

TESIS DE LICENCIATURA

Mecanismo de reacción e inhibición de la glicosiltransferasa 'MshA' de *M.* *tuberculosis*

Carrera: Licenciatura en Ciencias Biológicas.

Tesista: Juan Ignacio Blanco Capurro.

Director: Dr. Marcelo Adrián Martí

Unidad Académica: Departamento de Química Biológica

Mes/año: Mayo 2014.



FACULTAD DE CIENCIAS EXACTAS Y NATURALES
UNIVERSIDAD DE BUENOS AIRES

A Jorge Alfredo Blanco y Graciela Mónica Capurro,

ÍNDICE

	Pags.
RESUMEN -----	4
INTRODUCCIÓN -----	5-10,
i) <u>Micotiol: biosíntesis, función y rol de 'MshA'</u> -----	5-6,
ii) <u>MshA: características generales y clasificación</u> -----	7-8,
iii) <u>Las Glicosiltransferasas de retención y el mecanismo 'SNI'</u> -----	9-10,
iv) <u>Simulaciones computacionales en Biología</u> -----	10,
MATERIALES Y MÉTODOS -----	11-21,
i) <u>Estructuras de partida</u> -----	11-12,
ii) <u>Cálculos de Docking</u> -----	13-14,
iii) <u>Simulaciones de Dinámica Molecular Clásica</u> -----	14-16,
iv) <u>Métodos Híbridos Cuántico-Clásicos (QM-MM)</u> -----	16-18,
v) <u>Búsqueda del camino de Reacción. Umbrella Sampling</u> -----	18-19,
vi) <u>Umbrella Sampling "bidimensional"</u> -----	19-20,
vii) <u>Definición de las Coordenadas de Reacción</u> -----	19-21,
RESULTADOS -----	22-32
i) <u>Cálculos de Docking</u> -----	22-24,
ii) <u>Umbrella Sampling monodimensional</u> -----	24-28,
iii) <u>Umbrella Sampling bidimensional</u> -----	28-32,
DISCUSIÓN -----	35-37
REFERENCIAS BIBLIOGRÁFICAS -----	38-39
ÍNDICE DE FIGURAS -----	40
AGRADECIMIENTOS -----	41

RESUMEN

El Micotiol es un compuesto clave en la resistencia del patógeno *Mycobacterium tuberculosis* a los ataques oxidativos que sufre por parte del sistema inmune del huésped cuando se halla en 'fase latente', lo que sugiere que cumple un rol central en su supervivencia y patogenicidad. En la década pasada, los genes que codifican las 5 enzimas que participan en la biosíntesis de Micotiol en las Micobacterias han sido debidamente identificados y las estructuras proteicas correspondientes resueltas por cristalografía. Entre estas enzimas, la glicosiltransferasa de retención *MshA* cataliza el primer paso de la vía metabólica, y ha demostrado tener numerosas características que la hacen un promisorio blanco terapéutico. Si bien la estructura cristalina de *MshA* de *M. Tuberculosis* no ha sido reportada aún, se encuentra disponible la estructura de *MshA* de *Corynebacterium glutamicum*, la cual posee un 51% de analogía de secuencia con la de *M. tuberculosis*. Las glicosiltransferasas en general son enzimas muy ubicuas y de suma importancia en casi todos los sistemas biológicos, y en particular, las glicosiltransferasas de retención son las que actualmente se encuentran en el centro de atención de muchos glicobiólogos y bioquímicos dado que el producto de reacción retiene la configuración estereoquímica de los reactivos (de ahí su nombre "de retención"), lo cual es incompatible con una sustitución nucleofílica de tipo SN_2 convencional. En los últimos años, un creciente número de hipótesis se han ido inclinando en favor de un mecanismo de "desplazamiento simple frontal" (SN_i). En este contexto, la presente Tesis de Licenciatura tiene como objetivo utilizar las técnicas de simulación computacional conocidas como QM-MM Umbrella Sampling mono y bidimensional, con el propósito de caracterizar estructural y termodinámicamente el mecanismo de reacción de *MshA*, e identificar cual es el orden de eventos clave que determinan el avance de la reacción. Los resultados fueron satisfactorios, pudiéndose reproducir *in silico* y con detalle atómico un mecanismo consistente con el SN_i , y arrojando una barrera de energía de libre de 19 kcal/mol, en excelente correspondencia con cálculos derivados de datos cinéticos experimentales.

INTRODUCCION

i) Micotiol: biosíntesis, función y rol de 'MshA'

El mantenimiento del equilibrio óxido-reducción (redox) en el medio intracelular es un punto central en el desarrollo de la vida. Particularmente en el filo *Actinobacteria*, un grupo de bacterias gram positivas al cual pertenece el llamado “bacilo de Koch” *Mycobacterium tuberculosis*, causante de la enfermedad conocida como tuberculosis, el equilibrio redox está regulado en gran medida por un producto metabólico llamado **Micotiol** (MSH). Desde el punto de vista funcional, el Micotiol es un análogo del "Glutati6n" en bacterias gram negativas, organismos eucariotas y eucariotas superiores; se trata de un agente altamente reductor, y considerado clave en la defensa del bacilo frente a condiciones de stress oxidativo, como las que experimenta luego de ser fagocitado por las células del sistema inmune ^[1].

El Micotiol es el producto final de una vía metabólica de 5 etapas ^[2], en cada una de las cuales interviene una enzima catalizadora (figura 1). De las 5 enzimas involucradas, la primera en actuar es la **MshA**, una enzima proteica que tomando como sustratos **UDP-N-acetilglucosamina** y **L-inositol-1-fosfato** (figura 2), forma el primer precursor del Micotiol, el 3-fosfo-1-D-*myo*-inositol-2-acetamido-2-deoxy- α -D-glucopiranosido ^[3].

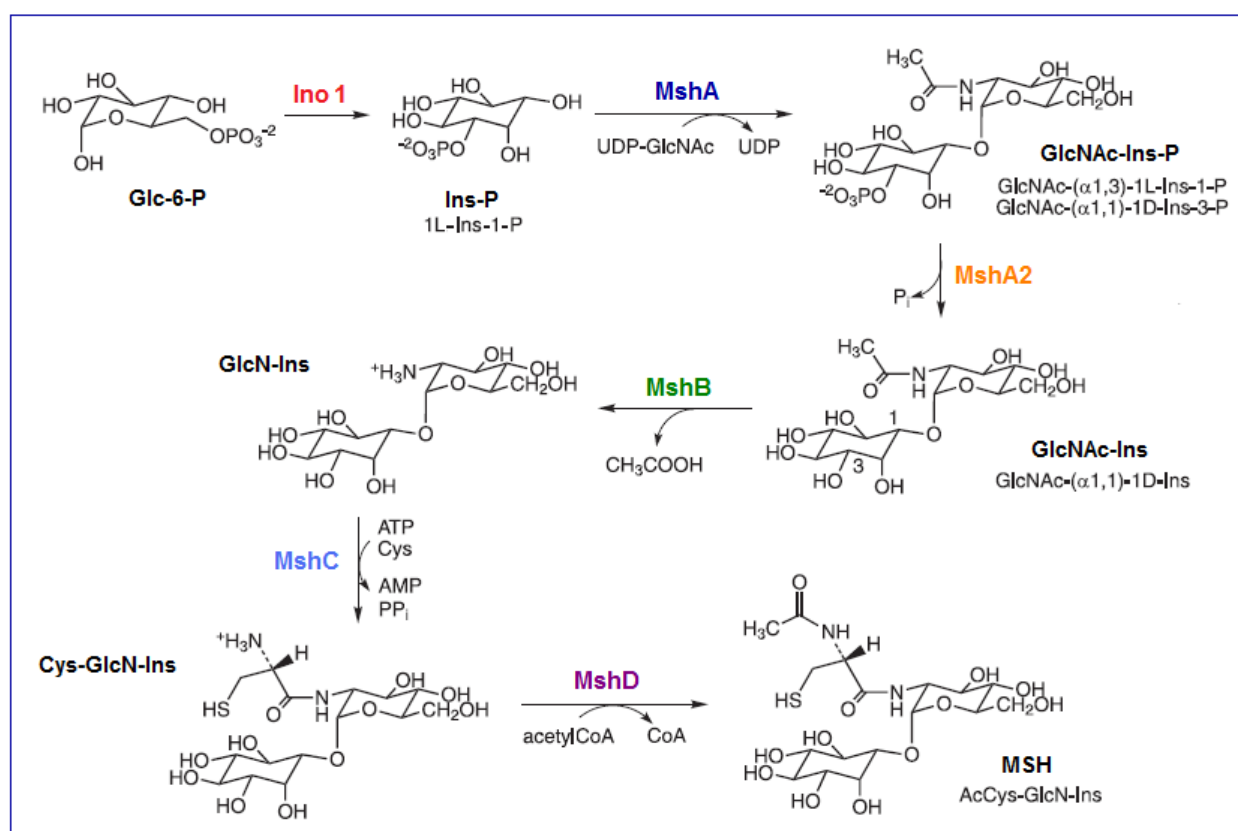


Figura 1. Vía de Biosíntesis del Micotiol (MSH), compuesto formado por Acetil-Cisteína, Glucosamina e inositol.

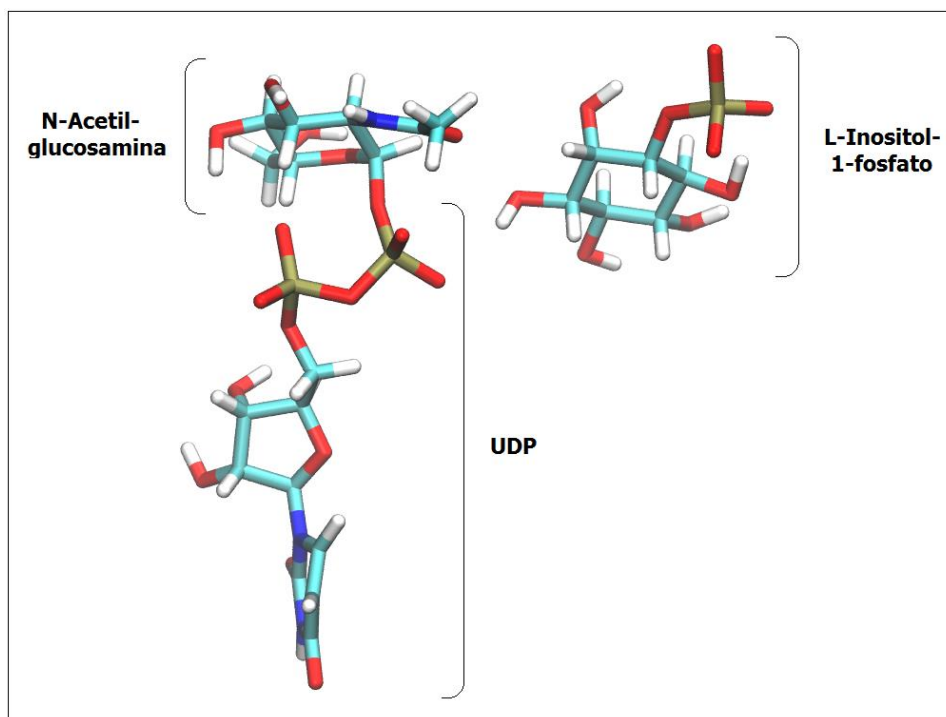


Figura 2. Estructura de las moléculas sustrato de la glicosil-transferasa MshA: el azúcar donador "N-Acetil-Glucosamina" unido al nucleótido UDP, y el azúcar aceptor "L-inositol-1-fosfato".

En cuanto a su utilidad como blanco terapéutico, *MshA* mostró ser un candidato idóneo para inhibir la biosíntesis de Micotiol y combatir a la bacteria patógena por varios motivos, entre los que se pueden destacar:

- Es la primera enzima de la vía, y a diferencia de las demás (que cumplen funciones análogas que también realizan otras enzimas de otras vías, como ser fosfatasas, acetilasas, etc) su función es muy específica y no puede ser llevada a cabo por otra enzima.
- Ha sido sometida a diversos estudios de mutagénesis y se ha comprobado experimentalmente un aumento en la vulnerabilidad de los organismos knock-out a condiciones de stress oxidativo, así como también a distintos tratamientos con fármacos comerciales ^[4].

ii) MshA: características generales y clasificación

Las glicosil-transferasas (nomenclatura general EC 2.4.x.y) son enzimas ubicuas y muy diversas que han cobrado importancia en la última década por sus múltiples aplicaciones en biotecnología y como blanco para el diseño racional de drogas para combatir diversas enfermedades. En términos químicos, su labor consiste en catalizar la hidrólisis de una molécula donora que consta de un hidrato de carbono generalmente unido a un nucleótido por su carbono anomérico, y la transferencia de dicho hidrato de carbono a una molécula aceptora (generalmente otro carbohidrato, lípido u otra molécula orgánica) con la formación de un enlace glicosídico.

Desde el punto de vista estructural, generalmente se las clasifica en 2 grandes grupos según el tipo de plegado, y la presencia o no de cationes divalentes en el sitio activo que participan en la reacción: las superfamilias GT-A y GT-B, respectivamente. *MshA* pertenece a esta última: exhibe un plegado de tipo 'Rossmann' en sus dos dominios y no presenta cationes en el sitio activo. Por otra parte, desde el punto de vista funcional, se las clasifica en las categorías "de retención" o "de inversión" de acuerdo si el carbohidrato transferido retiene o invierte su configuración estereoquímica en el producto de la reacción (figura 3).

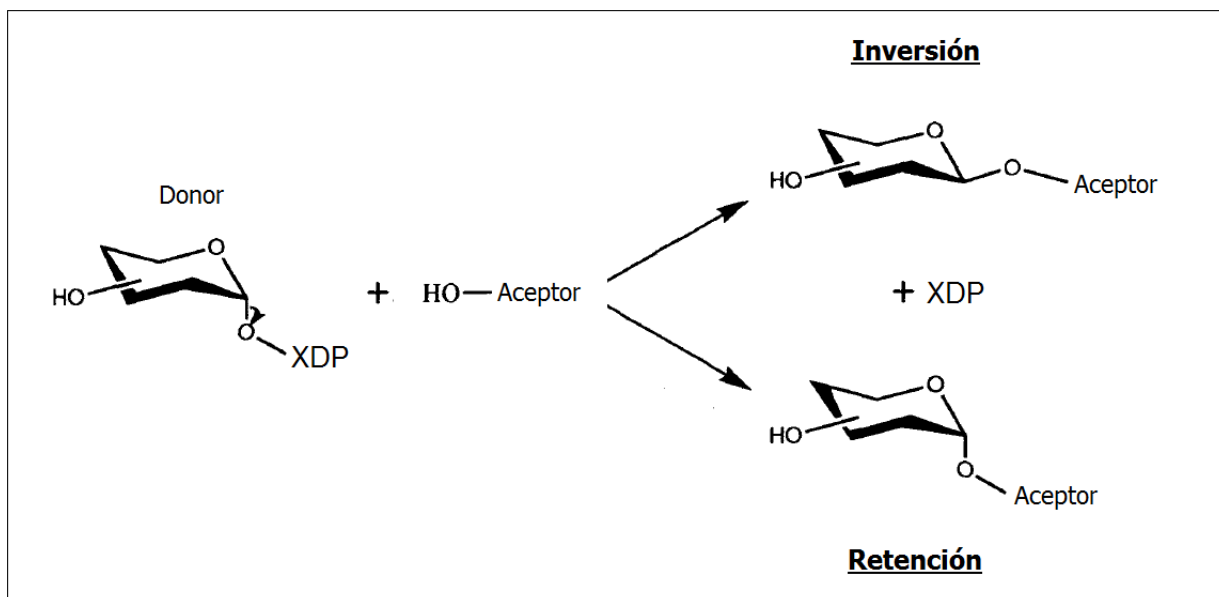


Figura 3. Esquema que ilustra una reacción de transferencia con inversión o retención de la configuración estereoquímica.

MshA es una **glicosil-transferasa de retención**, el grupo que históricamente siempre ha sido el más controversial debido a la gran diversidad de mecanismos de reacción propuestos para racionalizar las características mencionadas del producto en un contexto de relación estructura-función ^[26, 27, 28].

En el presente sólo se encuentra reportada la estructura cristalográfica de la *MshA* de la actinobacteria *Corinebacterium glutamicum* (Cg) ^[6], que presenta una identidad de secuencia del 51 % con la de *M. tuberculosis*, y 20 de los 22 residuos que participan en el anclaje de los sustratos se hallan conservados. La proteína fue cristalizada en 3 situaciones diferentes: en ausencia de sustratos (PDB id: 3C48), en presencia de UDP (PDB id: 3C4Q), y en presencia de UDP e Inositol-1-fosfato (PDB id: 3C4V). Es decir que se desconoce la estructura del complejo ternario completo, ya que falta la N-acetil-glucosamina unida al nucleótido UDP.

Su estructura en solución es un homodímero (figura 4), en el que cada monómero está formado por dos dominios globulares separados por una región bisagra, cerca de la cual se encuentra el sitio de unión de los sustratos. Los monómeros interactúan entre sí a través del dominio N terminal. Tras la unión del sustrato más voluminoso (UDP-N-acetil-glucosamina), la estructura sufre un cambio conformacional en el que los dominios C terminales se aproximan entre sí y giran aproximadamente 90 grados sobre sus ejes, generando el "bolsillo" para la unión del L-inositol-1-Fosfato ^[6].

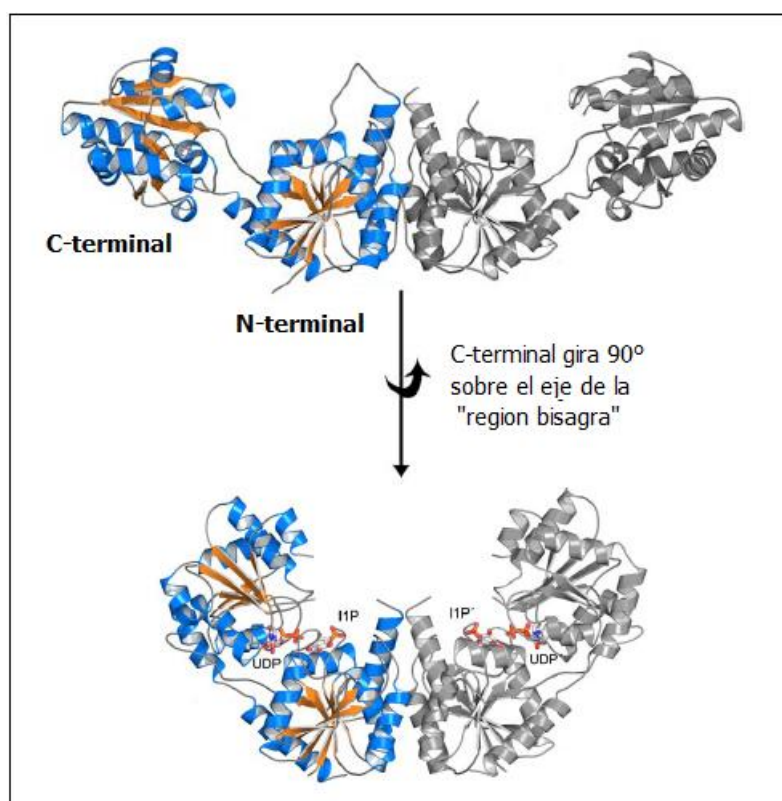


Figura 4. Estructura tridimensional de la glicosiltransferasa de retención *MshA*. En la figura se puede apreciar la estructura del dímero, con 1 subunidad coloreada, y el cambio conformacional que ocurre tras la unión del azúcar donador.

iii) Las Glicosiltransferasas de retención y el mecanismo 'S_Ni'

La estrategia adoptada por las glicosiltransferasas para lograr un producto que retiene la configuración estereoquímica tradicionalmente se intentó explicar por dos posibles mecanismos ^[31], conocidos respectivamente como "desplazamiento doble" y "desplazamiento simple frontal" (figura 5).

El mecanismo "desplazamiento doble" es el primero en ser postulado, en un contexto donde la cantidad de estructuras de glicosiltransferasas disponible era muy limitada. Este mecanismo es postulado a partir de la analogía con el mecanismo de las 'Glicosilhidrolasas' (enzimas que se encontraban abundantemente estudiadas en la época), e involucra una reacción de dos pasos: en el primer paso participa de un residuo que actúa como "base catalítica" y ejerce un ataque nucleofílico sobre el donador, formando un "intermediario covalente" enzima-sustrato en el que la configuración estereoquímica se invierte, y en segundo paso el aceptor ataca nucleofílicamente al intermediario, con lo que la configuración estereoquímica se vuelve a invertir, retomando su estado inicial.

En la última década, con la resolución de nuevas estructuras cristalográficas de glicosiltransferasas de retención, co-cristalizadas con sustratos y análogos, pudo estudiarse con mejor detalle la estructura de los sitios activos, y la aparición cada vez más frecuente de glicosiltransferasas que no poseen residuos capaces de actuar como base catalítica en las cercanías del azúcar donador, como es el caso de *MshA*, llevó a los investigadores a postular un mecanismo alternativo, el '**desplazamiento simple frontal**' (S_Ni) ^[6, 7, 32]. Este mecanismo consiste en la desprotonación de un grupo oxhidrilo del azúcar aceptor, la transferencia de dicho protón al nucleótido (que actúa como base general y grupo saliente), y el ataque nucleofílico del oxhidrilo desprotonado sobre el carbono anomérico del azúcar donador, dando lugar a un estado de transición de tipo "ión oxocarbenio". Dada la ubicación espacial particular de los átomos intervinientes en el centro de reacción, el nucleófilo lleva a cabo el ataque por la misma cara del azúcar por la que el grupo saliente lo abandona, resultando en un producto que retiene la configuración estereoquímica del

sustrato. Los eventos descriptos ocurren en forma concertada, pero necesariamente asincrónica. En este contexto, el gran interrogante que gira en torno a este mecanismo reside justamente en determinar que tan marcada es esa asincronía, y cual de todos los eventos constituye la verdadera fuerza impulsora de la reacción.

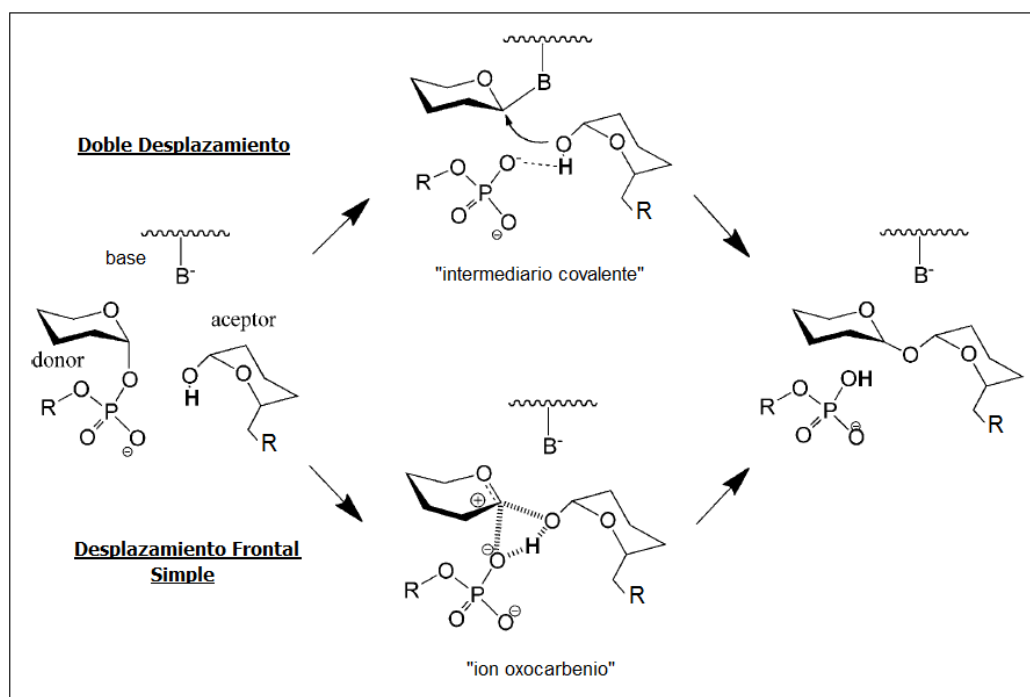


Figura 5. Mecanismos de reacción propuestos en la literatura para las Glicosiltransferasas de retención.

La ausencia de una base catalítica en el sitio activo de MshA y la ausencia de evidencia directa de la existencia de un intermediario covalente ‘enzima-sustrato’, sumado a datos experimentales cinéticos consistentes con un mecanismo concertado ^[6], son indicios que la reacción podría ocurrir a través de un mecanismo ‘S_Ni’ (figura 6).

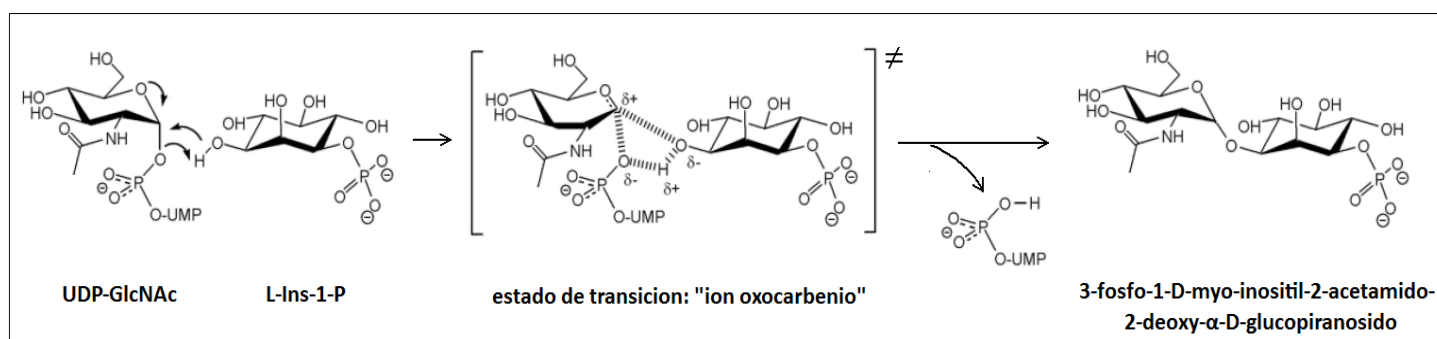


Figura 6. Esquema del mecanismo de Reacción propuesto para la reacción catalizada por la *MshA*. Implica la desprotonación del oxhidrilo 3 del L-Inositol-1-fosfato, la transferencia de dicho protón a la base UDP, y el ataque nucleofílico del oxhidrilo desprotonado sobre el carbono anomérico de la N-acetil-glucosamina.

Actualmente se conocen 2 inhibidores para la *MshA*, el 2,3-dialdehído-UDP-*N*-acetilglucosamina, (que proviene de la oxidación de su propio sustrato), y su propio sustrato con un grupo fluoruro sustituyente en el carbono 5 de la acetil-glucosamina ^[5], cuyo mecanismo de inhibición es actualmente desconocido. Si bien como se mencionó con anterioridad es un blanco prometedor para combatir la enfermedad, es una enzima muy difícil de purificar en estado activo, lo cual hace un sistema idóneo para abordar su estudio por técnicas *in silico*.

iv) Simulaciones Computacionales en Biología

En disciplinas como la Química, Físico-Química y Bioquímica, suele ser muy recurrente que la información microscópica detallada este fuera del alcance de los métodos experimentales actuales. Para suplir esta deficiencia, en las últimas décadas del siglo XX, con el amplio avance de la computación se han desarrollado y perfeccionado diversos métodos basados en la **simulación computacional**.

La simulación computacional de biomoléculas consiste en el empleo de modelos matemáticos implementados en un código de computadora con el objetivo de realizar predicciones acerca del sistema en estudio. Estas técnicas se utilizaron tradicionalmente como herramientas complementarias de los datos experimentales. Sin embargo, en años recientes han ido aportando cada vez mayor valor agregado a los resultados experimentales, en particular con significativo éxito en el estudio estructural y químico de biomoléculas. En la actualidad, estos métodos permiten obtener información estructural, dinámica y termodinámica precisa asociada a la función proteica, información que en muchos casos resulta inaccesible por los medios experimentales conocidos.

Como se explicara en el transcurso de este trabajo, particularmente interesantes resultan los métodos de simulación cuánticos (QM, por 'Quantum Mechanics') aplicables al estudio de reacciones químicas, porque permiten estudiar la evolución de la estructura electrónica a lo largo del proceso reactivo. Estos métodos resultan computacionalmente muy costosos dada la compleja naturaleza de los modelos matemáticos que los fundamentan, por lo que se restringen a sistemas moleculares de no mas de 100 átomos. Por otra parte, existen métodos basados en la mecánica clásica, (MM, por 'Molecular Mechanics') que dada su relativa simpleza y eficiencia computacional son aplicables a sistemas del orden de los miles de átomos como las proteínas, los ácidos nucleicos y las membranas lipídicas. Estos últimos resultan útiles en estudios estructurales de procesos como la unión de ligando, el plegamiento proteico, la desnaturalización y los cambios conformacionales, pero no permiten el estudio de reacciones químicas. Es en este contexto que tuvo lugar el desarrollo de métodos híbridos (QM/MM), que permiten estudiar procesos químicos tratando los sustratos y el sitio activo enzimático de manera cuántica, considerando a la vez de manera explícita el entorno proteico tratándolo de manera clásica. Los métodos QM-MM han recorrido un largo camino y evolución desde que fueron introducidos por primera vez en 1976^[9], y le han valido el Premio Nobel de Química 2013 a sus desarrolladores y principales impulsores. Entre otras aplicaciones, han sido aplicados satisfactoriamente al estudio de numerosas reacciones enzimáticas, quedando demostrada la invaluable importancia de sus aportes en el campo.

El objetivo de este trabajo de tesis es, utilizando las estructuras y la información descriptas previamente como punto de partida, comprender desde un punto de vista termodinámico y con detalle atómico-molecular el mecanismo de reacción de la enzima MshA, aplicando métodos de simulación QM/MM.

Palabras clave: *Micotiol ; Mycobacterium tuberculosis ; glicosil-transferasa de retención ; mecanismo S_Ni ; ión oxocarbenio ; simulación computacional ; QM-MM*

MATERIALES Y MÉTODOS.

i) Estructuras de partida

En este primer apartado se describen las técnicas utilizadas para generar el complejo proteína-ligando que constituyeron el punto de partida para todas las simulaciones computacionales llevadas a cabo en la presente Tesis.

Se partió de 2 estructuras cristalográficas, resueltas por Vetting y colaboradores, tomadas de la base de datos de estructuras Protein Data Bank (PDB). Estas estructuras son:

- PDB ID 3C4V: proteína en complejo con UDP y L-inositol-1-fosfato, en adelante "conformación cerrada".
- PDB ID 3C48: apo-proteína, en adelante "conformación abierta".

Cada una de las estructuras anteriores presentan las coordenadas de 2 monómeros (cadenas A y B), y poseen zonas no resueltas, es decir, zonas internas de la estructura donde el mapa de densidad electrónica obtenida no permitió resolver con precisión las coordenadas de las entidades atómicas correspondientes:

- PDB ID 3C4V 3 Gaps: - residuos 139 a 146
(figura 7) - residuos 266 a 273
- residuos 410 a 418
- PDB ID 3C48 4 Gaps: - residuos 16 a 22
- residuos 143 a 146
- residuos 206 a 209
- residuos 267 a 272

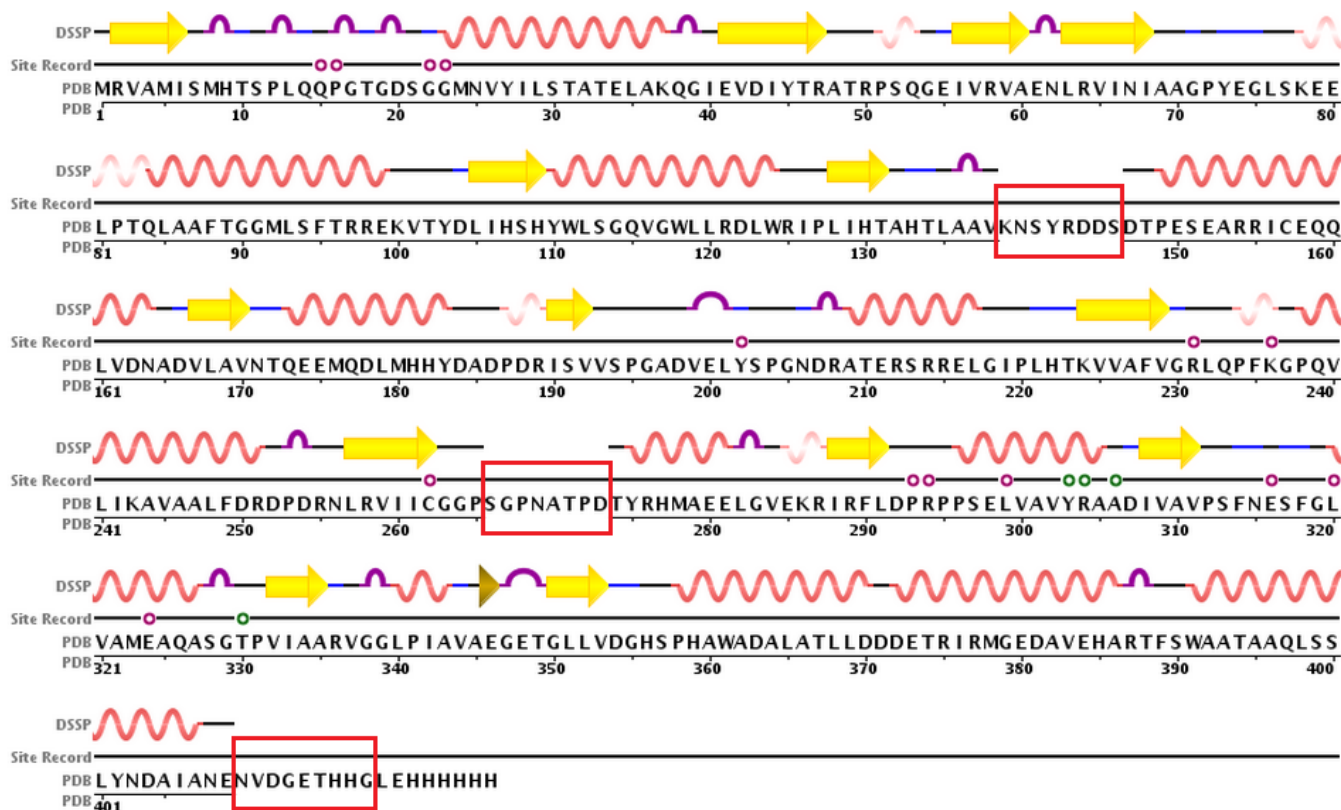


Figura 7. Estructuras primaria y secundaria de la MshA de *C. glutamicum*, tomada del sitio web de PDB [8], (PDB ID 3C4V). Los recuadros rojos indican la posición de las 3 zonas no resueltas de la estructura tridimensional;

Frente a esta problemática, fue necesario completar estas estructuras utilizando herramientas bioinformáticas para lograr una total correspondencia entre secuencia y estructura.

Puesto que ambos cristales reportan las coordenadas de 2 cadenas ('A' y 'B') con distinto grado de completitud, se disponía en total de 4 monómeros como posibles moldes para lograr un monómero completo definitivo. Afortunadamente, la mayoría de estos "gaps" en las estructuras corresponden a loops expuestos al solvente y que no participan directamente en el anclaje del sustrato en el sitio activo, a excepción el loop que comprende los residuos 16 a 23 (que esta ausente en la conformación abierta, pero esta presente en la conformación cerrada), el cual luego de la unión del UDP y el cambio conformacional sufre un reordenamiento para interactuar con el sustrato que limita su movilidad.

Utilizando la herramienta de alineamiento estructural Multiseq del programa de visualización de biomoléculas VMD^[10], fue posible alinear las coordenadas de las 4 cadenas y combinarlas de manera tal de generar un primer prototipo más completo en la conformación cerrada (figura 8), para proceder luego a la colocación del sustrato faltante y el posterior estudio del mecanismo de reacción.

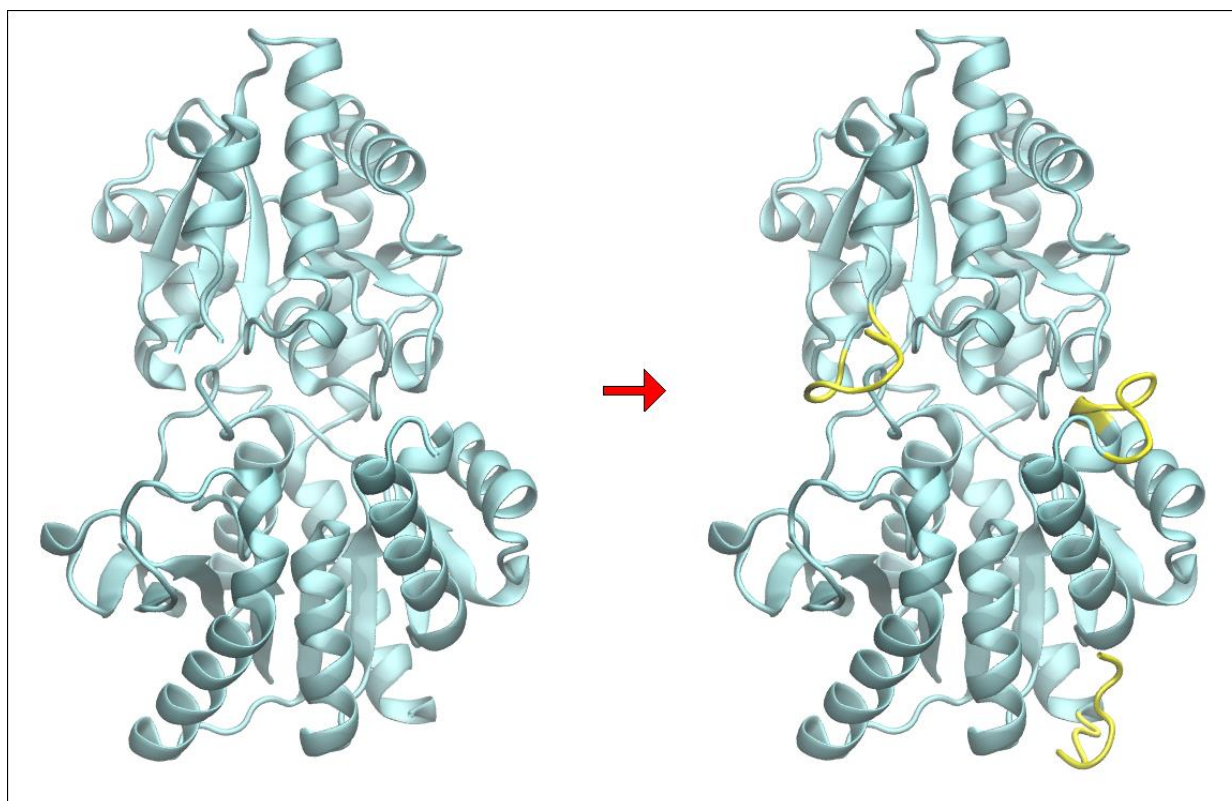


Figura 8. Estructura terciaria de la *MshA* de *C. glutamicum*. Izquierda: estructura cristalográfica incompleta (PDB ID 3C4V); derecha: Modelo final completo de la estructura "apo-cerrada".

Para generar el complejo ternario proteína-ligandos, ya que el cristal 3C4V contenía las coordenadas de los sustratos UDP e inositol-fosfato, se utilizó nuevamente la herramienta de alineamiento estructural del programa VMD para superimponer las coordenadas de este cristal con las del modelo generado, y luego de esta manera poder copiar las coordenadas de los sustratos UDP e inositol fosfato al PDB del modelo. Este modelo de "complejo ternario incompleto", fue sometido a un protocolo de simulación MM (que se describe en el capítulo siguiente) con el objeto de optimizar la estructura y analizar su comportamiento dinámico en presencia de los sustratos, y a la vez para obtener una estructura estable para los cálculos de docking del sustrato faltante, la N-acetil-glucosamina.

ii) Cálculos de Docking

El cristal 3C4V constituye el registro más completo de la proteína MshA cristalizada con sustratos. No obstante, en lugar de presentar el sustrato UDP-N-acetil-glucosamina completo, presenta solo el UDP. Por este motivo, fue necesario incorporar a los modelos las coordenadas de la N-acetil-glucosamina faltante mediante la técnica de Docking.

Como propusieron Vetting y colaboradores, uno de los modelos mas plausibles de orientación de la N-acetil-glucosamina en el sitio activo de la *MshA* es la llamada conformación “bent-back”, una orientación observada previamente en otras estructuras de glicosil-transferasas de retención, en la que el azúcar parece adoptar una inclinación particular respecto del nucleótido (figura 9).

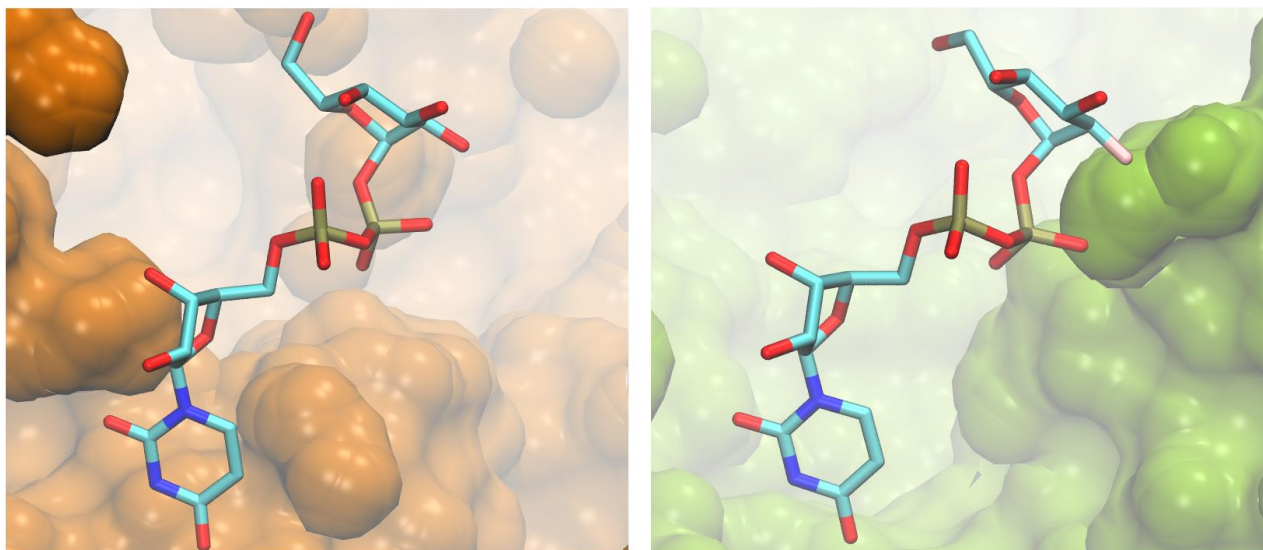


Figura 9. conformación “bent back” en estructuras cristalográficas de otras glicosil-transferasas: izquierda, OtsA de *E. coli* en complejo con UDP-glucosa (pdb ID: 1UQU). derecha, WaaG de *E. coli* unida a UDP-2-fluoro-glucosa (pdb ID: 2IW1);

Partiendo de esta base, el primer objetivo específico fue validar mediante cálculos de Docking la hipótesis de Vetting. Todos los cálculos fueron efectuados con el programa Autodock 4 ^[11].

El programa Autodock se basa en la utilización de un algoritmo genético y una Función de Scoring (figura 10.(a)) para evaluar la energía de distintas conformaciones de un complejo proteína-ligando. Para esto utiliza un potencial semiempírico (figura 10.(b)), con el cual construye una serie de grillas (una por cada tipo de átomo del ligando, mas una grilla adicional electrostática y una de desolvatación), que utilizara para calcular el costo energético de colocar cualquier átomo del ligando en un determinado punto del espacio del entorno proteico. Un algoritmo genético permite explorar rápidamente un elevado número de conformaciones y descartar las energéticamente desfavorables. El resultado del cálculo de Docking es una lista de las conformaciones más favorables energéticamente, rankeadas por su población (entendiéndose como población a la frecuencia en que esa dada conformación es obtenida).

Es importante destacar que no se trató un cálculo de Docking convencional, sino que el objetivo conllevaba posar la N-acetil-glucosamina directamente sobre el grupo pirofosfato de la molécula de UDP, ya que en realidad el nucleótido y el azúcar donador se encuentran unidos por un enlace glicosídico. En este sentido, fue necesario realizar lo que se conoce como **docking sesgado** (“bias-docking”), para asegurar que oxígeno anomérico del azúcar quedara siempre lo más cerca posible del fosfato beta del UDP y así poder reproducir correctamente el enlace. La técnica de bias docking consiste sencillamente en modificar la grilla de energía correspondiente al átomo particular cuya ubicación y orientación se desea sesgar (en este caso el oxígeno), colocando un pozo de potencial sobre las coordenadas espaciales donde se desea colocar preferencialmente dicho átomo de modo tal de premiar energéticamente esa ubicación, que de otro modo sin el sesgo estaría obstaculizada por la repulsión estérica entre ambas moléculas.

$$\begin{aligned}
 \text{(a)} \quad \Delta G &= (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf}) \\
 \text{(b)} \quad V &= \underbrace{W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)}_{1.} + \underbrace{W_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right)}_{2.} + \underbrace{W_{elec} \sum_{i,j} \frac{q_i q_j}{e(r_{ij}) r_{ij}}}_{3.} + \underbrace{W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2 \sigma^2)}}_{4.}
 \end{aligned}$$

Figura 10. (a) Función de scoring energética del Programa Autodock 4.2. **(b)** Campo de Fuerzas Semiempírico de Autodock 4.2, en el que se detallan: 1. Término de Dispersión/Repulsión (con parámetros A y B tomados del Campo de Fuerzas de AMBER); 2. Término de Puentes de Hidrógeno direccionales (con E(t) Función de direccionalidad, y parámetros C y D calibrados empíricamente); 3. Término electrostático; y 4. Término de desolvatación (con parámetros de desolvatación S, pesados exponencialmente por la distancia y el factor Sigma). Las Constantes empíricas W están basadas en datos de Constantes de Binding medidas Experimentalmente para una serie de complejos y son utilizadas para calibrar el Potencial.

iii) Simulaciones de Dinámica Molecular Clásica

Las simulaciones computacionales consisten esencialmente en modelos matemáticos implementados en un software que permiten estudiar el comportamiento de distintos sistemas y realizar predicciones sobre su evolución. Todo modelo consta de aproximaciones, y en este sentido la complejidad del modelo estará estrechamente ligada al costo computacional que el mismo demanda.

La mayoría de los métodos de simulación empleados en química computacional se basan esencialmente en una **Función de Energía Potencial**, también llamada en la literatura simplemente como "potencial". La derivada del potencial respecto de la posición de un determinado ente (ej: átomo) permite obtener las fuerzas que actúan sobre dicho ente. Para el estudio del comportamiento de biomoléculas es muy común el uso de modelos de escala atómica, es decir, que representan a los distintos átomos como entes individuales con sus propiedades particulares^[12]. Dentro de este grupo, existen 3 tipos muy comúnmente utilizados y que poseen distintas aplicaciones, y estos son los basados en la Mecánica Cuántica (QM), la Mecánica Clásica (MM) y los modelos híbridos (QM-MM). Este trabajo de tesis se fundamenta principalmente en un amplio uso tanto de los métodos MM como los QM-MM.

Los métodos MM^[12] parten de la siguiente aproximación: la naturaleza cuántica de los electrones (e incluso los electrones per se) es ignorada, y a cada átomo se le asigna una carga atómica fija durante toda la simulación, por lo cual estos métodos no permiten analizar la evolución temporal de la distribución de carga en una molécula, ni tampoco describir procesos reactivos como ser formación o ruptura de enlaces químicos. La energía potencial del sistema, entonces, depende exclusivamente de las posiciones de los núcleos atómicos. Esta aproximación permite el empleo de términos matemáticos sencillos, y como consecuencia se logra abaratar significativamente el costo computacional de la simulación, lo cual hace a los métodos MM viables para simular sistemas del orden de miles de átomos, como ser una proteína rodeada de agua. En los métodos MM, las interacciones entre los núcleos se agrupan en dos tipos:

- 1) de unión: Describen las interacciones entre dos átomos unidos entre sí directamente o hasta 2 enlaces de distancia. Consisten en aquellos términos del potencial cuya energía se ve afectada por los estiramientos de los enlaces, las flexiones de los ángulos entre 2 átomos, y la rotación de dos átomos adyacentes sobre un eje (ángulos dihedros). Los estiramientos y las flexiones angulares son modeladas mediante un **oscilador armónico**, mientras que las rotaciones de los enlaces en el plano son modeladas mediante una función trigonométrica.

2) **de no unión**: describen la interacción entre átomos ubicados a más de 3 enlaces de distancia de la misma molécula, o bien entre átomos de moléculas distintas, y consisten en un término para la contribución electrostática computada con la **Ley de Coulomb**, y un término para la contribución de Vaan Der Waals modelada por un **potencial de Leenard-Jones 12-6**.

En este trabajo de Tesis, se utilizó el potencial de AMBER ^[13], el cual es descrito en la figura 11.

$$V(\mathbf{r}) = \underbrace{\sum K_b(b - b_0)^2 + \sum K_\theta(\theta - \theta_0)^2 + \sum (V_n/2)(1 + \cos[n\phi - \delta])}_{\text{interacciones de union}} + \underbrace{\sum (A_{ij}/r_{ij}^{12}) - (B_{ij}/r_{ij}^6) + (q_i q_j / r_{ij})}_{\text{interacciones de no union}}$$

Figura 11. Potencial de Amber. En azul se representan los términos matemáticos que modelan las interacciones de unión (1. Término de estiramiento de enlaces; 2. Término de flexión angular; 3. Término de rotación de ángulos diedros), y en rojo las interacciones de no unión (Potencial de Leenard-Jones 12-6 e interacción Coulómbica).

Los parámetros que se observan en la figura 11 (las constantes K_b , K_θ y V_n , los valores de equilibrio b_0 , θ_0 y γ , las constantes A y B del potencial Leenard-Jones, y las cargas q) deben ser ajustados especialmente para cada molécula individual o residuo que se desee incluir en la simulación (cada aminoácido, monosacárido, bases nitrogenada, etc, debe contar con sus propios parámetros). Estos parámetros se suelen obtener tanto de datos experimentales de espectroscopía o cristalografía, como así también de rigurosos cálculos cuánticos, y en ocasiones emplean correcciones empíricas. Al conjunto o "set" de parámetros derivados para utilizarse de forma conjunta en una simulación se lo conoce como **Campo de Fuerzas**.

Una vez descriptos los conceptos de "Potencial" y "Campo de Fuerzas", y aplicados los mismos a un sistema de moléculas, es posible calcular la propagación temporal de las posiciones de los átomos del sistema mediante la técnica conocida como **Dinámica Molecular** ^[12]. La misma se basa en resolver las Ecuaciones de Movimiento de Newton para cada átomo del sistema; así, en cada "paso" de la simulación se calcula la energía y luego las fuerzas que actúan sobre cada átomo, y junto con las posiciones y velocidades atómicas de dicho paso, se calculan las nuevas posiciones y velocidades en un intervalo de tiempo posterior ("dt"). De este modo se genera una trayectoria que describe como cambia la conformación espacial del sistema a lo largo del tiempo. La integración de las Ecuaciones de Newton se calcula computacionalmente, mediante el llamado **Algoritmo de Verlet** ^[14]. La elección del "dt" es una situación de compromiso, ya que un dt muy chico, si bien representa la propagación del movimiento de forma "mas exacta", provoca que se requiera una mayor cantidad de cálculos para alcanzar la longitud de una dada trayectoria total. Generalmente, se utiliza un dt ligeramente menor al movimiento más rápido del sistema, que en caso de biomoléculas corresponde a la vibración de los enlaces CH y está en el orden de los femptosegundos.

La Dinámica Molecular permite no solo obtener una trayectoria temporal de distintas conformaciones del sistema, sino también otras propiedades termodinámicas como la temperatura. La temperatura de un sistema puede ser entendida a nivel microscópico como el promedio de la energía cinética de las partículas, lo que implica no solo que conociendo las velocidades de las partículas se puede calcular la temperatura, sino que además controlando las velocidades se puede tener control de la temperatura del sistema (por ejemplo, para estudiar un proceso a temperatura constante, o estudiar como el mismo sistema se comporta a distintas temperaturas, etc). La temperatura se controla a través de los Termostatos, que son algoritmos que actúan como si el sistema estuviese acoplado a un "baño térmico" como fuente de temperatura. En cada paso, este algoritmo re-escala las velocidades de los átomos, de modo tal que el cambio sea proporcional a la diferencia entre la temperatura del sistema y la temperatura del baño térmico.

Existen distintos tipos de termostatos, siendo el utilizado en este trabajo el "**Termostato de Berendsen**"^[15].

Protocolo de una simulación MM

A continuación, se describe brevemente el protocolo al que fueron sometidas cada una de las estructuras generadas y descriptas previamente en las secciones i) y ii) de este capítulo:

1- Preparación: Las estructuras en formato PDB son en primer lugar entregadas al modulo LEaP del paquete AMBER, el cual es un software capaz de separar la información contenida en el PDB en dos archivos separados, por un lado la '**topología**', y por el otro las '**coordenadas**' iniciales. Para ello, se empleo el campo de fuerzas para aminoácidos ff99SB^[16] incluido en el paquete AMBER 12, y el campo de fuerzas para carbohidratos GLYCAM-06_b^[17] para la N-acetil-glucosamina. Parámetros para el UDP y el 1-inositol-Fosfato fueron generados a partir de cálculos cuánticos y el módulo Antechamber del paquete AMBER, cuyo procedimiento se omite detallar por razones de brevedad. Cada una de las estructuras fue solvatada en una caja de aguas octaédrica según el modelo TIP3P^[18]. La topología final resultante, conteniendo la proteína en agua, alcanzó un total de aproximadamente 36.000 átomos.

2- Optimización: consiste en someter las estructuras solvatadas a un algoritmo de "Minimización de la energía", el cual se encargara de explorar distintas conformaciones geométricas del sistema con el objetivo de minimizar la energía del mismo. Esta etapa del protocolo tiene un doble propósito: por un lado, eliminar contactos geométricos desfavorables provenientes de la propia estructura cristalográfica o que sean producto de las técnicas de modelado por homología que fueron utilizadas para rellenar los gaps, y por otro lado, distribuir homogéneamente las aguas (que son añadidas por LEaP de manera espacialmente aleatoria), para evitar la formación de "burbujas de vacío" al calentar el sistema en el paso siguiente. Se emplearon 500 ciclos de minimización utilizado el **algoritmo de gradiente conjugado**.

3- Termalización: consiste en calentar el sistema de 0 a 300 grados Kelvin. Para esto, se utilizó el Termostato de Berendsen, el cual asigna de manera controlada velocidades aleatorias a los átomos (tomadas de la distribución de Boltzman) hasta estabilizar la temperatura en torno al valor deseado. Esto se conoce como ensamble NTV o 'canónico', dado que el número de partículas, la temperatura y el volumen del sistema se mantienen constantes durante la simulación. Este proceso se llevó a cabo con un 'dt' de 0.001 ps, durante 200.000 pasos, obteniendo una trayectoria de 200 ps.

4- Equilibrio de la presión: dado que el volumen se mantiene constante durante la etapa de equilibración de la temperatura, el sistema puede sufrir cambios de presión. El objetivo de esta cuarta etapa es equilibrar la presión del sistema en torno a un valor equivalente a un 1 atm. Esto se realizó con una simulación de 500 ps. Al final de esta etapa, se grafican la temperatura, la presión, y la energía total versus el tiempo, para tener un control global del proceso antes de proceder a la siguiente etapa. Asimismo, se revisan las trayectorias con un programa visualizador (VMD), para chequear que la estructura se mantiene estable.

5- Dinámica Molecular de Producción: se trata de la simulación a presión constante, de una duración de 10 a 100 veces más extensa que las anteriores, es decir, en el orden de los nanosegundos. Esta es la simulación a partir de la cual se extraen los parámetros termodinámicos deseados y se estudia la estabilidad de las estructuras.

En todas las simulaciones de las etapas 3, 4 y 5, se utilizaron condiciones periódicas de contorno, y el método de sumas de Ewald con un valor de corte de 8 Å°. El algoritmo SHAKE se aplicó a todos los enlaces conteniendo hidrógeno.

iv) Métodos Híbridos Cuántico-Clásicos (QM-MM).

Hasta aquí hemos explicado brevemente los fundamentos teóricos de los métodos clásicos (MM), y se han mencionado sus puntos fuertes y débiles: por un lado la capacidad de estudiar sistemas del orden de decenas de miles de átomos con un costo computacional bajo, pero por otro lado la imposibilidad de estudiar procesos reactivos que involucren transferencias de electrones.

Para el estudio de las reacciones químicas, existen los ya mencionados métodos cuánticos (QM), basados en la resolución de la Ecuación de Schrödinger, que permiten describir en detalle la estructura electrónica de los sistemas. La formulación matemática de los métodos QM excede el alcance de la presente Tesis, motivo por el cual no será descrita en detalle. Debido a su alto costo computacional, se torna muy difícil su aplicación a sistemas de más de 100 átomos. Es por estas razones que surgieron los métodos híbridos (QM-MM) ^[9, 19], que nuclea las ventajas de las dos grandes familias de métodos ya mencionadas: por un lado la versatilidad y el relativo bajo costo de los métodos MM aplicable a sistemas de decenas de miles de átomos, y por otro lado la capacidad de estudio de fenómenos reactivos de los métodos QM.

Los métodos híbridos QM-MM se aplican a los sistemas estableciendo dentro de ellos una división en sub-sistemas, por un lado el 'subsistema QM', donde todos los átomos contenidos en él son tratados de manera cuántica (típicamente en el caso de una proteína, el sitio activo o los residuos catalíticos del mismo y los sustratos), y por otro lado un 'subsistema MM' que es tratado de manera clásica (el resto de la matriz proteica, más las moléculas de agua y iones que solvatan el sistema) ^[20]. De este modo, la reacción que se desea estudiar quedará circunscrita al subsistema QM, mientras que la influencia del entorno sobre las propiedades reactivas será contemplada en el subsistema MM. La función de energía potencial es un Hamiltoniano que posee entonces 3 componentes (figura X), a saber: un término de energía cuántica, un término de energía clásica, y un tercer término de acoplamiento, que describe la interacción electrostática y de Van Der Waals que existe entre ambos subsistemas.

Todos los cálculos QM-MM de este trabajo de Tesis se efectuaron utilizando los métodos QM semi-empíricos incluidos en el paquete AMBER ^[21], especialmente el método de **Self-Consistent-Charge Density Functional Tight-Binding (SCC-DFTB)** ^[22, 30, 34].

Brevemente, los métodos QM semiempíricos se basan en la parametrización previa de ciertas variables cuánticas relacionadas a la función de onda y los orbitales moleculares, que tiene como objetivo reducir o eliminar por completo el número de integrales que los programas deben realizar (con lo que se ahorra mucho tiempo de cómputo). Particularmente, la parametrización del método SCC-DFTB está basada en una expansión de segundo orden de la energía de Kohn-Sham en la **Teoría del Funcional de la Densidad (DFT)** con respecto a la fluctuación de la densidad de cargas. El orden 'cero' es equivalente al esquema 'Tight Binding' (TB) que no es auto-consistente, mientras que el segundo orden permite derivar una expresión transparente para los elementos de una Matriz de Hamiltoniano. Estos elementos son luego modificados por una redistribución auto-consistente de las cargas de Mulliken (SCC) ^[22, 34].

En otras palabras, el método SCC-DFTB es un método cuántico, mucho más veloz que los llamados *ab initio* o 'de primeros principios', pero mucho menos preciso dadas las aproximaciones que utiliza. Analizar la calidad y capacidad de este método dentro de un esquema híbrido QM-MM para el estudio de reacciones químicas que involucran carbohidratos es parte de los objetivos de esta Tesis.

Protocolo de una simulación QM-MM

El protocolo básico consistió en todos los casos en tomar una estructura pre-termalizada y con presión equilibrada mediante Dinámica Molecular MM como estructura de partida. Todas las simulaciones híbridas QM-MM fueron realizadas con un paso de integración de 0.001 picosegundos a una temperatura de 300 grados Kelvin controlada por termostato Berendsen, y equilibradas a 1 atm de presión. Se utilizaron condiciones periódicas de contorno, un valor de corte de 10 Ångström para las interacciones de no unión, y el algoritmo SHAKE a todos los enlaces conteniendo hidrógeno.

El subsistema QM elegido comprendió un total de **64 átomos cuánticos** y una carga total de -2 (figura 12), y comprendió la porción pirofosfato y el azúcar N-acetil-glucosamina del UDP-N-acetil-glucosamina (es

decir, se excluyó la porción del nucleótido Uracilo), y el L-inositol-1-fosfato sin el grupo ortofosfato. En principio, no se incluyó ningún residuo proteico dentro del subsistema QM.

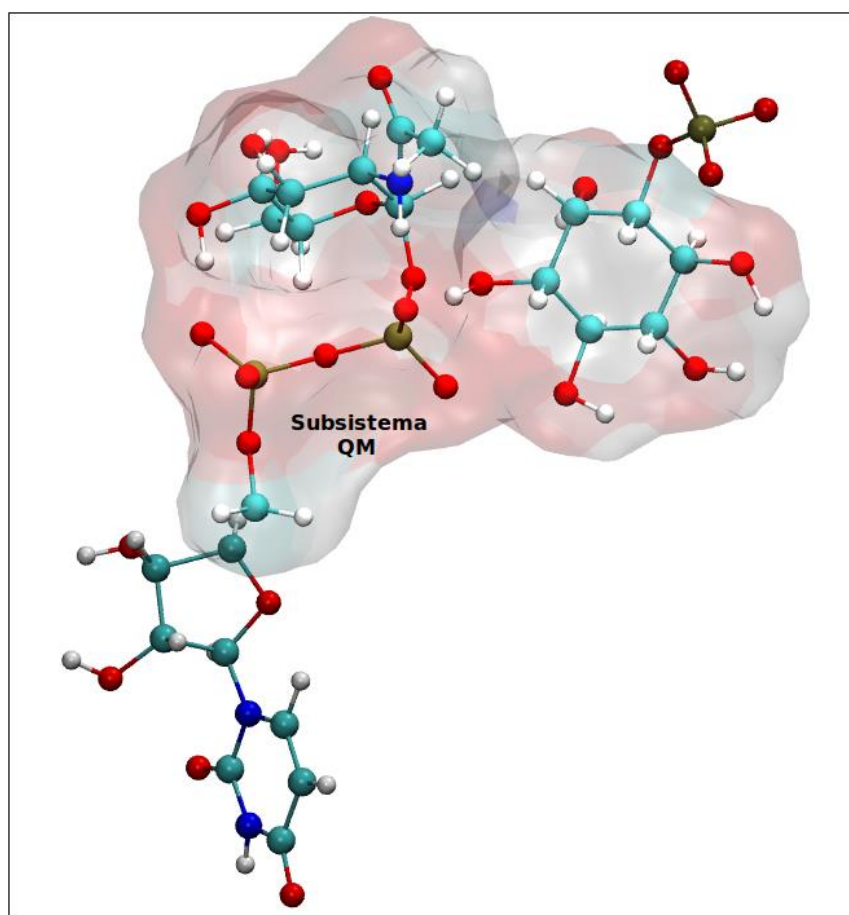


Figura 12. Esquema de los reactivos involucrados en la reacción. Encerrados en la superficie transparente se indican aquellos átomos de los reactivos que forman parte del Subsistema cuántico elegido.

v) Búsqueda del camino de Reacción. Umbrella Sampling

La función que verdaderamente describe la energía de un sistema es naturalmente una función multidimensional, que se conoce con el nombre de Superficie de Energía Libre (SEL) ^[12]. Conocer la SEL en su totalidad nos permitiría inferir a partir de ella todas las propiedades termodinámicas y cinéticas del sistema, pero ésta es una tarea computacionalmente imposible. Por este motivo, se desarrollaron métodos de exploración de la SEL que permiten obtener los datos termodinámicos y cinéticos de interés (mínimos de energía, barreras energéticas, estados de transición, etc) de un modo eficiente en términos de costo computacional. Uno de ellos, especialmente útil en el estudio de reacciones químicas como el que aborda esta Tesis, es la "**Búsqueda del camino de reacción**".

Mientras que las propiedades termodinámicas están relacionadas con los estados de mínima energía del sistema (en el caso del estudio de una reacción, los "reactivos" y "productos") y la energía relativa entre éstos, la cinética de un proceso está determinada por la altura de la "barrera de energía libre" que separa dichos mínimos, también llamada **energía de activación (ΔG^\ddagger)**. El camino de mínima energía de activación sobre la SEL que conecta estos dos mínimos es lo que se conoce como "camino de reacción". El camino de reacción de un proceso químico puede ser descrito mediante una coordenada simple relacionada con parámetros geométricos del centro reactivo, como ser las distancias interatómicas más

relevantes y/o el valor de un ángulo. A la combinación algebraica de todos los parámetros geométricos que permiten describir un proceso reactivo se la denomina "**coordenada de reacción**".

De la mecánica estadística se sabe que la **energía libre de Gibbs (G)** de un proceso esta relacionada a la **probabilidad** de acuerdo con la Ecuación de Boltzman (ecuación 1). Entonces, una aproximación sencilla para obtener la energía libre es muestrear todas las conformaciones del sistema a lo largo de la coordenada de reacción con Dinámica Molecular Canonica (NVT constantes) y calcular la probabilidad de cada conformación.

$$G(x) = -k_B T \ln P(x) \quad (\text{ec. 1})$$

Donde k_B es la 'Constante de Boltzman' y $P(x)$ es la 'probabilidad' de un evento o conformación X.

El problema es que la probabilidad $P(x)$ converge muy lento (o dicho con otras palabras, requiere un elevadísimo costo computacional alcanzar su convergencia) debido a las barreras energéticas naturales del sistema a lo largo del camino de reacción. Frente a este problema de muestreo es que surgen técnicas como el **Umbrella Sampling**, que permiten un muestreo más eficiente y parejo sesgando al sistema.

El Umbrella Sampling consiste en añadir restricciones artificiales de la forma de potenciales armónicos (con constantes de fuerza ajustables por el usuario) para sesgar al sistema a muestrear la coordenada de reacción elegida en torno a un cierto rango de valores de interés. Luego, repetidamente se mueve el mínimo de dicho potencial armónico a distintos valores de coordenada de reacción (en la dirección que esta tomaría cuando eventualmente se llegue al estado de "producto/s"), y se llevan a cabo simulaciones. Cada uno de estos minimos donde se coloca el potencial y donde se efectúa una simulación para muestrear el valor que adquiere la coordenada de reacción constituye una "ventana" o punto en el camino de reacción. Estas ventanas deben tener cierto "solapamiento", es decir, cada valor particular de la coordenada de reacción debe ser muestreada en más de una ventana. A cada respectiva ventana se le calcula un histograma (que representa la **probabilidad sesgada $P'(x)$** de cada valor de la coordenada), y posteriormente se transforma dicho histograma en un perfil Energía libre aplicando la siguiente ecuación:

$$G(x) = -k_B T \cdot \ln P'(x) - U'(x) + F \quad (\text{ec. 2})$$

Donde $U'(x)$ es el potencial armónico aplicado para el sesgo, y es igual a:

$$U'(x) = -k_u (x - x_0)^2$$

con " k_u " la constante de Fuerza utilizada, " x " es cada valor del histograma y " x_0 " es el mínimo donde se centro el potencial correspondiente a esa ventana.

Se obtienen entonces distintos fragmentos $G(x)$, cada uno correspondiente a una ventana de muestreo, que luego se pueden "empalmar" para construir una curva definitiva de "Energía Libre" en función de la "Coordenada de Reacción". En este sentido, cabe hacer notar la constante F de la ecuación 1 depende de cada ventana particular y se pierde al normalizar cada histograma, pero representa en términos prácticos cuanto se debe sumar al fragmento de $G(x)$ de cada ventana para que alcance su altura relativa en la curva definitiva, por lo que es fácilmente determinable.

vi) **Umbrella Sampling "bidimensional"**

Hasta ahora se describió como obtener el perfil de Energía libre de una reacción química utilizando una única coordenada de reacción y el método de muestreo 'Umbrella Sampling'. Pero, ¿que ocurriría si, en lugar de concentrar todas las distancias relevantes en una única coordenada, se utilizaran dos coordenadas de reacción? La respuesta obvia e inmediata es que se transformaría en un problema bidimensional más

complejo, de cuya resolución se obtendría ahora una 'superficie' de energía en lugar de un perfil unidimensional. La verdadera pregunta es, entonces, ¿es realmente ventajosa la complejización del problema, se obtienen los mismos resultados o es posible obtener información adicional? Esa es una de las preguntas clave que este trabajo de Tesis intento explorar.

La forma de llevar a cabo una superficie bi-dimensional es establecer un rango de valores de Reactivos a Productos para dos combinaciones de parámetros geométricos que definan dos coordenadas de reacción, y generar todos los pares posibles de simulaciones de Dinámica Molecular. Típicamente, se elije primero una coordenada de reacción y se llevan a cabo todas las simulaciones correspondientes, es decir, se confecciona la primera fila (o columna) de la superficie variando sólo una coordenada y dejando la otra constante en su valor inicial. A continuación, cada "punto" en esta primer fila generada se toma como "estructura inicial" y se llevan a cabo las simulaciones correspondientes variando la segunda coordenada (dejando la primer coordenada constante) para extender la superficie en la dimensión faltante. El resultado es una superficie conformada por 'N x M' simulaciones, siendo N la cantidad de puntos o ventanas contenidas en el rango que se hizo variar una coordenada y M la cantidad de puntos en que se hizo variar la otra coordenada.

De esto se desprende que se obtiene un total de 'N x M' ventanas con valores de muestreo, y su procesamiento para obtener la Superficie de Energía Libre es mucho mas complejo que el caso monodimensional, ya que determinar P(x) y F ahora consiste en resolver un sistema de ecuaciones, y requiere de un software específico. En este trabajo de tesis, tanto para el cálculo de los perfiles monodimensionales como de las superficies, se utilizó el método Weighted Histogram Analysis Method (WHAM) ^[24, 25], que resuelve las incógnitas por el método iterativo hasta alcanzar la autoconsistencia.

La constante de fuerza utilizada en todos los cálculos monodimensionales y bidimensionales de Umbrella Sampling fue siempre de 200 kcal/mol, salvo que se indique lo contrario. Este valor fue determinado como el más plausible a partir de numerosas pruebas con otros valores (50, 100, 300, 350, 500).

vii) Definición de las Coordenadas de Reacción

Todas las coordenadas de reacción (RC) que se utilizaron en este trabajo, tanto para los cálculos monodimensionales como para los bidimensionales, fueron definidas a partir de las 4 distancias indicadas en la figura 13:

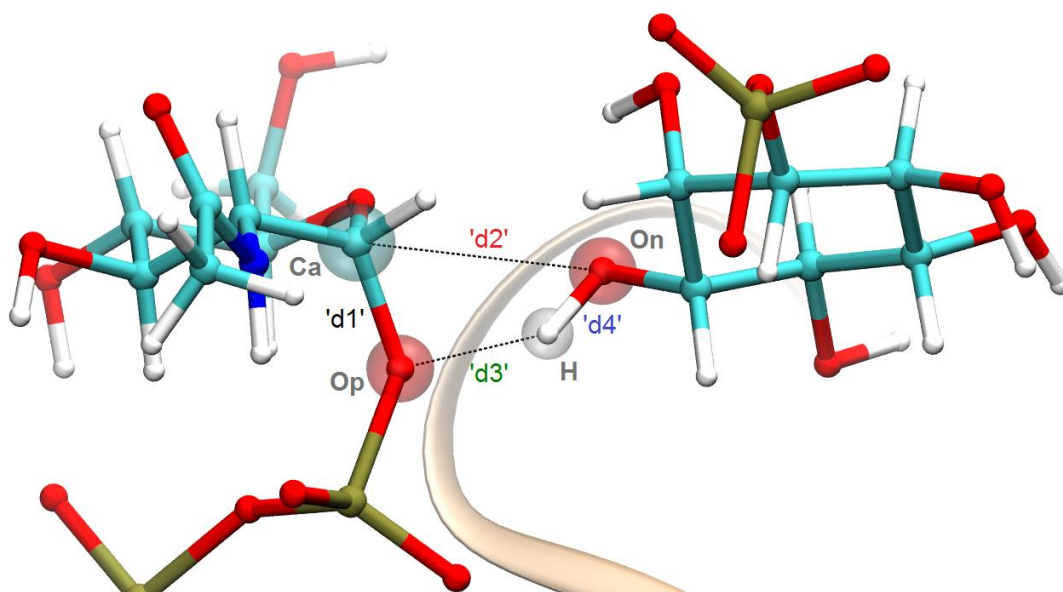


Figura 13. Centros atómicos involucrados en el mecanismo de reacción y distancias relevantes utilizadas para definir las distintas coordenadas de reacción. $d1 = d(O3B-C1)$; $d2 = d(C1-O4')$; $d3 = d(O3B-H4')$; $d4 = d(H4'-O4')$

Casos monodimensionales: $RC\ 1 = (d1 - d2 - d3 + d4)$

$RC\ 2 = (d1 - d2)$

'RC 1' es una combinación lineal de las 4 distancias que conectan los centros atómicos reactivos, y es una de las primeras opciones lógicas para poner a prueba. Los signos de cada distancia dentro de la coordenada se determinan de manera tal de asignar el mismo signo a todos los enlaces que se rompen (en este caso, d1 y d4), y el signo opuesto a todos los enlaces que se forman (d2 y d3).

El segundo caso, 'RC 2', resulta la coordenada mas simple con la que se podría describir el proceso global, por lo que es muy común observarla en los primeros trabajos de simulación computacional de glicosiltransferasas tanto de retención como de inversión, y aun en trabajos actuales se la sigue utilizando por su simpleza.

La longitud de los enlaces involucrados en la coordenada de reacción en la estructura de reactivos definieron los valores iniciales de cada coordenada de reacción. Para determinar los valores finales, se asumió que en el estado "productos", la distancia $d2=d(Ca-On)$ debiera ser igual a la distancia de equilibrio aproximada de un enlace glicosidico (~1.5 Ang) y la distancia $d3=d(Op-H)$ fuese igual a la distancia de equilibrio de un enlace oxígeno-hidrógeno (~1 Ang). Las distancias d1 y d4, que en el estado "productos" corresponden a enlaces que se rompen, se fijaron arbitrariamente. Toda esta información se detalla en la siguiente tabla:

<i>Estado</i>	<i>Distancia (Enlace)</i>				<i>Coordenada</i>	
	d1 (Ca - Op)	d2 (Ca - On)	d3 (H - Op)	d4 (H - On)	RC1	RC 2
Inicial (Reactivos)	1.43	3.02	2.36	0.97	-2.98	-1.59
Final (Productos)	3	1.5	1	2.5	3.02	2.01

Tabla 1. Definición de las coordenadas de reacción RC1 y RC2 para el Umbrella Sampling monodimensional. Todos los valores en Ångstrom.

Las "ventanas" de muestreo se efectuaron mediante incrementos de 0.1 Ångstrom para el caso de RC1, e incrementos de 0.1 Angstrom en el caso de RC2. Cada "ventana" consistió en una simulación QM-MM de 4 picosegundos (4000 pasos de integración a $dt=0.001$ ps) con un valor de coordenada de reacción fijo, utilizando siempre como estructura inicial la última estructura generada por la ventana anterior. El muestreo de la coordenada de reacción se realizó cada 4 pasos de integración, con lo cual cada ventana contó con 1000 datos de muestreo para el cálculo de probabilidades.

Caso bidimensional: $RC\ X = (d1 - d2)$; $RC\ Y = (d3 - d4)$

Definimos la coordenada 'RC X' (que se representara en todas las superficies en el eje de las abcisas) como la resta de las distancias de los enlaces que involucran al carbono anomérico 'Ca', y por otra parte la coordenada 'RC Y' (que sera en lo sucesivo representada en el eje de las ordenadas) que agrupa las distancias de los dos enlaces que involucran al átomo de hidrógeno. Es decir, mientras el progreso de la coordenada 'RC X' describe los enlaces involucrados en los eventos de 'hidrólisis y ataque nucleofílico', el progreso de la coordenada 'RC Y' describe la 'transferencia del protón'.

De un modo análogo al caso monodimensional, se determinaron los valores iniciales y finales de ambas coordenadas RCX y RCY según muestra la siguiente tabla:

<i>Estado</i>	<i>Distancia (Enlace)</i>				<i>Coordenada</i>	
	d1 (Ca -Op)	d2 (Ca - On)	d3 (H - Op)	d4 (H - On)	RCX= (d1-d2)	RCY= (d3-d4)
Inicial (Reactivos)	1.53	3.01	1.68	0.99	-1.48	0.69
Final (Productos)	3.5	1.5	1	2.05	1.96	-1.05

Tabla 2. Definición de las coordenadas de reacción RCX y RCY para el Umbrella Sampling bidimensional.
Todos los valores en Angstrom.

Como se puede apreciar en esta tabla, la estructura de partida utilizada es distinta a la utilizada en los cálculos monodimensionales, principalmente presenta el enlace "d3" más cortos. La razón es que al tener que realizar N x M simulaciones (las "ventanas" de muestreo se efectuaron mediante incrementos de 0.8 Ångstrom para ambas coordenadas, generando un N=44 y M=23), el costo computacional se incrementa considerablemente con cada simulación, por lo que se optó por buscar una estructura inicial donde el protón estuviese en una orientación mas favorable para su transferencia y mas cercano a Op (que se refleja en un menor "d3"). De este modo, se reduce la cantidad de simulaciones necesarias para RCY.

Cada "ventana" consistió en una simulación QM(SCC-DFTB)-MM de 4.5 picosegundos (4500 pasos de integración a $dt=0.001$ ps) con un valor de coordenada de reacción fijo, utilizando siempre como estructura inicial la última estructura generada por la ventana anterior. El muestreo de la coordenada de reacción se realizó cada 4 pasos de integración, con lo cual cada ventana contó con 4.500 datos de muestreo. Las constantes de fuerza fueron de 200 kcal/mol para ambas dimensiones.

El procedimiento para construir la SEL definitiva consistió en primero construir una "superficie preliminar" de estructuras "equilibradas" mediante Simulaciones QM(SCC-DFTB)-MM cortas de 0.5 ps, y una vez esta fuese chequeada y depurada, extender cada punto de la superficie en el tiempo unos 4 ps mas para realizar el muestreo propiamente dicho.

La "superficie preliminar" se construyo del siguiente modo: se extendió en primer lugar la coordenada RCY (dejando RCX siempre fija), es decir, primero permitiendo que el protón se transfiriera. Una vez confeccionada esta primer columna, con el protón en distintos "estados" de transferencia, cada una de estas ventanas sirvió como punto de partida para extender la superficie en el sentido horizontal (dejando RCY fija y extendiendo RCX) para avanzar con el muestro de la hidrólisis y el ataque nucleofílico. En aquellos casos en que una determinada ventana "XY" tuviese problemas de convergencia o el cálculo no llegara a término, el cálculo era re-lanzado tomando como estructura inicial una estructura vecina de una ventana adyacente.

RESULTADOS

i) Cálculos de Docking

Para lograr obtener la estructura del "**complejo ternario definitivo**", en una primera instancia se intentó dockear el azúcar donador N-acetil-glucosamina directamente sobre el modelo del "complejo ternario incompleto" (que contenía solo UDP e Inositol-fosfato). Como resultado de 100 corridas de Docking, se obtuvieron 3 grupos o 'clusters' de conformaciones posibles, siendo el cluster mejor posicionado en el ranking aquel que mas se asemejaba al modelo de orientación "bent back" propuesto por Vetting y colaboradores, con una población del 42% y que se corresponde con el de menor energía de unión. Los resultados se presentan en la tabla a continuación, y los dos mejores se ilustran en la figura 14.

Cluster	Menor Energía de Binding	Energía de Binding Media	Población
1	-4.29	-4.26	42
2	-4.27	-4.25	34
3	-4.25	-4.24	24

Tabla 3. Resultados del Bias docking de la N-acetil-glucosamina sobre el "complejo ternario incompleto".

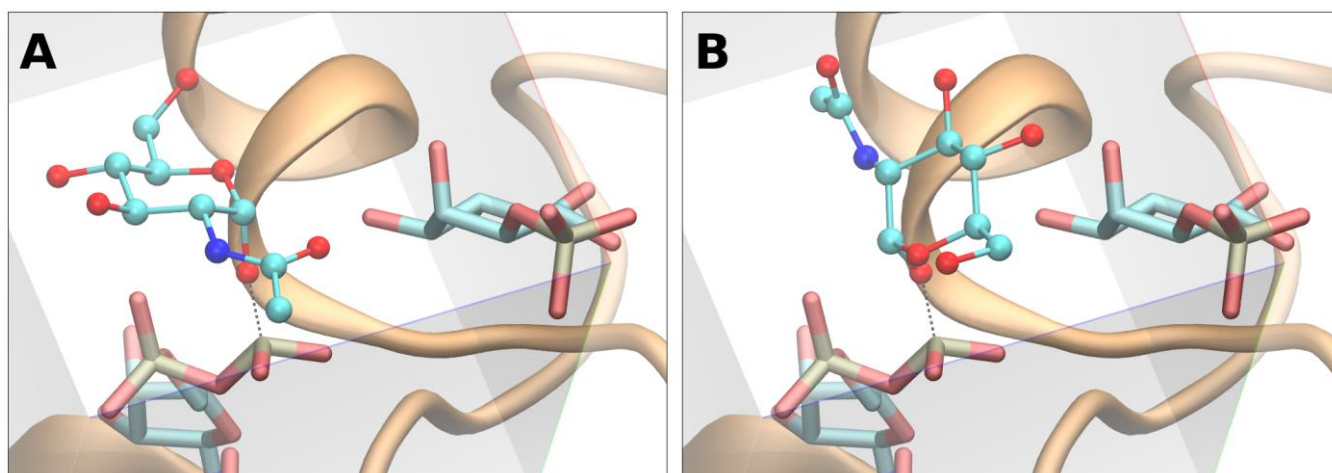


Figura 14. Resultados del cálculo de Docking de la N-acetil-glucosamina sobre el modelo del complejo ternario incompleto. **A.** Conformación mejor posicionada en el ranking poblacional, que reproduce la orientación bent back y orienta al carbono anomérico en una posición favorable para la reacción. **B** Segunda conformación en el ranking (tercer conformación es similar), en la que la molécula se orienta en el sentido inverso y poco favorable para la reacción. El prisma transparente ilustra a modo de referencia la posición y las dimensiones de la grilla utilizada en el cálculo de Docking. Los átomos de hidrógeno se omiten para mayor claridad.

En segunda instancia, se decidió evaluar si era posible dockear la N-acetil-glucosamina en la conformación “bent-back” a partir de un complejo con UDP e Inositol-fosfato previamente sometido a una simulación de Dinámica Molecular, para evaluar si el sitio activo sufría algún tipo de deformación tras la simulación que modificara significativamente su geometría y en consecuencia privilegiara algún otro tipo de conformación alternativa del nucleótido que no fuera la obtenida hasta entonces. De este modo, El complejo cerrado con UDP e inositol fosfato fue sometido a una Dinámica Molecular corta (5 nanosegundos), y al observar las trayectorias generadas, se detectó la siguiente situación: los oxígenos cargados negativamente de los grupo pirofosfato del UDP provocaban que a los pocos picosegundos de simulación los grupos aminos de los residuos cercanos colapsaran en torno a esta región, reduciendo por tanto el espacio físico para que poder luego dockear la N-acetil-glucosamina (situación ilustrada en la figura 15).

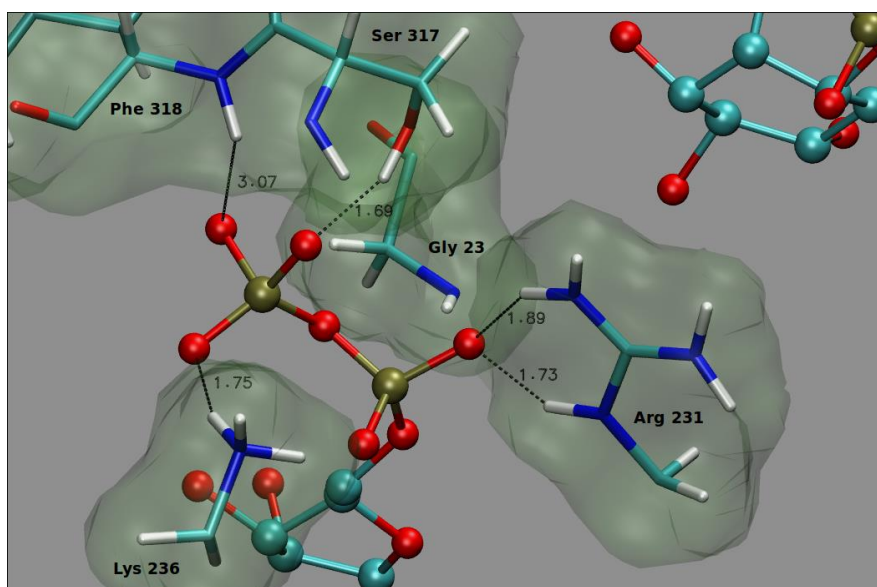


Figura 15. Grupos aminos de los residuos Gly 23, Arg 231, Lys 236, Ser 317 y Phe 318 colapsados en torno al grupo pirofosfato del UDP, ocupando parte del bolsillo que debería ocupar la N-acetil-glucosamina.

Es por eso que se realizó una nueva simulación de Dinámica Molecular, pero esta vez, aplicando una constante de fuerza de 100 Kcal/mol que limitara la movilidad de los residuos mencionados durante las etapas de Termalización y Equilibrio de la Densidad, para permitir a los fosfatos "solvatarse" adecuadamente antes de comenzar con la etapa de Producción. Durante la etapa de Producción, dicha constante de fuerza fue retirada, y ya no se observó la situación de colapso. Se tomó entonces una foto al azar de entre todas trayectorias generadas, y se realizó el cálculo de bias-docking sobre la misma. El resultado de 200 corridas de bias-docking sobre esta nueva foto permitió reproducir la orientación "bent-back" con una población del 91 % y la menor energía de unión. Quedó entonces claro que, dada la estructura particular del azúcar y la geometría del bolsillo proteico, esta estructura es la más plausible y por lo tanto un adecuado punto de partida para el estudio del mecanismo de reacción.

ii) Umbrella Sampling monodimensional

Una vez obtenida la estructura del complejo ternario definitivo, la misma fue solvatada en una caja de aguas octaédrica, luego termalizada y equilibrada su presión de acuerdo con el protocolo descrito en la sección anterior, y a partir de este tratamiento se seleccionó una estructura adecuada como estructura inicial, que sería el punto de partida para los cálculos de QM(SCC-DFTB)-MM Umbrella Sampling. Los resultados y la performance de cada cálculo/método será luego analizada y comparada a la luz de la barrera energética de la reacción (ΔG^\ddagger) obtenida, y el mecanismo será interpretado en base a la evolución de las distancias involucradas en la/s coordenada/s de reacción.

En primer lugar, la utilización de la coordenada de reacción 'RC 1' no logró resultados satisfactorios con ninguno de los métodos semiempíricos empleados, a pesar de numerosas y muy diversas pruebas. Concretamente, las ventanas individuales del protocolo de Umbrella Sampling no permitían un empalme adecuado para construir el perfil de energía libre resultante. Si bien las simulaciones en la mayoría de los casos llegaban hasta el final y se lograba transferir el protón y formar el enlace glicosídico alcanzando el estado de "producto", fue el muestreo el que demostró ser en algún punto ineficiente (las razones se discutirán mas adelante). Por otra parte, la utilización de la RC 2 permitió obtener los primeros resultados positivos concretos. El muestreo de la coordenada de reacción RC 2 para cada una de las ventanas de simulación en función del tiempo se muestran en el gráfico de la figura 16:

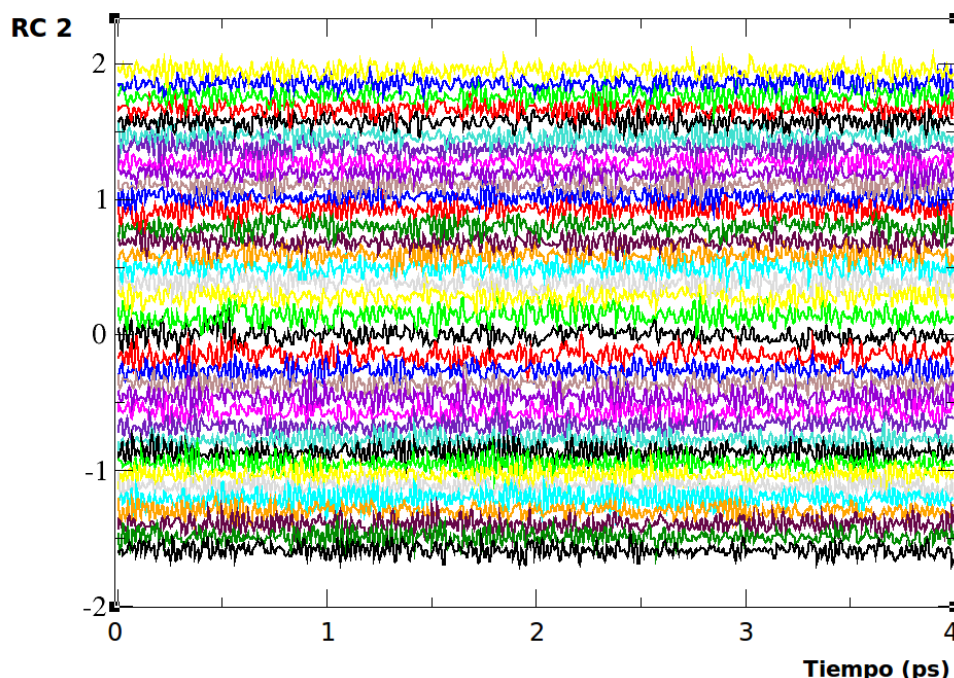


Figura 16. Valores muestreados de la coordenada RC 2 en función del tiempo. Cada color representa una ventana. Puede apreciarse que el grado de solapamiento entre las ventanas es adecuado.

Los datos del muestreo fueron posteriormente transformados en un perfil energía libre con el programa WHAM. EL perfil arroja una barrera de activación (ΔG^\ddagger) de aproximadamente **56.6 kcal/mol**. En la figura 17 se muestra el perfil de energía libre, junto con la evolución ventana a ventana de cada una de las 4 distancias que definen la coordenada de reacción.

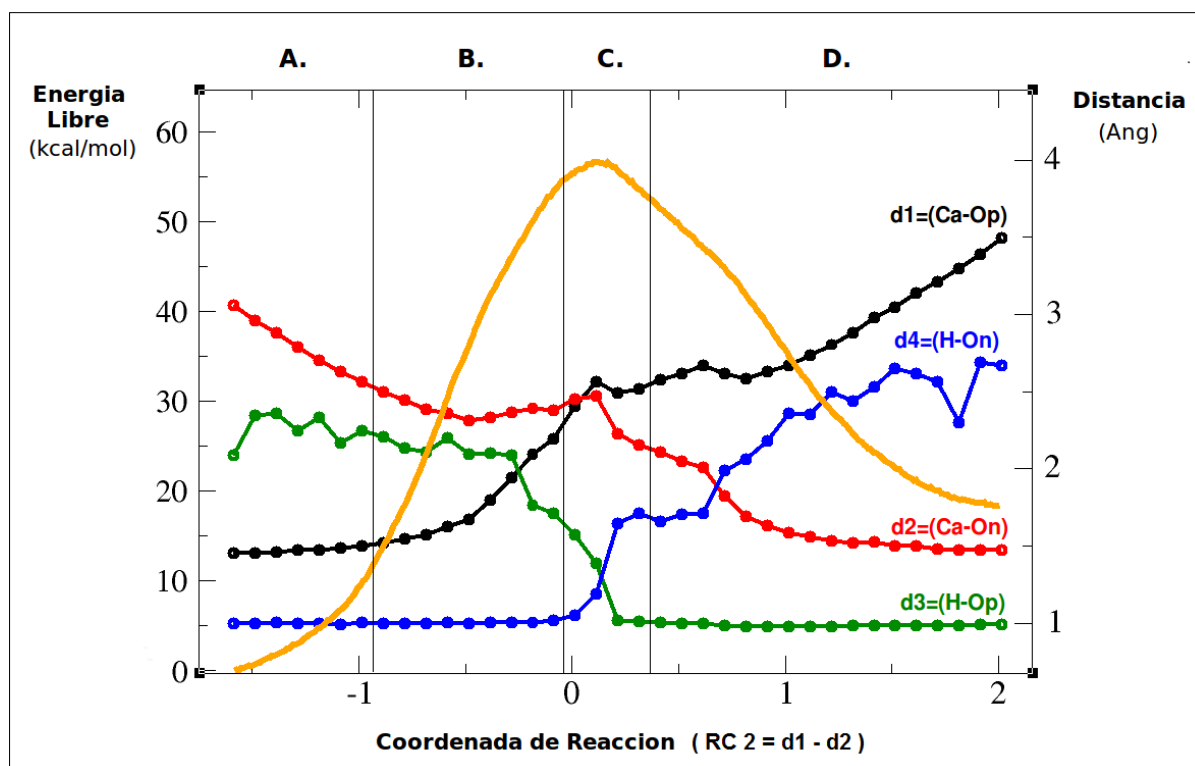


Figura 17. Perfil de energía libre (naranja), y fluctuación de las 4 distancias relevantes en función de la Coordenada de Reacción RC 2. Cada valor de distancia corresponde al promedio de 100 valores extraídos de 100 fotos equiespaciadas de la trayectoria correspondiente. ($\Delta G^\ddagger = 56.6$ Kcal/mol; $\Delta G = 18.5$ kcal/mol).

Un análisis detallado del gráfico permite distinguir 4 zonas o etapas:

A. En una primer instancia, el Oxígeno nucleofílico del aceptor se acerca al centro anomérico del donador, reflejado en una disminución de "d2", mientras las otras 3 distancias permanecen relativamente constantes. El aumento en la energía libre es gradual y moderado.

B. El acercamiento paulatino del nucleófilo al carbono, que aún se encuentra considerablemente lejos, ($d2 \sim 2.5$ Ang) comienza a inducir el estiramiento progresivo del enlace del nucleótido-azúcar donador, reflejado en el incremento de "d1", lo cual produce un crecimiento abrupto de la energía. El protón comienza a acercarse al oxígeno del UDP ("d3" comienza a disminuir).

C. Esta etapa corresponde a la formación de la **zona de transición**. Al comienzo, la hidrólisis se completa ($d1 > 2.5$ Ang) y el protón se acerca rápidamente al pirofosfato del nucleótido ("d3" decae súbitamente). Cuando el sistema finalmente alcanza el punto máximo en la energía (el "estado de transición"), el protón se encuentra en el medio del proceso de transferencia, equidistante entre ambos oxígenos ($d3=d4$).

D. La energía libre desciende, debido por un lado a la formación del enlace glicosídico ("d2" progresivamente disminuye hasta alcanzar un valor de ~ 1.5 Ang, que es la longitud de equilibrio del

enlace glicosídico), y por otro lado, al grupo saliente alejándose del centro reactivo (reflejado por el crecimiento de d1 y d4).

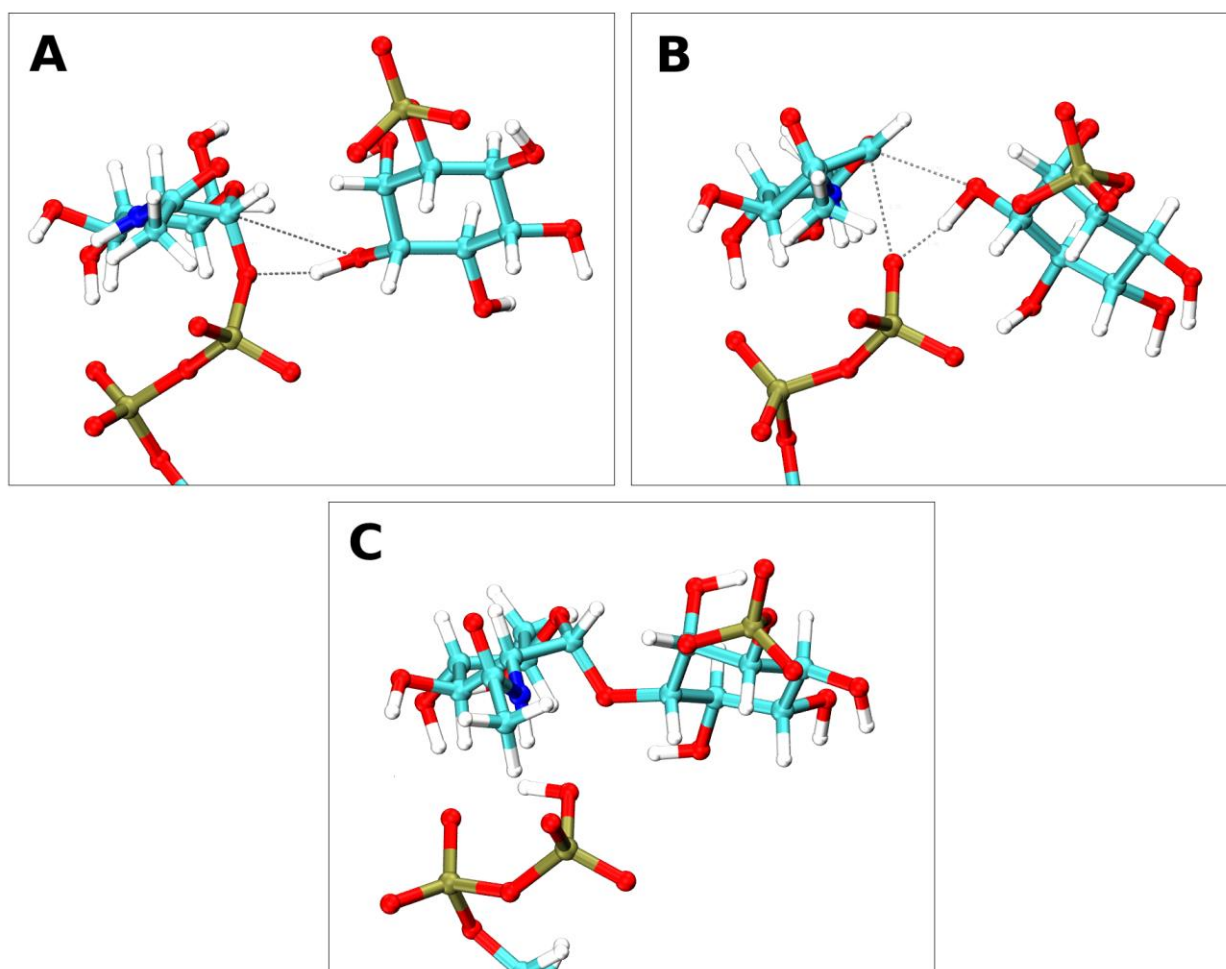


Figura 18. Estructuras del cálculo SCC-DFTB monodimensional para: **A)** Reactivos; **B)** Estado de Transición; **C)** Productos.

La barrera energética obtenida resulto mucho mas elevada de lo esperado. Datos experimentales en otras glicosiltransferasas y valores calculados *in silico* en otros trabajos sugieren valores entre 10 - 25 kcal/mol. Por este motivo, y para determinar si la elevada barrera se debía a un problema de muestreo o de definición del sistema reactivo, se optó por efectuar el mismo cálculo, es decir, utilizando la misma estructura de partida y el mismo método ("SCC-DFTB"), pero en 3 situaciones diferentes alterando distintos parámetros de la simulación (resultados en la figura 19):

- 1) Aumentando el tiempo de muestreo 1 orden de magnitud (40 ps en lugar de 4 ps), para asegurar que el tiempo de muestreo fue suficiente.
- 2) Aumentando la temperatura electrónica al triple, para ver si mejoraba aún más la convergencia.
- 3) Incorporando al subsistema QM las cadenas laterales de la Lisina 236 y Arginina 231, que interaccionan con el el pirofosfato del UDP (se ha hipotitizado un rol 'estabilizador del grupo saliente' este tipo de residuos en otras glicosiltranferasas). Por ser aminoácidos con carga positiva, aportan 2 unidades de carga positiva al subsistema QM, por lo tanto la carga cuántica total del sistema pasa a ser 0 en lugar de -2.

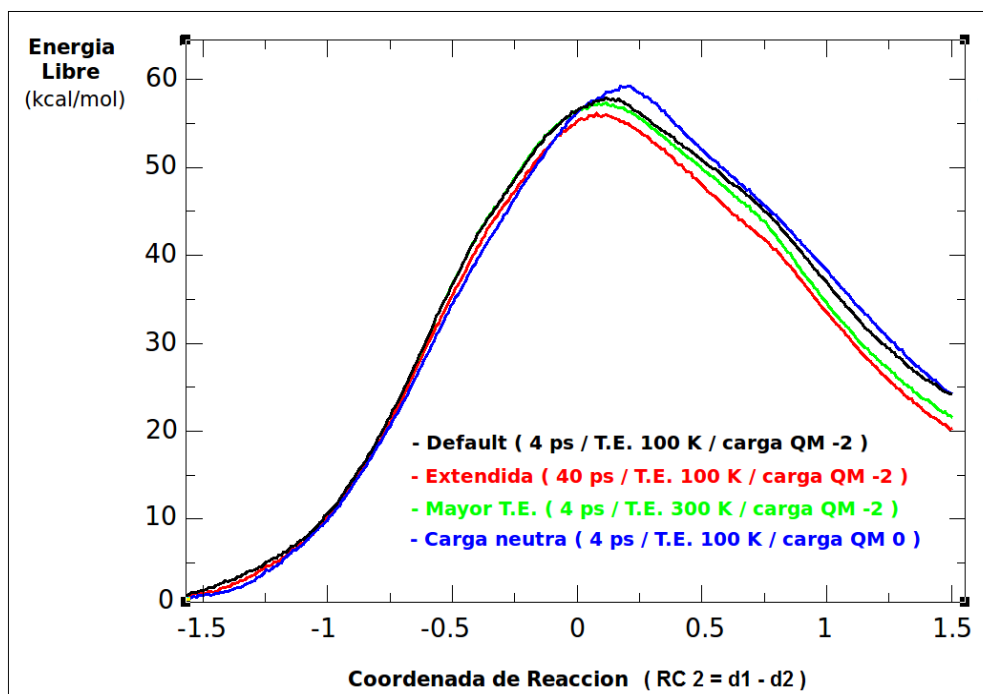


Figura 19. Perfiles de energía libre de diferentes cálculos con el método SCC-DFTB, utilizando distintos parámetros indicados entre paréntesis (tiempo muestreado en picosegundos, "Temperatura Electrónica", y carga total de la región QM).

Como se puede deducir de la figura 19, los parámetros alterados no aportaron cambios de mayor significancia en los resultados finales: La curva de energía conserva la misma forma, y apenas se puede observar una diferencia de 4 kcal/mol entre los máximos más extremos. La inclusión de las cadenas laterales de los residuos de Arginina y Lisina interactuantes con el UDP en la región QM, al contrario de lo que se esperaba, arrojaron una barrera mayor ($\sim +2$ kcal/mol respecto a la curva inicial), lo cual es indicio que el subsistema QM elegido inicialmente es adecuado y agregar más residuos a la parte cuántica no tiene sentido, ya que sólo aumentaría el tiempo de computo. Por otra parte, como resultado satisfactorio si puede notarse que el tiempo de muestreo de 4 ps empleado hasta ahora es adecuado (y computacionalmente eficiente), ya que muestrear un tiempo 10 veces mayor no arroja una mejoría sustancial en la altura de la barrera (~ -2 kcal/mol respecto a la curva 'Default').

Para determinar si la magnitud de la barrera obtenida era debido a un problema en la descripción del sistema por el nivel de teoría cuántica propio del método SCC-DFTB, se probó a continuación repetir el cálculo con todos los demás métodos semiempíricos disponibles en el paquete AMBER (que incluyen AM1, AM1d, RM1, PM3, PM6, MNDO y MNDOd), y las **correcciones de tercer orden para transferencia de protones** implementadas en el paquete AMBER para SCC-DFTB^[30] (denotadas en la bibliografía con el acrónimo 'PA', por 'Proton Affinty'). Todos los métodos semiempíricos a excepción de SCC-DFTB-PA fallaron en reproducir la reacción satisfactoriamente: en la mayoría de los casos la estructura de los reactivos se deformaba, o bien se formaban enlaces incorrectos, o directamente ciertos átomos de hidrógeno no involucrados en la reacción rompían sus enlaces y salían despedidos del subsistema QM, provocando el colapso de los cálculos. Esto demuestra que casi la gran mayoría de métodos semiempíricos presentan dificultades para describir este tipo de reacciones, en tanto que SCC-DFTB parece ser de las pocas excepciones.

Por su parte, el cálculo SCC-DFTB-PA arrojó una barrera de **40.6 kcal/mol** (figura 20), esto es 16 kcal/mol por debajo de lo calculado por el método SCC-DFTB estándar.

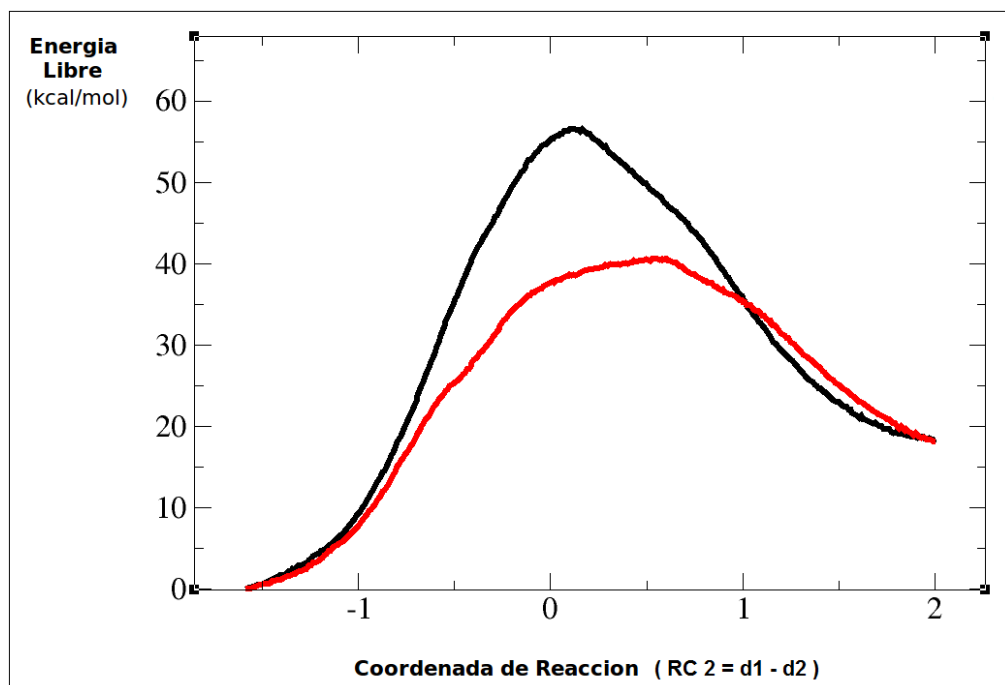


Figura 20. Perfil de energía libre con los métodos QM empleados. **Negro:** SCC-DFTB, **Rojo:** SCC-DFTB-PA.
($\Delta G^\ddagger = 40.6$ kcal/mol; $\Delta G = 18.5$ kcal/mol)

La curva de energía de SCC-DFTB-PA parece tener una forma mas achatada, en un principio crece con menos pendiente, y posee el máximo desplazado hacia la derecha. Pese a estas diferencias, el orden de todos los eventos implicados en el mecanismo de reacción y las 4 etapas anteriormente descriptas para el cálculo SCC-DFTB inicial continúan siendo los mismos. También es importante destacar que pese a la diferencia entre las barreras, la diferencia de energía libre entre productos y reactivos (el ΔG de reacción) es el mismo en ambos métodos.

Posiblemente el cambio en la barrera este asociado a un aumento en la afinidad del oxígeno del fosfato por el protón del nucleófilo, ya que esta interacción es la que se desencadena para compensar la ruptura del enlace oxígeno-carbono anomérico (d1) al comienzo de la reacción.

iii) Umbrella Sampling bidimensional

No conforme con los resultados del Umbrella Sampling convencional, se opto por poner a prueba un enfoque distinto: el Umbrella Sampling Bidimensional. Con este método, se buscó explorar más a fondo y desde una perspectiva distinta el sistema en estudio, manteniendo la misma estructura de partida y los mismos métodos cuánticos (SCC-DFTB y SCC-DFTB-PA).

Una Vez obtenida una "superficie preliminar" y chequeada la correcta finalización de cada una de las 1012 estructuras que la componían, se procedió a extender cada simulación un total de 4 ps para efectuar el muestreo propiamente dicho y así determinar la SEL definitiva. Esto se realizó por duplicado, empleando los dos métodos utilizados exitosamente en los cálculos monodimensionales: SCC-DFTB y SCC-DFTB-PA. Se obtuvieron así las dos superficies que se muestran en la figuras 21 y 22, respectivamente.

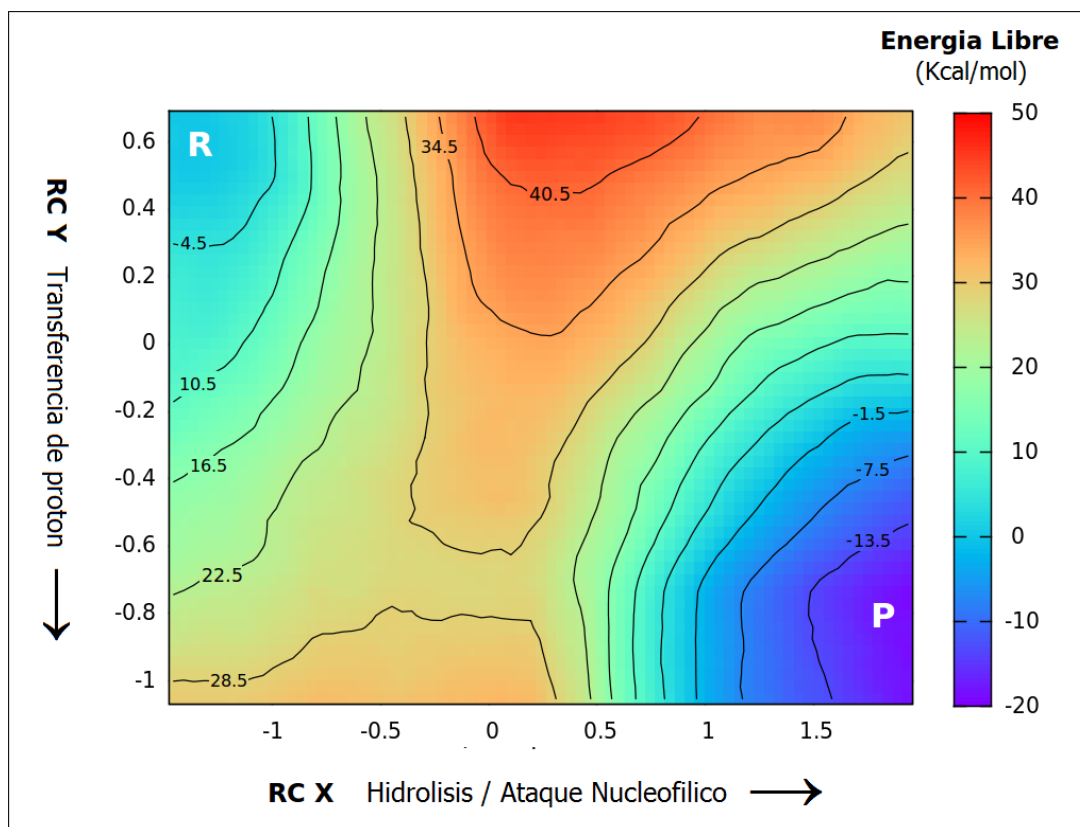


Figura 21. Superficie de Energía Libre obtenida al nivel SCC-DFTB. ($\Delta G^\ddagger = 28$ kcal/mol; $\Delta G = -18$ kcal/mol).
Cada línea de contorno equivale a 6 kcal/mol.

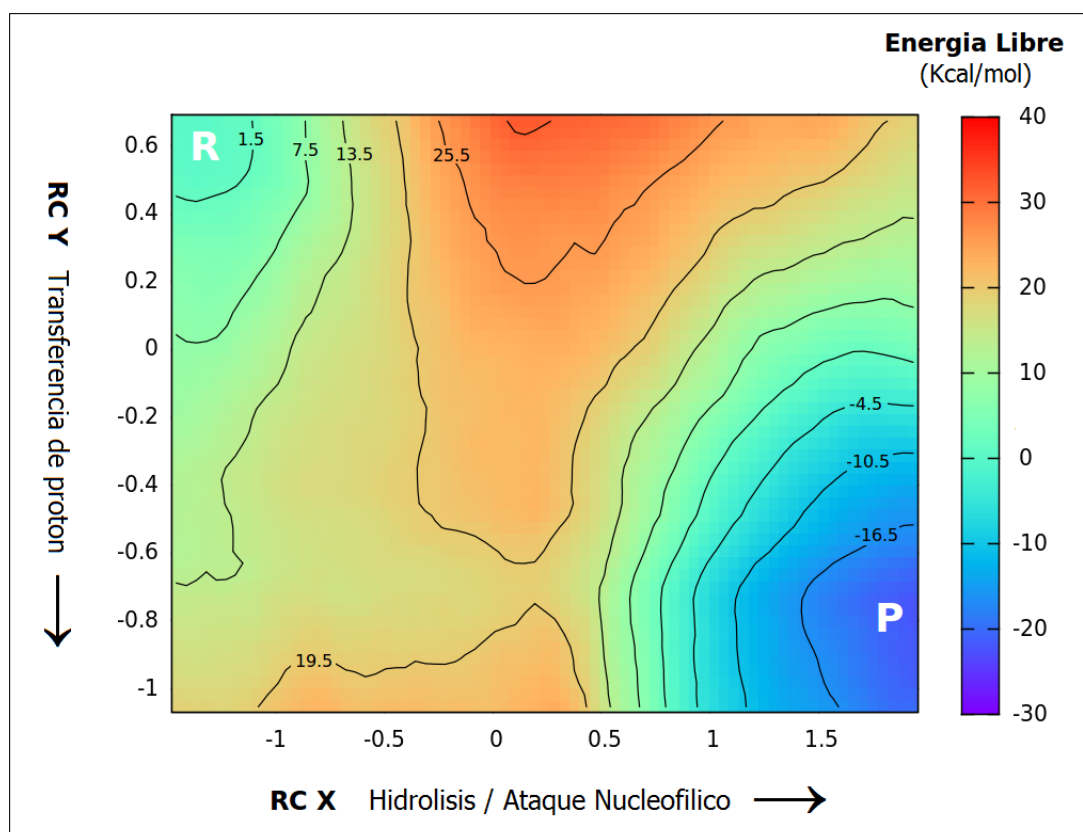


Figura 22. Superficie de Energía Libre obtenida al nivel SCC-DFTB-PA. ($\Delta G^\ddagger = 19$ Kcal/mol; $\Delta G = -23$ kcal/mol).
Cada línea de contorno equivale a 6 kcal/mol.

Los 3 escenarios posibles para el mecanismo quedan entonces expuestos a partir de estas superficies energéticas. Mientras que cualquier camino que una reactivos y productos trazado por la diagonal estaría apuntando a un proceso concertado, en el que la transferencia de proton y la hidrólisis/ataque nucleofílico se dan de manera simultánea, aquellos caminos que pasen por las esquinas describen el proceso por etapas (figura 23). Por un lado, un camino que pase por la esquina superior deja explícito el costo de que primero ocurra la hidrólisis y el ataque nucleofílico, y el protón se transfiera en último lugar. Contrariamente, cualquier camino que pase por la esquina inferior sugiere una transferencia del protón que antecede a la hidrólisis, lo cual contrasta con los resultados obtenidos en los cálculos monodimensionales (en los que la hidrólisis sucedía en primer lugar para dar luego paso a una abrupta transferencia de protón en el punto máximo de la curva).

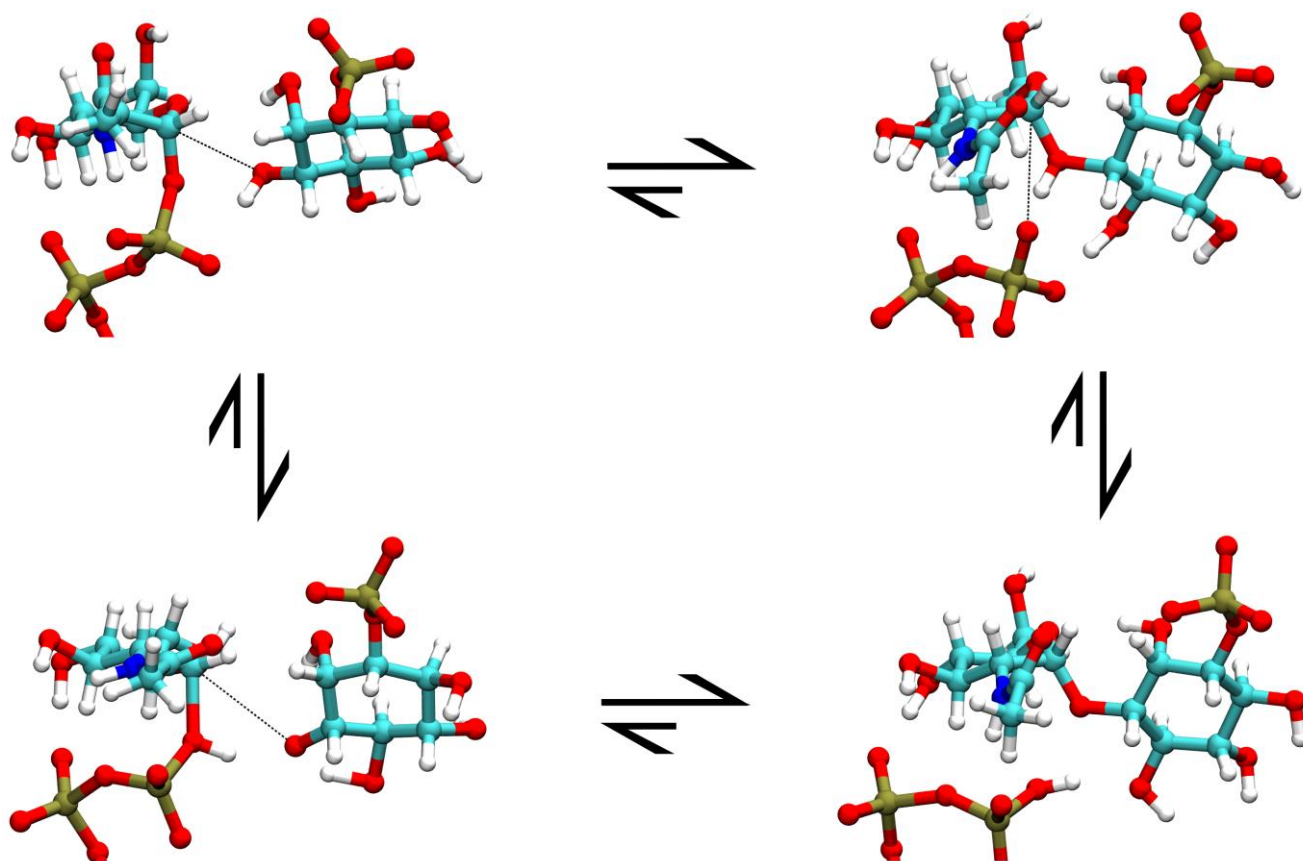


Figura 23. Representación de las estructuras correspondientes a las 4 esquinas de la Superficie de energía libre

Se observa que para ambas superficies en las figuras 18 y 19, el camino energéticamente mas favorable es aquel que bordea la esquina inferior. En la figura 24 a continuación se muestra, para la superficie SCC-DFTB, el camino de menor energía y la fluctuación de las 4 distancias a lo largo de dicho camino, y el gráfico respectivo para SCC-DFTB-PA se muestra en la figura 25.

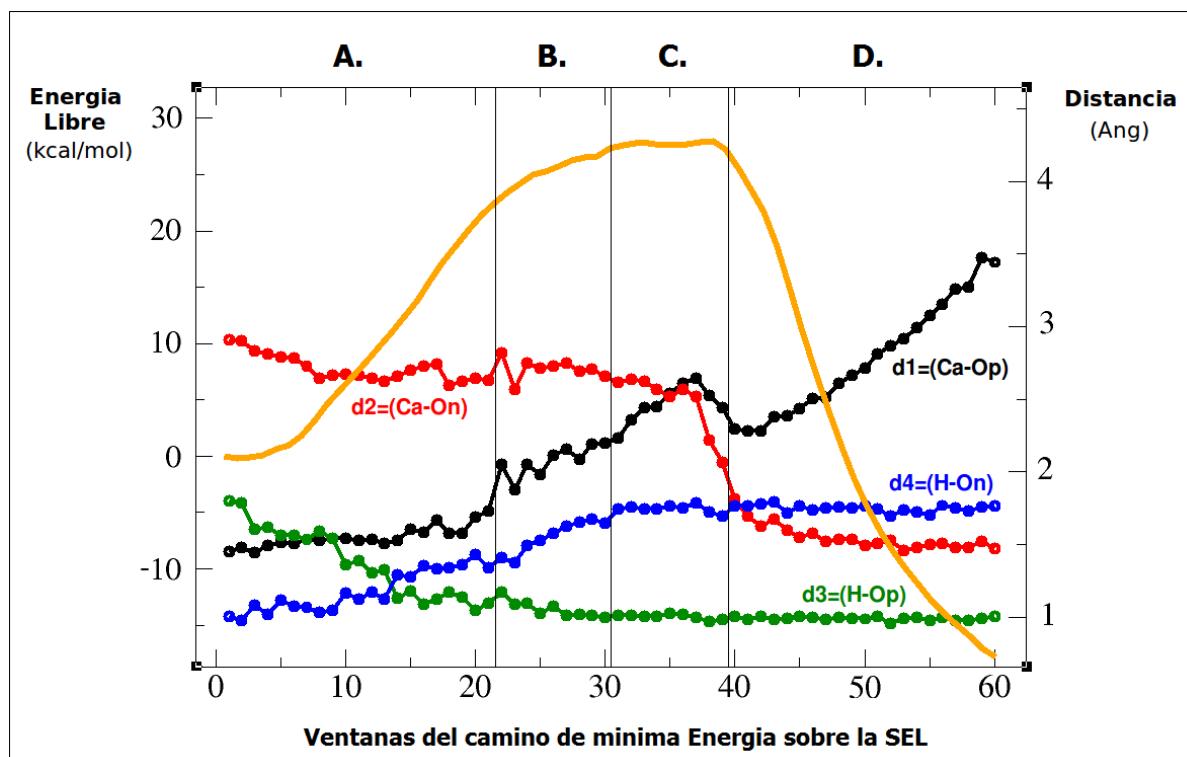


Figura 24. Perfil de energía libre y la fluctuación de las 4 distancias relevantes para el camino de menor energía de la superficie bidimensional SCC-DFTB presentada en la figura 21. Cada valor de distancia corresponde al promedio de 20 valores extraídos de 20 fotos equiespaciadas de cada trayectoria correspondiente.

Nuevamente, a los fines prácticos se puede analizar la evolución del perfil en 4 etapas, que por otra parte contrastan en gran medida con los cálculos monodimensionales:

A. En la primer etapa, las distancias que vinculan el carbono anomérico con los oxígenos saliente y nucleofílico ("d1" y "d2", respectivamente) varían escasamente. El protón se transfiere en esta etapa, de una manera muy gradual, y es este proceso en mayor medida el responsable de la crecida en la energía libre.

B. El aumento gradual de "d1" desde 2 a 2.5 Ångstrom es indicio que se comienza a producir la hidrólisis. Por otra parte, el grupo saliente comienza a alejarse mas marcadamente (d4 aumenta). El resto de las distancias se mantienen relativamente constantes. La energía libre, si bien aun no alcanza su valor máximo, muestra un marcado desaceleramiento en su crecimiento, debido a que el nucleófilo se encuentra muy lejos aún del sitio anomérico (no se observan cambios importantes en d2).

C. La energía alcanza su punto máximo, y se mantiene constante. La hidrólisis se completa, como muestra el crecimiento en "d1". Por otra parte, el oxígeno nucleofílico, que en las dos etapas anteriores sufrió un acercamiento muy gradual y paulatino, sobre el final de esta etapa se acerca rápidamente al carbono anomérico hasta un valor de aproximadamente 2 Ångstrom como muestra "d2". Por último, "d3" y "d4" alcanzan su máxima separación y se mantienen constantes, lo que indica que el protón finalmente se estabiliza.

D. La energía libre desciende abruptamente, impulsada por la formación del enlace glicosídico (d1 ~ 1.5 Å) y el creciente alejamiento del grupo saliente.

Por otra parte, el camino de menor energía de la superficie SCC-DFTB-PA presenta una curva ligeramente diferente (figura 25), en la que la región análoga a las "etapas B y C" de la figura 24 presenta variaciones de energía muy pequeñas, creando una extensa zona plana en la que se vuelve dificultoso

identificar un claro estado de transición único. De todas maneras, los eventos ocurren en el mismo orden que en el caso SCC-DFTB, por lo que los mecanismos obtenidos son los mismos.

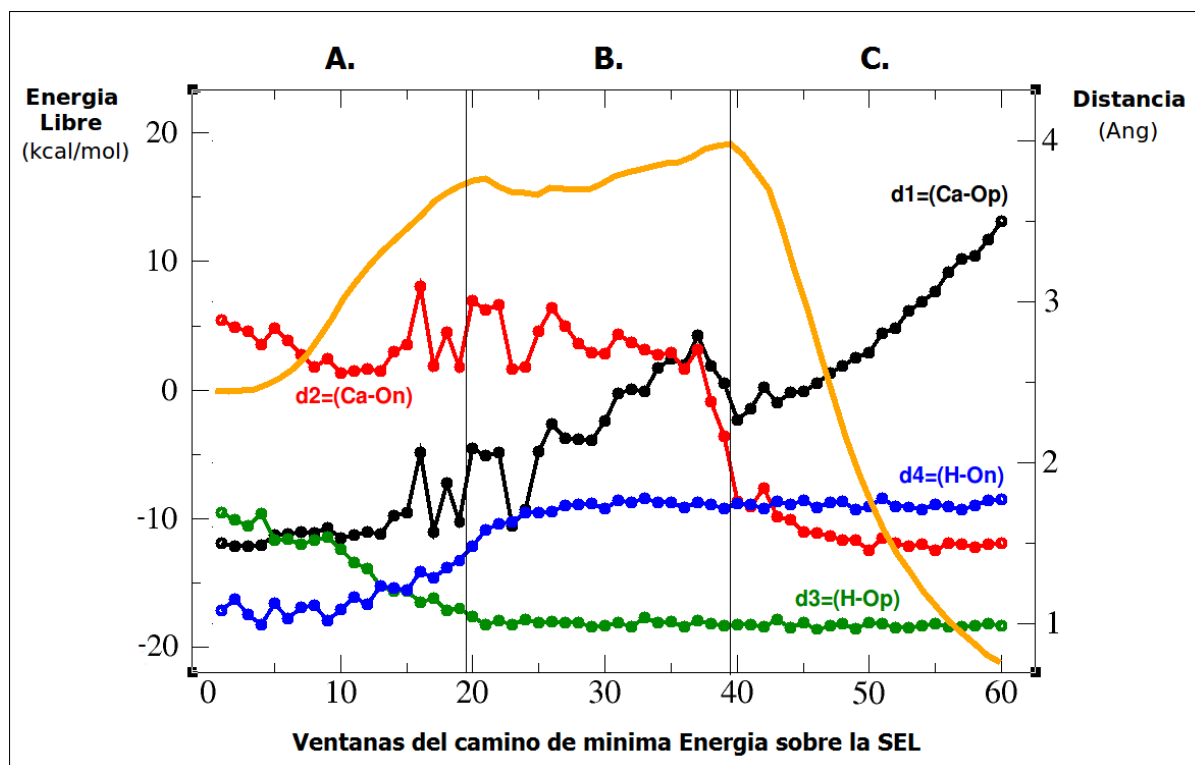


Figura 25. Perfil de energía libre y la fluctuación de las 4 distancias relevantes para el camino de menor energía de la superficie bidimensional SCC-DFTB presentada en la figura 22. Cada valor de distancia corresponde al promedio de 20 valores extraídos de 20 fotos equiespaciadas de cada trayectoria correspondiente.

Por último, para el caso SCC-DFTB bidimensional, se muestran a continuación la evolución a lo largo del camino de reacción de la distribución de carga electrónica (poblaciones de Mulliken) de cada uno de las moléculas o fragmentos individuales incorporados en la región del subsistema cuántico (figura 26) y de los principales centros atómicos involucrados relevantes en la reacción (figura 27), y una tabla que resume toda la información relevante de distancias y carga electrónica para caracterizar los 3 estados relevantes del proceso (Tabla 4).

	Reactivos	E. de Transición	Productos
d1 = d(Op-Ca)	1.46 (0.04)	2.67 (0.09)	3.39 (0.05)
d2 = d(Ca-On)	2.92 (0.06)	2.55 (0.08)	1.49 (0.04)
d3 = d(Op-H)	1.73 (0.11)	1.00 (0.04)	0.99 (0.02)
d4 = d(H-On)	1.09 (0.12)	1.76 (0.05)	1.76 (0.04)
d(O5-Ca)	1.46 (0.05)	1.29 (0.03)	1.46 (0.03)
q O5	-0.41 (0.02)	-0.13 (0.03)	-0.41 (0.01)
q Ca	0.40 (0.02)	0.10 (0.05)	0.38 (0.02)
q HCa	0.03 (0.015)	0.25 (0.03)	0.04 (0.01)
q Op	-0.47 (0.03)	-0.58 (0.02)	-0.62 (0.02)
q On	-0.54 (0.03)	-0.72 (0.03)	-0.40 (0.02)

Tabla 4. Distancias relevantes involucradas en la reacción y cargas de Mulliken de los átomos involucrados (Desvío estándar entre paréntesis). Todas las distancias en Ångstrom. El átomo "O5" representa el oxígeno del anillo del azúcar N-GlcNAc.

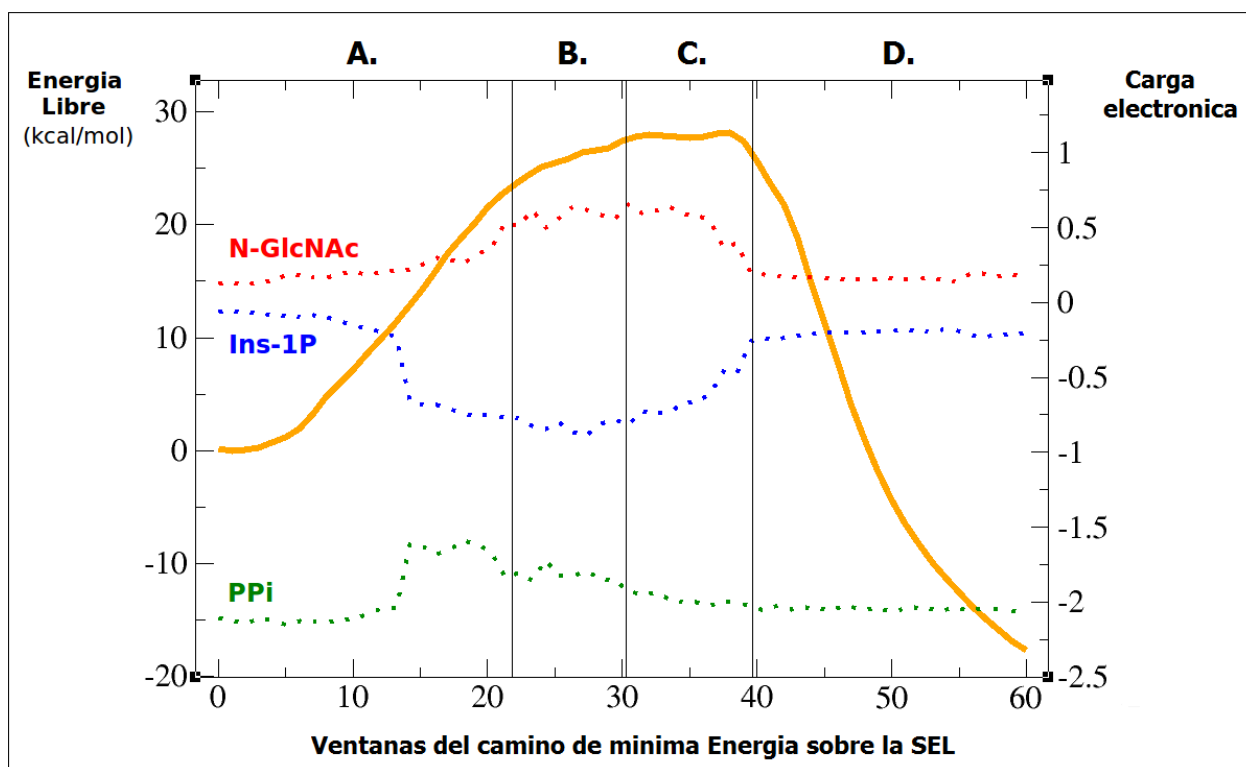


Figura 26. Distribución de la carga electrónica (poblaciones de Mulliken) en cada punto del camino de reacción para el grupo pirofosfato, el azúcar donador N-Acetil-glucosamina (N-GlcNAc) y la molécula aceptora 1-Inositol-fosfato.

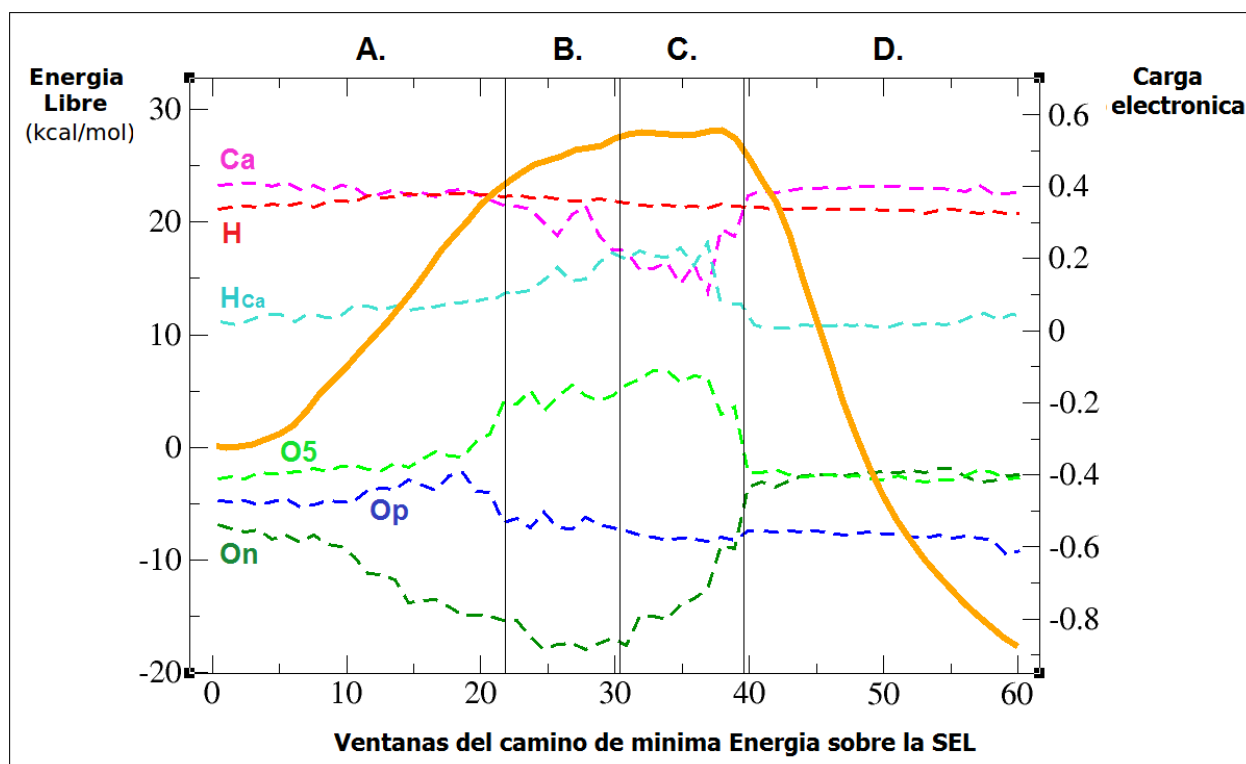


Figura 27. Distribución de la carga electrónica (poblaciones de Mulliken) en cada punto del camino de reacción, para los principales centros atómicos relevantes en la descripción del proceso Reactivo. (Ca = carbono anomérico; HCa = Hidrógeno unido a Ca ; H = protón transferido; O5 = oxígeno del anillo del azúcar donador; Op = Oxígeno del pirofosfato; On = oxígeno nucleofílico del aceptor)

En la figura 26 se observa que la estructura electrónica de los entes moleculares incluidos en la región cuántica se comporta correctamente de acuerdo a lo observado. En las zonas de transición (B y C) la carga en la molécula donora aumenta (de 0.15 a 0.6, aprox.) consistente con la formación de un carbocatión, en tanto que en la molécula aceptora disminuye (de -0.1 a -0.9, aprox.), consistente con la desprotonación del nucleófilo y el alejamiento del protón. Esta situación se revierte en la etapa final una vez alcanzado el estado de los productos. La carga del grupo pirofosfato, por su parte, se vuelve ligeramente mas positiva durante el proceso de hidrólisis y transferencia, hasta estabilizarse en torno a un valor constante a fines de la zona de transición C.

Por otra parte, la figura 27 permite un análisis más detallado y a nivel atómico de lo que sucede con la distribución de cargas. Mientras que la carga del protón transferido (H) y del oxígeno del enlace hidrolizado (Op) se mantiene relativamente constante, el resto de los átomos muestran importantes fluctuaciones, las cuales en principio pueden ser explicadas de manera simple a la luz de lo que ocurre con el oxígeno del anillo del azúcar donador (O5): en la zona de transición, al formarse el carbocatión se genera una acumulación de carga positiva en torno al centro anomérico. Es aquí donde el par libre de electrones de O5 entra en juego, y se "deslocaliza" para compensar dicho fenómeno, provocando que el carbono anomérico (Ca) se vuelva más negativo y la carga positiva de la especie carbocatiónica se concentre marcadamente en el hidrógeno alifático unido a Ca (HCa). La evidencia de esta deslocalización se ve sustentada tanto en la figura 27 (donde se aprecia que O5 se vuelve más "positivo" en la zona de transición), como así también en el hecho de que en toda la zona de transición, el enlace Ca-O5 se acorta (Tabla 4) casi 0.2 Ångstrom respecto de su valor de equilibrio tanto en Reactivos como Productos, lo cual es consistente con el fenómeno de deslocalización electrónica. Una vez alcanzado los productos, tanto las cargas de O5 como de Ca y HCa vuelven a los valores iniciales. Por último, se observa que la carga del oxígeno nucleofílico On se vuelve más negativa al principio de la reacción, lo cual es consistente con el proceso de desprotonación teniendo lugar como primer evento del mecanismo, y que luego de formarse el enlace glicosídico (a fines de la etapa C) retoma su valor inicial.

En síntesis, para los dos métodos semiempíricos utilizados SCC-DFTB y SCC-DFTB-PA, las barreras de activación a lo largo del camino de menor energía obtenidas a partir de los cálculos bidimensionales (**28 kcal/mol** y **19 kcal/mol** respectivamente) son más consistentes con lo esperado. Por otra parte, el orden de eventos que describen el mecanismo son marcadamente diferentes que en los cálculos monodimensionales, así como también la forma de las curvas. Finalmente, el analisis correspondiente de la distribución de cargas a nivel molecular y atómico es consistente con todo lo descripto.

CAPITULO IV: DISCUSIÓN

Mecanismo de Reacción

El objetivo de la presente tesis se centró en caracterizar a escala atómica el mecanismo de reacción de la glicosiltransferasa de retención *MshA*, mediante técnicas de simulación computacional, concretamente, QM-MM Umbrella Sampling mono y bidimensional.

El uso de distintas coordenadas de reacción, distintos métodos semiempíricos para tratar la región QM, y por último el empleo de distintas variantes de la técnica Umbrella Sampling, derivó en un amplio abanico de resultados que, si bien en algunos casos ligeramente diferentes en cuanto al orden puntual de los eventos reactivos, todos concuerdan en que se trata de un mecanismo de **"Desplazamiento simple frontal"** (SNI).

Los resultados obtenidos con el Umbrella bidimensional en el mejor de los casos muestran que la reacción en su totalidad tendría con el nivel de teoría SCC-DFTB-PA una energía de activación de **19 kcal/mol**. Estos resultados se pueden analizar a la luz de datos experimentales reportados para la enzima *MshA*. Utilizando la definición fenomenológica que asocia la constante catalítica ' K_{cat} ' con la energía de activación, la misma puede ser aproximada mediante la ecuación:

$$\Delta G^\ddagger = -RT \ln (h \cdot K_{cat}/k_B T)$$

Donde R es la constante de gases ideales, h es la Constante de Dirac y k_B es la constante de Boltzman.

Vetting y colaboradores reportaron para *MshA* una constante catalítica de $12.5 \pm 0,2 \text{ s}^{-1}$ (a $T = 298 \text{ K}$), lo cual nos permite estimar un $\Delta G^\ddagger \approx 17 \text{ kcal/mol}$, lo que muestra la excelente estimación del método en el caso bidimensional SCC-DFTB con correcciones de tercer orden para afinidad de protones.

En consecuencia, se puede concluir que el mecanismo de reacción más plausible es aquel descrito por el cálculo de Umbrella Sampling bidimensional al nivel SCC-DFTB-PA, y el orden de eventos observado en dicho mecanismo se puede puntualizar de la siguiente manera:

- A. Transferencia lenta y gradual del protón.
- B. Debilitamiento y posterior ruptura del enlace entre el nucleótido y el azúcar donador, inducida por el protón transferido.
- C. Leve y gradual acercamiento del nucleófilo desprotonado, en contraposición a un marcado y rápido alejamiento del grupo saliente.
- D. Formación abrupta del enlace glicosídico.

Es importante destacar que, si bien ambas superficies de energía libre exhiben barreras de energía que difieren en $\sim 10 \text{ kcal/mol}$, el camino de reacción descrito por ambas es el mismo. La diferencia de energía se podría atribuir, en principio, a un muestreo mas preciso en favor de SCC-DFTB-PA, sobretodo en aquellos enlaces y geometrías conteniendo hidrogeno, en los cuales actuó la corrección de tercer orden 'Proton Affinities'.

Umbrella Sampling monodimensional y Bidimensional

En primer lugar, la utilización de la coordenada de reacción RC 1, que representaba la combinación algebraica de las 4 distancias relevantes en una única coordenada, mostró ser ineficiente. Esto fue comprobado ya que, al emplear el WHAM para obtener los perfiles de energía, se obtuvieron curvas con forma irregular y barreras de activación inesperablemente altas, lo cual fue el primer indicio de un posible problema en el muestreo. Esto fue luego corroborado al construir los histogramas "a mano" de un grupo representativo de ventanas; se observó que los mismos eran irregulares (con mínimos en el centro y con los valores máximos distribuidos hacia los extremos). Debido al empleo exitoso de la coordenada RC 2, se podría en principio hipotetizar que la razón principal de este muestreo defectuoso en RC 1 se debe a la

inclusión del protón en la coordenada de reacción; muy probablemente estas características en el muestreo estuviesen relacionadas a la naturaleza "lábil" propia de los protones y su baja masa, que tras aplicarle una constante de fuerza del orden de la requerida para forzar al resto de los enlaces responden con un comportamiento caótico.

En segundo lugar, la utilización de la coordenada RC 2 permitió obtener los primeros resultados contundentes, ya que la misma excluye al protón del potencial sesgado y deja que el mismo se mueva y se oriente con libertad. La transferencia ocurre de todos modos, sin necesidad de forzar el enlace On-H ("d4"). No obstante, se podría argumentar que el uso de RC 2 es insuficiente, ya que describe el proceso sólo a partir de dos enlaces, no teniendo en cuenta (al menos no de forma directa) lo que sucede con el protón en la energética del proceso. Mas aun, la energía de activación (ΔG^\ddagger) sobreestimaba demasiado los valores experimentales reportados para otras glicosiltransferasas y los valores teóricos calculados *in silico* en trabajos similares.

Analizando con detenimiento el camino de menor energía que conecta Reactivos y Productos a lo largo de las Superficies de Energía Libre, y comparando estos con los Perfiles de Energía monodimensional obtenidos en primera instancia, quedan en evidencia dos mecanismos diferentes:

- 1) El mecanismo "bidimensional" apunta hacia una temprana transferencia de protón, acompañada de una hidrólisis gradual. El estado de transición, en lugar de exhibir un máximo puntual, corresponde a una zona extensa de la curva donde la energía permanece relativamente constante (y esta observación es aún mas pronunciada en SCC-DFTB-PA), y se caracteriza por el azúcar donador en estado de carbocatión, el nucleófilo desprotonado, y el grupo saliente protonado.
- 2) El mecanismo "monodimensional" prioriza la hidrólisis como primer evento, seguida de un acercamiento gradual y paulatino del protón. El "estado de transición", entendido como el punto de maxima energía sobre la curva, se alcanza cuando el carbono anomérico se encuentra aproximadamente equidistante del oxígeno saliente y el oxígeno nucleofílico, y a su vez el protón se encuentra equidistante entre ambos oxígenos.

A pesar de las diferencias entre ambos mecanismos, se puede apreciar una similitud: el ataque nucleofílico se produce al final y es el responsable de que la energía descienda. Esto estaría indicando que el nucleófilo no es el responsable directo de la hidrólisis, es decir, queda descartado que la ruptura del enlace Op-Ca sea desencadenada o estimulada por el ataque de On sobre Ca.

Como complemento, el análisis de las poblaciones de Mulliken en uno de los casos bidimensionales permitió avanzar en la comprensión del mecanismo de reacción, y a su vez permite hacer ciertas conjeturas sobre el proceso de inhibición. Como se menciona en la introducción de este trabajo, tanto *MshA* como otras glicosiltransferasas se inhiben competitivamente en presencia de un azúcar donador fluorado (es decir, con un grupo fluor reemplazando un grupo oxhidrilo, generalmente el unido al carbono 2 del azúcar, o en el caso puntual del azúcar GlcNAc, el 5 ya que la posición 2 esta ocupada por el grupo N-acetilo). Teniendo en cuenta este dato con lo observado en estos cálculos, puede hipotetizarse que la inhibición podría estar relacionada a la marcada naturaleza atractora del electrones del flúor; Este sustituyente podría estar dificultando el fenómeno de deslocalización de los electrones del oxígeno del anillo (O5), atrayéndolos para sí, y en consecuencia generando un estado de transición altamente inestable. Como la estabilidad del estado de transición esta directamente relacionada con la altura de la barrera de la reacción, podría ser esta entonces una explicación plausible para este fenómeno experimental.

Conclusiones Finales

Queda no obstante esperando una respuesta definitiva uno de los interrogantes iniciales: "cuál es la fuerza impulsora del mecanismo?" En este sentido, en la literatura hasta el momento hay solo dos trabajos publicados sobre mecanismo de reacción S_Ni en glicosiltransferasas de retención, el primero de ellos sobre la enzima *OtSA* de la familia GT-B^[27], que utiliza un enfoque bidimensional de simulaciones QM/MM basadas en el método de "metadinamica" de Laio et al^[33] (este método utiliza Dinámica Molecular de Car

Parinello basada en la Teoría del Funcional de la Densidad para la parte QM). El segundo es sobre la enzima *LgtC* de la familia GT-A^[28], que utiliza un enfoque monodimensional de Umbrella Sampling al nivel SCC-DFTB, como así también un enfoque monodimensional de la técnica de "Optimización Restringida" a nivel SCC-DFTB y a otros niveles QM mucho más complejos y costosos computacionalmente (BP86, B3LYP y M05-2X). Los resultados obtenidos por ambos trabajos, si bien utilizan enfoques y técnicas muy deferentes, son consistentes entre si, y sorpresivamente guardan cierta similitud con los resultados del caso monodimensional para *MshA* aquí presentados. Quizá la mayor diferencia sea una transferencia de protón mas tardía (especialmente en el caso de *OtSA*, en la que la transferencia total del protón es el último evento en ocurrir, casi en simultáneo la formación del enlace glicosídico). Por otra parte, en particular en el trabajo sobre *LgtC*, niveles crecientes de teoría parecen no alterar el mecanismo en lo que al orden de eventos respecta, lo cual, del mismo modo que en este trabajo, demuestra que el nivel de teoría QM utilizado tiene en realidad mas impacto en el cálculo de las energías relativas y barreras de reacción que en proceso mecanístico en si. Esto deja en claro que, si bien los métodos semiempíricos como SCC-DFTB se encuentran entre los de mas bajo nivel de teoría, en la practica continúan siendo muy útiles (por su bajo costo computacional) como primeras aproximaciones en el estudio de reacciones químicas complejas en sistemas bioquímicos, y queda demostrado que su potencial no debe ser subestimado.

Como conclusión final, puede afirmarse que se logró reproducir satisfactoriamente *in silico* la reacción de transferencia catalizada por la enzima *MshA*, y se logró caracterizar el mecanismo de reacción SNi por distintos enfoques, poniendo a prueba una gran cantidad de técnicas y métodos computacionales.

REFERENCIAS BIBLIOGRAFICAS

1. **Gerald L. Newton, Nancy Buchmeier, and Robert C. Fahey***. Biosynthesis and Functions of Mycothiol, the Unique Protective Thiol of *Actinobacteria*. *Microbiology and Molecular Biology Reviews*, Sept. 2008
2. **Fan Fan, Matthew W. Vetting, Patrick A. Frantom, and John S. Blanchard**. Structures and Mechanisms of the Mycothiol Biosynthetic Enzymes. *Curr Opin Chem Biol*. 2009.
3. **Newton GL, Ta P, Bzymek KP, Fahey RC**. Biochemistry of the initial steps of mycothiol biosynthesis. *J Biol Chem* 2006;281:33910–33920. [PubMed: 16940050]
4. **Newton GL, Unson MD, Anderberg SJ, Aguilera JA, Oh NN, delCardayre SB, Av-Gay Y, Fahey RC**. Characterization of *Mycobacterium smegmatis* mutants defective in 1-d-myo-inosityl-2-amino-2-deoxy- α -d-glucopyranoside and mycothiol biosynthesis. [PubMed: 10049692]
5. **Patrick A. Frantom, James K. Coward, and John S. Blanchard**. UDP-(5F)-GlcNAc Acts as a Slow-Binding Inhibitor of MshA, a Retaining Glycosyltransferase
6. **Vetting MW, Frantom PA, Blanchard JS**. Structural and enzymatic analysis of MshA from *Corynebacterium glutamicum*: substrate-assisted catalysis. *J Biol Chem* 2008;283:15834–15844. [PubMed: 18390549].
7. **Sinnott, M. L.; Jencks, W. P.** *J. Am. Chem. Soc.* 1980, 102, 2026– 2032.
8. Research Collaboratory for Structural Bioinformatics, Universidad de Rutgers, New Brunswick, NJ (<http://www.rcsb.org/>)
9. **Warshel A, Levitt M**. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol*. 1976; 103:227.
10. **W. Humphrey, A. Dalke, K. Schulten**, *J. Mol. Graph.* 1996, 14, 33 – 38.
11. **Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J.** (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J Comput Chem* 30, 2785-2791.
12. **Leach A.R.**, Molecular Modeling. Eds; Addison Wesley Longman, 1997.
13. **Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA** (1995). "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". *J. Am. Chem. Soc.* **117**: 5179–5197
14. **Verlet L**. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules *Phys. Rev.* 1967, 159, 98.
15. **Berendsen H.J.C, Postma J.P.M, van Gunsteren W, DiNola A., Haak J.R**, Molecular dynamics with coupling to an external bath. *J. Chem, Phys.* 1984, 81, 3684.
16. **Cheatham, T. E., III, P. Cieplak, and P. A. Kollman**. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* 16:845–862.
17. **Kirschner, K.N., Yongye, A.B., Tschampel, S.M., Daniels, C.R., Foley, B.L, Woods, R.J.** *J. Comput. Chem.* 2008, 29, pg 622-655
18. **W. L. Jorgensen and J. D. Madura**, *J. Am. Chem. Soc.* 1983. 105, 1407.
19. **Gao J**. Methods and applications of combined quantum mechanical and molecular mechanical Potentials. *Reviews in Computational Chemistry*; Lipkowitz K.B, Boyd D. B, Eds. VCH Publishers: New York, **1996**; Vol. 7.
20. **Monard G, Merz K. M. Jr.**, Combined Quantum Mechanical/Molecular Mechanical Methodologies Applied to Biomolecular Systems *Acc. Chem. Res.* **1999**, 32, 904.

21. **R.C. Walker, M.F. Crowley and D.A. Case, J. COMP. CHEM.** 29:1019, (2008)
22. **Seabra, G.M., Walker, R.C. et al., J. PHYS. CHEM. A,** 111, 5655, (2007)
23. **Valleau, J.P.; Torrie, G.M.** en *Modern Theoretical Chemistry, Vol. 5: Statistical Mechanics, Part A*, Berne, B.J., Ed. Plenum Press, New York, 1977.
24. **Kumar, S.; Bouzida, D.; Swendsen, R.H.; Kollman, P.A.; Rosenberg, J.M.** The weighted histogram analysis method for free-energy calculations on biomolecules. *J. Comput. Chem.*, 1992, 13, 1011–1021.
25. **Kumar, S.; Rosenberg, J.M.; Bouzida, D.; Swendsen, R.H.; Kollman, P.A.** Multidimensional free-energy calculations using the weighted histogram analysis method. *J. Comput. Chem.*, 1995, 16, 1339–1350.
26. **Tvaroska I.** Molecular modeling insights into the catalytic mechanism of the retaining galactosyltransferase LgtC. *Carbohydr Res.* 2004 Apr 2; 339(5):1007-14.
27. **Ardèvol A, Rovira C.** The molecular mechanism of enzymatic glycosyl transfer with retention of configuration: evidence for a short-lived oxocarbenium-like species. *Angew Chem Int Ed Engl.* 2011 Nov 11;50(46):10897-901.
28. **Hansel Gómez, Iakov Polyak, Walter Thiel, JoséM. Lluch, and Laura Masgrau** Retaining Glycosyltransferase Mechanism Studied by QM/MM Methods: Lipopolysaccharyl- α -1,4-galactosyltransferase C Transfers α -Galactose via an Oxocarbenium Ion-like Transition State. *J. Am. Chem. Soc.*
29. **Seung Seo Lee, Sung You Hong, James C Errey, Atsushi Izumi, Gideon J Davies & Benjamin G Davis.** Mechanistic evidence for a front-side, S_Ni-type reaction in a retaining glycosyltransferase. *Nature Chemical Biology*, Vol 7 | Sep 2011
30. **Yang, Y., Yu, H., York, D., Elstner, M., and Cui, Q.** Description of phosphate hydrolysis reactions with the Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB) theory. 1. Parameterization., *J. Chem. Theory Comput.* 4, 2067–2084. (2008)
31. **Tvaroska I.** Structural Insights into the Catalytic Mechanism and Transition State of Glycosyltransferases Using ab initio Molecular Modeling. *Trends Glycosci. Glycotechnol.* 2005, 17, 177-190.
32. **Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C.** Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat. Struct. Biol.* 2001, 8, 166–175.
33. **A. Laio, J. VandeVondele and U. Rothlisberger, J. Chem. Phys.** 2002, 116, 6941–6947.
34. **Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G.** Self Consistent Charge density functional tight binding method for simulation of complex materials properties. *Phys Rev B.* 1998; 58 : 7260–7268.

INDICE DE FIGURAS

Figura	Descripción	Pág.
1	Via de Biosíntesis del Micotiol.	2
2	Representación de los sustratos de 'MshA'.	3
3	Esquema de reacción con inversión o retención de la configuración isométrica.	4
4	Estructura tridimensional de 'MshA' y cambio conformacional	5
5	Esquema de los mecanismos de reacción propuestos para las Glicosiltransferasas de retención.	6
6	Mecanismos de reacción propuesto para 'MshA'.	6
7	Estructura primaria y secundaria de 'MshA'	8
8	Estructura terciaria del modelo sin gaps de 'MshA'.	9
9	Conformación 'bent-back' del UDP-azúcar en glicosiltransferasas	10
10	Función de Scoring y campo de Fuerzas de Autodock	11
11	Campo de fuerzas de AMBER	12
12	Esquema del conjunto de átomos que integran el subsistema QM	14
13	Distancias relevantes para la definición de las coordenadas de reacción	17
14	Resultados del cálculo de Docking de la N-acetil-glucosamina	19
15	Colapso de residuos en torno a la región de Docking en la MD del complejo ternario incompleto.	20
16	Valores muestreados de la coordenada RC 2 en función del tiempo.	
17	Perfil de energía libre y fluctuación de distancias relevantes en función de la coordenada de reacción 'RC 2' para el calculo de Umbrella Sampling monodimensional a nivel SCC-DFTB	
18	Estructura de Reactivos, Estado de Transición y Productos para el cálculo de Umbrella Sampling monodimensional a nivel SCC-DFTB	
19	Perfiles de energía libre de cálculos monodimensionales a nivel SCC-DFTB con diferentes parámetros de simulación.	
20	Perfiles de energía libre de cálculos monodimensionales para los métodos SCC-DFTB y SCC-DFTB-PA	
21	Superficie de energía libre para el cálculo de Umbrella Sampling bidimensional a nivel SCC-DFTB	26
22	Superficie de energía libre para el cálculo de Umbrella Sampling bidimensional a nivel SCC-DFTB-PA	26
23	Representación de las estructuras correspondientes a las 4 esquinas de la Superficie de energía libre	27
24	Mínimo camino energético para la Superficie de energía libre SCC-DFTB, y fluctuación de las distancias relevantes.	28
25	Mínimo camino energético para la Superficie de energía libre SCC-DFTB-PA, y fluctuación de las distancias relevantes.	29
26	Poblaciones de Mulliken para distintos entes moleculares incluidos en el subsistema cuántico.	
27	Poblaciones de Mulliken para los principales centros atómicos involucrados en la reacción.	

AGRADECIMIENTOS

Como toda actividad que se desarrolla en un ámbito social, esta tesis es en menor o mayor medida el fruto de un esfuerzo colectivo. Quiero agradecer a todas aquellas personas que contribuyeron a transmitirme una enseñanza, a facilitarme una herramienta, a compartirme una idea...

- Gracias a Santi Di Lella, Lucia Álvarez, Juan Pablo Arcón y Victoria Dumas por haberme recibido tan cálidamente en el grupo.

- Gracias a Juan Pablo Bustamante, Lucas Defelipe, Ramiro Rodríguez Limardo, Elias Lopez, Uriel Morzan, Claudia Ramírez, Natalí Di Russo y Johan Galindo Cruz por haberme disipado miles de dudas info y bioinformaticas.

- Gracias a Diego Gauto y Carlos Modenutti, (los miembros del "Glycoteam"), que me permitieron siempre participar de sus proyectos y de los que aprendí mucho, con muchas risas incluidas.

- Gracias a Ariel Petruk por los parámetros del UDP que se utilizaron en este trabajo, y por (sin saberlo) ser una inspiración para dar mis primeros pasos programando en TCL.

- Gracias a Fernando Boubeta, por haber pasado toda una tarde ayudándome a hacer funcionar Antechamber.

- Gracias a Gabriel Jara, por haberme enseñado como hacer los gráficos con doble eje Y en Grace.

- Gracias a Charly Guardia, por enseñarme a hacer un perfil de Umbrella Sampling "a mano".

- Gracias a Ari Zeida por clarificarme muchas cuestiones inherentes al Umbrella Sampling en el trayecto final de este trabajo.

- Gracias a Leonardo Boechi por darme una mano con las Dinámicas Aceleradas.

- Gracias a Gustavo (vos mostro) por haber participado activamente en este trabajo.

- Gracias a Darío Estrín, Adrian Turjanski y Damián Scherlis, por ser excelentes docentes.

- Gracias a Adrian E. Rotiberg, por su inmensa generosidad, por lo mucho que me enseñó, y por proveer la mayoría de los recursos para este trabajo.

- Gracias especialmente a Marcelo Adrián Martí por haberme dado esta oportunidad, por haber confiado en mí, y fundamentalmente, por tenerme paciencia.

A todos ustedes, sinceramente, muchas gracias. Un pedazo de este trabajo les pertence.