sistema
$\hat{H}$	Operador Hamiltoniano
M	Masa del núcleo
Z	Número atómico de los núcleos
$g(q)$	Gradiente
core	Considerando los efectos de correlación electrónica de los electrones de valencia
r_{ij}	Repulsión interelectrónica
r_i	Repulsión nuclear
r_j	Atracción electrón-núcleo
$\hat{F}$	Operador de Fock
PES	Superficies de Energía Potencial
φ_i	Orbitales moleculares
V^{av}	Evaluación del potencial promedio
ε_i	Energía de los orbitales
c	Coeficientes de desarrollo de los orbitales moleculares
E_{HF}	Energía óptima de Hartree-Fock
HMO	Método de Hückel
TFD	Teoría de los funcionales de la Densidad
$\rho(r)$	Densidad electrónica
$E[\rho]$	Energía total del sistema
$v(r)$	Potencial externo impuesto por los núcleos
$F[\rho]$	Funcional universal
$T[\rho(r)]$	Funcional de la energía cinética
μ	Funcional de la energía de interacción electrón-electrón
μ	Potencial químico
I	Potencial de ionización
A	Afinidad electrónica
η	Dureza absoluta
q_i	Carga electrostática
s	Blandura local
S	Blandura global del sistema
EE	Estructura electrónica
E_{ONIOM}	Método ONIOM
E_3	Energía del sistema entero con los cálculos del nivel bajo
E_2	Energía del sistema modelo con los cálculos del alto nivel

E_1	Energía del sistema modelo con cálculos del nivel bajo
f_i^+	Para un ataque electrofílico
f_i^-	Para un ataque nucleofílico
f_i^0	Para un ataque por radicales libres
$C_6H_{12}O_6$	Glucosa
Gly	Glicina
Gln	Glutamina
Ile	Isoleucina
Val	Valina
Glu	Acido glutámico
Ala	Alanina
Leu	Leucina
Phe	Fenilalanina
Tyr	Tirosina
HOMO	Orbital más alto ocupado
LUMO	Orbital más bajo desocupado

Lista de Tablas

Tabla 2.1	Estructuras primarias de la glucosa	44
Tabla 2.2	Selección de capas utilizando el método <i>ONIOM</i>	47
Tabla 3.2	Propiedades electrónicas de la glucosa	57
Tabla 3.3	Mapeo del HOMO	59
Tabla 3.4	Mapeo del LUMO	59
Tabla 3.5	Análisis vibracional de la glucosa	60
Tabla 4.1	Átomos electrofílicos mas reactivos obtenidos a partir de la estructura 9INS optimizada en el nivel PM6.	67
Tabla 4.2	Átomos nucleofílicos más reactivos obtenidos a partir de la estructura 9INS optimizada en el nivel PM6	69
Tabla 5.1	Átomos electrofílicos y nucleofílicos más reactivos obtenidos a partir del <i>Docking</i>	78
Tabla A.1	Mapeo del HOMO y LUMO	84

Lista de Figuras

Fig. 1.1	Rutas principales de la utilización de la glucosa	4
Fig. 1.2	Glucólisis	6
Fig. 1.3	Reacciones del ciclo del ácido cítrico	7
Fig. 1.4	Glucogénesis	8
Fig. 1.5	Superficie de energía potencial	17
Fig. 1.6	Modelo <i>ON/OM</i>	36
Fig. 2.1	Estructura de la insulina (9INS). Modelo <i>ON/OM</i> . Sección 1.	46
Fig. 3.1	Ciclación de la glucosa	52
Fig. 3.2	La mutarrotación de la glucosa	53
Fig. 3.3	Estructura primaria de la celulosa	53
Fig. 3.4	Modelos utilizados (glucosa)	54
Fig. 4.1	Insulina. Código de acceso 9INS	65
Fig. 4.2	Distribución del orbital HOMO de la insulina monomérica 9INS optimizada en el nivel PM6	67
Fig. 4.3	Distribución del orbital LUMO de la insulina monomérica 9INS optimizada en el nivel PM6	69
Fig. 5.1	Esquema del método de muestreo en Docking de los métodos fast shape matching (<i>DOCK</i>).	74
Fig. 5.2	<i>Docking</i> (glucosa-insulina)	77

Capítulo 1

Antecedentes

1.1 Introducción

Root-Bernstein y Dillon han sugerido que la complementariedad molecular es una presión selectiva que dirige la evolución en una dirección determinada mediante la vinculación de la estructura molecular y la función. En esencia, las moléculas que se unen entre sí tienen más probabilidades de sobrevivir a los procesos de degradación, y se convierten en los materiales sobre los que se construyen los sistemas vivos. Además, los complejos que estos pares de moléculas forman tienen propiedades emergentes que sus componentes individuales no; lo que proporciona una fuente constante de novedad en los que a una complejidad mayor se puede construir. Recientemente se aplicó esta teoría a la elucidación de la evolución del receptor de la insulina. En particular, el receptor de la insulina tiene similares múltiplos a la insulina y como el glucagón las secuencias situadas en las regiones asociadas con la insulina vinculante. ^[1-3]

Otro aspecto interesante del comportamiento de la insulina es que la glucosa se une directamente a ella, lo que sugiere que la glucosa y la insulina estaban vinculadas funcionalmente antes de su evolución. Varios métodos físico-químicos, cristalografía de rayos X y programas de modelado molecular han demostrado que en la D-glucosa existen sitios de unión a la molécula de insulina, dos con un enlace constante de 1×10^{-3} M y seis con un solo enlace. Esta unión constante de alrededor de 6×10^{-2} M. significa que en la glucosa en sangre, en condiciones normales, es de 5×10^{-3} M, los dos mayores sitios de afinidad tienden a ser ocupados, mientras que los sitios más bajos

son probables condiciones de hiperglucemia. Uno puede formular la hipótesis de que antes de los orígenes los receptores de insulina y de las proteínas de transporte de la glucosa, una molécula similar a la insulina puede haber llevado a la glucosa en las células en sí. Además, en presencia de la glucosa, la insulina es rápida y no enzimáticamente glucosilada, que puede tener una variedad de efectos que se analizan a continuación. ^[3-5]

La unión de la glucosa a la insulina, combinada con la teoría de la complementariedad molecular de la evolución, sugirió que tal vez los transportadores de glucosa (GLUT) podrían haber evolucionado de una molécula similar a la insulina. Si ese fuera el caso, lo lógico sería que la afinidad modular será similar a la insulina en el núcleo de las regiones del transportador de GLUT, pero no en los núcleos de los transportadores de otros tipos de azúcar. ^[5]

1.1.1 Tipos de Diabetes

La diabetes es un padecimiento crónico de gran importancia médico social porque afecta a muchos adultos y su frecuencia está aumentando con el envejecimiento de las poblaciones. Puede ocasionar complicaciones graves como ceguera, nefropatías, gangrena, entre otras. ^[6]

El metabolismo se define como el proceso que convierte el alimento que ingerimos en energía. Dentro de este proceso, el transporte de la glucosa hacia la membrana celular es de vital importancia. Cuando dicho transporte falla empiezan a aparecer enfermedades tales como la diabetes. ^[6]

Cuando el páncreas produce poca insulina se le denomina diabetes tipo I y si las células del cuerpo no responden a la insulina que produce es de tipo II. ^[6]

La *diabetes mellitus* es una alteración en el metabolismo de los carbohidratos asociada con una insuficiencia insulínica debido a una anomalía en los islotes de Langerhans del páncreas. [7]

Hay también otros factores ambientales muy asociados a la diabetes por ejemplo la obesidad, en el sentido de que los obesos presentan diabetes más que las personas delgadas, entre otras enfermedades como: arteriosclerosis y enfermedad coronaria; parece ser que los diabéticos desarrollan más frecuente y tempranamente estas dos enfermedades que los no diabéticos. [8]

El tratamiento dietético, el uso de la insulina y la prevención de las infecciones reducen mucho las complicaciones; sin embargo, a pesar de un adecuado tratamiento, los enfermos pueden desarrollar arteriosclerosis, hipertensión y catarata. La expectativa de vida de los diabéticos ha aumentado mucho desde la introducción farmacéutica de la insulina. La insulina ha mitigado el padecimiento pero no se cura la diabetes, sólo mantiene al enfermo con vida previniendo las consecuencias. El tratamiento, entonces, dura toda la vida. [8]

La incapacidad física que produce la diabetes es grave cuando se presentan complicaciones tales como la ceguera parcial o total, amputaciones por gangrena o diabetes muy avanzadas que exigen un tratamiento muy rígido. En estos casos los procedimientos de rehabilitación pueden ayudar al individuo a valerse por sí mismo.

1.1.2 Glucosa e Insulina

El metabolismo de los carbohidratos se inicia con el movimiento de glucosa a través de las membranas celulares. Inmediatamente después de llegar al interior de la célula, la glucosa reacciona con el ATP para producir glucosa-6-fosfato. La reacción se conoce como fosforilación de la glucosa, y es catalizada por la enzima glucocinasa. El

metabolismo de la glucosa consiste en dos procesos sucesivos: glucólisis y el ciclo del ácido cítrico (también llamado ciclo del ácido tricarboxílico, ciclo de Krebs y respiración celular).^[9]

Las células beta de los islotes de Langerhans en el páncreas secretan la hormona más conocida que participa en la regulación del metabolismo de la glucosa: la insulina. Aunque todavía no se ha dilucidado con exactitud cómo actúa la glucosa, varios de sus efectos se conocen plenamente.^[9]

Los islotes de Langerhans secretan dos hormonas que regulan la glucosa: la insulina por las células beta y glucagon por las células alfa. La insulina tiende a disminuir la glucosa en la sangre, mientras que el glucagon tiende a aumentarla.

1.1.3 Mecanismos de regulación metabólica

La glucosa no es solo un combustible excelente sino también un precursor extremadamente versátil, capaz de suministrar una gran cantidad de intermedios metabólicos, que son materiales de partida necesarios para reacciones biosintéticas (ver *figura 1.1*). El estudio de los numerosos destinos metabólicos necesarios para el crecimiento.^[10]

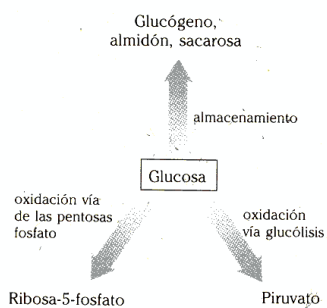
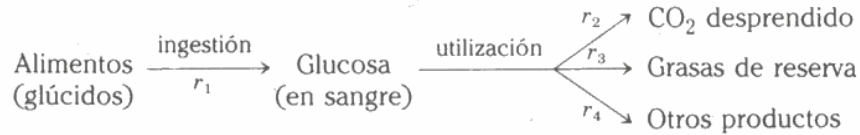


Figura 1.1 Rutas principales de utilización de la glucosa en las células de plantas y animales superiores. Aunque no son los únicos destinos posibles de la glucosa, estas tres vías son las más significativas en términos de la cantidad de glucosa que fluye a través de las mismas en la mayoría de células.^[10]

En las plantas y animales superiores la glucosa tiene destinos principales:



La glucosa procedente de los alimentos (o de reservas de glúcidos) penetra en el torrente sanguíneo y ser consumida como combustible para los procesos metabólicos de otros tejidos (corazón, cerebro, músculo esquelético). En este esquema r_1 , etc. representan las velocidades de los diversos procesos. Las concentraciones de hemoglobina y glucosa en el estado estacionario dinámico se mantienen mediante complejos mecanismos regulares de las velocidades relativas de los procesos que aquí se comentan.^[10]

La *glucólisis* es el proceso que cambia en varios pasos la glucosa en dos moléculas de piruvato y por lo tanto libera algo de energía almacenada en las moléculas de la glucosa (ATP, NADH, FADH₂) esto ocurre en el citoplasma de las células (*figura 1.2*).^[8]

El *ciclo del ácido cítrico* está constituido por una serie de reacciones químicas catalizadas por enzimas acompañadas por cambios energéticos, consiste en el cambio producido por el ácido pirúvico, (*figura 1.3*) el cual se desdobla y se oxida para formar CO₂ y agua. Genera muchas más moléculas energéticas que la glucólisis.^[8]

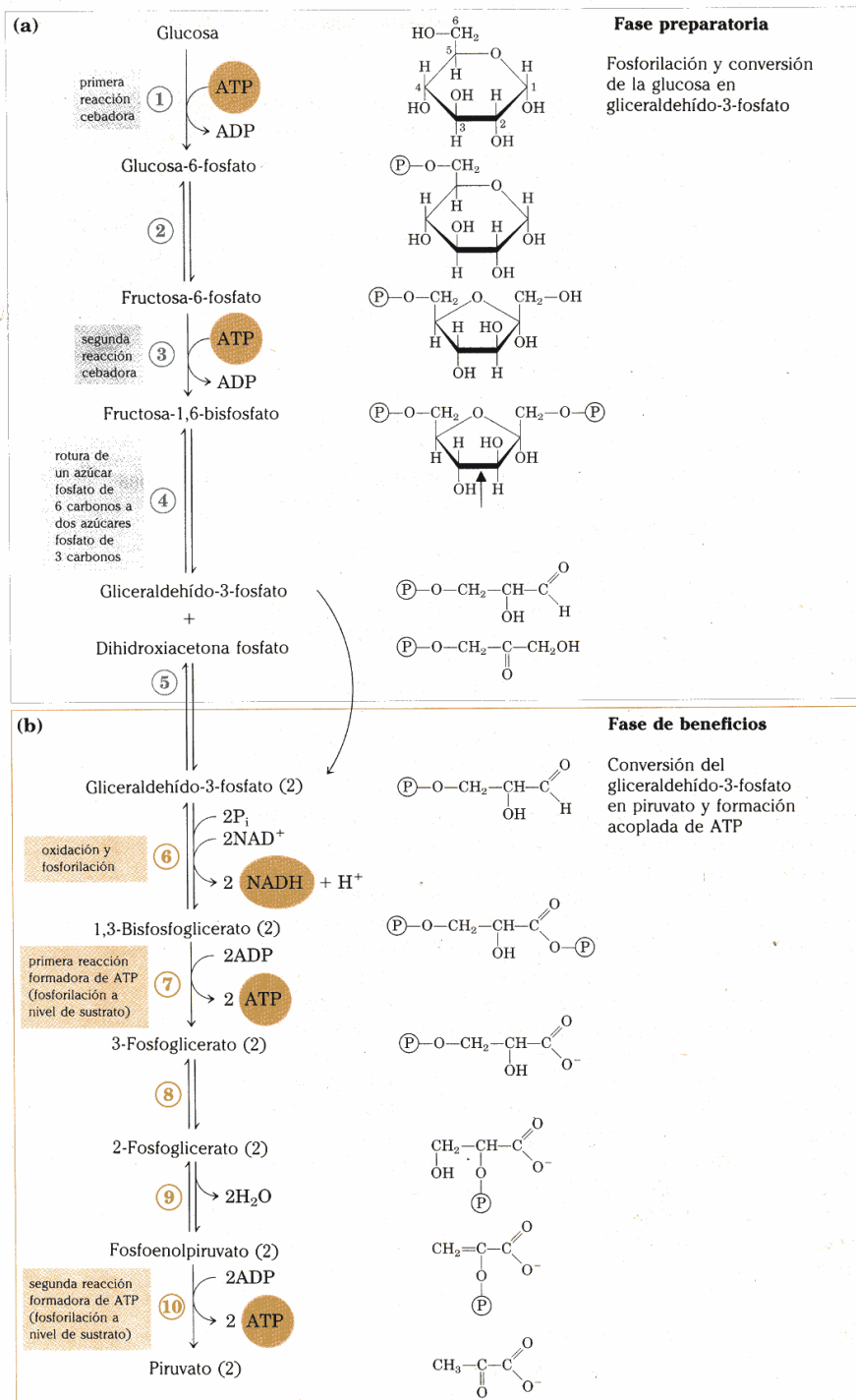


Figura 1.2 Glucólisis. Las dos fases de la glucólisis. Por cada molécula de glucosa que pasa a través de la fase preparatoria (a), se forman dos moléculas de gliceraldehído-3-fosfato: ambas pasan por la fase de beneficios (b). El piruvato es el producto final de la segunda fase en condiciones aeróbicas, pero en condiciones anaeróbicas se reduce a lactato para generar el NAD⁺. Por cada molécula de glucosa se consumen dos ATP en la fase preparatoria y se producen cuatro ATP en la fase de beneficios dando un rendimiento neto de dos moléculas de ATP por una glucosa convertida en piruvato. ^[10]

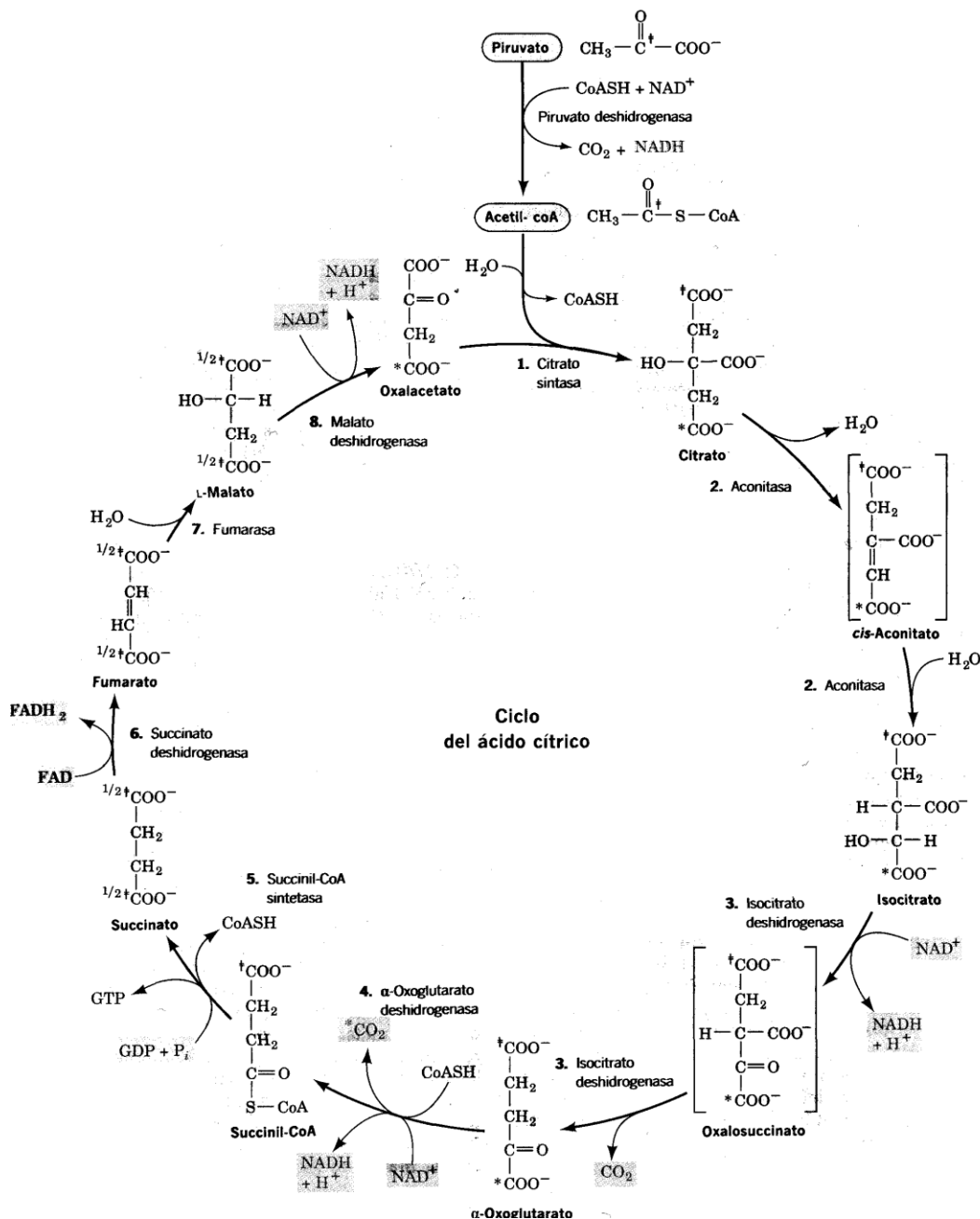


Figura 1.3 Reacciones del ciclo del ácido cítrico. Los reactivos y los productos de este ciclo aparecen sombreados. La reacción piruvato $\rightarrow$ acetil-CoA (arriba) aporta el sustrato del ciclo, que proviene del metabolismo de carbohidratos, aunque no se considera parte del ciclo. Los compuestos que aparecen entre corchetes son intermediarios unidos a enzimas. Cuando se marca el C(4) del oxalacetato(*), la marca pasa al C(1) del α -oxoglutarato y se libera en forma de CO_2 en la reacción 4. Cuando se marca el C(1) del acetil-CoA (≠), la marca aparece el C(5) del α -oxoglutarato y se reparte entre el C(1) y el (4) del succinato ($1/2\neq$).^[9]

La *glucogénesis* es la formación de glucógeno a partir de la glucosa. Consiste en reacciones químicas, cada una de ellas catalizada por enzimas diferentes. Las moléculas de glucógeno no quedan permanentemente en las células sino se desdoblan o hidrolizan a glucosa-6- fosfato, o en algunas células, a glucosa. ^[11]

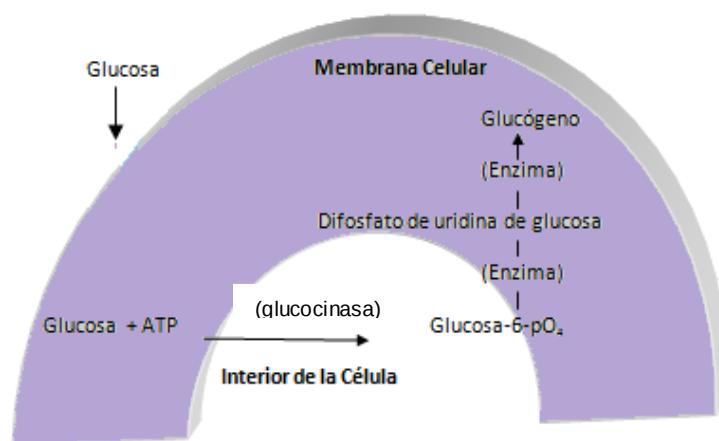


Figura 1.4 Glucogénesis. El hígado y el músculo son los dos principales tejidos de almacenamiento de glucógeno. En el músculo, la demanda de ATP provoca conversión del glucógeno en glucosa-6-fosfato (G6P) que entrará en la glucólisis. En el hígado una concentración baja de glucosa en la sangre pone en funcionamiento la degradación del glucógeno a G6P que, a su vez, es hidrolizada glucosa y liberada al torrente sanguíneo ^[9,11]

1.2 Introducción a los métodos de cálculo

La química computacional es una rama de la química teórica y de la química cuántica. Su objetivo es producir y utilizar programas informáticos para el estudio de las propiedades (como propiedades electrónicas, energía, momento, frecuencias de vibración) de moléculas. En química teórica, los químicos y los físicos desarrollan algoritmos y teorías que permiten hacer predicciones precisas de propiedades atómicas o moleculares, o caminos para explicar las reacciones químicas. Los químicos computacionales usan los programas y metodologías existentes y los aplican para resolver problemas químicos específicos. ^[12]

Hay dos aproximaciones diferentes a esto:

1. Estudios computacionales para encontrar un punto de partida para la síntesis en laboratorio.
2. Estudios computacionales para explorar mecanismos de reacción y explicar observaciones en reacciones ya vistas en el laboratorio.

Hay varias áreas de desarrollo en la química computacional:

- La representación computacional de átomos y moléculas
- Almacenamiento, organización y búsqueda de datos sobre entidades químicas (bases de datos)
- Identificación de patrones, tendencias y correlaciones entre estructuras químicas y sus propiedades (QSPR)
- Elucidación de estructuras basadas en la simulación de campos de fuerzas
- Modelos para ayudar a la síntesis eficiente de compuestos
- Diseño de moléculas que interaccionen con otras de forma efectiva, especialmente en diseño de fármacos

Los programas usados en química computacional se basan en diferentes métodos mecanocuánticos que resuelven la ecuación de *Schrödinger*. Los métodos que no incluyen parámetros empíricos ni *semi-empíricos* en sus ecuaciones se llaman métodos *ab initio*. Las clases más populares de métodos *ab initio* son: *Hartree-Fock*, teoría de perturbaciones de Moller-Plesset, interacción de configuraciones, *coupled cluster* y teoría del funcional de la densidad (existen ciertas diferencias de opinión sobre si la teoría de los funcionales de la densidad ha de considerarse puramente *ab initio* o *semiempírico*).^[12]

En principio, sería posible usar un método "exacto" (por ejemplo, interacción completa de configuraciones con una base lo suficientemente grande) y aplicarlo a todas las moléculas. Sin embargo, aunque estos métodos son bien conocidos y están

disponibles en muchos programas, el costo analítico computacional crece factorialmente (más rápido que exponencialmente) con el número de electrones de la molécula. Por este motivo, hay un gran número de métodos aproximados que buscan el mejor compromiso en cada caso entre exactitud y coste computacional. ^[12]

Actualmente, la química computacional puede calcular con gran precisión y sin grandes costes las propiedades de moléculas de hasta una decena de electrones. El tratamiento de moléculas que contienen unas pocas docenas de electrones prácticamente sólo es abordable por métodos más aproximados, como la teoría del funcional de la densidad. Los sistemas mayores, como moléculas con varios metales pesados, superficies, sistemas tridimensionales o enzimas, sólo se pueden tratar de forma heterogénea, focalizando el interés en una zona del espacio o en algunos de los electrones y dejando que el resto se reproduzca por pseudopotenciales, cargas eléctricas puntuales, o campos de fuerzas clásicos: los llamados métodos híbridos QM/MM (siglas de las palabras inglesas *Quantum Mechanics / Molecular Mechanics*, *Mecánica Cuántica / Mecánica Molecular*, en español). ^[12]

La química computacional está compuesta de dos grandes áreas basadas en principios físicos distintos. Estas son:

- I). La mecánica molecular (MM) y
- II). Los métodos de estructura electrónica (MEE).

La mecánica molecular (MM)

Se basa en la mecánica clásica, considerando a los átomos como partículas puntuales (dotadas de masa y carga) interaccionando unos con otros mediante enlaces y ángulos que se modelan como resortes, en donde las constantes de estos se ajustan empíricamente.

Los métodos de estructura electrónica (MEE).

Se basan en la aplicación de la Mecánica Cuántica a los sistemas atómicos y moleculares.

1.2.1 Mecánica molecular

La *MM* utiliza las leyes de la física clásica para predecir estructuras y propiedades moleculares.^[12]

Los cálculos en *MM* no tratan los electrones de un sistema explícitamente, pero está optimizada en el cálculo basado en la interacción electrón-núcleo. Los efectos electrónicos se encuentran implícitos en los campos de fuerza. Esto hace que *MM* sea muy eficiente computacionalmente y pueda ser utilizado para hacer cálculos en moléculas muy grandes.

Cada método de *MM* está caracterizado por:

- A.** Un campo de fuerza (FF) particular, compuesto por un conjunto de ecuaciones que define cómo varía el potencial con la posición de los átomos en la molécula.
- B.** Una serie de átomos tipo que definen las características de un elemento en un contexto químico específico. Por ejemplo, el átomo de carbono en un carbonilo es tratado de diferente forma al átomo de carbono enlazado a tres hidrógenos. Por tanto, el tipo de átomo depende de su hibridación, carga y tipo de átomos a los que se encuentra enlazado.
- C.** Uno o más conjuntos de parámetros que ajustan las ecuaciones y tipos de átomos a los datos experimentales. Estos parámetros están definidos como *constantes de fuerza*, las cuales dan los valores utilizados en las ecuaciones que relacionan las

características atómicas con las componentes de energía y los datos estructurales como longitudes y ángulos de enlace.

Estos tres componentes constituyen lo que se conoce con el nombre de campo de fuerza y definen completamente un método de *MM*.

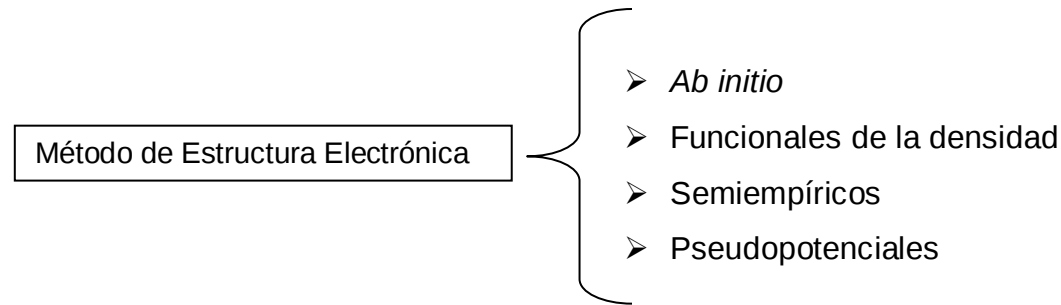
Las aproximaciones contenidas en la *MM* hacen que este método sea sumamente rápido comparado con los *MEE* y, por consiguiente, aplicable a moléculas de un tamaño muy superior a las estudiadas usando métodos químico-cuánticos. Sin embargo, no debe perderse de vista que la *MM* tiene limitaciones importantes, tales como las siguientes:

- Los campos de fuerza particulares consiguen buenos resultados sólo para una clase limitada de moléculas, relacionadas con aquéllas para las que el método fue parametrizado (es decir, no son de aplicación general a cualquier caso de interés).
- El hecho de que los electrones no se consideran explícitamente implica que no pueden tratarse problemas para los cuales los efectos electrónicos sean predominantes. Por ejemplo, la *MM* no puede describir fenómenos de ruptura o formación de enlaces, ni propiedades moleculares que dependan de interacciones entre orbitales moleculares.

1.2.2 Métodos de estructura electrónica

La teoría de los funcionales de la densidad es una reformulación de la mecánica cuántica y utiliza a la densidad electrónica como variable básica, de acuerdo con los teoremas de Hohenberg y Kohn. Por otro lado, los otros *MEE* se basan en la solución de la ecuación de *Schrödinger*. Ambas teorías son exactas en su formulación; sin embargo en el aspecto práctico es necesario realizar aproximaciones en cada una de ella. ^[12]

Por otro lado, los métodos de estructura electrónica se pueden dividir en:



Ecuación de Schrödinger

La ecuación que describe la dinámica de sistemas microscópicos, como átomos y moléculas, es la ecuación de *Schrödinger*^[13]

$$\left\{ \frac{-h^2}{8\pi^2m} \nabla^2 + V \right\} \psi(\vec{r}, t) = \frac{ih}{2\pi} \frac{\partial \psi(\vec{r}, t)}{\partial t}, \quad (1.1)$$

Esta ecuación tiene la forma de una ecuación de onda y surge de la dualidad de comportamiento (como partícula y como onda) que presentan las partículas microscópicas. La función ψ , función de onda del sistema, no es un observable, sino una construcción matemática con la que puede obtenerse la función:

$$|\psi|^2 = \psi \psi^*, \quad (1.2)$$

que se interpreta como la densidad de probabilidad de distribución de las partículas en el espacio. Esta función, a diferencia de ψ sí es un observable.

De aquí en adelante, se considerará que $\mathbf{r}$ es un vector que representa todas las coordenadas de las diferentes partículas que componen el sistema^[13]

$$\psi(\vec{r}, t) = \psi(\vec{r}) \gamma(t) \quad (1.3)$$

Introduciendo la expresión (1.3) en la (1.2) nos permite obtener dos ecuaciones independientes (una para la función espacial y otra para la función temporal). De ellas, la más importante es la ecuación que describe el comportamiento de la función de onda independiente del tiempo

$$\hat{H}\psi(\vec{r}) = E\psi(\vec{r}), \quad (1.4)$$

donde **E** aparece como la constante de separación de las dos ecuaciones en distintas variables, representa la energía total del sistema y $\hat{H}$ es el operador Hamiltoniano del sistema, que se define en *unidades atómicas* como

$$\hat{H} = \frac{1}{2}\nabla^2 + V. \quad (1.5)$$

La ecuación (1.5) presenta una familia de soluciones que corresponden (salvo en caso de degeneración) a distintas energías y estados del sistema. La función de onda que corresponde a la energía más baja posible es la que representa el estado fundamental o basal del sistema. En estas consideraciones, no se han tomado en cuenta los efectos relativistas, que no se revisarán aquí. ^[13]

El Hamiltoniano molecular

Para un sistema molecular, la función de onda depende de las coordenadas de todas las partículas en el sistema. Núcleos y electrones intervienen en el Hamiltoniano molecular (los núcleos se consideran, al igual que los electrones, partículas puntuales sin estructura interna), pero por comodidad en la notación y para luego discutir la aproximación *Born-Oppenheimer*, es usual denominar **r** al vector de las 3n coordenadas de los n electrones y **R** al vector de las 3N coordenadas de los N núcleos. Los elementos individuales de estos vectores serán denotados como **r_i** y **R_i** respectivamente. ^[14]

El Hamiltoniano molecular se expresa como la suma de un operador de energía cinética y otro de energía potencial

$$\hat{H} = \hat{T} + \hat{V} = -\frac{1}{2}\nabla^2 + V. \quad (1.6)$$

En forma explícita los dos operadores tienen, en unidades atómicas, la forma

$$\hat{T} = -\frac{1}{2}\sum_k \left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) - \frac{1}{2}\sum_l M_l^{-1} \left(\frac{\partial^2}{\partial X_l^2} + \frac{\partial^2}{\partial Y_l^2} + \frac{\partial^2}{\partial Z_l^2} \right), \quad (1.7)$$

donde M y Z representan la masa y el número atómico de los núcleos. Los dos términos en el operador de energía cinética representan respectivamente la energía cinética de los electrones y los núcleos, mientras que los tres términos del operador de energía potencial (siguiente ecuación) representan respectivamente la atracción electrón-núcleo, y las repulsiones electrón-electrón y núcleo-núcleo:

$$\hat{V} = -\sum_i^{el} \sum_l^{nuc} \frac{Z_l}{|r_i - R_l|} + \sum_i^{el} \sum_{j < i}^{nuc} \frac{1}{|r_i - R_j|} + \sum_l^{nuc} \sum_{m < l}^{nuc} \frac{Z_l Z_m}{|R_l - R_m|}. \quad (1.8)$$

La aproximación *Born-Oppenheimer*

Núcleos y electrones participan en las ecuaciones que se han mostrado hasta el momento. La masa de los núcleos es mucho mayor que la masa de los electrones (el protón es 1836.1 más masivo que el electrón). Consecuentemente se puede extender la analogía clásica de un movimiento muy rápido (el de los electrones) frente a otro bastante más lento (el de los núcleos) para intentar desacoplar ambos movimientos. Físicamente, lo que esto implica es que los electrones reaccionan rápidamente ante cualquier cambio de la configuración nuclear y que, por lo tanto, la distribución electrónica dentro de un sistema molecular dado depende esencialmente de la posición

de los núcleos y no de su velocidad. En otras palabras, básicamente los electrones se mueven en una superficie de energía potencial generada por los núcleos. ^[15]

Matemáticamente, lo anterior implica la posibilidad de separar ambos movimientos y así podemos escribir la función de onda del sistema como:

$$\psi_{total}(r, R) = \psi_{el}(r; R) \chi_{nuc}(R), \quad (1.9)$$

donde la notación $(r; R)$ implica que la función de onda electrónica depende paramétricamente de la posición de los núcleos. Introduciendo la expresión (1.9) en la ecuación de *Schrödinger* molecular, podemos obtener las ecuaciones en la aproximación de *Born-Oppenheimer*.

Las ecuaciones que se obtienen no son triviales, debido a que se debe realizar un cambio de variables que permita separar el centro de masas del sistema y realizar una serie de manipulaciones. Lo que se acostumbra hacer y que muchas veces es confundido con la propia aproximación de *Born y Oppenheimer*, es la llamada aproximación de los núcleos fijos donde lo que efectivamente se realiza es despreciar la componente de energía cinética de los núcleos. Hecho esto, la ecuación que entonces queda para resolver tiene la forma:

$$H_{el} \psi_{el}(r; R_o) = E_{el}(R_o) \psi_{el}(r; R_o), \quad (1.10)$$

Superficies de Energía Potencial (PES)

La energía electrónica del sistema, obtenida mediante la solución de la ecuación de *Schrödinger* electrónica del sistema, es una función que depende paramétricamente de las coordenadas nucleares y determina la superficie de energía potencial. Es claro, que para cada función de onda electrónica que sea solución del problema de valores propios de la ecuación (1.10), habrá un punto en la superficie de energía potencial. ^[16]

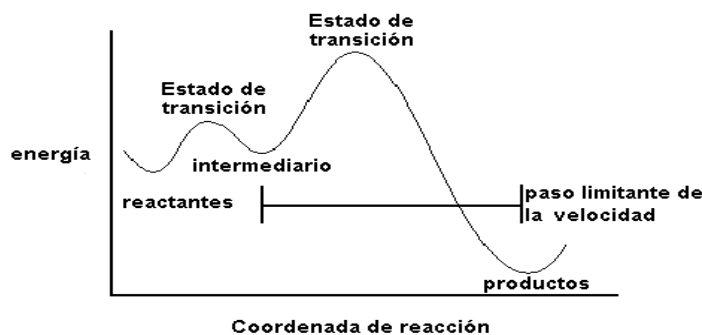


Figura. 1.5 Superficie de energía potencial

La PES de un sistema permite investigar las modificaciones nucleares, ya sea en una reacción química, en un rearrreglo conformacional o simplemente para el estudio de las propiedades vibro-rotacionales del sistema (*figura1.5*). Existen tres temas fundamentales relacionados con las PES y que tienen mucha importancia en química:

- La localización de puntos estacionarios y puntos críticos en la PES;
- La determinación de caminos de reacción;
- El cálculo de trayectorias

Para un sistema molecular de N átomos, luego de la aplicación de la aproximación de *Born-Oppenheimer* (núcleos fijos), nos quedan únicamente $3N-6$ coordenadas independientes ($3N-5$ para moléculas lineales) que representan vibraciones del sistema. Las otras seis (o cinco) coordenadas que teníamos en nuestro conjunto original de $3N$ coordenadas q_i (correspondientes a los N núcleos con vectores de posición $\{R_o\}$) han sido usadas para representar el movimiento del centro de masas del sistema (tres de ellas) y la rotación alrededor de los tres ejes espaciales (dos si se trata de una molécula lineal). La energía electrónica en función de las $3N-6(5)$ coordenadas independientes del sistema, es una hipersuperficie (superficie en $3N-6(5)$ dimensiones en lugar de dos) que presenta ciertas características importantes asociadas a la química-física del sistema considerado. ^[16]

Dos magnitudes matemáticas son de extrema importancia en el análisis de las PES. Una de ellas es el vector gradiente de la energía respecto a las coordenadas nucleares, $\mathbf{q}_i$.

$$g(q) = \left[\frac{\partial E(q)}{\partial q_1} \dots \frac{\partial E(q)}{\partial q_{3N-6(5)}} \right]. \quad (1.11)$$

El cálculo del gradiente no es trivial, de cualquier forma, es claro que en principio este gradiente podría calcularse mediante diferencias finitas de la energía en dos conformaciones nucleares relativamente próximas. El gradiente $g(q)$ indica, como para cualquier otra función, la pendiente de la misma en un punto determinado. Los puntos estacionarios y los puntos críticos es decir, aquellos puntos para los cuales $g(q)=0$ tienen especial importancia en química dado que se asocian con las estructuras de equilibrio y los estados de transición.^[16]

Antes de hablar de los tipos de punto estacionario existentes, necesitamos el segundo elemento matemático que es la matriz Hessiana, la matriz de segundas derivadas de la energía respecto a las $3N-6(5)$ coordenadas nucleares independientes

$$H(\mathbf{q}) = \begin{bmatrix} \frac{\partial^2 E(\mathbf{q})}{\partial q_1 \partial q_1} & \dots & \frac{\partial^2 E(\mathbf{q})}{\partial q_1 \partial q_{3N-6(5)}} \\ \dots & \dots & \dots \\ \frac{\partial^2 E(\mathbf{q})}{\partial q_{3N-6(5)} \partial q_1} & \dots & \frac{\partial^2 E(\mathbf{q})}{\partial q_{3N-6(5)} \partial q_{3N-6(5)}} \end{bmatrix} \quad (1.12)$$

La matriz Hessiana actúa en la misma forma que la segunda derivada de una función de una variable, permitiéndonos determinar la curvatura de la función en el espacio $3N-6(5)$ dimensional. Dado que se trata de una matriz, la forma de obtener la curvatura (en cada una de las $3N-6(5)$ direcciones) consiste en diagonalizar la matriz, con lo cual se obtienen los valores propios, que pueden ser positivos, negativos o nulos y que caracterizan la magnitud de la curvatura en un punto determinado. Los vectores propios de la matriz Hessiana dan las direcciones en que se miden las respectivas curvaturas.^[16]

Para un punto estacionario existe una multitud de posibilidades según los signos de los valores propios de la Hessiana. En particular, un punto para el cual todos los valores propios sean positivos implica que nos encontramos en presencia de un mínimo de la $E(q)$. Este mínimo no es necesariamente el más bajo posible en la hipersuperficie, por lo cual se llama usualmente mínimo local. El mínimo local más bajo de toda la PES, será el mínimo global del sistema. Los mínimos locales en una PES corresponden en química-física a estructuras estables del sistema químico. ^[16]

Dado que todos los valores propios de la Hessiana en un mínimo son positivos, ello implica que al alejarnos de ese mínimo en cualquier dirección sobre la PES, la energía sube. Si analizamos el caso del HCN, es obvio que al haber dos isómeros correspondientes a mínimos locales y al crecer la energía al alejarnos de cualquiera de ellos, debe existir un punto medio en el cual la PES alcance un cierto valor máximo en la dirección que estamos siguiendo y luego vuelva a disminuir. Tal punto existe y constituye un punto de ensilladura de primer orden. Puntos de ensilladura se llaman a todos los puntos estacionarios tales que algunos de los valores propios de la Hessiana sean positivos y otros negativos. El orden del punto de ensilladura está dado por el número de valores propios negativos de la Hessiana, por lo cual el punto de ensilladura de primer orden que mencionamos antes tiene uno y sólo un valor propio negativo. Es fácil demostrar que entre dos mínimos existe siempre un punto de ensilladura de primer orden y, en general, estos puntos de ensilladura se asocian a los estados de transición en la Teoría del Estado de Transición. Es importante señalar, sin embargo, que el verdadero estado de transición para una transformación química, es el máximo de energía en el camino que conecta los mínimos, pero no en la PES, sino en la superficie de energía libre. En forma bastante laxa, sin embargo, se utilizan el término estado de transición en lugar de punto de ensilladura de primer orden. La localización de mínimos y de puntos de ensilladura de primer orden es uno de los problemas más importantes de la química computacional. ^[16]

Los mínimos y el punto de ensilladura de primer orden que los separa se encuentran sobre una curva unidimensional en la PES, que recibe el nombre de camino de reacción. El camino de reacción no es invariante a la definición que se haga del mismo y, en particular, a la elección de las coordenadas sobre las que se expresa la PES. Su importancia radica fundamentalmente en que, de alguna forma, marca el fondo del valle por el que se produce la modificación de las relaciones geométricas de los núcleos en un sistema molecular. Este camino de reacción no es una trayectoria, dado que no tiene componentes dinámicas. ^[16]

Es preciso señalar que hay puntos estacionarios para los cuales uno o más de los valores propios de la Hessiana son nulos. Estos puntos estacionarios reciben el nombre de puntos críticos y tienen importancia en la teoría de las reacciones químicas. ^[16]

Método Hartree-Fock

A partir de la ecuación (1.13) puede verse que la energía potencial incluida en el Hamiltoniano electrónico tiene un término que es independiente de las coordenadas electrónicas (la repulsión nuclear), un término que depende solamente de las coordenadas de un único electrón (la atracción electrón-núcleo) y un término que depende de las coordenadas de todas las parejas de electrones (repulsión interelectrónica). Este último término, utilizando la notación $r_{ij} = |r_i - r_j|$ y para n electrones, se puede representar como:

$$\sum_i^n \sum_{j<i}^n \hat{H}_{ij}^2 = \sum_i^n \sum_{j<i}^n \frac{1}{r_{ij}} \quad (1.13)$$

es el más difícil de tratar desde el punto de vista matemático. Para evitar este problema se recurre a la conocida como aproximación de las partículas independientes. Desde el punto de vista físico, esto implica que en lugar de considerar que los electrones interaccionan uno a uno mediante un potencial coulombico, tal cual se muestra en la

ecuación (1.13), consideramos que los electrones mantienen interacción promedio con todos los otros electrones del sistema. Matemáticamente, esto equivale a sustituir la ecuación (1.13) con la (1.14).^[18]

$$\sum_i^n \sum_{j<i}^n \hat{H}_{ij}^2 = \sum_i^n \hat{V}_i^{av}, \quad (1.14)$$

y, por lo tanto, podemos escribir el Hamiltoniano total del sistema, en la aproximación de los núcleos fijos, en la siguiente forma

$$\hat{H}_{tot} = \hat{H}_{el} + \sum_{a<b}^{núcleos} \frac{Z_a Z_b}{|R_a - R_b|} \quad (1.15)$$

donde, el Hamiltoniano electrónico lo podemos escribir como

$$\hat{H}_{el} = \sum_i^n (\hat{H}_i^l + \hat{V}_i^{av}) = \sum_i^n \hat{F}_i, \quad (1.16)$$

introduciendo así el operador de Fock $\hat{F}$ ^[16-17]. La parte monoelectrónica exacta del Hamiltoniano contiene sólo la energía cinética y la interacción del electrón con todos los N núcleos del sistema, en la forma.^[18]

$$\hat{H}_i^1 = -\frac{1}{2} \nabla_i^2 - \sum_{a=1}^N \frac{Z_a}{|R_a - r_i|}. \quad (1.17)$$

Por lo tanto,

$$\hat{F}(1) = \frac{1}{2} \nabla_1^2 - \sum_{a=1}^N \frac{Z_a}{|R_a - r_1|} + \hat{V}_1^{av}, \quad (1.18)$$

En el método de Hartree-Fock

$$\hat{V}_1^{av} = \sum_{j=1}^n (\hat{J}_j(1) - \hat{K}_j(1)), \quad (1.19)$$

donde $\hat{J}_j(1) = \int \varphi_j^*(2) r_{12}^{-1} \varphi_j(2) d\vec{r}_2$ es el operador de Coulomb y el de intercambio $\hat{K}_j$ se define por $\hat{K}_j(1)\varphi_i(1) = \int \varphi_j^*(2) r_{12}^{-1} \varphi_i(2) \varphi_j(1) d\vec{r}_2$ donde los $\{\varphi_i\}$ son espín orbitales moleculares (Funciones de onda monoeléctricas)

Lo que se ha hecho con la ecuación (1.16) es transformar el problema de tal manera que ahora los operadores de Fock dependen (formalmente) sólo de una partícula. Es claro que la evaluación del potencial promedio V^{av} depende de todos los electrones del sistema y que, debemos realizar el procedimiento descrito para determinarlo. Por el momento, notemos que ahora tenemos el operador de Fock dividido en partes formalmente monoeléctricas, no sucede lo mismo con la función de onda, por lo que debemos ahora analizar su forma y como representarla para sacar provecho de la nueva forma del operador Fock. Esto se logra justamente si la función de onda multieletrónica es un determinante de los espín orbitales $\{\varphi_i\}$. Puede demostrarse, que entonces el problema a resolver está dado por las ecuaciones pseudo Schrödinger monoeléctricas (ecuaciones de Hartree-Fock), ^[16-18]

$$\hat{F}\varphi_i = \varepsilon_i \varphi_i, \quad (1.20)$$

donde los $\{\varphi_i\}$ son los espín orbitales que se determinan computacionalmente con las energías orbitales $\{\varepsilon_i\}$

Conjuntos de funciones de base

Con lo visto hasta el momento es posible inferir que, en principio, es posible resolver esta serie de ecuaciones integrodiferenciales (las ecuaciones de *Hartree-Fock*) cuyas soluciones, las funciones desconocidas φ_i (orbitales moleculares) con energías orbitales ε_i , forman un determinante de Slater que es la función de onda electrónica del sistema, en base a la cual podemos determinar todas las propiedades del mismo. La forma determinantal es requerida para cumplir con el principio de antisimetría y al expandir el determinante se obtiene una serie de productos de espín orbitales. ^[18]

El problema es, sin embargo, todavía muy complicado. Una simplificación notable, introducida por Roothaan^[8], dió origen al método LCAO (Linear Combination of Atomic Orbitals) que es la base de toda la química cuántica computacional basada en la teoría de orbitales moleculares de hoy en día. El método LCAO se basa simplemente en representar cada uno de los orbitales moleculares en la forma ^[16]

$$\varphi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} . \quad (1.21)$$

los coeficientes c en esta expresión se conocen como los coeficientes de desarrollo de los orbitales moleculares y las funciones χ_{μ} (que se eligen normalizadas) representan orbitales atómicos que a su vez se desarrollan en un conjunto de funciones base. Como el nombre del método indica, la idea inicial fue expresar los orbitales moleculares en un desarrollo sobre orbitales atómicos. Sin embargo, esto no es necesario y actualmente se emplean dos tipos de funciones base, las funciones de Slater (exponenciales) y las funciones gaussianas. Las primeras son empleadas fundamentalmente en los métodos *semiempíricos*, mientras que las segundas se emplean en los métodos *ab initio*. Es necesario tener al menos una idea de la estructura de los conjuntos de funciones base gaussianas, porque la elección correcta del conjunto base es uno de los requisitos para el empleo adecuado de los métodos de cálculo más sofisticados. ^[16]

Las funciones de Slater o las funciones gaussianas constituyen propiamente sólo la parte radial de las funciones base. La dependencia angular de las funciones base está dada por los armónicos esféricos en el caso de las funciones de Slater y por el empleo de las llamadas funciones gaussianas cartesianas en el segundo caso. Las funciones gaussianas cartesianas no son estrictamente equivalentes a las funciones gaussianas polares y existen algunos problemas de dependencia lineal que hay que considerar. ^[16]

Las funciones gaussianas que se emplean como funciones base reciben usualmente el nombre de primitivas y a partir de ellas pueden formarse combinaciones lineales fijas que se llaman funciones contraídas. Las funciones contraídas son las que normalmente se emplean como conjuntos de funciones base para los cálculos moleculares. Entonces los orbitales atómicos χ_μ pueden desarrollarse en términos de las gaussianas primitivas del conjunto base seleccionado,

$$\chi_\mu(r) = \sum_p d_{\mu p} g_p(\alpha, r). \quad (1.22)$$

Las funciones primitivas para un determinado átomo suelen expresarse con la notación (is,jp,kd,lf) o, más concisamente, (isjpkdlf). Así, un conjunto base consistente en 14 primitivas de tipo s, 9 de tipo p, 3 de tipo d y 1 de tipo f, se expresaría como (14s9p3d1f). Las funciones contraídas se expresan de la misma forma, pero usando corchetes en lugar de paréntesis, como, por ejemplo, [5s3p2d1f].^[16]

Los conjuntos de funciones base pueden clasificarse con determinados nombres de acuerdo al número y tipo de funciones que incluyan. Enseguida se hace referencia a algunas funciones base y lo que significan.^[16]

Base mínima (MBS)

Contiene exactamente el número necesario de funciones para representar los orbitales atómicos ocupados de los átomos en las moléculas (i.e. 1 para el H, 5 para C, O, etc)

Base Doble Zeta (DZ)

Contiene exactamente dos veces el número de funciones de la MBS (para el H, por ejemplo, contiene 2 funciones 1s)

Base "Split-Valence" (DZV)

Es una MBS para los electrones del core de un átomo y una DZ para los electrones de la capa de valencia (por ejemplo para el carbono tendrá una única función contraída para representar el orbital 1s y dos contraídas para representar cada uno de los orbitales 2s y 2p)

Base extendida

Contiene cualquier número mayor que 2 de funciones para cada orbital atómico. Los ejemplos incluyen la Triple Zeta (TZ) y Cuádruple Zeta (QZ).

Base DZ con Polarización (DZP)

Incluye una base DZ para los orbitales atómicos de core y de valencia e incluye funciones de polarización, es decir, funciones que no están ocupadas en el estado fundamental del átomo considerado, pero que permiten la polarización de la densidad electrónica hacia las zonas del enlace en la molécula. Estas funciones tienen momento angular superior a la de la última función de la capa de valencia (funciones 2p para el H, 3d para los elementos de la primera fila, etc)

Bases con funciones difusas

Estas bases están construidas como las bases normales, pero incluyen funciones con el mismo número angular que las ocupadas pero con exponentes mas pequeños. Estas funciones son llamadas difusas porque se extienden a zonas lejanas al núcleo. Sirven para representar la densidad electrónica en aniones, complejos débiles y estados excitados de moléculas.

Dos familias sistemáticas de conjuntos de funciones de base muy populares son las de Huzinaga-Dunning y las de Pople. Las bases de Huzinaga-Dunning se denotan generalmente como DZ, TZ y DZP, TZP cuando se incluyen funciones de polarización. Los conjuntos de funciones de base de Pople tienen una notación más complicada que se resume a continuación

STO-NG

Son MBS con N gaussianas en combinaciones fijas para cada una de las funciones de la base. (p.ej.: STO-3G).3-21G

6-31G

Son DZV con tres o seis funciones primitivas en cada una de las funciones de la MBS que describe el core de los átomos y 2 o 3 primitivas en la primera de las dos contraídas en la DZ de valencia.

3-21G*, 3-21G (*), 6-31G*, 6-31G(d)

Igual que la anterior pero incluyendo funciones de polarización en los átomos de la primera fila (pero no en los hidrógenos) cuando el asterisco está sin paréntesis o aparece como (d). El asterisco entre paréntesis indica que se usan funciones de polarización desde el Na hacia arriba

6-31G**, 6-31G (d,p), 6-31G(2d,2p)

Igual que la anterior, pero incluyendo también funciones de polarización en los átomos hidrógenos (segundo asterisco). Cuando se incluyen varias funciones de polarización (en lugar de sólo 1) se indica como (nd,mp) con n y m el número de funciones de polarización d y p.

6-311++G (3df, 3pd)

Indica una MBS de seis gaussianas para el core, tres funciones contraídas a partir de 5 primitivas para cada número angular en la capa de valencia, funciones difusas sp en los átomos de la primera fila (el primer +), funciones difusas en los H (segundo +), 3 funciones de polarización d y una f en los átomos de la primera fila y 3 funciones de polarización p y una d en los hidrógenos.

No todos los conjuntos de funciones base mencionados existen para todos los átomos y debe consultarse el programa que se está empleando para conocer cuál es aplicable al caso específico. Existen procedimientos y conjuntos base especiales que se emplean para los métodos de funcionales de la densidad y para tratar metales o átomos de la 3a y 4a filas de la tabla periódica (tales como los potenciales efectivos de core (ECP)) que pertenecen a los pseudopotenciales.

Con esto pueden finalmente resolverse las ecuaciones Hartree-Fock-Roothan, para obtener los orbitales moleculares $\{\phi_i\}$, con la precisión numérica impuesta por el conjunto base seleccionado. Este procedimiento permite determinar el conjunto de coeficientes óptimos $\{c_{\mu i}\}$ para cada orbital molecular.

Una vez que se tiene el conjunto óptimo de $\{\phi_i\}$ es necesario evaluar a partir de estos y del operador de Fock, la energía óptima de *Hartree-Fock*, E_{HF} . El método variacional permite obtenerla para un sistema de capa cerrada mediante ^[16-18]

$$E_{HF} = 2 \sum_{i=1}^{n/2} H_{ii} + \sum_{i>j}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij}) + \sum_{a<b}^N \frac{Z_a Z_b}{|R_a - R_b|}, \quad (1.23)$$

donde

$$H_{ii} = \langle \phi_i(1) | \hat{H}_i^1 | \phi_i(1) \rangle \quad (1.24)$$

$\hat{H}_i^1$ se encuentra dado por la ecuación (1.17) en la notación de Dirac.

$$J_{ij} = \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) \right\rangle, \quad (1.25)$$

y

$$K_{ij} = \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_j(1) \phi_i(2) \right\rangle. \quad (1.26)$$

1.3 Métodos de cálculo

1.3.1 *Semiempíricos*

Estos métodos, utilizan parámetros experimentales para simplificar el cálculo computacional. Estos métodos resuelven en forma aproximada la ecuación de Schrödinger que depende de parámetros apropiados al tipo de sistema químico bajo estudio.

Los métodos *semiempíricos* tienen un bajo costo computacional y proveen una descripción cualitativa razonable, pero la exactitud en la predicción cuantitativa de la energía y las estructuras moleculares depende de qué tan buenos sean los conjuntos de parámetros, el tamaño del sistema y tipo de átomos que lo conformen. Los métodos *ab initio*, a diferencia de los *semiempíricos*, predicen con mayor exactitud desde el punto de vista cuantitativo y cualitativo, pero el costo computacional aumenta sustancialmente.

1.3.2 *Ab initio*

Los MEE *ab initio* que hemos visto hasta aquí (*Hartree-Fock*) tienen el enorme inconveniente de que son muy costosos desde el punto de vista computacional y, consecuentemente, aplicables sólo a moléculas de tamaño reducido. El factor limitante en estos métodos es el tiempo necesario para calcular el gran número de integrales bieletrónicas sobre funciones base (en el caso *Hartree-Fock*) y/o para transformarlas en integrales sobre orbitales moleculares (en los casos post-*Hartree-Fock*). Los métodos de *Mecánica Molecular*, por el contrario, son aplicables a grandes moléculas, pero no tienen en cuenta en forma explícita la estructura electrónica. Consecuentemente, es necesario encontrar alguna metodología intermedia que permita tratar la estructura electrónica de las moléculas pero en forma aproximada, sin los

costosos requisitos de los métodos *ab initio*. Este camino lo proporcionan los métodos *semiempíricos* y los métodos de los *funcionales de la densidad*.

Todos los métodos *semiempíricos* están basados en la teoría *Hartree-Fock* pero realizan aproximaciones que permiten eliminar todas, gran parte o alguna parte de las integrales que es necesario calcular. Las aproximaciones que se realizan modifican la matriz de Fock, que es la forma matricial de las ecuaciones *Hartree-Fock-Roothan* discutidas anteriormente y que expresada en términos de funciones base en la aproximación LCAO tiene la forma

$$F_{\mu\nu} = \langle \mu | \hat{H}_I | \nu \rangle + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\lambda | \nu\sigma) \right], \quad (1.27)$$

donde P es la matriz densidad

$$P = [P_{\mu\nu}] = \left[\sum_{i=1}^N n_i c_{\mu i}^* c_{\nu i} \right], \quad (1.28)$$

y H_I es la parte monoelectrónica del Hamiltoniano. Los métodos *semiempíricos* utilizan un Hamiltoniano más simple que el Hamiltoniano molecular correcto expresado en la ecuación (1.15) y utilizan parámetros cuyos valores se ajustan a los datos experimentales o los resultados de los cálculos *ab-initio*. El más antiguo de todos los métodos *semiempíricos*, antecesor de todos los métodos modernos de estructura molecular es el método de Hückel (HMO). ^[16,18]

En este caso se realizan las siguientes aproximaciones:

- Se tratan únicamente los electrones π
- Se ignora la repulsión Coulómbica entre los electrones π
- Se asumen valores fijos para las integrales restantes

$$\langle \mu | \hat{\mathbf{H}}_1 | \nu \rangle = \begin{cases} \alpha, & \mu = \nu \\ \beta, & \mu \text{ y } \nu \text{ en átomos vecinos} \\ 0, & \text{en cualquier otro caso} \end{cases} \quad (1.29)$$

donde las constantes α y β son, respectivamente, la integral coulombica de atracción electrón-núcleo y la integral de resonancia, ambas obtenidas empíricamente. La solución de las ecuaciones del HMO se reduce a la diagonalización de una matriz compuesta exclusivamente por 0, α y β y las energías orbitales quedan expresadas exclusivamente en términos de estas constantes. Hoffman extendió el método HMO a lo que es conocido hoy en día con el nombre de EHT (Extended Hückel Theory). A diferencia del HMO, EHT trata también los electrones σ e incluye la geometría de la molécula. ^[16]

Los métodos *semiempíricos* actuales se basan en la aproximación ZDO (zero-differential overlap) que implica

$$\varphi_\mu(r)\varphi_\nu(r) = 0 \quad \text{si } \mu \neq \nu \quad (1.30)$$

lo que origina no solamente la ortonormalidad de los orbitales (la matriz de traslape es igual a la matriz unidad) sino también la anulación de muchas integrales bielectrónicas

$$(\mu\nu|\lambda\sigma) = \delta_{\mu\nu}\delta_{\lambda\sigma}(\mu\mu|\lambda\lambda), \quad (1.31)$$

esta aproximación y el hecho común de tratar únicamente los electrones de valencia son la base de todos los métodos *semiempíricos* modernos, de los cuales se puede mencionar:

- CNDO: Complete Neglect of Differential Overlap, en desuso^[19]
- INDO: Intermediate Neglect of Differential Overlap, sobrevive con dos parametrizaciones diferentes, INDO/1 para geometrías (no usada frecuentemente) e

INDO/S parametrizado para reproducir propiedades espectroscópicas, sobrevive mayormente en la versión de Zerner, ZINDO/S^[16]

- MINDO: Modified Intermediate Neglect of Differential Overlap, actualmente en desuso.^[16]

- MNDO: Modified Neglect of Differential Overlap, actualmente en desuso excepto cuando no existe parametrización específica en los métodos más modernos ^[16]

- AM1: Austin Model 1, uno de los dos métodos más modernos en uso, debido a Dejar^[19]

- PM3: Parametric Method Number 3, el segundo de los dos métodos más modernos en uso, debido a Stewart^[18]

Además de las derivadas de la ZDO, todos estos métodos usan parámetros específicos que les permiten determinar más o menos correctamente distintas propiedades de distintos grupos de moléculas. Así, MNDO, AM1 y PM3 están parametrizados para reproducir calores de formación, mientras que INDO/S (e incluso una versión del CNDO, CNDO/S) está parametrizado para reproducir transiciones electrónicas en la región UV-Vis mediante cálculos CI sobre la función de onda calculada semiempíricamente. Además, MNDO, AM1 y PM3 son distintas parametrizaciones (con algunas diferencias formales) de aproximadamente la misma metodología.

1.3.3 Teoría de los Funcionales de la densidad

La teoría de los Funcionales de la Densidad (DFT) es similar a los métodos *ab initio* en muchas formas; requiere fuentes similares de cálculo a la teoría *Hartree-Fock*, y tiene un costo computacional bastante cercano a los métodos *ab initio*.^[13]

La DFT estudia la estructura de átomos y moléculas utilizando a la densidad electrónica $\rho(r)$ como variable básica, de acuerdo al teorema de Hohenberg y Kohn.

Puede llegar a obtenerse la ecuación básica de la DFT en función de la densidad electrónica.^[14]

$$E[\rho] = F[\rho] + \int \rho(r) v(r) dr \quad (1.32)$$

siendo:

$E[\rho]$ la energía total del sistema como un funcional de la densidad. $v(r)$ el potencial externo impuesto por los núcleos y $F[\rho]$ es el llamado funcional universal y está definido como

$$F[\rho] = T[\rho(r)] + V_{ee}[\rho(r)], \quad (1.33)$$

donde:

$T[\rho(r)]$ es el funcional de la energía cinética y $V_{ee}[\rho(r)]$ es el funcional de la energía de interacción electrón-electrón.

Las ecuaciones fundamentales de la TFD son:

$$dE = \mu dN + \int \rho(r) dv(r) dr, \quad (1.34)$$

$$d\mu = \eta dN + \int f(r) dv(r) dr, \quad (1.35)$$

Estas ecuaciones son bastante importantes pues en ellas aparecen los parámetros μ , η y $f(r)$ necesarios para el estudio de reactividad.

Potencial químico (μ)

Se define como la tendencia de escape de los electrones de una especie molecular a otra^[17] y se encuentra que está dado por $\mu = (\partial E / \partial N)_v$. El potencial químico se puede evaluar de manera aproximada por diferencias finitas obteniéndose.^[15]

$$\mu = -(I + A)/2, \quad (1.36)$$

donde I es el potencial de ionización y A es la afinidad electrónica

Dureza absoluta (η)

La dureza se define como la separación que existe entre el orbital más alto ocupado y el más bajo desocupado. Si la dureza es grande la especie es estable y si es pequeña la especie es reactiva. La dureza se puede determinar de la siguiente forma.^[14]

$$2\eta = (\partial\mu/\partial N)_v = (\partial^2 E / \partial N^2)_v, \quad (1.37)$$

y se puede evaluar de manera aproximada por diferencias finitas y se obtiene $\eta = I - A$.

Función Fukui

Esta función se encuentra relacionada con los orbitales frontera cuando se utiliza la aproximación de core congelado, dando información de cómo se comporta la densidad del sistema cuando existe un cambio en el número de electrones.^[16-17]

$$f(r) = (\partial\rho(r)/\partial N)_v, \quad (1.38)$$

$f(r)$ tiene valores punto a punto y es por lo tanto una propiedad local

La integración de la función de Fukui cuando se lleva a cabo en regiones atómicas nos lleva a las llamadas Fukui condensadas, o sea a una Fukui evaluada en el i -ésimo átomo:

$$f_i^+ = q_i(N+1) - q_i(N), \quad \text{relacionada con los ataques electrofílicos}$$

$$f_i^- = q_i(N) - q_i(N-1), \quad \text{relacionada con los ataques nucleofílicos}$$

$$f_i^0 = \frac{1}{2} [q_i(N+1) - q_i(N-1)], \quad \text{relacionada con los ataques por radicales libres.}$$

siendo q_i la carga electrostática en el i -ésimo átomo

Blandura local (s)

El inverso de la dureza es la blandura y se puede determinar como:

$$S = (\partial N / \partial \mu)_v, \quad (1.39)$$

si se define la blandura local como^[19]

$$s(r) = [\partial \rho(r) / \partial \mu] v, \quad (1.40)$$

se obtiene junto con la función Fukui

$$s(r) = f(r) S, \quad (1.41)$$

Donde **S** es blandura global del sistema $S = \mathbf{1} \left(\frac{1}{I - A} \right)$

I primer potencial de la ionización y *A* afinidad electrónica, uno puede definir semejantemente la blandura local condensada usando las Eqs (1.36 y 1.37) en (1.38) obteniendo respectivamente $s^-_j(r)$ y $s^+_j(r)$.

1.4 Utilización del método *ONIOM*

Desarrollo y uso del método de *ONIOM*

A partir de los métodos *ab initio* de la Mecánica Cuántica, se pueden predecir las energías, estructuras y frecuencias vibracionales de los sistemas moleculares.^[28]

El método se basa en una selección (sistema modelo) de una parte reactiva en un sistema grande, que es delimitado. En el sistema modelo, la parte externa es substituida por los átomos del acoplamiento (generalmente hidrógeno) ubicados a una distancia conocida (*R*). Los métodos tipo QM/MM, utilizan una aproximación armónica para analizar la dependencia de *ONIOM* hasta cierta distancia del átomo del acoplamiento, también llamada *R*₂. La que es proporcional a la distancia del átomo que es sustituido por el hidrógeno (*R*₃) con un factor (*g*).^[28]

$$R_2 = g R_3 \quad (1.42)$$

El sistema se divide según el propósito de cada región y el nivel del interés en la región dicha.

La energía obtenida mediante el método *ONIOM* es una ecuación simple:

$$E_{\text{ONIOM}} = E_3 + E_2 - E_1 \quad (1.43)$$

Donde E_3 = energía del sistema entero con los cálculos del nivel bajo, E_2 = energía del sistema modelo con los cálculos del alto nivel, E_1 = energía del sistema modelo con cálculos del nivel bajo. [28]

La aproximación armónica se refiere a la variación de la energía del cálculo modelo de los sistemas ($E_i = E_2$ o E_1) en función de R_2 . La energía se supone que varía cuadráticamente (con una geometría del equilibrio para el átomo del acoplamiento ($R_{2,0,i} = x_{0,i}$) y las constantes de una fuerza (k_i). [28]

El método de mecánica molecular es el más simple dentro de ONIOM. Se optimiza la región de QM mientras que consideran los efectos estéricos, las cargas formales, y las interacciones de Van der Waals de la región. La interacción electrostática entre la región descrita con mecánica molecular y la región descrita con mecánica cuántica permanece igual durante la optimización [28]

$$V_c^{\text{real-model}} = V_c^{\text{MM,real}} - V_c^{\text{MM,model}} = \sum_{k=1}^{N^{\text{real}}} \sum_{l>k} \frac{q_k^{\text{real}} q_l^{\text{real}}}{r_{kl}} - \sum_{m=1}^{N^{\text{model}}} \sum_{n>m} \frac{q_m^{\text{model}} q_n^{\text{model}}}{r_{mn}} \quad (1.44)$$

Donde q son las cargas atómicas parciales de MM, el subíndice i indica los electrones, el subíndice N el punto carga y el subíndice J los átomos de QM en la ecuación (1.44). [28]

El método de estructura electrónica

Los métodos de estructura electrónica (EE) es una intensificación de la mecánica molecular. La interacción electrostática entre las regiones cambia sobre el curso de la optimización. Esto es posible modificando al operador Hamiltoniano de la región de QM para explicar cargas parciales.

$$\hat{H}_V^{model} = \hat{H}_0^{model} - \sum_i \sum_N \frac{q_N}{r_{iN}} + \sum_J \sum_N \frac{Z_J q_N}{r_{JN}} \quad (1.45)$$

Esquema de ONIOM.

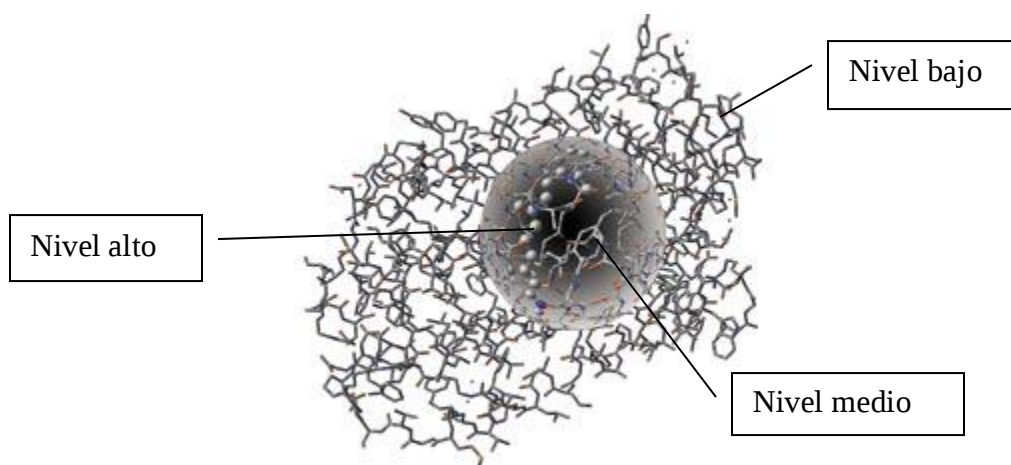


Figura 1.6 Modelo ONIOM

El sistema, como se muestra en la figura 1.6 proporciona la representación de como se selecciona cada una de las capas y posteriormente asignan los niveles de teoría y se escoge los métodos que deseamos utilizar. ^[29]

ONIOM puede ser subdividido más a fondo dependiendo de cómo se configure. Se ha utilizado tradicionalmente dos tipos de método: Lo métodos de Mecánica Molecular (ME) y los de Estructura Electrónica (EE). Recientemente se ha reportado un tercer

método conocido como Mecánica Molecular Polarizable desarrollado por el Dr. Khavrutskii y Vreven.^[29]

1.5 Justificación

En la actualidad ha surgido una enorme necesidad de conocer los factores electrónicos que ocasionan la diabetes, con la intención de desarrollar mejores tratamientos de los que disponemos actualmente. A pesar de que se cuenta con una gran cantidad de información experimental, hasta donde sabemos no se ha encontrado la causa precisa de este padecimiento o desorden metabólico. Pocos son los estudios de los factores que intervienen a nivel molecular en la aparición de la diabetes. Un conocimiento a nivel fundamental de cómo se desarrolla el metabolismo de los carbohidratos y cual es el papel de la insulina en este proceso, proveerá información importante de cuales son los factores que provocan este padecimiento. Así mismo el conocimiento de cómo estos factores influyen en la modificación de la cantidad de azúcar en la sangre, brindará nuevas oportunidades para desarrollar nuevos tratamientos más eficaces contra la diabetes.

Un estudio experimental de los factores que intervienen a nivel molecular de cómo se lleva a cabo el metabolismo de los carbohidratos y su relación con la insulina 9INS, actualmente es muy complicado. A este respecto las teorías aceptadas de la química cuántica computacional, proveen métodos eficaces que se pueden utilizar ampliamente en el estudio de procesos que se desarrollan a nivel molecular, tal como el de diabetes. Por lo tanto, en el presente trabajo se propone el estudio a nivel molecular (teórico- cuántico), utilizando los fundamentos de la química cuántica, en la interacción entre la glucosa y la insulina.

1.6 Objetivos

1.6.1 Objetivo general

Realizar un estudio en la determinación de los sitios activos en la interacción glucosa- insulina utilizando métodos de química cuántica computacional.

1.6.1.1 Objetivos específicos

1. Estudiar los factores electrónicos entre glucosa y la insulina.
2. Estudiar las propiedades electrónicas de la glucosa.
3. Estudiar las propiedades electrónicas de la insulina.
4. Estudiar la interacción entre la glucosa y la insulina.

1.7 Referencias

- [1] R.S Root-Bernstein, P.F and Dillon,. (1997). *Molecular complementarity, I: The molecular complementarity theory of the origin and evolution of life*. J. Theor. Biol., Salud y Enfermedad. Ecología humana, Epidemiología, Salud pública, Medicina preventiva. Sociología y Economía de la salud. México. Pág. 188
- [2] A.Hunding, D.Lancet, , Kepes, F., Minsky, A., Norris, V., Raine, D., Sriram, K. and Root-Bernstein, R. (2006). *Compositional complementarity and prebiotic ecosystems in the origin of life*. Bioessays, Pág. 28
- [3] R.S. Root-Bernstein, (2005). *Peptide self-aggregation and peptide complementarity as bases for the evolution of peptide receptors: A review*. J. Molec. Recognit., 18 Murray
- [4] P. Anzenbacher, and V. Kalous, (1975). *Binding of D-glucose to insulin*. Biochim. Biophys. Acta., 386
- [5] G.B. Yu, and D.L. Caspar, (1998). *Structure of cubic insulin crystals in glucose solutions*. Biophys. J., 74
- [6] V.Zoete, M. Meuwly, and Karplus, M. (2004). *Investigation of glucose binding sites on insulin*. Proteins. Pág 55
- [7] J. A. Lozano Teruel (1995) *Bioquímica para ciencias de la salud*. España: McGraw-Hill Interamericana
- [8] Shiry Rashid, D Kristine. Uffelman, F Gary. Lewis. (2002). *The mechanism of HDL lowering in hypertriglyceridemic, insulin-resistant states*. Journal of Diabetes and Its Complications 1. Pág. 24–28
- [9] Catherine Parker Antony, Norma Jane Kolthoff (1997). *Anatomía y Fisiología*. 9^{na} Edición. Ed. Interamericana. México. Pág. 421-425
- [10] L.Nelson M. David, Cox Michael. (1992). *Principios de bioquímica Leninger*. Madison, Wisconsin. Pág. 400-402
- [11] Mayes, Granner,Rodwell. *Bioquímica de Harper*. Pág. 178
- [12] M. Toporek. *Bioquímica*. 3^a Edición. Ed. Interamericana. Pág. 184-196
- [13] D. B. Boyd. y K. B. Lipkowitz. (1982). *J. Chemistry. Education*. 59 pág. 269

- [14] R. G. Parr y Weitao Yang. (1989). *Density-Functional Theory of Atoms and Molecules*. Ed. Oxford Science Publications. USA.
- [15] Ira N. Levine. (1991) *Quantum Chemistry*. 4ta. Edición. Ed. Prentice Hall.. USA
- [16] S. M. Blinder. (1965) *American Journal of Physics*. 33. pág 431.
- [17] J.J.P. Stewart. (1989) *Journal of Computational Chemistry*. 10 pág 209.
- [18] J. A. Pople y G. A. Segal, (1965) *J. Chemistry. Physics.*, 43 S136; (1966) pág 289.
- [19] M. J. S. Dewar y W. Thiel, (1977) *J. Am. Chemistry. Soc.*, 99. 4899.
- [20] W. E. Pickett. (1989) *Pseudopotential Methods In Condensed Matter Applications. Complex Systems Theory Branch, Naval Research Laboratory. Washington, DC 20375-5000, . USA. North-Holland.*
- [21] D. Vanderbilt. (1985) *Phys. Rev. B* 32. 8412.
- [22] P. J. Hay y Wadt, (1985) *J. Physics. Chemistry*. 82. 270.
- [23] W. Khon y L.J. Sham. (1965) *Physics. Rev.* 140 A1133.
- [24] R. G. Parr y R. A. Donnelly. (1964) *Physics Rev. B*. 136. 864.
- [25] R. G. Parr y R. G. Pearson, (1983) *J. Am. Chemistry. Soc.* 105.7512.
- [26] R. G. Parr y W. Yang, (1981) *J. Am. Chemistry. Soc.* 105. 4049.
- [27] W. Yang, R. G. Parr y R.J. Pucci. (1984) *Chemistry. Phys.* 81. 2862.
- [28] www.softwarecientifico.com/paginas/gaussian.htm. Consultado el 21 octubre 2009
- [29] spice.chem.emory.edu/students/rahim/ONIOM.html. Consultado el 11 julio 2007

Capítulo 2

Metodología

2.1 Introducción

Se denomina modelado molecular al conjunto de métodos que pretenden predecir las propiedades y el comportamiento de los sistemas químicos en forma numérica apoyándose en entornos gráficos para la manipulación de moléculas. ^[1]

El modelado molecular es una técnica de estudio relativamente nueva que relaciona la estructura molecular y las propiedades químicas macroscópicas de la materia. La base del modelado molecular es el uso de métodos computacionales para simular, explicar o predecir la estructura tridimensional y las propiedades fisicoquímicas de la molécula con la ayuda esencial de métodos gráficos. Estas técnicas se encuentran dentro de la química computacional.

Es bueno señalar que los términos Modelado Molecular, Química computacional, Química cuántica y Química Teórica no tienen el mismo significado:

Química teórica:

Es la disciplina que estudia todos los métodos teóricos que pueden aplicarse a la química. Contiene entonces a las otras tres disciplinas, pero contiene también otras disciplinas o técnicas que no se comprenden en ellas.

Química cuántica:

Es la disciplina que estudia la aplicación de la Mecánica Cuántica al estudio de las moléculas. Técnicas tales como la Mecánica Molecular no constituyen parte de ella.

Química Computacional:

Es la disciplina que estudia el uso de métodos numéricos aplicados a la Química. Temas tales como la solución de sistemas diferenciales que se obtienen de la solución de sistemas de ecuaciones químicas están incluidos aquí. Temas como la aplicación de teoría de grafos o topología a la consideración de procesos químicos no están incluidos aquí.

Modelado molecular:

En principio, modelado molecular fue el nombre dado a la aplicación de la Mecánica Molecular para la determinación estructural. Hoy en día se incluye también el uso de los métodos de la Química cuántica y de la Mecánica estadística.

Se caracteriza por el uso de técnicas gráficas para la representación de la estructura y reactividad de la molécula

Para realizar el modelado molecular se recurre a programas de computación más o menos sofisticados que incluyen una o varias de las siguientes funcionalidades:

- a) Permiten el cálculo de energías moleculares y propiedades asociadas con ella
- b) Permiten el cálculo de la estructura tridimensional de las moléculas
- c) Permiten la visualización de las moléculas y de sus propiedades
- d) Permiten el análisis cuantitativo de los resultados del modelado
- e) Permiten la aplicación de todos los métodos antedichos a problemas reales

De aquí que los componentes del modelado molecular sean

- a) Los métodos de cálculo (*empíricos*^[2], *semiempíricos*^[3], *ab initio*^[4-5], *funcionales de la densidad*^[6], etc.) forman el corazón de un verdadero sistema de modelado molecular. Con ellos se calculan las energías (calores de formación, energías electrónicas, electrostáticas, etc.) y otras propiedades tales como momentos dipolares, densidad electrónica, etc.
- b) Los procedimientos de manipulación de la geometría molecular permiten el estudio de las diferentes conformaciones de un sistema. Al acoplarse con los métodos de cálculo, permiten el estudio de las propiedades moleculares dependientes de la conformación
- c) Los métodos gráficos, son el tercer gran pilar en que se asienta el modelado molecular. Permiten la interpretación visual de las conformaciones y las propiedades moleculares en forma atractiva.
- d) Finalmente, los procedimientos de análisis permiten interpretar los datos producidos por algunos componentes del sistema. Por ejemplo analizar el grado de flexibilidad o de rigidez que presentan las distintas zonas de la macromolécula
- e) Todas las funcionalidades expresadas pueden utilizarse en el estudio de sistemas químicos o bioquímicos complicados, como puede ser; por ejemplo, los estudios de la farmacología teórica o los estudios de reacciones químicas.

2.2 Método de calibración

Las moléculas de glucosa e insulina se obtuvieron de la base de datos libre protein databank.

En la tabla 2.1 se puede observar las moléculas de la glucosa, se les realizó una optimización de geometría utilizando un nivel de teoría: AM1, MM, HF/6-31G con el programa de Spartan^[7]. Para analizar las diferentes estructuras moleculares de la

glucosa y cual podría interaccionar con la insulina, se identificaron las zonas más reactivas realizando un mapeo del HOMO y LUMO de cinco moléculas: 6-desoxi-D-glucosa (forma lineal), dos moléculas de β -D-glucosa y dos α -D-glucosa la diferencia entre ellas es espacialmente por lo que tienen propiedades distintas (ver tabla 2.1 propiedades fueron visualizadas con el software Spartan^[7]).

Tabla 2.1 Estructuras primarias de la glucosa. Código de colores: el color blanco átomos de hidrógenos, el rojo para el átomo oxígeno y el gris para el átomo de carbono.

H	O	C
---	---	---

Modelos utilizados	Propiedades
 6-desoxi-D-glucosa (forma lineal)	gas: -608.472 au HOMO: -11.35 eV LUMO: 3.99 eV dipolo: 2.43 debye centro quirál: 4 masa: 164.16 g/mol área de superficie: 190.18 Å² volumen: 151.24 Å³ www.biocheminfo.org/klotho/pdb/6-deoxy.D-g.pdb
 α-D-glucosa	gas: -683.319 au HOMO: -10.97eV LUMO: 5.10 eV dipolo: 4.86 debye centro quirál: 5 masa: 180.16 g/mol área de superficie: 188.68 Å² volumen: 155.36 Å³ www.chem.ccu.edu.tw/cyhung/wwwhighschool/chapter-9/a-glucose.pdb

 α-D-glucosa	gas: -683.326 au HOMO: -11.44 eV LUMO: 5.14 eV dipolo: 2.17 debye centro quiral: 5 masa: 180.16 g/mol área de superficie: 191.90 Å² volumen: 155.52 Å³ www.ch.cam.ac.uk/magnus/molecules/misc/glucose.pdb
 β-D-glucosa	gas: -683.324 au HOMO: -11.43 eV LUMO: 5.21 eV dipolo: 2.43 debye centro quiral: 5 masa: 180.16 g/mol área de superficie: 195.94 Å² volumen: 155.79 Å³ www.chem.ccu.edu.tw/cyhung/wwwhighschool/chapter-9/b-glucose.pdb
 β-D-glucosa	gas: -683.318 au HOMO: -11.08 eV LUMO: 4.96 eV dipolo: 5.44 debye centro quiral: 5 masa: 180.16 g/mol área de superficie: 184.29 Å² volumen: 154.79 Å³ www.rpi.edu/dept/molbiochem/MBweb/pdb/beta-d-glucose.pdb

2.2.1 Mecánica Molecular

La molécula de insulina está formada por aproximadamente 778 átomos lo que hace imposible realizar un cálculo de estructura electrónica y con un alto nivel de teoría con recursos computacionales modestos. Si consideramos que la molécula a estudiar es muy grande, podemos hacer la suponer de que la interacción con una molécula pequeña se realizará solo en una sección muy localizada de la misma. Bajo este supuesto es claro que la modificación de las propiedades electrónicas se realizará a un nivel local. En tal situación considerar cada vez más átomos en esta región local hará más exacto el cálculo. Sin embargo, el costo computacional aumentara drásticamente, A este respecto el método *ONIOM* ofrece ventajas importantes para resolver la situación anterior.

La determinación de las propiedades electrónicas de la insulina 9INS, se realizó utilizando métodos de química cuántica computacional seleccionando 20 átomos para el nivel alto y 20 átomos para el nivel medio como se muestra en la figura 2.1, el resto de la molécula representa el nivel bajo. Con la intención de determinar el nivel de teoría a utilizar en este estudio, se realizó el proceso de calibración utilizando tres niveles.

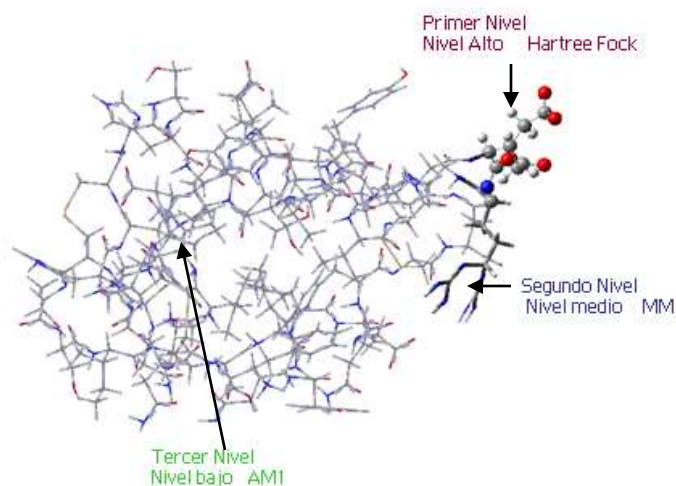
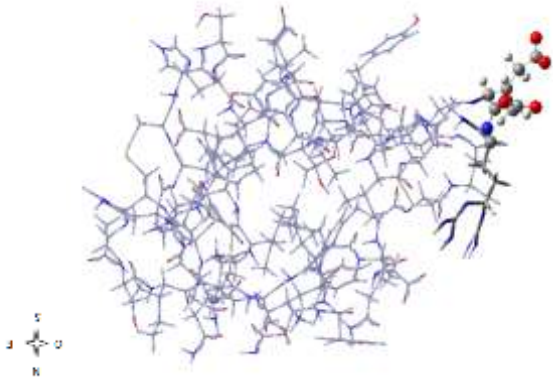
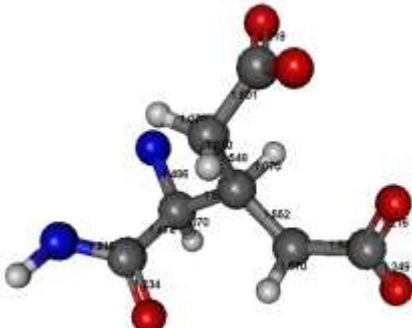
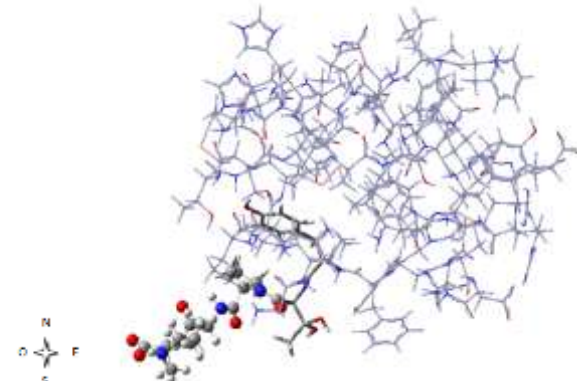
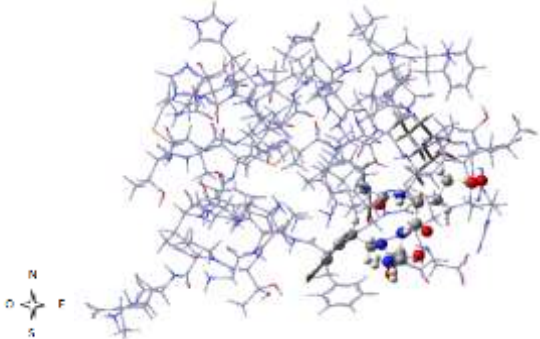
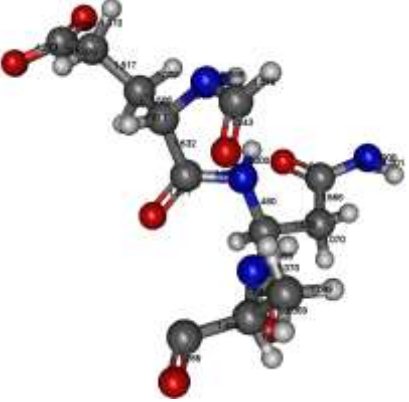
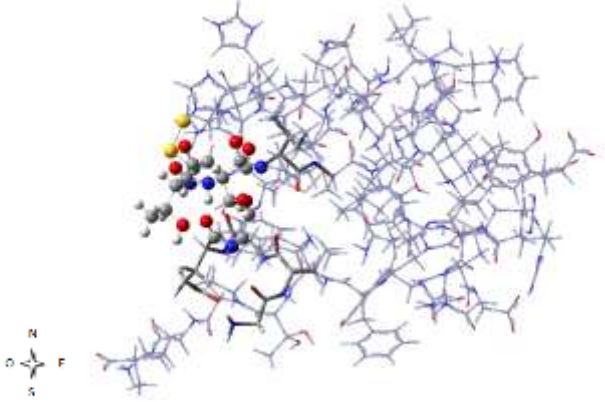
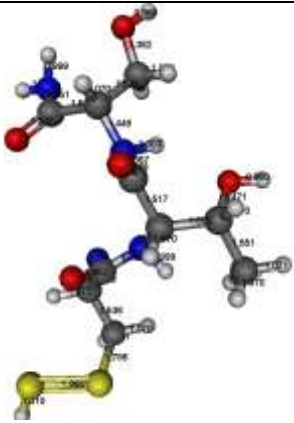
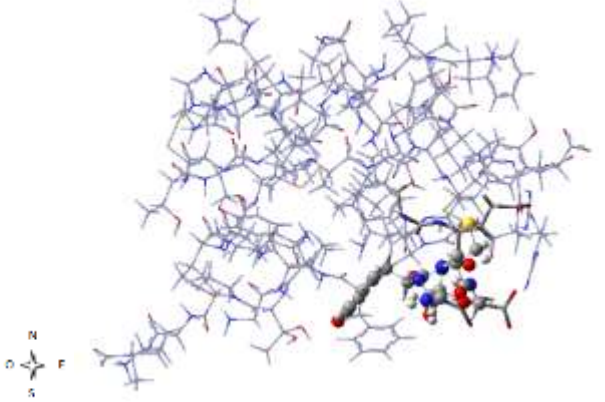
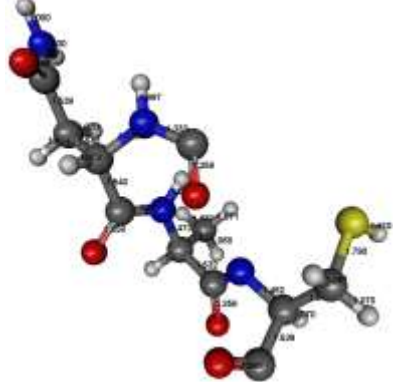


Fig. 2.1 Estructura tridimensional de la insulina (9INS). Modelo ONIOM. Sección 1.

Se identificaron las zonas más reactivas y se asignaron en 5 secciones diferentes de la molécula como se muestra en la tabla 2.2. Utilizando tres regiones, la zona de interés por el método *Hartree-Fock* y para la zona cercana *MM* y por ultimo para el nivel más bajo *AM1*. El método ONIOM calcula la energía para la región del nivel bajo, recalcula la región del nivel alto usando *Hartree-Fock*, y después compara estas energías y encuentra la relación entre ellas. Entonces el programa calcula la energía para el sistema entero (las regiones del nivel alto, y bajo juntos). Toma esta energía y aplica la relación encontrada en el primer cálculo extrapolando la energía que debe haber en el modelo entero calculado con *Hartree-Fock*.

Tabla 2.2 Selección de capas utilizando el método ONIOM para el nivel alto HF/6-31, para el nivel medio *MM* y por último para el nivel bajo *AM1*.

No. Sección	Modelo ONIOM	Estructura optimizada
1		
2		

3		
4		
5		

Código de colores: el átomo de nitrógeno está representado por el color azul, el amarillo para el átomo de azufre, gris átomo de carbono, rojo átomo de oxígeno y blanco para el átomo de hidrógeno.

H	O	C	N	S
---	---	---	---	---

2.2.2 Semiempíricos

La insulina 9ins fue optimizada en el nivel *PM6*; durante la optimización de geometría no se consideró restricción de ningún tipo. Para todos los cálculos se utilizó una estación de trabajo con 2 procesadores Intel Xeon a 2.4 GHZ, con 1GB de memoria RAM. Los cálculos se realizaron con el software *MOPAC*^[9] 2009 versión Linux y fueron visualizados con *Aguslab* y *Gausview*.

2.2.3 Ab initio

Todos los cálculos de la glucosa se realizaron a un nivel de teoría HF/6-31G. la optimización de las cinco estructuras moleculares se realizó con el software *Spartan*^[7], Las propiedades electrónicas como mapeo de HOMO y LUMO, análisis vibracional y la dureza, se determinaron con *Single point* y fueron visualizados con *GausView*^[8].

2.2.4 Funcionales de la densidad

Describen los índices de reactivad química en un rango muy amplio de aplicaciones. Parr y Yang mostraron que los sitios atómicos en las especies químicas con los valores más grandes de $f(r)$ son aquellos con la reactivad más alta. Uno puede definir las correspondientes funciones condensadas o funciones Fukui atómicas en el j -ésimo átomo como. ^[9-11]

$$f_i^+ = q_i(N+1) - q_i(N) \quad (2.1)$$

$$f_i^- = q_i(N) - q_i(N-1) \quad (2.2)$$

$$f_i^0 = \frac{1}{2}[q_i(N+1) - q_i(N-1)] \quad (2.3)$$

Para un ataque nucleofílico, electrofílico o por radicales libres, respectivamente, para la molécula en estudio. [12-13]

Estos índices locales de reactividad, se han aplicado con éxito para racionalizar la reactividad en diferentes situaciones químicas. En este trabajo usamos dos aproximaciones a la función Fukui para evaluar la reactividad de los sitios activos. Los mapas del HOMO (aproximación de core congelado a la $f(r)$) y la blandura local condensada. [14-15]

Para obtener conocimiento sobre las observaciones (factores electrónicos como: dureza, funciones Fukui entre otros) experimentales mencionadas y reducir el costo computacional, hemos realizado un estudio para el caso específico de 5 secciones de la insulina 9INS sin tomar en cuenta el resto de la molécula, para encontrar el sitios activos calculando las funciones Fukui obtendremos los valores de f_i^+ y f_i^- . [16] Ver anexo 1 comparándolo con un mapeo del HOMO y LUMO de cada sección de la estructura molecular.

Recursos computacionales

En todos los cálculos se empleó un cluster Beowulf con 16 procesadores de 3.0 GHz cada uno y 1 GB de memoria RAM.

2.3 Referencias

- [1] T. Clarck A. (1985) *Handbook of Computational Chemistry*. John Wiley and Sons.USA.
- [2] D.B. Boyd. y K. B Lipkowitz.(1982) *J. Chem. Education*. 59. pág 269
- [3] J.J.P. Stewart. (1989) *Journal of Computational Chemistry*. 10 pág 209
- [4] S.M. Blinder.(1965) *American Journal of Physics*.33 pág 431
- [5] W. J. Hehre, L. Radom, P.V.R. Shleyer y J.A. Pople. (1986) *Ab initio Molecular Orbital Theory*, Ed Wiley, New York.
- [6] R.G. Parr y Weitao Yang.(1989) *Density- Functional Theory of Atoms and Molecules*. Ed. Oxford Science Publications. USA.
- [7] Spartan version 5.1.1,(1999), *Wavefunction*. Inc. 18401 Von Karman Ave., Suite 370. Irvine, CA.
- [8] Gaussview Rev. 3.09 windows version. Gaussview Inc., Pittsburgh, PA.
- [9] MOPAC2009, James J. P. Stewart Computational chemistry, version 8.032L
- [10] R. G. Parr and W. (1989)*Yang Density Functional Theory of atoms and molecules*. New York Oxford University Press,
- [11] B.G. Backeback, A. Cedillo y R. G. Parr (1952). *J. Chem. Phys.* 20
- [12] R.G. Parr and W. Yang. (1984) *J. Am. Chem. Soc* 106
- [13] W. Yang y W. Morber (1986) *J. Am.Chem. Soc.* 108
- [14] F. Güardoni, J. Weber y T.R. Ward (1998) *J. Phys.Chem. A*. 102
- [15] C. Lee, W. Yang y RG. Parr (1998) *J. Mol. Struct. (Teochem)*, 163
- [16] R. G. Parr and W. Yang. Ann. Rev (1995). *Phys. Chem.*46

Capítulo 3

Propiedades electrónicas de la glucosa

3.1 Introducción

La *glucosa* ($C_6H_{12}O_6$) es un azúcar utilizada por las células para obtener energía al combinarse con el oxígeno de la respiración. Se obtiene fundamentalmente a través de la alimentación y se almacena en el hígado, lo cual mantiene los niveles de glucosa en la sangre, para ello hay una hormona reguladora que es la insulina producida por el páncreas (islotos pancreáticos). Esta hormona hace que la glucosa de la sangre entre en las células y sea utilizada en glucosa y se almacena en forma de glucógeno y ácidos grasos. La glucosa es el segundo compuesto más abundante en la naturaleza y es la fuente principal de energía en las células, mediante la degradación catabólica, y es el componente principal de polímeros de importancia estructural como la celulosa y de polímeros de almacenamiento energético como el almidón. ^[1]

En su forma la D-glucosa sufre ciclación como se muestra en la *figura 3.1* hacia su forma hemiacetálica para lograr sus formas de furano y pirano (D-glucofuranosa y glucopiranosas) que a su vez presentan anómeros Alfa y Beta. Estos anómeros no presentan diferencias de composición estructural, pero sí difieren de características físicas y químicas. La D-(+)-glucosa es uno de los compuestos más importantes para los seres vivos, incluyendo a seres humanos. ^[1]

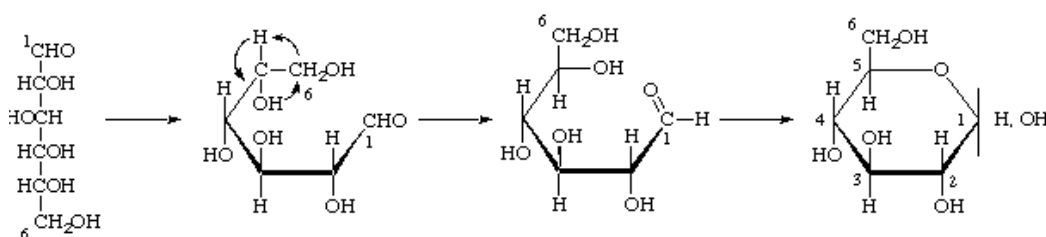


Fig 3.1 Ciclación de la glucosa

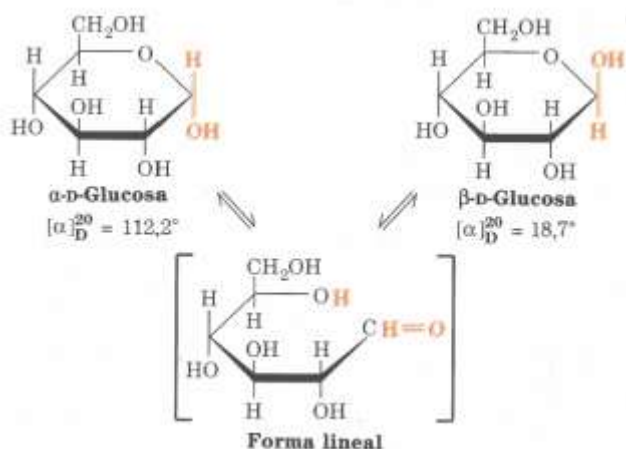


Figura 3.2. La mutarrotación de la glucosa proporciona un ejemplo instructivo de catálisis ácido-base. Adquiere cualquiera de las dos formas anoméricas cíclicas, a través de una forma lineal intermedia.

Estas son formas isoméricas, sin embargo no son las únicas, ya que como podemos ver *figura 3.2*, los monosacáridos presentan carbono asimétrico y por lo tanto también poseen isómeros ópticos

En su forma β - D- glucopiranosica, una molécula de glucosa se une a otra gracias a los –OH de sus carbonos [1-4] para formar Celobiosica a través de un enlace β , y al unirse varias de estas moléculas, forman celulosa.

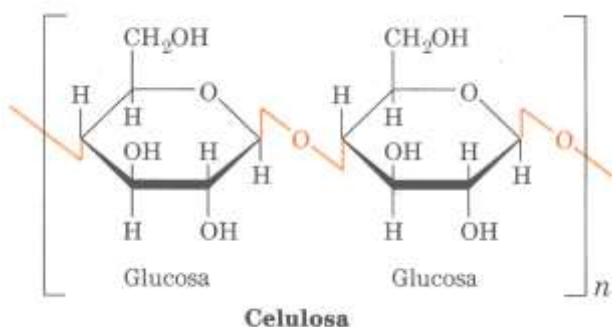
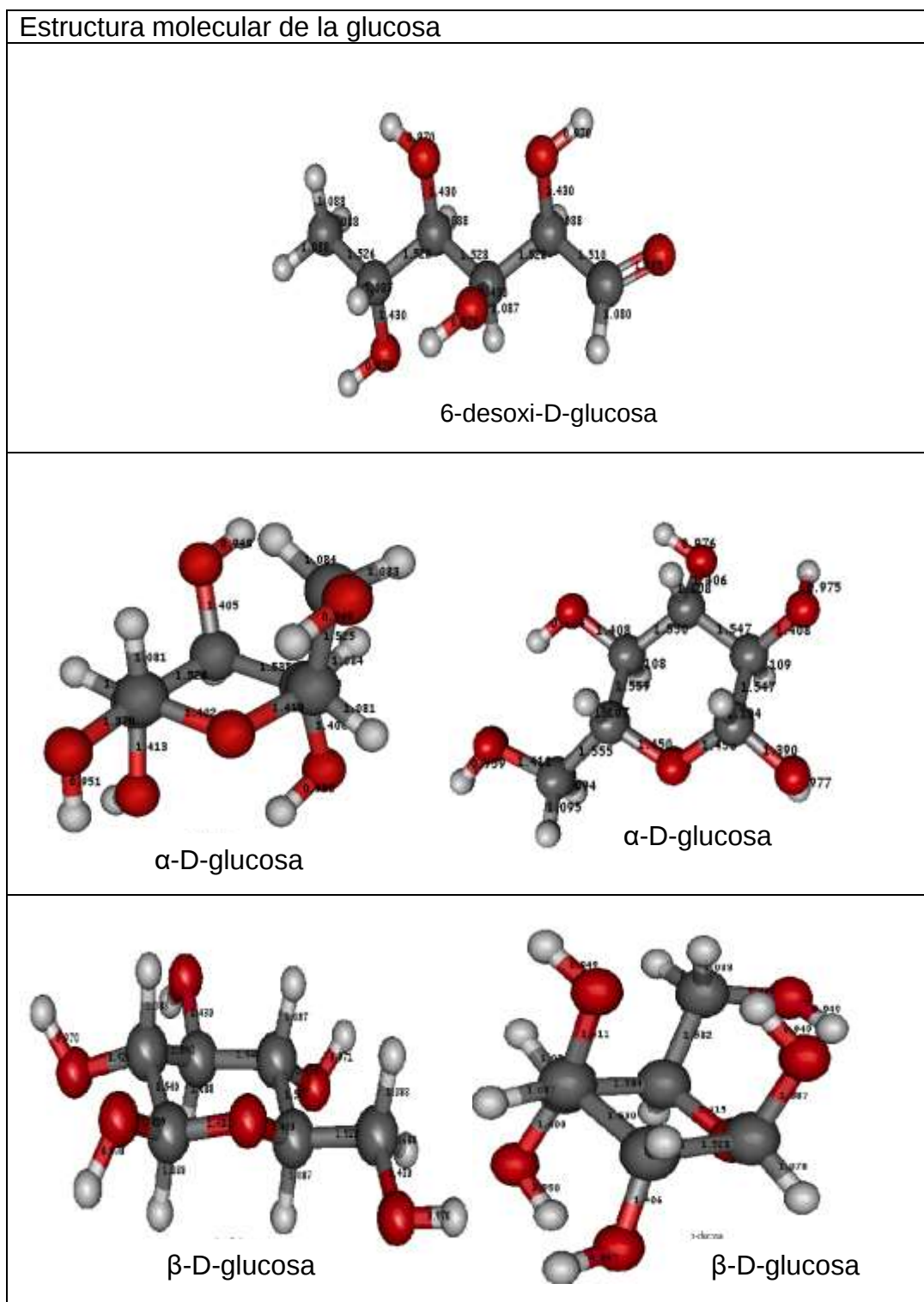


Figura 3.3 Estructura primaria de la celulosa. En ella n puede representar varios millares.

Figura 3.4 Modelos utilizados (glucosa)



Con la intención de seleccionar la glucosa a estudiar en el presente trabajo se realizó un proceso de calibración:

En la *figura 3.4* se muestran las diferentes estructuras moleculares de la glucosa que para fines de este trabajo se analizaron para determinar la molécula más reactiva en la interacción glucosa-insulina, utilizamos estaciones de trabajo Silicon Graphics, OCTANE para todos los cálculos. Se utilizó además la interfase gráfica *Spartan* ^[2] y los software *Gaussian 94* ^[3] y *Gaussian 98* ^[4].

3.2 Dureza

Teoría funcional de la densidad (DFT) es una forma de la mecánica cuántica que utiliza la función de densidad de electrones, ρ , en lugar de la función de onda más habitual, Ψ , para describir un sistema químico. Este sistema es una colección de núcleos y electrones. ^[5]

Durante la última década, el papel de la DFT en la química cuántica ha mostrado un rápido aumento. Con estos métodos, es posible incluir la correlación electrónica a un coste computacional mucho menor que utilizando métodos *ab initio* convencionales de correlación. Otro aspecto importante de la DFT, además de las ventajas de cálculo, es el hecho de que muchos conceptos químicos de uso común, como la electronegatividad (χ) ^[6], la dureza global (η) ^[7], y la blandura global (S) ^[8] reciban una definición matemática precisa.

En este sentido, DFT proporciona un puente que conecta a algunos de los conceptos empíricos de la mecánica cuántica ^[5-9]. Entre estos conceptos, la dureza es especialmente significativa ya que es la piedra angular de algunos principios

importantes de la reactividad química, tales como los ácidos duros y blandos y las bases (HSAB) ^[7-9] y el principio de máxima dureza (MHP).^[9-13]

Además de los parámetros globales (χ , η , y S), una serie de contrapartes locales (por ejemplo, la dureza local ^[10-11], la blandura locales, y la función de Fukui ^[12]) se han desarrollado.^[9-13]

La definición original de la dureza de Pearson introdujo los conceptos de los ácidos duros y blandos / bases y el principio de HSAB. Este principio afirma que los ácidos duros prefieren coordinar con las bases duras, mientras que los ácidos blandos prefieren reaccionar con bases blandos tanto desde los puntos de vista termodinámicos y cinéticos.

Dureza ^[14-16] es una medida de la resistencia de una especie química para cambiar su configuración electrónica. El recíproco de la dureza es la blandura ^[14-15], que mide la facilidad de transferencia de carga y que se asocia con alta polarizabilidad ^[17].

$$S = \frac{1}{\eta}. \quad (3.1)$$

Mediante la aplicación de la aproximación por diferencias finitas a la ecuación (3.1), obtenemos la definición operativa de η como ^[14-16],

$$\eta \cong I - A \quad (3.2)$$

donde I y A son la energía de ionización vertical y la afinidad electrónica, respectivamente. Estas cantidades pueden ser obtenidas directamente de *Hartree-Fock* (HF).^[18]

$$\eta \cong E_{\text{LUMO}} - E_{\text{HOMO}} \quad (3.3)$$

donde E_{HOMO} y E_{LUMO} son las energías de los orbitales moleculares ocupados más alto y más bajo el orbital molecular ocupado, respectivamente.

En la tabla 3.2 se reportan los valores de energía electrónica de formación (Kcal/mol), energía del HOMO y LUMO (eV), dureza (η). De acuerdo con datos termodinámicos la β -glucosa la más estable con respecto α -D-glucosa y la 6-desoxi-D-glucosa (forma lineal) podemos ver, el análisis de dureza indica que la β -glucosa donde $\eta >$ especie estable, $\eta <$ reactiva lo anterior sugiere utilizar que la 6-desoxi-D-glucosa (forma lineal) es la más reactiva.

Tabla 3.2 Propiedades electrónicas. Todos los cálculos se realizaron a un nivel de teoría HF/ 6-31G

Molécula	Energía electrónica de formación Kcal/mol	Energía de HOMO /eV	Energía de LUMO eV	Dureza absoluta $\eta = (E_{\text{HOMO}} - E_{\text{LUMO}})$
6-desoxi-D-glucosa (forma lineal)	-230.425	-11.353	3.991	-15.345
α -D-glucosa	-279.882	-10.971	5.103	-16.074
α -D-glucosa	-271.651	-11.447	5.138	-16.584
β -D-glucosa	-277.942	-11.429	5.206	-16.635
β -D-glucosa	-280.755	-11.082	4.964	-16.046

3.3 Distribución de sitios activos

La distribución de sitios activos electrofílicos en un sistema puede derivarse a partir de la teoría de los orbitales frontera dentro de la aproximación de core congelado. Con la intención de determinar esta distribución, analizamos los sitios donde el orbital HOMO tiene el valor absoluto más grande. En la Tabla 3.3 se ha mapeado el orbital HOMO. Puede observarse que el HOMO se encuentra localizado una densidad electrónica mayor en tres regiones de la molécula 6-desoxi-D-glucosa en donde un ataque electrofílico podría llevarse a cabo así como siendo la más reactiva.

En la tabla 3.4 puede observarse, el mapeo del LUMO sobre una isosuperficie de densidad. Puede observarse que la 6-desoxi-D-glucosa para un ataque nucleofílico podría llevarse a cabo en la mayor parte de la molécula donde puede notarse la presencia de sitios activos al observar que los sitios azules son las posiciones donadoras de carga, mientras que las posiciones más rojas son aceptoras de carga. Lo que podría relacionarse con sitios nucleofílicos y electrofílicos respectivamente.

Es bueno señalar los términos nucleofílo, electrofílo y sitio activo:

Nucleofílo:

Es un átomo que tiene un par de electrones sin compartir que puede utilizarlos para formar un enlace con un carbono.^[19]

Electrofílo:

Es una especie (ión o compuesto) que puede actuar como un aceptor de pares de electrones, es decir un buscador de electrones^[19]

Sitio activo:

Lo consideramos como el punto más probable para que exista una transferencia de electrones.

Tabla 3.3. Mapeo del HOMO (valor absoluto) a un nivel de teoría Single Point sobre una isosuperficie de densidad ($\rho=0.002$ e/a.u.³). Las zonas azules indican los sitios más reactivos para un ataque electrofílico.

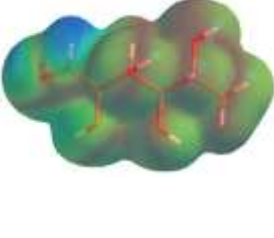
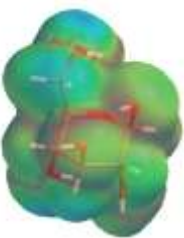
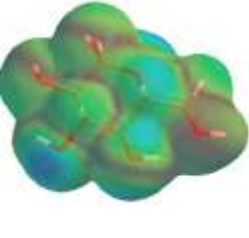
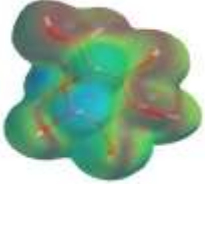
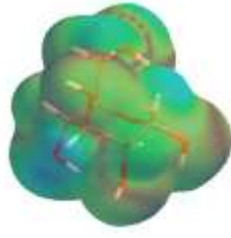
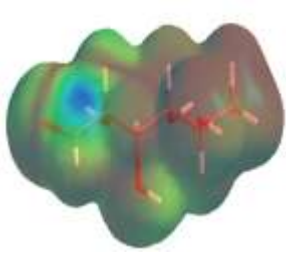
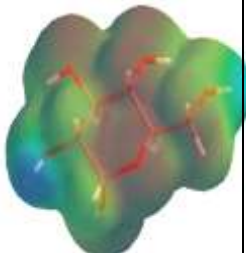
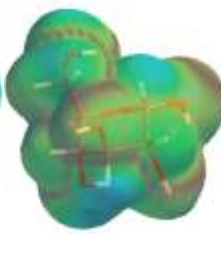
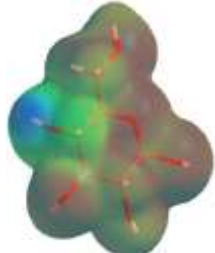
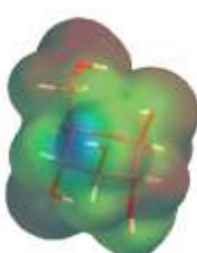
				
6-desoxi-D-glucosa	α -D-glucosa	α -D-glucosa	β -D-glucosa	β -D-glucosa

Tabla 3.4 Mapeo del LUMO (valor absoluto) a un nivel de teoría Single Point sobre una isosuperficie de densidad ($\rho=0.002$ e/a.u.³). El zonas más rojas corresponden al valor más alto del LUMO.

				
6-desoxi-D-glucosa	α -D-glucosa	α -D-glucosa	β -D-glucosa	β -D-glucosa

3.4 Análisis Vibracional

En la siguiente tabla 3.5 nos indica todas las frecuencias son positivas eso quiere decir que están en un mínimo local para su asignación como intermedios o estado de transición a lo largo del camino de reacción.

Tabla 3.5 Análisis vibracional de la glucosa

Vibraciones A ° beta-d-glucosa	Vibraciones A ° 6-deoxy - D-glucosa	Vibraciones A ° β-glucosa	Vibraciones A ° α-glucosa	Vibraciones A ° glucosa
75.03	60.99	75.93	95.56	78.99
114.12	82.30	117.79	103.15	104.49
132.01	84.82	162.52	157.25	131.31
158.60	126.84	207.25	199.23	151.67
196.80	171.27	230.96	206.93	184.98
268.05	240.64	246.50	237.44	250.28
276.93	250.64	254.54	259.89	258.12
285.47	273.92	294.22	287.72	278.81
307.90	281.65	341.90	310.07	290.20
312.77	303.39	345.75	328.70	305.78
350.97	305.69	358.22	363.46	369.65
365.58	350.41	383.23	371.96	405.19
395.99	372.79	420.14	416.89	421.49
427.42	411.34	433.80	433.58	430.47
440.35	467.56	491.01	466.18	443.66
447.26	479.72	534.76	503.04	467.32
460.16	494.22	553.18	516.58	512.72
484.45	540.95	569.36	540.51	549.46
543.16	620.08	602.78	581.82	567.75
598.59	736.48	615.34	619.41	607.46
654.44	791.01	672.02	649.96	639.79
690.25	847.17	731.34	772.50	771.67
701.10	912.81	856.36	892.35	824.69
977.57	940.00	943.56	944.15	956.10
1095.37	972.41	983.48	978.53	989.32
1107.00	1036.08	991.47	996.97	1007.52
1124.31	1078.89	1104.92	1078.76	1108.22
1165.54	1140.08	1164.84	1150.46	1130.71
1196.91	1159.71	1171.35	1167.45	1172.81
1207.25	1196.95	1189.38	1179.80	1187.46
1212.43	1215.99	1194.34	1207.62	1220.08
1230.33	1239.15	1212.36	1214.95	1233.56
1244.91	1259.13	1213.93	1218.33	1240.58
1260.96	1312.68	1221.43	1235.83	1265.12
1275.34	1323.02	1245.03	1269.98	1270.70
1296.55	1360.20	1282.26	1274.21	1287.29
1305.06	1399.43	1330.31	1330.30	1314.45
1342.01	1434.66	1349.33	1340.64	1346.09
1354.30	1452.61	1366.76	1356.59	1367.41
1367.27	1485.19	1380.49	1375.65	1379.63
1380.05	1503.97	1404.02	1424.29	1406.36
1428.60	1524.96	1440.32	1440.58	1440.06
1474.94	1531.03	1459.69	1457.10	1460.56
1483.04	1545.05	1494.17	1492.39	1494.39
1509.22	1567.21	1516.06	1500.75	1504.90
1513.11	1573.75	1528.33	1516.42	1519.41
1532.83	1583.48	1535.62	1538.51	1528.10

1540.64	1619.88	1555.56	1556.32	1565.52
1564.49	1637.77	1575.44	1564.26	1574.23
1588.78	1653.87	1588.32	1590.67	1580.16
1597.56	2017.49	1594.87	1598.47	1587.05
1604.40	3171.37	1599.71	1599.91	1603.62
1655.80	3199.36	1627.00	1623.71	1608.66
1661.28	3206.68	1681.08	1681.10	1656.15
3183.16	3219.86	3215.79	3207.50	3158.57
3183.67	3222.71	3226.25	3230.52	3182.40
3191.01	3260.81	3240.04	3240.16	3214.13
3200.69	3292.18	3246.24	3254.42	3241.22
3215.38	3312.15	3268.34	3269.55	3267.13
3229.59	4029.44	3292.82	3289.16	3321.98
3284.60	4067.16	3333.83	3298.48	3327.64
4094.09	4074.52	4067.69	4072.13	4070.57
4098.09	4116.50	4077.37	4075.80	4084.93
4098.92	-	4091.70	4087.46	4093.56
4106.38	-	4103.41	4106.99	4100.54
4124.10	-	4104.63	4112.59	4115.34

3.5 Conclusiones

Los resultados obtenidos indican que la β -D-glucosa es la más estable. Por otro lado, el análisis de dureza indica que la 6-desoxi-D-glucosa es la más reactiva que la α -D-glucosa o β -D-glucosa, al mostrar un sitio activo bastante extendido para el caso de un ataque nucleofílico o electrofílico.

3.6 Referencias

- [1] www.ferato.com/wiki/index.php/Glucosa. Consultado el 11 Diciembre 2009
- [2] Spartan versión 5.1.1, 1999, Wavefunction, Inc. 18401 Von Karman Ave. Suite 370. Irvine, CA.
- [3]. Gaussian 94. Revision E.2, M.J. Frish, G. W. Trucks, H.B Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J. R. Cheeseman, T. Keith, G.A. Peterson, J.A. Montgomery, K. Raghavachari, M. A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J Cioslowski, B.B. Stefanov, A. Nanayakkara, M Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M. W. Chen, M.W. Wong, J. L. Andres, E.S. Replogle, R. Gomperts, R.L Martín, D.J. Fox, J. S Binkley, D.J. Defrees, J. Baker, J.P. Stewart, M. Head- Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburg PA.1995.
- [4] Gaussian 98.(Revision A.1.) , M. J. Frish, G.W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb J. R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millan, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P. Y. Ayala,Q. Cui, K. Morokuma, D. K. Malick, A.D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. L. Fox, T. Keith, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head- Gordon, E.S. Replogle and J. A. Szabo, N.S. J. A. Pople, Gaussian, Inc., Pittsburg PA.1998.
- [5] R.G. Parr and W. Yang, (1989).*Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York
- [6] R.G. Pearson, (1989)*J. Org. Chem.* 54
- [7]. R.G. Pearson, (1997).*Chemical Hardness: Applications from Molecules to Solids*, Wiley-VCH, Weinheim
- [8] R.G. Parr and Z. Zhou, (1993)*Acc. Chemistry. Res.* 26
- [9] (a)H. Chermette, 20 (1999) *J. Comput. Chemistry*.
(b)P. Geerlings, F. De Proft and W. Langenaeker, (2003) *Chemistry. Rev.* 103

- [10] R.G. Parr and R.G.J. Pearson, (1983) *J. Am. Chem. Soc.* 105
- [11] (a) P.K. Chattaraj, H. Lee and R.G. Parr, (1991) *J. Am. Chemistry. Soc.* 113
(b) Y. Li and J.N.S. Evans, (1995) *J. Am. Chemistry. Soc.* 117
- [12] (a) R.G. Pearson, (1987) *J. Chemistry. Educ.* 64
(b) R.G. Pearson, (1999) *J. Chemistry. Educ.* 76
- [13] M. Torrent-Sucarrat, J.M. Luis, M. Duran and M. Solà, (2001) *J. Am. Chemistry. Soc.* 123
- [14] R.G. Pearson, (1985) *J. Am. Chemistry. Soc.* 107
- [15] R.G. Parr and Z. Zhou. (1993) *Acc. Chemistry. Res.* 26
- [16] R.G. Parr and P.K. Chattaraj, (1991) *J. Am. Chemistry. Soc.* 131
- [17] P.J. Politzer, (1987) *J. Chemistry. Phys.* 86
- [18] T. Koopmans, (1934) *Physica* 1
- [19] Francis A. Carey. (1999) *Química orgánica*. Tercera edición. Ed Mc Graw Hill. España pág. 12

Capítulo 4

Propiedades electrónicas de la insulina

4.1 Estudio a nivel semiempírico *PM6*

El desarrollo de nuevas drogas para la diabetes ha sido un proceso continuo debido al alarmante aumento de esta enfermedad y complicaciones asociadas. Desafortunadamente la velocidad con la que avanza nuestro conocimiento sobre la enfermedad y sus efectos no están a la par del desarrollo de nuevas drogas. La identificación del receptor en la insulina (RI) y su actividad quinasa intrínseca han sido señaladas como el punto clave en el desarrollo de nuevos medicamentos. Por otro lado, en estudios experimentales se ha elucidado que la activación molecular del RI se lleva a cabo en el sitio de enlace del receptor con la membrana celular. Así, un primer paso para entender la interacción de la insulina a través de su receptor implica un buen conocimiento de las zonas reactivas de esta molécula. ^[1]

Desde su descubrimiento, la insulina ha sido objeto de una intensa investigación bioquímica y biofísica. Se ha reportado, que existe una alta probabilidad de que la insulina forme agregados tales como dímeros, tetrámeros, hexámeros, etc, así como estructuras fibrosas en función del pH, de la temperatura, concentración, entre otras. Sin embargo, varios estudios experimentales indican que la actividad biológica de la insulina monomérica, (*figura 4.1*), consiste de dos cadenas polipeptídicas, la cadena α (21 aminoácidos) y la cadena β (30 aminoácidos). Varias regiones de la molécula se han asociado a actividad biológica, tal como la posición de las cisteínas que forman los puentes de bisulfuro, las regiones terminales de la cadena A en el nitrógeno y carbono y los residuos hidrofóbicos en el carbono terminal de la cadena β . ^[1-4]

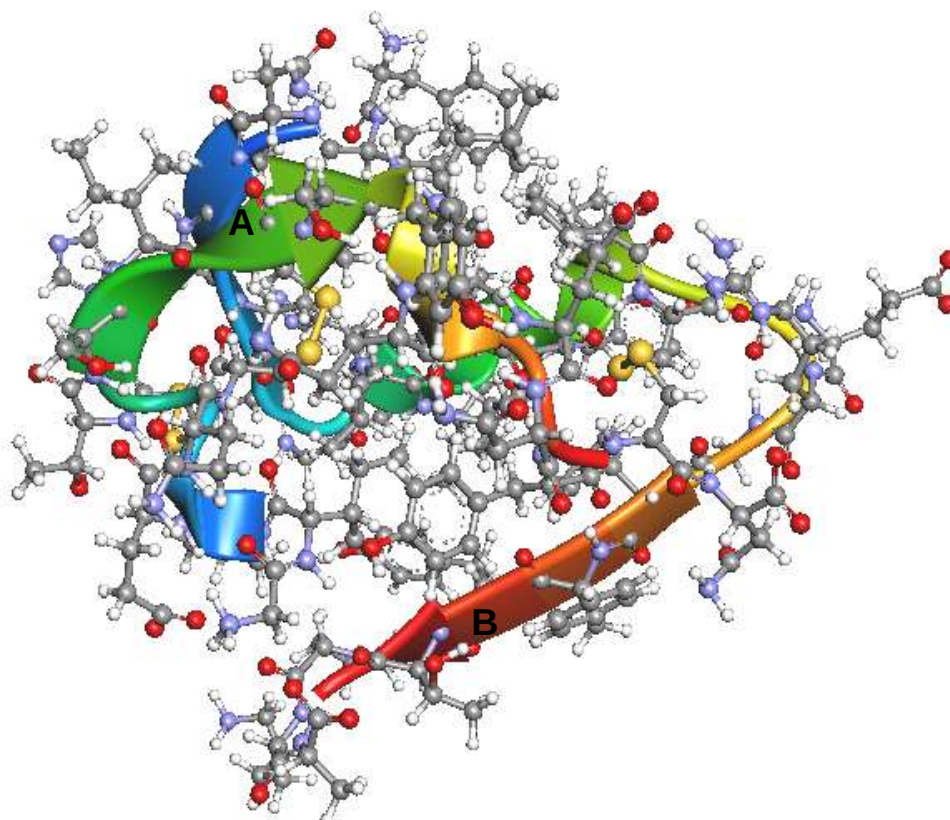


Figura 4.1

Insulina 9INS. En la base de datos
protein data bank

Debido a la complejidad y tamaño de la insulina, hasta donde sabemos, no ha sido posible estudiar su reactividad desde el punto de vista de la mecánica cuántica. Actualmente, el estudio de las propiedades electrónicas de algunas proteínas es difícil; es por esto que se requieren de nuevos métodos y habilidades para estudiar sistemas moleculares compuestos de una gran cantidad de átomos. Los métodos semiempíricos PM6 desarrollados por el grupo de investigación de Stewart^[5], ofrecen nuevas posibilidades de modelar proteínas en estaciones de trabajo convencionales. Adicionalmente, la predicción de las propiedades estructurales y electrónicas de las bases proteicas que constituyen a la insulina han sido predichas por el método PM6 y son comparables con resultados obtenidos por medio de la Teoría de los Funcionales de la Densidad (TFD)^[6]. Dado lo anterior, en el presente trabajo analizamos las capacidades del método PM6 para estudiar la reactividad global de la insulina 9INS.^[6]

4.1.1. Optimización de geometría

La insulina monomérica 9ins fue optimizada en el nivel PM6, durante la optimización de geometría no se consideró restricción de ningún tipo. Para todos los cálculos se utilizó una estación de trabajo con 2 procesadores Intel Xeón a 2.4 GHZ, con 1 GB de memoria RAM. Los cálculos se realizaron con el software *MOPAC 2009* versión Linux^[10] y fueron visualizados con (*Gaussview*, 2007)^[7]

4.2 Estudio a nivel *ab initio* utilizando el método *ONIOM*

4.2.1 Distribución de sitios activos

La distribución de sitios activos electrofílicos en un sistema puede derivarse a partir de la teoría de los orbitales frontera dentro de la aproximación de core congelado, (ecuación 4.1). Con la intención de determinar esta distribución, analizamos los sitios donde el orbital HOMO tiene el valor absoluto más grande. En la *figura 4.2* se ha mapeado el orbital HOMO, en el estado basal. Es interesante observar que el HOMO se encuentra localizado en una sola región de la molécula en donde un ataque electrofílico podría llevarse a cabo. El HOMO se encontró exclusivamente en la cadena α de la insulina, una lista de los átomos más reactivos se reportan en la Tabla 4.1.

$$f(r) = (\partial p(r) / \partial N)_v \quad (4.1)$$

4.2.1.1 Electofílicos (Mapa del orbital HOMO).

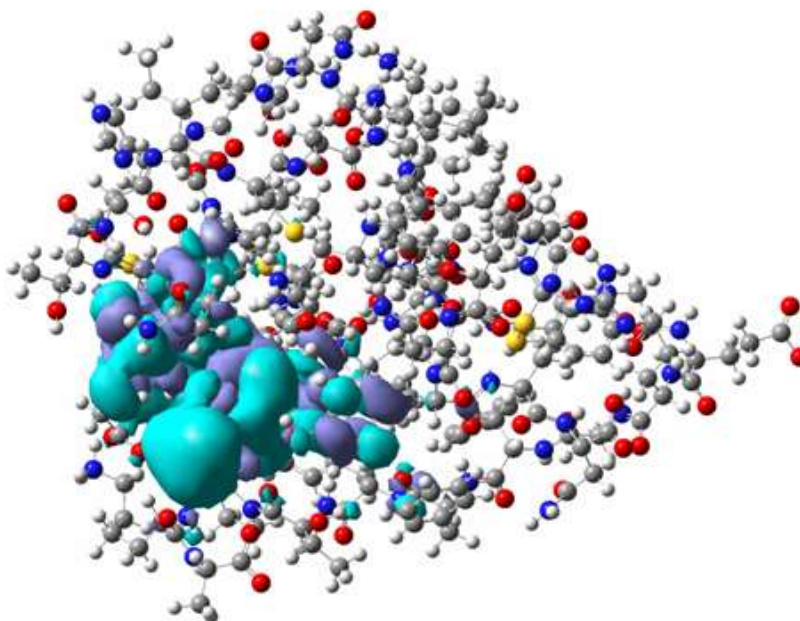


Figura 4.2 Distribución del orbital HOMO de la insulina monomérica 9INS optimizada en el nivel PM6.

A partir de la tabla 4.1 es posible observar que los residuos 1 Gly, 2 Ile, 3 Val, 4 Glu, 5 Gln y 19 Tyr en la insulina son los que exhiben la mayor reactividad ante un ataque de tipo electrofílico.

Tabla 4.1. Átomos electrofílicos más reactivos obtenidos a partir de la estructura de la insulina 9INS optimizada en el nivel PM6.

Tipo de átomo	Residuo	Cadena
N	1 Gly	α
CA	1 Gly	α
C	1 Gly	α
O	1 Gly	α
N	2 Ile	α
CA	2 Ile	α
C	2 Ile	α
O	2 Ile	α

CB	2 Ile	α
CG1	2 Ile	α
N	3 Val	α
CA	3 Val	α
C	3 Val	α
O	3 Val	α
CB	3 Val	α
CG1	3 Val	α
CG2	3 Val	α
N	4 Glu	α
CA	4 Glu	α
C	4 Glu	α
O	4 Glu	α
CB	4 Glu	α
N	5 Gln	α
C	5 Gln	α
O	5 Gln	α
CE1	19 Tyr	α
CE2	19 Tyr	α
CZ	19 Tyr	α
OH	19 Tyr	α
H	19 Tyr	α

Nomenclatura basada en ref. [11] donde: C es átomo de carbono carbonilo; CA es átomo de carbono α ; CB es átomo de carbono β ; CG átomo de carbono γ ; CD átomo de carbono δ .

4.2.1.2 Nucleofílicos (Mapa del orbital LUMO).

En la figura 4.3 se muestra el mapeo del orbital LUMO, obsérvese que al igual que el orbital HOMO, el orbital LUMO se encuentra bien localizado; sin embargo, en este caso se encontró exclusivamente en la cadena β , involucrando a los residuos 33 Val, 34 Glu, 35 Ala, 36 Leu, 37 Tyr y 45 Phe B.

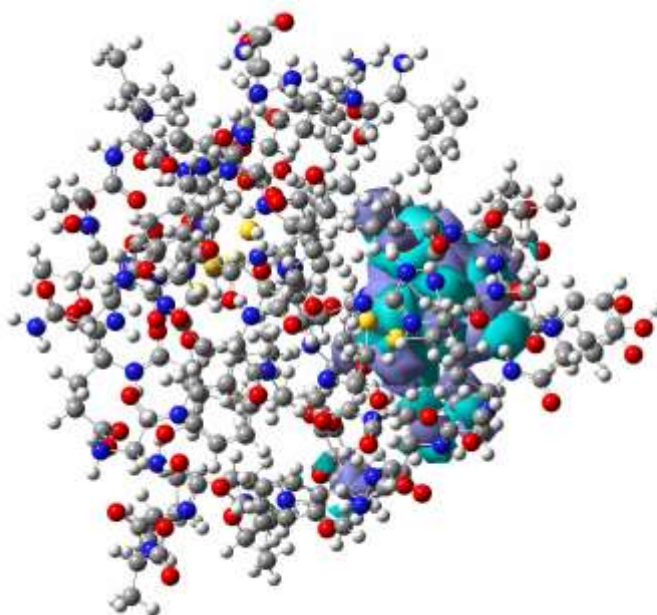


Figura 4.3

Distribución del orbital LUMO de la insulina monomérica 9INS optimizada en el nivel PM6.

Tabla 4.2 Átomos nucleofílicos más reactivos obtenidos a partir de la estructura de la insulina 9INS optimizada en el nivel PM6.

Tipo de átomo	Residuo	Cadena
C	33 Val	β
O	33 Val	β
N	34 Glu	β
CA	34 Glu	β
C	34 Glu	β
O	34 Glu	β
CB	34 Glu	β
H	34 Glu	β
H	34 Glu	β
N	35 Ala	β
CA	35 Ala	β
C	35 Ala	β
O	35 Ala	β
CB	35 Ala	β
H	35 Ala	β

H	35 Ala	β
H	35 Ala	β
H	35 Ala	β
H	35 Ala	β
N	36 Leu	β
CA	36 Leu	β
C	36 Leu	β
O	36 Leu	β
CB	36 Leu	β
CG	36 Leu	β
H	36 Leu	β
H	36 Leu	β
H	36 Leu	β
H	36 Leu	β
N	37 Tyr	β
CA	37 Tyr	β
H	37 Tyr	β
H	37 Tyr	β
CG	45 Phe B	β
CD1	45 Phe B	β
CZ	45 Phe B	β
CD2	45 Phe B	β

Nomenclatura basada en ref. [11] donde: C es átomo de carbono carbonilo; CA es átomo de carbono α ; CB es átomo de carbono β ; CG átomo de carbono γ ; CD átomo de carbono δ .

4.3 Conclusiones

Es bien conocido que la insulina es una proteína que regula los niveles de azúcar en la sangre. Sin embargo, en la insulina como en cualquier otra proteína, no todas las regiones tienen la misma importancia y su actividad concreta se localiza en determinadas zonas llamadas “sitios activos”. Por eso, si alguna molécula extraña se acopla a este sitio activo puede inhibirlo y la proteína perderá su actividad. Un buen conocimiento del tipo y distribución de los sitios activos en la molécula permitirán establecer mecanismos que impidan este proceso de inhibición. Los resultados experimentales reportados en la literatura sugieren que el RI de la insulina está localizado en las regiones terminales de las cadenas α y β . Sin embargo, hasta donde sabemos no existe un estudio a nivel molecular que permita establecer el tipo y

distribución de sitios activos en la molécula de la insulina. A este respecto, las teorías aceptadas de la mecánica cuántica, en combinación con los criterios de reactividad derivados a partir de la teoría de los funcionales de la densidad permiten determinar el tipo y distribución de sitios activos en la molécula. Sin embargo actualmente la aplicación de métodos *ab initio* para estudiar las propiedades electrónicas de pequeñas proteínas es prohibitiva en sistemas de súper cómputo comunes. Recientemente, fue desarrollado el método semiempírico PM6 que ha permitido mejorar las predicciones del proceso de *Docking* ampliamente utilizado en el diseño de nuevos fármacos. ^[9]

4.4 Referencias

- [1] M.Balasubramanyam, V. Mohan. (2001). Review. *Orally active insulin mimics: where do we stand now?* J. Biosci. 26(3).
- [2] G.Dodson, D. Steiner (1998). *The role of assembly in insulin's biosynthesis*. Curr Opin Struct Biol.8(2).
- [3] J.Dong, Z. Wan,. Popov M., Carey P., Weiss M. (2003) *Insulin assembly damps conformational fluctuations: Raman analysis of amide I linewidths in native states and fibrils*. J. Mol. Biol. 330
- [4] E. Nettleton, P. Tito, Sunde M., Bouchard M. Dobson C, Robinson c. (2000) *Characterization of the oligometric states of insulin in self- assembly and amyloid fibril formation by mass spectrometry*. Biophys J.79(2)
- [5] J. Stewart (2007) *Optimization of Parameters for semiempirical Methods V: Modification of NDDO Approximations and application to 70 Elements* J. mol. Mod 13(12).
- [6] J. Stewart (2009) *Application of the PM6 method to modeling proteins*. J Mol Model 15(7).
- [7]. Gaussview Rev. 3.09, windows version. Gaussian. Inc Pittsburgh, PA.
- [8] R. Parr and W. Yang (1989). *Density functional Theory of Atoms and Molecules*. Primera edición Oxford Univ Press. Oxford.
- [9] Z. Bikadi, E. Hazai (2009) *Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock* J. of Cheminformatics, 1:15. doi:10.1186/1758-2946-1-15.
- [10] Mopac2009, J.P. James, Stewart Computational Chemistry. Version 8.032L.
- [11] R.J. Feldam, (1976) *Atlas of Macromolecular Structure on Microfiche (AMSOM)* 1st Ed, Tracor Jitco, Inc., Rockville, Md.

Capítulo 5

Interacción glucosa-insulina utilizando *Docking*

5.1 ¿En qué consiste el *Docking*?

Una quimioteca de compuestos orgánicos se posiciona en el sitio de unión y se evalúa la actividad potencial de estos compuestos a partir de la energía de interacción proteína-ligando. Aquellos ligandos con mayor actividad calculada son candidatos a síntesis o puede comprarse en la base de datos. Este tipo de aproximación es bastante frecuente cuando se dispone de la estructura resuelta de la proteína o bien de un modelo de la misma construido por homología. Así, se han identificado ligandos para más de 50 receptores, tanto de estructura conocida como a partir de modelos teóricos. Otro de los usos establecidos del *Docking* es la identificación del modo de unión, es decir, la orientación y conformación que el ligando adopta en la cavidad de la proteína y, menos frecuentemente, se utiliza para identificar el sitio de unión.^[1]

El *Docking*, es un método para muestrear el espacio conformacional del complejo ligando-receptor utilizada para evaluar la afinidad de la interacción ligando-macromolécula. Existen diferentes implementaciones de algoritmos para encontrar configuraciones proteína-ligando (denominadas *poses*) próximas a la conformación nativa del complejo cristalizado (una RMSD inferior a 2 Å es el criterio de aceptación de una *pose* particular).^[1]

Actualmente, todos los algoritmos modernos de *Docking* modelan el ligando como flexible, dejando de lado las aproximaciones más primitivas en las que el ligando se consideraba rígido (*Docking rígido*). Los métodos más comunes son: *fast shape matching* (DOCK, EUDOCK, LIGANDFIT), construcción incremental del ligando en la cavidad de la proteína (FLEXX, HAMMERHEAD), búsquedas tabú (PRO_LEADS,

SFDOCK), algoritmos genéticos (GOLD, AUTODOCK3.0, GAMBLER), algoritmos genéticos acoplados a búsqueda local o Lamarckianos (AUTODOCK3.0), programación evolutiva⁸³, *simulated annealing* (AUTODOCK2.4, GLIDE), métodos de Monte Carlo (MCDOCK, QXP, ICM-DOCK) y geometría de distancias (DOCKIT). También existen combinaciones de estos métodos. ^[1]

Los métodos *fast shape matching*, como el implementado en DOCK, caracterizan el sitio activo del receptor mediante esferas, cuyos centros se ajustan a los centros del ligando (átomos pesados o esferas) sobre la base de una comparación de las distancias internas ligando-ligando y receptor-receptor (*figura 5.1*). Los métodos de construcción incremental del ligando utilizan en muchos casos, como HAMMERHEAD, una caracterización del sitio activo similar a la de los métodos *fast shape matching*. En este caso, acoplan progresivamente fragmentos del ligando que contengan como mínimo dos enlaces rotables, explorando para cada uno de ellos las conformaciones posibles. ^[1]

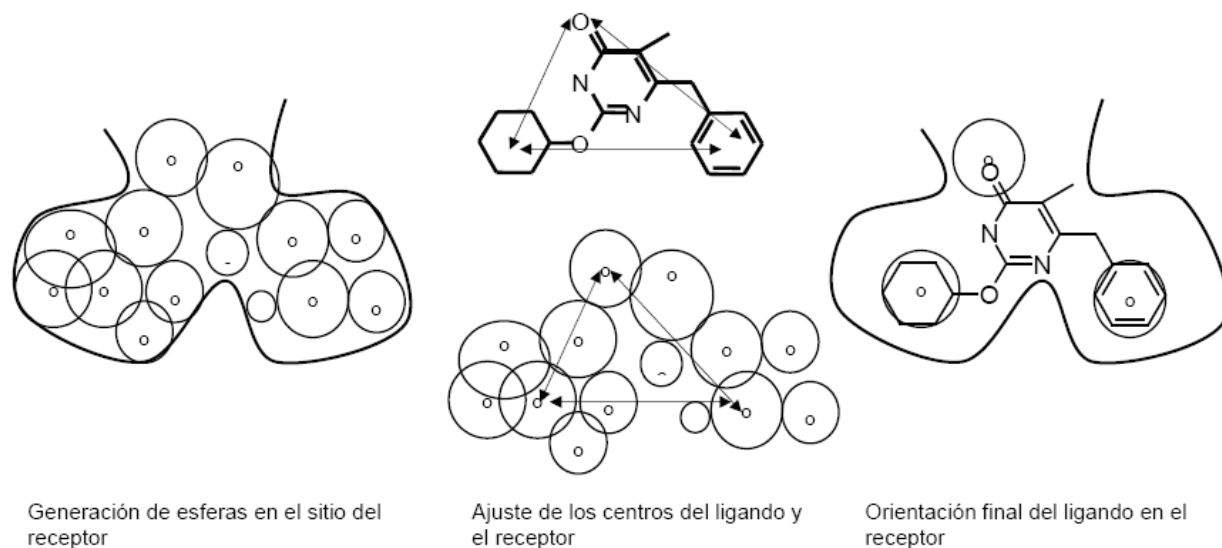


Figura 5.1. Esquema del método de muestreo en *docking* de los métodos *fast shape matching* (DOCK). ^[1]

El *Docking* es la parte que requiere más tiempo computacional, por lo que los algoritmos que tardan más de tres minutos por ligando por procesador, se consideran

demasiado lentos para ser utilizados en VS. La parte más conflictiva es la función de *scoring* para predecir la afinidad de la unión proteína (o cualquier otra macromolécula)–ligando. Las funciones tradicionalmente aplicadas se clasifican en:

- ❖ *Basadas en campos de fuerza (Force field-Based)*: a partir de mecánica molecular, aproximan la energía libre de unión a partir de la suma de interacciones electrostáticas y de van der Waals. Frecuentemente, incluyen también términos empíricos que incluyan la entropía y solvatación. Destacan las funciones DOCK y CHARMM.^[1]
- ❖ *Empíricas (Empirical)*: estiman la energía libre de unión sumando términos de interacción derivados de la contribución ponderada de parámetros estructurales (número de puentes de hidrógeno, interacciones iónicas, contactos apolares, entropía). Los pesos de cada parámetro se obtienen por ajuste a constantes de unión experimentales de un conjunto de complejos proteína-ligando. Las más conocidas son LUDI, CHEMSCORE, SCORE, FRESNO, FLEXX, PLP, AUTODOCK Y GOLDScore.^[1]
- ❖ *Knowledge-Based*: representan la afinidad como suma de interacciones de pares de átomos proteína-ligando. Estos potenciales se derivan a partir de complejos de estructuras conocidas del *Protein Data Bank*, donde las distribuciones de probabilidad de distancias interatómicas entre diferentes pares de tipos de átomo proteína-ligando se convierten, asumiendo distribuciones energéticas tipo Boltzmann, en funciones de potencial. La energía libre de interacción se calcula sumando las contribuciones de los pares de átomos dentro de una cierta distancia. Destacan PMF, DrugScore SMOG, BLEEP y SMOG2001. Las funciones empíricas son las más usadas en los programas de diseño de fármacos, aunque no hay ninguna función superior al resto, ya que diferentes funciones se comportan mejor para determinados complejos proteína-ligando. De hecho, debido a la falta de fiabilidad general, normalmente se utiliza una combinación de funciones (*consensus scoring*). Con ello,

se combinan varias funciones y solo aquellas conformaciones (*poses*) que reciben altos *scores* por dos o más funciones de *scoring* son consideradas favorables.^[1]

A pesar de que, por fundamento teórico, el *Docking* es uno de los filtros más precisos de VS, existen tres grandes problemas asociados a él:

- ❖ El gran número de posibles ligandos y sus posibles orientaciones y conformaciones exceden la capacidad computacional.^[1]
- ❖ La flexibilidad se introduce únicamente en el tratamiento del ligando, mientras que el receptor se considera rígido en la mayor parte de casos. El tratamiento flexible del receptor tiene un coste computacional todavía demasiado alto, por lo que se usan aproximaciones como el introducir movilidad en las cadenas laterales de algunos aminoácidos a partir de bibliotecas de rotámeros (GOLD), uso en simulaciones paralelas de distintos conformeros de la proteína o la construcción de una geometría difusa que engloba distintas conformaciones (módulo FlexE de FlexX, AUTODOCK). Sin embargo, numerosas proteínas muestran fenómenos de inducción (*induced-fit effects*) de las cadenas laterales y cambian de forma y estructura del solvente tras la unión del ligando.^[1]
- ❖ El cálculo de la afinidad proteína-ligando no es exacto, principalmente en lo referente al cálculo de energías de solvatación y a la consideración de cambios en la entropía. Pese a la reducción del espacio conformacional proteína-ligando muestreado al considerar el receptor rígido, se asume que los algoritmos de *Docking* funcionan adecuadamente en esta parte del *Docking*¹⁰⁵, siendo la función de *scoring* la parte más débil.^[1]

5.2 Optimización de geometría después de utilizar el proceso de *Docking*

El cálculo de energías libres de unión aparece a comienzos del siglo pasado, basado en simulaciones de mecánica molecular con dinámica molecular y métodos de Monte Carlo. Los dos grandes métodos: el de perturbación de energía libre (*Free Energy Perturbation*, FEP) e Integración Termodinámica (*Thermodynamic Integration*, TI), se presentaron como posibilidades fáciles y fiables. Se basan en que los cambios de energía libre relacionados con pequeñas perturbaciones de un sistema molecular se pueden determinar a partir de una simulación. Así, estos métodos realizan un tratamiento riguroso de todos los grados de libertad de complejos ligando-proteína, incluyendo modelos de solvatación adecuados. ^[1]

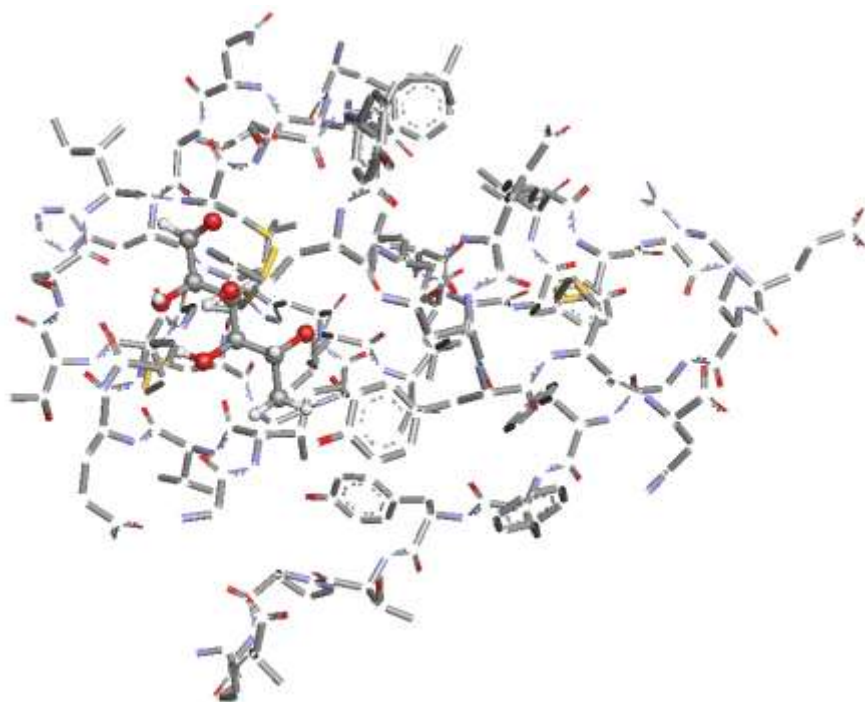


Figura 5.2 *Docking* (glucosa-insulina)

Los métodos de *Docking* como se muestra en la figura 5.2 están optimizados para encontrar el modo de unión de la glucosa-insulina, pero no están dirigidos, en principio, a determinar el sitio de unión. Por lo anterior nos servirá de ayuda para compararlos con los sitios activos de la insulina como se reporta en el capítulo anterior.

5.3 Resultados y discusión

En la tabla 5.1 se muestran los resultados obtenidos al estudiar la interacción entre la glucosa 6 desoxi-D-glucosa. Se mapeo el HOMO sobre una isosuperficie de densidad para determinar los sitios más susceptibles para un ataque por electrófilos (ácidos de Lewis) y el LUMO para un ataque nucleofílico como se muestran en el anterior capítulo. Puede observarse que no existe una gran diferencia en la interacción de ambas estructuras. Lo que indica que la determinación con el *Docking* (glucosa-insulina) nos ayudó para visualizar algunos puntos de unión. Por otro lado, resultados experimentales a partir de los cálculos de la función Fukui han demostrado el tipo de átomo, residuo y cadena con mayor exactitud.

Tabla 5.1 Átomos nucleofílicos más reactivos obtenidos a partir del *Docking* en la molécula 6-desoxi-D-glucosa

Tipo de átomo	Residuo	Cadena
N	34 Glu	β
CA	34 Glu	β
O	34 Glu	β
N	35 Ala	β
O	35 Ala	β
N	36 Leu	β
O	36 Leu	β
CB	36 Leu	β
CG	45 Phe B	β
H	34 Glu	β
O	37 Tyr	β
CA	37 Tyr	β

Nomenclatura basada en ref. [2] donde: C es átomo de carbono carbonilo; CA es átomo de carbono α ; CB es átomo de carbono β ; CG átomo de carbono γ ; CD átomo de carbono δ .

5.4 Conclusiones

Normalmente, se puede aceptar que el modo de unión se conserva entre distintos ligandos, aunque no siempre esta afirmación se cumple, como se verá en ciertos casos en este trabajo. Esto permite reducir en gran parte la búsqueda conformacional y eliminar aquellos ligandos (glucosas) para los que la unión sería físicamente imposible (por ejemplo, si se sabe que la interacción se establece a través de un aceptor de puente de hidrógeno, se puede prescindir de intentar realizar el *Docking* de una molécula que carezca de grupos aceptores). La introducción de restricciones de interacción se ha mostrado útil en el VS de los receptores acoplados a la insulina (9INS).

5.5 Referencias

[1] Obdulia Rabal Gracia. Tesis Doctoral Escola Tècnica Superior IQS. Departamento Química Orgánica y Bioquímica. Barcelona *Herramientas de cribado virtual aplicadas a inhibidores de tirosina quinasas. Contribución al desarrollo del programa PRALINS para el diseño de quimiotecas combinatorias*. C.I.F. G: 59069740 Universitat Ramon Lull Fundació Privada. Rgtr. Fund. Generalitat de Catalunya núm. 472 (28-02-90).

[2] R.J. Feldam, (1976) *Atlas of Macromolecular Structure on Microfiche (AMSOM) 1st* Ed, Tracor Jitco, Inc., Rockville, Md.

Capítulo 6

Conclusiones Generales

1. Se estudió la molécula de la glucosa desde un punto de vista teórico. El análisis de blandura y dureza absoluta, indica que la 6-desoxi-D-glucosa es más reactiva que β -D-glucosa, α -D-glucosa, al mostrar una zona activa bastante extendida para el caso de un ataque nucleofílico o electrofílico.
2. Dada la importancia de localizar los sitios activos en la insulina 9INS, proponemos varios métodos para determinar sitios activos, utilizando técnicas de la química cuántica computacional. Realizamos el estudio de la reactividad proveniente de la Teoría de los Funcionales de la Densidad. Nuestra metodología arrojó resultados que permiten identificar y señalar en donde se encuentran los sitios activos en la molécula de la insulina 9INS estudiada. Se encontró que la distribución y posición de los sitios activos se encuentra en la cadena β .
3. Se estudio la reactividad de la glucosa con la insulina en función con el número de sitios activos. El método *semiempírico PM6*, que introduce correcciones mayores a los métodos *PM3* y *AM1* el cual ha mostrado ser útil en el cálculo de propiedades electrónicas de proteínas debido a la inclusión de métodos que involucran escalamiento lineal.
4. Analizamos y comparamos el método *ONIOM* ha permitido mejorar las predicciones del proceso de *Docking*. Los resultados obtenidos en el nivel PM6

permiten sugerir que en la molécula 9INS existen dos zonas activas una electrofílica y una nucleofílica localizadas en la cadena α y β respectivamente.

5. En nuestro modelo encontramos valores parecidos entre el proceso de *Docking* donde arrojó resultados que permiten identificar dónde se lleva a cabo la interacción entre la 6-desoxi-D-glucosa y la insulina.

Capítulo 7

Perspectivas

En el presente trabajo se llevaron a cabo cálculos teórico-cuánticos de la interacción de la glucosa con la Insulina 9INS, puesto que ésta última presenta características similares a la insulina humana, por lo que ahora sería interesante realizar los cálculos empleando esta molécula con las estructuras α -D-glucosa que estimula la secreción de la insulina y β -D-glucosa es transportada a las células.

Por otra parte, debería realizarse un estudio teórico-cuántico de otras posibles interacciones moleculares de la insulina 9INS como GLUT, Glucógeno, o otros posibles transportadores de la glucosa.

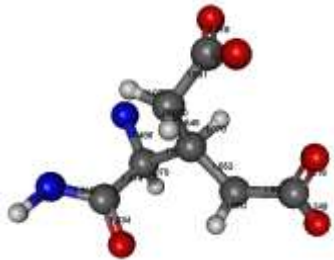
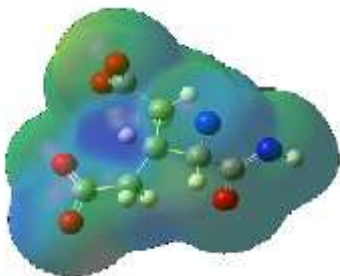
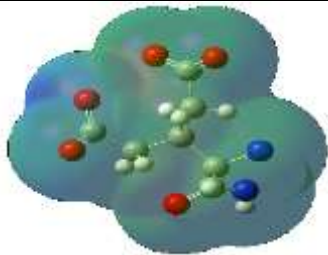
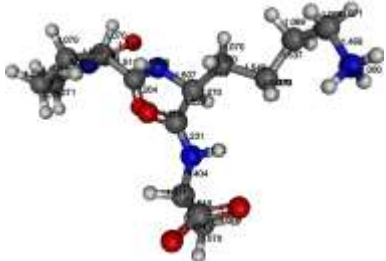
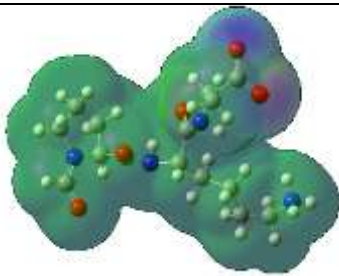
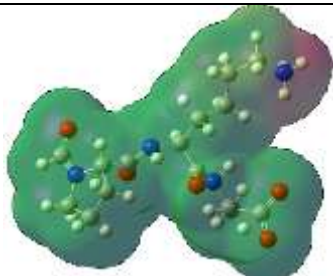
Finalmente, valdría la pena analizar que otros factores a nivel molecular intervienen en la diabetes tipo I y tipo II.

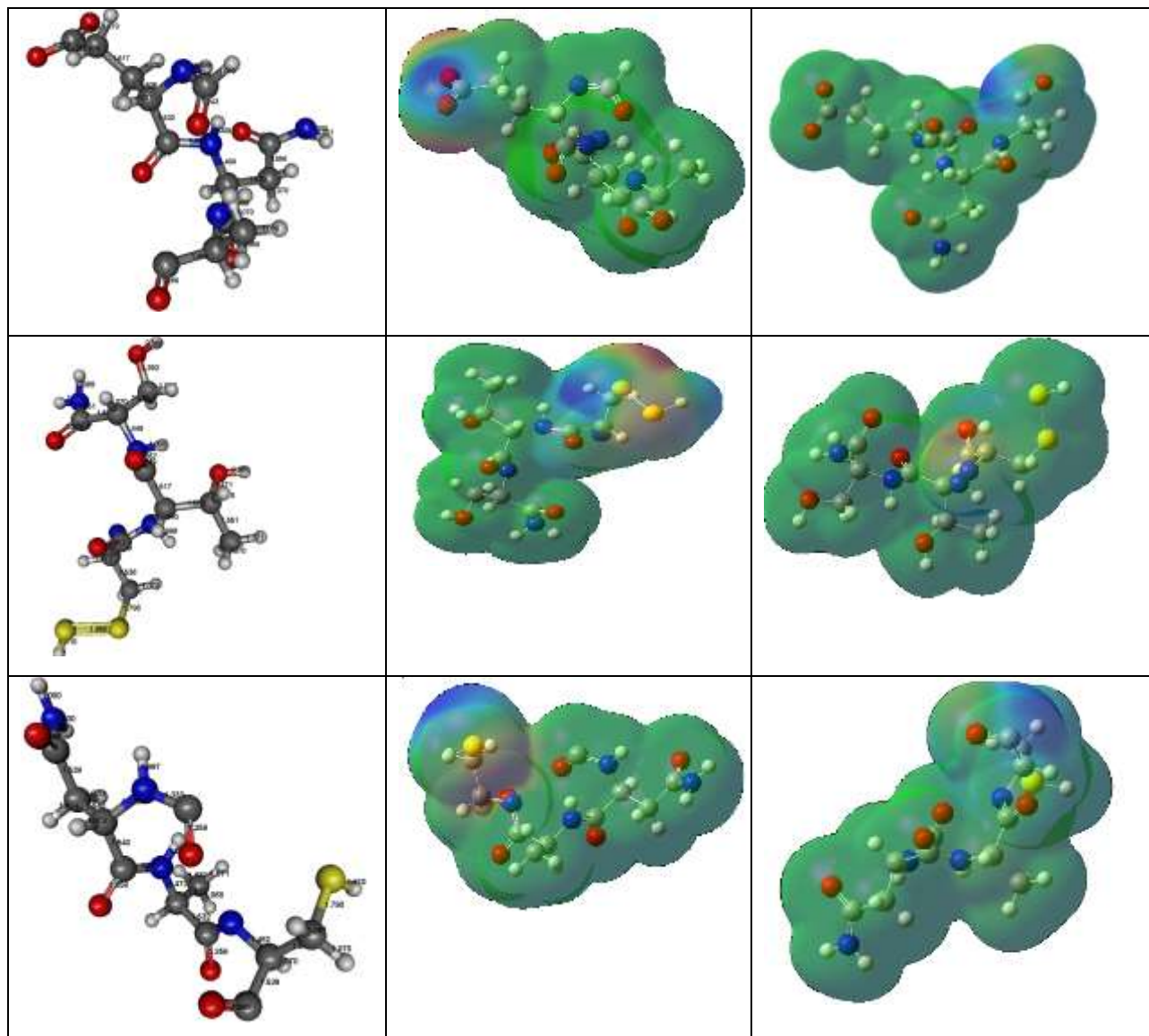
Anexo

En la tabla el HOMO en el estado basal es mapeado para cada una de las secciones de la insulina sobre una isosuperficie de densidad ($\rho = 0.02 \text{ e/a.u.}^3$).

El código de colores indica las posiciones, a lo largo de esta isosuperficie, en donde el HOMO logra sus valores mínimos (rojos) y máximos (azules). Así las zonas más azules indican el lugar más probable de ataque por aceptores de carga, mientras que las regiones rojas (zonas nodales) indican los lugares más probables para un ataque por donadores de carga.

Tabla A.1 Mapeo del HOMO (valor absoluto) a un nivel de teoría Hartree Fock sobre una isosuperficie de densidad ($\rho = 0.02 \text{ e/a.u.}^3$). Las zonas rojas corresponden al valor más bajo del HOMO, las zonas más azules al valor más alto.

Modelos utilizados	Mapeo del HOMO	Mapeo del LUMO
		
		



El efecto en la distribución de sitios activos cuando se introduce la molécula de la glucosa (6 desoxi-D-glucosa) al realizar un *Docking* como se muestra en la *Figura 5.2 del capítulo 5*. Corresponde a la sección 4 del método *ONIOM*.

Determinación de sitios activos en la insulina 9INS

Luis Humberto Mendoza Huizar*, Ana María Vera Meneses, Clara Hilda Rios Reyes.
Electroquímica y Química Teórica. Área Académica de Química. Universidad Autónoma del Estado de Hidalgo. Mineral de la Reforma, Hgo. México.

*Autor a quien la correspondencia debe ser enviada: hhuizar@uaeh.edu.mx

Recibido _____; aceptado _____

RESUMEN

En el presente trabajo analizamos la distribución de sitios activos en la insulina 9INS utilizando el método semiempírico PM6; como un primer paso para entender el papel de esta hormona durante el mecanismo de transporte de la glucosa. Los resultados obtenidos indicaron que un ataque electrofílico a la insulina se encuentra localizado exclusivamente en la cadena de aminoácidos α y los residuos responsables de esta reactividad son 1 Gly, 2 Ile, 3 Val, 4 Glu, 5 Gln y 19 Tyr. Por otro lado, el comportamiento nucleofílico se ubicó en la cadena β y se encuentra dado por los residuos 33 Val, 34 Glu, 35 Ala, 36 Leu, 37 Tyr y 45 Phe β .

Palabras clave: Insulina, 9INS, PM6, sitio activo, Fukui.

INTRODUCCIÓN

La guerra contra la diabetes a través del desarrollo de nuevas drogas ha sido un proceso continuo debido al alarmante aumento de esta enfermedad y complicaciones asociadas. Desafortunadamente la velocidad con la que avanza nuestro conocimiento sobre la enfermedad y sus efectos no están a la par del desarrollo de nuevas drogas. La identificación del receptor en la insulina (RI) y su actividad quinasa intrínseca han sido señalados como el punto clave en el desarrollo de nuevos medicamentos (Balasubramanyam et al., 2001). Por otro lado, se han reportado estudios experimentales en donde se ha elucidado que la activación molecular del RI se lleva a cabo en el sitio de enlace del receptor con la pared celular (Balasubramanyam et al., 2001). Así, un primer paso para entender la interacción de la insulina a través de su receptor implica un buen conocimiento de las zonas reactivas de esta molécula.

Desde su descubrimiento, la insulina ha sido objeto de una intensa investigación bioquímica y biofísica. Se ha reportado, que existe una alta probabilidad de que la insulina forme agregados tales como dímeros, tetrámeros, hexámeros, etc, así como estructuras fibrosas en función del pH, de la temperatura, concentración, etc. (Whittigham et al., 1998, Dodson et al., 1998; Dong et al., 2003; Nielsen et al, 2001; Nettleton et al., 2000). Sin embargo, varios estudios experimentales indican

que la actividad biológica de la insulina monomérica, ver Figura 1, consiste de dos cadenas polipéptidas, la cadena α (21 aminoácidos) y la cadena β (30 aminoácidos). Así, varias regiones de la molécula se han asociado a actividad biológica, tal como la posición de las cisteínas que forman los puentes de bisulfuro, las regiones terminales de la cadena A en el nitrógeno y carbono y los residuos hidrofóbicos en el carbono terminal de la cadena β (Balasubramanyam et al., 2001).

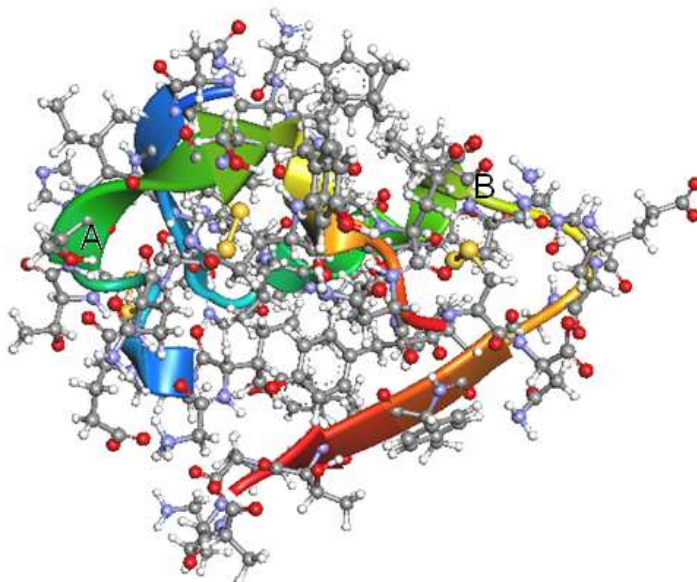


Figura 1

Insulina. Código de acceso 9INS.

Debido a la complejidad y tamaño de la insulina, hasta donde sabemos, no ha sido posible estudiar su reactividad desde el punto de vista de la mecánica cuántica. Actualmente, el estudio de las propiedades electrónicas de algunas proteínas es difícil; es por esto que se requieren de nuevos métodos y habilidades para estudiar sistemas moleculares compuestos de una gran cantidad de átomos. Recientemente los métodos semiempíricos PM6 (Stewart, 2007) desarrollados por el grupo de investigación de Stewart, ofrecen nuevas posibilidades de modelar proteínas en estaciones de trabajo convencionales (Stewart, 2009). Adicionalmente, la predicción de las propiedades estructurales y electrónicas de las bases proteicas que constituyen a la insulina han sido predichas por el método PM6 y son comparables en resultados a aquellos obtenidos por medio de la Teoría de los Funcionales de la Densidad (TFD) (Stewart, 2009). Dado lo anterior, en el presente trabajo analizamos las capacidades del método PM6 para estudiar la reactividad global de la insulina 9INS.

MATERIALES Y MÉTODO

La insulina monomérica 9INS fue optimizada en el nivel PM6, durante la optimización de geometría no se consideró restricción de ningún tipo. Para todos los cálculos se utilizó una estación de trabajo con 2 procesadores Intel Xeon a 2.4 GHZ, con 1 GB de memoria RAM. Los cálculos se realizaron con el software MOPAC 2009 versión Linux (Stewart, 2009) y fueron visualizados con Gaussview (Gaussian Inc, 2007).

TEORÍA

A partir de los conceptos TFD, Parr y Yang mostraron que los sitios en las especies químicas con los valores más grandes de la Función Fukui ($f(r)$) son aquéllos con mayor reactividad. La Función Fukui se define como (Parr et al., 1989)

$$f(r) = \left(\frac{\partial \rho(r)}{\partial N} \right)_v, \quad (1)$$

donde ρ es la densidad electrónica, N es el número de electrones en el sistema y v es el potencial externo ejercido por los núcleos. Es posible evaluar la Función Fukui utilizando orbitales frontera (aproximación de core congelado) (Parr et al., 1989), En esta aproximación, $f(r)$, puede evaluarse relacionando esta con los orbitales frontera en la aproximación de core congelado. Bajo esta aproximación se considera que cuando existe una variación del número de electrones, los orbitales frontera son los únicos afectados. Así, cuando N aumenta a $N+dN$ se obtiene la Función Fukui para un ataque electrofílico:

$$f^-(r) \cong \phi_H^*(r)\phi_H(r) = \rho_H(r), \quad (2)$$

donde $\rho_H(r)$ es la densidad electrónica del orbital más alto ocupado (HOMO).

Cuando N disminuye a $N-dN$ se obtiene la Función Fukui para un ataque nucleofílico,

$$f^+(r) \cong \phi_L^*(r)\phi_L(r) = \rho_L(r), \quad (3)$$

donde $\rho_L(r)$ es la densidad electrónica del orbital molecular más bajo desocupado.

Estas aproximaciones en términos de los orbitales frontera son muy útiles, ya que permiten estudiar problemas de reactividad química considerando únicamente propiedades en el estado basal.

RESULTADOS Y DISCUSIÓN

Ha sido bien establecido en la literatura que la insulina es una proteína que regula los niveles de azúcar en la sangre y en la célula. Sin embargo, en la insulina como en cualquier otra proteína, no todas las regiones tienen la misma importancia y su actividad concreta se localiza en determinadas zonas llamadas “sitios activos”. Si alguna molécula extraña se acopla a este sitio activo puede inhibirlo y la proteína perderá su actividad. Por lo tanto, un buen conocimiento del tipo y distribución de los sitios activos en la molécula permitirán establecer mecanismos que impidan este proceso de inhibición. Los resultados experimentales reportados en la literatura sugieren que el RI de la insulina esta localizado en las regiones terminales de la cadena A y B. Sin embargo, hasta donde sabemos no existe un estudio a nivel molecular que permita establecer el tipo y distribución de sitios activos en la molécula de la insulina. Por otro lado, una distribución puntual de estos sitios activos no se han reportado aun. A este respecto, las teorías aceptadas de la mecánica cuántica en combinación con los criterios de reactividad derivados a partir de la teoría de los funcionales de la densidad permiten determinar el tipo y distribución de sitios activos en la molécula. Sin embargo, actualmente la aplicación de métodos *ab initio* para estudiar las propiedades electrónicas de pequeñas proteínas es prohibitiva en sistemas de súper computo comunes. Recientemente, fue desarrollado el método semiempírico PM6, que introduce correcciones mayores a los métodos PM3 y AM1. Inclusive ha sido reportado que el método PM6 ha permitido mejorar las predicciones del proceso de docking ampliamente utilizado en el diseño de nuevos fármacos (**Bikadi et al., 2009**). Por lo tanto en el presente trabajo utilizamos las capacidades del método semiempírico PM6 para determinar la distribución de sitios activos en la insulina 9INS.

La distribución de sitios activos electrofílicos en un sistema puede derivarse a partir de la teoría de los orbitales frontera dentro de la aproximación de core congelado (**Parr et al., 1989**), ecuación 1 en este trabajo. Con la intención de determinar esta distribución, analizamos los sitios donde el orbital HOMO tiene el valor absoluto más grande. En la Figura 2 se ha mapeado el orbital HOMO, en el estado basal en la molécula 9INS optimizada en el nivel PM6. Es interesante observar que el HOMO se encuentra localizado en una sola región de la molécula, sobre la cadena A, en donde un ataque de tipo electrofílico podría llevarse a cabo. El HOMO se encontró ubicado exclusivamente en la cadena A de la insulina en donde los residuos 1 Gly, 2 Ile, 3 Val, 4 Glu, 5 Gln y 19 Tyr

exhibieron la mayor reactividad ante un ataque de tipo electrofílico. Una lista detallada de los átomos electrofílicos más reactivos en los residuos antes mencionados se reportan en la Tabla 1.

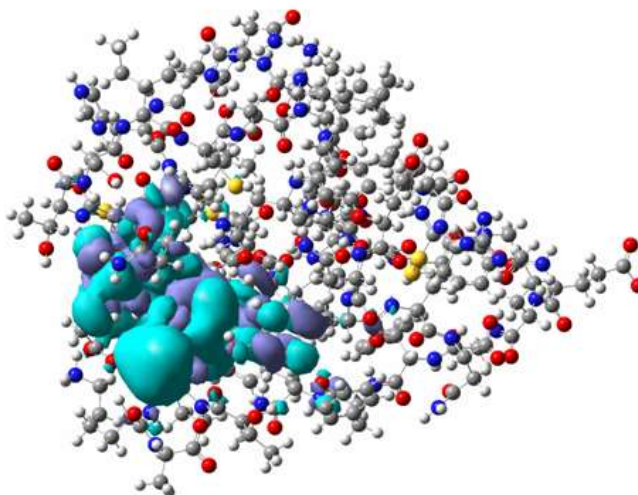


Figura 2 Distribución del orbital HOMO de la insulina monomérica 9INS optimizada en el nivel PM6.

Tabla 1: Átomos electrofílicos mas reactivos obtenidos a partir de la estructura 9INS optimizada en el nivel PM6.

Tipo de átomo	Residuo	Cadena
N	1 GLY	A
CA	1 GLY	A
C	1 GLY	A
O	1 GLY	A
N	2 ILE	A
CA	2 ILE	A
C	2 ILE	A
O	2 ILE	A
CB	2 ILE	A
CG1	2 ILE	A
N	3 VAL	A
CA	3 VAL	A
C	3 VAL	A
O	3 VAL	A
CB	3 VAL	A
CG1	3 VAL	A
CG2	3 VAL	A
N	4 GLU	A
CA	4 GLU	A
C	4 GLU	A
O	4 GLU	A
CB	4 GLU	A
N	5 GLN	A

C	5 GLN	A
O	5 GLN	A
CE1	19 TYR	A
CE2	19 TYR	A
CZ	19 TYR	A
OH	19 TYR	A
H	19 TYR	A

La distribución de sitios activos nucleofílicos se obtuvo empleando la ecuación (2). En la Figura 2 se muestra el mapeo del orbital LUMO, observe que al igual que el orbital HOMO, el orbital LUMO se encuentra bien localizado. Sin embargo, en este caso se encontró localizado exclusivamente en la cadena B, involucrando a los residuos 33 Val, 34 Glu, 35 Ala, 36 Leu, 37 Tyr y 45 Phe B. Los rótulos de los átomos nucleofílicos más reactivos se detallan en la tabla 2.

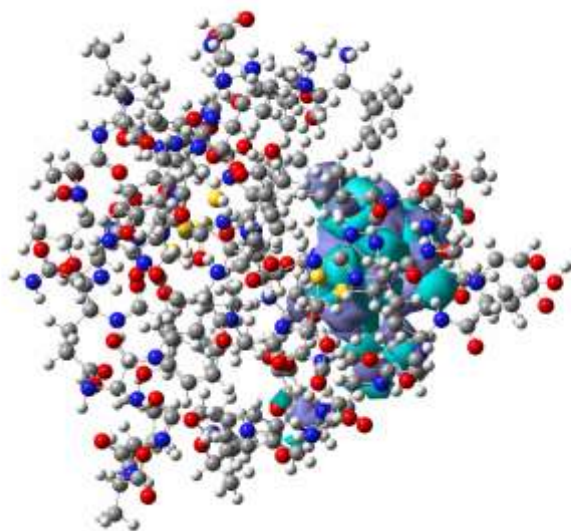


Figura 2

Distribución del orbital LUMO de la insulina monomérica 9INS optimizada en el nivel PM6.

Tabla 2: Átomos nucleofílicos mas reactivos obtenidos a partir de la estructura 9INS optimizada en el nivel PM6.

Tipo de átomo	Residuo	Cadena
C	33 VAL	B
O	33 VAL	B
N	34 GLU	B
CA	34 GLU	B
C	34 GLU	B
O	34 GLU	B
CB	34 GLU	B
H	34 GLU	B
H	34 GLU	B
N	35 ALA	B
CA	35 ALA	B
C	35 ALA	B
O	35 ALA	B
CB	35 ALA	B
H	35 ALA	B
H	35 ALA	B
H	35 ALA	B
H	35 ALA	B
H	35 ALA	B
N	36 LEU	B
CA	36 LEU	B
C	36 LEU	B
O	36 LEU	B
CB	36 LEU	B
CG	36 LEU	B
H	36 LEU	B
H	36 LEU	B
H	36 LEU	B
H	36 LEU	B
N	37 TYR	B
CA	37 TYR	B
H	37 TYR	B
H	37 TYR	B
CG	45 PHE B	B
CD1	45 PHE B	B
CZ	45 PHE B	B
CD2	45 PHE B	B

Es interesante notar que la reactividad obtenida a partir de nuestros cálculos se encuentra ligeramente desplazada de la posición de las cisteínas que forman los puentes disulfuro, tanto en el ataque de tipo electrofílico como nucleofílico. Por otro lado, observe que la reactividad determinada en este trabajo se encuentra localizada en partes externas de la molécula, lo anterior discrepa con los resultados obtenidos experimentalmente en donde la

posición mas reactiva se ha asociado a sitios internos en la molécula en donde el acoplamiento con la glucosa durante el transporte podría ser difícil de llevar a cabo.

AGRADECIMIENTOS

Agradecemos el apoyo económico otorgado a través del proyecto CONACyT APOY-COMPL-2008 No. 91261 y a la Universidad Autónoma del Estado de Hidalgo. También agradecemos al profesor J.P.P. Stewart por proveer una copia gratuita del software MOPAC2009.

ABSTRACT

At present work we determined the distribution of the actives sites of 9INS insulin employing the semiempirical method PM6 like a first step to understand the role of this hormone during the transport mechanism of the glucose. The results obtained indicated that an electrophilic attack on the insulin is located exclusively on the chain A and the residues 1 gly, 2 ile, 3 val, 4 glu, 5 gln y 19 tyr. On the other hand, the nucleophilic attack was located on chain B and it is given by the residues 33 val, 34 glu, 35 ala, 36 leu, 37 tyr y 45 Phe B.

Palabras clave: Insulin, PM6, active sites, fukui.

REFERENCIAS

- Balasubramanyam M., Mohan V. (2001). *Review, Orally active insulin mimics: where do we stand now?*. J. Biosci. 26(3), 383-390.
- Bikadi Z., Hazai E. (2009) *Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock*. J. of Cheminformatics, 1:15. doi:10.1186/1758-2946-1-15
- Dodson G., Steiner D. (1998). *The role of assembly in insulin's biosynthesis*. Curr Opin Struct Biol. 8(2):189–194.
- Dong J., Wan Z., Popov M., Carey P., Weiss M. (2003). *Insulin assembly damps conformational fluctuations: Raman analysis of amide I linewidths in native states and fibrils*. J. Mol. Biol. 330 431–442.
- Gaussview Rev. 3.09, windows version. Gaussian Inc., Pittsburgh, PA.
- MOPAC2009, James J. P. Stewart, Stewart Computational Chemistry, Version 8.032L

Nettleton E., Tito P., Sunde M., Bouchard M., Dobson C., Robinson C. (2000) *Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry*. Biophys J. **79**(2):1053–1065.

Nielsen L, Khurana R, Coats A, Brange S., Vyas S., Uversky V., Fink A. (2001). *Effect of Environmental Factors on the Kinetics of Insulin Fibril Formation: Elucidation of the Molecular Mechanism*, Biochemistry, 40(20), 6036–6064.

Parr R. and Yang W. (1989). *Density Functional Theory of Atoms and Molecules*, Primera edición. Oxford Univ. Press. Oxford, p. 99

Stewart J. (2007) *Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements* J. Mol. Mod. 13(12) 1173-1213.

Stewart J. (2009) *Application of the PM6 method to modeling proteins*. J Mol Model 15(7):765-805

Whittigham J., Scott D., Chance K., Wilson A., Finch J., Brange J. and Dodson G., (2002) *Insulin at ph 2: Structural analysis of the conditions promoting insulin fibre formation*. J. Mol. Biol., 318 (2), 479–490.