

INSTITUTO DE CIENCIA MOLECULAR



VNIVERSITAT[̄] DE VALÈNCIA

ESTUDIO TEÓRICO DE MECANISMOS DE REACCIONES
ENZIMÁTICAS CON TRANSFERENCIA DE HIDRÓGENO:
LACTATO DESHIDROGENASA

SILVIA FERRER CASTILLO
Valencia, 8 Marzo 2006

Ignacio Tuñón García de Vicuña, profesor titular de Química Física del departamento de Química Física de la Universidad de Valencia y **Vicent Moliner Ibáñez**, profesor titular de Química Física del departamento de Ciencias Experimentales de la Universidad Jaume I de Castellón,

CERTIFICAN:

Que el trabajo con título: '**Estudio Teórico de Mecanismos de Reacciones Enzimáticas con Transferencia de Hidrógeno: Lactato Deshidrogenasa**' ha sido realizado por Silvia Ferrer Castillo bajo nuestra dirección, para optar al grado de doctora en Química.

Así, autorizan la presentación de este trabajo a efectos de seguir los trámites correspondientes de la Universidad de Valencia.

Valencia, 8 Marzo de 2005.

Dr. Ignacio Tuñón García de Vicuña

Dr. Vicent Moliner Ibáñez

*A mis padres
y a Rober*

*“... La ciència es una senda per a caminar,
no és una estació per a arribar i quedar-se parat.”
V. Sos Baynat*

*“Tengo una pregunta que a veces me tortura:
¿Estoy loco yo? o los locos son los demás”
Albert Einstein*

ÍNDICE GENERAL

PRÓLOGO	I
ACRÓNIMOS	V
CAPÍTULO 1. INTRODUCCIÓN	1
1 BIBLIOGRAFÍA	6
CAPÍTULO 2. OBJETIVOS	9
CAPÍTULO 3. FUNDAMENTOS Y MÉTODOS	13
1 SUPERFICIES DE ENERGÍA POTENCIAL	16
2 MÉTODOS HÍBRIDOS QM/MM	23
2.1 FUNCIÓN ENERGÍA	23
2.2 ÁTOMOS FRONTERA	26
2.3 LIMITACIONES Y MEJORAS DEL MÉTODO QM/MM	28
2.3.1 MÉTODOS DE ENLACE DE VALENCIA SIMPLE	30
2.3.2 MÉTODO DE CORRECCIONES INTERPOLADAS	31
3 LOCALIZACIÓN DE PUNTOS ESTACIONARIOS EN SISTEMAS DE ALTA DIMENSIONALIDAD	34
4 DINÁMICA MOLECULAR Y ENERGÍAS LIBRES	37
4.1 DINÁMICA MOLECULAR	37
4.2 ENERGÍAS LIBRES EN TERMODINÁMICA ESTADÍSTICA	42
4.3 MÉTODO DE PERTURBACIÓN DE LA ENERGÍA LIBRE	45
4.4 MÉTODO DE INTEGRACIÓN TERMODINÁMICA	47
4.5 MÉTODO DE CRECIMIENTO LENTO	48
4.6 MÉTODO DE POTENCIALES DE FUERZA MEDIA	49
4.6.1 CONCEPTO DE POTENCIAL DE FUERZA MEDIA	49
4.6.2 UMBRELLA SAMPLING	52
4.6.3 MÉTODO DE ANÁLISIS DE HISTOGRAMAS PONDERADOS (WHAM)	53
5 TEORÍA DEL ESTADO DE TRANSICIÓN	57
5.1 TEORÍA CONVENCIONAL DEL ESTADO DE TRANSICIÓN	58
5.2 TEORÍA VARIACIONAL DEL ESTADO DEL TRANSICIÓN	64
5.3 EFECTOS CUÁNTICOS	66
5.4 TEORÍA VARIACIONAL DEL ESTADO DE TRANSICIÓN PARA MEDIOS ENZIMÁTICOS	70
6 ENZIMOLOGÍA	75
6.1 CINÉTICA ENZIMÁTICA	76
6.2 ESTADO DE PROTONACIÓN Y REACTIVIDAD	81
6.3 MUTAGÉNESIS	86

7	BIBLIOGRAFÍA	89
<hr/>		
	CAPÍTULO 4. LACTATO DESHIDROGENASA	99
1	BIOQUÍMICA	104
2	RESULTADOS	109
2.1	CÁLCULOS PRELIMINARES	109
2.2	PARÁMETROS DE INTERACCIÓN QM/MM	113
2.3	COORDENADAS DE REACCIÓN	116
2.4	CONSTRUCCIÓN DE UN NUEVO MODELO	117
2.5	EXPLORACIÓN DE LA SEP PARA EL TETRÁMERO	119
2.6	CORRECCIONES A LA DESCRIPCIÓN CUÁNTICA	121
2.6.1	CORRECCIONES SVB	122
2.6.2	CORRECCIONES INTERPOLADAS UNIDIMENSIONALES	124
2.6.3	CORRECCIONES INTERPOLADAS BIDIMENSIONALES	126
2.7	AMPLIACIÓN DE LA PARTE QM	129
2.8	ESTADO DE PROTONACIÓN DE LOS RESIDUOS	134
2.9	ESTUDIO DE LA CINÉTICA DE LA REACCIÓN QUÍMICA	140
2.9.1	EXPLORACIÓN DE LA REGIÓN DEL ET	145
2.9.2	ANÁLISIS DEL ESTADO DE REACTIVOS	149
2.9.3	CÁLCULOS DE ENERGÍA LIBRE	156
2.9.4	EFECTO TÚNEL	161
2.9.5	KIES	162
2.10	MUTACIONES	163
2.10.1	ARG109GLN	166
2.10.2	ASP168ALA	170
2.10.3	DOBLE MUTACIÓN ARG109GLN Y ASN14ASP	173
2.10.4	DOBLE MUTACIÓN ARG109GLN Y ASN140GLU	176
3	CONCLUSIONES Y PERSPECTIVAS DE FUTURO	178
4	BIBLIOGRAFÍA	181
	CAPÍTULO 5. DIHIDROFOLATO REDUCTASA	187
1	BIOQUÍMICA	191
2	RESULTADOS	194
2.1	RESULTADOS EN FASE GAS	195
2.2	RESULTADOS EN EL SENO DE LA ENZIMA DHFR	200
3	PERSPECTIVAS DE FUTURO	208
4	BIBLIOGRAFÍA	210
	ANEXO	215

PRÓLOGO

Me resulta muy difícil resumir el trabajo de una Tesis Doctoral e incluir todos los pasos que se han dado a lo largo del camino en mi formación investigadora. Durante estos años he tenido la oportunidad de trabajar con personas y visitar lugares increíbles. El conjunto de toda esta experiencia me ha traído a este tramo final en el que hay que recopilar todo el trabajo y plasmarlo en un texto escrito.

Además de la necesidad de transcribir todo el trabajo doctoral en un único libro, he tratado que este trabajo quedara, dentro de su complejidad, como herramienta útil para generaciones venideras de nuevos jóvenes investigadores y así, con mi granito de arena, ayudarles a comprender un poco mejor el mundo de la Química Computacional.

El prólogo de esta Tesis Doctoral sirve como sección de agradecimientos ya que no quiero olvidar a la gente que me ha ayudado a llegar a este punto y final (aunque espero que en realidad sea un punto y seguido). No quisiera dejar pasar la oportunidad de agradecer a las personas que han influido en mi trabajo tanto directa como indirectamente. No obstante, adelanto que no voy a utilizar ningún nombre propio. No por el hecho de que quizás enumerándolos uno a uno la extensión del prólogo superaría la de la totalidad de la Tesis, sino porque durante este tiempo he intentando personalizar mi agradecimiento individualmente, con una sonrisa, un gesto, un guiño; esas personas ya sabéis quienes sois.

Para comenzar, tengo un especial agradecimiento, con mucho cariño, para mis directores de tesis que me han conducido por la senda

todos estos años, sin dejar que me perdiera en el camino. En ellos he encontrado un grado de comprensión y ayuda que yo misma no he sabido valorar en ocasiones. Me han enseñado una filosofía del mundo de la ciencia que agradezco de corazón y espero no haberles dado demasiados quebraderos de cabeza con mi manera de ser y afrontar las cosas.

Adicionalmente, quiero agradecer a los directores no oficiales de mi trabajo, aquellos que me han transferido sus conocimientos de manera extraordinaria y genial. Con los que he compartido artículos, discusiones y análisis de resultados y mucha paciencia por su parte.

Para continuar mis agradecimientos, quiero destacar que mi trabajo de tesis lo he llevado a cabo entre dos grupos de investigación en su mayor parte (entre ciudades cercanas) y de otros dos donde he realizado estancias cortas pero intensas (un poco más lejos de mi hogar). Me siento privilegiada por haber podido compartir despacho con mis compañeros. Me he sentido arropada y querida. Y he podido compartir con ellos muy buenos momentos. La carga de los instantes malos: en los que no te da un resultado o se te ha borrado el “*scratch*”; la he llevado mucho más liviana. Y sé que he podido contar con vuestra ayuda desinteresada por lo que quiero que conste que os lo agradezco, desde los que ya se han ido hasta los que llevan menos tiempo. Da gusto ir a trabajar con el ambiente de *buen rollo*. Por supuesto, en este párrafo incluyo los que han compartido conmigo durante las estancias, ha sido menos tiempo pero les tengo un gran afecto.

También agradezco a la gente del departamento, o de la Universidad en general, con la que no he compartido despacho directamente pero que hemos compartido muchos momentos; me han disipado dudas existenciales (o me las han generado) y me han animado (incluso a veces desanimado) en mi trabajo. Con ellos he compartido

profundos debates en el café o charlas banales; cenas y celebraciones, y siempre percibido un especial cariño y amabilidad.

Agradezco el apoyo de mis amigos (de los que tenía antes y los que han ido apareciendo por el camino, los que han ido dejando de llamar y los que continúan llamando). Agradezco el apoyo mis familiares cercanos y lejanos, que muestran un respeto especial y admiración por mi trabajo aunque no están muy seguros de lo que hago. Familia ajena a la Universidad que me pregunta cómo llevo la tesis. A los amigos y familiares, os lo agradezco porque durante los últimos meses me habéis ‘dejado’ quedar menos, vernos poco y aceptado mis excusas de que estaba liada con la tesis y no podía quedar. Ahora, por lo menos veréis que no era una excusa y que realmente estaba haciendo lo que os decía: *escribiendo la tesis*.

Agradezco a mis padres su amor y su entrega incondicional, insistiendo siempre en que siguiera mi camino, que estudiara, que yo servía para esto. Estos años me han cuidado (mimado en exceso) sin reprocharme mis olvidos, mis manías, mi desorden y mis cambios de humor. Sé que estáis orgullosos de mí pero no sé si sabéis que yo estoy orgullosa de vosotros.

A mi compañero en la vida, que me ha influido estos años para ser mejor persona, para comprender ciertas actitudes de los demás, aceptando las críticas con mesura. Gracias, porque tu manía por el orden me ha hecho ser más ordenada y sistemática, porque me has animado en los momentos en que pensaba que no podía seguir, porque siempre estás dispuesto a ayudar.

Espero que al leer estas líneas os sintáis identificados en algún momento, he intentado no olvidar a nadie. Tanto si leéis esta tesis en su totalidad como si sólo le vais a dedicar el tiempo necesario del prólogo:

Gracias a todos.

ACRÓNIMOS

Por orden alfabético.

AM1	Austin Model 1
BsLDH	Bacillus stearothermophilus Lactato Deshidrogenasa
CPU	Unidad Central de Proceso
CR	Coordenada de Reacción.
CSUT	Cubic Splines Under Tension
CTST	Teoría Convencional del Estado de Transición
CVT	Teoría Variacional del Estado de Transición Canónica
DFT	Teoría del Funcional de la Densidad
DHFR	Dihidrofolato Reductasa
EcDHFR	Escherichia coli Dihidrofotalo Reductasa
EF	Eigenvector Folloging
EsT	Estructura de Transición
ET	Estado de Transición.
FEP	Perturbación de la Energía Libre
GHO	<i>Generalized Hybrid Orbital</i>
GTS	Estado de Transición Generalizado
IC	Correcciones Interpoladas Unidimensionales
IC2D	Correcciones Interpoladas Bidimensionales
IRC	Camino Intrínseco de Reacción.
LCT	<i>Large-Curvature Tunneling</i>
LDH	Lactato Deshidrogenasa
LSCF	Campo Local Auto-Consistente

MD	Dinámica Molecular
MM	Mecánica Molecular
NADH	Nicotinamida Adenín Dinucleótido
OMT	<i>Optimized Multidimensional Tunneling</i>
QM	Mecánica cuántica (<i>Quantum Mechanics</i>)
QM/MM	Mecánica Cuántica / Mecánica Clásica
RFO	Rational Function Optimization
SCT	<i>Small-curvature Tunneling</i>
SEP	Superficie de Energía Potencial
SRP	Parámetros específicos para una reacción (<i>Specific Reaction Parameters</i>)
SVB	Métodos de Enlace de Valencia Simple (<i>Simple Valence Bond</i>)
TI	Integración Termodinámica (Thermodynamic Integration)
TST	Teoría del Estado de Transición
VTST	Teoría Variacional del Estado de Transición
VTST/OMT	Variational Transition State Theory with <i>Optimized Multidimensional Tunneling</i>
WHAM	Método de Análisis de Histogramas Ponderados
ZCT	Túnel de curvatura cero (<i>Zero-Curvature Tunneling</i>)

CAPÍTULO 1

INTRODUCCIÓN

CAPÍTULO 1. INTRODUCCIÓN	3
---------------------------------	----------

1 BIBLIOGRAFÍA	6
-----------------------	----------

Las enzimas son proteínas que catalizan las reacciones biológicas. La eficacia y especificidad de las enzimas permite que las reacciones que tienen lugar en los seres vivos se lleven a cabo a una escala de tiempo compatible con la vida.^[1] Las reacciones químicas catalizadas por enzimas son aceleradas en un orden de magnitud de 10^6 a 10^{23} con respecto a sus respectivas reacciones en disolución.^[1] Por lo tanto, el principal objetivo de muchos estudios científicos, dentro del campo de la química y de la bioquímica, es conocer el mecanismo de reacciones catalíticas y entender el motivo del poder catalítico de las enzimas, tanto desde un punto de vista experimental^[2-6] como teórico.^[7-13]

La Química Computacional tiene como objetivo el análisis de problemas químicos mediante programas informáticos basados en la Química Teórica. Los resultados permiten mejorar el entendimiento de datos experimentales y predecir el comportamiento desconocido de sistemas químicos. La evolución en las últimas décadas de la Química Computacional ha ido estrechamente ligada a la evolución de la capacidad de cálculo de los ordenadores. A medida que la tecnología lo ha ido permitiendo han surgido metodologías más potentes y complejas capaces de abordar el estudio de problemas que hasta hace poco tiempo

eran impensables; a su vez, estas metodologías requieren cada vez de más recursos. Afortunadamente, los procesadores aumentan su capacidad muy rápidamente lo que ha permitido que la Química Computacional se aplique a un número de problemas cada vez mayor. Paralelamente, el crecimiento de la red *internet* ha proporcionado un recurso mediante el cual compartir metodologías, datos y teorías en fracciones de segundo.

A pesar de este gran avance, el estudio mecanocuántico desde el punto de vista computacional es muy costoso en sistemas con un elevado número de átomos. Por ello en los estudios teóricos sobre sistemas biológicos como las enzimas se tiende a utilizar nuevas metodologías que permiten estudiar la reacción química teniendo en cuenta la interacción con el medio que la rodea, ya sea en una proteína o un disolvente, a un costo no excesivo. La aproximación mayormente aceptada para abordar este tipo de problemas se basa en los métodos híbridos QM/MM^[14-22] que permiten una descripción de los átomos involucrados directamente en la formación o ruptura de enlaces de una reacción utilizando métodos mecánico cuánticos (*quantum mechanical*: QM), mientras los átomos del entorno se describen mediante campos de fuerza clásicos de la mecánica molecular (*molecular mechanics*: MM).

La presente Tesis Doctoral es un estudio teórico de reacciones enzimáticas que implican la transferencia de átomos de hidrógeno, ya sea en su forma protonada o en forma de ión hidruro. La memoria consta de 5 capítulos más un anexo en el que se presentan las principales aportaciones publicadas en revistas científicas.

En la primera parte del manuscrito se presentan los objetivos del trabajo doctoral y los fundamentos de las metodologías de la Química Computacional necesarias para entender el desarrollo y progresión de los

trabajos realizados en esta Tesis Doctoral (Capítulos 2 y 3) incluyendo una introducción breve a la enzimología, donde se hará hincapié en aspectos relacionados con la reactividad química de estos complejos catalizadores biológicos.

En la segunda parte (Capítulos 4 y 5) se presentan los estudios realizados en las dos enzimas objeto de estudio. El Capítulo 4 está dedicado al estudio realizado en la enzima Lactato Deshidrogenasa (LDH). El recorrido de este estudio va desde la exploración de la superficie de energía potencial con distintas metodologías hasta el estudio de perfiles de energía libre del mecanismo de reacción, incorporando efectos cuánticos como el efecto túnel. Al final de este capítulo se analizan diferentes mutaciones de la enzima. En el Capítulo 5 se presenta un estudio preliminar de la enzima Dihidrofolato Reductasa (DHFR) donde, a partir de la exploración de las superficies de energía potencial, se pretende discernir el origen del protón que es transferido, previo a la etapa cinéticamente limitante en la reacción química.

Finalmente, debido al carácter europeo de la Tesis Doctoral, en el Anexo se presentan los artículos científicos publicados (o en proceso de publicación) en inglés.

En esta memoria se ha querido comentar todas las metodologías y teorías implicadas directa o indirectamente en el trabajo realizado. En general, las metodologías son publicadas en revistas de carácter internacional por lo que muchos de los nombres quedan en su acrónimo inglés. Para facilitar la lectura de la tesis se ha creado una sección de acrónimos en las páginas V y VI. Al final de cada capítulo aparecen las referencias bibliográficas. Las ecuaciones se numeran por orden dentro de cada sección de cada capítulo, mientras que las figuras se numeran a lo largo de cada capítulo. Los encabezados de página dan

información del capítulo (páginas impares) y de la sección dentro del capítulo (páginas pares).

1 Bibliografía

- [1] R. Wolfenden, *Acc. Chem. Res.* **1972**, *5*, 10-18.
- [2] A. Cooper, J. L. Houben and L. C. Chien, *The Enzyme Catalysis Process*, Plenum, New York, **1989**.
- [3] J. R. Knowles, *Science* **1987**, *236*, 1252-1258.
- [4] L. Stryer, *Biochemistry*, W. H. Freeman and Company, New York, **1995**.
- [5] A. R. Fersht, *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman and Company, New York, **1999**.
- [6] C. Walsh, *Nature* **2001**, *409*, 226-231.
- [7] A. Yadav, R. M. Jackson, J. J. Holbrook and A. Warshel, *J. Am. Chem. Soc.* **1991**, *113*, 4800-4805.
- [8] A. Warshel, *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250-5254.
- [9] M. Garcia-Viloca, J. Gao, M. Karplus and D. G. Truhlar, *Science* **2004**, *303*, 186-195.

- [10] K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann and A. J. Mulholland, *Org. Biomol. Chem.* **2004**, *2*, 968-980.
- [11] A. J. Mulholland, *Drug Discovery Today* **2005**, *10*, 1393-1402.
- [12] M. Oliva, O. Dideberg and M. J. Field, *Proteins-Structure Function and Genetics* **2003**, *53*, 88-100.
- [13] D. Riccardi, P. Schaefer and Q. Cui, *J. Phys. Chem. B* **2005**, *109*, 17715-17733.
- [14] A. Warshel and M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227-249.
- [15] U. C. Singh and P. A. Kollman, *J. Comput. Chem.* **1986**, *7*, 718-730.
- [16] M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700-733.
- [17] P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe and M. Karplus, *Biochemistry* **1991**, *30*, 5826-5832.
- [18] T. C. Bruice and K. Kahn, *Curr. Opin. Chem. Biol.* **2000**, *4*, 540-544.
- [19] M. García-Viloca, A. Gonzalez-Lafont and J. M. Lluch, *J. Am. Chem. Soc.* **2001**, *123*, 709-721.
- [20] J. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.* **2002**, *53*, 467-505.
- [21] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Theor. Chem. Acc.* **2001**, *3*, 207-212.
- [22] P. A. Kollman, B. Kuhn, O. Donini, M. Peräkylä, R. V. Stanton and D. Bakowies, *Acc. Chem. Res.* **2001**, *34*, 72-79.

CAPÍTULO 2

OBJETIVOS

La presente Tesis Doctoral tiene como objetivo la aplicación de los métodos híbridos que combinan la mecánica cuántica con la mecánica molecular, métodos QM/MM, al estudio teórico de mecanismos de reacción en medios enzimáticos. Concretamente, se estudiarán reacciones de transferencia de átomos de hidrógeno. Los dos sistemas enzimáticos objeto de estudio son: la Lactato Deshidrogenasa (LDH) y la Dihidrofolato Reductasa (DHFR).

Para la LDH los objetivos son:

- Estudiar el mecanismo de reacción de la LDH a partir de la superficie de energía potencial aplicando los métodos híbridos QM/MM.
- Estudiar las correcciones necesarias para subsanar las limitaciones de los métodos semiempíricos utilizados en la descripción del subsistema cuántico.
- Hacer un estudio del estado de protonación de los residuos que componen la enzima a pH fisiológico, analizando las consecuencias de cada estado de protonación sobre la reactividad.
- Estudiar mediante métodos de Dinámica Molecular los perfiles de energía libre de la reacción.

- Aplicar la Teoría Variacional del Estado de Transición, incluyendo el efecto túnel, para calcular la constante cinética de la reacción. Calcular los Efectos Isotópicos Cinéticos.
- Estudiar la importancia y el papel concreto de ciertos aminoácidos mediante el análisis del efecto de las mutaciones de dichos aminoácidos.
- Establecer un protocolo de trabajo que pueda ser aplicado a otras enzimas como la DHFR.

Para la DHFR los objetivos son:

- Analizar la bonanza del método semiempírico AM1 que se va a aplicar en el estudio QM/MM. Para ello se llevará a cabo el estudio en fase gas con un modelo que represente la reacción catalizada por la enzima a diferentes niveles de cálculo.
- Estudiar el mecanismo de reacción de la etapa de protonación del N5 del anillo pteridino a partir de la superficie de energía potencial obtenida mediante los métodos híbridos QM/MM.

CAPÍTULO 3

FUNDAMENTOS Y MÉTODOS

CAPÍTULO 3. FUNDAMENTOS Y MÉTODOS	13
1 SUPERFICIES DE ENERGÍA POTENCIAL	16
2 MÉTODOS HÍBRIDOS QM/MM	23
2.1 FUNCIÓN ENERGÍA	23
2.2 ÁTOMOS FRONTERA	26
2.3 LIMITACIONES Y MEJORAS DEL MÉTODO QM/MM	28
2.3.1 Métodos de Enlace de Valencia Simple	30
2.3.2 Método de Correcciones Interpoladas	31
3 LOCALIZACIÓN DE PUNTOS ESTACIONARIOS EN SISTEMAS DE ALTA DIMENSIONALIDAD	34
4 DINÁMICA MOLECULAR Y ENERGÍAS LIBRES	37
4.1 DINÁMICA MOLECULAR	37
4.2 ENERGÍAS LIBRES EN TERMODINÁMICA ESTADÍSTICA	42
4.3 MÉTODO DE PERTURBACIÓN DE LA ENERGÍA LIBRE	46
4.4 MÉTODO DE INTEGRACIÓN TERMODINÁMICA	48
4.5 MÉTODO DE CRECIMIENTO LENTO	48
4.6 MÉTODO DE POTENCIALES DE FUERZA MEDIA	49
4.6.1 Concepto de Potencial de Fuerza Media	49
4.6.2 Umbrella Sampling	52
4.6.3 Método de Análisis de Histogramas Ponderados (WHAM)	53
5 TEORÍA DEL ESTADO DE TRANSICIÓN	57
5.1 TEORÍA CONVENCIONAL DEL ESTADO DE TRANSICIÓN	58
5.1.1 Teoría Canónica	59
5.2 TEORÍA VARIACIONAL DEL ESTADO DEL TRANSICIÓN	64
5.3 EFECTOS CUÁNTICOS	67
5.4 TEORÍA VARIACIONAL DEL ESTADO DE TRANSICIÓN PARA MEDIOS ENZIMÁTICOS	71
6 ENZIMOLOGÍA	75
6.1 CINÉTICA ENZIMÁTICA	76
6.2 ESTADO DE PROTONACIÓN Y REACTIVIDAD	81
6.3 MUTAGÉNESIS	86
7 BIBLIOGRAFÍA	89

La Química Teórica y Computacional presenta un amplio abanico de posibilidades tanto en metodologías como en teorías. Consecuentemente, es importante introducir las teorías, fundamentos y herramientas de cálculo empleados en nuestros ‘experimentos’ computacionales, así como describir las aproximaciones en las que se basan.

La Química Teórica y Computacional permite abordar el estudio de la reactividad química. Ésta implica la reorganización de los núcleos y los electrones que constituyen el sistema químico dando lugar a cambios de la estructura molecular desde reactivos hasta productos. Dicha transformación puede llevarse a cabo en más de un paso y por caminos diferentes, constituyendo cada uno de estos caminos un posible mecanismo de la reacción. Los mecanismos de reacción pueden describirse a través de la localización de los diferentes puntos estacionarios sobre la superficie de energía potencial (reactivos, productos, estructura de transición e intermedios). Desde un punto de vista teórico, la caracterización de la estructura de transición de la etapa limitante de la velocidad es clave para la comprensión de la reactividad química.^[1]

1 Superficies de Energía Potencial

La Energía Potencial de una molécula es la energía correspondiente a una configuración fija de los núcleos. Una superficie de energía potencial (SEP) es una representación de la energía potencial del sistema en función de un número determinado de coordenadas internas donde cada punto representa la posición relativa de los núcleos del sistema y su energía. Las coordenadas generalizadas internas q_i , que son distancias de enlace, ángulos y ángulos diedros, pueden usarse para describir la posición de los núcleos, de manera que la energía se exprese en función de dichas coordenadas.^[2-4] Si el sistema tiene N núcleos, el número de coordenadas internas independientes o grados de libertad necesarios para determinar una SEP es, en general, $3N-6$. La SEP, dentro de la aproximación de Born-Oppenheimer,^[5] es la herramienta fundamental para explicar las propiedades de las moléculas, así como también para describir las reacciones en las que intervienen. Una vez obtenida la SEP se pueden estudiar sobre ella propiedades de las moléculas como su estabilidad, su estructura o sus posibles isómeros, y las reacciones que puede sufrir dicho sistema desde el punto de vista cinético y termodinámico.^[6] Al representar en la SEP la energía frente a las coordenadas internas se obtendría un diagrama n -dimensional imposible de visualizar. La situación más favorable de visualización de una superficie corresponde al caso tridimensional (3D) donde se representa la energía frente a dos coordenadas o combinación de coordenadas. Para facilitar la caracterización del mecanismo de la reacción en la práctica es habitual trabajar con la proyección de estas

superficies tridimensionales, donde la energía se indica mediante curvas isoenergéticas. (Figura 1).

La SEP contiene mucha información del sistema reactivo, pero no todas las regiones de la superficie tienen la misma importancia desde el punto de vista químico. Puntos de especial consideración son los reactivos, productos, intermedios reactivos y estructuras de transición (EsT). Todos ellos son puntos estacionarios o críticos, puntos particulares sobre la SEP donde el gradiente es nulo.

$$\vec{g} = 0 \quad (1.1)$$

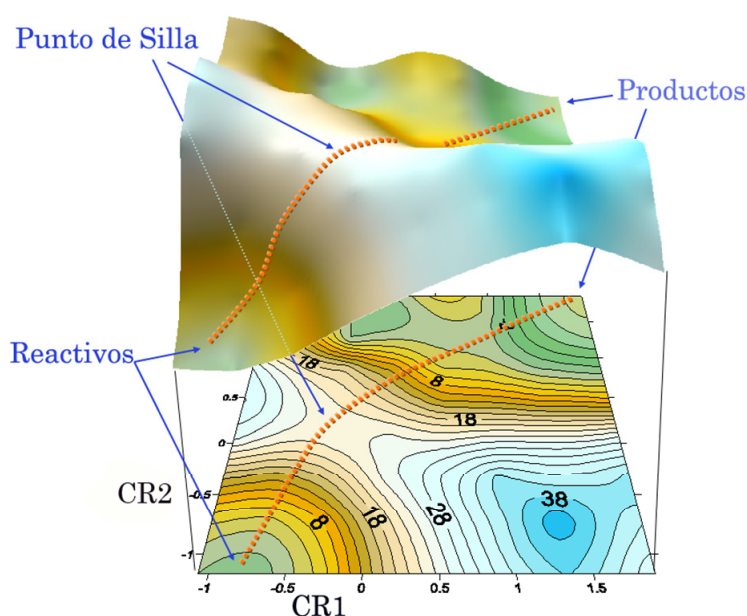


Figura 1. Superficie de energía potencial en 3D y su proyección en 2D. La energía relativa se representa en función de dos combinaciones de coordenadas (CR1 y CR2). Los puntos estacionarios localizados en la SEP están señalados con flechas azules.

Los puntos estacionarios presentan unas determinadas características: corresponden a estructuras moleculares estables o metaestables; la fuerza sobre los núcleos es nula y su posición es

independiente del sistema de coordenadas elegido.^[7] Matemáticamente, los puntos estacionarios se caracterizan porque la primera derivada de la función energía respecto a las coordenadas internas es igual a cero. La naturaleza de estos puntos viene dada por el número de valores propios positivos y negativos de la matriz Hessiana (matriz de constantes de fuerza o de derivadas segundas). Si todos los valores propios son positivos significa que al desplazarnos del punto crítico la energía siempre aumentará. Esta condición corresponde, por tanto, a un mínimo sobre la SEP. Entre los puntos mínimos podemos tener reactivos, productos e intermedios reactivos. Si la matriz Hessiana diagonalizada posee un único valor propio negativo significa que este punto es un mínimo en todas las direcciones excepto en una, para la que es un máximo.^[8] Si hay n valores propios negativos tendremos un punto de silla de orden n , y si todos los valores propios son negativos presenta un máximo.

La estructura del estado de transición de una reacción química se define como el punto sobre la SEP que cumple las siguientes cuatro condiciones: a) es un punto estacionario, es decir, de gradiente cero; b) la matriz diagonalizada de constantes de fuerza en el punto debe tener un único valor propio negativo; c) debe ser el punto de más alta energía potencial en una línea continua que conecte reactivos y productos; d) debe ser el punto de menor energía potencial que cumpla las anteriores tres condiciones. Debido a las dos primeras condiciones matemáticas la estructura de transición (EsT) recibe también el nombre de punto de silla.^[9]

Además de las coordenadas cartesianas e internas, en algunas ocasiones puede resultar útil la definición de otros sistemas de coordenadas que faciliten el tratamiento algebraico o aritmético en los

estudios cinéticos. Si llamamos m_i a la masa del átomo i (donde $i=1,2, \dots, N$) y $R_{i\gamma}$ a las coordenadas cartesianas del átomo i (siendo $\gamma = x, y, z$), definidas con respecto al centro de masas del sistema, las coordenadas cartesianas ponderadas en peso, $x_{i\gamma}$, se definen mediante la relación:

$$x_{i\gamma} = m_i^{1/2} R_{i\gamma} \quad (1.2)$$

con unidades de (masa)^{1/2} longitud.

En principio, el camino de reacción debe corresponder al camino de menor energía sobre la SEP que une reactivos y productos (ver Figura 1). Sin embargo, cuando se desea precisar este concepto y describirlo cuantitativamente aparecen dificultades, por lo que se debe distinguir entre coordenada de reacción distinguida y camino de reacción intrínseco.

Coordenada de Reacción Distinguida.^[10] La forma más intuitiva de definir el camino de reacción es relacionarlo con una o varias coordenadas (habitualmente coordenadas internas) que tengan una especial importancia en la reacción (por ejemplo el enlace que se forma o se rompe, o bien la combinación antisimétrica de ambas), a la que se le suele denominar coordenada de reacción distinguida (*distinguished reaction coordinate*, RC). Moviéndonos sobre esta coordenada, se optimizan las restantes coordenadas internas. El conjunto de coordenadas resultante define el camino de reacción. La altura de la barrera dependerá de la posición relativa del punto de mayor energía sobre el perfil de energía potencial.

Aunque este concepto es fácil de manejar presenta dos importantes inconvenientes. En primer lugar, no siempre es posible identificar una coordenada de reacción apropiada con una de las

coordenadas internas. En segundo lugar, este método no garantiza que el camino de reacción pase sobre la barrera de energía más baja posible conectando reactivos y productos; es decir, no garantiza que pase sobre el verdadero punto de silla de la reacción. A la coordenada de reacción distinguida^[10] se le ha llamado también camino de reacción intuitivo. ^[3, 11]

Camino de reacción intrínseco.^[12, 13] (*Intrinsic Reaction Coordinate*, IRC). Para solucionar los anteriores inconvenientes se propuso reemplazar el concepto de camino de mínima energía por camino de máxima pendiente. En lugar de partir desde los reactivos hasta los productos, como en el método anterior, se parte del punto de silla y se sigue el camino de máxima pendiente. Este método de máxima pendiente pasó a referirse a un sistema de coordenadas único, el sistema de coordenadas ponderadas en peso, ecuación (1.2). Fukui y colaboradores^[13] denominaron a este método coordenada de reacción intrínseca, (*Intrinsic Reaction Coordinate*, IRC). En la Figura 2 se observa el corte la una SEP en el plano que describe la coordenada de reacción intrínseca. Una vez identificado el camino de reacción de forma inequívoca por medio del IRC, se puede describir cualquier punto usando como única variable su localización en este camino. Para ello se define la coordenada de reacción intrínseca, habitualmente representada por la letra **s**. Al punto de silla se le asigna de forma arbitraria un valor de **s=0**. El valor de **s** para un punto del camino se calcula como la diferencia a lo largo del camino entre dicho punto y el punto de silla. Por convenio, si este punto se encuentra entre los reactivos y el punto de silla, su valor será negativo, mientras que si se encuentra en la zona de productos será positivo.

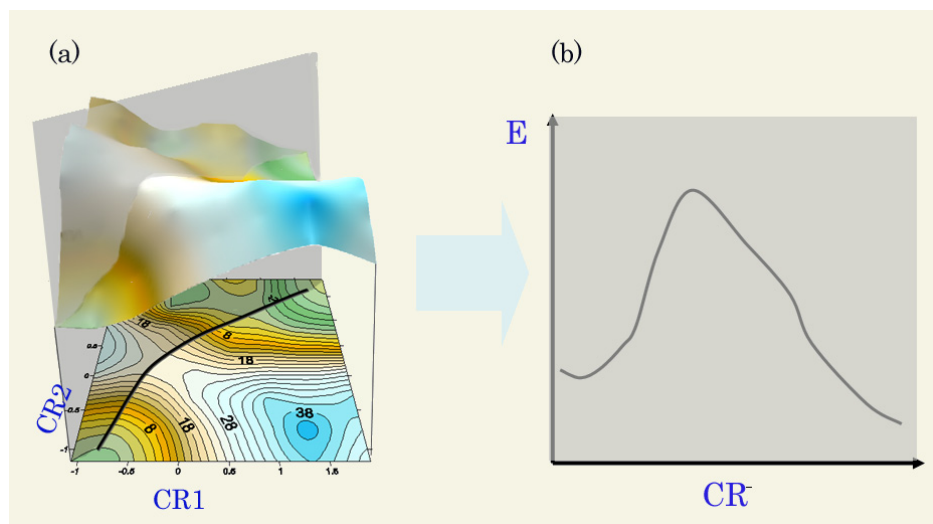


Figura 2. (a) Superficie de energía potencial en 3D, donde la energía se representa en función de dos coordenadas de reacción (CR1 y CR2). (b) Corte de la SEP 3D por el camino de reacción intrínseco. La energía se representa frente a la coordenada de reacción.

El cálculo del IRC consiste en resolver la ecuación (1.3) de forma que se obtengan las coordenadas del punto $n+1$, x_{n+1} , a partir de las coordenadas del punto anterior, x_n . Para cualquier punto $x(s_0)$ sobre el camino de reacción (excepto en el punto de silla donde el gradiente es nulo), el camino de reacción se puede expresar como una serie de Taylor hasta orden 2, ^[14, 15]

$$x(s) = x(s_0) + c_1(s - s_0) + \frac{1}{2}c_2(s - s_0)^2 \quad (1.3)$$

donde el coeficiente c_1 es el vector gradiente normalizado (primera derivada de la energía),

$$c_1 = \frac{dx}{ds} = -\frac{g(s)}{|g(s)|} \quad (1.4)$$

siendo x el vector de posición en coordenadas ponderadas en peso, s la coordenada de reacción intrínseca, g el gradiente de energía potencial, es

decir, el vector de las fuerzas sobre los átomos en cada una de las direcciones cartesianas ponderadas en peso, y $|g(s)|$ la norma del gradiente que asegura que el vector g sea unitario aún cuando la fuerza sobre cada núcleo cambie a lo largo del camino. Así, c_1 es un vector unitario tangente al IRC. Y c_2 , representa la curvatura del camino (segunda derivada de la energía),

$$c_2 = \frac{dc_1}{ds} = -\frac{d^2x}{ds^2} = -\frac{Hc_1(s) - [c_1(s)^T Hc_1(s)]}{|g(s)|} \quad (1.5)$$

donde H es la matriz Hessiana en coordenadas cartesianas ponderadas en peso.

Como en el caso de los puntos estacionarios, la efectividad del método va a depender de la disponibilidad de las derivadas de la energía. Así, se ha propuesto un gran número de métodos para el seguimiento del IRC, desde los que sólo usan las derivadas primeras de la energía a los más efectivos que usan información de las derivadas primeras y segundas de la energía.^[14, 15]

2 Métodos Híbridos QM/MM

2.1 *Función energía*

La descripción mecanocuántica de una molécula pasa por la resolución de la ecuación de Schrödinger. Sin embargo, hasta el momento únicamente se ha conseguido resolver exactamente esta ecuación para un problema de dos partículas no relativista. En el resto de sistemas deberemos utilizar métodos que proporcionen soluciones aproximadas. Generalmente, cuanto mayor sea el tamaño del sistema los métodos que se aplican son menos rigurosos, ya que al requerir mayores recursos computacionales se introducen más aproximaciones.^[4, 15]

Para sistemas con un gran número de átomos, como pueden ser las enzimas, se encuentra el problema de no poder tratar todo el sistema mediante la mecánica cuántica (QM). Por otro lado la mecánica molecular (MM) no permite estudiar los procesos que implican una reacción química, ya que no describe electrones. La metodología híbrida que combina la mecánica cuántica y la mecánica molecular (QM/MM) utiliza la mecánica cuántica (QM), para describir la parte del sistema donde se producen los procesos de rotura y formación de enlaces para un número limitado de átomos, mientras que el resto del sistema se describe con potenciales clásicos de la mecánica molecular, incluyendo así el efecto del medio.^[16-33] En consecuencia, el sistema a estudiar se divide en dos regiones tal y como se representa en la Figura 3.

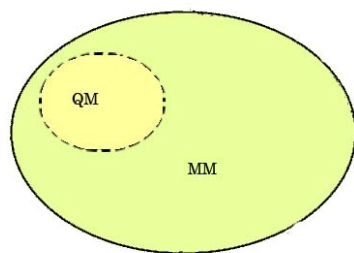


Figura 3. Representación esquemática de la división de un sistema molecular en zona QM y zona MM, es decir, los dos subsistemas presentes en los métodos QM/MM.

La energía potencial total del sistema será la suma de la energía de la parte cuántica, la energía de la parte clásica y la energía de interacción entre ambos subsistemas.

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{mm}} + E_{\text{qm/mm}} \quad (2.1)$$

Los átomos de la región QM se representan como electrones y núcleos. El Hamiltoniano asociado a esta región, en la aproximación de Born-Oppenheimer, se escribe:

$$H_{\text{QM}} = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{ij} \frac{1}{r_{ij}} - \sum_{i\alpha} \frac{Z_\alpha}{r_{i\alpha}} + \sum_{\alpha\beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \quad (2.2)$$

donde i y j representan las coordenadas electrónicas, α y β las coordenadas nucleares, r la distancia electrón-electrón o electrón-núcleo, R las distancia núcleo-núcleo, ∇ es el operador energía cinética electrónica y Z es la carga nuclear.

La zona MM contiene el resto de los átomos del sistema. Éstos se describen clásicamente, y sus interacciones se determinan mediante funciones de energía potencial empíricas, de modo que dichos átomos no pueden estar implicados en procesos de formación o ruptura de enlaces, ni de transferencia de carga. La energía de la región MM depende inicialmente de las posiciones de los átomos o centros tratados mediante mecánica molecular, si ignoramos la polarización de este subsistema:

$$E_{MM} = E_{\text{enlace}} + E_{\text{no enlace}} \quad (2.3)$$

$$E_{\text{enlace}} = E_{\text{enlaces}} + E_{\text{ángulos}} + E_{\text{diedros}} \quad (2.4)$$

$$E_{\text{no enlace}} = E_{\text{electrostática}} + E_{\text{van del Waals}} \quad (2.5)$$

Por último, el término $E_{QM/MM}$ describe cómo interaccionan los átomos QM con los centros MM. El Hamiltoniano, en unidades atómicas, viene dado por la siguiente ecuación:

$$H_{QM/MM} = -\sum_{iM} \frac{q_M}{r_{iM}} + \sum_{\alpha M} \frac{Z_{\alpha} q_M}{R_{\alpha M}} + \sum \left\{ \frac{A_{\alpha M}}{R_{\alpha M}^{12}} - \frac{B_{\alpha M}}{R_{\alpha M}^6} \right\} \quad (2.6)$$

donde los subíndices i, j hacen referencia a los electrones de los átomos QM, los subíndices α a los núcleos de los átomos QM y los subíndices M a los átomos MM. El primer término de la ecuación (2.6) representa las interacciones electrostáticas entre los átomos MM y los electrones de los átomos QM. El segundo término representa las interacciones electrostáticas entre los átomos MM y los núcleos de los átomos QM, y el tercer término describe interacciones de van der Waals entre centros QM y MM. Los dos últimos términos de esta ecuación no incluyen coordenadas electrónicas, de modo que pueden ser calculados y sumados directamente a la energía total. Sin embargo, el primer término sí que incluye coordenadas electrónicas por lo que debe tratarse en el procedimiento autoconsistente.

2.2 Átomos frontera

La partición del sistema en subsistemas QM y MM resulta especialmente problemática cuando la división tiene lugar, necesariamente, a través de la rotura de un enlace covalente entre ambos subsistemas (Figura 4). En estos casos hay que recurrir a algún tipo de aproximación que permita satisfacer las valencias de los átomos cuánticos que definen el enlace entre la región cuántica y la región clásica por donde se producirá la división. Las metodologías más utilizadas se presentan a continuación.

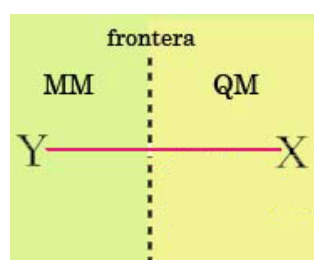


Figura 4. Representación esquemática de la división de un sistema molecular en subsistemas en los métodos QM/MM cuando la frontera corta necesariamente un enlace covalente.

Átomos de Unión. En este método se saturan las valencias de los átomos cuánticos afectados por el corte empleando átomos de unión (*link atoms*) de modo que sustituyen el átomo MM del enlace frontera. El átomo más empleado con este propósito es el átomo de hidrógeno. Este átomo se describe a nivel QM y es invisible para los átomos MM porque no se tienen en cuenta las interacciones entre el átomo *link* y los átomos del subsistema clásico. [18, 19]

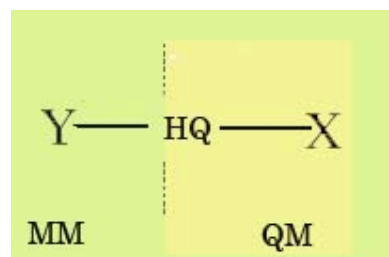


Figura 5. Representación esquemática de la aproximación de átomo link para el tratamiento de los átomos frontera en los métodos QM/MM.

LSCF. Este método está basado en el formalismo del campo local auto-consistente (*Local Self Consistent Field*, LSCF)^[29, 34-36] donde los enlaces QM/MM frontera son descritos mediante orbitales de enlace localizados. Estos orbitales, llamados orbitales congelados, son excluidos del cálculo SCF y se encuentran definidos por sus coeficientes de hibridación y por su población electrónica. Estos dos parámetros son determinados mediante cálculos QM sobre sistemas pequeños y se supone que estos resultados son transferibles a los sistemas de interés.

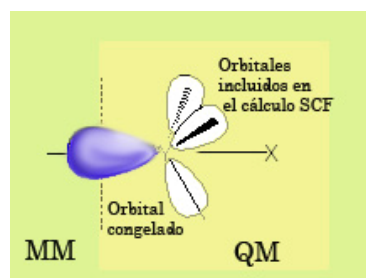
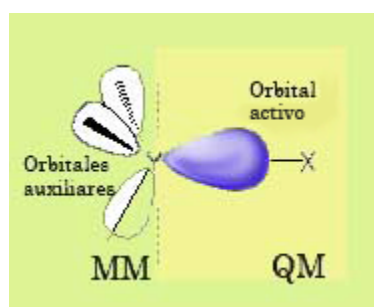


Figura 6. Representación esquemática de la aproximación de LSCF para los átomos frontera en los métodos QM/MM.

GHO. (*Generalized Hybrid Orbital*). En este método se hace uso del orbital atómico híbrido, similar al método anterior LSCF, pero en este caso los orbitales pueden dividirse en orbitales auxiliares y orbitales activos. Estos últimos se optimizan junto con el resto de orbitales de la región QM en el cálculo SCF, mientras que la unión entre el átomo

frontera y la región MM se define mediante los campos de fuerza MM. A diferencia del método basado en el formalismo LSCF, el método GHO no necesita de un proceso de parametrización cada vez que se estudia un



nuevo sistema.^[29, 37]

Figura 7. Representación esquemática de la aproximación de GHO para los átomos frontera en los métodos QM/MM.

En general se recomienda que, en el caso de ser necesario, la partición entre el sistema QM y MM esté lo más alejada posible de los átomos implicados directamente en la reacción. De esta manera se evita que las aproximaciones utilizadas en las metodologías influyan sobre el resultado final. En la presente tesis doctoral los dos métodos elegidos para tratar los enlaces frontera han sido el átomo *link* y el GHO.

2.3 Limitaciones y Mejoras del Método QM/MM

En los estudios QM/MM que se presentan en esta Tesis Doctoral el subsistema QM es tratado con un nivel semiempírico (*semiempirical Austin Model 1*, AM1).^[38] Esta descripción semiempírica puede ser bastante imprecisa en algunos casos y su uso requiere la previa

calibración del Hamiltoniano. Por supuesto, sería ideal usar un nivel más alto como el de orbitales moleculares *ab initio* o la teoría del funcional de la densidad (*Density Functional Theory*, DFT) para representar la parte reactiva del sistema. El problema es que estos cálculos son todavía demasiado costosos y consumen muchísimo tiempo de unidad central de proceso (*Central Processing Unit*, CPU). En sistemas grandes como son las enzimas, las exploraciones de la SEP (localización y caracterización de estructuras estacionarias) y las simulaciones de dinámica molecular pueden requerir miles o millones de evaluaciones de energía y gradiente. Como alternativa se podría reducir el tamaño de la región QM pero ello sería una fuente de error considerable ya que nuestro modelo no podría reproducir las posibles transferencias de carga dentro del sistema reactivo o incluso limitaría la descripción de la reacción química. Hay diferentes alternativas para mejorar la descripción semiempírica de la región QM con un coste computacional aceptable. En principio se podría utilizar parámetros específicos para la reacción estudiada (*Specific Reaction Parameters*, SRP),^[39] diseñados para reproducir resultados energéticos a mejor nivel. El problema de esta estrategia es que requiere una reparametrización previa y además ésta puede ser diferente para las diversas etapas que pueden aparecer en una misma reacción. En este trabajo vamos a incluir términos de corrección a la energía asociados a la coordenada de reacción elegida. Para utilizar estos términos de corrección necesitamos definir una coordenada de reacción. En este trabajo las reacciones estudiadas son transferencias de hidrógeno. En general para describir una transferencia se puede usar como coordenada de reacción (R_i) la combinación antisimétrica de las distancias que definen el enlace que se rompe y el que se forma. Esto es, la distancia al átomo dador

menos la distancia al átomo aceptor. A continuación describiremos los métodos de corrección utilizados en el presente trabajo.

2.3.1 Métodos de Enlace de Valencia Simple

El primer método de corrección utilizado para corregir la SEP es el *Simple Valence Bond* (SVB).^[40, 41] En este método la energía potencial total del sistema está dada por la siguiente ecuación.

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{mm}} + E_{\text{qm/mm}} + E_{\text{SVB},R_i} \quad (2.7)$$

Los tres primeros términos están explicados al principio de este capítulo mientras que el término E_{SVB,R_i} representa la corrección al término de energía semiempírica, E_{qm} , para un valor particular de la coordenada de reacción R_i . Este término de SVB viene definido como una función que depende de las distancias entre el hidrógeno y sus átomos dador y aceptor, así como la distancia entre el dador y aceptor,

$$E_{\text{SVB},R_i}(r_{\text{DH}}, r_{\text{DA}}, r_{\text{AH}}) = \frac{1}{2} [M_1(r_{\text{DH}}) + M_2(r_{\text{AH}})] - \frac{1}{2} ([M_1(r_{\text{DH}}) - M_2(r_{\text{AH}})]^2 + 4[V_{12}(r_{\text{DA}})]^2)^{1/2} + \Delta D_{\text{DH}} \quad (2.8)$$

donde r_{DA} es la distancia entre el átomo dador y el aceptor, y las funciones V_{12} y M_i vienen dadas por:

$$V_{12}(r_{\text{DA}}) = D_{\text{DA}} \exp \left[-\alpha_{12}(r_{\text{DA}} - r_{\text{DA}}^0) \right] \quad (2.9)$$

$$M_i(r_{\text{XH}}) = \Delta D_{\text{XH}} \left(\exp \left[-2\alpha_{\text{XH}}(r_{\text{XH}} - r_{\text{XH}}^0) \right] - 2 \exp \left[-\alpha_{\text{XH}}(r_{\text{XH}} - r_{\text{XH}}^0) \right] \right) \quad (2.10)$$

En la ecuación (2.10) XH representa los diferentes pares de átomos, dador-hidrógeno (DH) o aceptor-hidrógeno (AH) en los potenciales de Morse M_1 y M_2 , respectivamente, y r_{XH}^0 es la distancia de equilibrio del correspondiente enlace en una EsT. Los términos ΔD_{XH} se

calculan como la diferencia en energía de disociación de modelos que representen nuestra reacción entre un nivel alto de cálculo y el nivel bajo; α_{XH} se calcula como:

$$\alpha_{XH} = \left(\frac{k_{XH}}{2D_{XH}^0} \right)^{1/2} = u_{XH} \left(\frac{2\pi^2\mu}{D_{XH}^0} \right)^{1/2} \quad (2.11)$$

donde D_{XH}^0 es la energía de enlace, μ es la masa reducida de los átomos X y H, u_{XH} es el valor de la frecuencia del enlace XH y k es la constante de fuerza. Los parámetros D_{DA} y α_{12} que encontramos en la definición de la ecuación (2.9) son ajustadas para obtener la altura de la barrera deseada, la cual es la altura de la barrera calculada en fase gas con un Hamiltoniano de alto nivel.

En el caso de que se utilice más de una coordenada distinguida para seguir la reacción, la energía total incluirá tantos términos de SVB como sean necesarios:

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{mm}} + E_{\text{qm/mm}} + \sum_i E_{\text{SVB,R}_i} \quad (2.12)$$

2.3.2 Método de Correcciones Interpoladas

2.3.2.1 Correcciones Unidimensionales

El segundo método de corrección es el de las Correcciones Interpoladas (*Interpolated Corrections, IC*)^[42] basados en el uso de *splines* cúbicos tensionados (*Cubic Splines Under Tension, CSUT*).^[43-45] La energía potencial total es dada por la siguiente ecuación:

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{mm}} + E_{\text{qm/mm}} + \Delta E_{\text{IC}}(\text{R}_i) \quad (2.13)$$

El término de corrección es la diferencia entre la energía calculada con un método de alto nivel y la calculada con uno de bajo nivel

(semiempírico) y corrige la energía dada por la parte cuántica en el cálculo QM/MM. Esta diferencia se ajusta mediante el uso de *splines* cúbicos tensionados^[43, 44] con el objeto de hacerla continua y derivable.

$$\Delta E_{IC}(R_i) = \text{spline} \left[E_{QM}^{HL}(R_i) - E_{QM}^{LL}(R_i) \right] \quad (2.14)$$

Se seleccionan diferentes estructuras obtenidas en una SEP a lo largo de la coordenada R_i . Se elige un modelo que contiene el subsistema QM o parte de él para calcular la diferencia de energía entre un método de alto nivel y el semiempírico a lo largo de la coordenada de reacción R_i . Entonces, siguiendo el trabajo de Truhlar y colaboradores,^[43, 44] se hace uso de las funciones de *spline* cúbicas para interpolar su término de corrección a cualquier valor de R_i . De este modo se obtiene una función de R_i , con primeras y segundas derivadas continuas, necesarias para localizar puntos estacionarios y para la realización de simulaciones de dinámica molecular. En el caso de que haya más de una coordenada de reacción se seguirá el mismo protocolo pero utilizando las coordenadas de reacción a lo largo del camino de reacción. Así, en el caso de dos coordenadas la energía total vendrá dada por la ecuación siguiente:

$$E_{tot} = E_{qm} + E_{mm} + E_{qm/mm} + \Delta E_{IC}(R_1) + \Delta E_{IC}(R_2) \quad (2.15)$$

2.3.2.2 Correcciones Bidimensionales

Los métodos de corrección anteriores pueden ser extendidos al caso de más de una coordenada de reacción asumiendo que las correcciones aplicadas a cada una de ellas son independientes. Esto, sin embargo, no tiene carácter general. Las correcciones incluidas a lo largo de una coordenada de reacción pueden depender de forma importante de

los valores de otra. Por ello utilizamos otro método, generalización del anterior, que permite expresar los términos correctivos, como una función continua y derivable de dos coordenadas, utilizando *splines* bidimensionales.^[46]

Este método de corrección es una mejora cuando tenemos una reacción que tiene que ser estudiada utilizando dos coordenadas de reacción distinguidas, obteniendo una SEP bidimensional. Las Correcciones Interpoladas bidimensionales (IC2D) requieren un cálculo a dos niveles de estructuras localizadas en toda la SEP, en función de las dos coordenadas elegidas. La nueva SEP se obtiene mediante la función de energía potencial:

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{qm/mm}} + E_{\text{mm}} + \Delta E_{\text{IC2D}}(R_1, R_2) \quad (2.16)$$

donde,

$$\Delta E_{\text{IC2D}}(R_1, R_2) = \text{spline2D} \left[E_{\text{HL}}(R_1, R_2) - E_{\text{LL}}(R_1, R_2) \right] \quad (2.17)$$

y $E_{\text{XL}}(R_1, R_2)$ es el resultado del cálculo de la energía sin optimizar en fase gas para un modelo del subsistema QM con el nivel de cálculo alto (*High Level*, X=H) y el nivel bajo (*Low Level*, X=L) y que dependen de las dos coordenadas seleccionadas.

3 Localización de Puntos Estacionarios en Sistemas de Alta Dimensionalidad

Para localizar puntos estacionarios en una SEP existen diversos algoritmos. Algunos usan únicamente los valores de la energía, mientras que otros usan los gradientes de energía y las derivadas segundas. De cualquier modo, para caracterizar la naturaleza de los puntos estacionarios es necesario calcular el conjunto de constantes de fuerza del sistema, como ya se ha comentado anteriormente. La dificultad de los sistemas a estudiar en la presente Tesis Doctoral es el altísimo grado de dimensionalidad de las SEP (típicamente con valores del orden de 10^4 - 10^5 de grados de libertad). Esto hace totalmente inviable el empleo de algoritmos tradicionales de segundo orden de localización y caracterización de puntos estacionarios, debido al problema de calcular, almacenar y diagonalizar la matriz Hessiana resultante. [\[8, 47-50\]](#)

Para evitar el inconveniente de la gran cantidad de grados de libertad presentes en estos sistemas se puede utilizar una estrategia basada en la partición del conjunto total de coordenadas. Las coordenadas del sistema se dividen en un “espacio de control” (llamado también CORE) y un “espacio complementario”. Las coordenadas de los átomos que están directamente involucrados en la reacción química se incluyen en el CORE y el resto define el “espacio complementario”. Para localizar un punto estacionario se realiza el movimiento del espacio de control utilizando un algoritmo basado en el uso de primeras y segundas derivadas. Esto permite asegurar que la naturaleza del punto estacionario responde a los requisitos deseados en el subespacio considerado. A cada paso realizado en el espacio de control le sigue una

optimización completa del espacio complementario utilizando un algoritmo basado únicamente en el gradiente. Cuando se alcanza la convergencia la estructura obtenida es un punto estacionario (todos los elementos del vector gradiente son nulos) y además la matriz Hessiana del espacio de control contiene el número deseado de valores propios negativos en el espacio de control. Esta estrategia suele denominarse método de micro/macro iteraciones. (Figura 8).

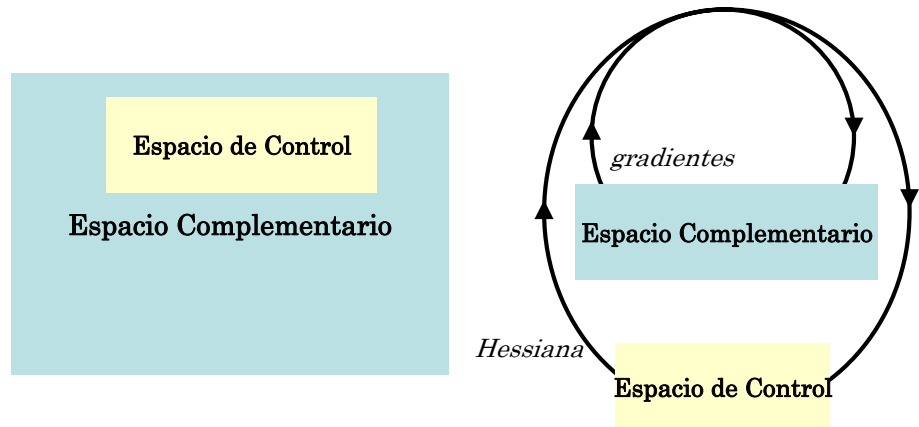


Figura 8. Representación esquemática del algoritmo de micro/macro iteraciones. El espacio de coordenadas se divide en dos subconjuntos: espacio de control y espacio complementario.

La combinación de programas CHARMM/GRACE, es uno de los paquetes utilizados para localizar puntos estacionarios en esta Tesis Doctoral. El programa GRACE^[51] permite llevar a cabo cálculos de álgebra de matrices utilizando los datos generados por el programa CHARMM^[52, 53] en el proceso de exploración de la superficie de energía potencial. GRACE calcula la hessiana numéricamente sobre el espacio de control a partir de los datos de posición, energía y gradiente que proporciona CHARMM. El resto de coordenadas, el espacio complementario, se minimiza en cada una de las etapas del cálculo con el programa CHARMM, de modo que el resultado final corresponda a un

punto estacionario. El método de optimización que incorpora GRACE está basado en los algoritmos *Eigenvector Following* (EF) y *Rational Function Optimization* (RFO),^[53] donde los vectores propios positivos en la matriz Hessiana son minimizados, mientras el vector propio negativo se maximiza.

Recientemente, en nuestro grupo de investigación, se ha realizado una actualización de la aplicación anterior bajo el nombre de GRACEFUL.^[54] Las microiteraciones del espacio control en GRACEFUL son guiadas por el algoritmo de BAKER.^[55, 56] Se introduce un control sobre pasos secuenciales de optimización, permitiendo que el algoritmo compruebe si se ha perdido o no el vector de búsqueda. GRACEFUL, escrito con el lenguaje de programación PYTHON^[57], se aplica al programa CHARMM, o cualquier programa que proporcione energías y gradientes de manera que además de implementar algoritmos más eficaces se convierte en una herramienta relativamente sencilla de utilizar.^[54]

4 Dinámica Molecular y Energías Libres

4.1 *Dinámica Molecular*

Cuando el proceso químico que se estudia ocurre en fase condensada es necesario tener en cuenta las múltiples configuraciones a las que puede acceder el sistema a una temperatura dada. La influencia del medio se traduce en la presencia de multitud de estados accesibles a bajo coste energético. Estos estados accesibles se agrupan en familias de conformaciones que se diferencian en pequeños cambios geométricos y energéticos que definen los estados de reactivos, estado de transición (ET) y estados de productos. Una de las consecuencias de este aumento de puntos estacionarios es la dependencia de los resultados obtenidos según las estructuras que se escojan como representativas de reactivos, ET y productos. Por tanto, las propiedades del sistema, tanto termodinámicas como cinéticas no se pueden obtener a partir del cálculo de las funciones de partición sobre el conjunto de niveles correspondientes a una única estructura estacionaria. Es esencial el cálculo de valores promedio de dichas propiedades sobre diferentes configuraciones que puedan ser comparados con datos experimentales.

La Dinámica Molecular (*Molecular Dynamic*, MD) es uno de los métodos empleados para explorar las distintas configuraciones de un sistema químico separadas por barreras de energía potencial de mayor o menor magnitud.^[58] La otra estrategia más extendida en su uso para

llevar a cabo exploraciones conformacionales moleculares está basada, en simulaciones de Monte Carlo,^[59] que hace uso de procesos estocásticos*.

La MD recurre a la solución de las ecuaciones del movimiento de las partículas. Hay varias formulaciones para el estudio dinámico de un sistema, sin embargo, cualquier descripción empieza con la definición del Hamiltoniano del sistema. En su forma clásica puede ser expresado como la suma de la energía cinética más la potencial en función de las coordenadas atómicas:

$$\hat{H}(\vec{p}_i, \vec{r}_i) = \sum_{i=1}^N \left(\frac{1}{2m_i} \vec{p}_i^2 + V(\vec{r}_i) \right) = \sum_{i=1}^N (K(\vec{p}_i) + V(\vec{r}_i)) \quad (4.1)$$

donde p_i es el momento cinético de la partícula i , y V es el potencial efectivo. Es posible derivar las ecuaciones de movimiento para las variables de la ecuación anterior utilizando las ecuaciones de Hamilton,

$$\dot{p}_i = -\frac{\partial \hat{H}}{\partial \vec{r}_i} = -\frac{\partial V}{\partial \vec{r}_i} = \vec{F}_i \quad (4.2)$$

$$\dot{r}_i = \frac{\partial \hat{H}}{\partial \vec{p}_i} = \frac{\vec{p}_i}{m_i} \quad (4.3)$$

La sustitución de la ecuación (4.2) en la (4.3) da lugar a la ecuación de Newton, que describe el movimiento nuclear de nuestro sistema.

$$\vec{F}_i(t) = m_i \vec{a}_i = m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} \quad (4.4)$$

Para estudiar la dinámica de un sistema, las ecuaciones de movimiento son resueltas para cada partícula del sistema. Una vez

* Este método genera aleatoriamente configuraciones del sistema utilizando un conjunto de criterios especiales para decidir si se acepta o no se acepta una nueva configuración, pero no proporciona de una manera explícita información dinámica acerca del sistema.

conocida la fuerza que actúa sobre un determinado núcleo $\vec{F}_i(t)$, así como su posición (\vec{r}_i) y velocidad (\vec{v}_i) en el instante t , es posible calcular la nueva posición que tendrá transcurrido un incremento finito de tiempo (Δt). Para ello hay que integrar la ecuación (4.4). De entre los diversos métodos se ha escogido el algoritmo Velocity-Verlet,^[58, 60] en el cual las ecuaciones para la nueva posición y la velocidad se expresan como:

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t \vec{v}_i(t) + \frac{\Delta t^2}{2m_i} \vec{F}_i(t) \quad (4.5)$$

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{\Delta t}{2m_i} (\vec{F}_i(t) + \vec{F}_i(t + \Delta t)) \quad (4.6)$$

Sea cual sea el algoritmo o método empleado para llevar a cabo las dinámicas, es importante que a lo largo de la dinámica se conserven ciertas propiedades del sistema, como el momento lineal total (\vec{M}) y el momento angular total (\vec{L}).

$$\vec{M} = \sum_{i=1}^N \vec{p}_i = \text{cte} \quad \vec{L} = \sum_{i=1}^N \vec{r}_i \wedge \vec{p}_i = \text{cte} \quad (4.7)$$

Una vez dispuestas las expresiones necesarias para tratar el movimiento de los núcleos, se escogen las condiciones en las que la simulación tendrá lugar. En función de cuales sean las variables que permanezcan constantes durante la simulación se obtendrá un observable u otro. Esta elección puede, en ocasiones, venir impuesta por la naturaleza del sistema que se esté estudiando. En general, en la mayoría de los procesos químicos existe conservación en el número total de partículas (N) que definen el sistema. Otro factor decisivo sobre las condiciones de trabajo será la disponibilidad de datos experimentales con los que comparar resultados. De entre las distintas opciones disponibles, una comúnmente adoptada es la que permite evolucionar al sistema

manteniendo el número de partículas (N), el volumen (V) y la temperatura constante (T), conocida termodinámicamente como el colectivo canónico o NVT. La razón de que normalmente se elija este colectivo se encuentra en que la mayoría de los procesos químicos en fase condensada tienen lugar bajo estas circunstancias.

Según la Termodinámica Estadística es conocido que las velocidades de traslación de los átomos en un sistema clásico se encuentran distribuidas de acuerdo a una distribución de Maxwell-Boltzmann. Entonces, si la temperatura del sistema es T , la probabilidad de que la velocidad del átomo i tenga un valor entre v_r y $(v_r + dv_r)$ es:

$$f(v_r)dv_r = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left(-\frac{m_i}{2k_B T} v_r^2\right) dv_r \quad r = x, y, z \quad (4.8)$$

Los valores de las velocidades iniciales de los átomos pueden ser asignados tratándolas como variables aleatorias *gaussianas* (*Gaussian random variables*) para la distribución definida en la ecuación (4.8), cuyo valor medio es cero y la desviación estándar $\sqrt{k_B T/m_i}$. Existe una expresión de la Termodinámica Estadística que relaciona la energía cinética del sistema con la temperatura:

$$T = \frac{2}{N_{df} k_B} \langle K \rangle \quad (4.9)$$

donde N_{df} corresponde al número de grados de libertad del sistema y $\langle K \rangle$ es el promedio de la energía cinética. Para una temperatura instantánea, $T(t)$, puede utilizarse la misma expresión pero eliminando el promedio, por lo tanto:

$$T(t) = \frac{2}{N_{df} k_B} K(t) \quad (4.10)$$

Esta expresión permite que la temperatura instantánea sea definida una vez las velocidades iniciales hayan sido elegidas.

Si se escoge una simulación bajo las condiciones de trabajo del colectivo canónico (NVT), para mantener constante la temperatura se hace uso de un baño termostático externo en contacto con el sistema, como el propuesto originalmente por S. Nosé.^[61, 62] Otro algoritmo para mantener la temperatura constante es el desarrollado por H. C. Andersen,^[63] quien sugirió que a intervalos de la simulación, las velocidades de un colectivo de núcleos elegidos aleatoriamente se reasignan haciendo uso de una distribución de Maxwell-Boltzmann, como consecuencia de una supuesta colisión elástica con las paredes del baño. Pero este algoritmo, a causa de la reasignación de las velocidades en los núcleos, provoca que la evolución de la posición de los mismos en el tiempo sea discontinua. Otro método introducido para evitar dichas discontinuidades, propuesto por Berendsen y colaboradores,^[64] incluye un término adicional al cálculo de las velocidades instantáneas (correspondientes a un tiempo t), asociado a una fuerza de fricción dependiente de la temperatura del baño que viene dada por la siguiente ecuación:

$$\bar{a}_i(t) = \frac{\bar{F}_i(t)}{m_i} + \bar{v}_i(t) \frac{1}{2\tau_i} \left(\frac{T_{\text{Baño}}}{T(t)} - 1 \right) \quad (4.11)$$

donde $T(t)$ es la temperatura instantánea definida en la ecuación (4.10). El término añadido a la ecuación de movimiento actúa como una fuerza de fricción. Cuando la temperatura del sistema es mayor que la temperatura del baño, la fuerza de fricción es negativa, y esto conlleva una disminución de la energía cinética y de la temperatura. En cambio, si la temperatura es menor que la temperatura del baño, la fuerza de

fricción es positiva, lo cual aporta energía al sistema aumentando la temperatura. De esta forma el control se ejerce automáticamente sobre todos los núcleos del sistema, obteniendo así trayectorias continuas en el tiempo.

Otro de los puntos importantes a tener en cuenta es qué valor del incremento del tiempo se debe escoger. El factor que limita el valor del incremento del tiempo es la naturaleza de los modos de vibración de mayor frecuencia del sistema. Se considera como criterio general que el Δt debe ser suficientemente pequeño como para que la vibración más rápida a describir (v_{\max}) pueda seguirse con suficiente detalle durante la integración de las ecuaciones de movimiento. En el caso de moléculas orgánicas, para un modo normal de vibración de alrededor de 3000 cm^{-1} (una vibración de *stretching* de un enlace carbono-hidrógeno) se obtendría un incremento de tiempo de unos pocos femtosegundos (fs). En la práctica, el incremento de tiempo escogido suele ser de 1fs para la mayoría de aplicaciones.

4.2 *Energías Libres en Termodinámica Estadística*

Cuando usamos mecánica cuántica para describir un sistema con N partículas ocupando un volumen V , el posible estado microscópico corresponde a una serie de estados propios ($\Psi_i(N, V)$) y sus correspondientes valores propios ($E_i(N, V)$). El número de estados microscópicos con la misma energía es conocido como degeneración de un nivel de energía, $\Omega_i(N, V, E_i)$. Para un sistema macroscópico, Ω_i es una función que aumenta con la energía y puede tomar valores muy grandes. Mientras el sistema esté en equilibrio termodinámico su estado

macroscópico permanecerá inalterado pero, en principio, el sistema puede ocupar todos los estados microscópicos compatibles con la descripción macroscópica. La probabilidad de encontrar el sistema en cada uno de los estados microscópicos durante su evolución dependerá de las restricciones impuestas en la evolución del sistema.

Para un sistema determinado podemos elegir el número de partículas, el volumen y la temperatura (N, V, T) , para especificar el estado termodinámico. Un colectivo de sistemas con NVT fijo se denomina el colectivo canónico y la probabilidad de encontrar un sistema en un estado microscópico particular i viene dado por:

$$P_i(N, V, T) = \frac{\exp\left(-\frac{E_i}{kT}\right)}{\sum_j \exp\left(-\frac{E_j}{kT}\right)} = \frac{\exp\left(-\frac{E_i}{kT}\right)}{Q(N, V, T)} \quad (4.12)$$

donde $Q(N, V, T)$ es la función de partición canónica, que puede ser alternativamente escrita como una suma sobre los niveles de energía en vez de sobre los estados microscópicos:

$$Q(N, V, T) = \sum_j \exp\left(-\frac{E_j}{kT}\right) = \sum_E \Omega(N, V, E) \exp\left(-\frac{E}{kT}\right) \quad (4.13)$$

En las expresiones previas hemos supuesto que uno puede sumar sobre diferentes niveles de energía, lo cual implica que la energía está cuantizada. Si usamos la mecánica clásica, las expresiones deben ser modificadas para integrar sobre una variación continua de la energía. Además, el estado microscópico no está ya especificado por números cuánticos. En su lugar, queda especificado por una serie compleja de coordenadas y momentos de todas las partículas $(r^N; p^N)$. Si la energía clásica es $H(r^N; p^N)$, la correspondiente expresión para una función de partición canónica es clásica:

$$Q(N, V, T) = C \int \dots \int \exp\left(-\frac{H(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N \quad (4.14)$$

donde $d\mathbf{r}^N = \prod_{i=1}^N dr_i$ y $d\mathbf{p}^N = \prod_{i=1}^N dp_i$

y C es una constante que tiene en cuenta dos principios cuánticos: el principio de exclusión de Pauli y el principio de incertidumbre de Heisenberg por lo que no es posible una derivación clásica pura.

$$C = (N! h^{3N})^{-1} \quad (4.15)$$

La probabilidad de encontrar un estado microscópico particular (esto es con coordenadas y momento en el rango $r^N + dr^N$ y $p^N + dp^N$) se define a través de su correspondiente densidad de probabilidad:

$$\rho_{NVT}(\mathbf{r}^N, \mathbf{p}^N) = \frac{\exp\left(-\frac{H}{kT}\right)}{\int \dots \int \exp\left(-\frac{H}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N} \quad (4.16)$$

Propiedades mecánicas (como la energía o la presión) tienen valores definidos en los estados microscópicos y sus correspondientes valores macroscópicos son simplemente el promedio sobre los estados microscópicos ocupados por el sistema durante el tiempo de la medida. De acuerdo con el principio ergódico esto es equivalente a un promedio sobre los estados microscópicos que aparecen en el colectivo. Así, en un conjunto canónico, una propiedad mecánica como la energía interna puede ser obtenida como:

$$U = \langle E \rangle = \int \dots \int \rho_{NVT} H(\mathbf{r}^N, \mathbf{p}^N) d\mathbf{r}^N d\mathbf{p}^N \quad (4.17)$$

Propiedades térmicas (entropía y energía libre) no pueden obtenerse de la misma manera ya que no están definidas para un estado

microscópico particular. De hecho, estas propiedades dependen de cómo el sistema evoluciona entre los diferentes estados microscópicos accesibles. Podemos obtener expresiones estadísticas de propiedades térmicas por comparación con la Termodinámica. La ecuación (4.17) puede ser fácilmente reordenada a:

$$U = - \left(\frac{\partial \ln Q(N, V, T)}{\partial 1/kT} \right)_{N, V} \quad (4.18)$$

y por comparación con la expresión $A=U-TS$, llegamos a:

$$A = -kT \ln Q(N, V, T) \quad (4.19)$$

La energía libre es difícil de calcular. Resulta mucho más sencillo de calcular la diferencia de energía libre entre dos estados. Hay tres métodos básicos para calcular diferencias entre energías libres: Perturbación de Energía Libre (*Free Energy Perturbation*, FEP), Crecimiento Lento (*Slow Growth*) e Integración Termodinámica (*Thermodynamic Integration*, TI); estando todas bastante relacionadas. A continuación se presentan estos métodos.

Debemos recordar que la energía libre de Helmholtz es el potencial termodinámico cuyas variables naturales independientes son las del conjunto canónico. En las siguientes secciones vamos a discutir sólo el caso para la energía libre de Helmholtz obtenida desde el conjunto canónico. Las energías libres de Gibbs y Helmholtz son diferentes cantidades para un mismo proceso, pero éstas tienen casi el mismo valor en las fases condensadas.

4.3 Método de Perturbación de la Energía Libre

Si consideramos dos estados I y II la diferencia de energía libre de Helmholtz entre ambos es:

$$\begin{aligned}\Delta A &= A(\text{II}) - A(\text{I}) = -kT \ln \frac{Q_{\text{II}}}{Q_{\text{I}}} = \\ &= -kT \ln \frac{\int \dots \int \exp\left(-\frac{H_{\text{II}}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N}{\int \dots \int \exp\left(-\frac{H_{\text{I}}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N}\end{aligned}\quad (4.20)$$

Podemos reordenar esta expresión quedando:

$$\begin{aligned}\Delta A &= \\ &= -kT \ln \int \dots \int \exp\left(-\frac{(H_{\text{II}} - H_{\text{I}})}{kT}\right) \rho_{\text{I}} d\mathbf{r}^N d\mathbf{p}^N = \\ &= -kT \ln \langle \exp[-(H_{\text{II}} - H_{\text{I}})/kT] \rangle_{\text{I}}\end{aligned}\quad (4.21)$$

La diferencia de energía libre entre los dos estados puede ser calculada a partir del valor promedio de la exponencial de la diferencia de energía entre ambos estados, evaluado sobre un conjunto de configuraciones de uno de ellos. El problema práctico aparece si durante la simulación de un estado dado, el espacio configuracional del otro estado no está correctamente muestreado. En ese caso, la cantidad $H_{\text{II}} - H_{\text{I}}$ sería muy grande y entonces las configuraciones seleccionadas contribuirían poco al promedio. No obstante, se puede introducir un número ilimitado de estados intermedios entre I y II y así calcular la diferencia de energía libre como la suma de pequeñas cantidades:

$$\begin{aligned}
 \Delta A &= A_{\text{II}} - A_{\text{I}} = (A_{\text{II}} - A_{N-1}) + (A_{N-1} - A_{N-1}) + \dots + (A_2 - A_{\text{I}}) = \\
 &= -kT \sum_{i=1}^{N-1} \ln \left\langle \exp[-(H_{i+1} - H_i)/kT] \right\rangle_i = \\
 &= kT \sum_{i=1}^{N-1} \ln \left\langle \exp[-(H_i - H_{i+1})/kT] \right\rangle_{i+1}
 \end{aligned}
 \tag{4.22}$$

Para definir estos estados, podemos introducir un parámetro de acoplamiento o perturbación, λ , que permite una transformación continua desde el Hamiltoniano del estado I al del estado II:

$$H(\lambda) = (1 - \lambda)H_{\text{I}} + \lambda H_{\text{II}}
 \tag{4.23}$$

donde $\lambda=0$ define el estado I y $\lambda=1$ al estado II. Es importante darse cuenta que el camino termodinámico trazado para ir del estado I al II no necesita ser un camino físico.

Una ventaja práctica del FEP es que las simulaciones llevadas a cabo por el estado I pueden ser usadas para calcular la diferencia de energía libre en ambos sentidos de la transformación química. Esta estrategia se denomina en inglés *Double Wide Sampling*.^[65] Idealmente, ambas cantidades deben ser equivalentes

$$\Delta G_{\text{I} \rightarrow \text{II}} = -\Delta G_{\text{II} \rightarrow \text{I}}
 \tag{4.24}$$

Los errores debidos a un muestreo incompleto introducen diferencias entre estos dos valores. Esta diferencia es una estimación razonable del error al calcular la diferencia de la energía libre, la cual se expresa como el promedio de las magnitudes del camino en el sentido directo e inverso.

4.4 Método de Integración Termodinámica

Si la energía libre de Helmholtz es una función continua de un parámetro, λ , podremos escribir:

$$\Delta A = \int \frac{\partial \Delta A(\lambda)}{\partial \lambda} d\lambda = \int \frac{\partial [-kT \ln Q(\lambda)]}{\partial \lambda} d\lambda = \int -\frac{kT}{Q(\lambda)} \frac{\partial Q(\lambda)}{\partial \lambda} d\lambda \quad (4.25)$$

A partir de la definición de la función de partición canónica, la derivada con respecto a λ es:

$$\frac{\partial Q(\lambda)}{\partial \lambda} = \frac{1}{N! h^{3N}} \int \dots \int \left(-\frac{1}{kT} \right) \frac{\partial H}{\partial \lambda} \exp\left(-\frac{H}{kT} \right) dr^N dp^N \quad (4.26)$$

Sustituyendo la expresión de la derivada en la expresión de la energía libre se obtiene:

$$\Delta A = \int \left[\int \dots \int \frac{\partial H}{\partial \lambda} \rho_{NVT} dr^N dp^N \right] d\lambda = \int \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (4.27)$$

Esta expresión permite obtener diferencias de energía libre a través del valor promedio de la derivada del Hamiltoniano con respecto a un parámetro de acoplamiento dado que define la transformación entre el estado inicial y el final. El valor promedio de la derivada se evalúa a diferentes valores de λ y finalmente la integración se realiza mediante técnicas numéricas estándar.

4.5 Método de Crecimiento Lento

La expresión del método de crecimiento lento (*Slow Growth*) de energía libre puede derivarse de la Perturbación de Energía Libre o de la Integración Termodinámica. Usando una expansión de Taylor para la ecuación (4.22) obtenemos:

$$\begin{aligned}
 \Delta A &= -kT \sum_{\lambda=0}^{\lambda=1} \ln \langle \exp[-(H(\lambda + \Delta\lambda) - H(\lambda))/kT] \rangle_{\lambda} \approx \\
 &\approx -kT \sum_{\lambda=0}^{\lambda=1} \ln \langle 1 - [(H(\lambda + \Delta\lambda) - H(\lambda))/kT] + \dots \rangle_{\lambda} = \\
 &= -kT \sum_{\lambda=0}^{\lambda=1} \ln \left\{ 1 - \frac{\langle H(\lambda + \Delta\lambda) - H(\lambda) \rangle_{\lambda}}{kT} + \dots \right\} \approx \sum_{\lambda=0}^{\lambda=1} \langle H(\lambda + \Delta\lambda) - H(\lambda) \rangle_{\lambda}
 \end{aligned}
 \tag{4.28}$$

En esta técnica, el parámetro de acoplamiento se cambia lentamente durante la simulación. Sin embargo, esta técnica ha sido criticada^[66] porque el sistema nunca tiene tiempo de estar equilibrado por el uso de un Hamiltoniano que es lenta, pero continuamente, variable. Además, si al final de la simulación se concluye que es necesario aumentar la exploración habría que repetir la simulación desde el principio. Con las técnicas de Perturbación de Energía Libre (*FEP*) o Integración Termodinámica (*TI*) se puede siempre añadir estados intermedios para recalculer la diferencia de energía, sin perder las simulaciones ya realizadas.

4.6 Método de Potenciales de Fuerza Media

4.6.1 Concepto de Potencial de Fuerza Media

Como hemos podido ver, la diferencia de energía libre entre dos estados I y II puede ser obtenida a partir de la relación entre las correspondientes funciones de partición:

$$\Delta A(I \rightarrow II) = -kT \ln \frac{Q_{II}}{Q_I} = -kT \ln \frac{\int \dots \int \exp\left(-\frac{H_{II}}{kT}\right) dr^N dp^N}{\int \dots \int \exp\left(-\frac{H_I}{kT}\right) dr^N dp^N}
 \tag{4.29}$$

En química frecuentemente estamos interesados en saber cómo cambia la energía libre en función de una coordenada del sistema en particular, ya sea una distancia, un ángulo, un ángulo diedro, una combinación de ellos, o bien una función de coordenadas cartesianas del sistema $\xi(\mathbf{r}^N)$. El cambio de la energía libre a lo largo de una coordenada particular es conocido como potencial de fuerza media (*Potential of Mean Force*, PMF) y lo presentamos como $W(\xi)$. En particular, propiedades de equilibrio conformacionales o la energía de activación de procesos activados puede ser expresada en términos del PMF.^[67, 68] Esta cantidad puede ser obtenida evaluando las integrales que aparecen en la expresión previa (4.29) para aquellas configuraciones del sistema que presentan un valor particular de la coordenada elegida. La selección de configuraciones puede llevarse a cabo formalmente introduciendo la función de Dirac (la cual restringe la integral a aquellas configuraciones con el valor deseado de la coordenada, ya que en otro caso ésta es igual a cero):

$$\Delta W(\xi_I \rightarrow \xi_{II}) = -kT \ln \frac{\int \dots \int \delta(\xi(\mathbf{r}^N) - \xi_{II}) \exp\left(-\frac{H}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N}{\int \dots \int \delta(\xi(\mathbf{r}^N) - \xi_I) \exp\left(-\frac{H}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N} \quad (4.30)$$

Esta cantidad puede ser alternativamente expresada como función particular de $\xi(\mathbf{r}^N)$ de la densidad de probabilidad de encontrar un valor particular en el sistema.

$$\begin{aligned} \langle \rho(\xi) \rangle &= \int \dots \int \rho_{NVT} \delta(\xi(\mathbf{r}^N) - \xi) d\mathbf{r}^N d\mathbf{p}^N = \\ &= \frac{\int \dots \int \delta(\xi(\mathbf{r}^N) - \xi) \exp\left(-\frac{H}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N}{\int \dots \int \exp\left(-\frac{H}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N} \end{aligned} \quad (4.31)$$

Introduciendo la función distribución en la ecuación (4.31) del PMF:

$$\Delta W(\xi_I \rightarrow \xi_{II}) = -kT \ln \frac{\langle \rho(\xi_{II}) \rangle}{\langle \rho(\xi_I) \rangle} \quad (4.32)$$

o, alternativamente:

$$W(\xi) = C' - kT \ln \langle \rho(\xi) \rangle \quad (4.33)$$

donde C' es una constante arbitraria. En principio, el PMF puede ser evaluado utilizando FEP o Integración Termodinámica, simplemente asignando el parámetro de acoplamiento a la coordenada ξ :

$$\xi = (1 - \lambda)\xi_I + \lambda\xi_{II} \quad (4.34)$$

Sin embargo, la ecuación (4.33) permite obtener el PMF a partir de la densidad de probabilidad de encontrar el sistema con un valor dado de la coordenada ξ . Durante una simulación la densidad de probabilidad puede ser evaluada determinando el número de veces que el sistema tiene un valor de la coordenada entre ξ y $\xi + \Delta\xi$: $N(\xi)$. Si el número total de configuraciones recogidas en la simulación es M , el histograma de la coordenada puede construirse como:

$$\langle \rho(\xi) \rangle \Delta\xi = \frac{\langle N(\xi) \rangle}{M} \quad (4.35)$$

El inconveniente de este procedimiento es que algunos valores de la coordenada ξ pueden no ser explorados o aparecer muy raramente, por lo que serían necesarios tiempos de simulación extraordinariamente largos para obtener valores convergidos del PMF.

4.6.2 Umbrella Sampling

Una de las técnicas más utilizadas para mejorar el muestreo de todas las configuraciones de interés es la conocida como *Umbrella Sampling*.^[69, 70] En este método la simulación se lleva a cabo en presencia de un potencial de sesgado adicional $V_{\text{umb}}(\xi)$, introducido para aumentar el muestreo en los alrededores de un valor particular de la coordenada ξ . Así, las simulaciones se generan utilizando una función nueva de energía:

$$H(\mathbf{r}^N, \mathbf{p}^N) + V_{\text{umb}}(\xi) \quad (4.36)$$

La distribución de la probabilidad modificada (*biased*) del sistema es:

$$\langle \rho(\xi) \rangle_{\text{biased}} = \frac{\int \dots \int \delta(\xi(\mathbf{r}^N) - \xi) \exp\left(-\frac{H + V_{\text{umb}}(\xi)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N}{\int \dots \int \exp\left(-\frac{H + V_{\text{umb}}(\xi)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N} \quad (4.37)$$

Esta distribución está relacionada con la del sistema sin modificar:

$$\langle \rho(\xi) \rangle_{\text{biased}} = \frac{\exp\left(-\frac{V_{\text{umb}}(\xi)}{kT}\right) \langle \rho(\xi) \rangle_u}{\left\langle \exp\left(-\frac{V_{\text{umb}}(\xi)}{kT}\right) \right\rangle_u} \quad (4.38)$$

donde el sufijo u indica el conjunto no perturbado (*unbiased*). El PMF está entonces relacionado con la distribución perturbada por la siguiente expresión:

$$W(\xi) = C' - kT \ln \langle \rho(\xi) \rangle_{\text{biased}} - V_{\text{umb}}(\xi) + F(\xi) \quad (4.39)$$

donde $F(\xi)$ es la energía libre asociada a la introducción del potencial de sesgado:

$$F(\xi) = -kT \ln \left\langle \exp \left(-\frac{V_{\text{umb}}(\xi)}{kT} \right) \right\rangle_{\text{u}} \quad (4.40)$$

El potencial se suele seleccionar de forma que restrinja la coordenada de interés dentro de un pequeño intervalo alrededor de un valor dado, ayudando a conseguir un mayor muestreo configuracional en un intervalo particular de valores de ξ que en condiciones normales no se exploraría. Estos intervalos de simulación se conocen como ventanas o *windows*. La forma del potencial *umbrella* es arbitraria. Una elección bastante común es una función armónica cuya forma es:

$$V_{\text{umb}}(\xi) = \frac{1}{2} K_{\text{umb}} \left(\xi - \xi_{\text{ref}}^{\alpha} \right)^2 \quad (4.41)$$

donde K_{umb} es la constante de fuerza y $\xi_{\text{ref}}^{\alpha}$ es el valor de referencia en la coordenada, cuyo valor se cambia en cada ventana α de simulación. De esta forma el intervalo entero de interés para la coordenada ξ se explora, no por una única simulación sino por una serie de ventanas.

4.6.3 Método de Análisis de Histogramas Ponderados (WHAM)

Se han propuesto diferentes aproximaciones para construir el potencial de fuerza media, $W(\xi)$, y la función de distribución, $\langle \rho(\xi) \rangle_{\text{u}}$, a partir de los histogramas de la coordenada ξ obtenida en las diferentes ventanas de simulación con *Umbrella Sampling*. El método de análisis

WHAM^[71] (*Weighted Histogram Analysis Method*) nos proporciona una forma óptima de combinar los datos obtenidos en las simulaciones. De acuerdo con este método, la función de distribución total se obtiene como una suma ponderada de las funciones de distribución de las N_w ventanas:

$$\langle \rho(\xi) \rangle_u = \sum_{\alpha=1}^{N_w} w^\alpha \langle \rho(\xi) \rangle_u^\alpha = \sum_{\alpha=1}^{N_w} w^\alpha \langle \rho(\xi) \rangle_{\text{biased}}^\alpha \exp\left(\frac{V_{\text{umb}}^\alpha - F^\alpha}{kT}\right) \quad (4.42)$$

Los pesos (w^α) cumplen la condición de normalización,

$$\sum_{\alpha=1}^{N_w} w^\alpha = 1 \quad (4.43)$$

y son introducidos con la condición de minimizar el error estadístico (σ^2) producido en la distribución de probabilidad total, esto es:

$$\frac{\partial \sigma^2[\rho(\xi)]}{\partial w^\alpha} = 0 \quad (4.44)$$

Se puede demostrar que los pesos que satisfacen estas condiciones anteriores son:^[71, 72]

$$w^\alpha = \frac{n^\alpha \exp\left(-\frac{(V_{\text{umb}}^\alpha(\xi)) - F^\alpha}{kT}\right)}{\sum_{\beta=1}^{N_w} n^\beta \exp\left(-\frac{(V_{\text{umb}}^\beta(\xi)) - F^\beta}{kT}\right)} \quad (4.45)$$

donde n^α es el número de los puntos de los datos independientes empleados para generar la función de distribución de la ventana α .

Para finalizar, es necesario estimar las constantes F^α :

$$\exp\left(-\frac{F^\alpha}{kT}\right) = \left\langle \exp\left(-\frac{V_{\text{umb}}^\alpha(\xi)}{kT}\right) \right\rangle_u = \int \dots \int \exp\left(-\frac{V_{\text{umb}}(\xi)}{kT}\right) \langle \rho(\xi) \rangle_u d\xi \quad (4.46)$$

Las ecuaciones (4.42) y (4.46) deben ser resueltas iterativamente porque tanto la función de distribución como las constantes F^α son desconocidas inicialmente. El procedimiento pasa por utilizar una serie de valores iniciales para F^α a partir de los que se evalúa $\langle \rho(\xi) \rangle_u$. Esta estimación de la función de distribución es entonces usada para determinar la constante de energía libre de la ventana repitiendo el proceso hasta llegar a convergencia.^[73, 74]

5 Teoría del Estado de Transición

La primera descripción de cómo varía la velocidad de una reacción química con la temperatura se debe a Arrhenius,^[75] quien en 1889, basándose en numerosas medidas experimentales y siguiendo el trabajo previo de van't Hoff,^[76] formuló la ecuación para la constante de velocidad de una reacción química que lleva su nombre:

$$k=Ae^{-E_a/RT} \quad (5.1)$$

donde A es el factor pre-exponencial y E_a es la energía de activación. La ecuación de Arrhenius, probablemente la más importante de la Cinética Química, es una ecuación macroscópica y estrictamente experimental.

La constante de velocidad k es el resultado promedio de la contribución de un número enorme de estados de características muy diferentes. Para entender cuales son los factores que determinan el valor de k, y por tanto de la velocidad de una reacción química, hay que entrar en el campo de la Dinámica de las reacciones químicas, el cual aporta una descripción microscópica de una reacción química y permite seguir el curso de cada colisión individual. En teoría, de la información que se obtiene de la Dinámica se podría calcular cualquier propiedad cinética, como la propia. Sin embargo, en la práctica, el enorme número de colisiones que habría que considerar para obtener un valor preciso de la constante hace inviable esta posibilidad para reacciones con más de 4 ó 5 núcleos, especialmente si existe una barrera energética, aunque sea pequeña. La única alternativa posible es la Teoría del Estado de Transición (*Transition State Theory*, TST)^[77-81], una aproximación estadística a la Dinámica formulada inicialmente en 1935 por Eyring, Evans y Polanyi.^[82-84] Desde entonces esta teoría ha recibido sucesivas

mejoras, de manera que ha llegado a ser probablemente la teoría más útil y extendida en el campo de la reactividad química, no sólo por su capacidad para predecir cuantitativamente la velocidad de las reacciones químicas con una exactitud razonable, sino también porque proporciona un marco de trabajo cualitativo mediante el cual es posible entender mejor la totalidad de las reacciones químicas. La TST, al igual que la ecuación de Arrhenius, proporciona una expresión de la constante de velocidad (una magnitud macroscópica) pero a partir de magnitudes moleculares. En los siguientes apartados de esta sección se exponen sucesivamente la teoría convencional del estado de transición (*Convencional Transition State Theory, CTST*),^[78, 80] la teoría variacional del estado de transición (*Variational Transition State Theory, VTST*),^[67, 85, 86] y la corrección que hay que introducir para tener en cuenta los efectos cuánticos en el movimiento de los núcleos a lo largo del camino de reacción.

5.1 Teoría Convencional del Estado de Transición

La teoría del estado de transición presupone que la aproximación de Born-Oppenheimer (B-O) es válida, es decir que los movimientos nuclear y electrónico pueden separarse por tener lugar en dos escalas de tiempo muy diferentes entre sí. La teoría convencional está ligada al concepto de estado de transición (ET), entendido como una ampliación del concepto de la estructura de transición (EsT) que se presenta en la sección 1 del presente capítulo. Así, el ET es el conjunto de estructuras que aparecen en una hipersuperficie en el espacio de las configuraciones nucleares, la llamada superficie de división, que pasa por

el punto de silla y divide la hipersuperficie de energía potencial de la reacción en dos partes: la hemihipersuperficie correspondiente a los reactivos y la hemihipersuperficie correspondiente a los productos. A partir de aquí se utilizará el término estado de transición (ET) como equivalente a la superficie de división, reservando el término estructura del estado de transición (EsT) para designar el punto de silla, que no es más que un punto particular de la superficie de división.

5.1.1 Teoría Canónica

La teoría canónica del estado de transición se aplica cuando los reactivos están en equilibrio térmico a una temperatura T , es decir, cuando las moléculas de reactivos se distribuyen entre todos los estados posibles según una distribución tipo Maxwell-Boltzmann. Para deducir la expresión de la constante de velocidad $k(T)$ se necesitan las siguientes dos hipótesis:

1) En la primera de ellas, que recibe el nombre de hipótesis fundamental, se supone que las supermoléculas que han cruzado el estado de transición en el sentido de reactivos a productos no pueden volver atrás (y formar otra vez los reactivos) y llegan finalmente a formar productos. Se entiende por supermolécula el resultado de considerar como un todo el conjunto de los núcleos y electrones de las moléculas que intervienen en la reacción.

2) La segunda hipótesis establece que en el estado de transición el movimiento a lo largo del camino de reacción (es decir, en la dirección que lleva desde los reactivos a los productos) puede ser separado de los demás movimientos y tratado como una traslación clásica.

Supongamos una reacción bimolecular entre los reactivos A y B. En el punto de silla se construye la superficie de división ortogonal al

camino de reacción. Hay que determinar la concentración de supermoléculas en el ET definidas como aquéllas que se encuentran en un pequeño entorno de la superficie de división. Por ejemplo, consideremos la concentración de supermoléculas que en un instante dado tienen configuraciones nucleares que se sitúan entre dos superficies paralelas a la superficie de división, la primera adelantada una distancia $\delta/2$ a lo largo del camino de reacción respecto a la superficie de división, y la segunda retrasada $\delta/2$, siendo δ una distancia arbitraria muy pequeña. Esto equivale a considerar una superficie de división con un grosor muy pequeño δ . En condiciones de equilibrio térmico de los reactivos, dicha concentración de supermoléculas en el estado de transición, N^\ddagger , puede calcularse a partir de la termodinámica estadística (el signo \ddagger se utiliza para marcar cualquier variable asociada al ET convencional):

$$K_{\text{tot}}^\ddagger = \frac{N^\ddagger}{[A][B]} = \frac{Q_{\text{tot}}^\ddagger}{Q_A Q_B} e^{-\Delta E_0/k_B T} \quad (5.2)$$

donde K_{tot}^\ddagger es la constante de equilibrio entre los reactivos y las supermoléculas del ET, Q_{tot}^\ddagger es la función de partición del estado de transición por unidad de volumen, Q_A y Q_B son las funciones de partición por unidad de volumen de los reactivos A y B respectivamente, ΔE_0 es la diferencia entre el nivel de más baja energía del estado de transición y el nivel de menor energía de los reactivos (es decir, la diferencia entre los respectivos niveles de energía de punto cero), y k_B es la constante de Boltzmann. El origen de energías de cada función de partición se sitúa en el nivel más bajo de cada especie, es decir en su nivel de punto cero.

Si se espera tiempo suficiente la reacción alcanza la situación de equilibrio entre reactivos y productos. En estas condiciones, en un instante dado existen dos tipos de moléculas en el estado de transición:

las que se mueven en el sentido de reactivos hacia productos (con una concentración $N_{R \rightarrow P}^\ddagger$) y las que se mueven en sentido contrario (con una concentración $N_{P \rightarrow R}^\ddagger$) de modo que:

$$N_{P \rightarrow R}^\ddagger + N_{P \rightarrow R}^\ddagger = N^\ddagger \quad (5.3)$$

Evidentemente, en esta situación de equilibrio la velocidad de la reacción directa debe ser igual a la velocidad de la reacción inversa y por tanto:

$$N_{P \rightarrow R}^\ddagger = N_{P \rightarrow R}^\ddagger = N^\ddagger / 2 \quad (5.4)$$

Supongamos que repetidamente se eliminan todos los productos. Si se cumple la hipótesis fundamental (no recruzamiento de la superficie de división), la ausencia de los productos no alterará en absoluto el valor de $N_{R \rightarrow P}^\ddagger$ (no es posible que ninguna molécula que contribuya a $N_{R \rightarrow P}^\ddagger$ provenga en realidad de los productos y esté contando dentro de $N_{R \rightarrow P}^\ddagger$ debido a una inversión del sentido de su trayectoria) y todas las supermoléculas que se incluyen en $N_{R \rightarrow P}^\ddagger$ llegarán a formar productos. Por tanto, la velocidad de la reacción (medida en moléculas por unidad de tiempo y unidad de volumen), es decir, la velocidad con la cual las supermoléculas atraviesan la superficie de división en el sentido adecuado, será:

$$\frac{dN}{dt} = \frac{N_{R \rightarrow P}^\ddagger}{\delta t} = \frac{N^\ddagger}{2\delta t} \quad (5.5)$$

donde δt es el tiempo medio que necesitan las supermoléculas para cruzar la superficie de división, el cual, a su vez, es igual a la anchura δ dividida por la velocidad media \bar{v} con la cual atraviesan el ET:

$$\delta t = \frac{\delta}{\bar{v}} \quad (5.6)$$

Suponiendo que las velocidades sigan una distribución de equilibrio térmico, éstas se podrán expresar como:

$$\bar{v} = \frac{\int_0^{\infty} v e^{-\frac{\mu v^2}{2k_B T}} dv}{\int_0^{\infty} e^{-\frac{\mu v^2}{2k_B T}} dv} = \left(\frac{2k_B T}{\pi \mu} \right)^{\frac{1}{2}} \quad (5.7)$$

donde v es la velocidad de cada supermolécula y μ es la masa reducida correspondiente al movimiento a lo largo del camino de reacción.

Si ahora se aplica la segunda hipótesis, la separabilidad de los movimientos conduce a:

$$Q_{\text{tot}}^{\ddagger} = Q_t Q^{\ddagger} \quad (5.8)$$

siendo Q_t la función de partición clásica correspondiente a un movimiento de traslación unidimensional que atraviesa la región del ET de anchura δ , y Q^{\ddagger} la función de partición para todos los restantes $3n-1$ grados de libertad en el estado de transición, donde n es el número total de núcleos de la supermolécula. De la termodinámica estadística se obtiene la expresión de Q_t :

$$Q_t = \frac{(2\pi\mu k_B T)^{\frac{1}{2}}}{h} \delta \quad (5.9)$$

donde h es la constante de Plank.

Combinando la ecuación (5.2) con las ecuaciones (5.5) a (5.9) se obtiene:

$$\frac{dN}{dt} = \frac{k_B T}{h} \frac{Q^{\ddagger}}{Q_A Q_B} e^{-\frac{\Delta E_0}{k_B T}} [A][B] \quad (5.10)$$

en la cual los parámetros indefinidos δ y μ ya no aparecen pues se han cancelado al combinar las ecuaciones.

Puesto que la ecuación de velocidad experimental (en moléculas por unidad de tiempo y unidad de volumen) para una reacción bimolecular es:

$$\frac{dN}{dt} = k(T)[A][B] \quad (5.11)$$

la comparación entre las ecuaciones (5.10) y (5.11) da la constante de velocidad canónica:

$$k(T) = \frac{k_B T}{h} \frac{Q^\ddagger}{Q_A Q_B} e^{-\frac{\Delta E_0}{k_B T}} \quad (5.12)$$

donde la constante de velocidad viene dada en s^{-1} .

Esta expresión puede escribirse alternativamente introduciendo una pseudoconstante de equilibrio entre el ET y reactivos donde se excluye la contribución del movimiento a lo largo del camino de reacción de la función de partición del ET:

$$k(T) = \frac{k_B T}{h} K_C^\ddagger \quad (5.13)$$

Si la constante de equilibrio se expresa en términos de energía libre de Gibbs estándar:

$$\Delta G_0^\ddagger = -RT \ln K_C^\ddagger \quad (5.14)$$

Podemos escribir la ecuación (5.14)

$$k(T) = \frac{k_B T}{h} e^{-\frac{\Delta G_0^\ddagger}{RT}} \quad (5.15)$$

5.2 Teoría Variacional del Estado del Transición

Tanto la teoría canónica del estado de transición^[78] como la microcanónica^[87] colocan la superficie de división en el punto de silla de la reacción dando lugar a la teoría convencional del estado de transición. De esta forma la constante de velocidad puede calcularse únicamente con información de la hipersuperficie de energía potencial de la reacción alrededor del punto de silla y de los reactivos, sin necesidad de conocer las características de la hipersuperficie en las regiones intermedias. Se consigue así un importante ahorro en cuanto a volumen de cálculo a costa de aceptar la hipótesis fundamental pero, al mismo tiempo, se exagera el valor de la constante de velocidad. La teoría variacional del estado de transición (VTST)^[88] es una generalización de la teoría convencional que elimina el requerimiento de que la superficie de división, sea el cuello de botella dinámico de la reacción; es decir, que la reacción tenga forzosamente que pasar por el punto de silla. Representa una clara mejora respecto a la teoría convencional, aunque requiere un esfuerzo computacional mucho mayor. La VTST se puede considerar tanto a partir de la versión canónica como de la versión microcanónica de la teoría del estado de transición.^[80] Vamos a desarrollar aquí la teoría variacional del estado de transición canónica, y al referirnos a ella utilizamos las siglas CVT.

El problema principal de la teoría convencional del ET es que no tiene en cuenta el recruzamiento de algunas trayectorias a través de la superficie de división que pasa por el punto de silla. La teoría variacional intenta mitigar el problema construyendo un conjunto de superficies de división generalizadas o estados de transición generalizados (*Generalized Transition State*, GTS), a lo largo de un camino de referencia que

normalmente es el IRC.^[67] El ET variacional será el que minimice el recruzamiento y por tanto exagere lo menos posible la constante de velocidad. De hecho el recruzamiento no se llega a eliminar por completo (para conseguir recruzamiento cero habría que construir superficies de división generalizadas en el espacio de fases, no sólo en el espacio de configuraciones). Por tanto, en lo que se refiere al recruzamiento, la teoría variacional sigue proporcionando un límite superior a la constante de velocidad, aunque el error es menor que en la teoría convencional.

En el caso de la teoría canónica, otra manera equivalente de abordar el problema es a través de la formulación termodinámica. La teoría convencional coloca la superficie de división en el punto de silla porque supone que allí está el cuello de botella dinámico de la reacción. De hecho, el punto de silla sólo es el punto del camino de reacción de más alta energía potencial. La constante de velocidad canónica no depende de la barrera de energía potencial sino de la energía libre de Gibbs de activación estándar. Por tanto, parece lógico que el cuello de botella dinámico para un reacción en equilibrio térmico lo establezca aquella superficie de división generalizada que haga máxima la energía libre de Gibbs de activación estándar (y por tanto minimice la correspondiente constante de velocidad), la cual, además de la contribución de la energía potencial, incluye las contribuciones de los niveles de punto cero, de las poblaciones térmicas de los niveles superiores y las contribuciones entrópicas.

Para clasificar los distintos estados de transición generalizados se utiliza la longitud del arco (s) a lo largo del camino de reacción de referencia, habitualmente el IRC. El valor $s=0$ se asigna al punto de silla, y los valores negativos o positivos corresponden al lado de los reactivos o de los productos, respectivamente. Cada superficie de división

generalizada en coordenadas cartesianas ponderadas se designa mediante el valor de s del punto en el cual intersecta al IRC. Además, en general, se requiere que la superficie divisoria que contiene a los GTS sea perpendicular al IRC en el punto de intersección. El GTS correspondiente a $s=0$ es precisamente el estado de transición convencional que da lugar a la constante de velocidad que dentro de la teoría canónica viene dada por la ecuación (5.12).

Para cada $s \neq 0$ se puede definir un GTS que intersecta al IRC en s , con lo que es posible definir una constante de velocidad generalizada para esta superficie de división generalizada. De todas las $k(T,s)$ se elige aquella que es menor, es decir, cuando $\Delta G_0^\ddagger(T,s)$ es máxima. Dicho de otra forma, se optimiza variacionalmente la posición del ET para minimizar los recruzamientos. Esta posición viene designada por $s^{CVT}(T)$, de modo que:

$$k^{CVT}(T, s^{CVT}) = \min \left[k(T, s) \right] \quad (5.16)$$

$$\Delta G_0^\ddagger(T, s^{CVT}) = \max \left[\Delta G_0^\ddagger(T, s) \right] \quad (5.17)$$

Por tanto, $s^{CVT}(T)$ indica la localización a lo largo del IRC del estado de transición canónico variacional para la temperatura T . Así definimos la constante variacional como:

$$k^{CVT}(T, s^{CVT}) = \frac{k_B T}{h} e^{-\frac{\Delta G_0^\ddagger(T, s^{CVT})}{RT}} \quad (5.18)$$

5.3 Efectos Cuánticos

La teoría del estado de transición es originalmente clásica. Sin embargo, la introducción de efectos mecano-cuánticos que dan lugar a la cuantización de los niveles de energía (y, especialmente, la existencia de las energías de punto cero) y de efectos mecano-cuántico sobre el movimiento a lo largo del camino de reacción es indispensable para obtener valores aceptables de las constantes de velocidad. En el apartado anterior se han considerado sólo los primeros, mientras que en este apartado podemos tener en cuenta la influencia de los segundos.

Hasta ahora hemos supuesto que una reacción sólo puede funcionar si los reactivos tiene una energía igual o superior al máximo de la curva de energía potencial adiabática, el cual corresponde al estado de transición canónico variacional a una temperatura de cero kelvin. Sin embargo, las propiedades ondulatorias de la materia permiten que la reacción pueda tener lugar mediante efecto túnel aún teniendo los reactivos una energía inferior a la de dicho máximo.^[89] Para tener en cuenta esta posibilidad se introduce en la teoría variacional el llamado coeficiente de transmisión semiclásico $\kappa^{\text{CVT}}(\text{T})$ ^[81] de tal modo que:

$$k^{\text{CVT}/\text{tun}}(\text{T}) = \kappa^{\text{CVT}}(\text{T})k^{\text{CTV}}(\text{T}) \quad (5.19)$$

La introducción del efecto túnel es absolutamente indispensable en reacciones que consisten en la transferencia de núcleos ligeros, como ocurre, por ejemplo, en la presente Tesis Doctoral donde tenemos reacciones de transferencia protónica y de hidruro.

Si suponemos que el camino de reacción tiene poca curvatura podemos aceptar que la corrección mecanocuántica al movimiento a lo largo del camino de reacción se debe fundamentalmente al túnel a través

de la región del punto de silla. Así, el movimiento cuántico de la supermolécula se describe aproximadamente mediante una ecuación de Schrödinger nuclear monodimensional correspondiente al movimiento de una partícula de masa reducida μ sobre un potencial dado por la curva de energía potencial adiabática en función de s . El operador Hamiltoniano es:

$$\hat{H} = \frac{-\hbar^2}{2\mu} \frac{d^2}{ds^2} + \hat{V}_a^G(s) \quad (5.20)$$

En coordenadas cartesianas ponderadas en peso las masas quedan incorporadas en las propias coordenadas por lo que la masa reducida de la ecuación (5.20) toma valor 1 sin dimensiones. El coeficiente de transmisión se calcula mediante el cociente entre la probabilidad de transmisión cuántica $P(E)$ promediada según una distribución de Maxwell-Boltzmann a la temperatura T , y la probabilidad de transmisión clásica promediada de la misma forma:

$$\kappa^{CVT/ZCT}(T) = \frac{\int_0^\infty P(E) e^{-\frac{E}{k_B T}} dE}{\int_{V^{AG}}^\infty e^{-\frac{E}{k_B T}} dE} \quad (5.21)$$

donde V^{AG} es el valor máximo de la energía potencial adiabática, es decir,

$$V^{AG} = V_a^G \left[s^{CVT}(T=0) \right] \quad (5.22)$$

En la integral del denominador se ha supuesto que la probabilidad de transmisión clásica es 0 para $E < V^{AG}$, pero 1 para $E \geq V^{AG}$. Los límites superiores de las integrales sólo deberían extenderse hasta la energía de disociación más baja de la molécula. Sin embargo, la presencia de la exponencial de la energía hace que los términos adicionales que se

añaden hasta el límite superior ∞ sean despreciables a las temperaturas que se consideran normalmente.

Una aproximación semiclásica a la probabilidad de transmisión cuántica permite llegar a las siguientes expresiones según un intervalo de energías en el cual esté comprendida la energía de la supermolécula:^[90]

$$P(E) = \begin{cases} 0, & E < E_0^{\text{tun}} \\ \frac{1}{1 + e^{2\theta(E)}}, & E_0^{\text{tun}} \leq E \leq V^{\text{AG}} \\ 1 - P(2V^{\text{AG}} - E), & V^{\text{AG}} \leq E \leq 2V^{\text{AG}} - E_0^{\text{tun}} \\ 1, & 2V^{\text{AG}} - E_0^{\text{tun}} < E \end{cases} \quad (5.23)$$

donde

$$E_0^{\text{tun}} = \max \left[V_a^G(s = -\infty), V_a^G(s = \infty) \right] \quad (5.24)$$

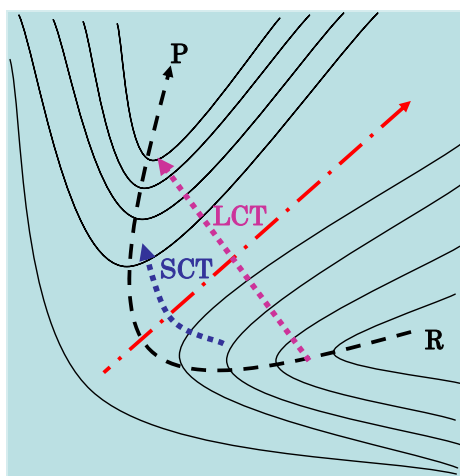
es la energía umbral cuántica y

$$\theta(E) = \frac{(2\mu)^{1/2}}{\hbar} \int_{s_1}^{s_2} \left[V_a^G(s) - E \right]^{1/2} ds \quad (5.25)$$

es la integral de acción siendo s_1 y s_2 los puntos de retorno clásicos, es decir, los puntos en los cuales $V_a^G(s) = E$. El segundo intervalo de la ecuación (5.23) corresponde al túnel, mientras que el tercero es el resultado de la reflexión cuántica. Como resultado de estas ecuaciones podemos decir que el túnel es mayor o igual que uno. Si la temperatura tiende a cero, entonces el coeficiente de transmisión tenderá a infinito; si la temperatura tiende a infinito, entonces, el coeficiente de transmisión (representado por $\kappa^{\text{CVT/ZCT}}(T)$) tiende a uno.

Si se quiere tener en cuenta la curvatura del camino de reacción en casos de curvatura moderada hay que introducir una masa reducida efectiva μ_{eff} en sustitución de μ . Debido a la curvatura $\mu_{\text{eff}} < 1$, lo que

provoca una disminución en la integral de acción y por tanto aumenta la probabilidad de transmisión cuántica y el factor de transmisión semiclásico (*Corner Cutting*).^[91, 92] Se obtiene entonces el método SCT, acrónimo de *Small-Curvature Tunneling*.^[93, 94] Sin embargo, si la curvatura del camino de reacción es muy grande hay que utilizar métodos bastante más apropiados como el LCT (*Large-Curvature Tunneling*).^[93, 95-97] Existe una metodología que optimiza el cálculo del efecto túnel (*Optimized Multidimensional Tunneling, OMT*)^[98, 99] que se basa en los métodos nombrados anteriormente y escoge el que da un



túnel mayor.

Figura 9. Representación esquemática del efecto túnel para una reacción de transferencia de hidruro. Las curvas negras son curvas isoenergéticas de la SEP. La figura muestra el camino de reacción entre reactivos (R) y productos (P). Se representa el túnel de curvatura pequeña (SCT) y el túnel de curvatura grande (LCT).

5.4 Teoría Variacional del Estado de Transición para Medios Enzimáticos

Incluir las fluctuaciones del entorno es muy importante para el estudio de procesos químicos que ocurren en fases condensadas. La reacción global que acontece en un sistema enzimático no transcurre a lo largo de un único camino de reacción de reactivos a productos atravesando una única estructura de transición o punto de silla (como sucede en reacciones en fase gas) (Figura 10). Al contrario, las reacciones son complejas e implican un incontable número de estructuras correspondientes al estado de reactivos, ET y estado de productos. Cada una de las estructuras de estos subconjuntos suelen diferir en pequeños cambios conformacionales en el entorno de la reacción.



Figura 10. Las reacciones enzimáticas pueden tener distintos valles de reactivos y productos que quedan unidos por distintas EsT, por lo tanto, tienen un gran número de caminos de reacción.

Truhlar y colaboradores^[100, 101] desarrollaron un método, llamado EA-VTST/OMT (*Ensemble-Averaged Variational Transition State Theory with Optimized Multidimensional Tunneling*), para incluir esta complejidad adicional de los procesos enzimáticos en el cálculo de la

constante de velocidad. La traducción al castellano del largo acrónimo que tiene este método sería Teoría Variacional del Estado de Transición Promediada en el Colectivo incluyendo el Efecto Túnel Multidimensional Optimizado. El método está diseñado para resolver las dificultades que presenta el cálculo de la constante de velocidad de una reacción catalizada por una enzima y permite incluir efectos cuánticos en la constante cinética de la reacción, por lo que es adecuada para procesos de transferencia de hidrógeno.

La metodología consta de dos etapas. En la primera etapa se calcula el perfil de energía libre de la reacción mediante un cálculo de PMF siguiendo una coordenada distinguida del sistema (z) que represente el progreso de la reacción química. En la segunda etapa se utiliza un conjunto de estructuras seleccionadas dentro de la zona del ET para determinar el coeficiente de transmisión y la contribución cuántica del efecto túnel.

En el primer paso de la primera etapa^[101] se calcula la energía libre en función de una coordenada de reacción (el PMF) de manera puramente clásica[†]. Las estructuras que aparecen a una distancia pequeña del máximo del PMF conforman el conjunto del ET (*the TS ensemble*). La energía libre de activación en este paso se obtiene como:

$$\Delta G_T^{(1,1)} = W^{CM}(T, z_*) - W^{CM}(T, z_R) + C(T, z_*) + W_{vib}(T, z_R) \quad (5.26)$$

donde $W^{CM}(T, z)$ es la energía clásica del PMF, z_* es el valor de z que tiene mayor energía en el perfil del PMF y z_R es el valor de z donde el PMF tiene el mínimo de energía y corresponde a reactivos. $C(T, z)$ es la corrección para una z curvilínea^[102] y $W_{vib}(T, z_R)$ es la contribución

[†] El término 'puramente clásica' se refiere a una dinámica nuclear, aunque el cálculo de la energía se realiza por medio del método híbrido QM/MM.

vibracional de la coordenada de reacción en el estado de reactivos, contribución no incluida en el PMF.

La energía libre vibracional del subsistema cuántico (tratada clásicamente en la dinámica molecular) debe corregirse para tener en cuenta la naturaleza discreta de los niveles vibracionales. Esto se realiza tomando estructuras de las dinámicas en las diferentes ventanas que constituyen el PMF y calculando las frecuencias, por diagonalización de la matriz Hessiana, utilizando un operador de proyección para eliminar el movimiento de la coordenada de reacción.^[100, 101] Las seis frecuencias menores restantes no se corrigen, admitiendo como correcto el tratamiento clásico (normalmente estos modos corresponden a pseudotraslaciones y pseudorotaciones del subsistema QM en el entorno MM). Para cada valor de la coordenada de reacción se hace un promedio de las frecuencias calculadas para las distintas estructuras que aparecen durante las dinámicas. Finalmente se obtiene la energía libre vibracional que corresponde a un valor dado de la coordenada de reacción, usando estas frecuencias promedio. La corrección a las vibraciones clásicas viene dado por:

$$\begin{aligned} \Delta W_{\text{vib}}^{3N-7}(T, z) &= \frac{1}{2} \hbar \sum_{m=1}^{3N-7} \langle \omega_m(z) \rangle + \\ &+ k_{\beta} T \sum_{m=1}^{3N-7} \ln(1 - e^{-\hbar \langle \omega_m(z) \rangle / k_{\beta} T}) - k_{\beta} T \sum \frac{\hbar \langle \omega_m(z) \rangle}{k_{\beta} T} \end{aligned} \quad (5.27)$$

Así, la energía libre de activación en el paso 2 vendría dada por:

$$\begin{aligned} \Delta G_{\text{T}}^{(1,2)} &= \left[W^{\text{CM}}(T, z_*) + C(T, z_*) + \Delta W_{\text{vib}}^{3N-7}(T, z_*) \right] - \\ &- \left[W^{\text{CM}}(T, z_{\text{R}}) + W_{\text{vib}}^z(T, z_{\text{R}}) + \Delta W_{\text{vib}}^{3N-7}(T, z_{\text{R}}) \right] \end{aligned} \quad (5.28)$$

La energía original del PMF con las correcciones vibracionales origina un ‘PMF semiclásico’ que proporciona la energía libre de activación $\Delta G_T^{(1,2)}$ correspondiente a la primera etapa.

En la segunda etapa se utilizan diferentes estructuras del colectivo del ET para calcular diferentes caminos de reacción y determinar el coeficiente de transmisión y la contribución de túnel en cada uno de ellos. El primer paso de esta segunda etapa sería elegir diferentes estructuras correspondientes a la zona del ET. Para cada estructura se traza un IRC manteniendo el entorno fijo y en cada uno de estos caminos se lleva a cabo la aplicación de la teoría variacional del estado de transición (VTST) aplicando el cálculo de efecto túnel multidimensional optimizado (VTST/OMT).^[103] El coeficiente de efecto túnel se calcula como un promedio sobre los diferentes caminos

$$\kappa^{(2)} = \left\langle \kappa_i^{(2)} \right\rangle \quad (5.29)$$

Finalmente, la constante de velocidad corregida quedará como:

$$k^{EA-VTST/OMT} = \kappa^{(2)}(T)k^{(1)}(T) \quad (5.30)$$

y la energía de activación:

$$\Delta G_T^{(2)} = \Delta G_T^{(1,2)} + \Delta G_{\text{tun}} \quad (5.31)$$

Por último, la correspondiente energía libre de corrección por el coeficiente de transmisión se obtendría como:

$$\Delta G_{\text{tun}} = k_B T \ln \kappa^{(2)}(T) \quad (5.32)$$

6 Enzimología

Tal como se ha indicado en el capítulo de Introducción, el presente trabajo representa un estudio teórico de dos reacciones enzimáticas: la LDH y la DHFR. Por ello, en este apartado haremos una breve introducción a los fundamentos fenomenológicos de la enzimología.

En todos los organismos vivos es preciso sintetizar macromoléculas a partir de moléculas más sencillas. La energía necesaria para el proceso se consigue rompiendo los enlaces químicos internos de otras moléculas, sustancias de reserva o alimentos. Todo ello comporta una serie de reacciones coordinadas cuyo conjunto se denomina metabolismo. Dado que las sustancias que intervienen en estas reacciones son generalmente muy estables se requeriría una gran cantidad de energía para que reaccionen entre sí. Para acelerar la reacción en un laboratorio bastaría con aumentar la temperatura o bien con añadir un catalizador. En los seres vivos, un aumento de la temperatura puede provocar la muerte, por lo que la Naturaleza opta por la otra posibilidad, el concurso de catalizadores biológicos o biocatalizadores. Las moléculas que desempeñan esta función son las enzimas.^[104]

Las enzimas son proteínas capaces de catalizar las reacciones metabólicas, las cuales difieren de la catálisis ordinaria en muchos aspectos: sus altas constantes de velocidad comparadas con reacciones no catalizadas; las relativamente suaves condiciones bajo las cuales tiene lugar la catálisis (i.e., temperatura sensiblemente menor de 100°C, presión atmosférica, pH prácticamente neutro, etc.);^[105] su gran especificidad, la cual les permite seleccionar reactivos y transformarlos en productos químicos bien definidos; y su capacidad de regulación en

respuesta a la concentración de otras sustancias. Las enzimas son macromoléculas de un peso molecular que oscila entre varios miles y varios millones de unidades atómicas, mientras que el sustrato sobre el cual actúa usualmente sólo pesa alrededor de varios cientos. En razón de esta diferencia de tamaño, sólo una fracción de la enzima entra en contacto con el sustrato, y la región donde se establece este contacto se llama sitio o centro activo de la enzima.^[106]

La actividad de algunas enzimas depende únicamente de su estructura como proteínas mientras que otras necesitan, además, uno o más componentes no proteicos llamados cofactores. El cofactor puede ser un ión metálico o bien una molécula orgánica llamada coenzima, aunque algunas enzimas necesitan ambos.^[107] Los coenzimas actúan, por lo general, como transportadores intermediarios de grupos funcionales, de átomos específicos o de electrones, los cuales son transferidos en la reacción enzimática global. El complejo enzima-cofactor catalíticamente activo recibe el nombre de holoenzima. Cuando la proteína se separa del cofactor, y queda por sí misma inactiva, se designa con el nombre de apoenzima.^[104, 106]

6.1 *Cinética Enzimática*

Para muchas enzimas, la velocidad de reacción del proceso catalizado, V , varía con la concentración de sustrato, $[S]$, de la forma que muestra la Figura 11. A concentración baja de sustrato, la velocidad inicial de la reacción es casi proporcional a la misma y la reacción es, por tanto, aproximadamente de primer orden. Sin embargo, a medida que la concentración de sustrato aumenta la velocidad inicial de la reacción deja

de ser directamente proporcional a la concentración de sustrato; en esta zona el orden de reacción es mixto. Finalmente se observa que a concentraciones elevadas de sustrato, la velocidad llega a ser independiente de ésta y se aproxima asintóticamente a una velocidad constante (V_{\max}). En este intervalo de concentraciones de sustrato la reacción es de orden cero y se dice entonces que la enzima está saturada con su sustrato.

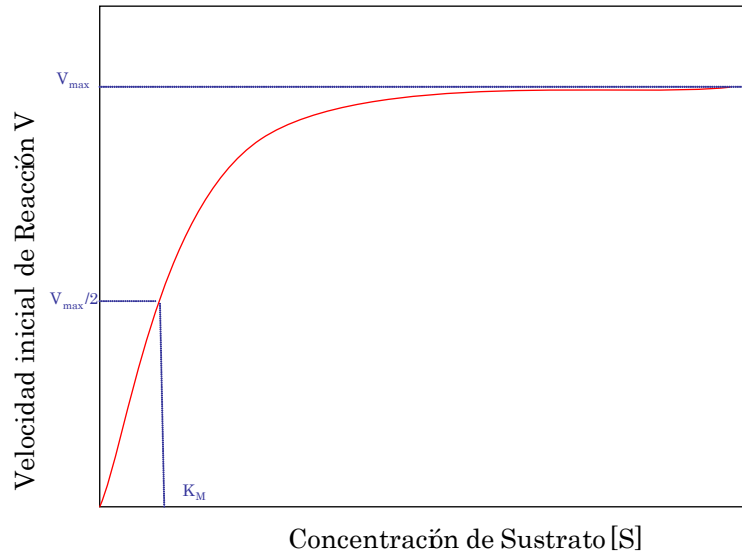
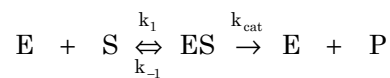


Figura 11. Representación de la velocidad inicial de reacción V , en función de la concentración de sustrato $[S]$, para una enzima que obedece la cinética de Michaelis-Menten. K_m es la constante de Michaelis.

En 1913, Leonor Michaelis y Maud Menten^[108] propusieron un modelo cinético sencillo que explica estas características. El modelo que propusieron es el más sencillo de los que pueden explicar las propiedades cinéticas de muchas enzimas:



(6.1)

Una enzima (E) se combina inicialmente con el sustrato (S) para formar un complejo enzima-sustrato (ES), el cual reacciona para formar productos (P) y la enzima libre, siendo $k_{\pm 1}$ las constantes de velocidad para los procesos directos e inversos de formación del complejo ES, y k_{cat} la constante de velocidad de la etapa de formación de productos que se supone irreversible.

La ecuación de Michaelis-Menten relaciona la velocidad de la reacción en la enzima y la concentración de sustrato,

$$V = \frac{V_{max}[S]}{K_M + [S]} \quad (6.2)$$

La representación gráfica de la V frente a [S] es la representada en la Figura 11. Cuando la concentración de sustrato es igual a K_M , V adquiere el valor de $V_{max}/2$ corresponde que se conoce como la constante de Michaelis Menten y viene dada por:

$$K_M = \frac{k_{-1} + k_{cat}}{k_1} \quad (6.3)$$

A partir de la ecuación (6.2) se puede deducir que la velocidad máxima viene determinada por:

$$V_{max} = k_{cat} [E]_0 \quad (6.4)$$

El concepto del complejo enzima-sustrato es la clave de la cinética enzimática. En honor a su introductora, este complejo suele ser llamado Complejo de Michaelis (Michaelis Complex, MC). Por supuesto que existe también un complejo enzima-producto (EP) pero en este mecanismo se supone que la disociación del EP es rápida y por lo tanto éste no se acumulará en concentración apreciable, permitiendo ignorar esta etapa.

Uno de los primeros estudios importantes sobre la especificidad de las enzimas fue el efectuado por Emil Fischer,^[109] quien descubrió que las enzimas que eran capaces de hidrolizar glucósidos podían distinguir entre las formas estereoisómeras de éstos. Esta observación le condujo a enunciar el principio de que la molécula del sustrato se adapta al centro activo de la enzima del mismo modo que lo haría una llave y una cerradura, es decir que tienen una relación complementaria. Unos años más tarde, J. B. S. Haldane^[104] sugirió la idea de que la enzima y el sustrato no son complementarios, sino que la enzima se une a una forma activada del sustrato. Esta idea fue propuesta previamente por Michael Polanyi,^[110] el cual sugirió que la aceleración de las reacciones químicas por parte de las enzimas podría depender de la especial afinidad de éstas por los sustratos no en sus formas normales, sino en sus formas distorsionadas, ya que agilizarían el proceso químico hacia productos. Unos 20 años más tarde, Linus Pauling^[111, 112] planteó la idea que una enzima está diseñada a través de la evolución para ser complementaria a la geometría de los reactivos en su estado de transición (ET), y no a la geometría de los reactivos en su estado fundamental.

El origen de la eficacia catalítica de las enzimas, en el aumento de la constante de velocidad ha sido, y es, motivo de numerosos debates destacando las siguientes hipótesis:

- **Efectos Electrostáticos:** ^[30, 113-115] La capacidad catalítica de las enzimas se atribuye a la estabilización electrostática de la estructura de transición, generada por los residuos que se encuentran alrededor del centro activo. Aparentemente las enzimas pueden crear micro-entornos, los cuales promueven, mediante su potencial electrostático, las transferencias de carga que tienen lugar durante las correspondientes reacciones. El dipolo creado por el entorno se encuentra preorientado para estabilizar las cargas de la EsT.^[113]

Existe, además, una reorganización mínima por parte de la enzima al pasar de reactivos hacia la EsT. En cambio, comparando la reacción catalizada con la misma reacción en disolución acuosa, la reorganización del disolvente desde reactivos al ET es mayor y esto lleva asociado un coste energético para orientar el disolvente hasta la configuración de la EsT.

- **Enlaces de Hidrógeno de Barrera Baja:** Un enlace de hidrógeno de barrera baja corresponde a un enlace de hidrógeno corto y fuerte entre dos bases o entre dos heteroátomos. El átomo de hidrógeno puede moverse libremente entre ambos, a través de enlaces de tipo covalente. La formación de este tipo de enlace no tienen lugar en disolución.^[116, 117] Por lo tanto, la estabilización energética del ET podría ser debida a este tipo de enlace y ser la causa de la catálisis enzimática.
- **Efecto *Near Attack Conformation* (NAC):**^[118, 119] Según esta hipótesis el estado de reactivos debe pasar a través de una conformación activada (NAC) para alcanzar el ET. El papel de la enzima sería disminuir la energía libre de activación mediante el incremento de la población de las estructuras de tipo NAC.
- **Factor de Tensión:** ^[120] La enzima puede inducir una tensión o una distorsión en un enlace de la molécula del sustrato, provocando una desestabilización del mismo en el MC y determinando que la ruptura de dicho enlace sea más fácil.
- **Desolvatación:**^[120-125] Aquellos centros activos enzimáticos no polares pueden catalizar reacciones por desolvatación de reactivos los cuales son estabilizados fuertemente en la correspondiente reacción en disolución.
- **Orientación orbital:**^[126-129] Desde el punto de vista de los orbitales moleculares, el sustrato y el grupo catalítico de la enzima no solamente deben hallarse próximos, sino también debe existir una orientación orbital adecuada. Es decir, debe haber una orientación

precisa del sustrato y del grupo catalítico de la enzima, uno respecto al otro, de modo que sus orbitales de enlace se alineen en la forma adecuada para la obtención del ET.

- **Factores entrópicos:**^[16, 120, 129-132] La enzima favorecería la reacción al reducir grados de libertad de reactivos restringiendo sus configuraciones de reactivos del tal forma que se fijarían en el centro activo mediante interacciones específicas. Dentro de estos factores podría encontrarse el término de Energía Crática, $\Delta G_{\text{crática}}$, definido como la energía libre necesaria para acercar y orientar los fragmentos que van a reaccionar.
- **Efecto Túnel:** ^[133, 134] Este efecto jugaría un papel importante en ciertos procesos enzimáticos como reacciones de transferencia de hidrógeno. La probabilidad de reacción por túnel en medio enzimático sería mucho mayor que para la reacción en disolución acuosa.
- **Efectos Dinámicos:** Según estos modelos, las enzimas pueden acelerar reacciones mediante fluctuaciones especiales que afectan al recruzamiento sobre la barrera energética y la difusión de reactivos o productos.^[135-137]

Posiblemente ninguno de estos factores puede, por sí solo, responder de la actividad catalítica completa de todas las enzimas, y quizás para cada enzima haya una combinación de factores específica, responsable de la aceleración global de la velocidad de reacción.

6.2 *Estado de protonación y reactividad*

Las enzimas son proteínas formadas por una secuencia ordenada de aminoácidos. Los aminoácidos son compuestos orgánicos que se caracterizan por poseer un grupo carboxilo y un grupo amino, además de tener una cadena, R, que los diferencia. Podemos clasificar los

aminoácidos según el grupo que tengan en la cadena lateral R en neutros, ácidos (cargados negativamente a pH fisiológico) y básicos (cargados positivamente a pH fisiológico).

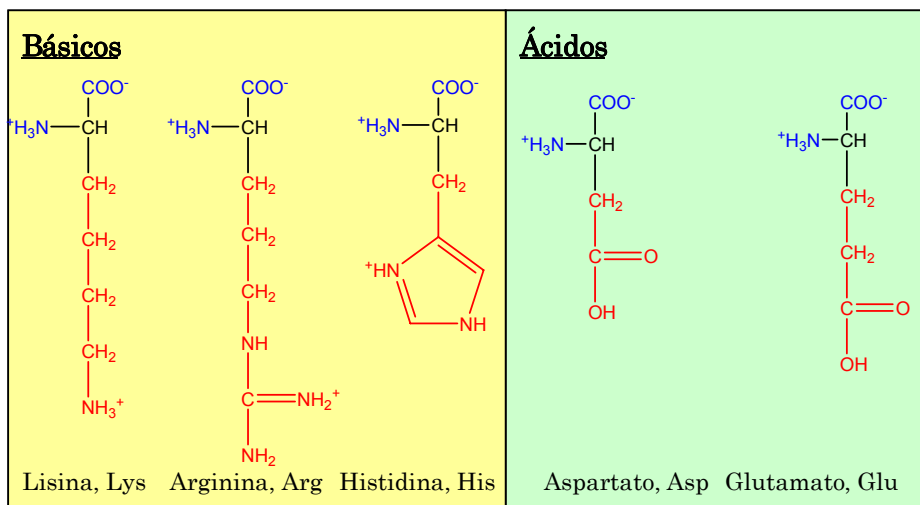


Figura 12. Forma de los aminoácidos con carácter básico y los aminoácidos con carácter ácido a pH 7. Se representa en color rojo la cadena lateral R, en negro el carbono alpha, en azul el grupo carboxilo y el grupo amino. El resto de aminoácidos tienen un comportamiento neutro aunque se pueden clasificar paralelamente como hidrófobos o polares.

La información estructural proveniente de técnicas de difracción de rayos X no suele ser capaz de discriminar el estado de protonación de los distintos aminoácidos. Su estado a un determinado pH, suele determinarse en base al valor del pK_a del aminoácido en disolución. Sin embargo, en el entorno proteico este pK_a puede desplazarse varias unidades debido a la influencia del entorno local en cada uno de los aminoácidos. En algunos casos es posible deducir experimentalmente el estado de protonación de algún aminoácido clave en la catálisis a partir de la dependencia de las constantes cinéticas con el pH.^[138-140] En otros casos es necesaria una aproximación teórica para calcular el pK_a del aminoácido en el entorno proteico. Aunque existen diversas técnicas para estimar el pK_a de un aminoácido,^[141] en el presente trabajo hemos

utilizado el método desarrollado por Gilson y colaboradores^[142] e implementado por Field y colaboradores^[143] que se basa en la dependencia del estado de protonación con la energía de interacción electrostática de las especies cargadas que aparecen en el equilibrio ácido-base. Así, el pK_a de un residuo de una proteína ($pK_{a,prot}$) se puede calcular a partir del pK_a en disolución acuosa ($pK_{a,dis}$) del aminoácido, perturbado por el entorno proteico, considerando la diferencia de energías libres electrostáticas entre ambos entornos:

$$pK_{a,prot} = pK_{a,dis} - \gamma \frac{\Delta\Delta G_i}{\ln(10)RT} \quad (6.5)$$

donde $pK_{a,prot}$ es el pK_a del residuo en una proteína neutra, es decir, a su alrededor no hay residuos cargados, el término γ es -1 para un aminoácido con el grupo $-R$ ácido y $+1$ para un aminoácido básico, y $\Delta\Delta G_i$ es la diferencia entre la energía libre electrostática de ionización del aminoácido i en la proteína con el resto de residuos neutros y el aminoácido en disolución acuosa. Podemos obtener la energía libre electrostática en agua y en la proteína a partir de las expresiones siguientes:^[143]

$$\Delta G_i^{dis} = \frac{1}{2} z_i^2 \phi(1,1) + z_i \sum_{a=1}^M q_a \phi(1,a) \quad (6.6)$$

$$\Delta G_i^{prot} = \frac{1}{2} z_i^2 \psi(1,1) + z_i \sum_{a=1}^N q_a \psi(1,a) \quad (6.7)$$

donde $\phi(i,a)$ y $\psi(i,a)$ son los potenciales electrostáticos creados por una unidad de carga situada en el átomo a sobre el centro i , en disolución y en la proteína respectivamente, M es el número de átomos en el residuo, N es el número de átomos en la proteína y q_a es la carga del átomo a . Los

cálculos electrostáticos se realizan mediante el método implementado dentro del programa UHBD^[144, 145] modelizando el disolvente como un continuo. En el caso de más de un grupo ionizable alrededor del residuo la determinación del pK_a en la proteína se complica ya que hay que tener en cuenta la interacción electrostática entre los residuos cargados. Esta energía libre electrostática se puede calcular como:

$$G_{\text{elec},X} = \sum_{a=1}^N X_i \Delta \Delta G_i + \sum_{1 \leq i \leq j \leq N} X_i X_j z_i z_j \psi(i, j) \quad (6.8)$$

donde $\Delta \Delta G_i$ es la energía de ionización de transferencia de un residuo i en disolución acuosa a la molécula. X_i vale 1 cuando el residuo está ionizado y vale 0 cuando el residuo no está ionizado; y z_i es la carga del átomo ionizado.

La ecuación (6.8) da la expresión de la energía que hay que aportar para pasar una molécula de un estado totalmente desprotonado al estado protonado para los N residuos ionizables. La probabilidad de encontrar al residuo i en un estado particular de protonación se obtiene a partir de la introducción de los correspondientes pesos de Boltzmann:

$$\theta_i = \frac{\sum_X X_i \exp(-\beta G_{\text{elec},X} - \ln(10)v(X)pH + \ln(10) \sum_j X_j pK_a(j))}{\sum_X \exp(-\beta G_{\text{elec},X} - \ln(10)v(X)pH + \ln(10) \sum_j X_j pK_a(j))} \quad (6.9)$$

donde $v(X)$ es el número de protones adjudicado a un residuo para conseguir su estado protonado y $\beta = 1/RT$.

La consideración de las interacciones entre residuos ionizables requiere un esfuerzo computacional significativo ya que para n grupos ionizables debemos calcular 2^n estados de ionización. Este número puede ser extraordinariamente grande en el caso de las proteínas, por lo que se requiere introducir aproximaciones para determinar el estado de

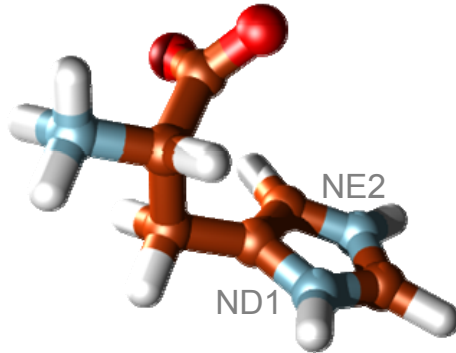
protonación. La primera de ellas consiste en suponer que hay un estado de protonación predominante para algunos residuos, y la segunda hacer una aproximación de campo medio combinada con un algoritmo de cluster.^[143]

La aproximación del estado de protonación predominante está basado en el hecho de que si el $pK_{a,dis}$ de un residuo está lejos del pH de la disolución, entonces el estado de ionización para este residuo puede considerarse fijo. Así, el número de residuos ionizables con estado de protonación a determinar se reduce. Para localizar estos residuos se calcula la diferencia de energía libre de ionización máxima y mínima para un residuo que interacciona con los otros sitios.

Para el resto de residuos se utiliza un algoritmo de cluster y los residuos son agrupados de acuerdo con su interacción. Las interacciones dentro del cluster se tratan explícitamente mientras que interacciones entre clusters se calculan usando la aproximación del campo medio. En la aproximación de cluster, propuesta por Gilson,^[142] la influencia de un aminoácido sobre otro es proporcional a la concentración de aminoácidos ionizados sin que tengan una posible correlación carga-carga entre ellos. En la aproximación de campo medio, se calcula la diferencia de energía libre entre estados de protonación medios.

Además de los posibles estados de protonación, es importante tener en cuenta el caso particular de las histidinas. Los residuos histidina tienen dos lugares de protonación en su estado neutro. Estos residuos pueden tener el protón localizado en el átomo ND1 o NE2 (Figura 13). Normalmente, las histidinas neutras presentan un protón en el ND1 mientras que el NE2 permanece desprotonado. Sin embargo, dado que la diferencia de pK_{as} es muy pequeña, la interacción con el entorno puede cambiar esta circunstancia. Para discernir entre las dos

posibilidades primero hay que hacer un cálculo del pK_a de estos residuos para saber si están en su forma neutra o protonada. En el caso de que estén neutros se calculan los potenciales electrostáticos sobre los átomos



ND1 y NE2. El protón se situará en aquel átomo con menor potencial.

Figura 13. Representación de una histidina protonada en los dos nitrógenos (ND1 y NE2).

6.3 *Mutagénesis*

La mutación de una enzima implica que al menos uno de los aminoácidos que componen su estructura sufre un cambio por otro aminoácido. Un cambio de una pequeña unidad en una enzima puede tener consecuencias impredecibles. En general, la mutación de un aminoácido que se encuentre relativamente alejado del centro activo no tiene que implicar un cambio en el comportamiento cinético de la enzima, aunque puede afectar a su estructura y con ella su función.^[146-148] Sin embargo, cambiar algún aminoácido que forme parte del centro activo de

la enzima puede afectar a la etapa de la reacción química e incluso puede desactivar la enzima.

El motivo de hacer mutaciones en el seno de una enzima es conocer a fondo la importancia de determinados aminoácidos; si una mutación conduce a una función catalítica significativamente reducida implica que el aminoácido mutado realizaba un papel fundamental para la catálisis. Esta es una práctica de extrema importancia conocida como mutagénesis dirigida (*site directed mutagenesis*) que se puede simular con las técnicas de la Química Teórica Computacional. En muchas reacciones enzimáticas los residuos del centro activo están directamente implicados en la reacción química: pueden participar como dadores o aceptores de átomos, formar enlaces con el sustrato durante la reacción y romperlos para dar el producto final, o romper enlaces del sustrato quedándose el aminoácido unido a un fragmento de éste. El resto de aminoácidos de la enzima puede jugar un papel importante en la catálisis sin participar directamente en la reacción química: estos aminoácidos pueden realizar el papel de polarizar enlaces, anclar el sustrato en el centro activo de manera que el producto sea un esteroisómero concreto, permitir o evitar el desplazamiento de algún bucle de la proteína, etc.

Tanto desde un punto de vista experimental como teórico, existe una gran variedad de posibilidades para elegir la mutación que se quiere estudiar, pero en ambos casos hay ciertas reglas a seguir de manera que se minimice la reorganización de la estructura de la enzima, tanto local como globalmente. La reorganización o distorsión de la estructura vendría acompañada de cambios energéticos difícilmente predecibles que complicarían el análisis de los cambios debidos a la mutación. Estas reglas son: [\[149\]](#)

1. Elegir una mutación que borre parte de una cadena. Borrar parte de una cadena es preferible a aumentar el volumen de ésta, especialmente en el interior de la enzima. Cualquier incremento en el volumen de una cadena implica directamente una distorsión en la estructura de la enzima, mientras que pequeñas cavidades dentro de la enzima son tolerables.
2. Evitar romper parejas de cargas. Borrar un grupo que solvata una carga es peligroso. Las energías de solvatación de iones son demasiado grandes a no ser que haya acceso abierto para que el disolvente se ocupe de solvatarlo.
3. Eliminar el mínimo número de interacciones. Es bastante complicado analizar el cambio de una única interacción, por lo que conviene evitar la eliminación de interacciones múltiples.
4. No añadir grupos funcionales que tengan comportamiento muy diferente a las cadenas sustituidas. La adición de un nuevo grupo en la cadena puede causar la total reorganización de la estructura si el nuevo grupo presenta nuevas interacciones.
5. Obviamente, todas estas reglas pueden ser desobedecidas cuando se crea apropiado.

En general, la mutación ideal es aquella en la que se elimina una interacción sin causar una 'disrupción' o reorganización importante de la estructura de la enzima. [\[149\]](#)

7 Bibliografía

- [1] P. Skancke, *Acta. Chem. Scand.* **1993**, *47*, 629.
- [2] K. Jug, *Theor. Chim. Acta* **1980**, *54*, 263-300.
- [3] Z. Havlas and R. Zaradnik, *Int. J. Quantum Chem.* **1984**, *26*, 607-619.
- [4] V. I. Minkin, B. Y. Simkin and R. M. Minyaev, *Quantum Chemistry of Organic Compounds*, New York, **1990**.
- [5] M. Born and R. Oppenheimer, *Ann. Phys. (Leipzig)* **1927**, *84*, 457-484.
- [6] P. G. Mezey, *Potencial energy hypersurfaces*, Amsterdam, **1987**.
- [7] J. E. Garcia, *Superficies de Energía Potencial y Reactividad Química*, Setem, Oviedo, **2001**.
- [8] P. G. Mezey, *Theoret. Chim. Acta (Berl.)* **1980**, *54*, 95.
- [9] J. M. Lluch in *Dinámica de las Reacciones Químicas II: Teoría del Estado de Transición*, Vol. 2, Universitat Jaume I, Castellon, **2000**, p. 314-349.
- [10] E. A. Mccullough and D. M. Silver, *J. Chem. Phys.* **1975**, *62*, 4050-4052.
- [11] E. Kraka and H. B. J. Dunning, *Advances in molecular electronic structure theory: the calculations and characterization of molecular potential energy surfaces*, Greenwich, C. N., **1989**.
- [12] K. Fukui, S. Kato and H. Fujimoto, *J. Am. Chem. Soc.* **1975**, *97*, 1-7.
- [13] K. Fukui, *Acc. Chem. Res.* **1981**, *14*, 363-368.
- [14] J. Bertran, V. Branchadell, M. Moreno and M. Sodupe, *Química Cuántica. Fundamentos y Aplicaciones Computacionales*, Editorial Síntesis, Madrid, **2000**.
- [15] R. D. Levine, *Quantum Mechanics of Molecular Rate Processes*, Oxford, **1969**.
- [16] P. A. Kollman, B. Kuhn, O. Donini, M. Peräkylä, R. V. Stanton and D. Bakowies, *Acc. Chem. Res.* **2001**, *34*, 72-79.
- [17] A. Warshel and M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227-249.

- [18] U. C. Singh and P. A. Kollman, *J. Comput. Chem.* **1986**, *7*, 718-730.
- [19] M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700-733.
- [20] A. J. Mullholland, G. H. Grant and W. G. Richards, *Prot. Eng.* **1993**, *6*, 133-147.
- [21] J. Gao, *Acc. Chem. Res.* **1996**, *29*, 298-305.
- [22] R. Castillo, V. Moliner and J. Andrés, *J. Am. Chem. Soc.* **1999**, *121*, 2140-2147.
- [23] G. Monard and K. M. Merz, *Acc. Chem. Res.* **1999**, *32*, 904-911.
- [24] C. Alhambra, J. Gao, J. C. Corchado, J. Villà and D. G. Truhlar, *J. Am. Chem. Soc.* **1999**, *121*, 2253-2258.
- [25] Y. K. Zhang, H. Y. Liu and W. T. Yang, *J. Chem. Phys.* **2000**, *112*, 3483-3492.
- [26] T. C. Bruice and K. Kahn, *Curr. Opin. Chem. Biol.* **2000**, *4*, 540-544.
- [27] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Theor. Chem. Acc.* **2001**, *3*, 207-212.
- [28] M. García-Viloca, A. Gonzalez-Lafont and J. M. Lluch, *J. Am. Chem. Soc.* **2001**, *123*, 709-721.
- [29] J. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.* **2002**, *53*, 467-505.
- [30] S. Martí, M. Roca, J. Andres, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Chem. Soc. Rev.* **2004**, *33*, 98-107.
- [31] O. Modesto and F. J. Luque, *Chem. Rev.* **2000**, *100*, 4187-4225.
- [32] K. E. Ranaghan, L. Ridder, B. Szeferczyk, W. A. Sokalski, J. C. Hermann and A. J. Mulholland, *Org. Biomol. Chem.* **2004**, *2*, 968-980.
- [33] S. Hammes-Schiffer, *J. Chem. Phys.* **1996**, *105*, 2236-2246.
- [34] N. Reuter, A. Dejaegere, B. Maigret and M. Karplus, *J. Phys. Chem. A* **2000**, *104*, 1720-1735.
- [35] V. Thery, D. Rinaldi, J.-L. Rivail, B. Maigret and J. J. Ferenczy, *J. Comput. Chem.* **1994**, *15*, 269-282.

- [36] G. Monard, M. Loos, V. Thery, K. Baka and J.-L. Rivail, *Int. J. Quantum Chem.* **1996**, *58*, 153-159.
- [37] J. Gao, P. Amara, C. Alhambra and M. J. Field, *J. Phys. Chem. A* **1998**, *102*, 4714-4721.
- [38] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.
- [39] A. Gonzalez-Lafont, T. N. Truong and D. G. Truhlar, *J. Chem. Phys.* **1991**, *95*, 8875-8894.
- [40] L. S. Devi-kesavan, M. Garcia-Viloca and J. Gao, *J. Theor. Chem. Acc.* **2003**, *109*, 133.
- [41] T. D. Poulsen, M. Garcia-Viloca, J. Gao and D. G. Truhlar, *J. Phys. Chem. B* **2003**, *107*, 9567-9578.
- [42] J. J. Ruiz-Pernia, E. Silla, I. Tuñón, S. Martí and V. Moliner, *J. Phys. Chem. B* **2004**, *108*, 8427-8433.
- [43] Y. Chuang, J. C. Corchado and D. Truhlar, *J. Phys. Chem. A* **1999**, *103*, 1140-1149.
- [44] J. C. Corchado, E. L. Coitiño, Y. Chuang, P. L. Fast and D. Truhlar, *J. Phys. Chem. A* **1998**, *102*, 2424-2438.
- [45] K. A. Ngunyen, I. Rossi and D. G. Truhlar, *J. Chem. Phys.* **1995**, *103*, 5522.
- [46] F. Proust-De Martin, R. Dumas and M. J. Field, *J. Am. Chem. Soc.* **2000**, *122*, 7688-7697.
- [47] S. S. Shaik, H. B. Schlengel and S. Wolfe, *Theoretical aspects of physical organic chemistry*, New York, **1992**.
- [48] B. D. Bunday, *Basic Optimization Methods*, London, **1984**.
- [49] R. Fletcher, *Practical Methods of Optimization*, New York, **1981**.
- [50] X. Prat-Resina, M. Garcia-Viloca, G. Monard, A. Gonzalez-Lafont, J. M. Lluch, J. M. Bofill and J. M. Anglada, *Theor. Chem. Acc.* **2002**, *107*, 147-153.
- [51] V. Moliner, A. J. Turner and I. H. Williams, *J. Chem. Soc. Chem. Commun.* **1997**, *14*, 1271.

- [52] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187-217.
- [53] J. Simons, P. Jorgensen, H. Taylor and J. Ozment, *J. Phys. Chem.* **1983**, *87*, 2745-2753.
- [54] S. Martí, V. Moliner and I. Tuñón, *J. Chem. Theory Comput.* **2005**, *1*, 1008-1016.
- [55] J. K. Baker, A.; Delley, B., *J. Chem. Phys.* **1996**, *105*, 192.
- [56] J. Baker, *J. Comput. Chem* **1997**, *18*, 1079.
- [57] J. Guido van Rossum and Fred L. Drake, *An Introduction to Python*, **2003**.
- [58] L. Verlet, *Phys. Rev.* **1967**, *159*, 98-103.
- [59] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.* **1953**, *21*, 1087-1092.
- [60] W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *J. Chem. Phys.* **1982**, *76*, 637-649.
- [61] S. Nosé, *J. Chem. Phys.* **1984**, *81*, 511-519.
- [62] S. Nosé and M. L. Klein, *Molec. Phys.* **1983**, *50*, 1055-1076.
- [63] S. E. Feller, Y. Zhang, R. W. Pastor and B. R. Brooks, *J. Chem. Phys.* **1995**, *103*, 4613-4621.
- [64] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola and J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684-3690.
- [65] W. F. v. Gunsteren in *Methods for Calculation of Free Energies and Binding Constants: Successes and Problems, Vol.*, Escom Science Publishers, Netherlands, **1989**, p.
- [66] A. R. Leach, *Molecular Modelling. Principles and Applications*, London, **1996**.
- [67] D. G. Truhlar and B. C. Garrett, *Acc. Chem. Res.* **1980**, *13*, 440-448.
- [68] B. C. Garrett and D. G. Truhlar, *Abstr. Pap. Am. Chem. S.* **1980**, *180*, 233.
- [69] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **1977**, *23*, 187-199.

- [70] J. P. Valleau and G. M. Torrie in *A Guide to Monte Carlo for Statistical Mechanics. 2. Byways, Vol. 5*, Plenum Press, New York, **1977**, p. 169-194.
- [71] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman and J. M. Rosenberg, *J. Comput. Chem.* **1992**, *13*, 1011-1021.
- [72] M. Souaille and B. Roux, *Comput. Phys. Commun.* **2001**, *135*, 40-57.
- [73] M. Field, *A Practical Introduction to the Simulation of Molecular Systems*, Cambridge University Press, Cambridge, **1999**.
- [74] D. Frenkel and B. Smit, *Understanding Molecular Simulation: from Algorithms to Applications*, Academic Press, **1996**.
- [75] S. Arrhenius, *Z. physic. Chem.* **1889**, *4*, 226.
- [76] J. H. Van't Hoff, *Études dynamique chimique*, **1884**.
- [77] S. Glasstone, K. J. Laidler and H. Eyring, *The Theory of Rate Processes*, McGraw-Hill, New York, **1941**.
- [78] B. C. Garrett and D. G. Truhlar, *J. Am. Chem. Soc.* **1979**, *101*, 5207-5217.
- [79] D. G. Truhlar, *J. Phys. Chem.* **1979**, *83*, 188-199.
- [80] B. C. Garrett and D. G. Truhlar, *J. Am. Chem. Soc.* **1979**, *101*, 4534-4548.
- [81] B. C. Garrett and D. G. Truhlar, *J. Phys. Chem.* **1991**, *95*, 10374-10379.
- [82] H. Eyring and A. E. Stern, *Chem. Rev.* **1939**, *24*, 253-270.
- [83] M. Polanyi and M. G. Evans, *Trans. Faraday Soc.* **1935**, *31*, 875.
- [84] H. Eyring, *J. Chem. Phys.* **1935**, *3*, 107.
- [85] B. C. Garrett and D. G. Truhlar, *Abstr. Pap. Am. Chem. S.* **1980**, *180*, 233-Phys.
- [86] R. T. Skodje, D. G. Truhlar and B. C. Garrett, *J. Phys. Chem.* **1981**, *85*, 3019-3023.
- [87] S. C. Tucker and E. Pollak, *J. Stat. Phys.* **1992**, *66*, 975-990.
- [88] E. Pollak, *J. Chem. Phys.* **1991**, *95*, 533-539.

- [89] D. G. Truhlar, A. D. Isaacson, R. T. Skodje and B. C. Garrett, *J. Phys. Chem.* **1982**, *86*, 2252-2261.
- [90] B. C. Garrett, D. G. Truhlar and G. C. Schatz, *J. Am. Chem. Soc.* **1986**, *108*, 2876-2881.
- [91] W. P. Hu, Y. P. Liu and D. G. Truhlar, *J. Chem. Soc. Faraday T.* **1994**, *90*, 1715-1725.
- [92] J. P. Klinman, *Pure. Appl. Chem.* **2003**, *75*, 601-608.
- [93] V. S. Melissas and D. G. Truhlar, *J. Chem. Phys.* **1993**, *99*, 1013-1027.
- [94] V. S. Melissas and D. G. Truhlar, *J. Phys. Chem.* **1994**, *98*, 875-886.
- [95] D. K. Bondi, J. N. L. Connor, B. C. Garrett and D. G. Truhlar, *J. Chem. Phys.* **1983**, *78*, 5981-5989.
- [96] C. Alhambra, J. Corchado, M. L. Sanchez, M. Garcia-Viloca, J. Gao and D. G. Truhlar, *J. Phys. Chem. B* **2001**, *105*, 11326-11340.
- [97] A. Fernandez-Ramos, D. G. Truhlar, J. C. Corchado and J. Espinosa-Garcia, *J. Phys. Chem. A* **2002**, *106*, 4957-4960.
- [98] T. C. Allison, B. Ramachandran, J. Senekowitsch, D. G. Truhlar and R. E. Wyatt, *J. Mol. Struct.-Theochem.* **1998**, *454*, 307-314.
- [99] J. Villa, J. C. Corchado, A. Gonzalez-Lafont, J. M. Lluch and D. G. Truhlar, *J. Phys. Chem. A* **1999**, *103*, 5061-5074.
- [100] D. G. Truhlar, J. L. Gao, M. Garcia-Viloca, C. Alhambra, J. Corchado, M. L. Sanchez and T. D. Poulsen, *Int. J. Quantum Chem.* **2004**, *100*, 1136-1152.
- [101] M. García-Viloca, C. Alhambra, D. G. Truhlar and J. Gao, *J. Chem. Phys.* **2001**, *114*, 9953-9958.
- [102] G. K. Schenter, B. C. Garret and D. G. Truhlar, *J. Chem. Phys.* **2003**, *119*, 5828-5833.
- [103] D. G. Truhlar, J. L. Gao, C. Alhambra, J. Corchado, M. Garcia-Viloca, M. L. Sanchez, L. D. Kesavan, T. Poulsen and J. Villa, *Abstracts of Papers of the American Chemical Society* **2002**, *224*, U480-U480.
- [104] J. B. S. Haldane, *Enzymes*, Longmans, Green and Co., **1930**.

- [105] N. C. Price and L. Stevens, *Fundamentals of Enzymology*, Oxford University Press, Oxford ; New York, **1984**.
- [106] N. C. Price and L. Stevens, *Fundamentals of Enzymology*, Oxford University Press, Oxford ; New York, **1989**.
- [107] L. Moran, *Biochemistry*, Englenwood, **1994**.
- [108] L. Michaelis and M. L. Menten, *Biochemistry* **1913**, *49*, 333-369.
- [109] E. Fischer, *Ber. Dt. Chem. Ges.* **1894**, *27*, 2985-2993.
- [110] M. Polanyi, *Zeitschrift für Elektrochemie* **1921**, *27*, 143.
- [111] L. Pauling, *Chem. Eng. News* **1946**, *24*, 1375-1377.
- [112] L. Pauling, *Am. Sci.* **1948**, *36*, 51.
- [113] A. Warshel, *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250-5254.
- [114] A. Warshel, *Acc. Chem. Res.* **1981**, *14*, 284-290.
- [115] J. C. Hermann, C. Hensen, L. Ridder, A. J. Mulholland and H. D. Holtje, *Journal of the American Chemical Society* **2005**, *127*, 4454-4465.
- [116] W. Cleland and M. M. Kreevoy, *Science* **1994**, *264*, 1887-1890.
- [117] P. A. Frey, S. A. Whitt and J. B. Tobin, *Science* **1994**, *264*, 1927-1930.
- [118] S. Hur and T. C. Bruice, *J. Am. Chem. Soc.* **2003**, *125*, 10540-10542.
- [119] S. Hur and T. C. Bruice, *Proc. Natl. Acad. Sci.* **2003**, *100*, 2215-2219.
- [120] W. P. Jencks, *Catalysis in Chemistry and Enzymology*, Dover Publication, New York, **1986**.
- [121] R. Wolfenden, *Science* **1983**, *222*, 1087-1093.
- [122] M. J. S. Dewar and D. M. Storch, *Proc. Natl. Acad. Sci.* **1985**, *82*, 2225-2229.
- [123] S. J. Weiner, U. C. Singh and P. A. Kollman, *J. Am. Chem. Soc.* **1985**, *107*, 2219-2229.
- [124] S. G. Cohen, V. M. Vaidya and R. M. Schultz, *Proc. Natl. Acad. Sci.* **1970**, *66*, 249-256.
- [125] J. Crosby, R. Stone and G. E. Lienhard, *J. Am. Chem. Soc.* **1970**, *92*, 2891-2900.

- [126] D. R. Storm and D. E. Koshland, *J. Am. Chem. Soc.* **1972**, *94*, 5805-5814.
- [127] A. Dafforn and D. E. J. Koshland, *Biochem. Biophys. Res. Commun.* **1973**, *52*, 779-785.
- [128] A. D. Mesecar, B. L. Stoddard and D. E. Koshland, *Science* **1997**, *277*, 202-206.
- [129] F. H. Westheimer, *Adv. Enzymol* **1962**, *24*, 441-482.
- [130] M. I. Page and W. J. Jencks, *Proc. Natl. Acad. Sci.* **1971**, *68*, 1678-1683.
- [131] J. Hermans and L. Wang, *J. Am. Chem. Soc.* **1997**, *119*, 2707-2714.
- [132] B. Kuhn and P. A. Kollman, *J. Am. Chem. Soc.* **2000**, *122*, 2586-2596.
- [133] Y. Cha, C. J. Murray and J. P. Klinman, *Science* **1989**, *243*, 1325-1330.
- [134] A. Kohen, R. Cannio, S. Bartolucci and J. P. Klinman, *Nature* **1999**, *393*, 496-499.
- [135] B. Gavish and M. M. Werber, *Biochemistry* **1979**, *18*, 1269-1275.
- [136] G. Careri, P. Fasella and E. Gratton, *Ann. Rev. Biophys. Bioeng.* **1979**, *8*, 69-97.
- [137] S. J. Benkovic and S. Hammes-Schiffer, *Science* **2003**, *301*, 1196-1202.
- [138] M. Jamin, B. Geierstanger and R. L. Baldwin, *Proc. Natl. Acad. Sci.* **2001**, *98*, 6127-6131.
- [139] M. Jamin and R. L. Baldwin, *Nat. Struct. Biol.* **1996**, *3*, 613-618.
- [140] G. Pettersson, *Eur. J Biochem.* **1987**, *166*, 163-165.
- [141] M. K. Gilson, *Proteins* **1993**, *15*, 266-282.
- [142] J. Antosiewicz, J. A. Mccammon and M. K. Gilson, *J. Mol. Biol.* **1994**, *238*, 415-436.
- [143] L. David, *Modélisation des interactions électrostatiques des biomolécules en solution.*, Université Joseph Fourier, Grenoble, **1996**

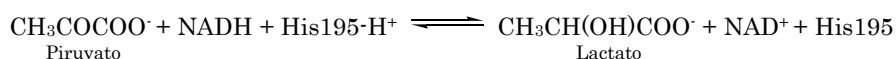
- [144] M. E. Davis, J. D. Madura, B. A. Luty and J. A. Mccammon, *Comput. Phys. Commun.* **1991**, *62*, 187-197.
- [145] M. E. Davis, J. D. Madura and J. A. Mccammon, *Biophys. J.* **1990**, *57*, A415-A415.
- [146] P. T. R. Rajagopalan, S. Lutz and S. J. Benkovic, *Biochemistry* **2002**, *41*, 12618-12628.
- [147] C. E. Cameron and S. J. Benkovic, *Biochemistry* **1997**, *36*, 15792-15800.
- [148] Z. Huang, C. R. Wagner and S. J. Benkovic, *Biochemistry* **1994**, *33*, 11576-11585.
- [149] A. Fersht, *Enzyme structure and mechanism*, W.H. Freeman, New York, **1985**.

CAPÍTULO 4

LACTATO DESHIDROGENASA

CAPÍTULO 4. LACTATO DESHIDROGENASA.....	99
1 BIOQUÍMICA	104
2 RESULTADOS	109
2.1 CÁLCULOS PRELIMINARES	109
2.2 PARÁMETROS DE INTERACCIÓN QM/MM	113
2.3 COORDENADAS DE REACCIÓN	116
2.4 CONSTRUCCIÓN DE UN NUEVO MODELO	117
2.5 EXPLORACIÓN DE LA SEP PARA EL TETRÁMERO	119
2.6 CORRECCIONES A LA DESCRIPCIÓN CUÁNTICA	121
2.6.1 Correcciones SVB.....	122
2.6.2 Correcciones Interpoladas Unidimensionales.....	124
2.6.3 Correcciones Interpoladas Bidimensionales	126
2.7 AMPLIACIÓN DE LA PARTE QM	129
2.8 ESTADO DE PROTONACIÓN DE LOS RESIDUOS	134
2.9 ESTUDIO DE LA CINÉTICA DE LA REACCIÓN QUÍMICA	140
2.9.1 Exploración de la región del ET.....	145
2.9.2 Análisis del Estado de Reactivos	149
2.9.3 Cálculos de energía libre	156
2.9.4 Efecto Túnel	161
2.9.5 KIEs	162
2.10 MUTACIONES	163
2.10.1 Arg109Gln	166
2.10.2 Asp168Ala	170
2.10.3 Doble Mutación Arg109Gln y Asn140Asp.....	173
2.10.4 Doble Mutación Arg109Gln y Asn140Glu.....	176
3 CONCLUSIONES Y PERSPECTIVAS DE FUTURO	178
4 BIBLIOGRAFÍA	181

La enzima L-Lactato Deshidrogenasa (LDH) juega un papel central en la glicólisis anaeróbica tanto en los organismos eucarióticos como en los procarióticos. Esta enzima cataliza la reacción de interconversión estereoselectiva entre piruvato y L-lactato en presencia del cofactor nicotinamida adenín dinucleótido (NADH), en la cual ocurre una transferencia de hidruro desde el cofactor y una transferencia de protón desde la histidina 195 de la cadena de aminoácidos de la enzima.



La LDH (EC 1.1.1.27)* es una enzima de considerables proporciones ya que consta de cuatro monómeros en su forma activa (Figura 1). En la literatura existe una gran cantidad de trabajos sobre la cinética de la LDH, especialmente en la que se encuentra en la bacteria *Bacillus stearothermophilus* (Bs).^[1-4] La enzima BsLDH está formada por una cadena polipeptídica de 316 aminoácidos en cada uno de sus monómeros.

* Nomenclatura y clasificación según el IUBMB (Internacional Union of Biochemistry and Biology).

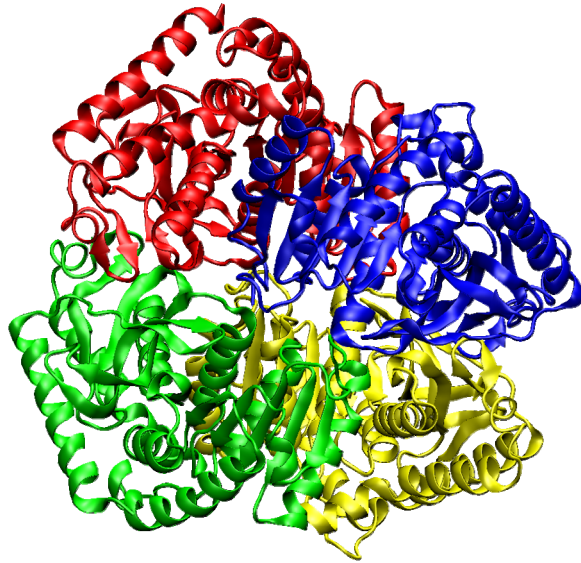


Figura 1. Representación del tetramero BsLDH. Cada uno de sus monómeros se ilustra en un color para su distinción.

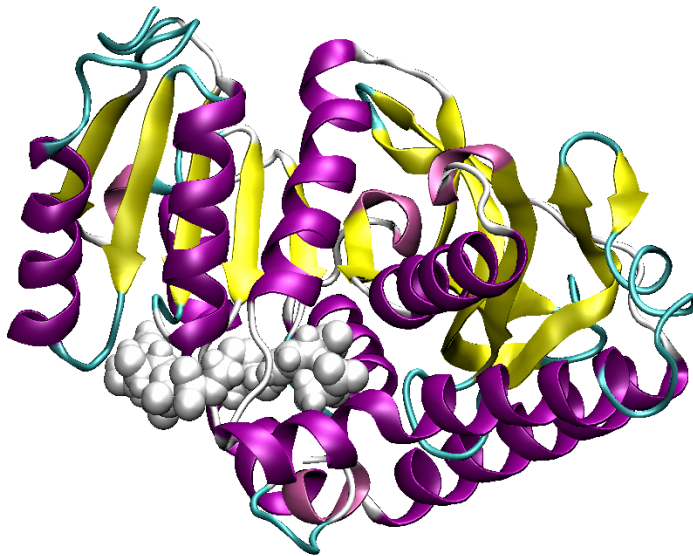


Figura 2. Representación de un monómero de la LDH donde se puede observar los dominios de láminas β en amarillo y las α -hélices en rosa, y el resto de hélices en azul y blanco. En forma de esferas blancas se representa el cofactor NAD.

La estructura de la LDH coincide con otras deshidrogenasas, como por ejemplo las enzimas Malato Deshidrogenasa y Alcohol Deshidrogenasa,^[5-7] en el lugar donde se une el cofactor. Este hecho sugiere que esta subestructura es general para unir nucleótidos, y en particular, el coenzima NAD. El lugar de unión del NAD(H) es un dominio paralelo y abierto, de 6 cadenas β con α hélices en cada lado las láminas β . Esta estructura α/β es simétrica y se ha observado que se construye a partir de dos mitades con topología idéntica y estructura similar (Figura 2).^[8]

1 Bioquímica

El piruvato es una molécula pequeña que constituye el producto final de glucólisis en el citosol[‡]. La glucólisis es el proceso por el que la glucosa se degrada hasta piruvato. El piruvato pasa a la mitocondria donde entra en el ciclo de Krebs y se produce la oxidación total de las moléculas combustibles de la célula. Antes de entrar en la mitocondria el piruvato puede convertirse en lactato en una reacción catalizada por la enzima LDH. Este proceso se conoce como glicólisis anaeróbica. El objetivo de la glicólisis anaeróbica no es formar el lactato sino reoxidar el NADH a NAD⁺ y así permitir la continua formación de ATP.

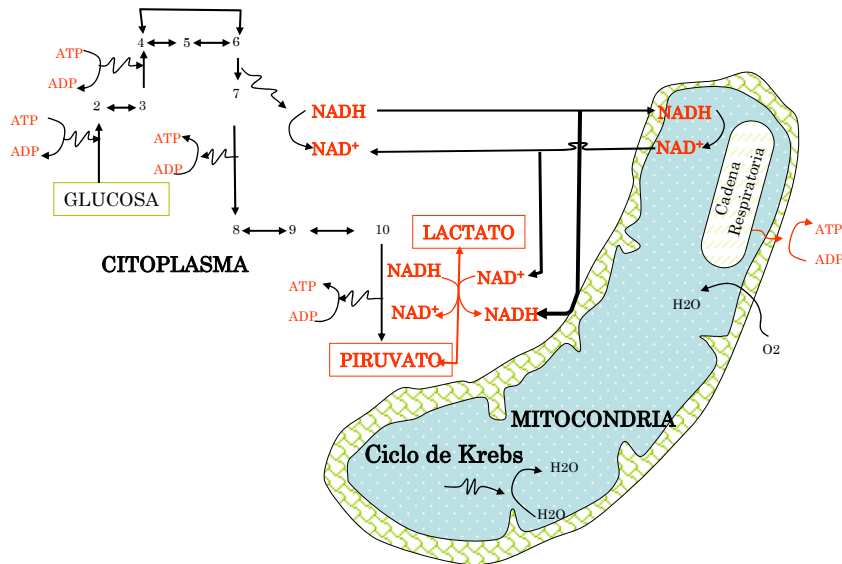


Figura 3. Metabolismo de la glucosa. En la glicólisis la glucosa se degrada en el citosol hasta llegar a piruvato. La LDH convierte el piruvato en lactato que entra en la mitocondria para ser degradado mediante el ciclo de Krebs.

[‡] El citosol es el medio interno del citoplasma. Se encuentra delimitado por la membrana citoplasmática y la envoltura nuclear. En él aparecen el citoesqueleto y los orgánulos celulares.

La LDH aumenta la constante de velocidad en la reacción de interconversión entre piruvato y L-lactato en un factor de 10^{14} comparado con la reacción en disolución.^[9] Cálculos de química cuántica confirman que el mecanismo incluye la donación de un ión hidruro^[10, 11], H^- , desde el cofactor NADH al átomo de carbono carbonílico del piruvato, combinado con la donación de un protón desde la His195 del centro activo^[10] al átomo de oxígeno carbonílico del piruvato. La constante catalítica medida experimentalmente, a $T=25^\circ C$, es de 250 s^{-1} .^[12] A partir de experimentos para calcular efectos isotópicos cinéticos (*Kinetic Isotope Effects*, KIEs),^[13] en los cuales se obtenía el valor de $KIE=1$, se dedujo que era el movimiento del bucle que incluye los aminoácidos 99-110, y no la reacción química, la etapa determinante de la reacción catalizada. El movimiento del bucle (Figura 4) permite aislar el centro activo del disolvente y situar adecuadamente algunos aminoácidos en las cercanías del sustrato. A partir de esta constante cinética y asumiendo un KIE intrínseco entre 3 y 6 (típico de enzimas de la familia de las deshidrogenasas)^[14, 15] se determinó que la etapa de la reacción química debía tener una constante de velocidad superior o igual a 750 s^{-1} .^[16]

Estudios de mutagénesis del centro activo de la LDH demuestran el importante papel jugado por la Arg109^[16, 17] y el Asp168^[18] en la constante catalítica de la reacción enzimática. La Arg109 provoca la polarización del enlace carbonílico del piruvato. Mientras que el Asp168 es responsable de la modulación del pK_a de la His195. Investigaciones similares muestran que la Arg171 es un residuo esencial para el anclaje correcto del sustrato, interaccionando con los oxígenos carboxílicos (ver Figura 4).^[10, 19]

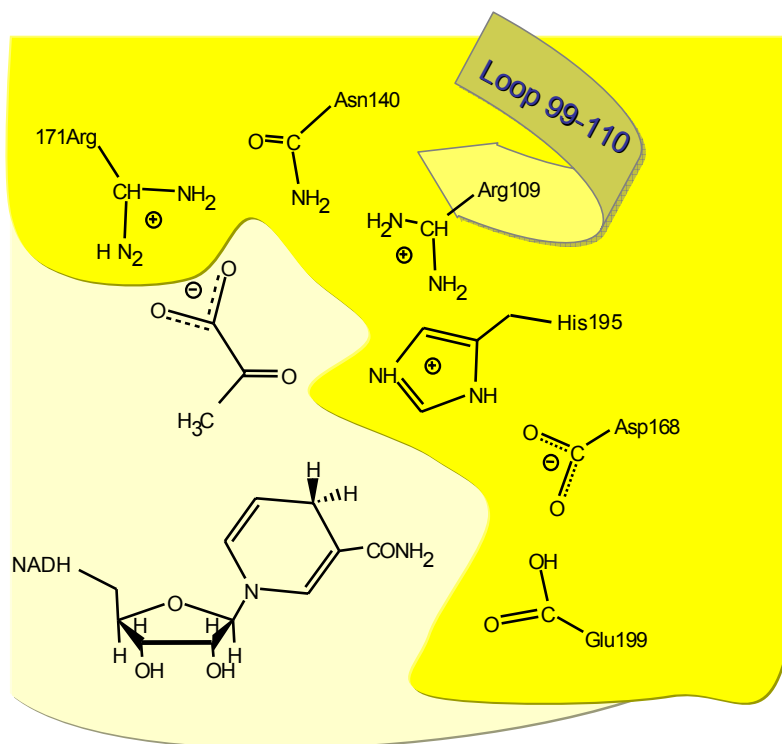


Figura 4. Representación esquemática del centro activo de la LDH. Se observa la situación del sustrato en el centro activo, así como parte del cofactor y los aminoácidos que forman el centro activo.

En los estudios cinéticos de algunas mutaciones de aminoácidos del centro activo se observa un descenso en la velocidad catalítica. Así, en la mutación de la Arg109 por una glutamina, Arg109Gln, se observa una constante cinética de 0.6 s^{-1} . Suponiendo que en este caso la etapa química es limitante, se concluye que la velocidad del proceso de transferencia de hidruro en la enzima mutada es al menos $750/0.6=1200$ veces menor que en la enzima nativa. Así pues, el efecto de la Arg109 sobre la energía de activación sería de más de 4.2 kcal.^[16] Hay un pequeño cambio de la termodinámica de la unión del coenzima (medida directamente por la anisotropía del NADH)^[1] y el pK_a de la histidina 195 que controla la unión del sustrato. El gran cambio en la constante

catalítica permite cuestionar si la eliminación de la Arg109 ha cambiado la naturaleza del proceso limitando la velocidad máxima de la enzima.^[20] Cuando se realiza el estudio con la enzima mutada Arg109Gln se obtiene un valor de KIE=3 para la sustitución del hidruro por un átomo de deuterio.^[16]

La influencia del Asp168 en la donación del protón de la His195 se evaluó realizando la mutación Asp168Ala y Asp168Asn.^[21-23] Incluso conociendo el hecho de que la asparagina (Asn) crea un enlace de hidrógeno con la His195 mientras la alanina (Ala) no lo hace, las dos mutaciones tienen comportamientos catalíticos similares; disminuyen la constante catalítica a 20 y 5.5 s⁻¹, respectivamente.^[17, 18] La conclusión que se extrae de este estudio es que el Asp168 estabiliza el estado de transición. Esta estabilización se atribuye a la interacción del Asp168 con la His195 que mantiene la His195 protonada en el momento que se transfiere el hidruro.^[17, 24]

La mutación Arg171Lys no afecta a la unión del coenzima, pero causa un incremento de la constante de Michaelis del piruvato. La causa de este comportamiento es el descenso de la energía de unión del sustrato en el sitio activo, de al menos 5.5 kcal/mol⁻¹ y, en parte, se explica por el aumento de la distancia de la lisina al sustrato. El efecto total en la constante catalítica de la reacción es un descenso a 18 s⁻¹. En general la estructura de la enzima no se ve perturbada ya que la influencia de la mutación es local.^[19]

En estudios teóricos previos^[25-36] se han propuesto diferentes mecanismos para la reacción química. Los estudios teóricos basados en la metodología híbrida QM/MM difieren en el carácter concertado o secuencial de ambas transferencias y, en el último caso, en el orden en que se producen cada una de ellas. Así, Ranganathan y Gready^[25, 28]

encontraron un mecanismo concertado en el cual la transferencia de hidruro precedía a la de protón, en concordancia con un estudio previo de Warshel y colaboradores,^[36] pero en desacuerdo con argumentos enzimáticos para procesos de transferencia de hidruro. Estudios teóricos previos llevados a cabo en nuestro grupo de investigación sobre esta enzima^[33, 34] predicen un mecanismo concertado pero asincrónico en el cual la transferencia de protón en la EsT está bastante más avanzada que la de hidruro, en concordancia con estudios previos en vacío.^[10, 11] Estudios más recientes, utilizando la misma metodología pero con un modelo molecular más grande y más flexible, demostraron que los dos mecanismos pueden ser localizados sobre la SEP, con valores de barreras (en términos de energía de la región cuántica y energía total) que no permitían discernir entre cual de ellos sería el más favorable.^[32]

2 Resultados

2.1 Cálculos Preliminares

La primera etapa en el estudio de la LDH fue la reoptimización de los puntos estacionarios implicados en la reacción química partiendo de estructuras obtenidas en un trabajo publicado por nuestro grupo de investigación.^[32] En este estudio la enzima se representó por un único monómero, el cofactor, el sustrato, y una esfera de moléculas de agua centrada en el sustrato (Figura 5).

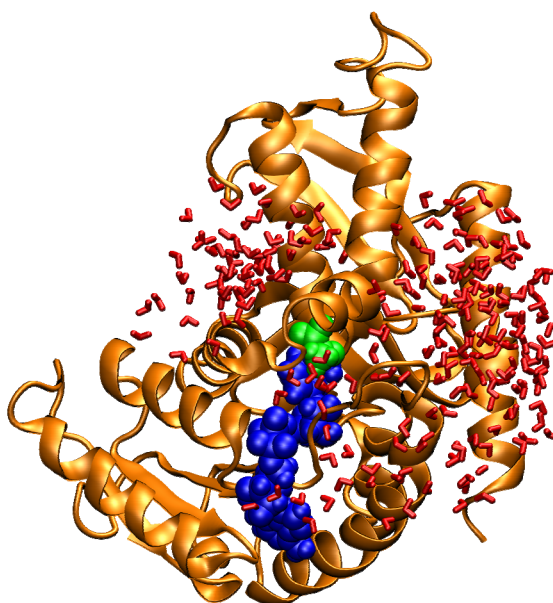


Figura 5. Representación del modelo utilizado en el estudio de la reacción en el monómero de la LDH. El sustrato y cofactor vienen representados con esferas verdes y azules, respectivamente, mientras que las moléculas de agua de solvatación están representadas con cilindros rojos.

En la Figura 6 se muestra la división QM/MM llevada a cabo en una primera exploración. La región QM, que será tratada mediante el Hamiltoniano AM1, incluye 52 átomos: piruvato, histidina hasta el

carbono α y el NADH hasta la ribosa. Al definir la región cuántica se cortan dos enlaces covalentes que son tratados con la aproximación de átomo *link*.^[37, 38] En el centro de masas de los átomos QM se centra una esfera de 14 Å de moléculas de agua TIP3P^[39] para solvatar el sistema. El modelo consta de un total de 5597 átomos MM de los cuales se congelan 3620 que son aquellos átomos que, seleccionados por residuo, quedan más allá de 14 Å de tres átomos: los dos que actúan como dador y aceptor del hidruro y el dador del protón (N_{his} , C_{pir} , C_{nic}).

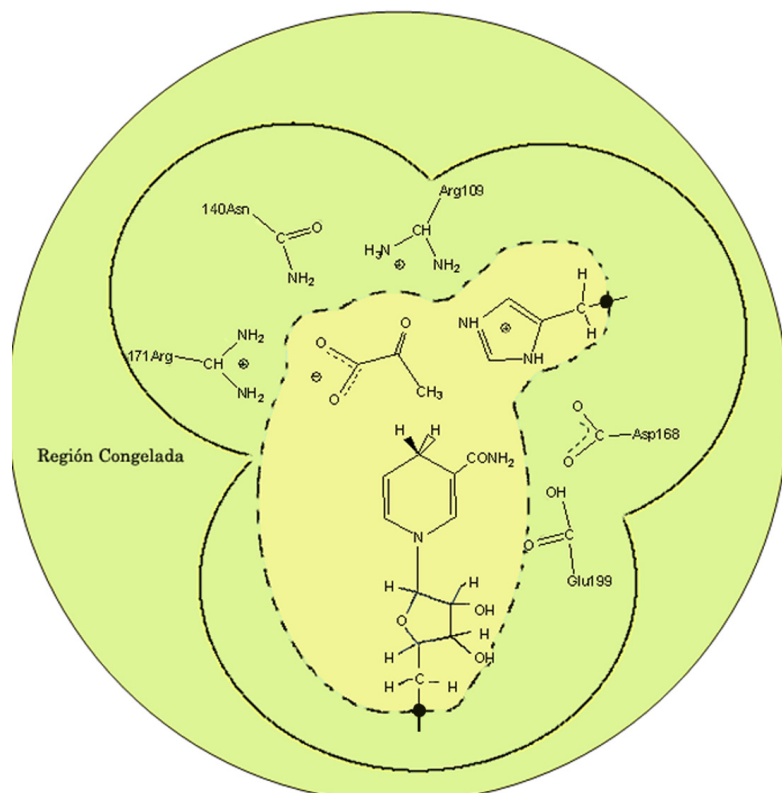


Figura 6. Esquema de la división QM/MM del primer modelo de la LDH. En amarillo se señala la región cuántica y en verde la región clásica. Los átomos link situados en los enlaces covalentes por los que pasa la división QM/MM se representan con un punto (•).

Con este primer modelo, utilizando la combinación de programas GRACE/CHARMM,^[34] localizamos tres estructuras de transición (EsT) para dos posibles mecanismos de reacción (Figura 7). El primer mecanismo es por etapas e implica primero la transferencia del hidruro, creándose un intermedio con dos cargas negativas, seguida por la transferencia del protón. El segundo mecanismo es concertado por lo que sólo se encuentra una EsT entre reactivos (piruvato) y productos (lactato).

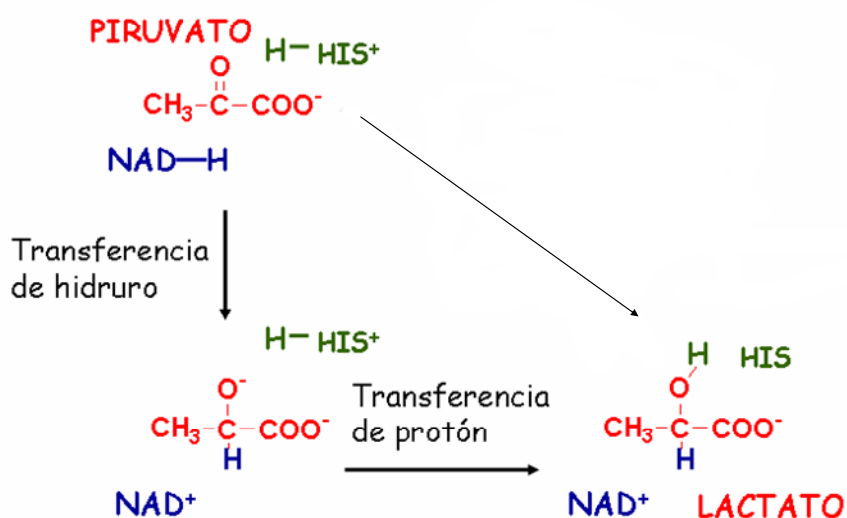


Figura 7. Representación de los dos mecanismos de reacción.

A partir de las estructuras de transición localizadas construimos los IRC^[40] hasta llegar a los valles de reactivos, intermedio o productos. La Figura 8 representa la evolución de la energía potencial del sistema con la coordenada de reacción para los dos mecanismos localizados. Los resultados reproducen los obtenidos en los estudios anteriores a esta tesis.^[32]

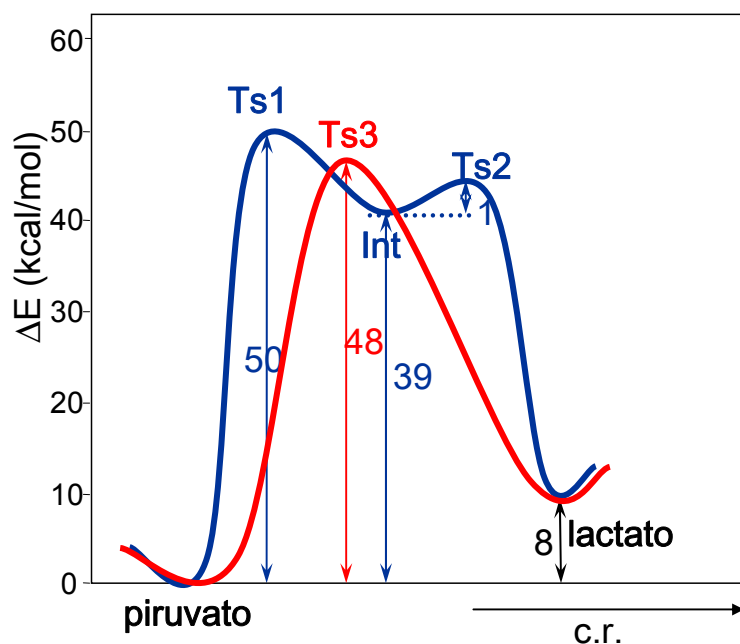


Figura 8. Representación esquemática de los perfiles de energía potencial frente a la coordenada de reacción. En azul se muestra el perfil del mecanismo por etapas, siendo Ts1 la EsT correspondiente a la transferencia de hidruro y Ts2 la EsT de transferencia de protón. En rojo el perfil del mecanismo concertado, Ts3 es la EsT correspondiente al mecanismo concertado.

Con estos resultados, la pequeña diferencia entre las barreras energéticas de los dos mecanismos no nos permite decidir cuál de ellos es el más representativo, teniendo en cuenta los errores asociados a los tratamientos semiempíricos en el subsistema QM y a la incertidumbre asociada a diferencias de valores de energías potenciales tan elevadas. Para obtener resultados más concluyentes es necesario introducir mejoras en nuestro tratamiento que permitan obtener resultados, al menos cualitativamente, comparables con los datos experimentales. En concreto, nos propusimos introducir varias mejoras al método QM/MM:

1. Incluir en el cálculo el tetrámero que constituye la BsLDH, de manera que se representen adecuadamente los efectos de larga distancia entre monómeros y el efecto cooperativo que pueda existir entre ellos.

2. Solvatar con una esfera de moléculas de agua de mayor radio y congelar los átomos situados más allá de esa distancia respecto al sustrato, ampliando así el área solvatada y la parte flexible.
3. Mejorar los parámetros que describen las interacciones QM/MM en nuestro sistema.
4. Hacer correcciones al método cuántico utilizado (AM1) para describir la reacción.
5. Ampliar la parte cuántica para considerar posibles efectos de transferencia de carga al entorno.
6. Mejorar la definición del átomo frontera entre la parte QM y MM utilizando la aproximación GHO.^[41]

2.2 *Parámetros de interacción QM/MM*

El primer paso en la construcción del nuevo modelo consiste en la calibración del potencial de interacción QM/MM ajustando, si fuera necesario, los parámetros de van der Waals de los átomos QM. Para ello empleamos el procedimiento descrito por Freindorf y Gao. Se calcula la energía de interacción a nivel cuántico *ab initio* (HF/6-31G(d)) entre moléculas similares al fragmento QM (cuyos parámetros se quieren mejorar) y moléculas de agua. A continuación se ajustan los parámetros de van der Waals para reproducir esta energía de interacción usando un modelo híbrido AM1/TIP3P.^[42] Así, consideramos complejos de agua con modelos para la histidina 195 y el anillo de nicotinamida del NADH, y con las moléculas de piruvato y etanol. En la Figura 9 se representan los complejos bimoleculares estudiados para calcular la energía de interacción intermolecular. Estos complejos se utilizan para optimizar los

parámetros de van der Waals para los átomos de nitrógeno, oxígenos hidroxílicos, hidrógenos polares enlazados a nitrógenos u oxígenos, y también los átomos de carbono y oxígeno del piruvato. En todos los casos se congelan los grados de libertad de la parte QM al valor correspondiente a la molécula optimizada en fase gas, mientras que la molécula de agua se optimiza completamente. En la Tabla 1 se presentan las energías de interacción de los modelos para los diferentes cálculos mientras que en la Tabla 2 se presentan los parámetros obtenidos de acuerdo con este procedimiento.

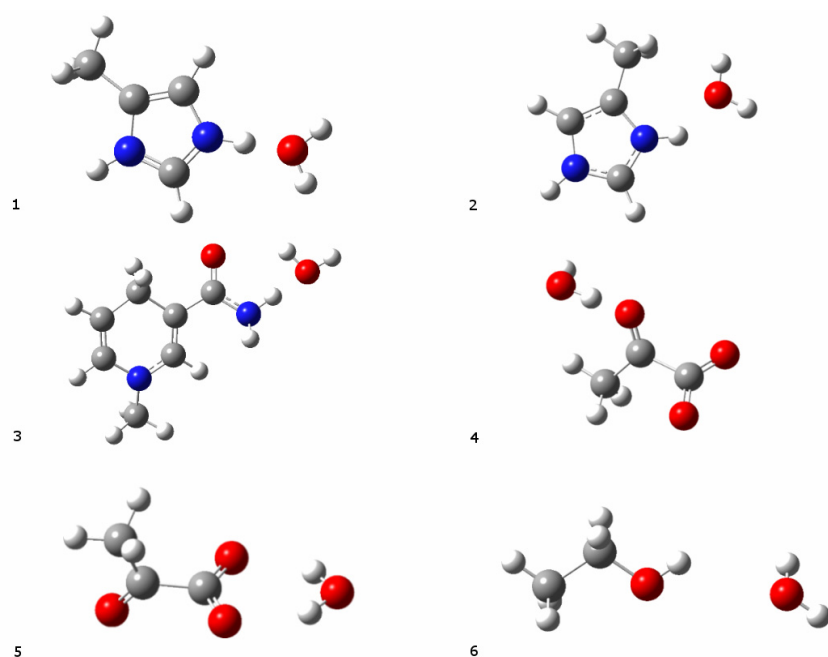


Figura 9. Complejos bimoleculares utilizados para calcular nuevos parámetros de Lennard-Jones QM/MM. Los complejos 1 y 2 representan la interacción de la histidina con una molécula de agua; el complejo 3 representa parte del anillo del NADH; los complejos 4 y 5 se usan para optimizar los parámetros del piruvato; y el complejo 6 para calcular la interacción de un alcohol con una molécula de agua y reparametrizar el grupo hidroxilo.

Número del complejo	$\Delta E(\text{HF})$	$\Delta E(1)$	$\Delta E(2)$	$\Delta E(3)$
1	-16.02	-12.92	-13.9	-15.82
2	-16.18	-13.03	-14.03	-15.98
3	-5.24	-5.03	-5.07	-5.62
4	-10.81	-14.74	-14.84	-14.84
5	-13.02	-17.63	-17.68	-17.68
6	-5.47	-3.17	-3.80	-3.89

Tabla 1. Energías de interacción (en kcal/mol) para los complejos que aparecen en la Figura 9. Cada columna muestra la diferencia entre los resultados obtenidos a nivel HF/6-31G(d) y AM1/MM utilizando los parámetros van der Waals del campo de fuerzas CHARMM (1), los parámetros dados en la referencia [42] (2) y los optimizados (3) que se dan en la Tabla 2.

Átomo	Este trabajo		Gao et al. ^a		CHARMM	
	σ	ϵ	σ	ϵ	σ	ϵ
N(H)/N(H2) ^b	1.00	-0.14	1.57	-0.15	1.85	-0.20
H(N) ^b	0.35	-0.08	0.45	-0.10	0.2245	-0.046
H(O) ^b	0.28	-0.10	0.45	-0.10	0.2245	-0.046
O(H) ^b	1.08	-0.08	1.65	-0.20	1.70	-0.12
C(=O) ^b	1.65	-0.20	1.65	-0.20	1.70	-0.12
C(H) ^b	1.96	-0.08	1.96	-0.08	1.99	-0.07
O(=C) ^b	1.65	-0.20	1.65	-0.20	1.70	-0.12
O ^c	1.65	-0.20	1.65	-0.20	1.70	-0.12
H ^d	1.12	-0.01	1.12	-0.01	1.32	-0.022

Tabla 2. Parámetros de Van der Waals para los átomos QM en Å, σ , y kcal/mol, ϵ . ^a Referencia [42] ^b A(B) significa el átomo A unido siempre al átomo B. ^c Oxígeno de un grupo carbonilo ^d Hidrógenos unidos a carbonos

Al realizar los cálculos para optimizar los parámetros de los átomos que contiene el piruvato (complejos 4 y 5 de la Figura 9) no se ha encontrado ningún valor de σ y ϵ que nos represente la energía de interacción QM/MM ajustable a los cálculos a nivel *ab initio*.

2.3 Coordenadas de reacción

Para definir las coordenadas de reacción y explorar las SEP nos basamos en el estudio realizado con el monómero. De esta manera definiremos una coordenada asociada a la transferencia de hidruro (H_1) desde el átomo de carbono del anillo de nicotinamida, C_{nic} (átomo dador), al átomo carbonílico del piruvato, C_{pyr} (átomo aceptor), y otra a la transferencia de protón (H_2), desde el átomo N de His195, N_{his} (átomo dador), al oxígeno carbonílico del piruvato, O_{pyr} (átomo aceptor). En concreto, aquí definimos la coordenada de reacción para la transferencia del hidruro, R_1 , como la combinación antisimétrica de la distancia del hidruro con el carbono dador menos la distancia de hidruro con el átomo de carbono aceptor (ecuación (2.1)). La coordenada de transferencia de protón, R_2 , la definimos como la combinación antisimétrica de las distancias del protón y sus átomos dador (N_{his}) y el aceptor (O_{pyr}) (ecuación (2.2)).

$$R_1 = r_{C_{\text{nic}}H_1} - r_{C_{\text{pyr}}H_1} \quad (2.1)$$

$$R_2 = r_{N_{\text{his}}H_2} - r_{O_{\text{pyr}}H_2} \quad (2.2)$$

La exploración de la SEP se lleva a cabo utilizando las coordenadas de reacción definidas según las ecuaciones (2.1) y (2.2), con el programa CHARMM (versión c29b2) utilizando el comando RESDISTANCE[§] para definir cada una de las coordenadas de reacción (R_i).

[§] *Restrained Distances*. Se aplica en general para restringir una o varias distancias permitiendo que se puedan especificar varias distancias. Facilita la exploración de una coordenada de reacción si ésta es una combinación lineal de distancias aplicando una fuerza armónica sobre la misma.

2.4 Construcción de un nuevo modelo

La estructura de RX de la bsLDH es un octámero, sin embargo utilizamos los cuatro monómeros que forman un tetrámero ya que esta es la forma activa de la LDH.^[13] Cada monómero de la estructura obtenida por difracción de RX contiene 316 residuos de aminoácidos, el cofactor NADH y un inhibidor utilizado para cristalizar la enzima. El inhibidor, oxamato, es sustituido por el piruvato en los cuatro centros activos para empezar nuestro estudio.

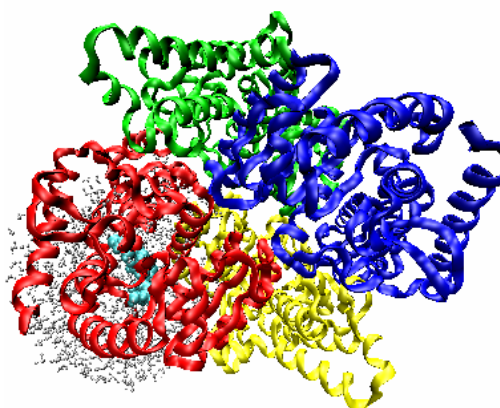


Figura 10. Modelo del tetrámero de la BsLDH. En el centro activo de uno de los monómeros (en rojo) se define la región QM (dibujado con esferas azul claro), que se solvata con una esfera de moléculas de agua de radio 24 Å.

Añadimos hidrógenos, tanto en la enzima, como en el cofactor y el sustrato utilizando el comando HBUILD implementado en el programa CHARMM. Todos los grupos ionizables de la enzima se protonan, inicialmente, según su estado de protonación normal a pH 7 a excepción de la His195, que modelamos en su estado protonado ya que así se requiere para que tenga lugar la reacción. En las histidinas que son neutras añadimos el hidrógeno en el sitio común ND1 (N δ).

Como en el estudio anterior, el sistema se divide en 52 átomos en la región QM y el resto de átomos definen la región MM. El subsistema QM se describe con un Hamiltoniano AM1 y los átomos frontera entre ambas regiones se describen con la metodología GHO. (Véase Figura 11). Antes de empezar la exploración de la SEP aplicamos un protocolo para relajar y solvatar bien el sistema. Le añadimos una esfera de moléculas de agua de radio 24 Å TIP3P^[39] centrada en el centro de masas del piruvato, quitando las moléculas de agua que están a menos de 2.8 Å de cualquier otro átomo distinto al hidrógeno. Después de hacer una optimización previa de las moléculas de agua (100 ciclos utilizando el método ABNR),^[43] realizamos una dinámica molecular de 5ps, con un paso de integración de 1 fs, de aquellos átomos que quedan dentro de una esfera de 24 Å de radio, con el fin de relajar el sistema y evitar contactos energéticamente desfavorables. Repetimos este ciclo de nuevo, añadiendo ahora la esfera de moléculas de agua en una orientación distinta, para asegurarnos que tenemos una buena solvatación y relajación del sistema. El modelo resultante consta de 22139 átomos, de los cuales 15191 permanecerán fuera de la esfera flexible y se congelan durante los posteriores cálculos.

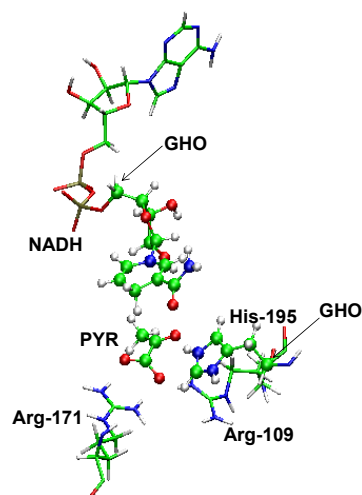


Figura 11. Centro activo de la LDH. La región QM se representa mediante un modelo de esferas y cilindros. Los átomos frontera (tratados con la metodología GHO) están marcados con flechas.

2.5 Exploración de la SEP para el tetrámero

La SEP AM1/MM obtenida se presenta en la Figura 12 en la que se pueden distinguir dos posibles mecanismos muy similares a los resultados de la exploración sobre el monómero.^[32] Como se puede observar, uno de ellos es por etapas e implica una transferencia de hidruro con la aparición de un intermedio de reacción y la posterior transferencia del protón. El otro mecanismo es concertado pero asincrónico. En la Tabla 3 se muestran las distancias más importantes y las energías relativas de los puntos singulares de la SEP.

Los resultados obtenidos serán referencia para las sucesivas mejoras que aplicaremos al método. En esta superficie nos encontramos un mecanismo concertado que pasa por la EsT llamada TS3 que se lleva

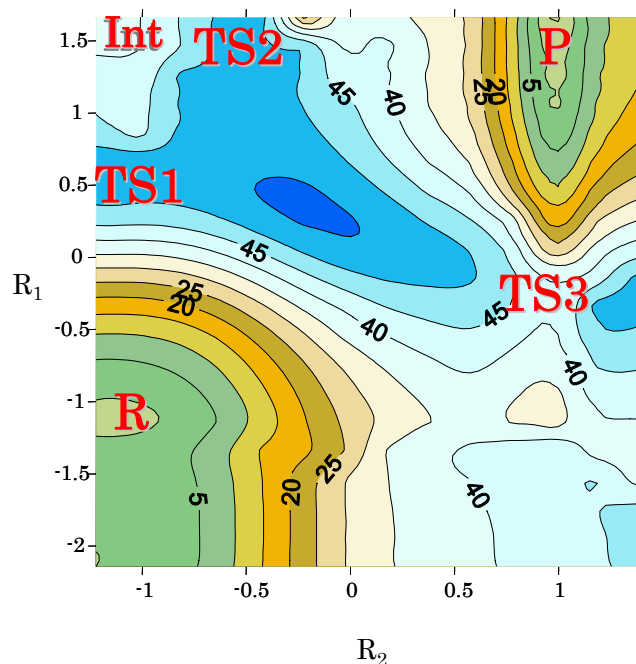


Figura 12. SEP obtenida a nivel AM1/MM. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), intermedio (Int), productos (P), y las EsT (TS1, TS2 y TS3) se indican cualitativamente

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.15	2.29	1.00	2.22	0
TS1	1.70	1.24	1.01	2.03	52
Intermedio	2.80	1.15	1.00	2.22	38
TS2	2.79	1.14	1.06	1.68	52
Productos	2.79	1.13	1.98	0.99	-5
TS3	1.30	1.44	1.96	0.99	41

Tabla 3. Distancias (Å) y energías relativas (kcal/mol) de los puntos singulares que se sitúan en la SEP AM1/MM.

a cabo con la transferencia de protón adelantada y cuya barrera es de 41 kcal/mol. Este es el mecanismo más favorecido ya que el mecanismo por etapas que pasa por TS1 y TS2 tiene una barrera de 52 kcal/mol.

2.6 Correcciones a la descripción cuántica

Para mejorar la descripción de la reacción obtenida a nivel AM1 empleamos las correcciones SVB o IC descritas en el Capítulo 3. Para ello primero realizamos cálculos en fase gas a nivel *ab initio* con modelos pequeños de las moléculas implicadas en la reacción química (Figura 13). En el modelo seleccionado tenemos 42 átomos descritos en la región QM donde se incluyen los átomos principales de las tres especies que están implicadas en la reacción química. Utilizamos las estructuras localizadas en el monómero, concretamente las estructuras de transición, reactivo, intermedio y producto, del camino por etapas donde tiene lugar primero la transferencia de hidruro y después la de protón. Así, obtenemos la energía de los modelos de 42 átomos sin reoptimizar a dos niveles de cálculo AM1 y MP2/6-31G(d,p). En la Tabla 4 se presentan las energías relativas calculadas a los dos niveles. Comparando entre los valores del método semiempírico y el MP2 podemos decir que en nuestra superficie AM1/MM tenemos un error importante y por lo tanto no estamos describiendo bien el proceso enzimático. En concreto, el uso del Hamiltoniano MP2 proporciona una energía de reacción unas 26 kcal/mol menor para la transferencia de hidruro, y 18 kcal/mol mayor en la transferencia de protón. De la misma manera la barrera también tiene una diferencia importante en la transferencia de hidruro (16.9 kcal/mol) mientras que para la transferencia de protón es de menor importancia (apenas 2 kcal/mol). La inclusión de estas correcciones, resulta ser por tanto, cualitativa y cuantitativamente significativa.

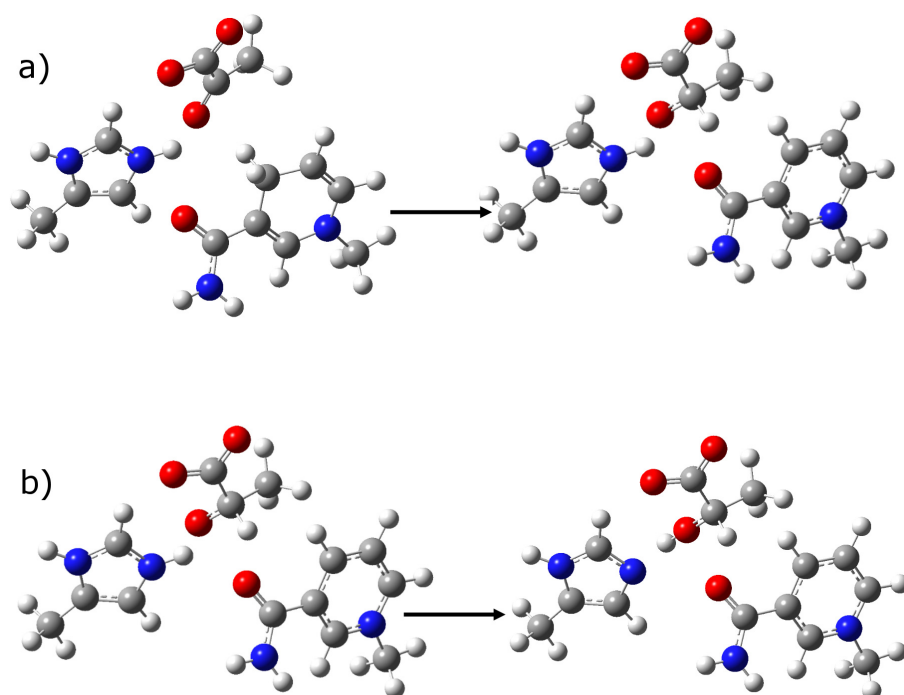


Figura 13 (a) Modelos moleculares utilizados para calibrar los cálculos AM1 versus MP2/6-31G(d,p): (a) modelo para la transferencia de hidruro: desde los reactivos (NADH y piruvato) al intermedio (NAD⁺ y PyrH), (b) modelo para la transferencia de protón: desde el intermedio (PyrH y la His195 protonada) a productos (lactato y la His195).

	ΔE^\ddagger	ΔE
Transferencia Hidruro		
AM1	76.27	93.11
MP2/6-31G(d,p)	59.38	66.96
Transferencia Protón		
AM1	-5.07	-61.57
MP2/6-31G(d,p)	-7.07	-43.14

Tabla 4. Energías de reacción en fase gas (ΔE) y contribuciones a la barrera (ΔE^\ddagger)(kcal/mol) tanto para la etapa de transferencia de hidruro como la etapa de la transferencia de protón presentado en la Figura 13 geometrías están tomadas de la SEP AM1/MM en el monómero.

2.6.1 Correcciones SVB

Con los datos presentamos podemos calibrar los parámetros a utilizar en las correcciones SVB tanto para la etapa de transferencia de

hidruro (Ecuación (2.3)) como para la del protón (Ecuación (2.4)). Los parámetros de los términos de corrección SVB se muestran en la Tabla 5.

$$E_{\text{SVB,R1}}(r_{\text{DH}}, r_{\text{DA}}, r_{\text{AH}}) = \frac{1}{2} [M_1(r_{\text{DH}}) + M_2(r_{\text{AH}})] - \frac{1}{2} ([M_1(r_{\text{DH}}) - M_2(r_{\text{AH}})]^2 + 4[V_{12}(r_{\text{DA}})]^2)^{1/2} + \Delta D_{\text{DH}} \quad (2.3)$$

$$E_{\text{SVB,R2}}(r_{\text{NH}}, r_{\text{NO}}, r_{\text{OH}}) = \frac{1}{2} [M_1(r_{\text{NH}}) + M_2(r_{\text{OH}})] - \frac{1}{2} ([M_1(r_{\text{NH}}) - M_2(r_{\text{OH}})]^2 + 4[V_{12}(r_{\text{NO}})]^2)^{1/2} + \Delta D_{\text{NH}} \quad (2.4)$$

(a)

ΔD_{DH}	α_{DH}	ΔD_{AH}	α_{AH}	D_{DA}	α_{DA}
0.01	1.36	26.15	6.43	1.5	0.50

(b)

ΔD_{NH}	α_{NH}	ΔD_{OH}	α_{OH}	D_{NO}	α_{NO}
21.27	1.49	1.46	1.09	7.1	1.00

Tabla 5. Parámetros para las funciones de corrección SVB para las etapas de reacción de transferencia de hidruro (a) y de protón (b). ΔD s están en kcal/mol y α en Å^{-1} .

Considerando la SEP representada en la Figura 12 como referencia, en este apartado se incluyen las correcciones SVB sobre la misma. En la Figura 14 se muestra la SEP corregida donde aparecen los dos mismos mecanismos descritos anteriormente con pequeñas diferencias. En la Tabla 6 se dan las distancias y energías relativas de los puntos singulares, observándose un importante descenso en la barrera del mecanismo por etapas. Así pues, en el mecanismo por etapas se pasa por una EsT (TS1) que supone una transferencia de hidruro y llega al intermedio Int y después el camino va por la EsT nombrada TS2. La EsT

etiquetada TS3 pertenece al mecanismo concertado y tiene la transferencia de protón más avanzada.

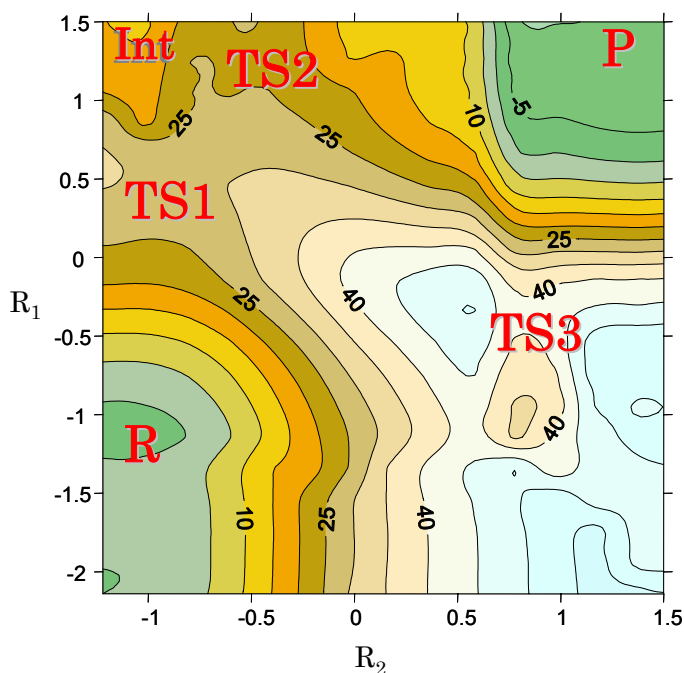


Figura 14. SEP corregida a nivel AM1-SVB/MM. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), intermedio (Int), producto (P), y las EsT (TS1, TS2 y TS3) se indican cualitativamente en la SEP.

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.14	2.29	1.01	2.03	0
TS1	1.62	1.16	1.00	2.02	29
Intermedio	2.79	1.13	1.01	2.03	14
TS2	2.79	1.13	1.06	1.67	25
Productos	2.79	1.13	1.96	0.98	-15
TS3	1.21	1.55	1.77	0.99	41

Tabla 6. Distancias (Å) y energías relativas (kcal/mol) de los puntos estacionarios localizados en la SEP corregida a nivel AM1-SVB/MM.

2.6.2 Correcciones Interpoladas Unidimensionales

Para las correcciones interpoladas (*Interpolated Corrections*, IC), utilizamos los mismos resultados de manera que también tendremos

un término de corrección para cada una de las coordenadas de reacción. Para un valor dado de la coordenada distinguida de reacción el término de corrección vendrá dado como la diferencia entre una energía potencial puntual en fase gas utilizando el nivel de cálculo alto (MP2/6-31(d,p)) y el bajo (AM1). Así, en función de cada coordenada de reacción, los *splines* cúbicos independientes ajustaran las energías de corrección.

$$\Delta E_{IC}(R_1) = \text{spline} \left[E_{QM}^{MP2}(R_1) - E_{QM}^{AM1}(R_1) \right] \quad (2.5)$$

$$\Delta E_{IC}(R_2) = \text{spline} \left[E_{QM}^{MP2}(R_2) - E_{QM}^{AM1}(R_2) \right] \quad (2.6)$$

El resultado de la inclusión de correcciones interpoladas unidimensionales, para las coordenadas R_1 y R_2 , se presenta en la SEP de la Figura 15. En ella vemos un comportamiento similar a la Figura 14, ya que obtenemos dos caminos posibles de reacción, y es en el mecanismo por etapas donde las barreras han disminuido notablemente respecto a la superficie AM1/MM. Precisamente el mecanismo por etapas es el utilizado para calibrar el método. De igual manera presentamos la Tabla 7 de distancias y energías relativas de los puntos estacionarios. Las geometrías de los puntos singulares son parecidas a las encontradas en la SEP SVB-AM1/MM, aunque las energías difieren un poco más. Por ejemplo en la EsT etiquetada TS3 en la SEP IC-AM1/MM tiene una barrera de 52 kcal/mol mientras que en la SVB-AM1/MM es de 41 kcal/mol. Si comparamos las geometrías, en la SEP IC-AM1/MM tenemos el hidruro menos transferido en comparación a la SVB-AM1/MM con una distancia al C_{pir} de 1.75 Å y 1.55 Å respectivamente.

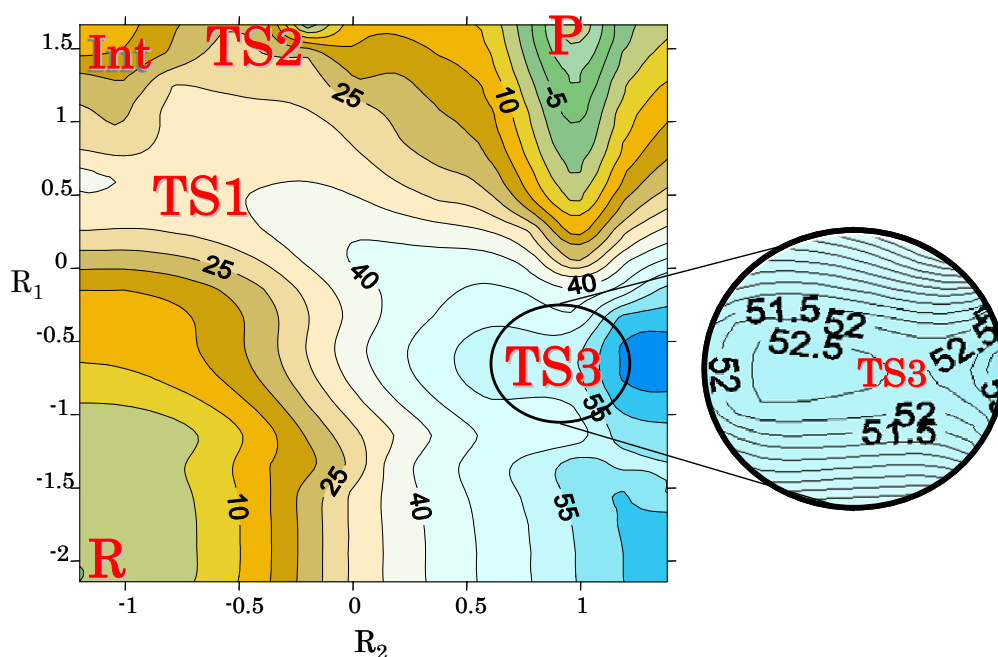


Figura 15. SEP corregida a nivel AM1-IC/MM. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), intermedio (Int), productos (P), y las EsT (TS1, TS2 y TS3) se indican sobre la SEP. La zona del TS3 está ampliada para mejor visualización.

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.13	3.27	1.00	2.22	0
TS1	1.86	1.20	1.03	1.85	34
Intermedio	2.80	1.15	1.00	2.22	13
TS2	2.79	1.13	1.11	1.53	24
Productos	2.79	1.13	1.98	0.99	-19
TS3	1.20	1.75	1.98	1.00	52

Tabla 7. Distancias (Å) y energías relativas (kcal/mol) de los puntos estacionarios que se sitúan en la SEP AM1-IC/MM.

2.6.3 Correcciones Interpoladas Bidimensionales

Hasta ahora las correcciones realizadas al Hamiltoniano AM1 han estado parametrizadas suponiendo independiente las contribuciones dependientes de las coordenadas R_1 y R_2 , es decir, de la transferencia de hidruro y protónica. Vamos a utilizar otro método en el que tenemos en

cuenta la dependencia entre ambas coordenadas. Para ello ajustamos *splines* bidimensionales dependientes simultáneamente de R_1 y R_2 , a las diferencias de energías relativas entre el método de alto nivel (MP2/6-31G(d,p) y el de bajo nivel (AM1) para 100 estructuras seleccionadas a lo largo de todo la SEP(R_1, R_2) AM1/MM representada en la Figura 12.

La nueva SEP con correcciones bidimensionales se presenta en la Figura 16 y las distancias y energías relativas de los puntos estacionarios en la Tabla 8. Los resultados son significativamente diferentes a los anteriores. La diferencia más importante es que ahora no aparece el intermedio en el camino que previamente era por etapas, y el que era concertado es ahora un camino por etapas con un intermedio nuevo (Int2) donde el piruvato tiene transferido el protón. En cuanto a las barreras de los caminos, se puede observar que en el caso del nuevo camino concertado donde la transferencia de hidruro se encuentra en un estado más avanzado, tiene una energía más baja comparada con la obtenida con las correcciones anteriores. El otro camino, donde la transferencia de protón es previa a la transferencia de hidruro, presenta una barrera significativamente más alta.

Es interesante recordar que para realizar las correcciones de SVB y las IC suponemos que las diferencias entre los niveles alto y bajo son independientes de la coordenada R_1 o R_2 . Esto quiere decir que se espera que el error en la descripción semiempírica sea aproximadamente el mismo cuando el protón se transfiere al sustrato (con una carga de -1 u.a.) o al intermedio (con una carga de -2 u.a.). Obviamente esto no es así. Sin embargo, aun teniendo en cuenta este defecto, estas correcciones mejoran significativamente la descripción energética de los dos mecanismos. Comparando la superficie ahora obtenida con las anteriores generadas con correcciones podemos observar: el método SVB presenta

unas barreras de energía similares en el camino que conlleva en primer lugar la transferencia de hidruro con la SEP obtenida con las correcciones bidimensionales. No obstante, el método de corrección IC genera una SEP en la cual, el camino donde la transferencia protónica está más adelantada, tiene barreras similares con el método de correcciones bidimensionales que hemos aplicado.

El camino que ha sido hasta ahora por etapas presenta ahora una EsT (TS4) que conlleva un mecanismo concertado con la transferencia ligeramente más adelantada que la de protón. Por contra, el mecanismo hasta ahora concertado presenta un comportamiento por etapas presentado dos EsT (TS5 y TS6) para llegar de R a P.

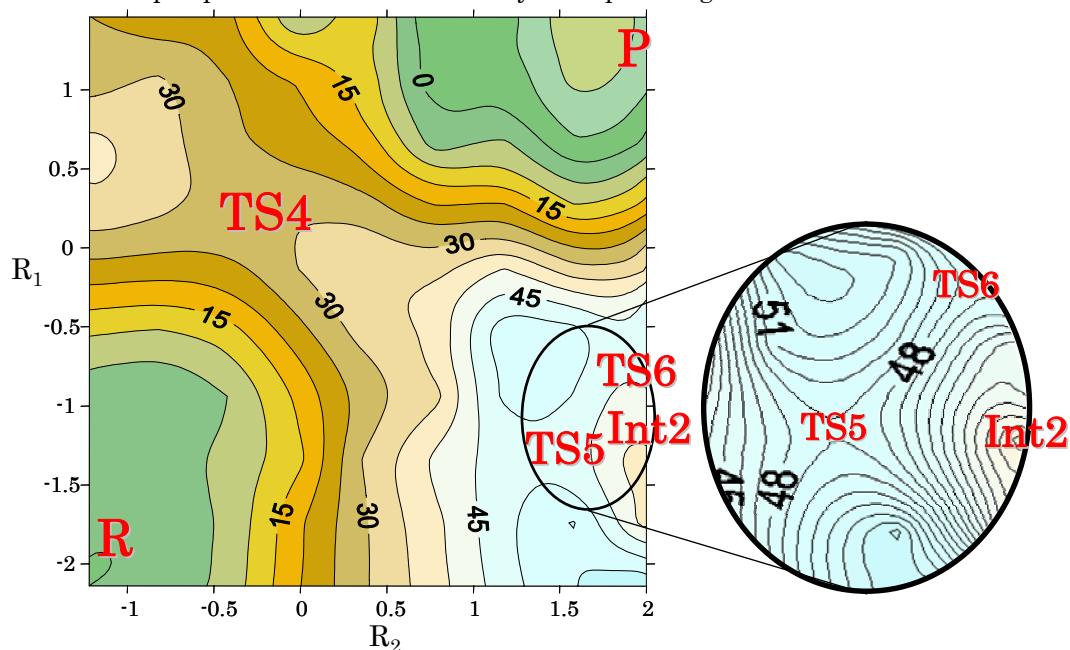


Figura 16. SEP corregida a nivel AM1-IC2D/MM. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), intermedio 2(Int2), productos (P), y las EsT (TS4, TS5 y TS6) se muestran sobre la figura. Se amplía la zona de TS5 y TS6 para facilitar la visualización.

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.13	3.27	1.01	2.22	0
TS4	1.53	1.28	1.10	1.53	29
TS5	1.14	2.49	2.39	1.01	49
Int2	1.15	2.50	2.98	1.01	37
TS6	1.20	1.75	2.77	1.00	48
Productos	2.59	1.13	2.56	0.98	-13

Tabla 8. Distancias (Å) y energías relativas (kcal/mol) de los puntos estacionarios que se sitúan en la SEP IC2D-AM1/MM.

2.7 Ampliación de la Parte QM

Al intentar optimizar los parámetro de van der Waals de la interacción QM/MM entre el piruvato y las moléculas de agua encontramos una desviación sistemática con respecto a los cálculos mecánico cuánticos en vacío (ver Tabla 1). Esta diferencia podría ser debida a efectos de transferencia de carga, término que no se incluye en los cálculos estándar QM/MM. De esta manera podemos pensar que nuestro modelo puede tener una fuente de error importante debido a una posible la transferencia de carga entre el piruvato y el entorno. Esta transferencia de carga con el entorno podría implicar a las dos argininas positivas presentes en el centro activo. El papel de la Arg109 (explicado en la sección 1 (Bioquímica) del presente capítulo) implica la polarización del enlace del grupo carbonilo del piruvato y como consecuencia estabilizar la EsT de la transferencia de hidruro. El papel de la Arg171 es, sin embargo, anclar el sustrato en el centro activo y orientarlo correctamente para la reacción química. Para ilustrarnos sobre esta posibilidad realizamos cálculos de la energía de interacción entre el piruvato y las dos argininas. Los cálculos se realizan en fase gas a nivel HF/6-31G(d), AM1 y AM1/MM. Los modelos utilizados para estos cálculos

se representan en la Figura 17 y las coordenadas han sido tomadas de los reactivos obtenidos en la SEP AM1/MM (Figura 12). Estas energías de interacción, que se pueden observar en la Tabla 9, se han calculado como la diferencia entre las energías de las moléculas por separado (piruvato y arginina) y del complejo entre las dos moléculas. Observamos que los resultados de los cálculos AM1/MM muestran una importante desviación respecto de los valores HF. La mayor contribución a este error no es el uso del Hamiltoniano AM1, ya que cuando todo el sistema es descrito con este método semiempírico la diferencia es mucho menor. La magnitud del error está directamente correlacionada con la magnitud de transferencia de carga (estimada por cargas de Mulliken, Tabla 9) por lo tanto el error en el tratamiento QM/MM lo atribuimos esencialmente a la falta de efectos de transferencia de carga.

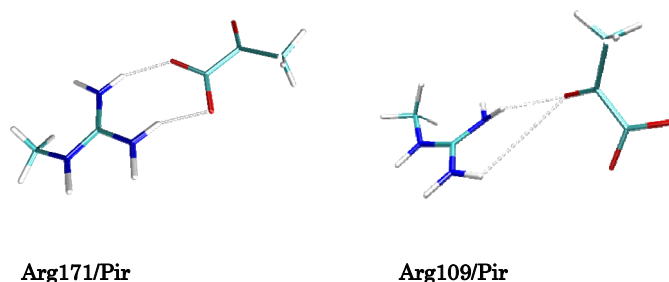


Figura 17. Estructura de los complejos piruvato con la arginina 171 (Arg171/Pir) y del piruvato con la arginina 109 (Arg109/Pir)

	Arg171/Pir		Arg109/Pir	
	E_{int}	Q	E_{int}	Q
HF/6-31G(d)	-111.21	-0.868	-68.76	-0.966
AM1	-98.37	-0.920	-62.80	-0.991
AM1/MM	-77.76	-1.000	-40.18	-1.000

Tabla 9. Energías de Interacción (E_{int} , en kcal/mol) y cargas de Mulliken en el piruvato (Q, en u.a.) para los complejos formados. Las geometrías se tomaron de la estructura de reactivos de la SEP AM1/MM.

Indiscutiblemente, si el error se conserva más o menos constante durante la SEP podríamos ignorar esta contribución. La carga

de los reactivos (piruvato) es la misma que la de los productos (lactato). No obstante, debe tenerse en cuenta que tras la transferencia de hidruro la carga del sustrato aumenta de -1 a -2, pudiendo ser el efecto de transferencia de carga diferente en esa región de la SEP. Por este motivo decidimos explorar de nuevo la superficie AM1/MM utilizando las mismas coordenadas de reacción pero aumentando ahora la región QM. Añadimos las argininas 109 y 171 a los átomos QM que teníamos en el primer modelo, resultando una región QM con 78 átomos, e introduciendo dos nuevos átomos frontera (GHO) en los C γ de cada arginina (teniendo 4 átomos frontera en total).

La nueva SEP se presenta en la Figura 18. Como en la SEP AM1/MM anterior, encontramos dos posibles caminos, con la diferencia de que en el camino donde ocurre primero la transferencia de hidruro no existe un intermedio de reacción. Por lo tanto en la SEP se observan dos caminos concertados. Las geometrías y energías de los puntos estacionarios se pueden encontrar en la Tabla 10. Las barreras de energía potencial no son significativamente diferentes respecto a la SEP obtenida con el subsistema QM pequeño. Podemos decir que la inclusión de las argininas en la región QM permite describir de manera más precisa las interacciones electrostáticas con el intermedio tras la transferencia de hidruro. La sobreestimación de la interacción en el modelo anterior da como resultado una exagerada estabilización del intermedio con carga -2, y también en una fuerte interacción con la His195, con carga positiva. Cuando estas argininas se describen cuánticamente la carga puede estar más deslocalizada, la sobreestimación desaparece y el intermedio deja de ser un punto estacionario. Esta interpretación se confirma con un análisis de la distribución de carga del subsistema QM cuando contiene 52 o 78 átomos.

El oxígeno del piruvato que recibe el protón tiene, en la geometría del intermedio, más carga negativa cuando se incluyen las argininas en la región QM.

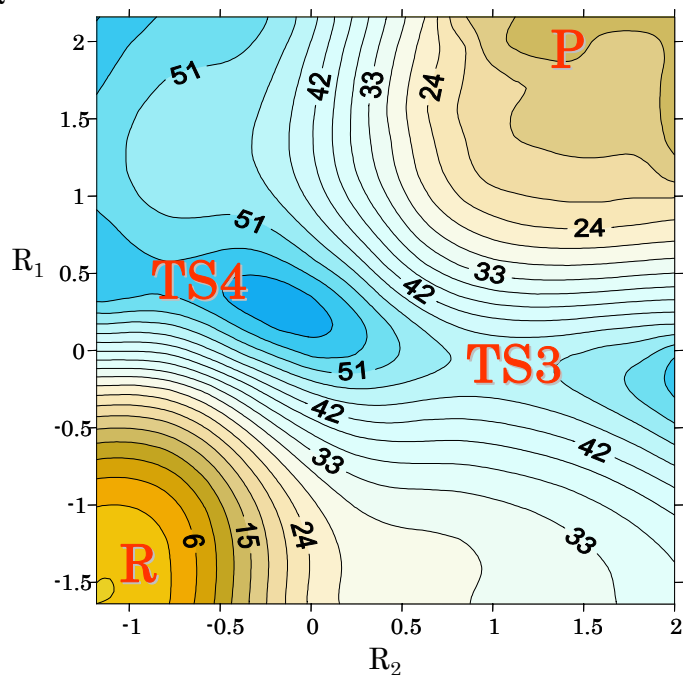


Figura 18. SEP obtenida a nivel AM1/MM con la región QM ampliada, que incluye la Arg109 y la Arg171. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), productos (P), y las EsT (TS3 y TS4) se indican sobre la SEP.

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.13	2.77	1.01	2.17	0
TS4	1.78	1.22	1.03	1.81	54
TS3	1.36	1.39	2.01	0.99	47
Productos	3.08	1.12	2.40	0.98	13

Tabla 10. Distancias (Å) y energías relativas (kcal/mol) de los puntos estacionarios localizados en la SEP AM1/MM obtenida con la región QM ampliada, incluyendo Arg109 y Arg171.

A esta última SEP le incluimos la corrección que hemos decidido que es la más adecuada para nuestro sistema, es decir, las correcciones interpoladas bidimensionales (IC2D). La SEP corregida se presenta en la

Figura 19. Las distancias y energías más relevantes de los puntos estacionarios de esta nueva SEP están recogidas en la Tabla 11. En este caso tenemos una diferencia importante con respecto a todas las SEP anteriormente exploradas; en ésta encontramos un único camino de reacción. La EsT describe un mecanismo concertado donde está considerablemente más avanzada la transferencia de hidruro que la de protón. Además, la reacción pasa a ser isoenergética, lo que concuerda con la evidencia experimental. El problema aparece en el valor de la barrera de energía potencial, mucho más alto del esperado.

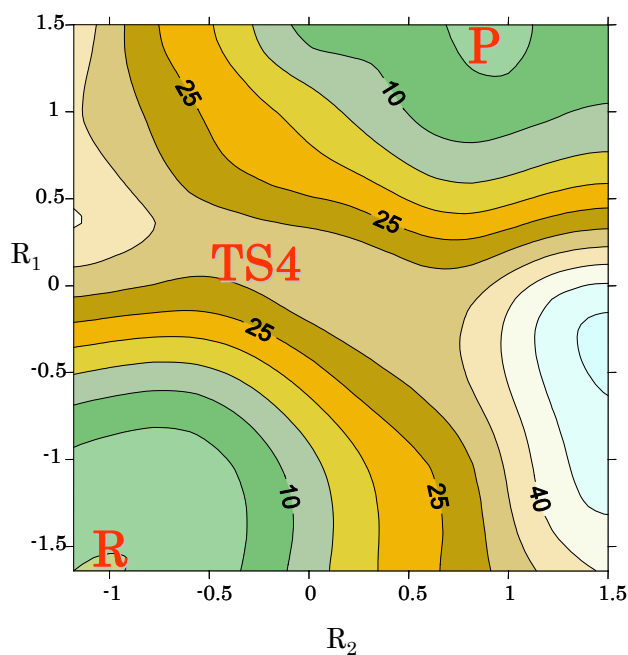


Figura 19. SEP obtenida a nivel IC2D-AM1/MM con la región QM ampliada incluyendo la Arg109 y la Arg171. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), productos (P), y la EsT (TS4) se muestran cualitativamente.

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.13	2.77	1.01	2.17	0
TS4	1.62	1.26	1.12	1.50	30
Productos	3.06	1.12	2.20	0.95	-1

Tabla 11. Distancias (Å) y energías relativas (kcal/mol) de las estructuras estacionarias que aparecen en la SEP IC2D-AM1/MM cuando la región QM incluye Arg109 y Arg171.

2.8 Estado de Protonación de los Residuos

A pesar de las correcciones realizadas al Hamiltoniano semiempírico y de la ampliación de la región QM, las barreras de energía potencial son muy altas, teniendo en cuenta que se trata de una reacción catalizada por una enzima. Tal y como se ha expuesto anteriormente en la sección de bioquímica del este capítulo, en la LDH la constante de velocidad de la etapa química es al menos de 750 s^{-1} . Aplicando la teoría variacional del estado de transición (VTST), esta constante catalítica corresponde a una energía libre de activación máxima de 15 kcal/mol a 298K, mientras que la barrera de energía potencial obtenida en nuestra mejor aproximación a la SEP es de 30 kcal/mol. Este dato nos indujo a pensar que el modelo utilizado para estudiar la reacción contenía algún error adicional a los considerados.

Al construir el modelo a estudiar hemos supuesto un estado de protonación de los aminoácidos que constituyen la enzima según su pK_a en disolución acuosa. No obstante, como se ha explicado en la Sección 6 del Capítulo 3 de la presente Tesis Doctoral, la influencia del entorno puede alterar el valor del pK_a de un residuo, afectando su estado de protonación en función del pH del medio. Para calcular el pK_a de los aminoácidos se parte de la estructura inicial de RX. El programa implementado por Field y colaboradores^[44] incluye en su cálculo la

adición de los hidrógenos polares de la proteína y su minimización. Estos cálculos se realizan con el cofactor NADH y el sustrato presentes con los hidrógenos polares ya dispuestos. Tras llevar a cabo el cálculo podemos observar una serie de aminoácidos cuyos pK_a calculados en el seno de la enzima varían notablemente respecto a los valores estándar en disolución (Tabla 12). El resultado más significativo es el cambio de pK_a del Glu199 que pasa de ser ácido a ser ligeramente básico y por lo tanto si tenemos el medio a pH 7 deberíamos protonar este residuo. En la Figura 20 se observa la situación de este aminoácido en las cercanías del centro activo. La adición de este hidrógeno permite establecer un enlace de hidrógeno fuerte entre el grupo carboxilo del Glu199 y el Asp168. Éste muestra un carácter más ácido que en disolución acuosa. El Asp168 juega a su vez un papel importante estabilizando la forma protonada de la His195, que presenta un pK_a de 8.7 en lugar del valor estándar en disolución que es 6.0. Otro cambio de estado de protonación a pH neutro lo encontramos en una histidina (His95) que se encuentra alejada del centro activo pero que resulta tener un pK_a de 7.6, sensiblemente superior al valor tomado como referencia. Por último, señalar que los cálculos para las histidinas neutras mostraron que cuatro de ellas (157, 188, 266, 304) tiene protonado el NE2 en lugar del ND1.

	Solución acuosa ^[45, 46]	LDH
His195	6.0	8.7
Asp168	3.9	1.2
Glu199	4.1	7.4
His95	6.0	7.6
Arg109	12.8	14.0

Tabla 12. pK_a calculados para algunos aminoácidos en el seno de la enzima comparados con el valor que presentan en disolución acuosa.

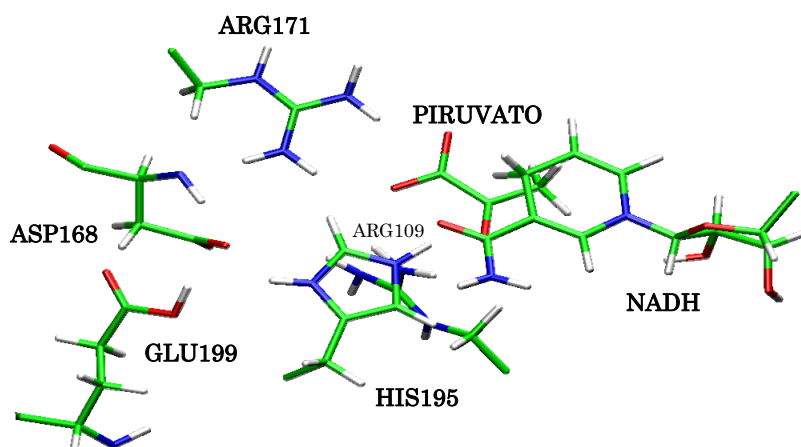


Figura 20. Detalle del centro activo de la LDH donde se incluye la posición del Asp168 y del Glu199 respecto a la His195.

El resto de aminoácidos conservan el mismo estado de protonación a pH 7 que el asumido inicialmente por valor estándar del pKa. La Arg109 conserva su carácter básico incluso se puede destacar que este residuo ve aumentado el valor de su pKa.

Una vez establecidas las diferencias de protonación y construido el nuevo modelo con las modificaciones, se solvata y relaja el sistema tal y como lo habíamos hecho para el anterior modelo. A continuación llevamos a cabo una exploración de la SEP con la parte QM ampliada (78 átomos) a nivel AM1/MM para comprobar si hay algún cambio en el comportamiento de la SEP. La Figura 21 muestra la nueva SEP. Las distancias y energías relativas correspondientes a los puntos estacionarios se dan en la Tabla 13. La nueva SEP presenta dos posibles mecanismos de reacción que difieren respecto a la SEP AM1/MM en la altura de las barreras (Figura 18). Mientras que en la Figura 18 encontrábamos barreras de 54 kcal/mol y 47 kcal/mol, en esta nueva

SEP, las barreras son de 44 y 33 kcal/mol, respectivamente. Así pues cambio de interacciones electrostáticas con la parte MM del sistema que no están directamente implicadas en la reacción química disminuye en más de 10 kcal/mol las barreras de energía potencial en la reacción química, siendo el descenso más importante para la transferencia protónica que para la de hidruro. Las EsT de esta SEP corresponden a procesos concertados asincrónicos en los dos mecanismos. Mientras que TS1 presenta la transferencia de hidruro más avanzada, en el TS2 más avanzada la transferencia de protón precede a la del hidruro.

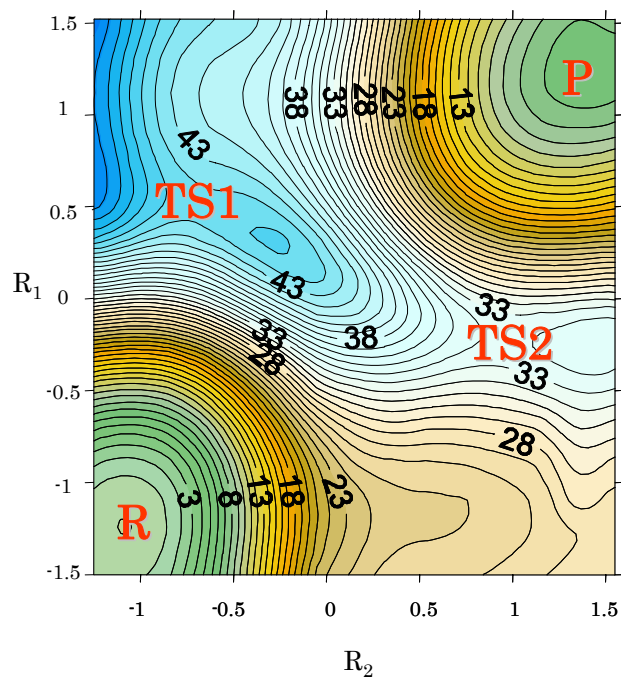


Figura 21. SEP obtenida a nivel AM1/MM (con cambio en la protonación de aminoácidos). La región QM incluye la Arg109 y la Arg171. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente.

	$r_{\text{C}_{\text{nic}}\text{H}_1}$	$r_{\text{C}_{\text{pir}}\text{H}_1}$	$r_{\text{N}_{\text{his}}\text{H}_2}$	$r_{\text{O}_{\text{pir}}\text{H}_2}$	Energía
Reactivos	1.12	2.34	1.05	1.90	0
TS1	1.69	1.21	1.04	1.74	44
TS2	1.27	1.47	1.92	0.99	33
Productos	2.13	1.13	2.00	0.99	-3

Tabla 13. Distancias (\AA) y energías relativas (kcal/mol) de los estados estacionarios localizados con el programa GRACEFUL^[47] en la SEP AM1/MM presentada en la Figura 21.

En la Figura 22 se presenta la SEP con los pK_a s recalculados e incluyendo las correcciones interpoladas bidimensionales (IC2D), mientras que en la Tabla 14 se dan las distancias y energías relativas de los puntos estacionarios localizados. En esta SEP se localiza una sola EsT que corresponde a un mecanismo concertado asincrónico con una barrera de 21 kcal/mol. Este valor de la barrera de energía potencial está más en concordancia con la que se extrae a partir de la constante de velocidad experimental. La EsT presenta una transferencia de hidruro sensiblemente más avanzada que la de protón, pero es un mecanismo mucho más sincronizado en comparación al resto de mecanismos concertados que hemos ido encontrando al explorar las diferentes SEPs. Concretamente el hidruro está a mitad camino entre los átomos dador y aceptor dando lugar a una coordenada distinguida R_1 de 0.12 \AA ; mientras que la coordenada distinguida que describe el proceso de la transferencia de protón, R_2 , es de -0.48 \AA ; por tanto está mucho más cerca del átomo dador que del aceptor.

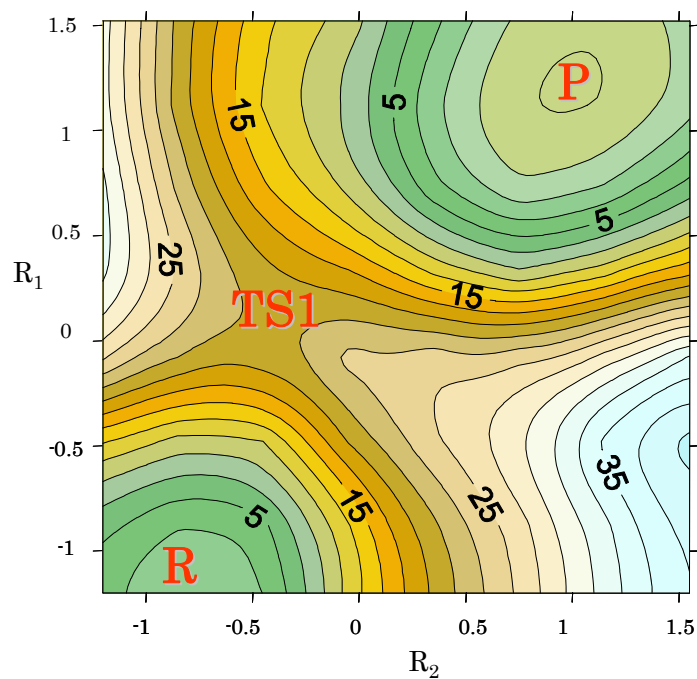


Figura 22. SEP obtenida con correcciones IC2D-AM1/MM y recalculando los pK_{as} . La región QM incluye Arg109 y Arg171. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. La posición de las estructuras de reactivos (R), productos (P), y la EsT (TS1) se muestran cualitativamente

	$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	Energía
Reactivos	1.12	2.36	1.05	1.77	0
TSA	1.43	1.31	1.09	1.57	21
Productos	2.35	1.13	1.93	0.94	-5

Tabla 14. Distancias (Å) y energías relativas (kcal/mol) de los estados estacionarios localizados en la SEP IC2D-AM1/MM con los pK_{as} de los aminoácidos recalculados con el programa GRACEFUL.^[47]

2.9 Estudio de la cinética de la reacción química

Tal como hemos puesto de manifiesto en los capítulos anteriores las enzimas son esenciales para los organismos vivos debido a que consiguen que las reacciones químicas ocurran en una escala de tiempo compatible con la vida.^[48] La habilidad de estas proteínas para aumentar la velocidad de los procesos químicos, disminuyendo la barrera de activación de energía libre, ha sido centro de estudio en un elevado número de aproximaciones teóricas durante los últimos años.^[49-59] Sin embargo, lo máximo que podemos afirmar sobre esta cuestión química fundamental es que tan sólo estamos arañando la superficie de un problema muy complejo. Aunque ciertas estrategias computacionales han tenido éxito en la reproducción de algunos valores experimentales, el conocimiento profundo de la maquinaria molecular que está implicada en las reacciones enzimáticas todavía requiere la solución de muchos desafíos conceptuales pero también a nivel práctico.

Los estudios computacionales de procesos enzimáticos normalmente implican la exploración del espacio configuracional del sistema a lo largo de una coordenada particular que define los cambios entre el estado reactivo y el estado producto. Esta coordenada (ζ) puede ser un parámetro geométrico simple relacionado con la rotura o formación de enlaces. Por otra parte, puede ser también definida una coordenada colectiva para seguir los cambios asociados no solo al sustrato sino también a la enzima. Una de estas coordenadas colectivas empleadas en el contexto de la teoría de enlace de valencia es la diferencia de energía entre curvas diabáticas definiendo los estados reactivo y producto. Independientemente de la elección realizada para una reacción particular, el objetivo es obtener la diferencia de energía

libre entre reactivos y productos; y entre reactivos y el máximo en el perfil de energía libre. Mientras la primera magnitud está relacionada con el proceso de equilibrio, la segunda nos permite obtener la constante de velocidad de la reacción (k_r). La teoría del estado de transición,^[60, 61] o su extensión variacional, relaciona la energía libre de reacción (ΔG^\ddagger) con la constante de velocidad mediante la ecuación (5.18). La energía de activación puede obtenerse de un modo aproximado como la diferencia entre el máximo y el mínimo (en el lado de reactivos) del perfil del energía libre trazado a lo largo de la coordenada de reacción (ζ), aunque algunos términos relacionados con la coordenada de reacción deben ser considerados, en particular cuando usamos coordenadas colectivas. Cuando una reacción tiene lugar en un entorno complejo, como puede ser el centro activo de una enzima, un gran número de cambios geométricos ocurren no sólo en el subsistema químico (alargando o acortando enlaces) sino también en el entorno (por ejemplo, enlaces de hidrógeno con el sustrato y/o entre residuos del centro activo pueden formarse o romperse durante el proceso). Todos estos cambios están acoplados y pueden no reflejarse completamente en una coordenada geométrica simple, asociada al sustrato. Cuando obtenemos el perfil de energía libre bajo el supuesto de un entorno en equilibrio con el sustrato, algunos movimientos del entorno están perfectamente estudiados. En muchos casos, esto se puede solucionar por medio del coeficiente de transmisión tomando valores cercanos a la unidad. Para estimar satisfactoriamente esta dependencia de la coordenada con el coeficiente de transmisión se puede emplear teorías fenomenológicas o simulaciones moleculares.^[62-66]

Sin embargo, resulta difícil de simular, cuando la transformación de reactivos a productos no puede ser descrita únicamente con una coordenada geométrica o una combinación simple.

Una característica de estas reacciones es que dos eventos químicos tienen lugar de modo concertado. Un ejemplo de estos casos es la doble transferencia de hidruro y protón en la LDH. Cada una de las transferencias puede describirse correctamente por medio de una combinación antisimétrica de la distancia entre el hidrógeno y los átomos dador y aceptor. Pero si se describe un proceso concertado utilizando una combinación de parámetros geométricos no se puede asegurar una correcta exploración de todas las configuraciones relevantes durante el camino concertado. En términos de coordenadas colectivas, esta situación puede corresponder al caso donde más de dos estados diabáticos participan en el proceso.

El problema de la selección de configuraciones relevantes del sistema total (sustrato más entorno) para los procesos objeto de análisis es el acoplamiento entre el paso químico y los cambios configuracionales.^[67-70] De acuerdo con estos resultados las diferentes configuraciones de la enzima pueden conducir a diferentes caminos de reacción, cada uno contribuyendo al flujo total de la reacción.^[67, 69, 71] La participación de los diferentes conformeros del complejo de Michaelis (complejo sustrato-enzima) a la constante de velocidad total dependería de las diferentes barreras de energía libre y la energía libre relativa entre ellos. Los casos limitantes corresponderían a: i) un control termodinámico, cuando la concentración de los diferentes complejos enzima-sustrato se encuentran en sus valores de equilibrio porque la constante de interconversión es mucho mayor que la constante del paso químico; ii) un control cinético, cuando las barreras de energía libre de interconversión entre los diferentes complejos es mayor que la barrera asociada al paso químico de la reacción. Por supuesto, también es posible que existan comportamientos intermedios, dando lugar a una enorme

variedad de posibles comportamientos dependiendo de peculiaridades del sistema. Como consecuencia se obtiene una distribución de constantes de reacción y la interpretación macroscópica de la constante catalítica puede ser más complicada de lo esperado.^[67] En concreto, cuando una interconversión entre diferentes complejos reactivos es más lenta comparada con la etapa química, la constante catalítica se obtiene como una ponderación de las constantes catalíticas correspondientes a cada conformero.^[67] Experimentos basados en la enzima β -galactosidasa han mostrado que una constante de velocidad de una molécula puede abarcar un amplio rango de escala de tiempo, desde 10 a 10^{-3} s.^[72] Esto significa que la barrera de energía libre puede variar hasta 5 kcal/mol dependiendo de la conformación desde la cual se ha obtenido. Estos resultados pueden relacionarse con resultados previos encontrados experimentalmente en la mioglobina, mostrando que los estados de la proteína presentan una estructura jerárquica y así es posible encontrar diferentes valles en los valles de reactivos.^[73] Cada uno de los diferentes estados de la proteína tiene una función diferente pero las diferentes configuraciones (valles dentro de los valles) tienen la misma función pero con diferentes constantes.

Desde un punto de vista teórico estos resultados significan que la simulación molecular de una reacción enzimática debe tener en cuenta diferentes caminos de reacción iniciándolos desde diferentes configuraciones de reactivos y también los caminos que conducen de un valle a otro.^[3, 71, 74-76] El uso de una única barrera de energía libre para caracterizar una reacción enzimática puede conducir a una incorrecta interpretación del proceso. Por ello, habrá que tener en cuenta que, primero, sólo aquellas simulaciones suficientemente largas podrán permitir un correcto muestreo de los diferentes complejos reactivos

implicados en el proceso y los caminos de reacción asociados a cada uno de los valles. En la práctica, la exploración de una región limitada del 'paisaje' de energías de la proteína no puede asegurar que se han incluido todos los conformeros relevantes. Segundo, este proceso no informa directamente sobre las transiciones entre los diferentes complejos. Como hemos indicado las fluctuaciones de la proteína pueden conducir el sistema de un valle de reactivos a otro desde el cual la barrera de la reacción química puede ser sustancialmente diferente. Una única barrera de energía libre puede ser adecuada para procesos enzimáticos en el que existe un número limitado de posibles configuraciones, mientras que para proteínas que presenten grandes fluctuaciones puede conducir a resultados no relevantes. Por el contrario, se ha sugerido que un número limitado de perfiles de energía libre puede ser promediado para generar una barrera de energía libre comparable con los valores experimentales.^[71] El problema en este sentido es la manera en la cual los perfiles deben ser ponderados, dependiendo si se supone el equilibrio entre los complejos reactivos o no.

La enzima LDH ofrece una excelente oportunidad para el análisis de estos factores ya que, por un lado la reacción química que cataliza supone una doble transferencia de hidrógeno, y por otro lado estudios previos de esta reacción enzimática en términos de energías potenciales han dejado indicios de la flexibilidad de esta proteína.^[77]

2.9.1 Exploración de la región del ET

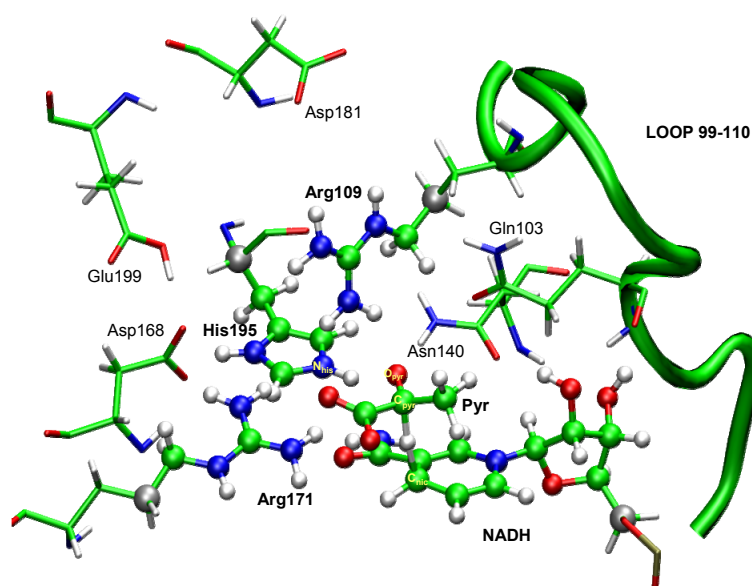


Figura 23. Estructura de transición de la transformación piruvato a lactato en la enzima LDH: La región QM se representa por medio de esferas y cilindros. La posición de los átomos frontera GHO son las esferas grises. Las etiquetas en negrita indican que la molécula está, al menos parcialmente definida en la región QM.

Nuestro estudio computacional comienza desde la región del ET. Para ello seleccionamos como configuración inicial del sistema la estructura de transición localizada y caracterizada después de la exploración de la SEP obtenida a nivel AM1-IC2/MM. Esta EsT pertenece a un camino de reacción concertado asincrónico donde la transferencia del hidruro está ligeramente más avanzada que la del protón. En esta estructura el piruvato está fijado por el grupo carboxilato a la Arg171 mientras que el oxígeno carbonílico del piruvato establece enlaces de hidrógeno con la Arg109. Otros residuos relevantes, como la His195 y el NADH se muestran en la Figura 23.

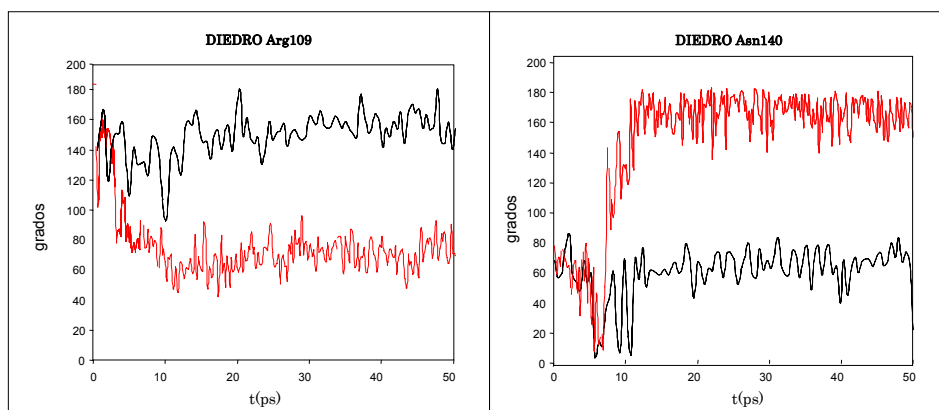


Figura 24. Evolución en el tiempo de simulación del diedro CA-CB-CD-CG de la Arg109 y el diedro CA-CB-CD-N de la Asn140. En negro se presenta la evolución de la dinámica a una temperatura de 298K y en rojo la dinámica a 300K.

Para explorar la región del ET llevamos a cabo una simulación de dinámica molecular del sistema fijando los valores de las coordenadas R_1 y R_2 descritos en las ecuaciones (2.1) y (2.2) del presente capítulo. Tras el calentamiento a 298K se lleva a cabo dos simulaciones paralelas a diferentes temperaturas (298 y 300K) para lograr un amplio rango de exploración de las configuraciones del sistema. Estas dos simulaciones exploran (de hecho) regiones del espacio configuracional con importantes diferencias desde el punto de vista de la catálisis. La Figura 24 muestra la evolución del ángulo diedro CA-CB-CD-CG y CA-CB-CD-N de la Arg109 y Asn140. La diferente orientación alcanzada en la DM de 300K después de algunos picosegundos implica diferentes interacciones de enlaces de hidrógeno con el oxígeno carbonílico del piruvato. En la DM de 298K la Arg109 establece un enlace de hidrógeno con el oxígeno carbonílico del sustrato (O_{pir}) pero en la DM de 300K tras varios picosegundos este residuo se sitúa a una distancia mayor y la interacción se establece con la Asn140. Esta competición entre estos dos residuos había sido ya observada en estudios computacionales de nuestro grupo de investigación^[77] y se propuso un mecanismo de estabilización alternativa

que puede explicar el hecho de que la mutación de la Arg109 no aniquile completamente el efecto catalítico de la enzima.^[16] Seleccionamos diferentes estructuras de las dos DM como puntos iniciales para localizar y caracterizar EsT del proceso sobre las superficies de energía potencial.

$r_{C_{nic}H_1}$	$r_{C_{pir}H_1}$	$r_{N_{his}H_2}$	$r_{O_{pir}H_2}$	$r(O_{pyr}HH_{Arg109})$	$r(O_{pyr}HD_{Arn140})$	Energía	Frec	ΔE
1.43	1.31	1.09	1.57	3.04	2.01	-17907	776i	17
1.39	1.36	1.10	1.59	2.20	1.88	-18213	1186i	23
1.58	1.24	1.20	1.35	4.49	1.88	-18382	630i	36
1.39	1.35	1.08	1.66	2.14	1.95	-18165	1095i	34
1.63	1.24	1.16	1.43	4.39	1.82	-18305	462i	32
1.69	1.23	1.15	1.43	4.03	1.88	-18329	356i	83
1.61	1.25	1.14	1.45	4.39	1.86	-18354	556i	37
1.58	1.25	1.17	1.41	4.60	1.88	-18383	352i	34
1.57	1.25	1.15	1.43	4.85	1.84	-18410	619i	22
1.57	1.24	1.18	1.38	3.85	1.89	-18419	222i	45
1.54	1.26	1.12	1.52	4.94	2.00	-18419	193i	23
1.52	1.27	1.11	1.55	4.99	1.85	-18418	257i	23
1.55	1.25	1.12	1.52	5.40	1.87	-18470	215i	40
1.53	1.26	1.12	1.53	4.81	1.85	-18446	284i	18
1.55	1.25	1.14	1.45	4.92	1.86	-18421	352i	22
1.57	1.25	1.15	1.43	5.14	1.85	-18426	501i	38
1.56	1.24	1.18	1.37	4.85	1.84	-18410	541i	22
1.54	1.25	1.18	1.37	4.75	1.84	-18429	284i	41
1.55	1.25	1.14	1.45	4.90	1.87	-18379	317i	38
1.53	1.26	1.13	1.48	5.19	1.85	-18426	476i	56
1.55	1.26	1.16	1.41	4.63	1.85	-18402	720i	21
1.40	1.34	1.08	1.65	2.10	1.97	-18183	1058i	19
1.41	1.32	1.08	1.62	2.29	1.94	-18187	850i	45
1.55	1.24	1.11	1.52	2.30	2.09	-18326	115i	62

Tabla 15. Estructuras de transición de la etapa química localizadas. Distancias en (Å), energía potencial (kcal/mol) y frecuencias imaginarias en cm^{-1} y barreras de energía potencial (kcal/mol).

La búsqueda de EsT se lleva a cabo con la combinación de programas CHARMM/GRACEFUL. Las 24 EsT localizadas y caracterizadas como puntos de silla de orden 1 se muestran en la Tabla 15. En la tabla se puede encontrar las distancias de los enlaces que se rompen y se forman

durante la reacción química. Las estructuras localizadas muestran una dispersión importante en términos geométricos, lo que claramente muestra la dependencia de las EsT con la configuración seleccionada de partida. En todos los casos la EsT obtenida describe un proceso concertado asincrónico donde la curvatura local parece extremadamente sensible a las fluctuaciones del sistema; la frecuencia imaginaria cambia en un factor de diez: desde $115i$ a $1188i$ cm^{-1} . El máximo de la función de distribución de la frecuencia imaginaria aparece en torno a $300i$ cm^{-1} , pero ésta tiene una fluctuación que alcanza valores muy superiores. Con respecto a la transferencia de hidruro, la distancia del enlace que se rompe se encuentra entre los valores de 1.39 y 1.60 Å mientras que el enlace que se forma tiene un intervalo menor, entre 1.24 y 1.36 Å. En el caso del protón el enlace que se rompe está entre valores de 1.08 y 1.66 Å y el que se forma entre 1.37 y 1.66 Å. La dispersión de las EsT observada en Tabla 15 se muestra gráficamente en Figura 25 donde las estructuras localizadas se representan en función de R_1 y R_2 . Para todas estas estructuras trazamos los IRC hasta llegar a los valles de reactivos y de productos. Las barreras se obtienen como la diferencia de energía potencial entre la EsT y la estructura de reactivos, y presentan valores entre 17 y 83 kcal/mol. Esta distribución tan amplia de barreras de energía no nos sorprende en absoluto. En el estudio de la transferencia de hidruro en la DHFR se obtuvieron resultados similares.^[71, 74]

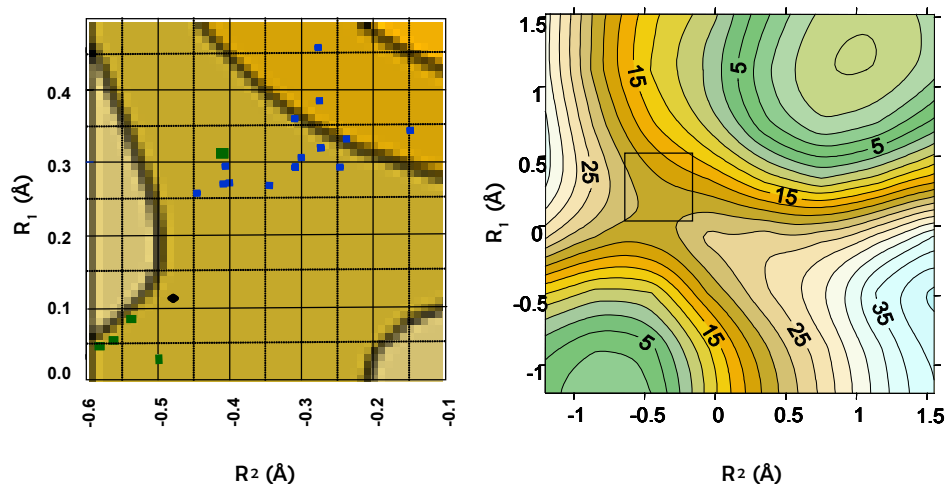


Figura 25. Representación de diferentes EsT localizadas sobre la SEP tras realizar dinámica molecular de la EsT inicial. El punto negro es la estructura de la cual parten las DM. En verde se representan las estructuras extraídas de la dinámica de 298K y en azul las estructuras de la dinámica de 300K.

2.9.2 Análisis del Estado de Reactivos

La idea básica que queremos expresar en este apartado es que el complejo reactivo enzima-sustrato presenta importantes diferencias que se obtienen al iniciar las simulaciones comenzado desde diferentes puntos iniciales. Efectivamente, las trayectorias obtenidas a partir del IRC trazadas desde las diferentes EsT y seguir direcciones ligeramente diferentes (los vectores de transición de cada EsT de partida) puede llevar a alcanzar valles de reactivos cualitativamente diferentes. Este proceso puede relacionarse como una especie de *simulated annealing* para encontrar esos mínimos del sistema relevantes para el proceso químico objeto de estudio. Para comprobar este punto seleccionamos 4 estructuras de reactivos obtenidas desde los IRCs de las EsT que aparecen en la Tabla 15. En concreto, utilizamos las estructuras de reactivos provenientes de las EsT marcadas en negrita en la Tabla 15.

Creemos que con esta selección se cubre un intervalo significativo de posibles barreras de reacción (desde 17 a 36 kcal/mol).

Se llevaron a cabo simulaciones de dinámica molecular (NVT) relativamente largas (500 ps) para cada uno de los complejos reactivos. El análisis de estas trayectorias muestra diferencias significativas durante el tiempo de simulación. Por supuesto, nos preguntamos si los tiempos de simulación muestran una superposición importante entre las cuatro trayectorias en el espacio configuracional. Sin embargo, desde un punto de vista práctico, y considerando los tiempos de simulación normalmente empleados para obtener energías libres, podemos considerar estos complejos como diferentes valles que contribuyen cada uno de ellos en el flujo macroscópico de la reacción. En la Tabla 16 se muestran los valores promedios y las fluctuaciones de algunos parámetros geométricos relacionados con el piruvato y obtenidos durante las dinámicas de estos cuatro complejos reactivos.

	REAC1	REAC2	REAC3	REAC4
d(C _{nic} H ₁)	1.13 (0.03)	1.13 (0.02)	1.13 (0.02)	1.13 (0.03)
d(C _{pyr} H ₁)	4.3 (0.7)	3.3 (0.3)	3.5 (0.4)	3.20 (0.2)
d(N _{his} H ₂)	1.00 (0.02)	1.01 (0.02)	1.01 (0.02)	1.01 (0.02)
d(O _{pyr} H ₂)	3.0 (0.2)	3.4 (0.4)	3.1 (0.2)	3.0 (0.2)
d(O _{pyr} H _{nic})	3.9 (1.0)	2.5 (0.3)	3.2 (0.7)	2.7 (0.4)
d(O _{pyr} HH _{Arg109})	7.1 (0.4)	6.9 (0.6)	6.0 (0.6)	6.7 (0.6)
d(O _{pyr} HD1 _{Asn140})	6.3 (1.4)	2.6 (0.3)	4.8 (0.7)	6.1 (0.9)
d(O _{pyr} HH _{Arg171})	5.4 (0.2)	5.5 (0.5)	4.0 (0.2)	5.2 (0.5)
d(O _{pyr} H _W)	5.2 (0.7)	6.0 (0.7)	5.8 (0.3)	4.6 (0.6)

Tabla 16. Geometrías promedio de los cuatro complejos reactivos analizados. Las distancias están dadas en Å y la desviación estándar está entre paréntesis.

Con respecto a los parámetros asociados directamente a las transferencias de hidrógeno, la distancia con dador-hidrógeno (C_{nic} - hidruro y N_{his} - protón) son iguales para las cuatro simulaciones. Este resultado es lógico considerando que se trata de enlaces formados.

Obviamente, y debido a una débil constante de fuerza asociada a estos parámetros, las fluctuaciones y las diferencias son mucho mayores para las distancias aceptor hidrógeno. En particular, el estado reactivos obtenido del TS1 (REAC1) muestra una distancia $C_{\text{pyr}}\text{H}$ más larga que en el resto de complejos reactivos. La desviación estándar también es significativamente grande reflejando el hecho de una fluctuación amplia observada en el REAC1. Esto parece mostrar que las diferencias en las distancias aceptor hidrógeno en el REAC1 deben ser atribuidas a un diferente comportamiento del piruvato en el centro activo. De hecho, la trayectoria REAC1 presenta las mayores fluctuaciones, expresadas como valor promediado respecto a su desplazamiento RMS. En las cuatro trayectorias reactivas analizadas el sustrato está anclado por medio de una interacción de enlace de hidrógeno entre los oxígenos carboxílicos del piruvato y la Arg171. Esta interacción se conserva durante las cuatro simulaciones. Las diferencias aparecen entre las interacciones del oxígeno carbonílico de piruvato (O_{pyr}). Las interacciones establecidas con este átomo en cada simulación son bastante diferentes. En todos los casos existe una fuerte interacción con el hidrógeno de la histidina protonada que va a ceder el protón (His195). La función y la posición de este residuo se modula por medio de la carga negativa del Asp168, el cual establece a su vez interacciones con la Arg171 y el Glu199 protonado. El pK_a de este último residuo, debido a su peculiar entorno, resultó ser significativamente diferente respecto a su valor estándar en disolución acuosa y en concreto por su proximidad al Asp168 (ver Tabla 16). Otro resultado que resulta interesante puntualizar es, que en ninguna simulación se establece enlace de hidrógeno entre el O_{pyr} con Arg109. Considerando que esta interacción puede aparecer en la EsT (al menos en alguna de ellas como discutimos en el apartado anterior) se puede

concluir que las coordenadas de este residuo deberían estar incluidas en la correcta definición de la coordenada de reacción, al menos para algunos posibles caminos de reacción. La interacción $O_{\text{pyr}} - \text{Arg109}$ observada en algunas EsT debe establecerse durante el proceso de la reacción química. Considerando que la transferencia de hidruro precede a la de protón en el proceso concertado asincrónico, debería ser un argumento para que la carga negativa que se transfiere al sustrato (o una fracción de la carga) pueda estar acoplada a la aproximación de este residuo. Por otra parte, la Asn140, que es el residuo que alternativamente puede hacer la función de estabilizar la EsT y mantener la actividad catalítica ante una mutación de la Arg109; forma un enlace de hidrógeno con el O_{pyr} en la trayectoria de REAC2 y se conserva durante todo el camino de reacción y también en la EsT. La existencia de este enlace de hidrógeno previene la formación de un enlace de hidrógeno adicional con la Arg109. Como hemos discutido en el análisis de la trayectoria del ET la posición de estos dos residuos está relacionada y parecen alternarse. Finalmente, es muy interesante comentar el caso particular observado en la simulación REAC2. En este caso el sustrato adopta una orientación diferente en el centro activo. Efectivamente, en esta simulación la molécula de piruvato se ve rotada a lo largo de los ejes perpendiculares a un plano molecular en el cual la Arg171 se aloja entre el oxígeno carbonílico del piruvato. En este caso podemos observar una interacción entre el O_{pyr} y la Arg171, como representamos en el promedio reducido de la distancia en la Tabla 16. Como resultado de todas estas interacciones descritas entre la molécula de piruvato y la enzima, podemos concluir que en el primer estado reactivo, REAC1, la molécula de piruvato está menos retenida y puede realizar más movimientos en el centro activo. Aparte de la interacción

con la His195, el oxígeno carbonílico muestra también interacciones con la Asn140 y Arg171. Esto puede ser la razón de las fluctuaciones mayores observadas para la distancia hidruro aceptor en el REAC1, como se confirma en el promedio del desplazamiento del piruvato (RMS) durante las simulaciones.

Para entender este comportamiento analizamos una posible compresión realizada por los residuos Ile253 y Val33 que están situados en la otra cara del anillo de nicotidamina perteneciente al NADH con respecto al piruvato. En un reciente estudio de la isoenzima presente en el corazón humano, h-H₄LDH, se sugería que el residuo que ocupaba una posición similar (Val31) podría jugar un papel favoreciendo la transferencia de hidruro.^[78] Sin embargo el análisis del promedio de las distancias de los residuos Ile253 y Val33 al C_{nic} en nuestras dinámicas no muestra estadísticamente diferencias significativas.

Además de las interacciones enzima-sustrato, hemos encontrado diferencias en las interacciones con moléculas de aguas. A partir del promedio el promedio del número de aguas que forman enlaces de hidrógeno con algún átomo de la región QM en nuestro sistema, se ha observado que este número promedio es próximo a 8 para las simulaciones correspondientes a REAC1, REAC2, REAC4, pero para REAC3 el valor promedio es 10.5. La desviación estándar en todos los casos es ± 1 . Una de las aguas adicionales en REAC3 forma un enlace de hidrógeno con el átomo del piruvato O_{pyr}. Esta interacción se encuentra en el 32% de las configuraciones analizadas de esta dinámica siendo el promedio de la distancia O_{pyr}-H_{wat} 2.1 Å (± 0.2). Una interacción de este tipo entre el sustrato y moléculas de agua en la simulación REAC1 se observó sólo en el 3% de las estructuras y con menor frecuencia para las simulaciones REAC2 y REAC4.

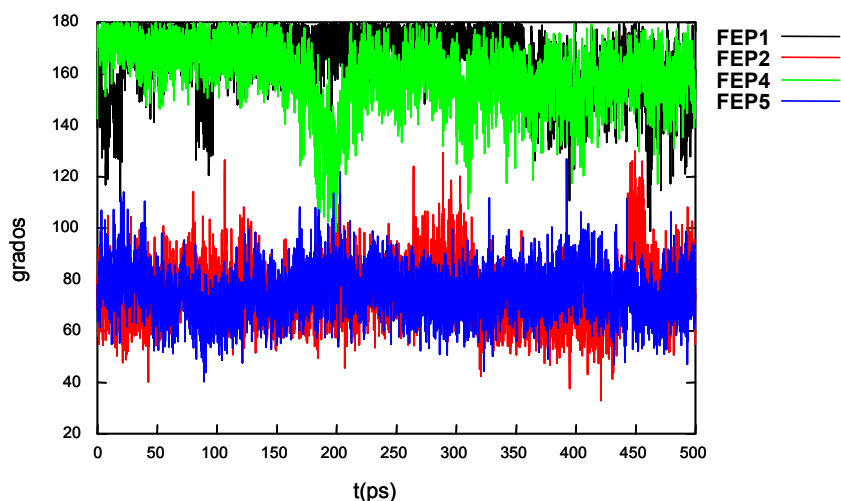


Figura 26. Evolución en el tiempo del diedro CA-CB-CD-CG de la Arg109 durante 500ps de simulación para los cuatro reactivos analizados.

También encontramos diferencias significativas entre las simulaciones analizando el comportamiento de residuos claves de la enzima, en particular la Arg109. Como se ha explicado, este residuo es parte de un bucle flexible cercano al centro activo. La Figura 26 muestra la evolución del diedro CA-CB-CD-CG de la Arg109 durante los 500 ps de simulación de los 4 complejos reactivos. Mientras en el REAC2 y REAC4 este diedro fluctúa hasta 90 grados, en el caso de REAC1 y REAC3 la fluctuación llega a 180 grados. Esta orientación de la Arg109 se conserva durante toda la simulación. Consecuentemente este residuo puede establecer diferentes interacciones en cada caso. La Tabla 17 muestra el promedio de las distancias de enlace de hidrógeno y desviación estándar entre la Arg109 y algunos residuos y moléculas de agua. Tanto en REAC2 como en REAC4 hemos observado una interacción de enlace de hidrógeno con una molécula de agua. Esta característica no se observa en las simulaciones REAC1 y REAC3, lo que puede relacionarse con la diferente disposición de la cadena lateral. Además, la Arg109 también forma enlaces de hidrógeno con la Gln103 y el Asp181 en

REAC2. En REAC3 se observa una mayor frecuencia de contacto con la Tyr224. En el caso de REAC1 la Arg109 forma un fuerte enlace de hidrógeno con el Asp181, el cual se conserva durante toda la simulación. En la simulación REAC3 se observa una interacción ocasional de la Arg109 tanto con el Asp181 como con la Tyr224.

	REAC1	REAC2	REAC3	REAC4
d(HH11 _{Arg109} -OD1 _{Asp181})	5.2 (0.2)	3.5 (1.3)	7.1 (0.8)	8.2 (0.8)
d(HH12 _{Arg109} -OD1)	5.4 (0.3)	3.2 (1.6)	6.5 (1.5)	8.1 (1.1)
d(HH21 _{Arg109} -OD1)	2.6 (0.7)	5.5 (0.6)	4.2 (0.6)	6.0 (0.5)
d(HH22 _{Arg109} -OD1)	4.2 (0.6)	4.8 (1.0)	4.8 (1.3)	6.9 (0.6)
d(HE _{Arg109} -OE1 _{Gln103})	5.1 (0.3)	3.7 (1.7)	5.4 (0.3)	6.2 (0.9)
d(HH _{Arg109} -OH _{Tyr239})	3.2 (0.4)	3.3 (0.8)	2.8 (0.8)	3.0 (0.4)
d(HH _{Arg109} -O _w)	6.5 (0.5)	2.5 (0.7)	4.7 (0.5)	2.7 (0.5)
ΔE_{stab}	9.7	1.0	6.7	8.2

Tabla 17. Distancias promediadas (en Å) entre la Arg109 y otros residuos y moléculas de agua. El efecto estabilizador de este residuo en la barrera de energía potencia (en kcal/mol) se presenta en la última fila de la tabla.

La interacción Arg109-Gln103, observada en la simulación correspondiente al complejo REAC2, requiere especial atención. Esta interacción aparece entre dos residuos del bucle flexible, cerrando el centro activo. La existencia de esta interacción lleva a cabo una importante disposición del bucle, el cual puede tener consecuencias importantes en las energías de reacción (Figura 27). La Tabla 17 también muestra el efecto en la barrera de energía, expresado como una diferencia entre las energías en fase gas calculadas a nivel AM1 con y sin la presencia de Arg109. Los cálculos se llevaron a cabo en la EsT y a sus correspondientes estructuras de reactivos localizadas a partir del IRC. En todos los casos la Arg109 juega un papel de estabilización de la EsT. Sin embargo hay diferencias importantes entre los 4 caminos de reacción. Mientras el primero consigue la energía de estabilización máxima (9.7 kcal/mol) en el segundo (el que corresponde a REAC2) la Arg109 juega un papel menos importante (sólo 1.0 kcal/mol). REAC3 y REAC4 tiene

valores intermedios (-6.7 y -8.2 kcal/mol respectivamente). Comparando este efecto de la Arg109 con el efecto total de la barrera de energía potencia (ver Tabla 15) podemos observar una correspondencia entre los dos conjuntos de valores. El efecto de máxima estabilización corresponde al mínimo de barrera de energía, mientras que el efecto mínimo de Arg109 corresponde a la barrera más alta. Así, la posición de la Arg109 es un elemento clave para determinar la constante de reacción. Este residuo pertenece al bucle flexible y parece que pueda tener diferentes posiciones, determinando así la existencias de diferentes valles de reacción.

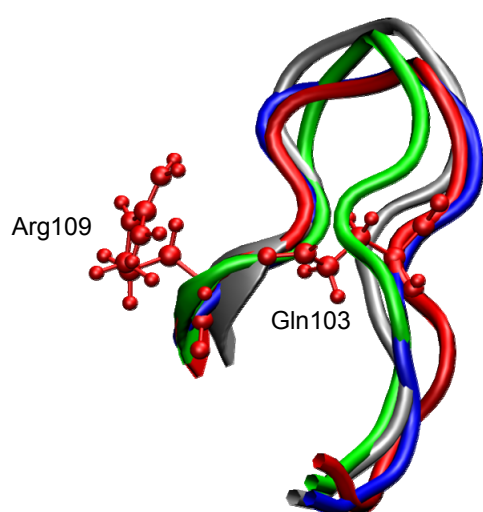


Figura 27. Superposición del bucle flexible a final de las simulaciones de los complejos reactivos. El código de color es el empleado en la figura anterior. Se muestra la situación en el bucle de la Arg109 y Gln103 en el REAC2.

2.9.3 Cálculos de energía libre

Como se discute en la introducción de esta sección, el hecho de que diferentes complejos reactivos puedan contribuir a la reacción puede tenerse en cuenta mediante el cálculo de perfiles de energía libre con

muestreos suficientemente largos. Considerando que la velocidad de conversión entre complejos reactivos diferentes puede ser incluso más lenta que la velocidad de reacción, este procedimiento resultaría completamente ineficaz excepto en casos muy simples. Otra posibilidad es identificar algunos de los complejos reactivos que contribuyen a la reacción obtener el perfil de energía libre correspondiente a cada caso mediante simulaciones cortas. Finalmente la obtención de la constante de velocidad microscópica requerirá la combinación de los diferentes perfiles dependiendo con los valores relativos entre la interconversión y las constantes de reacción.

Para obtener perfiles de energía libre correspondientes a la transformación de piruvato a lactato para cada uno de los valles reactivos analizados anteriormente, intentamos utilizar una técnica de *umbrella sampling*. Esta técnica permite, en principio, un muestreo correcto de todos los grados de libertad del sistema excepto, obviamente, la coordenada de reacción. Así, tratamos de obtener los perfiles de energía libre con diferentes definiciones de la coordenada de reacción del tipo:

$$RC = \alpha_1 r_{CnicH^+} - \alpha_2 r_{CpyrH^+} + \alpha_3 r_{NhisH^+} - \alpha_4 r_{Opyr} \quad (2.7)$$

Sin embargo, los resultados no nos llevaron a lo que esperábamos. Utilizando coordenadas del tipo definido en la ecuación (2.7) fue posible controlar la evolución de la transferencia de hidruro (analizando las distancias correspondientes a la transferencia obteníamos una evolución suave de las mismas con la coordenada de reacción). Por otra parte, las distancias asociadas a la transferencia protónica mostraban una evolución brusca y los perfiles de energía libre obtenidos presentaban discontinuidades. El cambio de los pesos α_x , que aparecen en la ecuación (2.7) no permitió superar este problema. Es

entonces cuando decidimos hacer uso de la metodología de Perturbación de Energía Libre (FEP).

En este procedimiento la estructura de reactivo puede expresarse aproximadamente como una función de la coordenada \mathbf{s} (explicada en el capítulo 3):

$$\begin{aligned} \Delta G_{\text{FEP}}(\mathbf{s}^R \rightarrow \mathbf{s}^j) &\approx \Delta E_{\text{QM}}^{0,R \rightarrow j} + \Delta G_{\text{QM/MM}}^{R \rightarrow j} = \\ &= \left(E_{\text{QM}}^0(\mathbf{s}^j) - E_{\text{QM}}^0(\mathbf{s}^R) \right) - k_B T \sum_{i=R}^{i=j-1} \ln \left\langle \exp \left(E_{\text{QM/MM}}(\mathbf{s}^{i+1}) - E_{\text{QM/MM}}(\mathbf{s}^i) \right) \right\rangle_{\text{MM},i} \end{aligned} \quad (2.8)$$

donde E_{QM}^0 es la energía QM en fase gas del subsistema QM, incluyendo la corrección bidimensional IC2D dependiente de las coordenadas R_1 y R_2 . La contribución de la interacción QM/MM a la energía libre entre dos configuraciones diferentes del sistema se obtiene promediando la energía de interacción QM/MM (incluyendo la energía de polarización) sobre todas las coordenadas MM del sistema para un valor particular de la coordenada \mathbf{s} . Por esta razón es necesario añadir un término para tener en cuenta la contribución de las fluctuaciones de todas las coordenadas del subsistema QM (excepto \mathbf{s}) al perfil de energía libre. Esta contribución se estima suponiendo un comportamiento armónico de los modos normales de vibración del subsistema QM en el campo creado por el resto de la enzima. Entonces, el perfil de energía libre cuasiclásico puede obtenerse añadiendo la contribución de estos grados de libertad vibracionales QM como:

$$\begin{aligned} G(\mathbf{s}^j) - G(\mathbf{s}^R) &= \Delta E_{\text{QM}}^{0,R \rightarrow j} + \Delta G_{\text{QM/MM}}^{R \rightarrow j} + \Delta G_{\text{fluc,QM}} = \\ &= \Delta E_{\text{QM}}^{0,R \rightarrow j} + \Delta G_{\text{QM/MM}}^{R \rightarrow j} + \sum_{x=1}^{3N-7} \frac{1}{2} h\nu_x^j + k_B T \ln \left(1 - e^{-\frac{h\nu_x^j}{k_B T}} \right) - G_{\text{fluc,QM}}(\mathbf{s}^R) \end{aligned} \quad (2.9)$$

Las seis frecuencias más bajas normalmente corresponden a pseudotraslaciones y pseudorotaciones del subsistema QM, las cuales no cabe esperar que contribuyan significativamente a las diferencias de energía libre y por lo tanto no las consideramos. Además, las frecuencias se han obtenido proyectando la contribución de la coordenada de reacción en la Hessiana para todas las estructuras excepto para reactivos y productos. Así, el término vibracional se obtiene sumando para de $3N-6$ modos de vibración (siendo N el número de átomos QM) en las estructuras de reactivos y productos y $3N-7$ ^[79] para el resto.** El perfil de energía libre así obtenido corresponde a un Potencial de Fuerza Media (PMF) porque, con las aproximaciones descritas, puede ser analizado en término de la probabilidad de encontrar el sistema en un determinado valor de la coordenada s . Excluyendo efectos túnel, la barrera de energía libre (cuasiclásica) se calcularía como:

$$\Delta G^\ddagger = G(s^\ddagger) - G(s^R) \quad (2.10)$$

donde s^\ddagger indica el máximo del perfil de energía libre de acuerdo con la ecuación (2.9). Para obtener el perfil de energía libre llevamos a cabo una simulación de DM para las estructuras a lo largo del camino de reacción. Cada ventana de simulación consiste en 5 ps de equilibrado y 15 ps de producción. Utilizamos un máximo de 50 ventanas por cada camino de reacción y empleamos una metodología de doble muestreo. Los perfiles de energía libre obtenidos para los 4 caminos analizados se presentan en la Figura 28 mientras que las barreras de activación se muestran en la Tabla 18. Las energías de activación obtenidas en esta aproximación cuasiclásica no se correlaciona exactamente con las energías potenciales.

** La contribución vibracional de los átomos QM al perfil de energía libre fue calculada para todas las estructuras de cada camino de reacción y después ajustada a una función polinómica en función de la coordenada s .

Además, de los cambios cualitativos hay un cambio cuantitativo importante ya que en los dos últimos caminos analizados hay más de 15 kcal/mol de diferencia entre la barrera de energía potencial y la de energía libre. Este hecho claramente resalta la importancia de considerar las fluctuaciones del entorno para evaluar la energética de la reacción. El término relativo a las fluctuaciones del subsistema QM no introduce cambios en el orden relativo a lo largo de los caminos de reacción pero, sin embargo, la importancia cuantitativa está fuera de discusión. Este término disminuye la barrera en más de 3 kcal/mol.

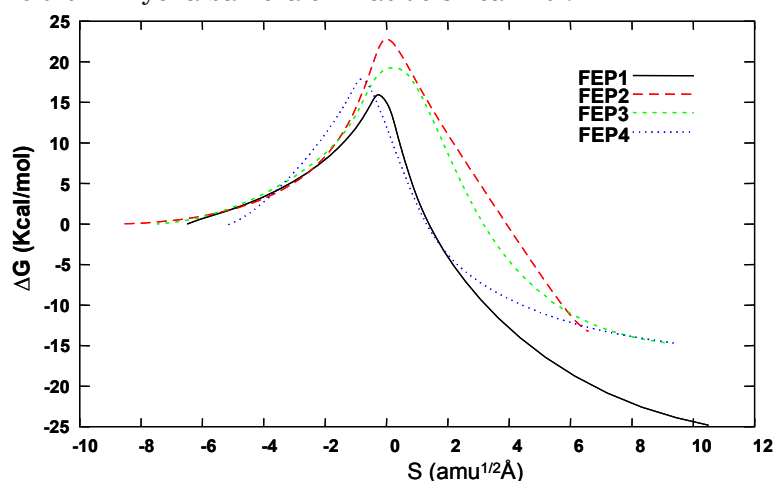


Figura 28. Representación de la energía libre frente a la coordenada intrínseca de reacción de los diferentes FEPs calculados a partir de las EsT seleccionadas.

	FEP1	FEP2	FEP3	FEP4
ΔG_{FEP}	16.7	23.5	19.4	20.6
$\Delta G_{\text{fluc,QM}}$	-3.1	-2.5	-2.6	-3.2
ΔG^\ddagger	13.6	21.0	16.8	17.4
κ	4.116	4.051	3.879	3.762
ΔG_{tun}	-0.84	-0.83	-0.80	-0.78
$\Delta G_{\text{eff}}^\ddagger$	12.8	20.2	16.0	16.6
k_{cat}	$2.77 \cdot 10^3$	$9.50 \cdot 10^{-3}$	$1.01 \cdot 10^1$	$4.02 \cdot 10^0$

Tabla 18. Contribuciones a la energía libre de activación (kcal/mol), coeficiente de transmisión y constantes catalíticas (en s^{-1}) obtenidas para cada uno de los caminos analizados.

2.9.4 Efecto Túnel

La reacción catalizada por la LDH tiene lugar por transferencia de átomos ligeros y, por tanto, tenemos que estudiar la posible contribución del efecto túnel en la constante de velocidad según la VTST. Para calcular esta contribución del efecto túnel utilizaremos el programa POLYRATE,^[80] combinado con el programa CHARMM.^[43] La metodología de POLYRATE consiste en congelar todos los átomos que definimos como zona secundaria (entorno), y realizar los cálculos de efecto túnel sólo para los átomos que definimos como zona primaria, en nuestro caso como zona primaria tendremos 37 átomos que incluyen el piruvato en su totalidad más una parte pequeña del cofactor y del residuo His195.

Seleccionamos la EsT del camino al que se le va a hacer el cálculo de túnel, congelando la zona secundaria, se traza un pequeño intervalo del IRC entorno a dicha EsT (en nuestro caso desde $s=-1.5$ y $s=1.3$, siendo $s=0$ la EsT) A continuación se realizan trayectorias que nos permiten calcular la probabilidad de túnel y de ahí calcular el coeficiente de transmisión de túnel (κ) utilizando las diferentes aproximaciones introducidas en el capítulo 3 (SCT, LCT y OMT).

En la Tabla 18 tenemos los resultados del valor del coeficiente de transmisión (κ), correspondiente al máximo valor de probabilidad de efecto túnel haciendo uso de la metodología de OMT. En nuestros cálculos el valor de κ obtenido con la metodología OMT coincide con el valor dado con la metodología SCT excepto para el FEP 2.

El valor de κ puede transformarse en energía libre mediante la ecuación (5.31) dada en el Capítulo 3. De este modo, podemos sumar la corrección de túnel a la energía libre total que habíamos calculado y obtener así una energía libre efectiva que sería comparable con los datos experimentales (ver Tabla 1). El valor promedio de la barrera de energía

libre aparente es 16.4 kcal/mol, número cercano al dado experimentalmente que es de 15 kcal/mol. Sin embargo, la forma correcta de promediar los resultados obtenidos dependerá de los valores relativos de las velocidades de reacción y de interconversión entre los distintos complejos considerados. Este punto deberá ser objeto de ulteriores estudios.

2.9.5 KIEs

Para realizar cálculos de efectos isotópicos cinéticos primarios (KIEs) tenemos que obtener la constante de velocidad para los dos isótopos del átomo que se transfiere, es decir, hidrógeno (H) y deuterio (D). Esto implica que se ha de realizar de nuevo el cálculo de la corrección vibracional cuántica con el átomo sustituido y del coeficiente de transmisión de túnel (κ). El valor de κ para los diferentes isótopos la presentamos en la Tabla 19 donde DHH significa que el hidruro de la reacción es sustituido por un deuterio y HHD significa que es el protón la especie sustituida por deuterio. Los resultados de los KIEs se presentan en la Tabla 19, donde inicialmente se presenta el efecto del KIE vibracional (k_H/k_D), seguidamente el efecto aislado del túnel (κ_H/κ_D) y por último el efecto global incluyendo ambos efectos.

DHH	FEP 1	FEP 2	FEP 3	FEP 4
k_H/k_D	2.75	3.70	3.54	2.66
κ_H/κ_D	1.52	1.37	1.40	1.31
$k^{\text{OMT}_H}/k^{\text{OMT}_D}$	4.18	5.06	4.98	4.47

HHD	FEP 1	FEP 2	FEP 3	FEP 4
k_H/k_D	1.87	1.22	1.13	1.50
κ_H/κ_D	1.44	1.27	1.63	1.58
$k^{\text{OMT}_H}/k^{\text{OMT}_D}$	2.69	1.56	1.84	2.36

Tabla 19. Valor de los KIEs calculados para el hidruro y el protón independientemente. El cálculo de las constantes incluye la CTV.

Se observa un mayor KIE en el caso de la transferencia de hidruro como indican los datos experimentales. En concreto, se ha estimado un valor de KIE 3-6 para la reacción en la enzima original. La incertidumbre de este valor es debida a que la etapa limitante de la reacción no es la reacción química,^[16] aunque Schowen y colaboradores han sido capaces de obtener un KIE intrínseco de 5.5 a 55°C.^[13] En el caso del protón obtenemos valores de KIEs más bajos. En una comunicación personal, el Profesor Clarke nos proporcionó el dato de un posible KIE del protón entorno al valor de 1.5.

2.10 Mutaciones

Como se ha explicado en el Capítulo 3 las mutaciones que se le hacen a una enzima y su posterior estudio nos permite conocer mejor la catálisis enzimática. Al principio de este capítulo hemos presentado estudios de mutaciones que se han realizado en la BsLDH. El motivo de esta sección es realizar el estudio teórico de alguna de las mutaciones estudiadas experimentalmente, e incluso algunas mutaciones que hemos propuesto nosotros, para comprender la importancia de ciertos aminoácidos en el centro activo de la enzima.

Para realizar el estudio de las mutaciones cambiamos el modelo de la región mecanocuántica debido a que una de las mutaciones implica a la Arg109 que tenemos en ésta región. Si definimos la región QM con 52 átomos (en los que no incluíamos las argininas del centro activo) podemos hacer una comparación cualitativa entre los estudios de las diferentes mutaciones. Por otra parte, para el estudio cualitativo de las mutaciones prescindimos de hacer las correcciones al Hamiltoniano

semiempírico. En principio, el estudio de las SEP AM1/MM nos permitirá comparar energética y estructuralmente la enzima original (wild type) con las distintas enzimas mutadas, analizando el papel de los aminoácidos del centro activo.

Hacemos la exploración de una nueva SEP de la enzima original donde incluimos 52 átomos QM y mantenemos el estado de protonación que hemos calculado. Esta nueva SEP la podemos comparar con los resultados de las mutaciones que, como hemos dicho, tendrán una región QM sin incluir las argininas. La SEP de la enzima original con la región QM pequeña la presentamos en la Figura 29. Cabe destacar que esta nueva superficie puede compararse directamente con la representada en la Figura 12, de modo que la diferencia entre ellas es únicamente el cambio en el estado de protonación de los aminoácidos, tal como se ha explicado en el apartado 2.8. La diferencia más destacada entre la SEP de la Figura 12 y la de la Figura 29 es el descenso en las barreras energéticas. En el camino que implica la transferencia de hidruro anterior a la de protón pasa de ser de 52 kcal/mol a 44.7 kcal/mol. En el camino que implica primero la transferencia del protón seguida por la de hidruro tenemos un descenso de 41 a 34 kcal/mol.

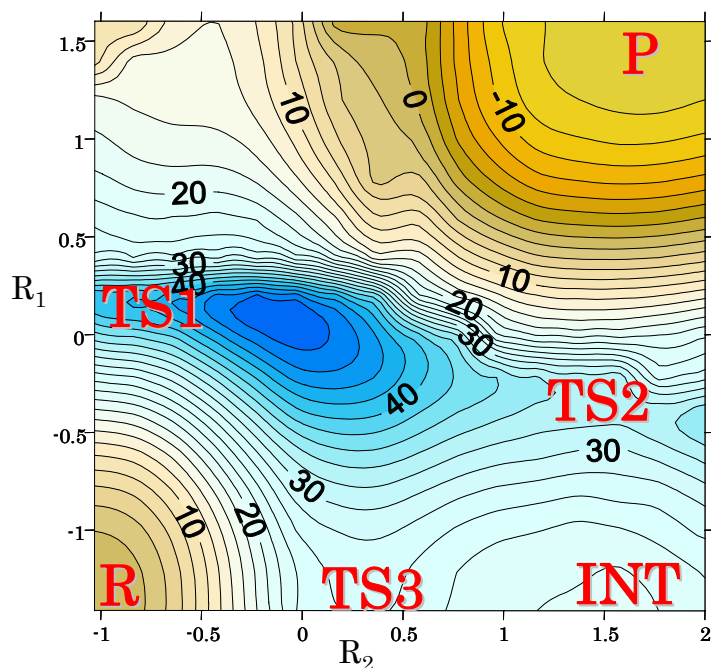


Figura 29. SEP AM1/MM correspondiente a la enzima nativa obtenida con la región QM de 52 átomos utilizando el estado de protonación calculado en el seno de la enzima original. R_1 y R_2 son las coordenadas de reacción (en Å) asociadas a la transferencia de hidruro y protón respectivamente. Las posiciones de los puntos singulares optimizados se indican sobre la SEP.

	Piruvato	TS1	TS2	TS3	Int	Lactato
Energía Relativa	0.0	44.7	33.8	26.2	23.0	-14.2
C _{NADH} ⋯H _t	1.13	1.64	1.28	1.13	1.13	2.67
H _t ⋯C _{pyr}	2.49	1.24	1.51	2.50	2.48	1.13
N _{His195} ⋯H _p	1.01	1.03	2.31	1.49	1.94	2.48
H _p ⋯O _{pyr}	2.05	1.84	0.99	1.01	1.02	0.97
O _{pyr} -HHres109	2.94	2.87	2.58	3.00	2.63	3.07
O _{pyr} -HDres140	2.05	1.91	2.93	2.89	2.91	2.83
O _{pyr} -Hwat	3.63	3.18	3.03	3.42	3.42	3.20
OD1Asp168-HD1His195	3.03	2.73	3.26	3.16	3.16	2.78
OE2Glu199-OD2Asp168	1.55	1.55	1.57	1.55	1.56	1.59
O _{pyr} -HHres109	2.60	2.59	2.58	2.59	2.59	2.59

Tabla 20. Distancias (en Å) más relevantes y energías relativas a reactivos (en kcal/mol) de los puntos estacionarios localizados en el sistema en la SEP AM1/MM con 52 átomos QM y el estado de protonación de los aminoácidos recalculados en el seno de la enzima.

En la SEP AM1/MM presentada en la Figura 29 se observan dos caminos de reacción donde se puede obtener unas barreras de 33 y 34 kcal/mol. Al optimizar los puntos de silla y calcular sus IRC, se observa que el IRC del TS1 se dirige directamente a productos sin pasar por el cuadrante superior izquierdo de la superficie. En la Tabla 20 se presenta las distancias y energías relativas para todas estructuras las optimizadas.

2.10.1 Arg109Gln

La mutación de la arginina 109 es la primera que vamos a estudiar. En los estudios cinéticos experimentales se observa que el cambio de un aminoácido con carga positiva, como es la arginina 109, por un aminoácido neutro, como es la glutamina, se traduce en una disminución en la constante de velocidad catalítica, que se puede atribuir a un aumento en la barrera de la reacción química.^[16] Como resultado del aumento en la barrera de la reacción química, ésta pasa a ser la etapa limitante de la reacción según los estudios experimentales,^[16] de manera que el cierre del giro ya no es la etapa limitante.

Para realizar el estudio de la primera mutación partimos de la estructura de RX inicial (el tetrámero) y cambiamos la arginina 109 por una glutamina. El siguiente paso consiste calcular el pK_a de los aminoácidos del sistema. El cambio del aminoácido repercute en un cambio en el valor del pK_a de los aminoácidos que están próximos al residuo mutado, tal y como se muestra en la Tabla 21. En concreto se observa un cambio de más de tres unidades en la His195 encargada de donar el protón en la reacción química mientras el Glu199 apenas aumenta seis décimas su valor.

	His195	Asp168	Glu199	His95
Enzima nativa	8.7	1.2	7.4	7.6
Arg109Gln	11.9	1.3	8.0	7.6

Tabla 21. pK_a de algunos aminoácidos s calculados para la enzima nativa y para la mutación Arg109Gln.

Una vez conocido el estado de ionización de los aminoácidos pasamos a saturar con hidrógenos todo el sistema. En este caso todos los aminoácidos tienen el sitio ionizable con las mismas condiciones que en la enzima original, es decir, queda protonado el Glu199 y la His95. El protocolo de solvatar el sistema y realizar dinámicas de relajación para iniciar la exploración de la SEP es idéntico al llevado a cabo en la obtención de las SEP anteriores.

La Figura 30 representa la SEP AM1/MM obtenida tras la mutación Arg109Gln. En la SEP podemos observar dos caminos de reacción concertados aunque asincrónicos. La barrera de menor energía la presenta el camino con una EsT que implica una transferencia del protón más adelantada que el hidruro. En la Tabla 22 se presentan las distancias de los puntos estacionarios calculados.

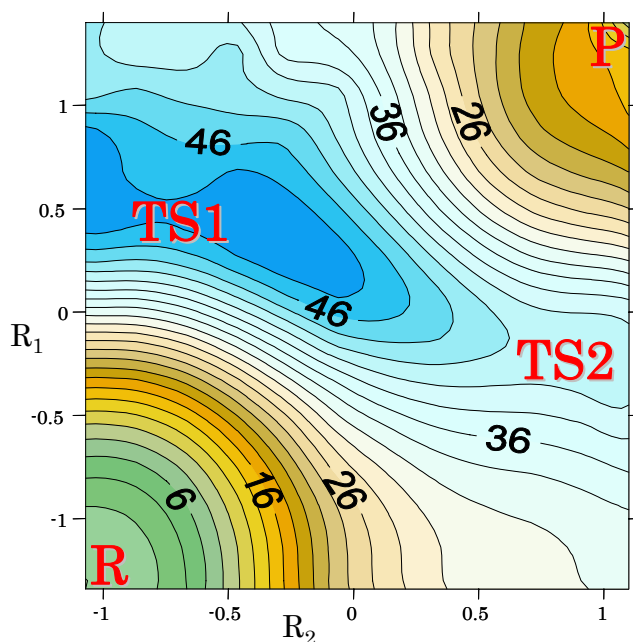


Figura 30. SEP AM1/MM con la región QM de 52 átomos de la mutación Arg109Gln. La posición de las estructuras optimizadas de reactivos (R), productos (P), y las EsT (TS1 y TS2) se muestran cualitativamente

	Piruvato	TS1	TS2	Lactato
Energía Relativa	0.0	48.1	42.7	17.5
C _{NADH} ⋯ H _t	1.12	1.75	1.29	2.47
H _t ⋯ C _{pyr}	2.45	1.21	1.45	1.13
N _{His195} ⋯ H _p	1.01	1.04	1.80	2.10
H _p ⋯ O _{pyr}	2.08	1.77	1.01	0.97
O _{pyr} -HHGln109	-	-	-	-
O _{pyr} -HDAsn140	2.20	2.01	2.15	2.41
O _{pyr} -Hwat	2.14	1.97	3.03	3.65
OD1Asp168-HD1His195	2.82	2.70	2.72	2.98
OE2Glu199-OD2Asp168	1.56	1.57	1.58	1.60
O _{pyr} -HHres109	2.58	2.58	2.58	2.58

Tabla 22. Distancias (en Å) más relevantes y energías relativas a reactivos (en kcal/mol) de los puntos estacionarios localizados la SEP AM1/MM con 52 átomos QM y con la protonación de los aminoácidos recalculados en el seno de la enzima en la mutación Arg109Gln.

La energía del cuadrante inferior derecho de la SEP aumenta en energía respecto a la SEP de la enzima nativa. Esto es consecuencia del

cambio de pK_a de la His195. El pK_a de la His195 aumenta de 8.7 a 11.9 (Tabla 21), este cambio de pK_a refleja el hecho de que la transferencia de protón desde la His195 al piruvato no es tan favorable como en la enzima original. La barrera para el camino de reacción que implica el paso por el TS2 está 9 kcal/mol sobre el valor para la enzima original. Sin embargo, es todavía menor que la barrera (48.1 kcal/mol) que presenta el camino que pasa por el TS1. El experimento de mutagénesis realizado por Clarke y colaboradores^[16] obtuvo un aumento en la barrera de unas 4.2 kcal/mol. Teniendo en cuenta las limitaciones del método AM1 para obtener cuantitativamente valores de energía precisos y el hecho de que estamos analizando un único camino de reacción, podemos afirmar que nuestros resultados están en concordancia con los experimentales ya que en nuestro caso también obtenemos un aumento en las barreras de los caminos de reacción.

Las distancias dadas en la Tabla 22 sugieren que en la mutación hay un aminoácido que sustituye las funciones del residuo mutado. El residuo Asn140 está a tan solo 2.20 Å del oxígeno carbonílico del piruvato en el complejo reactivo, mientras que en la enzima original se encontraba a 3.63 Å. Esta distancia se acorta cuando pasamos por las EsT. También encontramos una molécula de agua, que antes no se encontraba dentro del centro activo, que está formando enlace de hidrógeno con el oxígeno carbonílico del piruvato, especialmente en el TS1. Consecuentemente, la enzima mutada muestra un comportamiento similar a la enzima original, desde el punto de vista de la reacción química. Claro está, que otro efecto postulado por Clarke y colaboradores^[18] como es el cambio de las propiedades de la enzima afectando a la unión del sustrato con ésta y el cambio del comportamiento del bucle que cierra el centro activo, no puede ser

observado en nuestra SEP ya que nuestro estudio solo considera la etapa de la reacción química. Aún así, nuestros resultados apoyan en parte la evidencia experimental de que la enzima sigue conservando su actividad después de la mutación Arg109Gln. No obstante, existe una pérdida de efectividad, debido al hecho de que el residuo (Asn140), que se encarga de sustituir a la Arg109 en la estabilización de la EsT, es un residuo neutro.

El análisis de la distribución electrónica se llevó a cabo mediante análisis de cargas de Mulliken de los diferentes átomos contenidos en la región QM. Las cargas atómicas calculadas en el TS2 para el C y el O carbonílico del piruvato fueron +0.24 y -0.38, en la enzima original respectivamente, mientras que fueron +0.23 y -0.39 en la mutación Arg109Gln. Este resultado no está de acuerdo con los resultados de Clarke y colaboradores, que sugerían que una pérdida de carga positiva en el residuo 109 no permitiría la polarización del enlace del grupo carbonilo en los reactivos.^[16] La separación de cargas entre los átomos nombrados en el TS2 para ambos sistemas es prácticamente la misma. Quizás el efecto de polarización, si existe, no puede ser revelado por el análisis de Mulliken realizado.

2.10.2 Asp168Ala

La segunda mutación que estudiamos es el cambio del aspartato 168 por una alanina. El aspartato está próximo a la histidina 195 que se encarga de ceder el protón al piruvato y establece un puente de hidrógeno con la misma. La alanina es un residuo neutro incapaz de formar puentes de hidrógeno. Experimentalmente se encuentra que la mutación impide la formación del complejo reactivo por lo que se detendría la actividad catalítica de la enzima.^[21] En nuestro estudio teórico, tras realizar la

mutación, calculamos el pK_a de los aminoácidos llegando en esta ocasión a tener resultados significativamente diferentes respecto a la enzima nativa. El Glu199 cambia su estado de protonación pasando a estar cargado negativamente a $pH=7$, tal como se deduce de los valores de pK_a dados en la Tabla 23.

	His195	Asp168	Glu199	His95
Enzima nativa	8.7	1.2	7.4	7.6
Asp168Ala	7.3	-	4.2	7.6

Tabla 23. Valores de los pK_a s para algunos aminoácidos en la enzima nativa y en la mutación Asp168Ala.

La Figura 31 muestra la SEP obtenida para el caso de la mutación Asp168Ala. Como puede observarse el efecto de esta mutación en la SEP es, fundamentalmente, la desestabilización del cuadrante inferior derecho, es decir, el intermedio que resulta al transferir el protón. Al analizar las estructuras resultantes se observa la existencia de una molécula de agua entre la His195 y el Glu199 que los une por puente de hidrógeno sustituyendo la función del residuo mutado Asp168 (Figura 32). Esta molécula juega un papel importante moderando el descenso del pK_a de la His195 y el Glu199. Consecuentemente la SEP de la reacción química no cambia tan radicalmente respecto a la enzima nativa como se podía esperar.

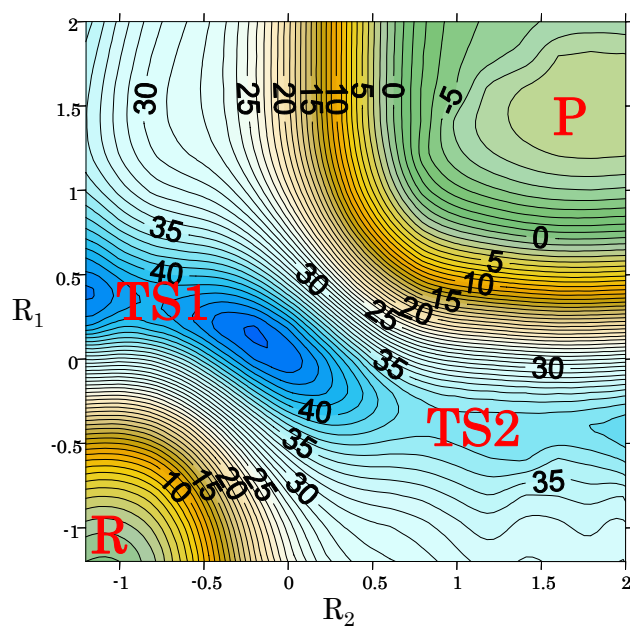


Figura 31. SEP AM1/MM de la mutación Asp168Ala. La posición de las estructuras de reactivos (R), productos (P), y las EsT (TS1 y TS2) se muestran cualitativamente

	Piruvato	TS1	TS2	Lactato
Energía Relativa	0.0	42.8	38.0	-5.4
C _{NADH} ···H _t	1.13	1.57	1.23	2.40
H _t ···C _{pyr}	2.55	1.28	1.61	1.13
N _{His195} ···H _p	1.01	1.04	1.98	2.65
H _p ···O _{pyr}	2.11	1.86	1.01	0.97
O _{pyr} -NE2His195	2.96	2.84	2.96	3.53
O _{pyr} -HHres109	2.28	2.04	2.54	2.13
O _{pyr} -HDres140	3.86	3.79	3.72	3.95
O _{pyr} -Hwat	3.31	2.63	3.50	3.43
OD1Asp168-HD1His195	-	-	-	-
His195-Hwat	1.65	1.68	1.69	1.66
Glu199-res168	1.62	1.63	1.63	1.63

Tabla 24. Distancias (en Å) más relevantes y energías relativas a reactivos (en kcal/mol) de los puntos estacionarios localizados la SEP AM1/MM con 52 átomos QM y con la protonación de los aminoácidos recalculados en el seno de la enzima en la mutación Asp168Ala.

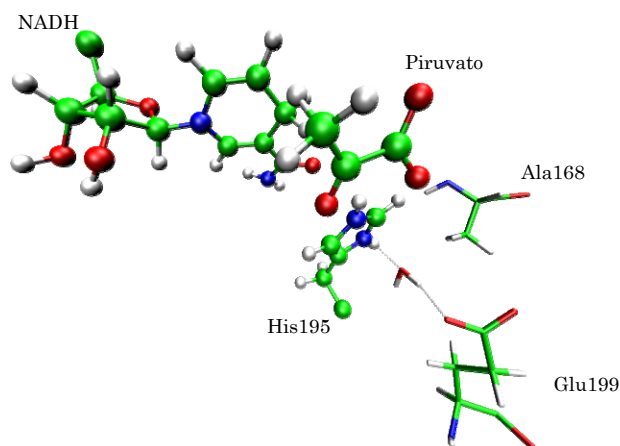


Figura 32. Detalle de los reactivos localizados para la mutación Asp168Ala. Se observa el papel de la molécula de agua tal como se explica en el texto.

2.10.3 Doble Mutación Arg109Gln y Asn140Asp

Con el objetivo de comprobar nuestra predicción sobre el papel que realiza la Asn140 polarizando y estabilizando las EsT como una alternativa a la Arg109, hemos realizado una doble mutación donde ambos residuos son sustituidos por otros con diferente comportamiento. La Arg109 es sustituida de nuevo por la Gln, mientras la Asn140 la reemplazamos por un aminoácido con carácter ácido como es el aspartato. Como se muestra en la Tabla 25, el cálculo de pK_a para los residuos ionizables revela que el nuevo Asp140 queda protonado a pH fisiológico. El otro efecto importante es el valor del pK_a de la His195, ahora mucho mayor que en la enzima original. El resultado de la exploración de la SEP se muestra en la Figura 33, la cual muestra dos caminos de reacción competitivos para la reducción del piruvato. Por comparación con la SEP de la enzima original (Figura 29 y Tabla 20), el camino que pasa por el TS1 (que ahora es un mecanismo por etapas) es todavía viable ya que tiene una barrera de unas 46.0 kcal/mol. El intermedio que aparece en la esquina del cuadrante superior izquierdo corresponde a un aducto en el

cual el O carbonílico del piruvato se enlaza con el C4 del cofactor (Figura 34).

	His195	Asp168	Glu199	Asp
Enzima nativa	8.7	1.2	7.4	-
DobleMutación	14.0	1.3	8.3	8.2

Tabla 25. Valores de los pK_{as} para algunos aminoácidos en la enzima nativa y en la doble mutación Arg109Gln-Asn140Asp.

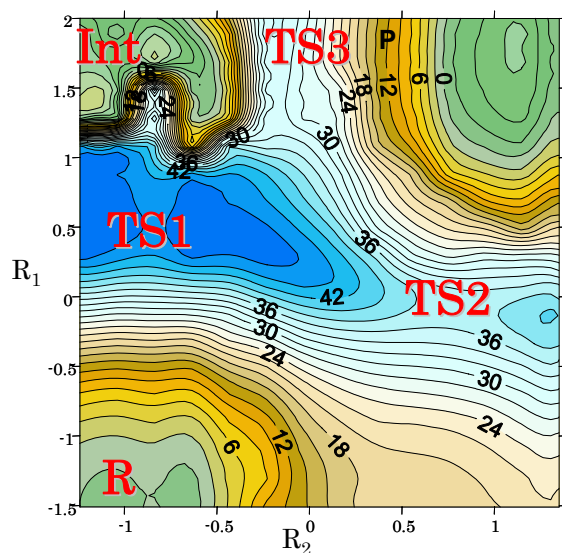


Figura 33. SEP obtenida AM1/MM de la doble mutación Arg109Gln y Asn140Asp. La posición de las estructuras de reactivos (R), productos (P), intermedio (Int) y las EsT (TS1, TS2) se muestran cualitativamente

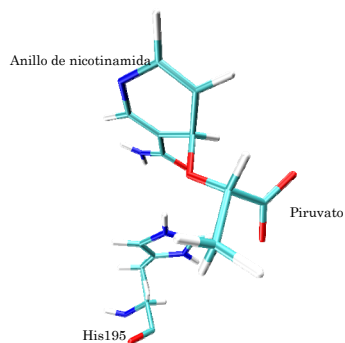


Figura 34. Detalle del aducto formado en el centro activo de la enzima doblemente mutada. Se puede observar el nuevo enlace entre el piruvato y el cofactor.

	Piruvato	TS1	TS2	TS3	Int	Lactato
Energía Relativa	0.0	46.0	37.9	24.0	0.0	-6.5
C _{NADH} ····H _t	1.13	1.67	1.33	2.63	2.64	2.82
H _t ····C _{pyr}	2.64	1.22	1.41	1.12	1.12	1.13
N _{His195} ····H _p	1.00	1.03	1.78	1.08	1.01	2.15
H _p ····O _{pyr}	2.04	1.78	1.00	1.57	2.06	0.98
O _{pyr} -NE2His195	2.86	2.73	2.76	2.64	2.91	2.87
O _{pyr} -HHres109	-	-	-	-	-	-
O _{pyr} -HDres140	3.40	1.89	2.53	2.84	3.11	3.17
O _{pyr} -Hwat	2.01	1.79	3.28	4.68	3.51	3.61
OD1Asp168-HD1His195	1.58	1.59	1.60	1.59	1.60	1.60
His195-Hwat	-	-	-	-	-	-
Glu199-res168	1.60	1.60	1.59	1.59	1.61	1.60

Tabla 26. Distancias (en Å) más relevantes y energías relativas a reactivos (en kcal/mol) de los puntos estacionarios localizados la SEP AM1/MM con 52 átomos QM y con la protonación de los aminoácidos recalculados en el seno de la enzima en la doble mutación Arg109Gln/Asn140Asp.

El aducto progresa hacia productos a través de una EsT, TS3, correspondiente a la ruptura del enlace formado (O C4) y la transferencia de protón en un mismo paso, con una barrera menor que la definida por el TS1. En dos experimentos realizados por Everse y colaboradores^[81] y por Burgner y Ray^[82] sobre inhibición, de la enzima LDH del ‘pez-perro’ (*dogfish*) se obtuvo un inesperado complejo que incluía un enlace covalente entre el cofactor y el sustrato. Sin embargo, el aducto localizado en nuestro trabajo no es exactamente el mismo que el obtenido por estos grupos, ya que ellos proponían que el enlace covalente entre el cofactor y el piruvato era entre el carbono del piruvato y el carbono del anillo de nicotinamida.

Respecto al camino de reacción en la Figura 33 que pasa por el TS2, el intermedio localizado en el cuadrante inferior derecho de la Figura 29 es desestabilizado, a pesar de que el mecanismo global presente una barrera similar (37.9 kcal/mol). Esta desestabilización del mencionado

cuadrante de la SEP está relacionado con el pK_a de la His195, la cual ya no tiene la misma predisposición para transferir el protón al piruvato. Sin embargo lo que es sorprendente es el hecho de que el TS1 no ha sido tan desestabilizado como se esperaba. El análisis de la geometría del TS1 revela la aparición de una molécula de agua que parece tener interacciones de hidrogeno fuertes con el oxígeno carbonílico del piruvato, con una distancia de enlace de hidrógeno de 1.8 Å. Esta molécula de agua se ha movido más de 1 Å desde su posición inicial en la enzima original que ahora realiza la función de la Arg109 jugando un papel importante en la estabilización del estado de transición.

2.10.4 Doble Mutación Arg109Gln y Asn140Glu

La última simulación de mutaciones en la enzima es muy similar a la anterior sólo que el residuo Asn140 lo sustituimos por Glu140. El glutamato tiene similar comportamiento al aspartato pero su cadena es más larga. De este modo predecimos que no habrá espacio para una molécula de agua en el centro activo y no podrá sustituir a la Arg109 en su función estabilizante.

A priori el efecto más importante en el cálculo de pK_a , Tabla 27, es el aumento a 14.0 en la His195 y el hecho de que el residuo Glu140 quede protonado a $pH=7$. El efecto en la SEP (Figura 35) es notable; las barreras se disparan a valores muy superiores a los obtenidos en todas las SEP calculadas en el presente trabajo. Como en la SEP anterior encontramos en el cuadrante superior izquierdo la estructura de un compuesto que corresponde a un aducto sustrato-cofactor. Esta estructura es similar a la que se muestra en la Figura 34. Estos resultados nos llevan a concluir que la capacidad catalítica de la enzima se ha anulado, corroborando la

hipótesis propuesta del papel de la Arg109 así como la alternativa del Asp140 e incluso la molécula de agua.

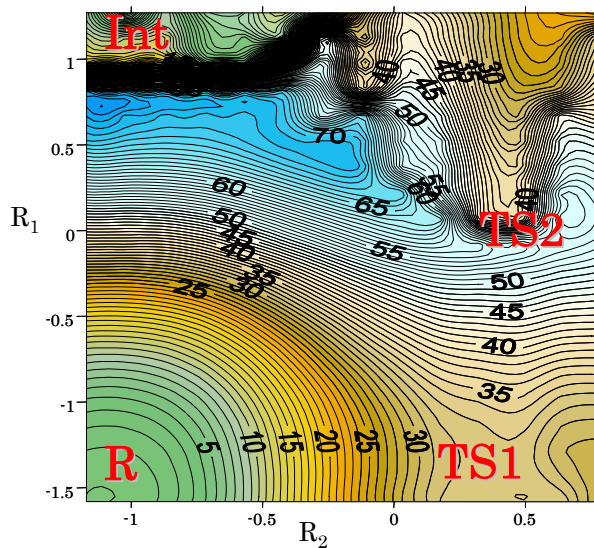


Figura 35. SEP AM1/MM de la doble mutación Arg109Asn y Asn140Glu. La posición de las estructuras de reactivos (R), productos (P), y las EsT (TS1 y TS2) se muestran cualitativamente.

	His195	Asp168	Glu199
Enzima nativa	8.7	1.2	7.4
DobleMutación II	14.0	1.4	8.4

Tabla 27. Algunos aminoácidos con los pK_a calculados para la enzima nativa y la doble mutación Arg109Gln-Asn140Glu.

3 Conclusiones y Perspectivas de Futuro

Dentro de este capítulo se ha realizado un detallado estudio de la reacción catalizada por la enzima LDH. El estudio comienza con la exploración de las SEPs, aplicando correcciones al error que genera el tratamiento semiempírico de la región definida con mecánica cuántica. Después de estudiar el estado de protonación de los aminoácidos que forman la estructura de la enzima y de clarificar el mecanismo de reacción, se obtuvieron los perfiles de energía libre mediante el cálculo de perturbaciones de energía libre. A estos perfiles de energía libre se les incluye una corrección vibracional a la región QM, que no tiene en cuenta la dinámica molecular semiclásica, y además se incluye la corrección del efecto túnel calculando el coeficiente de transmisión (κ) con la metodología ES-VTST/OMT. Con los resultados de las barreras de energía libre de la reacción realizamos cálculos de efectos isotópicos cinéticos (KIEs). Para finalizar el estudio se exploraron las SEPs de diferentes mutaciones que han sido estudiadas experimentalmente y otras que proponemos para corroborar el papel que se atribuye a algunos residuos.

En el estudio de las SEP hemos concluido que el método de corrección más adecuado para nuestra reacción es el de *splines* bidimensionales (IC2D). Estas correcciones tienen en cuenta las dos coordenadas de reacción distinguidas que describen la reacción química. Por el contrario, los otros dos métodos de corrección (SVB e IC) hacen uso de las dos coordenadas independientemente lo que da lugar a energías ‘ficticias’ en algunos cuadrantes de la SEP. Además, hemos mostrado la

importancia de utilizar cálculos electrostáticos para determinar el estado de protonación de los aminoácidos.

Los perfiles de energía libre han sido obtenidos para distintos caminos trazados desde distintas estructuras de transición. Estas estructuras se localizan a partir de simulaciones de dinámica molecular llevadas a cabo en la región del estado de transición. El análisis de los distintos perfiles muestra barreras de energía libre que difieren en más de 8 kcal/mol. Estos resultados pueden interpretarse teniendo en cuenta que distintas conformaciones de la enzima pueden contribuir de manera diferente a la reactividad observada. Distintos estudios experimentales recientes^[68, 71] muestran una dispersión importante en las constantes catalíticas medidas para una sola molécula. El acoplamiento conformación-reacción puede ser particularmente importante en la LDH ya que presenta un bucle flexible en el mismo centro activo. Los cálculos de los KIEs confirman el mismo comportamiento, ya que entre ellos difieren hasta una unidad. La forma de promediar correctamente las contribuciones de los diferentes complejos reactivos deberá ser objeto de estudios posteriores.

Por último, el estudio de las mutaciones ha mostrado la importancia de los aminoácidos del centro activo. La exploración de las SEPs en las diferentes mutaciones muestra cómo el efecto de sustituir un aminoácido (que tiene una función concreta) por otro (con características muy diferentes al sustituido) comporta una menor eficacia en la enzima aunque, en gran parte de las ocasiones, el entorno se adapta para seguir catalizando la reacción química. Después de realizar los estudios de mutaciones sobre las SEPs, el trabajo nos conducirá a calcular perfiles de energía libre con las correcciones aplicadas en la enzima nativa y

extender el estudio a otras mutaciones o bien a otras LDH que existen en otros organismos vivos.

4 Bibliografía

- [1] A. R. Clarke, T. Atkinson, J. W. Campbell and J. J. Holbrook, *Biochim. Biophys. Acta* **1985**, *829*, 387-396.
- [2] A. R. Clarke, A. D. B. Waldman, I. Munro and J. J. Holbrook, *Biochim. Biophys. Acta* **1985**, *828*, 375-379.
- [3] A. R. Clarke, D. B. Wigley, D. A. Barstow, W. N. Chia, A. D. B. Waldman, K. W. Hart, T. Atkinson and J. J. Holbrook, *Biochem. Soc. T.* **1987**, *15*, 152-153.
- [4] S. J. Cordwell, D. J. Basseal, J. D. Pollack and I. HumpherySmith, *Gene* **1997**, *195*, 113-120.
- [5] F. Basaglia, *Comp. Biochem. Phys. B* **1989**, *92*, 395-398.
- [6] J. Dudka, K. Dawidek-Pietryka, F. Burdan and B. Madej, *Drug. Metab. Rev.* **2002**, *34*, 170-170.
- [7] J. van Beek, H. Deng, R. Callender and J. Burgner, *J Raman. Spectrosc.* **2002**, *33*, 397-403.
- [8] W. Eventoff, M. G. Rossmann, S. S. Taylor, H. J. Torff, H. Meyer, W. Keil and H. H. Kiltz, *P. Natl. Acad. Sci. USA* **1977**, *74*, 2677-2681.
- [9] J. W. Burgner and W. J. Ray, *Biochemistry* **1984**, *23*, 3636-3648.
- [10] J. Andres, V. Moliner, J. Krechl and E. Silla, *Bioorg. Chem.* **1993**, *21*, 260-274.
- [11] J. Wilkie and I. H. Williams, *J. Am. Chem. Soc.* **1992**, *114*, 5423-5425.
- [12] A. R. Clarke, A. D. B. Waldman, K. W. Hart and J. J. Holbrook, *Biochim. Biophys. Acta* **1985**, *829*, 397-407.
- [13] J. Seravalli, W. P. Huskey, K. B. Schowen and R. L. Schowen, *Pure Appl. Chem.* **1994**, *66*, 695-702.
- [14] A. Lodola, J. D. Shore, D. M. Parker and J. J. Holbrook, *Biochem. J.* **1978**, *175*, 987-998.

- [15] P. Assa, M. Ozkan and G. Ozcengiz, *Ann. Microbiol.* **2005**, *55*, 193-197.
- [16] A. R. Clarke, D. B. Wigley, W. N. Chia, D. Barstow, T. Atkinson and J. J. Holbrook, *Nature* **1986**, *324*, 699-702.
- [17] H. Deng, J. Zheng, A. Clarke, J. J. Holbrook, R. Callender and J. W. Burgner, *Biochemistry* **1994**, *33*, 2297-2305.
- [18] A. R. Clarke, H. M. Wilks, D. A. Barstow, T. Atkinson, W. N. Chia and J. J. Holbrook, *Biochemistry* **1988**, *27*, 1617-1622.
- [19] K. W. Hart, A. R. Clarke, D. B. Wigley, W. N. Chia, D. A. Barstow, T. Atkinson and J. J. Holbrook, *Biochem. Bioph. Res. Co.* **1987**, *146*, 346-353.
- [20] J. J. Holbrook and Gutfreun, H., *Febs Letters* **1973**, *31*, 157-169.
- [21] T. J. Nobbs, A. Cortes, J. L. Gelpi, J. J. Holbrook, T. Atkinson, M. D. Scawen and D. J. Nicholls, *Biochem. J.* **1994**, *300*, 491-499.
- [22] S. Fushinobu, K. Kamata, S. Iwata, H. Sakai, T. Ohta and H. Matsuzawa, *J. Biol. Chem.* **1996**, *271*, 25611-25616.
- [23] N. Holmberg, U. Ryde and L. Bulow, *Protein Eng.* **1999**, *12*, 851-856.
- [24] J. L. Gelpi, R. M. Jackson and J. J. Holbrook, *J. Chem. Soc. Faraday T.* **1993**, *89*, 2707-2716.
- [25] S. Ranganathan and J. E. Gready, *J. Chem. Soc. Faraday T.* **1994**, *90*, 2047-2056.
- [26] K. E. Norris and J. E. Gready, *Theochem-J. Mol. Struc.* **1993**, *98*, 99-125.
- [27] K. E. Norris and J. E. Gready, *Theochem.-J. Mol. Struc.* **1992**, *90*, 109-138.
- [28] S. Ranganathan and J. E. Gready, *J. Phys. Chem. B* **1997**, *101*, 5614-5618.
- [29] R. K. Schmidt and J. E. Gready, *Abstr. Pap. Am. Chem. S.* **1997**, *214*, 116-Comp.
- [30] R. K. Schmidt and J. E. Gready, *J. Mol. Struc.-Theochem.* **2000**, *498*, 101-112.

- [31] R. K. Schmidt and J. E. Gready, *J. Mol. Model.* **1999**, *5*, 153-168.
- [32] V. Moliner and I. H. Williams, *Chem. Commun.* **2000**, 1843-1844.
- [33] A. J. Turner, V. Moliner and I. H. Williams, *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323-1331.
- [34] V. Moliner, A. J. Turner and I. H. Williams, *Chem. Commun.* **1997**, 1271-1272.
- [35] J. Andres, V. Moliner and V. S. Safont, *J. Phys. Org. Chem.* **1996**, *9*, 498-506.
- [36] A. Yadav, R. M. Jackson, J. J. Holbrook and A. Warshel, *J. Am. Chem. Soc.* **1991**, *113*, 4800-4805.
- [37] M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem* **1990**, *11*, 700-733.
- [38] U. C. Singh and P. A. Kollman, *J. Comput. Chem.* **1986**, *7*, 718-730.
- [39] W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [40] K. Fukui, *Acc. Chem. Res.* **1981**, *14*, 363-368.
- [41] J. Gao, P. Amara, C. Alhambra and M. J. Field, *J. Phys. Chem. A* **1998**, *102*, 4714-4721.
- [42] M. Freindorf and J. Gao, *J. Comp. Chem* **1996**, *17*, 386-395.
- [43] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187-217.
- [44] L. David, *Modélisation des interactions électrostatiques des biomolécules en solution*, Tesis Doctoral, Université Joseph Fourier, Grenoble, **1996**
- [45] A. Fersht, *Enzyme structure and mechanism*, W.H. Freeman, New York, **1985**.
- [46] A. R. Leach, *Molecular Modelling. Principles and Applications*, London, **1996**.
- [47] S. Martí, V. Moliner and I. Tuñón, *J. Chem. Theory Comput.* **2005**, *1*, 1008-1016.
- [48] R. Wolfenden and M. J. Snider, *Acc. Chem. Res.* **2001**, *34*, 938-945.

- [49] J. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.* **2002**, *53*, 467-505.
- [50] M. Garcia-Viloca, J. Gao, M. Karplus and D. G. Truhlar, *Science* **2004**, *303*, 186-195.
- [51] S. Martí, M. Roca, J. Andres, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Chem. Soc. Rev.* **2004**, *33*, 98-107.
- [52] A. Shurki, M. Strajbl, J. Villa and A. Warshel, *J. Am. Chem. Soc.* **2002**, *124*, 4097-4107.
- [53] J. Villa, M. Strajbl, T. M. Glennon, Y. Y. Sham, Z. T. Chu and A. Warshel, *P. Natl. Acad. Sci. USA* **2000**, *97*, 11899-11904.
- [54] H. Guo, Q. Cui, W. N. Lipscomb and M. Karplus, *P. Natl. Acad. Sci. USA* **2001**, *98*, 9032-9037.
- [55] J. P. Ma, X. F. Zheng, G. Schnappauf, G. Braus, M. Karplus and W. N. Lipscomb, *P. Natl. Acad. Sci. USA* **1998**, *95*, 14640-14645.
- [56] S. Hur and T. C. Bruice, *J. Am. Chem. Soc.* **2003**, *125*, 10540-10542.
- [57] S. Hur and T. C. Bruice, *Proc. Natl. Acad. Sci.* **2003**, *100*, 2215-2219.
- [58] X. Prat-Resina, A. Gonzalez-Lafont and J. M. Lluch, *J. Mol. Struct. - Theochem.* **2003**, *632*, 297-307.
- [59] A. J. Mulholland, *Drug Discov. Today* **2005**, *10*, 1393-1402.
- [60] B. C. Garrett and D. G. Truhlar, *J. Am. Chem. Soc.* **1979**, *101*, 5207-5217.
- [61] B. C. Garrett and D. G. Truhlar, *Abstr. Pap. Am. Chem. S.* **1980**, *180*, 233.
- [62] D. G. Truhlar and J. Gao, *Acc. Chem. Res.* **2006**, *en imprenta*.
- [63] M. Roca, J. Andres, V. Moliner, I. Tunon and J. Bertran, *Journal of the American Chemical Society* **2005**, *127*, 10648-10655.
- [64] A. Soriano, E. Silla, I. Tuñón and M. F. Ruiz-Lopez, *J. Am. Chem. Soc.* **2005**, *127*, 1946-1957.
- [65] K. Nam, X. Prat-Resina, M. García-Viloca, L. S. Devi-Kesavan and J. Gao, *J. Am. Chem. Soc.* **2004**, *126*, 1369-1376.

- [66] M. Roca, V. Moliner, I. Tuñón and J. T. Hynes, *J. Am. Chem. Soc.* **2006**, *ASAP*.
- [67] W. Min, B. P. English, G. B. Luo, B. J. Cherayil, S. C. Kou and X. S. Xie, *Accounts Chem. Res.* **2005**, *38*, 923-931.
- [68] N. M. Antikainen, R. D. Smiley, S. J. Benkovic and G. G. Hammes, *Biochemistry* **2005**, *44*, 16835-16843.
- [69] W. H. Tan and E. S. Yeung, *Anal. Chem.* **1997**, *69*, 4242-4248.
- [70] Q. F. Xue and E. S. Yeung, *Nature* **1995**, *373*, 681-683.
- [71] I. F. Thorpe and C. L. Brooks, *J. Am. Chem. Soc.* **2005**, *127*, 12997-13006.
- [72] B. P. English, W. Min, A. M. van Oijen, K. T. Lee, G. B. Luo, H. Y. Sun, B. J. Cherayil, S. C. Kou and X. S. Xie, *Nat. Chem. Biol.* **2006**, *2*, 87-94.
- [73] A. Ansari, J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder and R. D. Young, *P. Natl. Acad. Sci. USA* **1985**, *82*, 5000-5004.
- [74] I. F. Thorpe and C. L. Brooks, *J. Phys. Chem. B* **2003**, *107*, 14042-14051.
- [75] Y. K. Zhang, J. Kua and J. A. McCammon, *J. Phys. Chem. B* **2003**, *107*, 4459-4463.
- [76] M. Klahn, S. Braun-Sand, E. Rosta and A. Warshel, *J. Phys. Chem. B* **2005**, *109*, 15645-15650.
- [77] A. J. Turner, V. Moliner and I. H. Williams, *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323-1331.
- [78] J. E. Basner and S. D. Schwartz, *Journal of the American Chemical Society* **2005**, *127*, 13822-13831.
- [79] D. G. Truhlar, J. L. Gao, M. Garcia-Viloca, C. Alhambra, J. Corchado, M. L. Sanchez and T. D. Poulsen, *Int. J. Quantum Chem.* **2004**, *100*, 1136-1152.

- [80] A. D. Isaacson, D. G. Truhlar, S. N. Rai, R. Steckler, G. C. Hancock, B. C. Garrett and M. J. Redmon, *Comput. Phys. Commun.* **1987**, *47*, 91-102.
- [81] J. Everse, R. E. Barnett, C. J. R. Thorne and N. O. Kaplan, *Archives of Biochemistry and Biophysics* **1971**, *143*, 444.
- [82] J. W. Burgner and W. J. Ray, *Biochemistry* **1974**, *13*, 4229-4237.

CAPÍTULO 5

DIHIDROFOLATO REDUCTASA

<u>CAPÍTULO 5. DIHIDROFOLATO REDUCTASA</u>	<u>187</u>
<u>1 BIOQUÍMICA</u>	<u>191</u>
<u>2 RESULTADOS</u>	<u>194</u>
2.1 RESULTADOS EN FASE GAS	195
2.2 RESULTADOS EN EL SENO DE LA ENZIMA DHFR	200
<u>3 PERSPECTIVAS DE FUTURO</u>	<u>208</u>
<u>4 BIBLIOGRAFÍA</u>	<u>210</u>

La Dihidrofolato Reductasa (DHFR) es una enzima que actúa en el metabolismo celular normal tanto en células eucarióticas como procarióticas.^[1, 2] Su función es catalizar la reducción de 7,8-dihidrofolato (DHF) a 5,6,7,8-tetrahidrofolato (THF) utilizando el NADPH como cofactor (Figura 1). También es capaz de catalizar la reducción de folato a DHF, pero a una velocidad mucho menor y que varía según la especie.

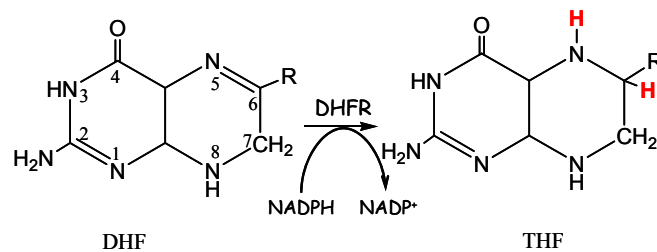


Figura 1. Reducción de 7,8-DHF a 5,6,7,8-THF catalizada por la DHFR utilizando el NADPH como cofactor.

La DHFR es una enzima pequeña, consta de un solo monómero, es estable y con un buen comportamiento cristalográfico. La DHFR de la *Escherichia coli* (EcDHFR) está formada por una sola cadena polipeptídica de 159 aminoácidos. La estructura cristalina de la EcDHFR sin modificar resuelta a 1,7 Å^[3] muestra una compleja estructura secundaria que consiste en 8 cadenas beta, 7 de las cuales son paralelas, 6 hélices alfa y 5 giros. Las cadenas β forman una lámina β hidrofóbica que actúa como el núcleo estructural de la DHFR (Figura 2). Estas propiedades, junto con su interés farmacológico debido a que forma parte del ciclo de síntesis de aminoácidos, están en el origen de los numerosos

estudios que se han llevado a cabo sobre esta enzima, en su mayor parte en la ecDHFR. Tanto los estudios experimentales, basados en métodos cinéticos,^[4-9] en espectroscopia de RMN,^[10-12] espectroscopia Raman,^[13, 14] difracción de rayos-X^[3, 15-17] y estudios de mutagénesis dirigida;^[18-24] como los estudios teóricos,^[25-33] han contribuido al conocimiento detallado de la estructura y función de la enzima.

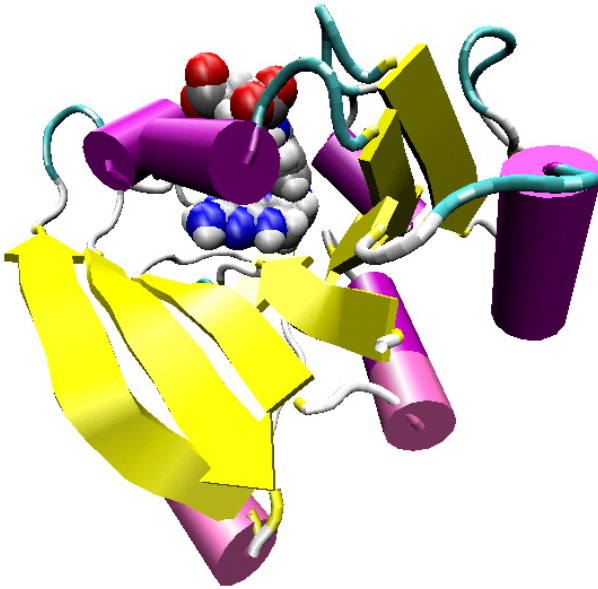


Figura 2. Estructura de la ecDHFR. Las cadenas beta están representadas con flechas, los cilindros y cadenas son giros y alfa-hélices y el centro activo está representado por esferas.

1 Bioquímica

El folato (Figura 3) está compuesto por tres partes:^[34] un anillo pteridino sustituido (6-metilpteridino), un segmento central de ácido p-aminobenzóico (pABA) y un residuo de glutamato. El pABA y el glutamato forman la parte llamada glutamato de p-aminobencilo (pABG). El anillo pteridino está constituido por un anillo 4-oxo-2-aminopirimidino fusionado con un anillo de pirazina, siendo esta última parte donde tiene lugar la reducción. El folato se convierte en 7,8-DHF al ser reducido el doble enlace C7-N8, y éste pasa, a su vez, a ser 5,6,7,8-THF al ser reducido el doble enlace N5-C6.

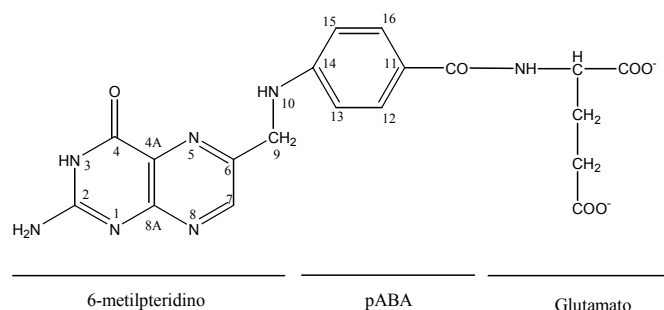


Figura 3. Estructura y numeración de los átomos que forman el folato

Dentro del metabolismo celular, la DHFR es de vital importancia debido a que mantiene la reserva de THF y de sus derivados.^[35] Éstos son cofactores esenciales en muchas reacciones metabólicas importantes, tales como la biosíntesis de timidilato, nucleótidos de purina y varias reacciones de movilización y transferencia de fragmentos de un solo carbono. En células que proliferan rápidamente, como son las células cancerígenas, la síntesis de ADN es elevada, de manera que existe una gran demanda del nucleótido timina, en cuya síntesis el coenzima es el THF, que actúa como dador de un carbono o

como fuente de electrones para la reducción que tiene lugar. Estos electrones provienen de la reducción del anillo de pteridina para dar como producto el DHF. Así, la DHFR debe actuar continuamente para regenerar el THF a partir del DHF formado. La inhibición de la enzima DHFR, por tanto, es el objetivo terapéutico de varias drogas cuya finalidad es detener la proliferación de células cancerígenas.^[36, 37]

Los estudios dedicados a esta enzima,^[6, 8, 9, 18, 19, 38-45] indican que el centro activo de la DHFR, presente en todos los organismos, es hidrófobo y que contienen un aminoácido con un sustituyente carboxilado ionizable, el Asp27 en el caso de la ecDHFR, que es importante para la actividad catalítica a pH neutro y se cree que juega un papel importante en la donación del protón durante la catálisis. El cofactor está próximo al sustrato, y la dihidronicotinamida está alineada con grupos de oxígenos coplanares. En la entrada del centro activo existe un aminoácido, Met20, cuyo giro controla la entrada y salida de los ligandos. Se piensa que los cambios conformacionales asociados a este aminoácido pueden desencadenar cambios conformacionales adicionales en otra parte de la proteína que transformen al complejo desde su estado fundamental a una estructura próxima a la EsT.^[46]

Un objetivo inmediato de la enzima es desarrollar una carga positiva parcial en el C6 del sustrato para promover la transferencia de un hidruro desde el C4 del anillo de dihidronicotinamida del cofactor. Esto puede ser asistido por la protonación del N5. En 1964, Huennekens y Scrimgeour^[47] propusieron la protonación previa como condición para la reducción. Un candidato obvio para la ceder un protón es el Asp27 en la ecDHFR, o el Glu30 en la DHFR humana. No obstante, el Asp27 está situado a 5 Å del N5 y del folato, por lo que transmisión del protón desde el N3 de la pteridina al O4 requeriría la ayuda de una molécula de agua

(ver Figura 4). Se ha observado que la estructura cristalina de la DHFR en todos los organismos presenta una molécula de agua próxima a un oxígeno carboxílico del Asp27 y al O4 del anillo pteridino. Por tanto es posible que el O4 enolizado y estabilizado sea el encargado de transmitir el protón vía una molécula de agua que se desliza bajo la Met20 y que es el dador último del protón al N5 (Figura 4). Una vez la protonación ocurre en el N5 se obtiene una carga positiva parcial en el C6 que promueve la transferencia de hidruro desde el cofactor.

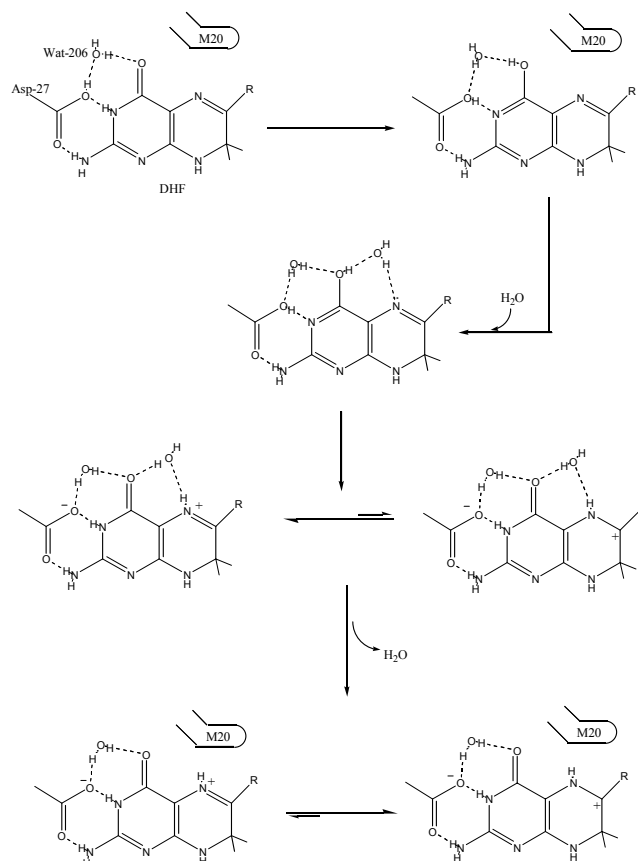


Figura 4. Mecanismo propuesto por Huennekens y Scrimgeour ^[47] para la protonación del N5 del 7,8-dihidrofolato con la intervención del Asp27. La metionina cercana está representada por M20.

2 Resultados

Los estudios computacionales previos al presente trabajo se centran en la transferencia de hidruro desde el cofactor NADPH al C6 del folato.^[25, 31, 33, 48-50] Por medio de un estudio teórico de las superficies de energía potencial procederemos a discernir el mecanismo de protonación del N5, etapa previa a la transferencia de hidruro que no ha sido estudiada anteriormente. Tal como se ha indicado anteriormente, (Figura 4) se han propuesto varios mecanismos para la transferencia de este protón desde el Asp27 al N5.

Tenemos en cuenta la transferencia de protón desde el Asp27 hasta el N3 del enol-DHF, de manera que se favorece la transferencia de un protón desde el O4 hasta el N5 a través de una molécula de agua (Figura 5) y la molécula queda preparada para la transferencia de hidruro. Estas transferencias pueden tener lugar simultáneamente o consecutivamente dando lugar a diferentes mecanismos.

Para poder estudiar la protonación del N5 del folato, partimos del enol dihidrofolato (enol-DHF) como reactivo. Los mecanismos que vamos a estudiar se muestran en la Figura 5, donde se indican las estructuras más relevantes para pasar desde el enol-DHF (estructura P1) hasta el trihidrofolato o H⁺DHF (estructura P4).

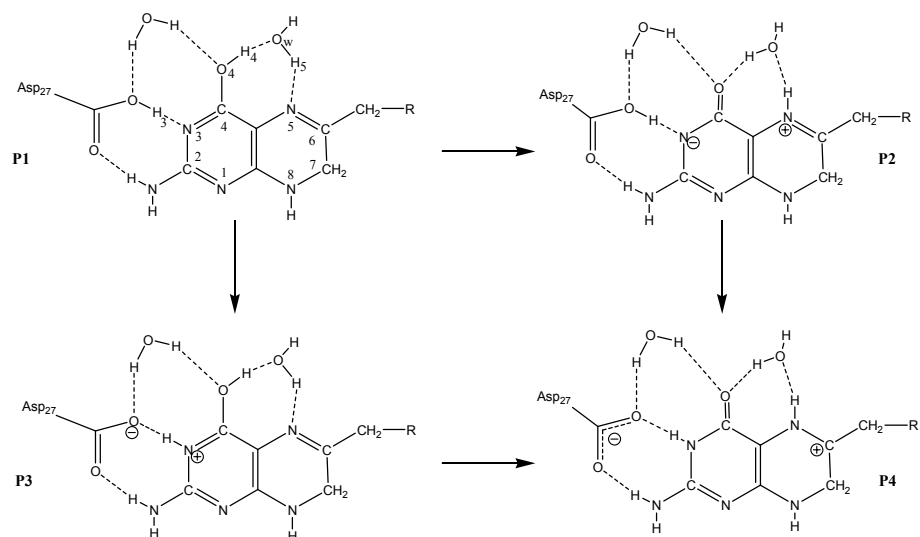


Figura 5. Posibles caminos estudiados para llegar del enol-DHF (P1) al producto con el nitrógeno protonado H⁺DHF(P4).

2.1 Resultados en fase gas

Para estudiar estos mecanismos en vacío simplificamos el sistema a estudiar sustituyendo la cadena lateral del folato (R) por un grupo metilo, y el aspartato por un ácido acético. Esta modelización se hace para que el coste computacional sea menor, especialmente cuando se emplean métodos basados en la teoría de funcional de la densidad (DFT). Como el aspártico está dentro de la secuencia de aminoácidos de la enzima en principio no debería tener libertad total de movimientos. Para imponer esta restricción en nuestros cálculos en vacío se congelaron las distancias, los ángulos y los ángulos diedros que definen la posición de los átomos de carbono del ácido acético.

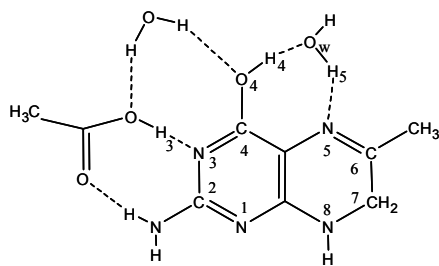


Figura 6. Representación del modelo utilizado en fase gas.

Utilizamos dos niveles de cálculo de manera que podremos comparar la bonanza del método semiempírico (AM1) frente a un cálculo de mayor nivel como es B3LYP/6-31G*. Las energías de los distintos puntos estacionarios localizados, relativas a P1, figuran en la Tabla 1. Cabe resaltar que a nivel semiempírico no ha sido posible localizar en vacío la estructura P3. Todo intento de localizar una estructura de este tipo llevaba a una transferencia espontánea del protón del ácido acético al anillo del folato. Así, la única estructura que une el reactivo P1 con el producto P4, es una estructura de transición (TS3) que conecta directamente P1 y P4. Por su parte, utilizando el funcional B3LYP no hemos sido capaces de localizar estructuras estacionarias correspondientes a P3 y P4. Este resultado confirma la importancia de la inclusión del efecto del entorno para tener una representación realista de los posibles mecanismos de reacción. La Tabla 1 y la Figura 7 resumen los resultados obtenidos para este sistema en vacío.

	<i>AM1</i>		<i>B3LYP</i> <i>6-31G*</i>	
	ΔE	ν	ΔE	ν
P1	0.00	--	0.00	--
TS1	37.24	1040i	11.73	927i
P2	9.65	--	6.84	--
TS2	30.21	955i	--	--
P4	27.58	--	--	--
TS3	54.95	1048i	--	--

Tabla 1 Energía relativa (kcal/mol) de los mínimos y estructuras de transición localizados en vacío, así como, las frecuencias imaginarias (en cm^{-1}) de las estructuras de transición.

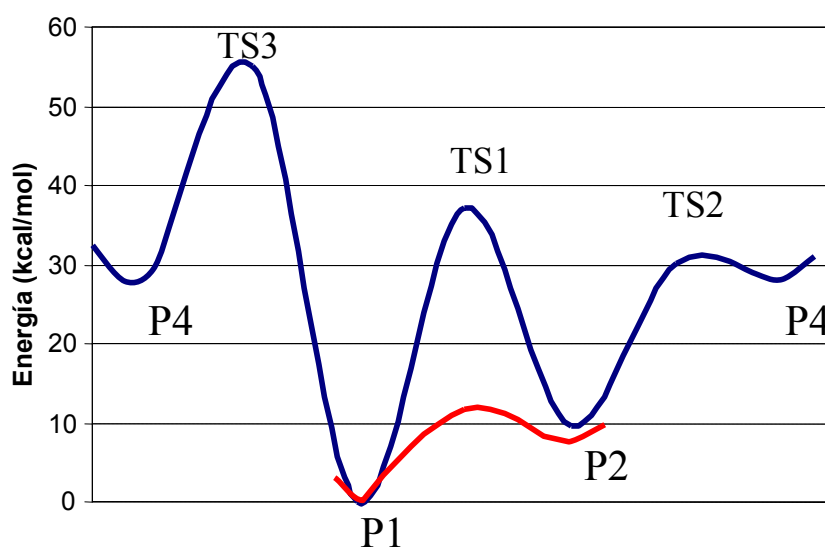


Figura 7. Perfil de energía potencial en vacío de los dos caminos de reacción obtenidos utilizando el método AM1 (en azul) y B3LYP (en rojo).

En la Figura 7 puede observarse como el método AM1 sobreestima de manera muy importante la barrera energética asociada a la EsT llamada TS1. Éste será un efecto que deberemos tener en consideración a la hora de analizar los resultados obtenidos en el seno de

la enzima, ya que en los métodos QM/MM, debido a las limitaciones en la capacidad de cálculo, nos vemos limitados a describir la región cuántica con un método semiempírico. Por otra parte, los métodos AM1 y B3LYP coinciden en la descripción termodinámica de la etapa P1-P2, dando valores muy parecidos de las energías de P2 relativas a P1.

	AM1			B3LYP/6-31G*		
	P1	TS1	P2	P1	TS1	P2
d(O4-H4)	0.976	1.519	2.186	0.997	1.253	1.685
d(H4-Ow)	2.084	1.076	0.964	1.710	1.177	1.003
d(Ow-H5)	0.962	1.717	1.970	0.992	1.273	1.690
d(N5-H5)	2.720	1.037	1.022	1.821	1.226	1.055
d(O_{asp}-H3)	0.976	0.982	0.980	1.007	1.011	1.014
d(N3-H3)	2.115	2.031	2.061	1.771	1.759	1.747

Tabla 2. Distancias (en Å) más importantes de los puntos estacionarios localizados en la etapa P1 a P2, obtenido a nivel semiempírico y a nivel B3LYP.

Además de comparar la energía entre los dos niveles de cálculo que nos muestra la Tabla 1, podemos hacer un análisis comparativo de las geometrías obtenidas con los dos métodos de cálculo. En la Tabla 2 podemos observar la diferencia entre las distancias de las estructuras que hemos podido localizar con el método B3LYP, pertenecientes a la primera etapa del camino de reacción, P1-P2-P4. Estas estructuras implican la transferencia del protón H5 desde el O4 al N5 a través de una molécula de agua. La diferencia más destacable reside en las distancias del O4-H4 en P2, que en AM1 vale 2.19 Å mientras que en B3LYP es 1.69 Å; y las distancias N5-H5 en P1 ya que en AM1 es 2.72 Å y en B3LYP es 1.82 Å. Ambas distancias corresponden a enlaces de hidrógeno entre la molécula de agua y el sustrato. Los valores geométricos sugieren que en B3LYP esta interacción es más intensa y el

agua en la estructura optimizada está más próxima a la molécula a la que va a asistir en la transferencia que en el caso obtenido en AM1.

2.2 Resultados en el seno de la enzima DHFR

Tal como hemos indicado en el apartado anterior, para estudiar el efecto del medio enzimático hemos utilizado la metodología híbrida QM/MM. El estudio lo hemos realizado tomando como modelo el complejo ternario de la ecDHFR con folato y NADPH a partir de una estructura obtenida por difracción de RX con una resolución de 1.7 Å^[35] Teniendo las coordenadas cartesianas de los átomos pesados que forman la enzima (obtenidos del Protein Data Bank, PDB), le añadimos hidrógenos a pH fisiológico, según los pK_a de los aminoácidos en disolución. Para simplificar el cálculo y optimizar el tiempo de CPU decidimos tomar una esfera de 17 Å con centro en el átomo C6 de la DHF y fijamos las coordenadas de los átomos localizados más allá de esta distancia. También se añadió una esfera de 17 Å de moléculas de agua (TIP3P) centrada en el C6 para simular que la enzima está en disolución y se eliminan aquellas moléculas de agua cuyo átomo de oxígeno se encuentra a menos de 2.8 Å de los átomos de nuestro sistema. El sistema resultante está formado por 3385 átomos de los cuales 608 permanecen con sus posiciones fijas.

El sistema total se dividió en una región QM de 36 átomos, donde se incluyen el anillo 6-metilpterino del dihidrofolato, el Asp27 hasta el carbono alpha y dos moléculas de agua. La región QM se describe mediante el Hamiltoniano semiempírico AM1, mientras que para la región MM formada por el resto de la proteína y moléculas de agua, se utiliza el potencial de CHARMM,^[51] y TIP3P.^[52] Al hacer la partición del sistema completo en las dos regiones (QM y MM) la frontera pasa necesariamente a través de dos enlaces covalentes, para lo cual hemos utilizado la metodología de átomos *link*. (ver Figura 8).

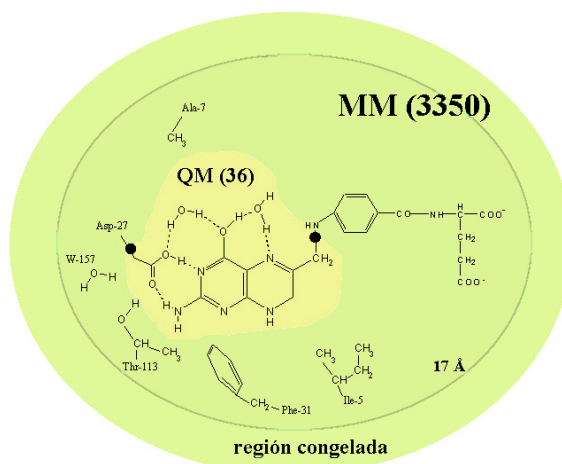


Figura 8. Representación esquemática del modelo utilizado en el estudio de la DHFR. La región QM y MM quedan distinguidas por los colores amarillo y verde respectivamente. Los átomos frontera se representan con puntos negros y son tratados con la metodología de átomo *link*.

Para localizar las estructuras de los puntos estacionarios de los mecanismos propuestos inicialmente (Figura 5) hemos tenido que estudiar las SEPs del sistema mediante el uso de diferentes coordenadas distinguidas de reacción. La exploración de las SEPs se realiza con el programa CHARMM versión 26b. Una vez obtenidas las SEPs se localizan las estructuras estacionarias con la combinación de programas GRACE/CHARMM.^[53, 54] De modo que podemos obtener las EsT mediante el cálculo de la hessiana y conocer la frecuencia imaginaria que posee cada una de ellas.

El esquema correspondiente al camino de reacción P1-P2-P4, donde inicialmente ocurre la transferencia de protón desde el O4 al N5 por medio de una molécula de agua y seguidamente se transfiere el protón desde OD2 del Asp27 al N3 del anillo pteridino se presenta en la Figura 9.

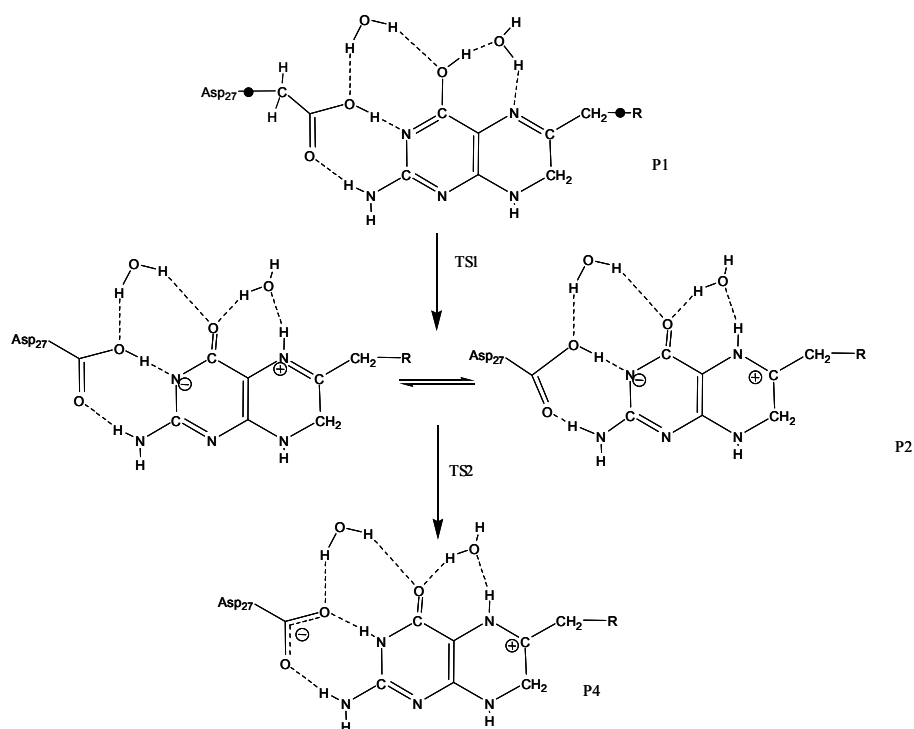


Figura 9. Esquema del camino de reacción P1-P2-P4. El dibujo sitúa los átomos link en el P1.

Para estudiar el paso de P1 a P2 definimos dos coordenadas distinguidas que son las combinaciones antisimétricas de las distancias entre el oxígeno del enol (O4) y el hidrógeno (H4) que se va a unir a la molécula de agua, y la distancia entre el hidrógeno del agua (H5) y el nitrógeno (N5) que se tiene que protonar. Así, se obtuvo una SEP tridimensional donde se localizó un camino que tiene una única estructura de transición (TS1). Para el paso de P2 a P4, la coordenada de reacción se define igualmente como la distancia entre el átomo que se transfiere y el átomo aceptor, es decir, la distancia entre el hidrógeno del aspartato y el N3, obteniendo una SEP 2D en la que aparece una EsT que llamamos TS2.

En la Tabla 3 se puede observar los resultados obtenidos tras la exploración y localización de los puntos estacionarios. Se muestra la energía relativa a los reactivos de modo que se aprecian las barreras de la reacción. En el seno de la enzima las estructuras P2 y P4 son más estables que la estructura de reactivos P1.

	ΔE	Distancias (Å)					ν (cm ⁻¹)
	(kcal/mol)	d(O4-H4)	d(Ow-H4)	d(Ow-H5)	D(N5-H5)	d(N3-H3)	
P1	0.0	0.98	1.96	0.97	2.84	2.36	
TS1	23.2	1.44	1.11	0.99	2.65	2.21	660i
P2	-2.8	3.6	0.96	2.04	1.03	2.22	
TS2	9.9	3.36	0.96	2.03	1.03	1.29	1279i
P4	-5.9	3.24	0.96	2.01	1.03	1.02	

Tabla 3. Energía total, energías relativas, distancias y frecuencias imaginarias de los puntos estacionarios localizados en el seno de la enzima DHFR en el camino P1-P2-P4.

El paso lento de la reacción por el camino P1-P2-P4 es el que implica el TS1 ya que tiene una barrera de 23.2 kcal/mol. La frecuencia imaginaria que caracteriza a esta EsT es de 660i cm⁻¹ asociada al movimiento de los dos átomos de hidrógeno y los átomos dador y aceptor.

Las SEPs asociadas al mecanismo P1-P3-P4 describen un proceso donde inicialmente ocurre la transferencia de protón desde el OD2 del Asp27 al N3 del anillo pteridino, seguido de la transferencia del protón del O4 al N5 por medio de una molécula de agua. Es decir, el mecanismo ocurre de manera inversa al camino P1-P2-P4. La Figura 10 representa esquemáticamente las estructuras más relevantes correspondientes a este camino.

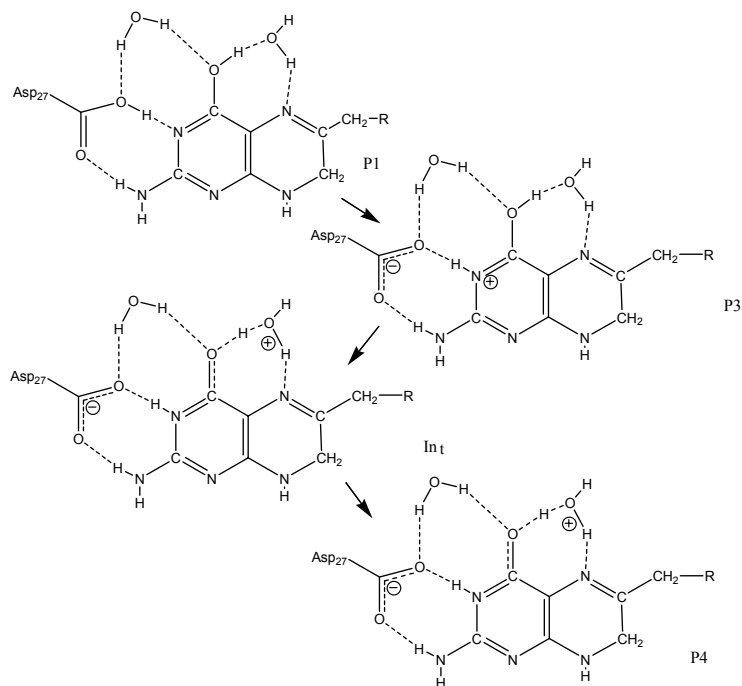


Figura 10. Esquema del camino P1-P3-P4. El paso de P3 a P4 implica la aparición de una especie intermedia (Int).

Entre P1 y P3 seleccionamos una única coordenada de reacción distinguida que define la distancia N3-H3, y obtenemos una única EsT (TS3). Finalmente, entre P3 y P4 definimos las mismas coordenadas que utilizamos para el estudio de la transformación de P1 a P2 obteniendo una SEP tridimensional que nos permite localizar un intermedio estable (Int) entre P3 y P4 y dos EsT diferentes (TSI y TSII). Este intermedio se corresponde con la formación de un ión hidronio en el seno de la enzima (Figura 11).

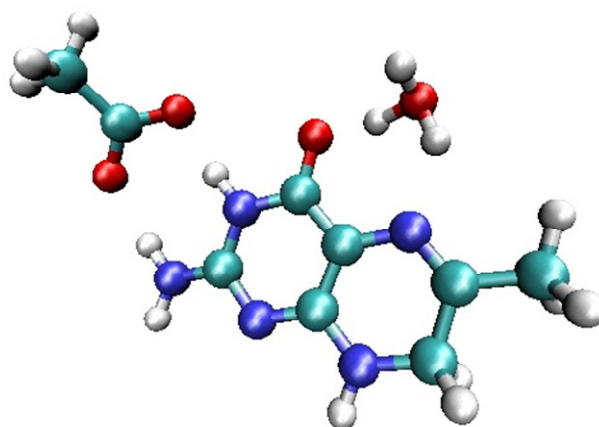


Figura 11. Intermedio (Int) formado en el camino P1-P3-P4. En la transferencia del protón desde el O4 al N5 se forma un intermedio estable en forma de ión hidronio.

En la Tabla 4 se presentan energías y parámetros geométricos obtenidos tras la localización de los puntos estacionarios, así como las frecuencias imaginarias que presentan las EsT. Se observa una frecuencia mayor en la EsT que implica la transferencia del protón desde el aminoácido.

	ΔE (kcal/mol)	Distancias (Å)					ν (cm ⁻¹)
		d(O4-H4)	d(Ow-H4)	d(Ow-H5)	d(N5-H5)	d(N3-H3)	
P1	0.0	0.98	1.96	0.97	2.84	2.36	--
TS3	13.0	0.98	1.94	0.97	2.89	1.20	1223i
P3	6.6	0.98	1.92	0.97	2.93	1.02	--
TSI	22.4	1.26	1.29	0.99	2.64	1.01	827i
Int	7.2	2.18	1.02	1.02	2.82	1.01	--
TSII	9.3	2.50	1.00	1.21	1.35	1.01	924i
P4	-5.9	3.24	0.96	2.01	1.03	1.02	--

Tabla 4. Energía relativa, distancias y frecuencias imaginarias de los puntos estacionarios localizados en el seno de la enzima DHFR en el camino P1-P3-P4.

En la Figura 12 se representa el perfil de la energía potencial frente a una coordenada de reacción genérica. Con esta representación

podemos discutir qué camino es el más favorecido en medio enzimático. En el centro de la gráfica se encuentra el reactivo P1 mientras que en los dos extremos aparece la estructura de productos P4 a los que se llega mediante los dos caminos especificados anteriormente.

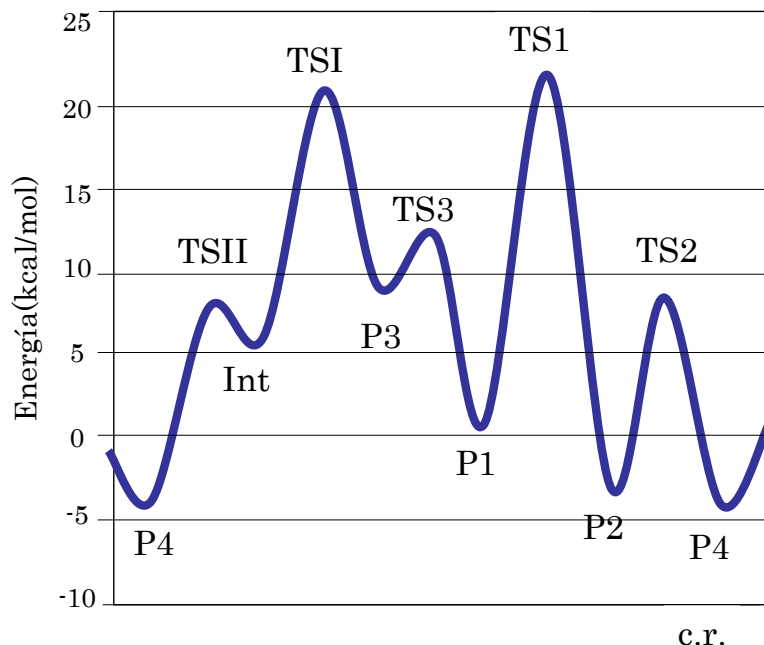


Figura 12. Representación de los dos caminos de reacción obtenidos mediante la metodología QM/MM. Energía relativa frente a coordenada de reacción.

Es interesante destacar que, frente a los resultados obtenidos en fase gas (Figura 7), la enzima favorece la formación de las estructuras propuestas en nuestro esquema inicial de trabajo (ver Figura 5). Esta diferencia indica la capacidad de la enzima de estabilizar la presencia de grupos cargados mediante interacciones específicas presentes en el centro activo. P4 resulta ser la estructura más estable en el seno de la enzima, mientras que en fase gas estaba situada casi 30 kcal/mol por encima de la estructura de reactivos. La barrera más alta aparece asociada al camino P1-P4, aunque, como ya hemos discutido, el error del método AM1 en la

estimación de las barreras nos impide alcanzar conclusiones definitivas. La utilización de Hamiltonianos de mayor nivel produciría, sin duda, unas barreras menores.

Con respecto a los residuos del centro activo (Figura 13), la mayoría de las distancias entre la proteína y el sustrato permanecen constantes yendo desde reactivos a productos, como se puede apreciar en la Tabla 5. Estos resultados pueden relacionarse con una pequeña reorganización de los alrededores de la enzima durante la reacción catalizada. El hecho de que la estructura de la enzima no cambie considerablemente es algo ya observado en otros estudios teóricos de otras enzimas [55, 56], pero que normalmente se asocia a procesos de un único paso.

	P1	P2	P3	In	P4
O(ILE5):H8(DHF)	1.86	1.87	1.84	1.86	1.81
O(w157):H91(DHF)	1.99	2.03	1.99	2.00	1.97
H61(THR113):OD2(ASP)	1.95	1.94	1.85	1.95	1.84
HD1(PHE31):N1(DHF)	2.69	2.68	2.70	2.70	2.71
HG12(ILE5):N8(DHF)	3.32	3.39	3.42	3.31	3.41
HB2(ALA7):N3(DHF)	2.71	2.73	2.78	2.72	2.86

Tabla 5. Distancias (en Å) más importantes de los átomos QM con los residuos del entorno en los mínimos locales de las SEPs localizados a nivel AM1/MM.

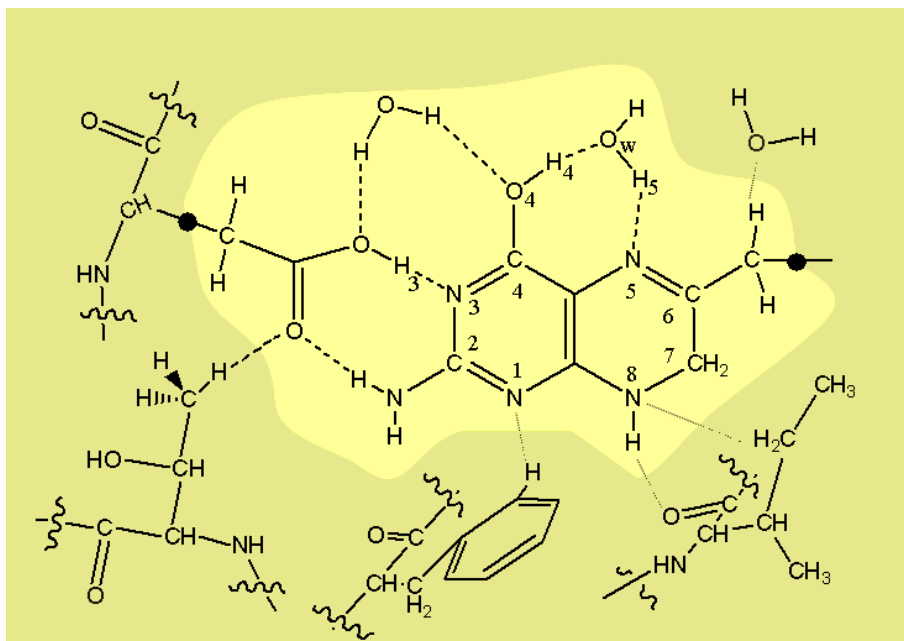


Figura 13. Esquema de las interacciones entre el sustrato y los residuos del centro activo de la enzima

3 Perspectivas de Futuro

El interés principal de esta enzima desde el punto de vista computacional reside en que se ha descrito el acoplamiento entre cambios conformacionales de la misma y la etapa química.^[31, 57] Este fenómeno ha sido estudiado recientemente para la etapa lenta, la transferencia de hidruro, pero no para las etapas previas. Tras haber realizado una exploración previa de la SEP AM1/MM y haber calibrado los errores inherentes a la descripción QM para la etapa de protonación del N5, nos proponemos aplicar un protocolo similar al desarrollado para la LDH. En concreto, los pasos a dar implicarían:

- Obtención de nuevas SEP incluyendo correcciones al Hamiltoniano semiempírico.

- Realización de dinámicas sobre la región del ET de la etapa lenta de la protonación del N5.
- Localización de estructuras de transición con características diferenciadoras y realización de los correspondientes IRCs.
- Obtención de los caminos de energía libre asociadas mediante técnicas de FEP o PMF.
- Estudio del efecto túnel.

4 Bibliografía

- [1] R. L. Blakley and S. J. Benkovic, *Chemistry and Biochemistry of Pterins*, Wiley, New York, **1985**.
- [2] R. L. Blakley and S. J. Benkovic, *Chemistry and Biochemistry of Folates*, Wiley, New York, **1984**.
- [3] J. T. Bolin, D. J. Filman, D. A. Matthews, R. C. Hamlin and J. Kraut, *J. Biol. Chem.* **1982**, *257*, 3650-3662.
- [4] S. R. Stone, A. Mark and J. F. Morrison, *Biochemistry* **1984**, *23*, 4340-4346.
- [5] S. R. Stone and J. F. Morrison, *Biochemistry* **1982**, *21*, 3757-3765.
- [6] C. A. Fierke, K. A. Johnson and S. J. Benkovic, *Biochemistry* **1987**, *26*, 4085-4092.
- [7] M. H. Penner and C. Frieden, *J. Biol. Chem.* **1987**, *262*, 15908-15914.
- [8] S. M. J. Dunn, T. M. Lanigan and E. E. Howell, *Biochemistry* **1990**, *29*, 8569-8576.
- [9] J. R. Appleman, E. E. Howell, J. Kraut and R. L. Blakley, *J. Biol. Chem.* **1990**, *265*, 5579-5584.
- [10] S. J. Hammond, B. Birdsall, M. S. Searle, G. C. K. Roberts and J. Feeney, *J. Mol. Biol.* **1986**, *188*, 81-97.
- [11] R. E. London, E. E. Howell, M. S. Warren, J. Kraut and R. L. Blakley, *Biochemistry* **1986**, *25*, 7229-7235.
- [12] R. L. Blakley, J. R. Appleman, J. H. Freisheim and M. J. Jablonsky, *Arch. Biochem. Biophys.* **1993**, *306*, 501-509.
- [13] Y. Q. Chen, J. Kraut and R. Callender, *Biophys. J.* **1997**, *72*, 936-941.
- [14] Y. Q. Chen, J. Kraut, R. L. Blakley and R. Callender, *Biochemistry* **1994**, *33*, 7021-7026.
- [15] D. A. Matthews, J. T. Bolin, J. M. Burrige, D. J. Filman, K. W. Volz, B. T. Kaufman, C. R. Beddell, J. N. Champness, D. K. Stammers and J. Kraut, *J. Biol. Chem.* **1985**, *260*, 381-391.

- [16] D. J. Filman, J. T. Bolin, D. A. Matthews and J. Kraut, *J. Biol. Chem.* **1982**, *257*, 3663-3672.
- [17] D. A. Matthews, R. A. Alden, J. T. Bolin, S. T. Freer, R. Hamlin, N. Xuong, J. Kraut, M. Poe, M. Williams and K. Hoogsteen, *Science* **1977**, *197*, 452-455.
- [18] E. E. Howell, J. E. Villafranca, M. S. Warren, S. J. Oatley and J. Kraut, *Science* **1986**, *231*, 1123-1128.
- [19] J. E. Villafranca, E. E. Howell, D. H. Voet, M. S. Strobel, R. C. Ogden, J. N. Abelson and J. Kraut, *Science* **1983**, *222*, 782-788.
- [20] J. T. Chen, K. Taira, C. P. D. Tu and S. J. Benkovic, *Biochemistry* **1987**, *26*, 4093-4100.
- [21] K. Taira, J. T. Chen, R. J. Mayer and S. J. Benkovic, *B. Chem. Soc. Jpn.* **1987**, *60*, 3017-3024.
- [22] S. J. Benkovic, C. A. Fierke and A. M. Naylor, *Science* **1988**, *239*, 1105-1110.
- [23] C. A. Fierke and S. J. Benkovic, *Biochemistry* **1989**, *28*, 478-486.
- [24] M. S. Warren, K. A. Brown, M. F. Farnum, E. E. Howell and J. Kraut, *Biochemistry* **1991**, *30*, 11092-11103.
- [25] J. Andres, V. Moliner, V. S. Safont, L. R. Domingo, M. T. Picher and J. Krechl, *Bioorg. Chem.* **1996**, *24*, 10-18.
- [26] J. Bajorath, J. Kraut, Z. Q. Li, D. H. Kitson and A. T. Hagler, *P. Natl. Acad. Sci. USA* **1991**, *88*, 6423-6426.
- [27] P. L. Cummins and J. E. Gready, *J. Comput. Aid. Mol. Des.* **1993**, *7*, 535-555.
- [28] P. L. Cummins and J. E. Gready, *Proteins* **1993**, *15*, 426-435.
- [29] P. L. Cummins and J. E. Gready, *Mol. Simulat.* **1995**, *15*, 155-175.
- [30] P. L. Cummins and J. E. Gready, *J. Phys. Chem. B* **2000**, *104*, 4503-4510.
- [31] M. Garcia-Viloca, D. G. Truhlar and J. L. Gao, *Biochemistry* **2003**, *42*, 13558-13575.

- [32] M. Garcia-Viloca, D. G. Truhlar and J. L. Gao, *J. Mol. Biol.* **2003**, *327*, 549-560.
- [33] R. Castillo, J. Andres and V. Moliner, *J. Am. Chem. Soc.* **1999**, *121*, 12140-12147.
- [34] H. Deng and R. Callender, *J. Am. Chem. Soc.* **1998**, *120*, 7730-7737.
- [35] M. R. Sawaya and J. Kraut, *Biochemistry* **1997**, *36*, 586-603.
- [36] G. H. Hitchings and B. Roth, *Enzyme Inhibitors as Drugs*, Macmillan, London, **1980**.
- [37] A. H. Calvert, T. R. Joner, D. P. J., B. Grzelakowska-Sztabert, R. Paine and G. A. Taylor, *Eur. J. Cancer* **1980**, *16*, 713.
- [38] J. R. Appleman, E. E. Howell, J. Kraut, M. Kuhl and R. L. Blakley, *J. Biol. Chem.* **1988**, *263*, 9187-9198.
- [39] E. E. Howell, C. Booth, M. Farnum, J. Kraut and M. S. Warren, *Biochemistry* **1990**, *29*, 8561-8569.
- [40] C. L. David, E. E. Howell, M. F. Farnum, J. E. Villafranca, S. J. Oatley and J. Kraut, *Biochemistry* **1992**, *31*, 9813-9822.
- [41] A. Dion, C. E. Linn, T. D. Bradrick, S. Georghiou and E. E. Howell, *Biochemistry* **1993**, *32*, 3479-3487.
- [42] V. A. Karginov, S. V. Mamaev, H. Y. An, M. D. VanCleve, S. M. Hecht, G. A. Komatsoulis and J. N. Abelson, *J. Am. Chem. Soc.* **1997**, *119*, 8166-8176.
- [43] S. J. Benkovic and S. Hammes-Schiffer, *Science* **2003**, *301*, 1196-1202.
- [44] S. Hammes-Schiffer, *J. Chem. Phys.* **1996**, *105*, 2236-2246.
- [45] P. T. R. Rajagopalan, S. Lutz and S. J. Benkovic, *Biochemistry* **2002**, *41*, 12618-12628.
- [46] L. Y. Li, C. J. Falzone, P. E. Wright and S. J. Benkovic, *Biochemistry* **1992**, *31*, 7826-7833.
- [47] F. M. S. Huennekens, K. G., *Pteridine chemistry : proceedings of the Third International Symposium held at the Institut fuer Organische Chemie der Technischen Hochschule, Stuttgart, September 1962*, Pergamon, **1964**.

- [48] P. L. Cummins, S. P. Greatbanks, A. P. Rendell and J. E. Gready, *J. Phys. Chem. B* **2002**, *106*, 9934-9944.
- [49] P. L. Cummins and J. E. Gready, *J. Am. Chem. Soc.* **2001**, *123*, 3418-3428.
- [50] P. L. Cummins and J. E. Gready, *J. Comput. Chem.* **1998**, *19*, 977-988.
- [51] B. R. Brooks, R. E. Brucoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187-217.
- [52] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926-935.
- [53] V. Moliner, A. J. Turner and I. H. Williams, *J. Chem. Soc. Chem. Commun.* **1997**, *14*, 1271.
- [54] A. J. Turner, V. Moliner and I. H. Williams, *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323-1331.
- [55] S. Marti, V. Moliner, I. Tunon and I. H. Williams, *Org. Biomol. Chem.* **2003**, *1*, 483-487.
- [56] R. A. Torres and T. C. Bruice, *J. Am. Chem. Soc.* **2000**, *122*, 781-791.
- [57] I. F. Thorpe and C. L. Brooks, *J. Am. Chem. Soc.* **2005**, *127*, 12997-13006.

ANEXO

Silvia Ferrer, Estanislao Silla, Iñaki Tuñón, Sergio Martí, Vicent Moliner
J. Phys. Chem, **2003**, 107, 14036-14041

Catalytic Mechanism of Dihydrofolate Reductase Enzyme. A Combined Quantum-Mechanical/Molecular-Mechanical Characterization of the N5 Protonation Step

Silvia Ferrer, Estanislao Silla, and Iñaki Tuñón*

Departament de Química Física; Universidad de Valencia, 46100 Burjassot, Valencia, Spain

Sergio Martí and Vicent Moliner*

Departament de Ciències Experimentals, Universitat Jaume I, 12071 Castelló, Spain

Received: May 28, 2003; In Final Form: September 22, 2003

The catalytic mechanism of dihydrofolate reductase (DHFR) requires the addition of both a proton and a hydride ion. In *Escherichia coli* dihydrofolate reductase (ecDHFR) Asp27 is considered as the only group in the active site capable of providing the proton for the reduction of the N5–C6 bond, although it is not clear enough if it takes place directly or through a series of water molecules. In this paper we present a theoretical study of the protonation of N5 of the hydrofolate, the prior step of the hydride transfer in dihydrofolate reductase. Hybrid quantum-mechanical/molecular-mechanical (QM/MM) calculations involving a flexible active-site region are used in combination with GRACE software. Two different reaction paths have been found to be feasible. The proposed mechanisms are in agreement with the ordered structure of the X-ray crystallographic water molecules supporting the hypothesis of an indirect proton transfer from Asp27 residue to the N5 atom of the substrate. Asp27 is found to play an important role stabilizing the cationic pteridine ring.

Introduction

Dihydrofolate reductase (5,6,7,8-tetrahydrofolate, NADP oxidoreductase, E.C. 1.5.1.3) catalyses the NADPH-dependent reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) in both bacterial and vertebrate cells. This enzyme is necessary for maintaining intracellular pools of THF and its derivatives, which are essential cofactors in the biosynthesis of several amino acids. This property has made DHFR a clinical target for antitumor and anti-infective therapy, and numerous inhibitors have been described.^{1,2} Nevertheless, the details of the enzyme molecular reaction mechanism are still unclear and can only be hypothesized at present.^{3,4}

The mechanism of the reaction requires the addition of both a proton and a hydride ion. In *E. coli* dihydrofolate reductase (ecDHFR) Asp27 is considered as the only residue in the active site region capable of providing the proton for the reduction of the N5–C6 bond,⁵ although it is not clear enough if it takes place directly or through a series of water molecules.⁶ In fact, mutation of this residue to a less acidic one (asparagine or serine) results in greatly diminished activity toward dihydrofolate at pH 7, but comparable activity to native DHFR at pH values sufficiently low to protonate dihydrofolate in solution.⁵ The presence of a solvent channel with a series of ordered water molecules hydrogen-bonded to the substrate and the protein, and the fact that Asp 27 was too far from the N5 (the proton acceptor atom) support the hypothesis of an indirect protonation to occur. Nevertheless, because various mechanisms for this proton shuttling between Asp27 and N5 have been proposed,^{4,7–11} and as a substrate binding role has also been conferred to the Asp 27 residue,¹² further studies are merited to clarify this regard. In fact, a recent contribution of Cummins et al.¹⁰ focused on the study of the effect of the protonated/unprotonated state of Asp27 on the hydride transfer step of DHFR arrived to the

conclusion that although ab initio and DFT gas phase cluster calculations suggested that the reduction took place when Asp27 was protonated, the same conclusions could not be drawn from hybrid PM3/molecular mechanical force fields plus MD model calculations. It is generally accepted that the protonation of the 7,8-dihydropteridine ring precedes hydride ion transfer and activates the substrate,^{13–16} thus being a chemical reaction strongly pH dependent.⁶

It is important to stress on the importance of the inclusion of environment effects to obtain a correct description of the process, not only from a quantitative point of view but also from a qualitative perspective. Hybrid quantum-mechanical/molecular-mechanical (QM/MM) methods provide a powerful way to deal with chemical reactions in very large systems. In these methods, only a small portion of the system is described using quantum mechanics, whereas for the remaining (typically several thousands of atoms) a molecular mechanics force field is employed. The exact characterization of molecular mechanisms requires the location of the different stationary points such as minima associated with reactants, products, and possible intermediates and saddle points associated with the transition structure (TS) on the potential energy surface (PES). Exploration of PESs with a very high dimensionality has become very recently affordable.¹⁷ This is a crucial step to know the molecular details of chemical reactions occurring in enzymatic media. In a previous study we reported the results of a combined classical/quantum modeling study for the reduction of 5,6-dihydrofolate catalyzed by ecDHFR, the key point being the characterization of the TS associated with the hydride transfer from the cofactor (nicotinamide adenine dinucleotide phosphate, NADPH).¹⁸ In the present paper, we report the results of the previous step, the proton transfer, presumably from Asp27 to the N5 atom of the substrate. This global transfer process from Asp27 to N5 implies

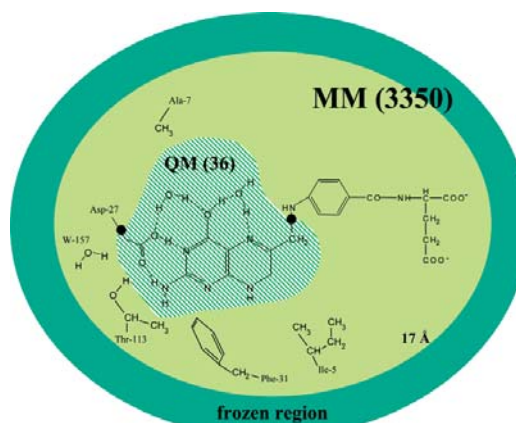
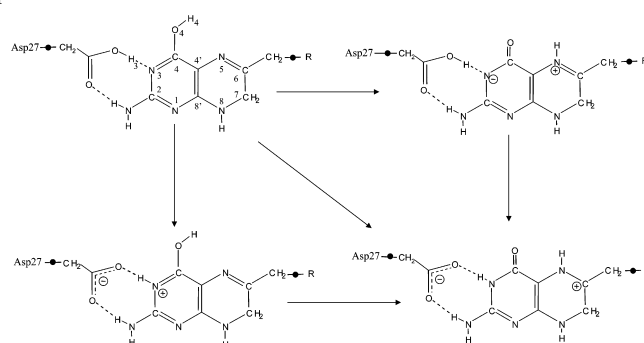


Figure 1. Schematic pictorial of the QMMM model. Hydrogen quantum link atoms, indicated as dots "*", divide the quantum (shaded area) and classical (solid color areas) regions. The darkest gray region contains the classical described atoms that were frozen during the optimizations.

SCHEME 1



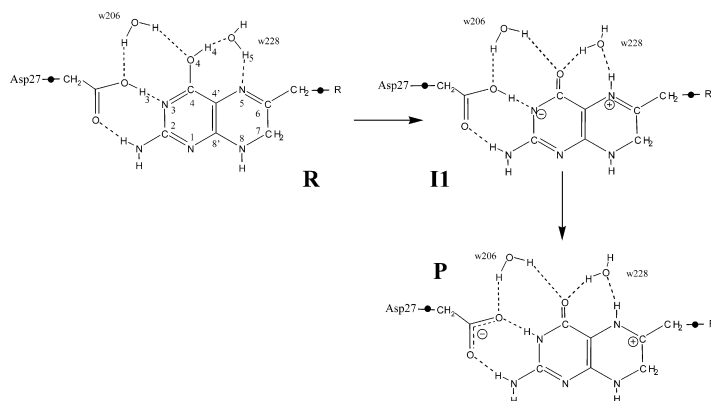
two different proton transfers: from Asp27 to the N3 atom of the substrate and from the enol oxygen (O4) to N5. These two transfers can take place either at the same time or in a stepwise manner, which would render two different paths depending on the order in which the two transfers occur. Thus, in principle, up to three different routes would be possible, as shown in Scheme 1.

Computational Details

The QM/MM energy hypersurface was obtained using the CHARMM24 program¹⁹ as described previously for reactions in aqueous solution²⁰ and in solvated enzyme active sites.^{17,20,21} The starting geometry comes from the 1.6 Å resolution ternary complex X-ray crystal structure of an Escherichia Coli DHFR (ecDHFR-NADPH-Folate).²² CHARMM24 was also used to add hydrogens to all titratable residues at a state complementary

to pH 7. The system was placed in a cavity deleted from a preformed 17 Å radius sphere of TIP3P²³ water molecules centered on the N5 nitrogen atom. Once the entire system was built, those atoms 17 Å away from the active site were frozen during the optimizations to maintain the water sphere and crystal structure. The resulting model, depicted in Figure 1, was a pseudosphere of 3385 atoms, 608 of them kept frozen. The full system was divided into a QM region (36 atoms consisting on the same atoms used for the model in gas phase plus an extra water molecule), treated by the AM1 semiempirical MO method, and a MM region comprising the rest of the protein (CHARMM24 potentials) and the water molecules. Of course, it would be ideal to use higher level methods, because semiempirical Hamiltonians can be problematic for describing bond breaking/forming processes. However, this is computationally impractical because our calculations typically involve

SCHEME 2



thousands of optimization steps and the evaluation of a large Hessian matrix several times along the stationary point search. Two hydrogen "link atoms",²⁴ where covalent bonds cross the boundary between the QM and the MM regions, were added to satisfy the valence of the QM fragments. The structural formula of O4 protonated DHF (DHF) is depicted in Figure 1, where the quantum link atoms are indicated as "w".

In the initial complex we placed the substrate in the product state. This structure was fully relaxed and then several scanning grids were used for the initial step of the PES generation. Once the final PESs were obtained, localized approximate TSs were refined using the GRACE software.¹⁷ A partitioned rational function optimizer combined with an adopted-basis-Newton-Raphson method was employed, utilizing a finite-difference Hessian matrix of order 108, describing the curvature of the QM/MM energy hypersurface for the quantum region subset of the system, together with a diagonal Hessian plus updates for the rest of the system. The rms residual gradient on the 36 atoms in the subset has been imposed to be less than 0.1 kcal/mol \AA^{-1} in the optimized structure, whereas on the remaining atoms (ca. 3000) it is less than 0.005 kcal/mol \AA^{-1} ; these residual gradients are lower than the commonly accepted convergence criterion for optimized geometries of small molecules in quantum chemistry.²⁵ Finally, intrinsic reaction coordinate (IRC)²⁶ paths were traced from the TSs in each direction to demonstrate conclusively that the reported structures are indeed the TS for the correct step under study.

Finally, once the QM/MM potential energy profiles were obtained, the correction of the zero point vibrational energy was estimated for the subset of atoms included in the Hessian calculation. Again, AMI presents some limitations as this method already contains zero point energies in its parametrization and then some contributions could be double counted. However, the neglect of the zero-point energy variation associated with the imaginary-frequency mode can lead to larger errors and then it may be preferable to include this term explicitly in the calculations.²⁷

Results and Discussion

From the exploration of the AMI/MM PES we obtained two possible reaction paths. These two paths correspond to two possible stepwise routes: (1) a proton transfer from enol O4

atom to N5 followed by the proton transfer from Asp27 to N3 and (2) a proton transfer from Asp27 to N3 followed by the transfer from O4 to N5 (this last transfer has been found to take place in two steps with the formation of an additional intermediate). The first path corresponds to the $R \rightarrow I1 \rightarrow P$ process (Scheme 2) through TS1 and TS2 transition structures and the second one to the $R \rightarrow I2 \rightarrow I3 \rightarrow P$ process (Scheme 3) through TS3, TS4, and TS5 transition structures. It is important to stress on the nature of the I3 structure: although, a priori, the step from I2 to P in path 2 could be expected to take place as a single step, we located a stationary structure where the H4 proton transfer is not coupled with the H5 proton transfer. A protonated water 228 is stabilized by the environment thus obtaining a new local minimum, I3 (see Scheme 3). Total and relative energies to R, obtained with the AM1/MM calculations, are reported in Tables 1 and 2 for paths 1 and 2, respectively. We also present some selected interatomic distances for all the stationary point structures and the imaginary frequencies characteristic of the transition structures. From an energetic point of view, a first insight into Tables 1 and 2, or the resulting energy profiles depicted in Figure 2, allows us to conclude that the global process is an exothermic reaction (-5.9 kcal/mol). Thus, the enzymatic environment stabilizes the final products when compared to gas-phase calculations.²⁸ From the kinetic point of view, the two mechanisms are competitive with activation energies of around 23 kcal/mol. It must be kept in mind that these results are obtained with a semiempirical description of the quantum region of our QM/MM model and thus nonnegligible errors could affect the predictions about the activation energies.

As can be seen in Tables 1 and 2, the crystal water molecules seem to remain in ordered positions during the reaction. In this regard (see the O_w-X columns), water 228 acts as a proton relay accepting a proton (H4) from O4 and transferring a proton (H5) to N5 (see Schemes 2 and 3). These events take place concertedly ($R \rightarrow I1$ step in path 1) or by steps ($I2 \rightarrow I3 \rightarrow P$ steps in path 2), but in all cases product P is hydrogen bonded, by the w228 molecule, to N5 and O4. The difference is that although w228 is acting as proton acceptor with respect to O4 and proton donor to N5 in R, we found the opposite situation; proton donor to O4 and proton acceptor from N5, in P. On the other hand, though w206 remains at H-bond distances from O4

TABLE 3: Selected Substrate–Enzyme Interatomic Distances (Å) for All the Stationary Point Structures Obtained from the QM/MM Calculations

	R	II	I2	I3	P
O(ILE5):H8(DHF)	1.86	1.87	1.84	1.86	1.81
O(w157):H91(DHF)	1.99	2.03	1.99	2.00	1.97
H61(THR113):OD2(ASP)	1.95	1.94	1.85	1.95	1.84
HD1(PHE31):N1(DHF)	2.69	2.68	2.70	2.70	2.71
HG12(ILE5):N8(DHF)	3.32	3.39	3.42	3.31	3.41
HB2(ALA7):N3(DHF)	2.71	2.73	2.78	2.72	2.86

would be required to reach the R structure. Therefore, both w206 and w228 play a direct role in the network of waters, which helps to the global transfer of a proton from Asp27 to the N5–C6 bond, as suggested previously by Cannon and co-workers.⁶ Nevertheless, a note of caution must be considered as these conclusions come from QM/MM potential energy calculations and not from dynamical exploration of a large number of structures. Calculation of free energies would also provide information related with entropic effects, which could be quite important for the proton relay mechanism.

With respect to the active site residues, as can be observed in Table 3, most of the distances between the protein and the substrate remain constant on going from R to P. These results could be interpreted as evidence of the small reorganization of the enzymatic surroundings during the chemical reaction. The fact that the structure of the enzyme does not change dramatically during the reaction has been also observed in other enzymatic systems,^{29,30} but usually associated with single-step processes. In other cases, when the enzyme catalyzes multistep chemical processes, the enzyme flexibility can be very important to adapt the active site to the reaction requirements. This is, for example, the case of β -lactamases where some residues can suffer important displacements, allowing them to play an active role even if they were initially located far away from the substrate.³¹ With respect to our system, molecular dynamics simulations performed on three ternary complexes from the DHFR catalytic cycle by Brooks and co-workers³² suggest the importance of correlated motions and structural perturbations. Although they observed that the protein maintained a core structure similar to the initial X-ray model, several loop conformational changes caused dramatic differences in the behavior of the protein dynamics. More recent calculations carried out by Hammes-Schiffer and co-workers,³³ based on hybrid quantum-classical molecular dynamics simulations and empirical valence bond (EVB) potentials,³⁴ for diverse wild-type DHFR species and mutants, suggest enzyme reorganization is significant for the following step; the hydride transfer reaction (the rate-limiting step). Motions of residues both in the active site and distal to the active site impact the free energy of activation and the degree of barrier reconfiguring for this chemical step.

The analysis of the normal-mode frequencies of the TSs yields relatively high imaginary frequencies (see Tables 1 and 2), associated with forming/breaking bonds. The dominant amplitudes of these vectors are related to the interatomic distances of donor and acceptor atoms in each step. Other normal modes, corresponding to the second, third, etc. eigenvalues are significant smaller (ca. 60 cm^{-1}), revealing quite soft movements. It can be considered that the active site helps the system in softening the vibrational degrees of freedom, as also shown in previous related works^{18,35} where comparative analysis of vacuum versus enzyme QM/MM studies were carried out.

In addition to the analysis of the geometrical aspects, it is also interesting to study the reaction from an electronic point

of view. For this purpose we have used the Mulliken charges summed for different subfragments of the reaction system³⁴ (the Asp27 residue, and the pyrazine and pyrimidine rings). Of course, the absolute value of these charges can be unreliable and other methods would be preferable, but the variation along a reaction path can give a reasonable picture of the electronic evolution of the process. According to this, the reaction can be viewed as a partial electron transfer from pyrazine and pyrimidine rings toward Asp27, thus preparing the system for the subsequent hydride transfer to C6. The calculations describe a neutral Asp27 in R, whereas its total charge in P is -0.96 au. The pyrazine ring follows the opposite trend: 0.03 au in R and 0.68 au in P. This is a very interesting result considering the fact that the C6, as mentioned above, is accepting a hydride anion in the following step. It must be here remembered that this hydride transfer is the rate-limiting step for the global process of DHF reduction to THF by DHFR. This electronic analysis also reveal the electrostatic stabilization of the protonated substrate by Asp 27, as pointed out by Callender et al. on the basis of their Raman studies.¹²

Conclusions

In this paper we present a theoretical study of the chemical reaction step prior to the hydride transfer in dihydrofolate reductase. The employed hybrid quantum-mechanical/molecular-mechanical (AM1/CHARMM24/TIP3P) technique has allowed us to include the protein environment and to explore structural features of significance on hypersurfaces spanning several thousand degrees of freedom. We have focused our interest in the stationary points on the PES, which have been located and characterized involving a fully flexible active-site region by means of GRACE software. Theoretical predictions have been compared with previous published experimental data. The analysis of the results can be summarized as follows.

We have explored two possible reaction paths for the activation of the substrate to proceed for the next step, the hydride transfer step. Both routes, which appear to be energetically competitive, explain the protonated state of the Asp27 and its role in stabilizing the cationic pyrazine ring. Furthermore, our calculations reveal the active role of the water molecules that have been detected in all DHFR crystal structures. The ordered structure of the X-ray crystallographic water molecules support the hypothesis of an indirect transfer of a proton from Asp27 residue to the N5 atom of the substrate. We have found that w206 serves to stabilize Asp27 whereas w228, together with the aforementioned w206, serve as a channel of hydrogen atoms from the Asp27 to the N5 of the pterin ring.

Acknowledgment. We are indebted to DGI for project DGI BQU2000-1425, BANCAIXA for project PIA99-03, and Generalitat Valenciana for project GV01-324, which supported this research, and the Serveis d'Informàtica of the Universitat de València and the Universitat Jaume I and for providing us with computer capabilities. S.F. thanks DGI for a Ministerio de Ciencia y Tecnología for a doctoral fellowship. E.S. thanks Prof. E. Longo for his hospitality and helpful discussions.

References and Notes

- (1) Blakley, R. L. *Folates and Pterins*; Wiley: New York, 1984.
- (2) Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4085.
- (3) Kraut, J.; Matthews, D. A. In *Biological Macromolecules and Assemblies*; Jurnak, F. A., McPherson, A., Eds.; Wiley: New York, 1987.

- (4) (a) Cummins, P. L.; Ramnarayan, K.; Singh, U. C.; Gready, J. E. *J. Am. Chem. Soc.* **1991**, *113*, 8247. (b) Cummins, P. L.; Greatbanks, S. O.; Rendell, A. P.; Gready, J. E. *J. Phys. Chem. B* **2002**, *106*, 9934.
- (5) Howell, E. E.; Villafranca, J. E.; Warren, M. S.; Oatley, S. J.; Kraut, J. *Science* **1986**, *231*, 1123.
- (6) Cannon, W. R.; Garrison, B. J.; Benkovic, J. *J. Am. Chem. Soc.* **1997**, *119*, 2386.
- (7) Howell, E. E.; Warren, M. S.; Booth, C. L. J.; Villafranca, J. E.; Kraut, J. *Biochemistry* **1987**, *26*, 8591.
- (8) Morrison, J. F.; Stone, S. R. *Biochemistry* **1988**, *27*, 5499.
- (9) Uchimaru, T.; Tsuzuki, S.; Tanabe, K.; Benkovic, J.; Furukawa, K.; Taita, K. *Biochem. Biophys. Res. Commun.* **1989**, *161*, 64.
- (10) Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263.
- (11) McTigue, M. A.; Davies, I. J. F.; Kaufman, B. T.; Kraut, J. *Biochemistry* **1992**, *31*, 7264.
- (12) Chen, Y.; Kraut, J.; Callender, R. *Biophys. J.* **1997**, *72*, 936.
- (13) Gready, J. E. *Biochemistry* **1985**, *24*, 4761.
- (14) Gready, J. E.; Cummins, P. L.; Wormell, P. *Adv. Exptl. Med. Biol.* **1993**, *338*, 487.
- (15) Ivery, M. T. G.; Gready, J. E. *Biochemistry* **1995**, *34*, 3724.
- (16) Jeong, S. S.; Gready, J. E. *Biochemistry* **1995**, *34*, 3734.
- (17) A good example of this kind of methodologies can be found in the following references: (a) Turner, A. J. Doctoral Thesis, University of Bath, 1997. (b) Moliner, V.; Turner, A. J.; Williams, I. H. *J. Chem. Soc. Chem. Commun.* **1997**, 1271. (c) Turner, A. J.; Moliner, V.; Williams, I. H. *J. Phys. Chem. Chem. Phys.* **1999**, *1*, 1323.
- (18) Castillo, R.; Andrés, J.; Moliner, V. *J. Am. Chem. Soc.* **1997**, *119*, 2386.
- (19) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
- (20) Barnes, J. A.; Williams, I. H. *J. Chem. Soc., Chem. Commun.* **1996**, 193.
- (21) Barnes, J. A.; Williams, I. H. *Biochem. Soc. Trans.* **1996**, *24*, 263.
- (22) Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586.
- (23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (24) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
- (25) Foresman, J. B.; Frisch, A. E. *Exploring Chemistry with Electronic Structure Methods*; Gaussian, Inc.: Pittsburgh, PA, 1993.
- (26) Fukui, K. *J. Phys. Chem.* **1970**, *74*, 4161.
- (27) García-Viloca, M.; Alhambra, C.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2001**, *114*, 9953.
- (28) AM1 gas-phase calculations of a model system (a molecule of acetic acid that mimics the Asp27, the 6-methyl-substituted pterin that represents the DHF and one water molecule) predict that the reaction is endothermic by about 28 kcal/mol.
- (29) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *Chem. Eur. J.* **2003**, *9*, 984–991.
- (30) Torres, R. A.; Schiott, B.; Bruce, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 8164–8173.
- (31) Castillo, R.; Silla, E.; Tuñón, I. *J. Am. Chem. Soc.* **2002**, *124*, 1809.
- (32) Radkiewicz, J. L.; Brooks, C. L., III. *J. Am. Chem. Soc.* **2000**, *122*, 225–231.
- (33) (a) Agarwal, P. K.; Billeter, S. R.; Rajagopalan, P. T. R.; Benkovic, S. J.; Hammes-Schiffer, S. *PNAS* **2002**, *99*, 2794–2799. (b) Agarwal, P. K.; Billeter, S. R.; Hammes-Schiffer, S. *J. Phys. Chem. B* **2002**, *106*, 3283–3293. (c) Watney, J. B.; Agarwal, P. K.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2003**, *125*, 3745–3750. (d) Rod T. H.; Radkiewicz, J. L.; Brooks, C. L., III. *Proc. Natl. Acad. Sci.* **2003**, *100*, 6980.
- (34) Warshel, A. *Computer Modelling of Chemical Reactions in Enzymes and Solutions*; John Wiley: New York, 1991.
- (35) Moliner, V.; Andrés, J.; Oliva, M.; Safont, V. S.; Tapia, O. *Theor. Chem. Acc.* **1999**, *101*, 228.
- (36) As pointed out by Alhambra and co-workers (Alhambra, C.; Wu, L.; Zhang, Z.; Gao, J. *J. Am. Chem. Soc.* **1998**, *120*, 3858) in the original implementation of the link-atom approach (see ref 24) electrostatic interactions between the link atom and the rest of the protein atoms were not included in the quantum calculation. This, however, introduces an imbalance in electrostatic interactions in the QM region due to the fact that molecular orbitals are delocalized and such a partition, which excludes some terms in the Fock matrix, results in an unrealistically large partial charges on link atoms and the atoms they are attached to. This deficiency, inherent in CHARMM link-atom treatment, has to be always kept in mind, and conclusions must be discussed with caution unless a correction is taken into account. In this work, because the quantum link atom charges are constant in all the states, no change in the polarization have been induced by the MM charges from R to P. This result means that a minimal artifact result is obtained by the presence of these virtual H atoms in the relative magnitudes, although the total charge on C6 can be polarized by the close presence of the negative charge in the link atom. Due to this reason, we have preferred to calculate and analyze the sum of the charges of the pyrazine ring, which renders a more realistic result. Furthermore, the Mulliken charge analysis approach has been shown to be highly basis set dependent. Nevertheless, considering the rest of the limitations related with the AM1/TIP3P optimization method implemented in CHARMM, the electronic analysis derived from these calculations can be considered qualitatively acceptable.

Silvia Ferrer, José Javier Ruiz-Pernia, Iñaki Tuñón, Vicent Moliner,
Mireia Garcia-Viloca, Àngels Gonzalez-Lafont, José Maria Lluch
J. Chem. Theory Comput., 2005, 1, 750-761

A QM/MM Exploration of the Potential Energy Surface of Pyruvate to Lactate Transformation Catalyzed by LDH. Improving the Accuracy of Semiempirical Descriptions

Silvia Ferrer,[†] J. Javier Ruiz-Pernía,[#] Iñaki Tuñón,^{*,#} Vicente Moliner,^{*,§}
Mireia Garcia-Viloca,[‡] Angels González-Lafont,^{‡,†} and José M. Lluch^{‡,†}

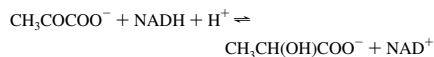
*Departament de Química Física/IcMol, Universitat de València, València, Spain,
Departament de Ciències Experimentals, Universitat Jaume I, Castelló, Spain, Institut
de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Spain, and
Departament de Química, Universitat Autònoma de Barcelona, Barcelona, Spain*

Received February 2, 2005

Abstract: We present a QM/MM study of the potential energy surface of the pyruvate to lactate transformation catalyzed by L-lactate dehydrogenase. The transformation involves a hydride and a proton transfer which are followed by means of the corresponding antisymmetric combination of the distances from the hydrogen atom to the donor and the acceptor atoms. To discriminate among the possible reaction mechanisms we have considered different improvements of the AM1/MM description: reoptimization of the van der Waals parameters and inclusion of corrections to the QM energy associated with both transfer coordinates. The QM subsystem has been also enlarged to include charge-transfer effects from the substrate to some specific residues. In our best treatment, the transformation is described as a concerted process through a single transition structure in which the hydride transfer is more advanced than the proton transfer. From the methodological point of view, the correction schemes tested here improve the quality of the semiempirical potential energy surface although they also present deficiencies attributed to consideration of the proton and hydride transfer corrections as separable ones.

1. Introduction

L-Lactate dehydrogenase (LDH) is a highly stereospecific metabolic enzyme which catalyzes the interconversion of pyruvate and L-lactate using the NADH/NAD⁺ pair as redox cofactor. Formally, the reaction in the pyruvate → L-lactate direction is achieved by reducing the carbonyl group with a hydride anion transferred from NADH and a proton transfer from a protonated His195 residue:¹



Much information on the mechanism has come from kinetic and site-directed mutagenesis experiments.² These studies characterize the enzyme (structure and details about the active site and mechanism) setting up that, while in the wild-type enzyme the rate-limiting step of both, the pyruvate reduction or the lactate oxidation process is a unimolecular rearrangement of the enzyme–NADH–pyruvate complex, and the rate of catalysis in the mutants is limited by the chemical reaction. In this regard, it is still unknown whether the reaction mechanism involves sequential proton and hydride ion transfers or if both transfers take place in a concerted way (see Scheme 1). The difference in the timing of the hydride transfer and the proton transfer is a matter of interest in mechanistic enzymology. Computational studies from different groups^{3–7} suggested that the sequential mechanism is more likely, but the proposals differ in the order in which the proton and hydride ions are transferred. Thus, Ranganathan and Gready^{6,7} found a mechanism in which the hydride transfer preceded proton transfer in a stepwise

* Corresponding authors e-mail: tunon@uv.es (I.T.) and moliner@exp.uji.es (V.M.).

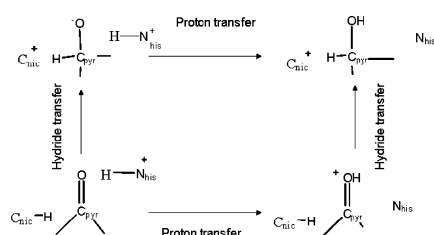
[†] Departament de Química, Universitat Autònoma de Barcelona.

[‡] Institut de Biotecnologia i de Biomedicina.

[§] Universitat Jaume I.

[#] Universitat de València.

Scheme 1



manner, in accordance with a previous study by Warshel et al. using an empirical valence bond study,⁸ but contrasting with the usual chemical and enzymatic arguments for hydride transfer processes and with other theoretical studies^{3,4,8} that predicted a concerted but asynchronous reaction in which the proton transfer was in a very advanced stage of the reaction. Finally, a more flexible quantum mechanical/molecular mechanical (QM/MM) treatment⁵ allowed tracing the two distinct reaction pathways across the energy hypersurface. Nevertheless, as it was pointed out in this study,⁵ the results at this point did not allow for making a definitive statement as to which mechanism was preferred. Without considering the fact that accurate determination of transition state properties would require statistical averaging over many configurations, each one individually being a transition state structure,^{5,9} the previous study presented a limitation since a semiempirical Hamiltonian was employed to describe the quantum region. This, together with the fact that both pathways presented similar energy barriers, prevented to answer the question of which mechanism was preferred. To get more reliable energetics thus allowing to answer this question, a higher level of theory should be used in the QM/MM calculations.

Because of the reasons mentioned above, this enzymatic reaction provides an excellent example to illustrate the need of improving the accuracy of the QM/MM methods where the QM region is described by means of semiempirical Hamiltonians. One of the final goals of a theoretical study is usually to obtain the free energy profile of the desired enzymatic reaction. This free energy profile can be compared to experimental data and is also a powerful tool to get a deeper insight into the different contributions to catalysis. Obviously, the quality of these calculations is determined by the accuracy of the potential energy surface employed to describe the process. In the current example the selection of a theoretical level can be decisive to favor one or another reaction path. The conclusions reached by a theoretical study will be relevant only if we can be confident on the quality of our theoretical method to describe with similar accuracy different possible reaction mechanisms. This purpose is obviously limited by the computational cost. High level quantum potential energy and gradient calculations for medium-sized systems interacting with a large classical environment are still prohibitive taking into account that optimizations in highly dimensional energy surfaces require hundreds or even thousands of steps. In this paper we test different correction schemes in order to improve semiem-

pirical descriptions at a low additional computational cost. These schemes are applied to discern between the two possible reaction mechanisms appearing on the semiempirical potential energy surface. We also address the other questions affecting the accuracy of the QM/MM calculations such as the optimization of the van der Waals parameters for the interaction between the QM and the MM subsystems and the size of the QM region. Considering all these aspects, and bearing in mind the limitations of the computational approach, the most favored reaction mechanism for the pyruvate to lactate transformation catalyzed by LDH in our treatment is a concerted but asynchronous mechanism where the hydride transfer is more advanced than the proton transfer at the transition state structure. Mechanistic details are dependent on the computational level chosen to describe the reaction.

2. Theory

In the present work, we have modeled the potential energy surface (PES) by a combined quantum-mechanical and molecular-mechanical (QM/MM) approach^{10,11} that includes the generalized hybrid orbital (GHO)¹² method to treat the boundary between the quantum-mechanical (QM) and the molecular-mechanical (MM) fragments of the system. We used the semiempirical Austin model 1 (AM1)¹³ to describe the QM subsystem. A semiempirical description can be quite inaccurate in some cases, and its use requires the previous calibration of the Hamiltonian. Obviously, it would be ideal to use a higher level *ab initio* molecular orbital or density functional theory (DFT) method to represent the reactive part of the system. However, these calculations are still too time-consuming to be practical at the present time for enzymatic systems considering that PES exploration, location, and characterization of stationary structures and molecular dynamics simulations (if needed) may require thousands of energy and gradient evaluations. An alternative dramatic reduction of the MM region size would not be realistic as the effect of long-range interactions as well as the flexibility of the system would be lost. There are several alternatives to improve the semiempirical description at a moderate or negligible computational cost. In principle one can develop specific reaction parameters (SRP)¹⁴ designed to reproduce higher level results or to include correction energy terms to the original PES. We have here employed two of this last kind of methods.

As explained before we are interested in a chemical process which is composed of two elementary chemical steps: a hydride transfer and a proton transfer. These hydrogen transfers are here described using as reaction coordinates the antisymmetric combination of the distances from the hydrogen atom to the donor (r_{DH}) and the acceptor (r_{AH}) atoms:

$$R = r_{DH} - r_{AH}$$

We will use the symbols R_1 for the hydride transfer and R_2 for the proton transfer (see below for details). We will consider in this work two independent corrections to the semiempirical electronic energy associated with each one of these two coordinates.

The two correction methods used in this work are carried out by introducing two energy terms, which are functions that depend exclusively on R_1 and R_2 reaction coordinates, respectively.

The first correction method is the Simple Valence Bond (SVB)^{15,16} In this method, the total potential energy of the system is given by the following equation

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{mm}} + E_{\text{qm/mm}} + E_{\text{SVB,R1}} + E_{\text{SVB,R2}} \quad (1)$$

where E_{qm} is the energy of the QM subsystem, E_{mm} is the energy of the MM subsystem, $E_{\text{qm/mm}}$ is the interaction energy between the QM and the MM regions, which include both electrostatic and van der Waals terms, and the E_{SVB} terms represent the SVB correction to the semiempirical E_{qm} energy in a particular reaction coordinate R_i . The following function has been proposed for this correction term

$$E_{\text{SVB,R}}(r_{\text{DH}}, r_{\text{DA}}, r_{\text{AH}}) = \frac{1}{2}[M_1(r_{\text{DH}}) + M_2(r_{\text{AH}})] - \frac{1}{2}[(M_1(r_{\text{DH}}) - M_2(r_{\text{AH}}))^2 + 4[V_{12}(r_{\text{DA}})]^2]^{1/2} + \Delta D_{\text{DH}} \quad (2)$$

where r_{DA} is the distance between the donor and the acceptor atoms, the function V_{12} is given by

$$V_{12}(r_{\text{DA}}) = D_{\text{DA}} \exp[-\alpha_{12}(r_{\text{DA}} - r_{\text{DA}}^0)] \quad (3)$$

and

$$M_i(r_{\text{XH}}) = \Delta D_{\text{XH}} (\exp[-2\alpha_{\text{XH}}(r_{\text{XH}} - r_{\text{XH}}^0)] - 2\exp[-\alpha_{\text{XH}}(r_{\text{XH}} - r_{\text{XH}}^0)]) \quad (4)$$

where XH represents the different pairs of atoms: donor-hydrogen (DH) or acceptor-hydrogen (AH) in the Morse potentials M_1 and M_2 , respectively, and r_{XH}^0 is the equilibrium distance of the corresponding bond. We set the two values ΔD_{XH} equal to the difference in dissociation energy between a high level theory calculation (in this work we use second-order Moller Plesset (MP2) level of calculation with a 6-31+G(d) basis set) and the lower level (the AM1 semiempirical method); α_{XH} is calculated as

$$\alpha_{\text{XH}} = \left(\frac{k_{\text{XH}}}{2D_{\text{XH}}^0} \right)^{1/2} = \nu_{\text{XH}} \left(\frac{2\pi^2\mu}{D_{\text{XH}}^0} \right)^{1/2} \quad (5)$$

where D_{XH}^0 is the bond energy, μ is the reduced mass of atoms X and H, and k is the force constant. The parameters D_{DA} and α_{12} in eq 3 are adjusted to obtain the desired barrier height, which is the barrier height calculated at the higher level in gas phase.

The second correction method is the interpolated corrections (IC) scheme based on the use of cubic splines under tension.¹⁷ The total potential energy is given by the following equation:

$$E_{\text{tot}} = E_{\text{qm}} + E_{\text{mm}} + E_{\text{qm/mm}} + \Delta E_{\text{IC}}(R_1) + \Delta E_{\text{IC}}(R_2) \quad (6)$$

This correction term is obtained as the difference between the qm energy provided by the high level method (HL) and the low level (LL) one for a particular configuration of the system obtained along the chosen reaction coordinate (R_i).

$$\Delta E(R_i) = E_{\text{QM}}^{\text{HL}}(R_i) - E_{\text{QM}}^{\text{LL}}(R_i) \quad (7)$$

Several structures are selected to calculate the energy difference between the high level and the low level methods at different values of the R_i coordinates. Then, following the work of Truhlar et al.,^{18,19} a spline under tension is used to interpolate this correction term at any value of R_i . In this way we obtain a continuous function in R_i , with continuous first and second derivatives, which are necessary to perform molecular dynamics simulations. To preserve the general applicability of the method the spline fit is carried out in terms of a mapping coordinate z , defined as

$$z_i = \frac{2}{\pi} \arctan\left(\frac{R_i - R_{i0}}{L}\right) \quad (8)$$

The new variable allows us to map the energy correction term onto the finite interval $[-1, +1]$. The constants R_{i0} and L are chosen for centering and scaling the mapping function in the range of interest.¹⁹ Then the final interpolated correction is obtained as

$$\Delta E_{\text{IC}}(R_i) = \text{spline}\{\Delta E[z_i(R_i)]\} \quad (9)$$

3. Calculations and Computational Details

3.1. Model of the Enzyme–Substrate–Coenzyme Complex. The X-ray structure of lactate dehydrogenase from *Bacillus stearothermophilus* comes from the PDB code 1LDN which is an octamer, although in this work only the tetramer is used (Figure 1). X-ray studies have proposed that the tetramer is the functional form of lactate dehydrogenase from *Bacillus stearothermophilus*.²⁰ In each monomer the crystal structure contains a total of 316 amino acid residues, the cofactor NADH, and the inhibitor oxamate (OXM), which is replaced by pyruvate in our study.

The coordinates of the hydrogen atoms of the protein and coenzyme were determined using the HBUILD facility of the CHARMM package.²¹ All the ionizable groups were set to their normal ionization state at pH 7 with the exception of His195, which was modeled in its protonated form. Other histidine residues of the protein were modeled as neutral with the proton at N δ .

Then, the system was partitioned into a quantum mechanical region consisting of 52 atoms and a molecular mechanical region containing the rest of the system. The QM subsystem includes 9 atoms of the pyruvate, 13 atoms of the His195 residue (including the C α as a boundary atom), and 30 atoms of the NADH, which include the dihydronicotinamide and ribose rings, and the C5' ribose atom as a boundary atom. The QM subsystem is described using the AM1 Hamiltonian and the boundary atoms are represented by the GHO method. A picture of the active site showing the partition into QM and MM subsystems is presented in Figure 2. First, the total energy of the system was minimized for 20 steps with the Adopted Basis Newton–Raphson (ABNR)²¹ method by moving only the QM part of the system. Then, the system was solvated with a 24 Å radius sphere of TIP3P²² water molecules centered on the pyruvate center of mass (see Figure 1). Water molecules that were within 2.5 Å of any non-hydrogen atom were removed. The

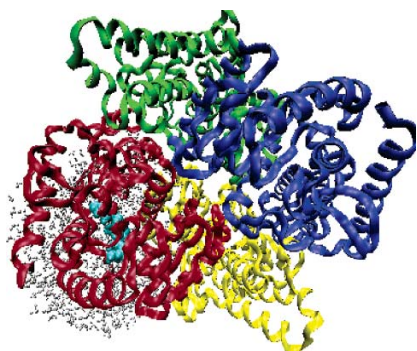


Figure 1. Tetramer of LDH with one of the active sites (NADH, pyruvate, and His-195) in blue balls. A sphere of water molecules with a radius of 24 Å centered on the center of mass of pyruvate.

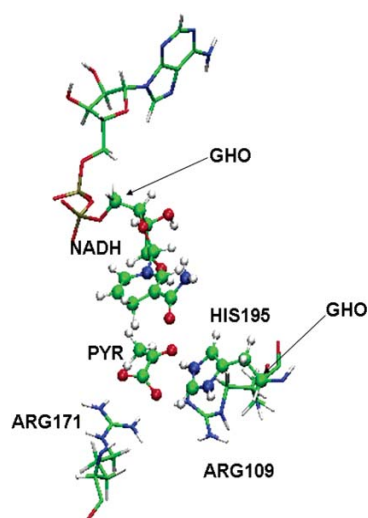


Figure 2. Snapshot of the LDH active center with the substrate (pyruvate) and the cofactor (NADH): The QM region is represented using balls and sticks. The position of the GHO atoms used to define the boundary with the MM region is also indicated.

resulting system was resolvated four more times using different relative orientations between the protein and the water sphere to ensure a good solvation of the system. Then, water positions were optimized for 20 steps followed by the optimization of the geometry of the part of the system included in a sphere of 24 Å. Finally, a molecular dynamics simulation of the waters (5 ps) was carried out to relax energetically unfavorable contacts, and the 3-fold cycle of superposition, deletion, and rotation was then repeated to fill in additional cavities generated during the dynamics

simulation. The final model has 22 159 atoms, 19 802 of them are atoms from the protein or the ligands, and the rest, 2277, are atoms from the water molecules.

3.2. Reaction Coordinates. As explained before two reaction paths have been proposed in previous studies⁵ (see Scheme 1). The first one is a mechanism where the hydride transfer precedes the proton transfer. This mechanism has been described as stepwise, where the initial step is the hydride transfer from a carbon atom of dihydronicotinamide, C_{nic} (donor atom), to the carbonyl carbon atom of pyruvate, C_{pyr} (acceptor atom), and in a second step a proton is transferred, from the N atom of His195, N_{his} (donor atom), to the carbonyl oxygen atom of pyruvate, O_{pyr} (the acceptor atom). The second proposed mechanism is a concerted but asynchronous mechanism where the proton and the hydride transfer take place through a single transition state (TS), being as how the proton transfer is now more advanced than the hydride transfer. For clarity purposes we will denote these two mechanisms as the hydride plus proton transfers pathway (the HP Pathway) and the proton plus hydride transfer pathway (the PH Pathway), respectively. Both mechanisms can be described on a single PES obtained as a function of two reaction coordinates (see Scheme 1). The reaction coordinate R_1 for the hydride transfer is defined in this work as the difference in the distances of the bonds between the transferring hydride-ion and the donor (C_{nic}) and the acceptor (C_{pyr}) atoms (eq 10). For the proton transfer a reaction coordinate R_2 is defined as the difference in the distance of the bonds between the transferring proton and the donor (N_{his}) and the acceptor (O_{pyr}) atoms (eq 11)

$$R_1 = r_{C_{\text{nic}}H_1} - r_{C_{\text{pyr}}H_1} \quad (10)$$

$$R_2 = r_{N_{\text{his}}H_2} - r_{O_{\text{pyr}}H_2} \quad (11)$$

The exploration of the PES was then carried out using the R_1 and R_2 coordinates defined before. These coordinates seem a reasonable choice to define a reduced PES, as they change smoothly along the surface. However, it must be taken into account that reduced surfaces give approximate descriptions of chemical processes. The CHARMM program was employed to carry out this exploration by means of the use of the RESDISTANCE keyword to define the reaction coordinates (R_i).

3.3. Exploration of the Potential Energy Surfaces. In this work, we have calculated two-dimensional (2D) potential energy surfaces (PES) at four different levels of theory: the uncorrected AM1/MM, the AM1-SVB/MM, the AM1-IC/MM, and the 2D-MP2sp/MM (bidimensional single-point calculations at the MP2 level). Gradient and energy calculations have been performed for the first three methods, whereas the 2D-MP2sp/MM is the result of single-point energy calculations on the AM1/MM optimized structures.

During all the optimizations, those atoms 24 Å away from the active site, 14 996 atoms, were kept frozen in order to reduce the computational cost, while 7143 atoms were allowed to move. In all the cases, the total energy of the system was minimized with the ABNR²¹ method until the norm of the gradient with R_1 and R_2 directions projected out was less than $0.001 \text{ kcal mol}^{-1} \cdot \text{Å}^{-1}$.

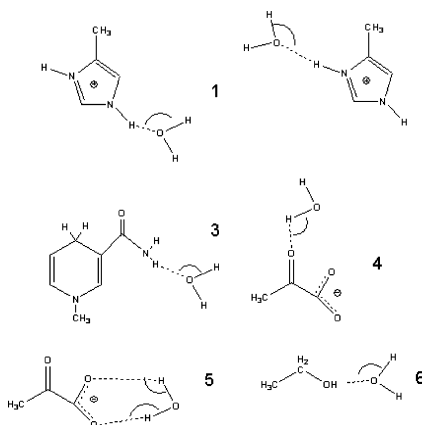


Figure 3. Bimolecular complexes of histidine fragment, nicotinamine ring model, pyruvate, and ethanol with water molecules.

3.4. QM/MM van der Waals Parameters. Prior to the PES exploration we have recalibrated the QM/MM potential by adjusting the van der Waals parameters of the QM atoms, employed to evaluate the interaction with the MM region. With this purpose we have obtained the interaction energy between some molecules that mimic the QM molecules and TIP3P water molecules by the procedure described by Freindorf and Gao.²³ We have considered complexes of water with models of the histidine and nicotinamide rings, with pyruvate and with ethanol to optimize the van der Waals parameters of the nitrogen atoms, hydroxyl oxygens, polar hydrogens bonded to nitrogen or oxygen, and also the carbon and oxygen atoms of pyruvate. Figure 3 shows the bimolecular complexes used to calculate the intermolecular interaction energy. The calculations were carried out describing all the complexes using quantum mechanics, at the HF/6-31G(d) level of theory and also using a hybrid QM/MM scheme where water molecules were described using TIP3P and the rest of the complex at the AM1 level. In all cases we have frozen the internal degrees of freedom of the QM fragments at gas-phase optimized values at the corresponding QM level, while the water intramolecular parameters have been fixed to the experimental geometry. To reproduce the HF/6-31G(d) results we have only modified the Lennard-Jones parameters, ϵ_{ab} and σ_{ab} , of selected QM atoms in the QM/MM interaction Hamiltonian $H_{qm/mm}$. As starting values for the Lennard-Jones parameters we used those provided by Gao and co-workers.²⁴ Table 1 provides the total interaction energies of the complexes at the HF level and also using QM/MM calculations with the van der Waals parameters obtained from the CHARMM force field, from Gao and co-workers, or after our optimization procedure. As it can be seen, in Table 1 we improved the energetic description of the complexes established with the histidine fragment and ethanol changing the parameters associated with nitrogen, oxygen, H(N), and H(O) atoms. However, we were not able to reproduce the full QM interaction with the

Table 1. Total Interaction Energies (in kcal/mol) for the Complexes Appearing in Figure 3^a

complex no.	$\Delta E(\text{HF})$	$\Delta E(1)$	$\Delta E(2)$	$\Delta E(3)$
1	-16.02	-12.92	-13.9	-15.82
2	-16.18	-13.03	-14.03	-15.98
3	-5.24	-5.03	-5.07	-5.62
4	-10.81	-14.74	-14.84	-14.84
5	-13.02	-17.63	-17.68	-17.68
6	-5.47	-3.17	-3.80	-3.89

^a Each column shows the results obtained at the HF/6-31G(d) level of theory and at the AM1/MM level using the van der Waals parameters of the CHARMM force field (1), those provided in ref 24 (2), and the optimized parameters (3) given in Table 2.

carbon and oxygen atoms of pyruvate following this scheme. The reason for this behavior could be related to the noninclusion of charge transfer effects in our QM/MM Hamiltonian between pyruvate and water molecules. The magnitude of the charge transfer term is probably too big to be accounted by fitting the van der Waals parameters. The importance of the charge-transfer effect in pyruvate will be further addressed below. The original CHARMM parameters, those of ref 24, and the optimized ones are listed in Table 2. Geometries and absolute energies at the HF/6-31G(d) level of the complexes are given as Supporting Information.

3.5. Evaluation of Corrections to the Potential Energy.

To improve the performance of the AM1 model through the use of the SVB or IC schemes, we carried out gas-phase ab initio calculations for smaller model reactions (Figure 4): In the selected model, 42 atoms constitute the three species that are involved in the two chemical processes: the hydride and the proton transfer. In a previous study³⁻⁵ of the same reaction, where we used a smaller model for the enzyme-substrate system (LDH monomer), the stationary points were localized on a AM1/MM potential energy surface. Here, we have used the stationary points of the HP Pathway to cut the Cartesian coordinates for the 42 atoms of the gas-phase model (Figure 4). The Cartesian coordinates for additional

Table 2. van der Waals Parameters for the Quantum Atoms^a

atom	this work		Gao et al. ^a		CHARMM	
	σ	ϵ	σ	ϵ	σ	ϵ
N(H)N(H2) ^b	1.00	-0.14	1.57	-0.15	1.85	-0.20
H(N) ^b	0.35	-0.08	0.45	-0.10	0.2245	-0.046
H(O) ^b	0.28	-0.10	0.45	-0.10	0.2245	-0.046
O(H) ^b	1.08	-0.08	1.65	-0.20	1.70	-0.12
C(=O) ^b	1.65	-0.20	1.65	-0.20	1.70	-0.12
C(H) ^b	1.96	-0.08	1.96	-0.08	1.99	-0.07
O(=C) ^b	1.65	-0.20	1.65	-0.20	1.70	-0.12
O ^c	1.65	-0.20	1.65	-0.20	1.70	-0.12
H ^d	1.12	-0.01	1.12	-0.01	1.32	-0.022

^a Reference 24. ^b A(B) stands for atoms A bound to atom or group B. ^c Oxygen of a carboxylate group. ^d On C. ^e The parameters are given according to the criteria of the CHARMM force field in Å and kcal/mol.

structures between have been obtained from AM1/MM optimizations at different values of the R_1 and R_2 coordinates on the LDH monomer model. Then, we have carried out single-point calculations on the 42 atoms gas-phase model at the AM1 level and by including electron correlation at the second-order Møller–Plesset perturbation theory with a polarized split valence basis set (MP2/6-31G(d,p)). The GAUSSIAN98²⁵ and CHARMM²¹ programs were used for these calculations. Table 3 gives the gas-phase relative energies calculated at the AM1 and MP2 levels. From comparison between semiempirical and MP2 values a non-negligible error can be expected when obtaining the PES of the enzymatic process using the AM1 Hamiltonian. In particular, the use of these Hamiltonian results in a single-point recalculated reaction energy about 26 kcal/mol higher for the hydride transfer and a single-point recalculated reaction energy about 18 kcal/mol lower for the proton transfer. It also produces an important error in the hydride transfer barrier (the difference in the single-point recalculated barriers is 16.9 kcal/mol). Thus, consideration of correction schemes seems to be necessary to improve the accuracy of our computational model. The parameters determined for the SVB correction terms (see eqs 1–5) using these values are shown in Table 4.

For the IC method, we used the same gas-phase results on the 42 atoms model (Figure 4) to determine the two correction terms as a function of R_1 and R_2 , respectively. For a given value of R_1 or R_2 the correction term is the difference between the single-point potential energy obtained using high level (MP2) and low level (AM1) methods. Two independent cubic splines are used to fit the correction energies as a function of the reaction coordinates (R_i).

It must be pointed out that according to the procedure described the structures selected to parametrize both correction schemes are obtained from the same region of the PES (if we divide the PES presented in Scheme 1 with a diagonal traced from reactants to products, the selected structures have been taken exclusively from the upper half part), and thus the correction procedures are expected to work more properly in this region. This important limitation will be discussed in more detail later on.

4. Results

4.1. Potential Energy Surfaces. Figure 5 shows the PES obtained using R_1 and R_2 coordinates at the AM1/MM level without corrections. The PES shows two distinct mechanistic pathways with very similar features to those found in a previous work on the LDH monomer using link atoms instead of GHO.⁵ The hydride plus proton transfer pathway (HP Pathway) is a stepwise mechanism involving first the hydride transfer from the dihydronicotinamide ring of the NADH to a carbon atom of pyruvate followed by a proton transfer from the protonated His195 to the carbonyl O atom of pyruvate. On the other hand, the proton plus hydride transfer pathway (PH Pathway) is a concerted mechanism in which the proton transfer is considerably more advanced than the hydride transfer in the transition state (TS3). Table 5 gives the most relevant bond distances and the relative energies of the stationary points located on the AM1/MM PES. From the values of the distances given in Table 5 it is evident that TS1 corresponds to the hydride transfer where the proton transfer has been only slightly advanced. TS2 is the transition state structure associated with the proton transfer once the hydride has been already transferred. TS3 is the transition state structure associated with a concerted but very asynchronous proton and hydride transfers, where the proton transfer is considerably more advanced than the hydride transfer. The energy barriers for the HP Pathway are about 52 kcal/mol from reactants (pyruvate) to the intermediate and 14 kcal/mol from the intermediate to products (lactate), while the energy barrier in the PH Pathway is about 41 kcal/mol. Thus, on the AM1/MM PES the PH Pathway is the favored mechanism but taking into account the expected errors of the AM1 description, these results are still inconclusive about the preferred mechanism for the pyruvate to lactate transformation.

Figure 6 shows the PES calculated with the inclusion of the two SVB correction terms. As it can be seen it is possible to define again the two same mechanisms with similar features. Thus, the HP Pathway is a stepwise mechanism presenting an intermediate after the hydride transfer, and the PH Pathway is a concerted but asynchronous process. The differences with respect to the uncorrected AM1/MM PES appear in the relative energies of the stationary structures. In the HP Pathway the energy barriers are of about 29 kcal/mol from reactants (pyruvate) to intermediate (TS1) and about 11 kcal/mol from intermediate to products (TS2)—see Table 6. In the PH Pathway the energy barrier is again about 41 kcal/mol. We can stress at least two important conclusions at this point. First, the energy barriers now found seem to be more reasonable for an enzymatic process as they are considerably lower than in the uncorrected exploration. Second, while the PH Pathway presents a lower energy barrier on the uncorrected AM1/MM PES, now the prevalence of the HP Pathway is quite clear, because it presents noticeably lower energy barriers. The preference for this mechanism was clearly established by means of PES explorations carried out using a symmetric combination of the R_1 and R_2 coordinates. Starting from both, the reactants or the products, and using this new distinguished reaction coordinate, the chemical transformation always took place

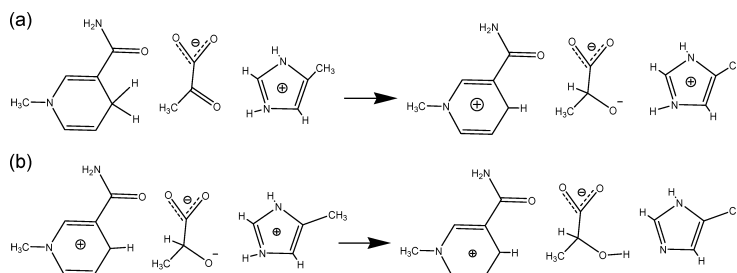


Figure 4. Molecular models used to calibrate the performance of the AM1 method versus MP2/6-31G(d,p) calculations: (a) model used for the hydride transfer reaction: from reactants (NADH and pyruvate) to intermediate (NAD⁺ and PyrH⁻) and (b) model used for the proton-transfer reaction: intermediate (PyrH⁻ plus protonated His195) to products (lactate and His195).

Table 3. Single-Point (SP) Recalculated Gas-Phase Reaction Energies and Barrier Heights (kcal/mol) for the Hydride-Transfer Reaction and the Proton-Transfer Reaction in the Molecular Model Presented in Figure 4^a

	SP recalculated barrier	SP recalculated reaction energy
Hydride Transfer		
AM1	76.27	93.11
MP2/6-31G(d,p)	59.38	66.96
Proton Transfer		
AM1	-5.07	-61.57
MP2/6-31G(d,p)	-7.07	-43.14

^a Recall that the geometries are taken from AM1/MM calculations.

Table 4. Parameters of the SVB Functions for the Hydride (a) and Proton Transfer Steps (b)^a

(a)					
ΔD_{DH}	α_{DH}	ΔD_{AH}	α_{AH}	D_{DA}	α_{DA}
0.01	1.36	26.15	6.43	1.5	0.50
(b)					
ΔD_{NH}	α_{NH}	ΔD_{OH}	α_{OH}	D_{NO}	α_{NO}
21.27	1.49	1.46	1.09	7.1	1.00

^a ΔD s are given in kcal/mol and α in \AA^{-1} .

through the HP Pathway. The energy profiles obtained using the R_1+R_2 coordinate are provided as Supporting Information.

Figure 7 shows the PES calculated with the inclusion of ICs through the use of two independent cubic splines associated with the R_1 and R_2 coordinates. The obtained PES is very similar to the previous one. In fact, by inspection of the PES it is possible to define again the two same reaction mechanisms. Thus, the HP Pathway is a stepwise mechanism and the PH Pathway is a concerted one. The energy barrier for the hydride transfer in the HP Pathway is of about 34 kcal/mol—see Table 7—while for the subsequent proton transfer is of about 11 kcal/mol (from the intermediate). In the PH Pathway the energy barrier is about 52 kcal/mol. Therefore, also within this correction methodology the preferred mechanism is the HP Pathway. Differences with respect to the previous methodology are found in the geometry of the stationary structures. Thus, reactants are now

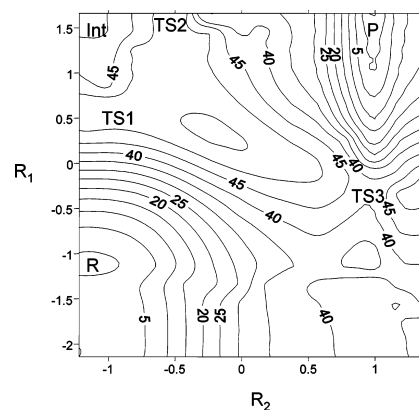


Figure 5. Potential energy surface at the AM1/MM level. R_1 and R_2 are the coordinates (in \AA) associated with the hydride and proton transfers (see eqs 10 and 11). The position of reactants (R), intermediate (Int), products (P), and transition state structures (TS1, TS2, and TS3) is qualitatively shown. Isoenergetic lines are depicted each 5 kcal/mol.

Table 5. Relevant Distances (\AA) and Relative Energies (kcal/mol) of the Stationary Structures Appearing on the AM1/MM PES

	$r_{C_{\text{nic}}H_1}$	$r_{C_{\text{pyr}}H_1}$	$r_{N_{\text{nic}}H_2}$	$r_{O_{\text{pyr}}H_2}$	energy
reactants	1.15	2.29	1.00	2.22	0
TS1	1.70	1.24	1.01	2.03	52
intermediate	2.80	1.15	1.00	2.22	38
TS2	2.79	1.14	1.06	1.68	52
products	2.79	1.13	1.98	0.99	-5
TS3	1.30	1.44	1.96	0.99	41

found at larger distances from pyruvate atoms to the hydride and the proton. However, these changes are not very relevant in energy as the reactants are found in a very flat valley, and then its exact position can change noticeably depending on the particularities of the energy function employed. Geometrical changes are also found in the structure of TS3 with respect to the SVB surface. On this new PES the TS3

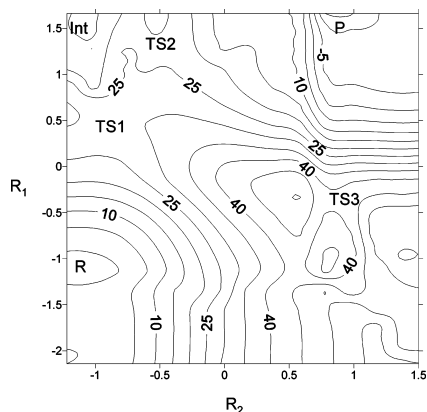


Figure 6. Potential energy surface at the AM1-SVB/MM level. R_1 and R_2 are the coordinates (in Å) associated with the hydride and proton transfers, respectively (see eqs 10 and 11). The position of reactants (R), intermediate (Int), products (P), and transition state structures (TS1, TS2, and TS3) is qualitatively shown. Isoenergetic lines are depicted each 5 kcal/mol.

Table 6. Relevant Distances (Å) and Relative Energies (kcal/mol) of the Stationary Structures Appearing on the PES Obtained at the AM1/MM Level with the Inclusion of SVB Correction Terms

	$r_{C_{\text{int}}H_1}$	$r_{C_{\text{pyr}}H_1}$	$r_{N_{\text{his}}H_2}$	$r_{O_{\text{pyr}}H_2}$	energy
reactants	1.14	2.29	1.01	2.03	0
TS1	1.62	1.16	1.00	2.02	29
intermediate	2.79	1.13	1.01	2.03	14
TS2	2.79	1.13	1.06	1.67	25
products	2.79	1.13	1.96	0.98	-15
TS3	1.21	1.55	1.77	0.99	41

structure appears at slightly less advanced values of the hydride and proton-transfer coordinates.

At this point it is important to stress the fact that the two correction schemes display very similar behavior because they have been parametrized to reproduce the same high level quantum method energies using structures taken from the same region of the uncorrected PES. As described above, the reference structures used to obtain the correction energies were selected from the HP Pathway of an uncorrected PES. Thus, it would be reasonable to question the ability of the two methods to give accurate predictions of the PES in the region corresponding to the PH Pathway. For this reason we decided to obtain a new corrected PES based on bidimensional energy single-point calculations at the MP2/6-31G(d,p) level using the gas-phase model in Figure 4 and taking the geometries from the AM1/MM-2D PES (Figure 5). This new PES was obtained using the following energy function

$$E = E_{\text{AM1}} + E_{\text{AM1/MM}} + E_{\text{MM}} + (E'_{\text{MP2}} - E'_{\text{AM1}}) \quad (12)$$

where E' are the result of single-point energy calculations

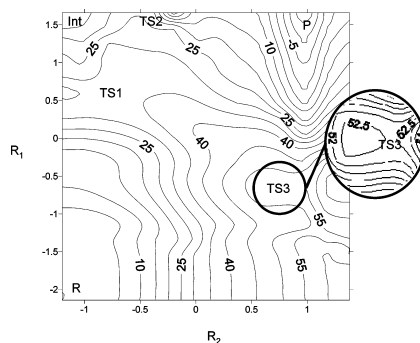


Figure 7. Potential energy surface at the AM1-IC/MM level. R_1 and R_2 are the coordinates (in Å) associated with the hydride and proton transfers (see text). The position of reactants (R), intermediate (Int), products (P), and transition state structures (TS1, TS2, and TS3) is qualitatively shown. Isoenergetic lines are depicted each 5 kcal/mol (0.5 kcal/mol in amplified zone).

Table 7. Relevant Distances (Å) and Relative Energies (kcal/mol) of the Stationary Structures Appearing on the PES Obtained at the AM1/MM Level with the Inclusion of Interpolated Corrections

	$r_{C_{\text{int}}H_1}$	$r_{C_{\text{pyr}}H_1}$	$r_{N_{\text{his}}H_2}$	$r_{O_{\text{pyr}}H_2}$	energy
reactants	1.13	3.27	1.00	2.22	0
TS1	1.86	1.20	1.03	1.85	34
intermediate	2.80	1.15	1.00	2.22	13
TS2	2.79	1.13	1.11	1.53	24
products	2.79	1.13	1.98	0.99	-19
TS3	1.20	1.75	1.98	1.00	52

in gas phase obtained for the reduced QM model corresponding to the molecules presented in Figure 4. This new PES is presented in Figure 8, and the coordinates and relative energies of the stationary structures are gathered in Table 8. There are significant differences with the previous PESs. The most evident is that the intermediate appearing after the hydride transfer to the pyruvate has disappeared and a new intermediate (Int2) is found after the proton transfer. Thus, on the new PES the HP Pathway is a concerted mechanism taking place through a single transition state structure (TS4). This new transition state structure roughly corresponds to TS1 found on the previous PESs although it is now located at more advanced values of the proton-transfer coordinate. The PH Pathway is now described as a stepwise mechanism with two transition states (TS5 and TS6). The first corresponds to a very advanced proton transfer and the second to an early hydride transfer to the protonated intermediate (Int2). The energy of this structure is of about 37 kcal/mol above the reactants. At the reactant structure the carbon atom of pyruvate is found at a larger distance from the hydride (as in the PES with interpolated corrections). The products are now found at much larger values of the proton-transfer coordinate than in any of the discussed PES. This is due to the larger distance between histidine to the transferred proton. From the kinetic point of view, the concerted HP Pathway

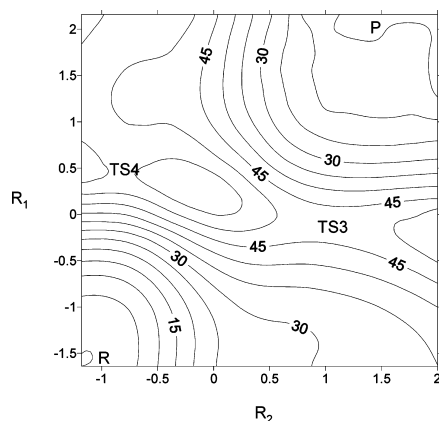


Figure 9. Potential energy surface obtained at the AM1/MM level with a QM region including Arg109 and Arg171. R_1 and R_2 are the coordinates (in Å) associated with the hydride and proton transfers (see text). The position of reactants (R), products (P), and transition state structures (TS3 and TS4) is qualitatively shown. Isoenergetic lines are depicted each 5 kcal/mol.

Table 10. Relevant Distances (Å) and Relative Energies (kcal/mol) of the Stationary Structures Appearing on the AM1/MM PES Obtained When the QM Region Includes Arg109 and Arg171

	$r_{C_{\alpha}H_1}$	$r_{C_{\beta}H_1}$	$r_{N_{\alpha}H_2}$	$r_{O_{\beta}H_2}$	energy
reactants	1.13	2.77	1.01	2.17	0
TS4	1.78	1.22	1.03	1.81	54
TS3	1.36	1.39	2.01	0.99	47
products	3.08	1.12	2.40	0.98	13

2D uncorrected PES using the same reaction coordinates but increasing now the size of the QM region to include both Arg109 and Arg171 (resulting in a total of 78 QM atoms). This requires the introduction of two new GHOs placed at C γ of each arginine. The new PES obtained for the larger QM subsystem is shown in Figure 9. As previously, we found two different mechanisms for the transformation of pyruvate into lactate: the HP Pathway and the PH Pathway in which the proton transfer precedes the hydride transfer. On this PES we only located two transition state structures (TS4 in the HP Pathway and TS3 in the PH Pathway), so both mechanisms are now described as concerted ones. The geometries and energies of the stationary structures are given in Table 10. The energy barrier in the HP Pathway (TS4) is about 54 kcal/mol and in the PH Pathway (TS3) is of about 47 kcal/mol. This description can be compared with the uncorrected AM1/MM PES with the smaller (52 atoms) QM subsystem. In that case the HP Pathway presented an energy barrier of about 52 kcal/mol associated with the first step (the hydride transfer), while the PH Pathway had a unique transition state structure with an associated energy barrier of about 41 kcal/mol (see Table 5). The reaction energy also changes noticeably, being of about -5 kcal/mol with the smaller QM

subsystem and about 13 kcal/mol with the larger one. Then, the main effect of the inclusion of the arginines into the QM subsystem is to destabilize the intermediate which appeared after the hydride transfer, increasing the energy of that region of the PES. As a consequence the reaction path associated with the HP Pathway is now considerably more concerted than in our previous PES. All these changes can be rationalized considering that when Arg109 and Arg171 are described classically, electrostatic interactions with the intermediate obtained after the hydride transfer and with the protonated His195 are overestimated. This results in too large a stabilization of the intermediate, with a charge of -2, and also in too large an interaction with the protonated His195, with a positive charge. When these arginines are described quantum mechanically and the charge can be more delocalized, this overestimation of electrostatic interactions disappears, and thus the intermediate is no longer a minimum on the PES. This interpretation is confirmed by analysis of the charge distribution on the QM subsystem when it contains 52 or 78 atoms. For the intermediate, the absolute charge on the substrate is about 0.1 electrons smaller when the QM region includes both arginines. The magnitude of this charge transfer is confirmed by analysis of Mulliken and NPA charges^{26,27} (a representation of the atomic charges for the stationary structures obtained at different levels is given as Supporting Information).

Because of the magnitude of the changes taking place when the size of the QM region was increased we decided to obtain a new PES for the large QM region including corrections to the AM1 description. According to the conclusions reached in the previous section we decided to include corrections by means of bidimensional single-point energy calculations at the MP2/6-31G(d,p) level, following the same procedure as before (see eq 12). The corrected PES is presented in Figure 10. The most relevant distances and the energies of the stationary structures corresponding to this PES are presented in Table 11. In this case we have an important difference with respect to the previously presented PESs: only one reaction mechanism is found. The unique transition state structure corresponds to a concerted hydride and proton transfers where the hydride transfer is considerably more advanced than the proton transfer, i.e., we have a concerted HP Pathway. The energy barrier is about 30 kcal/mol, and the reaction energy is close to zero (-0.7 kcal/mol).

5. Conclusions

We have here presented a QM/MM study of the PES corresponding to the pyruvate to lactate transformation catalyzed by LDH. This transformation takes place by means of a hydride and a proton transfer to the substrate. The order in which the hydride and the proton are transferred defines two different mechanistic routes for the process: a hydride transfer followed by a proton transfer (the so-called HP Pathway) and a proton transfer followed by a hydride transfer (the PH Pathway). Previous theoretical studies differ in the preferred mechanism as well as in their concerted or stepwise nature. For this reason we decided to improve the standard AM1/MM description of the reaction. In particular we have

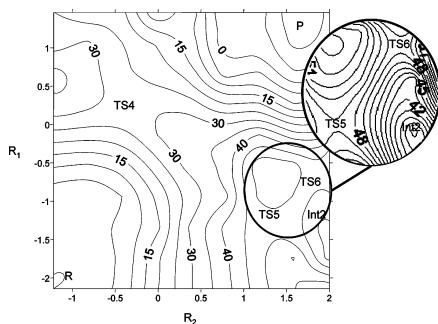


Figure 8. Potential energy surface obtained with bidimensional single-point MP2/6-31G(d,p) energy corrections. R_1 and R_2 are the coordinates (in Å) associated with the hydride and proton transfers (see text). The position of reactants (R), products (P), intermediate (Int2), and transition state structures (TS4, TS5, and TS6) is qualitatively shown. Isoenergetic lines are depicted each 5 kcal/mol (1 kcal/mol in amplified zone).

Table 8. Relevant Distances (Å) and Relative Energies (kcal/mol) of the Stationary Structures Appearing on the PES Obtained with Bidimensional MP2/6-31G(d,p) Single-Point Energy Corrections

	$r_{C_{\text{int}}H_1}$	$r_{C_{\text{py}}H_1}$	$r_{N_{\text{int}}H_2}$	$r_{O_{\text{py}}H_2}$	energy
reactants	1.13	3.27	1.01	2.22	0
TS4	1.53	1.28	1.10	1.53	29
TS5	1.14	2.49	2.39	1.01	49
Int2	1.15	2.50	2.98	1.01	37
TS6	1.20	1.75	2.77	1.00	48
products	2.59	1.13	2.56	0.98	-13

seems to be the preferred mechanism on the PES (29 versus 48 kcal/mol).

It is interesting to comment on the performance of both correction methods when compared to the PES obtained with single-point corrections. Both the SVB and the IC schemes assume that differences between the high and low computational levels (MP2 and AM1 in this case) in the R_1 and R_2 coordinates are mutually independent. This is, the errors associated with the semiempirical description are expected to be approximately the same when the proton is transferred to the substrate (with charge -1 au) or to the intermediate (with a charge of -2 au). This is obviously not the case. However, despite the inherent deficiency both correction methods considerably improve the energetic description of both mechanisms when compared to the uncorrected AM1/MM PES. As compared to the PES obtained with bidimensional MP2 corrections, the AM1 surface has an average error in barrier heights of 15 kcal/mol and an average error in the four key geometrical variables at the two transition states of 0.32 Å. The SVB correction reduces these average errors to 4 kcal/mol and 0.31 Å, and the IC correction reduces them to 4 kcal/mol and 0.25 Å, respectively. It seems that separable correction schemes do a good job correcting the energies but not as well for correcting geometries of dynamical bottlenecks. This feature was already reported for the IC

Table 9. Interaction Energies (E_{int} , in kcal/mol) and Mulliken Charge on the Pyruvate (Q , in au) for the Complexes Formed between Pyruvate and Arg171 and Pyruvate and Arg109^a

	Arg171/Pyr		Arg109/Pyr	
	E_{int}	Q	E_{int}	Q
HF/6-31G(d)	-111.21	-0.868	-68.76	-0.966
AM1	-98.37	-0.920	-62.80	-0.991
AM1/MM	-77.76	-1.000	-40.18	-1.000

^a Geometries are taken from the reactant structure of the AM1/MM PES.

methodology,¹⁸ and it seems that the SVB one presents a similar behavior. Anyway, the most important question is that both schemes are successfully predicting the HP Pathway as the preferred one, while on the uncorrected PES the PH Pathway displayed a lower energy barrier. The SVB method gives energy barriers in better agreement with the PES with bidimensional corrections for the HP Pathway, while the IC method agrees better with the bidimensional corrected PES in the PH Pathway.

4.2. The Size of the QM Subsystem. The interaction between the substrate (pyruvate) and some residues in the MM part obviously play an essential role defining the possible reaction mechanisms. It has been proposed that the role of the Arg109 is to enhance the polarization of the pyruvate carbonyl group in the ground state and to stabilize the transition state in the HP Pathway, whereas the role of Arg171 is to assist in the binding of the substrate.^{1,2} As previously described, when we tried to optimize the van der Waals parameters of the QM/MM interaction between pyruvate and water molecules we found a systematic deviation with respect to the QM values. This difference could be due to charge transfer effects, a term not included in standard QM/MM calculations. Obviously, this possible source of error requires some further studies in our system because the charge transfer could be even larger in the case of pyruvate surrounded by positively charged arginine residues. To address this issue we have compared the interaction energy between pyruvate and both arginines in the enzyme active site by means of gas-phase single-point energy calculations at the HF/6-31G(d) level, with those obtained at the AM1 and AM1/MM levels. These interaction energies are given in Table 9. As it can be seen, the AM1/MM calculations show important deviations from the HF values. The main contribution to this error is not the use of the AM1 Hamiltonian as the results are much improved when all the system is described by means of this semiempirical method. The magnitude of the error is directly correlated with the charge transfer (estimated from the Mulliken charges), and thus we attribute this error to the lack of charge-transfer effects in the QM/MM treatment.

Obviously, if the error is kept more or less constant all along the PES, then one could ignore this contribution. The charge of the substrate is the same in reactants (pyruvate) and products (lactate). However, it must be taken into account that after the hydride transfer the charge of the substrate increases from -1 to -2 , and then the charge-transfer effect on the energy could be different in that region of the PES. For this reason we decided to explore again the AM1/MM

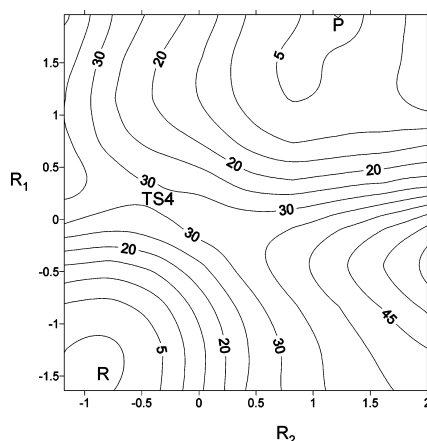


Figure 10. Potential energy surface obtained at the AM1/MM level with a QM region including Arg109 and Arg171 and bidimensional single-point MP2/6-31G(d,p) corrections. R_1 and R_2 are the coordinates (in Å) associated with the hydride and proton transfers (see eq 10). The position of reactants (R), products (P), and transition state structure (TS4) is qualitatively shown. Isoenergetic lines are depicted each 5 kcal/mol.

Table 11. Relevant Distances (Å) and Relative Energies (kcal/mol) of the Stationary Structures Appearing on the PES with Bidimensional MP2/6-31G(d,p) Single-Point Energy Corrections Obtained when the QM Region Includes Arg109 and Arg171

	$r_{C_{\alpha}H_1}$	$r_{C_{\gamma}H_1}$	$r_{N_{\text{His}}H_2}$	$r_{O_{\text{pyr}}H_2}$	energy
reactants	1.13	2.77	1.01	2.17	0
TS4	1.62	1.26	1.12	1.50	30
products	3.06	1.12	2.20	0.95	-1

considered the effect of the van der Waals parameters used in the QM subsystem, the correction of the QM energy by two different schemes (SVB and IC), and the size of the QM subsystem to take into account charge-transfer effects.

While the change in the van der Waals parameters is quite modest, the inclusion of correction terms evaluated at the MP2 level into the AM1 energy has important consequences on the studied PES. The PH Pathway presents a lower energy barrier on the uncorrected AM1/MM PES, but when both correction schemes are considered the HP Pathway presents noticeably lower energy barriers. However, both correction schemes display some deficiencies associated with the fact that the correction energies associated with the proton and the hydride transfer coordinates are taken as independent. The most important consequence is that both methods describe the HP Pathway as a stepwise mechanism, while it appears as a concerted one when corrections with single-point calculations on the whole PES at the MP2 level are carried out. This feature of the PES is stressed when the QM subsystem is enlarged to include also Arg109 and

Arg171. A classical description of these two residues results in an overstabilization of the structure appearing after the hydride transfer to the substrate. Note that the terms stepwise and concerted refer here just to the presence or absence, respectively, of an intermediate on the PES. Indeed, only a dynamical study could show the actual relevance of the intermediate for the reaction mechanism.

Thus, in our best treatment, the pyruvate to lactate transformation catalyzed by LDH is described as a concerted hydride and proton transfer mechanism with a single transition structure where the hydride transfer is considerably more advanced than the proton transfer. However, the mechanistic details depend on the particular computational level chosen to describe the process. The energy barrier found is of about 30 kcal/mol, a value that seems to be overestimated when compared to the experimental estimations of the activation free energy barrier (with an upper limit of about 15 kcal/mol²⁸). Entropic contributions, tunneling, higher QM level, and a better QM/MM description are expected to contribute to diminish the gap with the experimental estimations. The protonation state of some residues close to the active site, here assumed to be found in their standard protonation states at pH 7, could also be a source of discrepancy with experiments. Studies are underway to evaluate all these contributions. In any case, and as a methodological conclusion of the present work, dynamical treatments of this system should be carried out on a PES including high level corrections to the QM energy for the proton and the hydride transfers, and, importantly, these corrections should not be included independently in order to have an accurate description.

Acknowledgment. Authors thank Prof. D. G. Truhlar for reading and commenting the manuscript. We thank the DGI for projects DGI BQU2003-04168-C03 and BQU-2002-00301, BANCAIXA for project PIA99-03, and Generalitat Valenciana for project GV01-324, GV04B-21, GV04B-131, and GRUPOS04/28. S.F. and J.J.R.-P. acknowledge a doctoral fellowship of the Ministerio de Ciencia y Tecnología. S.F. also acknowledges the Ministerio for financial support during her visit to the UAB and the warm hospitality of all the people of the Departament de Química of the Universitat Autònoma de Barcelona.

Supporting Information Available: Geometries (in Å) and absolute energies at the HF/6-31G(d,p) level (in au) of the complexes shown in Figure 3, energy profiles (kcal/mol) obtained using the $R_1 + R_2 = R_4$ coordinate, and atomic charges (in au) obtained for stationary structures of the HP pathway. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Clarke, A. R.; Atkinson, T.; Holbrook, J. J. *Trends Biochem. Sci.* **1989**, *14*, 145–148.
- (2) Burger, J. W. I.; Ray, W. J. *J. Biochemistry* **1984**, *23*, 3620–3626.
- (3) Moliner, V.; Turner, A. J.; Williams, I. H. *Chem. Commun.* **1997**, 1271–1272.

- (4) Moliner, V.; Williams, I. H. *Chem. Commun.* **2000**, 1843–1844.
- (5) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323–1331.
- (6) Ranganathan, S.; Gready, J. E. *J. Phys. Chem. B* **1997**, *101*, 5614–5618.
- (7) Ranganathan, S.; Gready, J. E. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 2047–1456.
- (8) Yadav, A.; Jackson, R. M.; Holbrook, J. J.; Warshel, A. *J. Am. Chem. Soc.* **1991**, *113*, 4800–4805.
- (9) Truhlar, D. G.; Gao, J. L.; Garcia-Viloca, M.; Alhambra, C.; Corchado, J.; Sanchez, M. L.; Poulsen, T. D. *Int. J. Quantum Chem.* **2004**, *100*, 1136–1152.
- (10) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
- (11) Gao, J.; Thompson, M. A. *Combined quantum mechanical and molecular mechanical methods*; Washington, DC, 1988.
- (12) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*(24), 4714–4721.
- (13) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (14) Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618–4627.
- (15) Devi-kesavan, L. S.; Garcia-Viloca, M.; Gao, J. *J. Theor. Chem. Acc.* **2003**, *109*, 133–139.
- (16) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2003**, *107*, 9567–9578.
- (17) Ruiz-Pernia, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. *J. Phys. Chem. B* **2004**, *108*, 8427–8433.
- (18) Chuang, Y.; Corchado, J. C.; Truhlar, D. *J. Phys. Chem. A* **1999**, *103*, 1140–1149.
- (19) Corchado, J. C.; Coitiño, E. L.; Chuang, Y.; Fast, P. L.; Truhlar, D. *J. Phys. Chem. A* **1998**, *102*, 2424–2438.
- (20) Wigley, D. B.; Gamblin, S. J.; Turkenburg, J. P.; Dodson, E. J.; Piontek, K.; Muirhead, H.; Holbrook, J. J. *J. Mol. Biol.* **1992**, *223*, 317–335.
- (21) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (22) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (23) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386–395.
- (24) Gao, J.; Xia, X. *Science* **1992**, *258*, 631–635.
- (25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Peterson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Pittsburgh, PA, 1995.
- (26) Reed, A. E.; Curtis, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899–926.
- (27) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
- (28) Clarke, A. R.; Wigley, D. B.; Chia, W. N.; Barstow, D.; Atkinson, T.; Holbrook, J. *J. Nature* **1986**, *324*, 699–702.

CT050016L

Silvia Ferrer, Estanislao Silla, Iñaki Tuñón, Mónica Oliva, Vicent Moliner,
Ian H. Williams
Chem. Comm., 47, 5873-5875

Dependence of enzyme reaction mechanism on protonation state of titratable residues and QM level description: lactate dehydrogenase

Silvia Ferrer,^a Estanislao Silla,^a Iñaki Tuñón,^{*a} Mónica Oliva,^b Vicent Moliner^{*b} and Ian H. Williams^c

Received (in Cambridge, UK) 28th July 2005, Accepted 28th September 2005

First published as an Advance Article on the web 14th October 2005

DOI: 10.1039/b510735k

We have studied the dependence of the chemical reaction mechanism of L-lactate dehydrogenase (LDH) on the protonation state of titratable residues and on the level of the quantum mechanical (QM) description by means of hybrid quantum-mechanical/molecular-mechanical (QM/MM) methods; this methodology has allowed clarification of the timing of the hydride transfer and proton transfer components that hitherto had not been possible to state definitively.

It is observed macroscopically that proteins become unstable at extreme values of pH, but properties of enzymes are also pH dependent because their efficacy depends on the protonation state of their ionisable residues. The most expensive computational simulation may give meaningless results if an erroneous charge is selected for a single amino acid of the protein. Thus any theoretical analysis of protein structure or reactivity requires an adequate titration of all ionisable residues. Assignment of the protonation states of such residues is often based on the pK_a of the corresponding amino acids in aqueous solution. However, this practice may introduce significant artefacts into simulations because standard pK_a values of ionisable groups are shifted by local protein environments.¹

LDH is a highly stereospecific metabolic enzyme that catalyses the interconversion of pyruvate and lactate using the NADH/NAD⁺ pair as redox cofactor. This chemical step involves a hydride transfer (HT) from the dihydronicotinamide ring of NADH to the carbonyl C atom of pyruvate and a proton transfer (PT) to the carbonyl O atom from a N atom of the protonated His195 residue (Scheme 1).

Although the rate-limiting step of the full catalytic process in the wild-type enzyme is a conformational change involving closure of a surface loop (99–110) down over the active site,² nonetheless the chemical step catalyzed by LDH has been extensively studied both

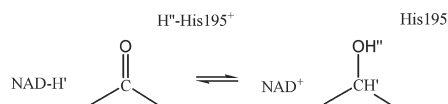
experimentally³ and theoretically.^{4–8} Although the relative timing of the HT and PT components is a matter of interest in mechanistic enzymology, it is not yet possible to state definitively which mechanism is preferred.⁸ While a stepwise mechanism is more likely than a concerted one, results from various computational groups differ in the order in which the HT and PT steps occur.^{4–8}

Initially the oxamate inhibitor in the ternary complex X-ray crystal structure⁹ of LDH tetramer from *B. stearothermophilus* with NADH cofactor was replaced by pyruvate. To assign protonation states for acidic and basic residues, two alternative strategies were considered. In the first, all ionisable groups were set at a state complementary to pH 7 using the pK_a values of the amino acids in aqueous solution, with the exception of His195, which was modelled in its doubly protonated form.

The second protocol involved recalculation of the pK_a values using the "cluster method",¹ as implemented by Field and co-workers,¹⁰ according to which each titratable residue in the protein is perturbed by the electrostatic effect of the protein environment. After adding hydrogen atoms, the entire system was hydrated with a preformed 24 Å radius sphere of TIP3P water molecules centred on the pyruvate center of mass. Atoms 24 Å away from this centre of mass were kept frozen during the simulations. The resulting model had 22139 atoms of enzyme, co-factor, substrate and solvent, 14996 of them kept frozen. The entire chemical system was divided into a QM region, treated by the AM1 semiempirical MO method, and a MM region comprising the rest of the protein (CHARMM26 potentials)¹¹ and the water molecules. The generalized hybrid orbital (GHO)¹² method was used to treat those covalent bonds crossing the boundary between the QM and the MM regions. The QM region, formed by the dihydronicotinamide and the ribose ring of the NADH, Arg171, Arg109, the pyruvate and the His195 contained a total of 78 atoms. The full system was relaxed by means of a combination of minimizations and molecular dynamic simulations.

The potential energy surfaces (PESs) were obtained using two antisymmetric combinations of distances defining the HT and PT coordinates. Fig. 1 shows the PES obtained when the ionisable residues of the enzyme were titrated using the standard pK_a values in aqueous solution. The resulting competitive mechanistic pathways across this energy hypersurface, as well as the geometries of the stationary points and the barrier heights (see first two columns of Table 1) are similar to the results of our previous study.⁸

When the pK_a values of the ionisable residues were recalculated under the effect of the protein environment, as explained above, the protonation state of several ionisable groups at pH = 7 was found to have changed. In particular, Glu199 and His82 became



Scheme 1

^aDepartamento de Química Física, Universidad de Valencia, 46100, Burjassot, Spain. E-mail: ignacio.tunon@uv.es; Fax: 34963544564; Tel: 34963544880

^bDepartament de Ciències Experimentals, Universitat Jaume I, Box 224, 12080, Castellón, Spain. E-mail: moliner@exp.uji.es; Fax: +34964728066; Tel: 34964728084

^cDepartment of Chemistry, University of Bath, Bath, UK BA2 7AY

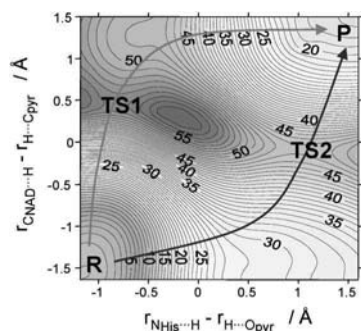


Fig. 1 PES 1: QM/MM contours (kcal mol^{-1}) for pyruvate (R) to lactate (P) LDH-catalyzed reaction. The protonation state of the ionizable groups in protein were calculated at pH = 7 using the standard $\text{p}K_a$ values in aqueous solution.

Table 1 QM/MM selected interatomic distances (\AA), and relative energies (kcal mol^{-1}) for the optimized geometries of the transition states of the reaction paths obtained at the two different protonation states of the ionizable groups of the protein

	PES 1		PES 2		PES 3
	TS1	TS2	TS1	TS2	TS1
$\text{C}_{\text{NAD}} \cdots \text{H}^+$	1.78	1.36	1.69	1.27	1.43
$\text{H}^+ \cdots \text{C}_{\text{pyr}}$	1.22	1.39	1.21	1.47	1.31
$\text{N}_{\text{His195}} \cdots \text{H}^+$	1.03	2.01	1.04	1.92	1.09
$\text{H}^+ \cdots \text{O}_{\text{pyr}}$	1.81	0.99	1.74	0.99	1.57
ΔH^\ddagger	54	47	44	33	21

protonated, thereby changing their charges to neutral and positive, respectively.¹³ The effect of this change in only two of the residues on the resulting PES is drastic, as shown in Fig. 2 and in the values listed in the third and fourth columns of Table 1. The main effect is

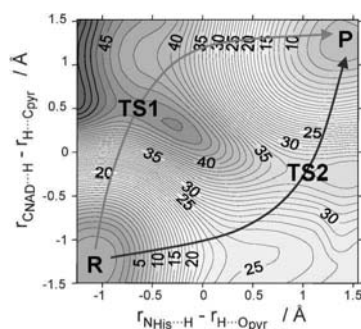


Fig. 2 PES 2: QM/MM contours (kcal mol^{-1}) for pyruvate (R) to lactate (P) LDH-catalyzed reaction. Protonation states of ionizable groups in protein were calculated following the method of Gilson *et al.*¹

to reduce the activation enthalpy by *ca.* 10 and 14 kcal mol^{-1} for TS1 and TS2 paths, respectively. This diminution in the reaction barriers is mainly a consequence of the added Glu199 proton close to Asp168, a residue that interacts directly to maintain His195 in a protonated state in preparation for substrate binding and turnover. Excess negative charge in this vicinity stabilizes the protonated imidazole of His195 in the reactant complex more than in the transition state or product, a situation exacerbated by negative charge on Glu199. The previously described ultra-simple electrostatic model, comprising only the substrate carbonyl group with the hydride and proton donors and suitably-located point-charges,⁴ also predicted reduction in the AM1 calculated reaction barrier by *ca.* 10 kcal mol^{-1} , due to annulment of a negative charge in the vicinity of Asp168.¹⁴

The last phase of this study was motivated by the need to improve the description of the QM region beyond the present AM1 semiempirical level. In view of the size of our system a high theoretical level is still prohibitive, so we have performed single point calculations at the MP2/6-31G(d,p) level on the same grid of AM1/MM optimized structures used to obtain the PES displayed on Fig. 2.¹⁵ The results (Fig. 3 and last column of Table 1) are impressive: the asynchronous concerted reaction pathway in which PT precedes HT disappears and the effective enthalpy barrier for the reaction mechanism now appears to be much more in accordance with the expected value for this biological reaction.¹⁶

We demonstrate in this work that reaction mechanisms and barrier heights of an enzymatic reaction can be wrong if protonation states for the titratable residues of the protein are selected on the basis of the $\text{p}K_a$ of the corresponding amino acids in aqueous solution. From our results, obtained at high level MP2//AM1/MM, we have arrived at some conclusions about the timing of the hydride transfer and proton transfer components in the chemical step catalyzed by LDH that had not been possible to state definitively up to now.

Finally, accurate determination of transition-state properties would require statistical averaging over many configurations, each one individually being a transition structure⁸ and, for this particular reaction, tunnel effects may also contribute to lowering the observed free energy barrier.

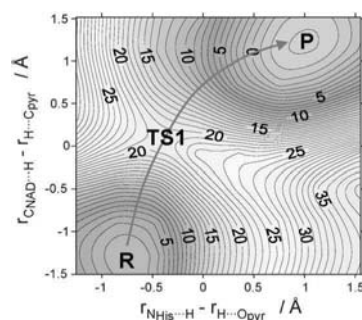


Fig. 3 PES 3: QM/MM contours (kcal mol^{-1}) for pyruvate (R) to lactate (P) LDH-catalyzed reaction, with MP2/6-31G(d,p) corrections. Protonation states of ionizable groups in protein are as for PES 2.

We thank the Royal Society for a Joint Project Grant, DGI for project BQU2003-04168-C03, BANCAIXA for project PIA99-03, Generalitat Valenciana for GV01-324, GV04B-21, GV04B-131 and GRUPOS04/28. S. F. acknowledges a doctoral fellowship of the Ministerio de Ciencia y Tecnología. We are also grateful to Dr M. Field and Dr P. Amara for supporting us in the pK_a calculations and discussions.

Notes and references

- 1 See the following references, where the method is explained and its accuracy compared with that obtained with other methods: (a) M. K. Gilson, *Proteins: Struct., Funct., Genet.*, 1993, **15**, 266–282; (b) J. Antosiewicz, J. A. McCammon and M. K. Gilson, *J. Mol. Biol.*, 1994, **238**, 415–436.
- 2 C. R. Dunn, H. M. Wilks, D. J. Halsall, T. Atkinson, A. R. Clarke, H. Muirhead and J. J. Holbrook, *Philos. Trans. R. Soc. London, Ser. B.*, 1991, **332**, 177–184.
- 3 See: P. Kedzierski, K. Moreton, A. R. Clarke and J. J. Holbrook, *Biochemistry*, 2001, **40**, 7247–7252 and references therein.
- 4 J. Wilkie and I. H. Williams, *J. Am. Chem. Soc.*, 1992, **114**, 5423–5425.
- 5 A. Yadav, R. M. Jackson, J. J. Holbrook and A. Warshel, *J. Am. Chem. Soc.*, 1991, **113**, 4800–4805.
- 6 S. Ranganathan and J. E. Gready, *J. Phys. Chem. B*, 1997, **101**, 5614–5618.
- 7 V. Moliner, A. J. Turner and I. H. Williams, *Chem. Commun.*, 1997, 1271–1272.
- 8 V. Moliner and I. H. Williams, *Chem. Commun.*, 2000, 1843–1844.
- 9 D. B. Wigley, S. J. Gamblin, J. P. Turkenburg, E. J. Dodson, K. Piontek, H. Muirhead and J. J. Holbrook, *J. Mol. Biol.*, 1992, **223**, 217–225.
- 10 M. J. Field, L. David and D. Rinaldo, personal communication.
- 11 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. J. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.
- 12 J. Gao, P. Amara, C. Alhambra and M. J. Field, *J. Phys. Chem. A*, 1998, **102**, 4714–4721.
- 13 pK_a calculations also showed changes in the protonated nitrogen (from N δ to N ϵ) in some neutral histidines (157, 188, 266, 318).
- 14 J. Wilkie and I. H. Williams, unpublished work.
- 15 For a detailed description of the computational method see: S. Ferrer, J. J. Ruiz-Pernia, I. Tunon, V. Moliner, M. Garcia-Viloca, A. Gonzalez-Lafont and J. M. Lluch, *J. Chem. Theor. Comput.*, 2005, **1**, 750–761.
- 16 An experimental k_{cat} of 250 s $^{-1}$, measured by Clarke *et al.* (ref. 17) is equivalent to a free energy barrier of 15 kcal mol $^{-1}$ which, considering that the chemical step is not the rate limiting step of LDH catalyzed reaction, would be the upper limit.
- 17 A. R. Clarke, D. B. Wigley, W. N. Chia, D. Barstow, T. Atkinson and J. J. Holbrook, *Nature*, 1986, **324**, 699–702.

Theoretical site-directed mutagenesis on LDH

Silvia Ferrer[‡], Estanislao Silla^{‡,§}, Iñaki Tuñón^{‡,}, Vicent Moliner^{§,*}
and Ian H. Williams[#]*

[‡]Departamento de Química Física; Universidad de Valencia, 46100 Burjassot,
Valencia,,Spain.

[§]Departament de Ciències Experimentals, Universitat Jaume I, 12071 Castelló, Spain.

[#]Department of Chemistry, University of Bath, BA2 7AY Bath, United Kingdom.

AUTHOR EMAIL ADDRESS: tunon@uv.es, moliner@exp.uji.es

ABSTRACT: The reduction of pyruvate to lactate catalyzed by the L-Lactate dehydrogenase has been studied in this paper by means of hybrid Quantum Mechanical / Molecular Mechanical simulations. A very flexible molecular model consisting on the full tetramer of the enzyme, together with the cofactor NADH, the substrate and solvent water molecules has allowed to theoretically mimic site directed mutagenesis studies, most of them previously experimentally performed. The potential energy surfaces obtained for every single mutation, compared with the one corresponding to the native enzyme, have been used to trace the possible reaction pathways and to locate and characterize the structures corresponding to the stationary points. The analysis of the results has been a very powerful tool to conclude about the role of key residues on the vacuole formed in the active site of the enzyme; Arg109, Asn140, Asp168, Glu199 and water molecules. Our results are in very good agreement with the previous conclusions derived from site directed mutagenesis, which has allowed to propose a never practice double mutation on Arg109 and Asn140 that, as our calculations suggest, would render an inactive enzyme. This double mutation is then an efficient way of inhibiting the pyruvate reduction by the LDH. This strategy can be extrapolated to other enzyme systems thus opening the door of a very promising methodology that, in combination with the appropriate experimental technique, will enable us to describe enzyme catalysis phenomenon and the particular role of the residues that form the protein. This knowledge placed us in a privileged position to modify the activity of enzymes and to propose efficient inhibitors.

MANUSCRIPT TEXT

L-Lactate dehydrogenase (LDH) is a highly stereospecific metabolic enzyme which catalyses the interconversion of pyruvate and L-lactate using the NADH/NAD⁺ pair as redox cofactor. Enzymes like LDH, involved in metabolic pathways, have been the subject of extensive research both experimentally and theoretically: while their properties are now relatively well-known, details of their reaction mechanisms are still not completely known.^[1]

The chemical step of LDH, in the pyruvate to L-lactate direction, involves a hydride transfer from the dihydronicotinamide ring of NADH to the carbonyl C atom of pyruvate and a proton transfer to the carbonyl O atom from a protonated histidine residue (the His195 residue).^[2]



The whole catalytic process goes through several stages: substrate binding, unimolecular rearrangement consisting on a closure of a surface loop (99-110) down over the active site, pyruvate reduction, reverse protein conformational change and products release.^[1] As shown by Dunn et al.^[3] and by Waldman et al.,^[4] closing of the 99-110 loop enforces isolation of the active site center from bulk solvent which is necessary for the reaction to proceed at a noticeable rate. In the wild-type enzyme this conformational rearrangement is in fact the rate limiting step.^[3]

During the last decades, the advent of protein engineering has enabled us to define the role of specific residues on the active site by substituting them by other aminoacids presenting differences in charge or/and size.^[5-10] Site-directed mutagenesis has been proved to be a good tool not only to change the rate limiting step of the full process, but also to demonstrate the feasibility of applying rational protein engineering to optimize and control the stereoselectivity of this particular enzyme when used in asymmetric synthesis.^[10]

In this sense, one of the most relevant experiments on the ionised residues that are part of the vacuole formed by the 99-110 loop was that carried out by Holbrook and Ingram,^[5] who prepared one of the first variants of LDH by site-directed mutagenesis, in which the histidine at position 195 was removed, thus demonstrating its role as the proton donor/acceptor in the

reaction. The influence of Asp168 on the reaction was evaluated by substituting the residue to Asn and Ala.^[6] This study supported the hypothesis that the Asp168 contributes to catalytic efficiency by raising the pK_a of His195 thus favouring the double protonation state of this histidine and by anchoring it to the most favourable orientation for enhancing its acid/base catalysis proton donating and accepting properties. It has been shown that LDHs cannot form the enzyme-cofactor-substrate ternary complex unless the imidazole group of His195 is protonated. The effect of replacing the negative charged Asp168 residue by neutral ones is to destabilize this ternary complex and, in contrast to the wild-type enzyme, the rate of catalysis becomes limited by the hydride transfer step.^[6]

Another important mutation was the Arg171Lys substitution that affected the catalytic properties of the enzyme by weakening the substrate binding.^[7]

The Arg109Gln mutation demonstrated the role of this arginine in polarizing the pyruvate carbonyl group in the ground state and stabilizing the transition state.^[8] This polarisation on the carbon-oxygen bond implies the development of a positive charge on the C2 thus facilitating the hydride transfer.

Holbrook and co-workers studied the importance of the electrostatic charge balance in the active site vacuole of NAD⁺-L-lactate dehydrogenase by introducing an aspartic acid at position 140 instead of the wild-type asparagine.^[9] They demonstrated that Asn140Asp mutant enzyme has the same k_{cat} as the native one at low pH, but at higher pH the K_m of the pyruvate increases. Experimental results and molecular modelling showed that the presence of an anion on the position of the Asn140 raises the pK_a of His195 to greater than 10, thus facilitating the opening of the loop and the entry of solvent molecules to the vacuole.^[9]

Recent mutations on other residues as Ala245Lys resulted in a qualitative change in the reaction mechanism by increasing the rigidity of the loop, thus provoking a relative acceleration of the rearrangement of the enzyme-NADH-pyruvate complex making the hydride transfer as the first slow step.^[1]

The chemical reaction catalyzed by the LDH has been also extensively studied by several theoretical methods, either in gas phase^[11, 12] and more recently by means of hybrid quantum mechanical / molecular mechanics methods (QM/MM).^[13-17] Nevertheless, although the sequential mechanism is more likely, different computational groups' proposals differ in the order in which the proton and hydride ions are transferred. Thus, some found a mechanism in which the hydride transfer preceded proton transfer in a stepwise manner,^[12, 13] while other studies^[14, 15, 18] predicted a concerted but asynchronous reaction in which the proton transfer was in a very advanced stage of the reaction. A very flexible QM/MM treatment^[17] found a megadimensional energy hypersurface that allowed tracing across the two distinct mechanism pathways. The difference in the timing of the hydride transfer and the proton transfer is a matter of interest in mechanistic enzymology but, as it was pointed out in this latter study, unfortunately the results at this point did not allow to make a definitive statement as to which mechanism is preferred. Finally, we have recently demonstrated the dependence of the chemical reaction mechanism of LDH on the protonation state of the ionisable residues of the protein and on the level of theory employed to describe the quantum region on a QM/MM.^[19] Corrections at MP2/6-31G(d,p) level of the semiempirical AM1 description QM region rendered a PES that describes an asynchronous concerted reaction pathway in which HT precedes PT with a remarkable reduction on the effective enthalpy barrier, much more in accordance with the expected value for this biological reaction.^[19]

In this paper we carry out an exhaustive analysis of the LDH catalyzed reaction by obtaining different potential energy surfaces of the native LDH as well as some mutated enzymes. Our results are compared with experimental studies already carried out in this system in order to test our methodology and to explain the predicted role of the ionized residues presented in the active site of the enzyme. Finally, we suggest new mutations that could be experimentally carried out which would definitely corroborate the proposed mechanism and the particular role of key residues on the enzyme active site.

COMPUTATIONAL METHODS

The QM/MM energy hypersurfaces were obtained using the CHARMM29b1 program.^[20] The starting geometry comes from the 2.5 Å resolution ternary complex X-ray crystal structure of the LDH tetramer from *Bacillus stearothermophilus* with NADH cofactor and oxamate inhibitor.^[21] Once the oxamate was replaced by pyruvate, CHARMM29b1 was also used to add hydrogens to all titratable residues at a state complementary to pH 7. Due to the fact that standard pK_a values of ionisable groups can be shifted by local protein environments,^[19] an accurate assignment of the protonation states of all these residues was carried out by recalculating the standard pK_a values of the aminoacids using the “cluster method”^[22, 23] as implemented by Field and co-workers.^[24] According to this method, each titratable residue in the protein is perturbed by the electrostatic effect of the protein environment. As demonstrated in our previous study on LDH,^[19] a non-accurate protonation state of only a few residues of the protein can render erroneous PESs and thus erroneous reaction mechanisms or/and activation barriers. After having properly added the hydrogen atoms, the system was placed in a cavity deleted from a pre-formed 24 Å radius spheres of TIP3P^[25] water molecules centred on the acceptor carbon and acceptor oxygen atoms. Water molecules within 2.5 Å of any non-hydrogen atoms of the protein, cofactor or substrate were removed. This procedure was repeated 3 times by randomly rotating the water spheres to solvate potential cavities of a single water configuration. Then after a short molecular dynamic simulation of 5 ps to relax unfavourable contacts, the solvation process was repeated to fill in additional cavities generated during the equilibration dynamics.^[18, 26] Once the entire system was built, those atoms 24 Å away from the centre of carbonyl C and O atoms of the pyruvate were kept frozen during the subsequent dynamics and optimizations. The resulting model was a pseudo-sphere of 22141 atoms of enzyme, co-factor, substrate and solvent, 15154 of them kept frozen. The entire chemical system was divided into a QM region, treated by the AM1 semiempirical MO method,^[27] and a MM region comprising the rest of the protein (CHARMM24 potentials) and the water molecules. Generalized hybrid orbital (GHO)^[28, 29] method has been used to treat the covalent bonds crossing the boundary between the QM and the MM regions, in order to satisfy the valence of the QM fragments. The QM region, formed by the dihydronicotinamide and the ribose ring of the NADH, the pyruvate and the His195 contained a total of 52 atoms. A picture of the active site showing the partition in the QM and MM subsystems is presented in Figure 1, where the QM region is shown as the shaded portion. The different mutations were carried out by substituting the original residue of the native structure by the new residue. Afterwards, the new tetramer was solvated by water molecules following the same simulation protocol as in the native LDH described above. Once the full system was relaxed, the same conditions were applied to all mutated and native LDH systems, slightly varying the number of MM atoms from one model to other, depending on the aminoacid substitution.

The PESs were obtained using two antisymmetric combinations of the distances between the hydrogen atoms with their donor (r_{DH}) and acceptor (r_{AH}) atoms. The reaction coordinate R_1 for the hydride transfer is defined as the difference in the distances of the bonds between the transferring hydride-ion and the donor and the acceptor atoms (Equation 2). For the proton transfer a reaction coordinate R_2 is defined as the difference in the distance of the bonds between the transferring hydrogen atom and the donor (N_{his}) and the acceptor (O_{pyr}) atoms (Equation 3)

$$R_1 = r_{C_{nic}H_1} - r_{C_{pyr}H_1} \quad \text{Equation 2}$$

$$R_2 = r_{N_{his}H_2} - r_{O_{pyr}H_2} \quad \text{Equation 3}$$

The CHARMM29b1 program was employed to carry out the grid scanning. Once the PESs were obtained, localised approximate transition state (TS) structures were refined using the recently modified version of the GRACE software.^[16] A partial-rational-function-

operator/adopted-basis-Newton-Raphson method was employed, utilising a Hessian matrix of order 156 x 156, describing the curvature of the QM/MM energy hypersurface for a subset of the system (the QM atoms), together with a diagonal Hessian plus updates for the rest of the system. The r.m.s. residual gradient on the 52 atoms in the subset was less than 0.1 kcal/mol Å⁻¹ in the optimised structure, while on the remaining atoms (ca. 3000) it was less than 0.005 kcal/mol Å⁻¹. Finally, the intrinsic reaction coordinate (IRC)^[30] path was traced from the TSs in order to demonstrate conclusively that the reported structure was indeed the TS for the correct chemical step.

RESULTS AND DISCUSSION

The PES obtained for the interconversion of pyruvate into lactate in the active site of the wild-type LDH enzyme is depicted on Figure 2, the energies for minima and saddle points are listed in Table 1 and selected QM/MM optimized interatomic distances of the different stationary structures located on the PES are reported on Table 2. Figure 2 clearly corresponds to a PES with two alternative pathways: an asynchronous concerted reaction pathway showing an effective enthalpy barrier of 44.7 kcal mol⁻¹ (blue path), and a stepwise pathway (red path) with an intermediate where the proton from His195 has been completely transferred. In this second mechanism the rate limiting step is the hydride transfer (TS2) with an energy barrier of 33.8 kcal mol⁻¹. It is important to point out that an appropriate description of the QM region in the QM/MM scheme may have important consequences on the topology of the PES. In this regard, we showed in our previous study that a more expensive MP2//AM1/MM method renders a single pathway with a lower barrier.^[19] Nevertheless, for the purpose of the present systematic study of the effect of mutations and analysis of key residues of the enzyme active site, a very CPU time consuming method would be prohibitive. Thus, keeping in mind the kind of error that can be introduced with this approach, the comparative analysis based on the obtained AM1/MM PESs can be accepted. A correction based on a high level *ab initio* Hamiltonian would render qualitative deviations from the results presented in this study, but similar conclusions.

From the analysis of the substrate-enzyme interactions listed on Table 2, it is important to emphasize the significantly short distance established between the carbonyl oxygen atom of pyruvate and the Arg109, observed in the TS1 and also in reactants (pyruvate). As previously reported in the literature, this interaction stabilizes the TS and polarizes the carbonyl bond in reactants thus facilitating the hydride transfer.^[8] Furthermore, it is important to keep in mind the presence of a water molecule in the vacuole of the active site (detected by X-ray diffraction analysis)^[21] together with some other important hydrogen bond interactions established between the substrate and the Arg171, between His195 and the Asp168, or the interaction between Glu199 and Asp168. This series of interactions favours the favourable orientation of the ternary complex for the reaction to proceed. Furthermore, as we demonstrated in our previous paper,^[19] this local environment has important consequences in the acidity of the ionisable residues of the active site. It can be observed in Table 3 how the electrostatic interactions in the enzyme shift the pK_a values of these residues from the standard pK_a values in aqueous solution. In particular, it is clear that the protein environment raises the pK_a of some residues with respect to aqueous solution (Glu199, His195 or Arg109), demonstrating their protonated state in the enzyme at pH 7, which favours the catalysis, and diminishes the value of Asp168, proving the deprotonated state of this residue which, as previously showed, favours the protonation state of His195.

As mentioned above, a deeper insight in the role of the different aminoacids of the active site of the enzyme is obtained by the analysis of the effect of different mutations on the topology of the PES of the wild type enzyme. Changes in pK_a values, in the reaction mechanisms, structures and barrier heights are used to define the catalytic role of the different residues. The results of the different simulations are discussed in turn.

Mutation Arg109Gln

The PES resulting from carrying out the substitution of Arg109 by glutamine is depicted in Figure 3. As for the wild-type, the reaction catalysed by this mutant can proceed through two competitive mechanisms although now both are concerted. The intermediate located in the bottom-right corner of the PES for the wild-type (Figure 2) is no longer present. The pK_a of His195 is increased from 8.7 to 11.9 (Table 3), making the proton transfer from this residue to pyruvate much less favourable and, as a consequence, raising the energy of the lower right quadrant of the PES. The barrier for the reaction path through TS2 is nearly 9 kcal mol⁻¹ higher than for the wild-type. However, at 42.7 kcal mol⁻¹, this is still lower than the 48.1 kcal mol⁻¹ barrier for the reaction path through TS1. In topographical terms, the upper-left quadrant of the PES is also raised with respect to Figure 2: the putative intermediate with hydride completely transferred from NADH to pyruvate would be destabilized. The mutagenesis experiment of Clarke et al.^[8] showed that the reaction was still feasible, although slower, and these authors predicted an increase of about 4.2 kcal mol⁻¹ in the barrier height. Keeping in mind the limitations of the AM1 semiempirical method for obtaining quantitatively reliable energies, there is reasonable agreement between this and the present computed estimates for the barrier-height increases resulting from the Arg109Gln mutation.

The distances given in Table 2 suggest that, in the Arg109Gln mutant, Asn140 fulfils the role of Arg109 to a significant extent. This residue is only 2.20 Å from the carbonyl O atom of pyruvate in the reactant complex, as compared with 3.63 Å in the wild-type, and it approaches closer still in both TS1 and TS2 for the mutant. Similarly, an active-site water molecule (Wat103) is also found to be much closer in the mutant, especially in TS1. Consequently, the mutated enzyme shows similar behaviour to the wild-type from the chemical reaction point of view. Obviously, the other effect postulated by Clarke et al.^[6], which is that the substitution changes the properties of the enzyme by affecting the binding of the substrate, triggering loop closure and exploiting the favourable hydrophobic contacts, cannot be observed on our PES since our study considers only the chemical reaction step. Nevertheless, our result supports in part the experimental evidence that the enzyme keeps its activity after this Arg109Gln mutation, although at a lower efficiency, due to the fact that the stabilizing role of a neutral glutamine residue or a water molecule cannot be as strong as the one produced by a positively charged arginine.

Analysis of the electron distribution has been carried out by computing the Mulliken charges on the different atoms of the quantum sub-system. The atomic charges calculated for the C and O carbonyl bond atoms of the pyruvate moiety were +0.24 and -0.38, respectively, in the wild type enzyme, but +0.23 and -0.39 in the Arg109Gln mutant. This result is not in agreement with Clarke et al. who suggested that loss of the positively charged Arg109 residue would prohibit polarization of the pyruvate carbonyl group in the ground state.^[8] The charge separation between these C and O atoms in the TS2 for both enzymes is almost the same for both wild-type and mutant; perhaps the polarization, if it exists, is a more subtle effect than can be revealed by Mulliken analysis.

Mutation Asp168Ala

Analysis of the role of the negatively charged Asp168 has been carried out by replacing this residue by the smaller, neutral alanine. The resulting PES is depicted in Figure 4. As can be observed, by comparison with the wild-type enzyme, the effect of this mutation on the PES is a slightly destabilization of the lower-right quadrant; i.e. the proton transferred intermediate. It has been proposed that LDHs cannot form an enzyme-cofactor-substrate complex unless the imidazole group of His195 is protonated. Thus the negatively charged Asp168 would contribute to catalytic efficiency by raising the pK_a of His195 to ensure that there is a proton available for transfer to the substrate from a suitably orientated positively charged imidazolium group. As observed in Table 3, in this mutant the pK_a values of His195 and Glu199 are decreased to 7.3 and 4.2, respectively. This means that positively charged imidazolium is slightly less favorable. Owing to the presence of a bridging water molecule between His195 and Glu199, as depicted in Figure 5, the loss of Asp168 is, to some extent, compensated by the negatively charge of the ionised Glu199. Consequently the PES of the chemical reaction does not change dramatically as the result of this mutant.

As previously mentioned, it has been suggested that this mutation hinders the conformational rearrangement of the protein to accommodate the substrate-cofactor-His195 complex to react.^[6] This rearrangement could be observed if long molecular dynamics calculations were carried out and the populations of the different ground state structures were analyzed. Nevertheless, apart from the dramatic effect on the pK_a shifts, our analysis of the electron density has demonstrated that, as expected, the negative charge of Asp168 polarizes the His195 residue and also the carbonyl bond of the substrate, which facilitate the orientation of the carbonyl oxygen of pyruvate with respect to the transferring proton of the His195. This effect is observed mainly in the pyruvate reactant (see Table 4), where the wild type enzyme presents a more positively charged carbonyl C together with a more negative O atom, as compared with the Asp168Ala mutant.

Double mutation Arg109Gln & Asn140Asp

In order to check our prediction concerning the role of Asn140 in polarizing and stabilizing the TS as an alternative to Arg109, we have performed a double mutation where both residues have been replaced, respectively, by glutamine and aspartate acid. In principle this avoids favourable interactions between the carbonyl O of the substrate and any positively charged residue of the active site that could polarize the carbonyl C=O bond and stabilizes the TS for hydride transfer. Nevertheless, as shown in Table 3, the pK_a calculations for the ionisable residues at pH=7 reveal that the new Asp140 is protonated at this pH. The other dramatic effect in this regard is the pK_a value of His195, now much higher than in the wild-type enzyme. The resulting PES is presented in Figure 6, which still describes two competitive pathways for the reduction of pyruvate. By comparison with the PES of the wild-type, the blue pathway - although now stepwise - is still feasible and with energy barrier (46.0 kcal·mol⁻¹) close to its counterpart in the wild type enzyme (44.7 kcal·mol⁻¹). The intermediate of the upper left corner corresponds to an adduct in which the carbonyl O of pyruvate is bonded to C4 of the cofactor (Figure 7). This species progresses to products through a transition structure, TS3, corresponding to C-O cleavage and proton transfer from His195 to the O atom, with a barrier height lower than that for TS1 (see Table 1). An unexpected complex involving a covalent bond between NAD⁺ and pyruvate was experimentally obtained in the early seventies by Everse et al.^[31] and by Burgner and Ray in a study of pyruvate-induced inhibition in the Dogfish LDH system.^[32] However, this adduct was not the same as the one obtained in our simulations, since it was proposed to involve a covalent link between the β-carbon of pyruvate and the 4-position of the nicotinamide ring.^[31]

Regarding the red pathway in Figure 6, the previous optimized intermediate of the lower right quadrant is destabilized, although the overall mechanism presents a similar barrier height (37.9 kcal mol⁻¹) to that in Figure 2. Obviously, this destabilization of the lower right quadrant of the PES is related with the dramatic effect on the pK_a value of His195, which is now less predisposed to transfer a proton to pyruvate. Nevertheless, what it is surprising is the fact that TS1 is not as destabilized as expected. Analysis of this TS1 geometry reveals the appearance of a water molecule that seems to be strongly hydrogen bonded to the carbonyl O of pyruvate, with an O...H distance of 1.8 Å. This water molecule has moved more than 1 Å from its initial position in the wild-type enzyme and now looks to be performing, to some extent, the role of Arg109.

Double mutation Arg109Gln & Asn140Glu

The last simulation of our study corresponds to a double substitution, similar to the previous one, but in which Asn140 is replaced by glutamic acid. In this case, we are trying to reproduce the expected effect of the previous double mutation but now introducing a bigger residue at position 140. This new residue would prevent the entrance of a water molecule into the vacuole formed in the proximity of the carbonyl group of pyruvate, thereby avoiding the stabilization of the polarized C=O bond. Again, the most important effects of this double mutation on the pK_a shifts are the increase to 14 in the value for His195 and the fact that the new Glu140 is protonated at pH=7. The effect on the PES (Figure 8) is dramatic. The impossibility of polarizing the pyruvate carbonyl group in the ground state and the consequent

stabilization of the hydride transfer TS by any close residue provokes a very significant rise of the left side of the PES that removes the possibility of any reaction path passing through this region. The only reaction pathway across the PES is concerted but very asynchronous: the TS shows a very advanced degree of proton transfer (see Table 2) and presents a very high energy barrier (see Table 1). The effect of this double mutation is also observed in the thermodynamics of the chemical reaction, with the equilibrium being shifted back towards pyruvate, thus, disfavoring its conversion to lactate. Finally, the adduct observed in the previous double mutation also appears on this PES, although it is now kinetically irrelevant.

CONCLUSIONS

The reduction of pyruvate to lactate catalyzed by the LDH has been studied by means of hybrid QM/MM simulations. A very flexible molecular model consisting on the full tetramer of the enzyme, together with the cofactor NADH, the substrate and solvent water molecules has allowed us to mimic computationally the effects of site directed mutagenesis studies, most of which have previously been performed experimentally. The PESs obtained for every single mutation, compared with the one corresponding to the native enzyme, have been used to trace the possible reaction pathways and to locate and characterize the structures corresponding to the stationary points. Analysis of the reaction mechanisms and the properties of the stationary structures is a very powerful tool for determining the roles of the key residues on the vacuole formed in the active site of the enzyme. Furthermore, some of the substitutions have dramatic effects on pK_a values of ionisable residues, with consequent implications for catalysis.

The role of Arg109 in polarizing the carbonyl bond of pyruvate, can also be fulfilled by Asn140 or even by a water molecule. Our results are in very good agreement with the previous conclusions derived from site directed mutagenesis, which allows us to propose a novel double mutation on Arg109 and Asn140 that, as our calculations suggest would render an inactive enzyme. This double mutation would be an efficient pyruvate-induced inhibition protocol. This strategy can be extrapolated to other enzyme systems thus opening the door to a very promising methodology that, in combination with the appropriate experimental technique, will enable us to describe the phenomenon of enzyme catalysis and the particular role of the residues that form the protein. This knowledge places us in a privileged position to modify the activity of enzymes.

ACKNOWLEDGEMENT

We thank the Royal Society for a Joint Project Grant, DGI for project BQU2003-04168-C03, BANCAIXA for project P1A99-03, Generalitat Valenciana for GV01-324, GV04B-21, GV04B-131 and GRUPOS04/28. S. F. acknowledges a doctoral fellowship of the Ministerio de Ciencia y Tecnología. We are also grateful to Dr M. Field and Dr P. Amara for supporting us in the pK_a calculations and discussions.

TABLE CAPTIONS

Table 1: Relative energies (in kcal mol⁻¹) of the different stationary point structures on the QM/MM PES of the native type LDH and the different mutated forms.

Table 2: Selected interatomic distances (in Å) of the different stationary point structures located on the QM/MM PESs of the native type LDH and the different mutated forms.

Table 3: pK_a values of the different ionisable aminoacids in aqueous solution and calculated for the native type LDH and the different mutated forms.

Table 4: Mulliken atomic charges calculated on carbonyl carbon and oxygen atoms, for the native type LDH and the different mutated forms, on the stationary points located on the PESs.

FIGURE CAPTIONS

Figure 1. Schematic covalent structure for cofactor (NADH) substrate (pyruvate) and some key residues of the active site. GHO are indicated as "•", dividing the quantum region (shaded area) and the classical region.

Figure 2. QM/MM contours (kcal mol⁻¹) for pyruvate (R) to lactate (P) in the wild-type LDH-catalyzed reaction.

Figure 3. QM/MM contours (kcal mol⁻¹) for pyruvate (R) to lactate (P) in the Arg109Gln mutated LDH-catalyzed reaction.

Figure 4. QM/MM contours (kcal mol⁻¹) for pyruvate (R) to lactate (P) in the Asp168Ala mutated LDH-catalyzed reaction.

Figure 5. Detail of the QM/MM reactants refined structure in the Asp168Ala mutated LDH-catalyzed reaction.

Figure 6. QM/MM contours (kcal mol⁻¹) for pyruvate (R) to lactate (P) in the Arg109Gln & Asn140Asp double mutated LDH-catalyzed reaction.

Figure 7. Detail of the QM/MM adduct complex refined structure in the Asp168Ala mutated LDH-catalyzed reaction. This structure corresponds to the upper right corner on the PES depicted in Figure 6.

Figure 8. QM/MM contours (kcal mol⁻¹) for pyruvate (R) to lactate (P) in the Arg109Gln & Asn140Gln double mutated LDH-catalyzed reaction.

REFERENCES

- [1] P. Kedzierski, K. Moreton, A. R. Clarke and J. J. Holbrook, *Biochemistry-Us* **2001**, *40*, 7247-7252.
- [2] J. J. Holbrook, A. Liljas, S. J. Steindel and M. G. Rossmann in *in The Enzymes*. (P. D. Boyer, Ed.), Academic Press, New York, **1975**, pp. 191-293.
- [3] C. R. Dunn, H. M. Wilks, D. J. Halsall, T. Atkinson, A. R. Clarke, H. Muirhead and J. J. Holbrook, *Philos T Roy Soc B* **1991**, *332*, 177-&.
- [4] A. D. B. Waldman, K. W. Hart, A. R. Clarke, D. B. Wigley, D. A. Barstow, T. Atkinson, W. N. Chia and J. J. Holbrook, *Biochem Bioph Res Co* **1988**, *150*, 752-759.
- [5] J. J. Holbrook and V. A. Ingram, *Biochem J* **1973**, *131*, 729-738.
- [6] A. R. Clarke, H. M. Wilks, D. A. Barstow, T. Atkinson, W. N. Chia and J. J. Holbrook, *Biochemistry-Us* **1988**, *27*, 1617-1622.
- [7] K. W. Hart, A. R. Clarke, D. B. Wigley, W. N. Chia, D. A. Barstow, T. Atkinson and J. J. Holbrook, *Biochem Bioph Res Co* **1987**, *146*, 346-353.
- [8] A. R. Clarke, D. B. Wigley, W. N. Chia, D. Barstow, T. Atkinson and J. J. Holbrook, *Nature* **1986**, *324*, 699-702.
- [9] A. Cortes, D. C. Emery, D. J. Halsall, R. M. Jackson, A. R. Clarke and J. J. Holbrook, *Protein Sci* **1992**, *1*, 892-901.
- [10] R. Sakowicz, M. Gold and J. B. Jones, *J Am Chem Soc* **1995**, *117*, 2387-2394.
- [11] J. Wilkie and I. H. Williams, *J Am Chem Soc* **1992**, *114*, 5423-5425.
- [12] A. Yadav, R. M. Jackson, J. J. Holbrook and A. Warshel, *J Am Chem Soc* **1991**, *113*, 4800-4805.
- [13] S. Ranganathan and J. E. Gready, *J Phys Chem B* **1997**, *101*, 5614-5618.
- [14] V. Moliner, A. J. Turner and I. H. Williams, *Chem Commun* **1997**, 1271-1272.
- [15] A. J. Turner, V. Moliner and I. H. Williams, *Phys Chem Chem Phys* **1999**, *1*, 1323-1331.
- [16] S. Marti and V. Moliner, *J Chem Theory Comput* **2005**, *1*, 1008-1016.
- [17] V. Moliner and I. H. Williams, *Chem Commun* **2000**, 1843-1844.
- [18] S. Ferrer, J. J. Ruiz-Pernia, I. Tunon, V. Moliner, M. Garcia-Viloca, A. Gonzalez-Lafont and J. M. Lluch, *J Chem Theory Comput* **2005**, *1*, 750-761.
- [19] S. Ferrer, E. Silla, I. Tunon, M. Oliva and V. Moliner, *Chem Commun* **2005**, 5873-5875.
- [20] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J Comput Chem* **1983**, *4*, 187-217.
- [21] D. B. Wigley, S. J. Gamblin, J. P. Turkenburg, E. J. Dodson, K. Piontek, H. Muirhead and J. J. Holbrook, *J Mol Biol* **1992**, *223*, 317-335.
- [22] M. K. Gilson, *Proteins* **1993**, *15*, 266-282.
- [23] J. Antosiewicz, J. A. Mccammon and M. K. Gilson, *J Mol Biol* **1994**, *238*, 415-436.
- [24] Field, M. J.; David, L.; Rinaldo, D., personal communication
- [25] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J Chem Phys* **1983**, *79*, 926-935.
- [26] M. Garcia-Viloca, K. Nam, C. Alhambra and J. L. Gao, *J Phys Chem B* **2004**, *108*, 13501-13512.
- [27] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J Am Chem Soc* **1985**, *107*, 3902-3909.
- [28] J. L. Gao, P. Amara, C. Alhambra and M. J. Field, *J Phys Chem A* **1998**, *102*, 4714-4721.

- [29] P. Amara, M. J. Field, C. Alhambra and J. L. Gao, *Theor Chem Acc* **2000**, *104*, 336-343.
- [30] K. Fukui, *J Phys Chem-Us* **1970**, *74*, 4161-&.
- [31] J. Everse, R. E. Barnett, C. J. R. Thorne and N. O. Kaplan, *Arch Biochem Biophys* **1971**, *143*, 444-&.
- [32] J. W. Burgner and W. J. Ray, *Biochemistry-Us* **1974**, *13*, 4229-4237.

Table 1: Relative energies (in kcal mol⁻¹) of the different stationary point structures on the QM/MM PES of the native type LDH and the different mutated forms.

	Pyruvate	TS1	TS2	TS3	Int	Lactate
Native LDH	0.0	44.7	33.8	26.2	23.0	-14.2
Arg109Gln	0.0	48.1	42.7	-	-	17.5
Asp168Ala	0.0	42.8	38.0	-	-	-5.4
Arg109Gln & Asn140Asp	0.0	46.0	37.9	24.0	-6.5	-6.0
Arg109Gln & Asn140Glu	0.0	-	58.0	-	0.0	22.5

Table 2: Selected interatomic distances (in Å) of the different stationary point structures located on the QM/MM PESs of the native type LDH and the different mutated forms.

	native LDH							Arg109Gln			
	Pyr	TS1	TS2	TS3	Int	Lac	Pyr	TS1	TS2	Lac	
C _{NADH} ···H _t	1.13	1.64	1.28	1.13	1.13	2.67	1.12	1.75	1.29	2.47	
H _t ···C _{pyr}	2.49	1.24	1.51	2.50	2.48	1.13	2.45	1.21	1.45	1.13	
N _{His195} ···H _p	1.01	1.03	2.31	1.49	1.94	2.48	1.01	1.04	1.80	2.10	
H _p ···O _{pyr}	2.05	1.84	0.99	1.01	1.02	0.97	2.08	1.77	1.01	0.97	
O _{pyr} -HHres109	2.94	2.87	2.58	3.00	2.63	3.07	2.82	2.70	2.72	2.98	
O _{pyr} -HDres140	2.05	1.91	2.93	2.89	2.91	2.83	-	-	-	-	
O _{pyr} -Hwat	3.63	3.18	3.03	3.42	3.42	3.20	2.20	2.01	2.15	2.41	
OD1Asp168-HD1His195	3.03	2.73	3.26	3.16	3.16	2.78	2.14	1.97	3.03	3.65	
OE2Glu199-OD2Asp168	1.55	1.55	1.57	1.55	1.56	1.59	1.56	1.57	1.58	1.60	
O _{pyr} -HHres109	2.60	2.59	2.58	2.59	2.59	2.59	2.58	2.58	2.58	2.58	

	Asp168Ala				Arg109Gln & Asn140Asp					
	Pyr	TS1	TS2	Lac	Pyr	TS1	TS2	TS3	Int	Lac
C _{NADH} ···H _t	1.13	1.57	1.23	2.40	1.13	1.67	1.33	2.63	2.64	2.82
H _t ···C _{pyr}	2.55	1.28	1.61	1.13	2.64	1.22	1.41	1.12	1.12	1.13
N _{His195} ···H _p	1.01	1.04	1.98	2.65	1.00	1.03	1.78	1.08	1.01	2.15
H _p ···O _{pyr}	2.11	1.86	1.01	0.97	2.04	1.78	1.00	1.57	2.06	0.98
O _{pyr} -NE2His195	2.96	2.84	2.96	3.53	2.86	2.73	2.76	2.64	2.91	2.87
O _{pyr} -HHres109	2.28	2.04	2.54	2.13	-	-	-	-	-	-
O _{pyr} -HDres140	3.86	3.79	3.72	3.95	3.40	1.89	2.53	2.84	3.11	3.17
O _{pyr} -Hwat	3.31	2.63	3.50	3.43	2.01	1.79	3.28	4.68	3.51	3.61
OD1Asp168-HD1His195	-	-	-	-	1.58	1.59	1.60	1.59	1.60	1.60
His195-Hwat	1.65	1.68	1.69	1.66	-	-	-	-	-	-
Glu199-res168	1.62	1.63	1.63	1.63	1.60	1.60	1.59	1.59	1.61	1.60

Table 3: pK_a values of the different ionisable aminoacids in aqueous solution and calculated for the native type LDH and the different mutated forms.

	Aqueous solution	Native LDH	Arg109Gln	Asp168Ala	Arg109Gln & Asn140Asp	Arg109Gln & Asn140Glu
His195	6.0	8.7	11.9	7.3	14.0	14.0
Asp168	3.9	1.2	1.3	-	1.3	1.4
Glu199	4.1	7.4	8.0	4.2	8.3	8.4
Arg109	12.8	14.0	-	14.0	-	-
Residue 140	-	-	-	-	8.2	8.2

Table 4: Mulliken atomic charges calculated on carbonyl carbon and oxygen atoms, for the native type LDH and the different mutated forms, on the stationary points located on the PESs.

	native LDH						Arg109Gln			
	Pyruvate	TS1	TS2	TS3	Int	Lactate	Pyruvate	TS1	TS2	Lactate
C_{pyr}	0.240	0.185	0.271	0.196	0.274	-0.006	0.234	0.173	0.193	0.012
O_{pyr}	-0.378	-0.701	-0.190	-0.206	-	-0.342	-0.390	-	-	-0.398
					0.103			0.720	0.283	

	Asp168Ala				Arg109Gln & Asn140Asp					
	Pyruvate	TS1	TS2	Lactate	Pyruvate	TS1	TS2	TS3	Int	Lactate
C_{pyr}	0.220	0.191	0.219	-0.001	0.236	0.186	0.183	-	-	-0.001
O_{pyr}	-0.353	-0.666	-0.349	-0.390	-0.360	-	-	-	-	-0.367
						0.713	0.265	0.221	0.314	

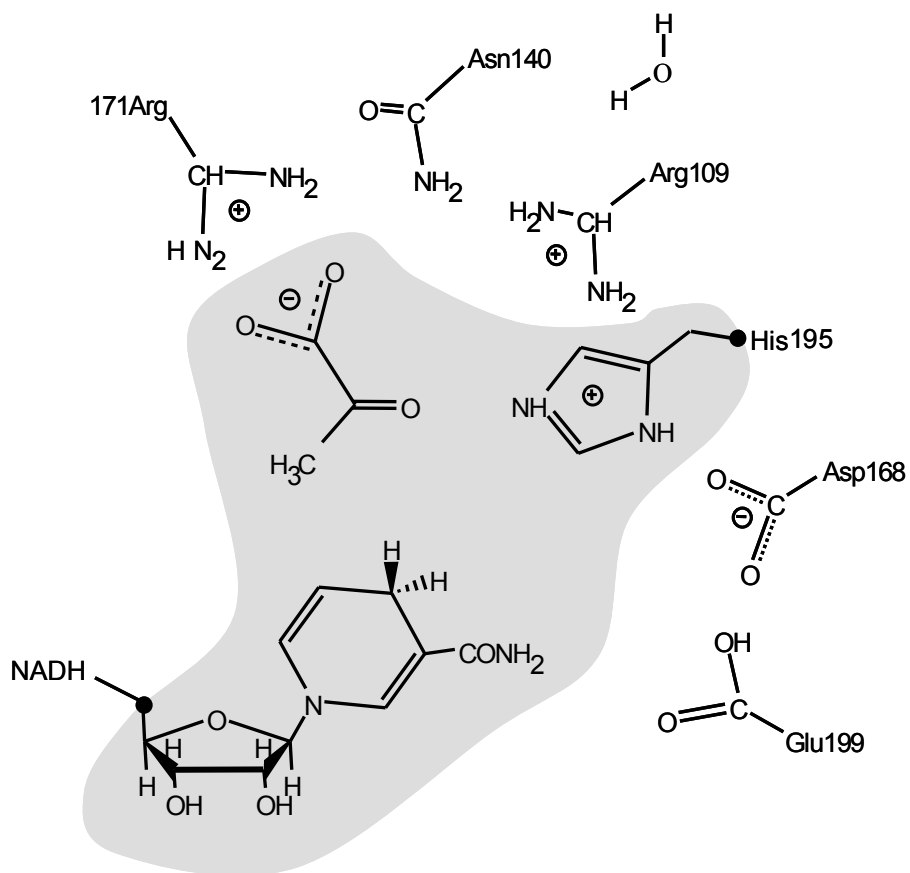


Figure 1

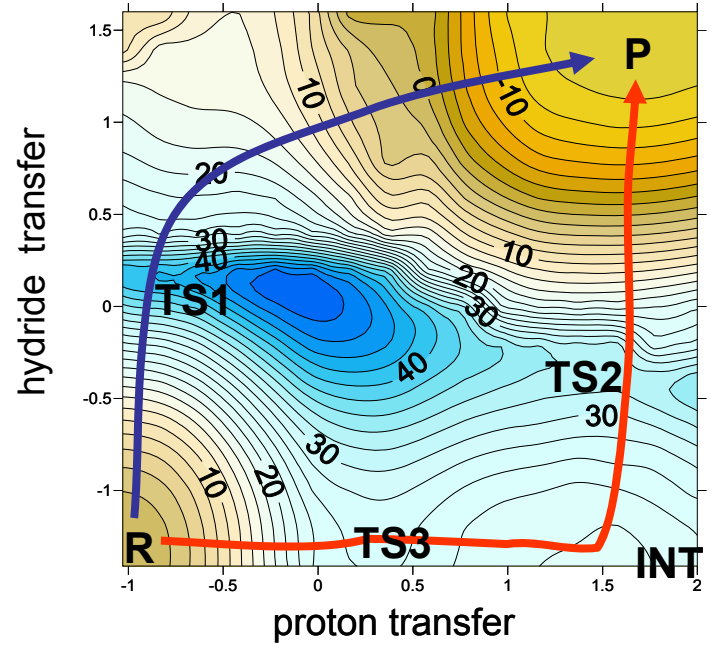


Figure 2

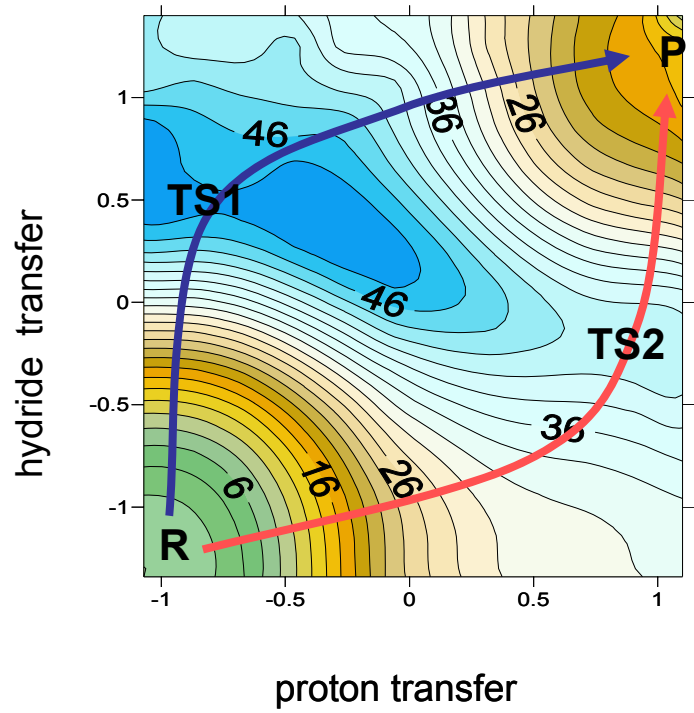


Figure 3

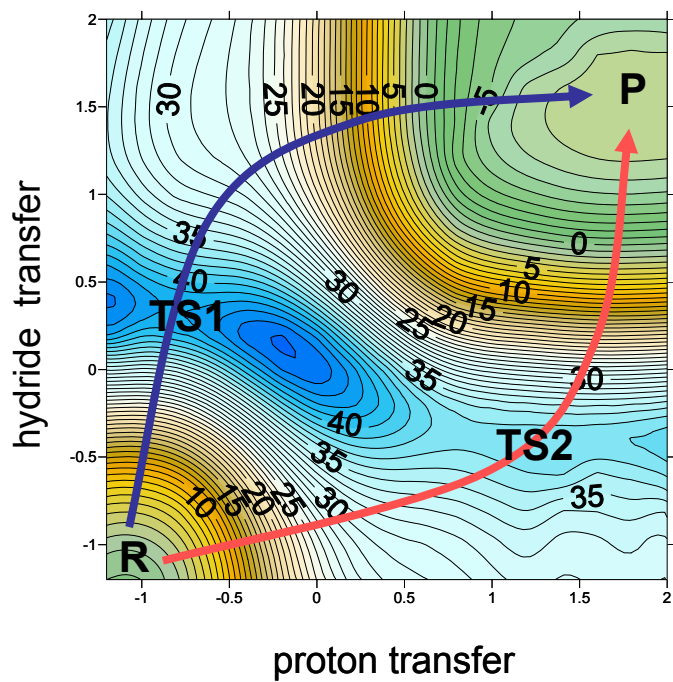


Figure 4

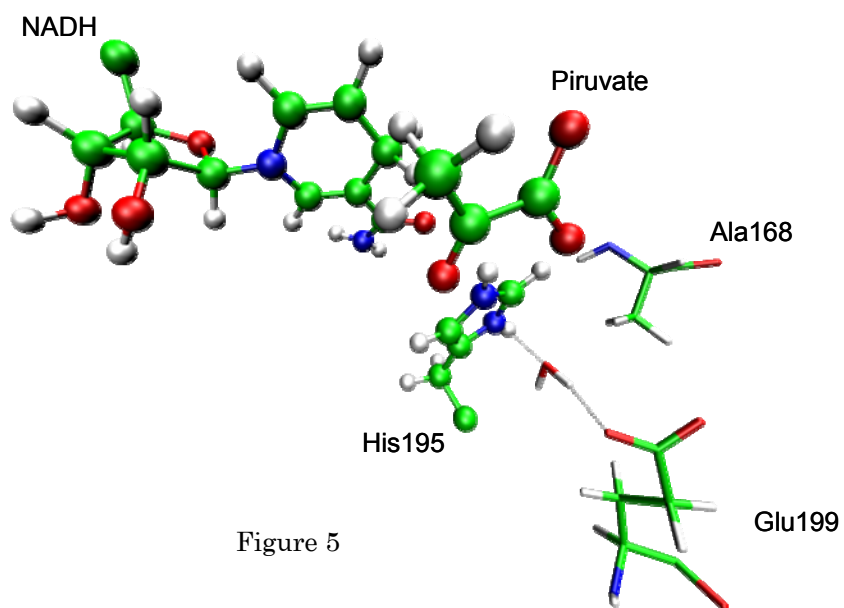


Figure 5

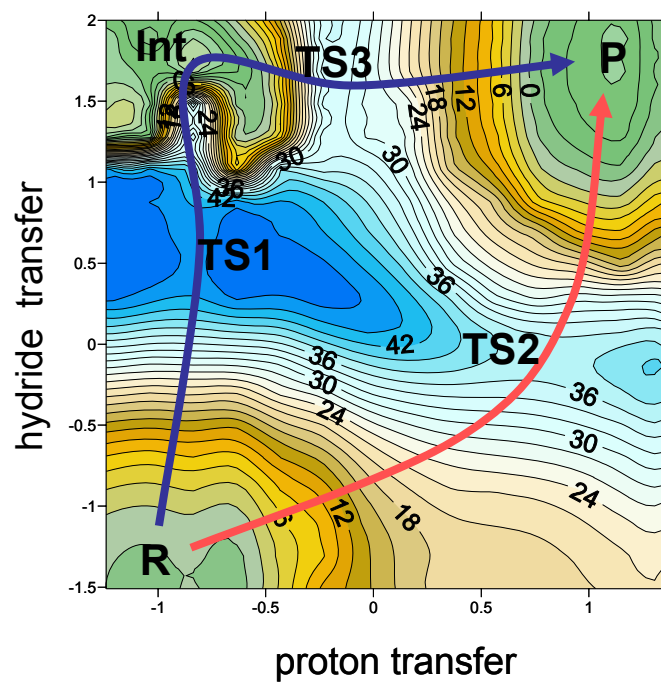


Figure 6

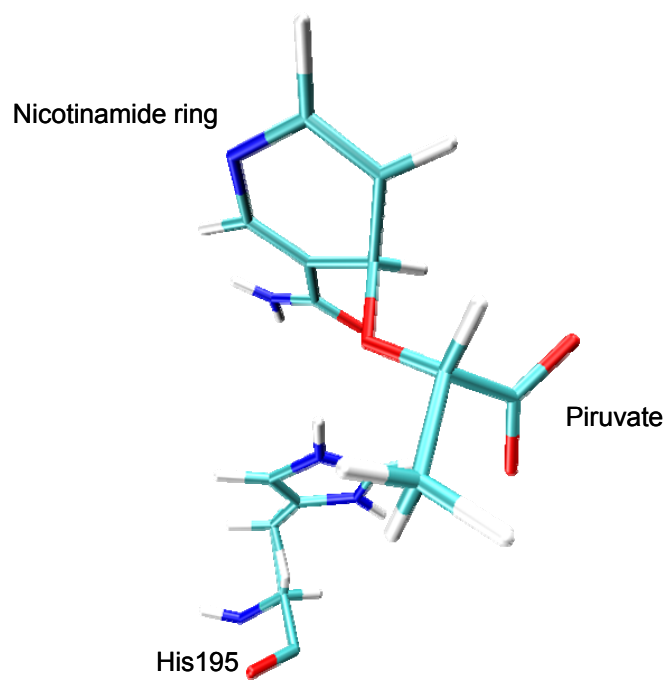


Figure 7

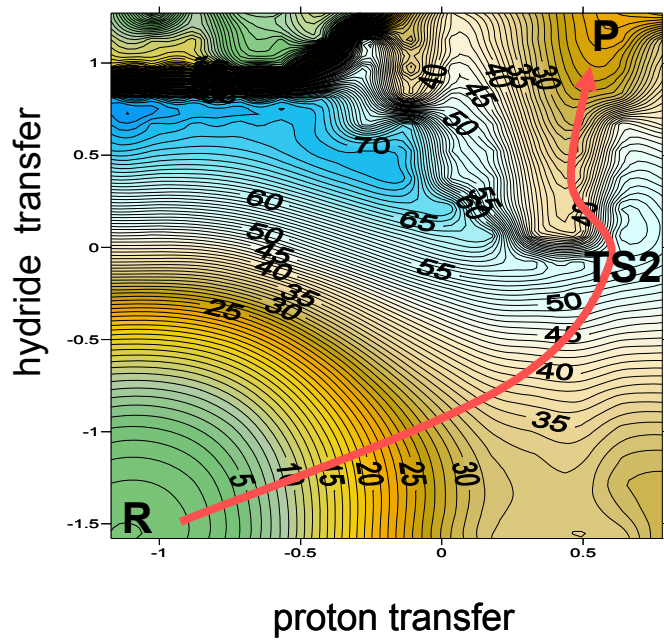


Figure 8.

A Theoretical Analysis of the Kinetic Behavior of Lactate Dehydrogenase. Determining Rate Constants and Kinetic Isotope Effects from Different Reaction Complexes.

Silvia Ferrer, Iñaki Tuñón*

Departament de Química Física, Universitat de València, València (SPAIN)

Sergio Martí, Vicente Moliner*

Departament de Ciències Experimentals, Universitat Jaume I, Castelló (SPAIN)

Mireia Garcia-Viloca, Angels González-Lafont, and José M. Lluch

Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona (SPAIN) and Departament de Química, Universitat Autònoma de Barcelona, Barcelona (SPAIN)

Universitat de Valencia

& Universitat Jaume I

‡ Institut de Biotecnologia i de Biomedicina

† Departament de Química, Universitat Autònoma de Barcelona

***e-mail: tunon@uv.es**

moliner@exp.uji.es

Abstract

In some enzymatic systems conformational changes are strongly coupled to the chemical step, in such a way that a large dispersion of rate constants can be observed in single-molecule experiments, each corresponding to different reactant complexes. Under this perspective here we present a computational study of pyruvate to lactate transformation catalyzed by Lactate Dehydrogenase. The reaction consists in a hydride plus proton transfer that take place concertedly. The degree of asynchronicity and the energy barrier depend on the particular reaction path analyzed. In order to estimate rate constants we used a free energy perturbation technique adapted to follow the intrinsic reaction coordinate for several possible reaction paths. Tunnelling contributions are also obtained with a slightly modified version of the Variational Transition State Theory-Optimized Multidimensional Tunneling methodology. According to our results the closure of the active site by means of a flexible loop, which is the rate-determining step, can lead to the formation of different reactant complexes displaying different features in the disposition of some key residues (such as Arg109), interactions with the substrate and number of water molecules in the active site. The chemical step of the reaction takes place with a different reaction rate from each of these complexes. Kinetic Isotope Effects corresponding to the use of NAD^2H as cofactor instead of NAD^1H are in good agreement with experimental observations.

1. Introduction

Enzymes are essential for living organisms, making chemical reactions to occur in a time scale compatible with life.¹ The ability of these proteins to speed up chemical processes lowering the activation free energy has been the subject of an increasing number of theoretical approaches during the last years.²⁻¹⁵ While a number of different computational strategies have been successful in reproducing some experimental values, the deep understanding of the molecular machinery involved in enzymatic reactions still requires the solution of many conceptual and practical challenges.

Computational studies of enzymatic processes usually involve the exploration of the configurational space of the system along a particular coordinate defining the changes between the reactant state and the product state. This coordinate (ζ) can be simply a geometrical parameter related to the broken or newly formed bonds (distances, angles or a combination of them, for example)^{3,4,16-20}. Otherwise, a collective coordinate can be also defined to follow the changes associated not only to the substrate but also to the enzyme. One of these collective coordinates used in the context of Valence Bond Theory is the energy difference between diabatic curves defining the reactant and product states.^{21,22} Irrespectively of the choice made for a particular reaction, the objective is to obtain the free energy differences between reactants and products and between reactants and the maximum in the free energy profile. While the first magnitude is related to the equilibrium of the process, the second one allows us obtaining the rate constant (k_r) of the reaction. Transition State Theory,^{23,24} or its variational extension, relates the activation free energy (ΔG^\ddagger) with the rate constant through the well-known equation:

$$k_r = \kappa(\zeta) \frac{k_B T}{h} e^{-\frac{\Delta G^\ddagger(\zeta)}{RT}} \quad (1)$$

where k_B is the Boltzmann's constant, T the temperature, h Planck's constant and κ the generalized transmission coefficient. We have explicitly shown in this equation that the activation free energy and the transmission

coefficient are dependent on the coordinate chosen to follow the reaction, while the rate constant is not. The activation free energy can be usually obtained in an approximate way as the difference between the maximum and the minimum (on the reactant side) of the free energy profile traced along ζ , although some additional terms related to the nature of the reaction coordinate should be considered, in particular when using collective coordinates.^{16,25} When a reaction takes place in a complex environment, such as the active site of an enzyme, a large number of geometrical changes can take place not only in the chemical subsystem (lengthening and shortening of bond distances) but also in the environment (for example, hydrogen bonds with the substrate and/or between residues of the active site can be formed or broken during the process). All these changes are coupled and can not be completely reflected in a simple geometrical coordinate, such as an antisymmetrical combination of the bond breaking and forming distances. When obtaining the free energy profile under the assumption of equilibrated environment, some strongly correlated motions of the environment are not perfectly sampled. In many cases, this can be solved by means of a transmission coefficient taking value close to unity. Phenomenological theories or molecular simulations can be employed to satisfactorily estimate this coordinate-dependent transmission coefficient.²⁶⁻³² This transmission coefficient can also be generalized to account for tunnelling contribution when light atoms transfers are taking place.^{18,33-37}

There are however, difficult cases for simulations, when the reactants to products transformations can not be satisfactorily described by a single geometrical coordinate or combination, a characteristic of those reactions in which two different chemical events take place in a concerted way. An example of this is provided by enzymatic reactions in which proton and hydride transfer occur concertedly.³⁸ Each of the two transfers can be satisfactorily described by the antisymmetric combination of the hydrogen-donor and hydrogen-acceptor distances but the description of the concerted process using a combination of geometrical parameters does not assure the correct mapping of those configurations relevant during the concerted path.

In terms of collective coordinates, this situation may correspond to a case where more than two diabatic states would be necessary to describe the process.

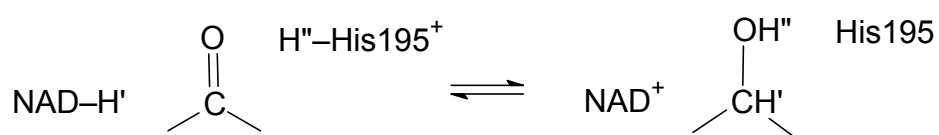
The problem of the selection of the relevant configurations of the whole system (substrate plus environment) for the process being analyzed can be considered in a broader perspective taking into account recent results of single-molecule enzyme catalysis that stress the coupling between chemical step and configurational changes.³⁹⁻⁴² According to these results different enzyme configurations can lead to different reaction paths, each contributing to the total reaction flux.^{39,40,43} The participation of the different enzyme-substrate complexes to the total reaction rate would depend on the free energy barriers and relative free energies among them. The limiting cases would correspond to i) thermodynamic control, when the concentration of the different enzyme-substrate complexes are found at their equilibrium values because the interconversion rate is much larger than the rate of the chemical step and ii) kinetic control, when the free energy barriers of interconversion among the different complexes is larger than the barrier associated to the chemical reaction. Of course, intermediate behaviours are also possible, leading to an enormous variety of possible behaviours depending on the peculiarities of the system. As a consequence a distribution of reaction rates is obtained and the interpretation of the macroscopic catalytic rate constant can be somewhat less straightforward than expected.³⁹ In particular, when interconversion among different reactant complexes is slow compared to the chemical step, the catalytic rate constant is obtained as a weighted mean of the catalytic rate constants corresponding to each of the conformers.³⁹ Experiments on β -galactosidase enzyme have shown that single molecule's rate constant can span over a broad range of time scales, from 10 to 10^{-3} s.⁴⁴ According to equation (1) this means that the free energy barrier that the reactants must overcome can vary as much as 5 Kcal/mol depending on the conformation in which they are found. These results can be related with previous experimental findings on myoglobin, showing that protein states present a hierarchical structure and thus that it is possible to find wells within

wells within wells.⁴⁵ Each of the different protein states have a different function but the different substrates or configurations (wells within wells) perform the same function but with different rates.

From a theoretical point of view these findings mean that molecular simulations of an enzymatic reaction should take into account different reaction paths starting from different reactants configurations and also the paths leading from one valley to another.^{43,46-49} Employing a single free energy barrier to characterize an enzymatic reaction may lead to an incomplete picture of the process. First, only very lengthy simulations would allow for a correct sampling of the different reactant complexes involved in the process and the reaction paths associated to each of these valleys. In practice, sampling is carried out for a limited region of the protein energy landscape and thus one can not be completely sure to include all the relevant reactants substrates. Second, this procedure does not directly inform about the transitions between these complexes. Protein fluctuations drive the system from a reactant valley to another from which the barrier height may be substantially different. Single free energy barrier procedures seem to be adequate for enzymatic processes with a limited number of possible configurations while may lead to non-converged results for proteins presenting large-scale fluctuations. Instead, it has been suggested that a limited number of free energy profiles could be averaged to generate a well-converged free energy barrier.⁴³ The problem is the way in which the profiles should be weighted, depending if equilibrium between reactant complexes is assumed or not.

Lactate dehydrogenases offer an excellent opportunity for the analysis of all these factors in the computational study of catalysis. L-Lactate dehydrogenases (LDH) are a family of 2-hydroxy acid oxidoreductases that simultaneously and stereospecifically interconvert pyruvate to lactate and the nicotinamide adenine dinucleotide cofactor from the NADH form to NAD⁺.⁵⁰⁻⁵³ This reaction, known as anaerobic homolactic fermentation is an important pathway for regenerating NAD⁺ enabling the continuation of

glycolysis.⁵⁴ Most mammals express several isozymes of LDH (up to five different forms in humans) and they have been also found in a range of species such as dogfish, *Bacillus stearothermophilus*, *Bifidobacterium longum* and *Plasmodium falciparum*.⁵⁵⁻⁵⁷ A high level of structural similarity is preserved among all these LDHs. In these enzymes binding of cofactor is followed by the substrate in an ordered process that includes the enclosure of the active site by means of a loop. This last step is the rate-limiting one.⁵⁸ In this movement of the active-site loop several key residues are brought to close contact with the substrate, including an invariant arginine that is believed to stabilize the transition state. The chemical step of the enzymatic process involves the transfer of a proton from a protonated histidine present in the active site to the carbonyl oxygen atom of pyruvate as well as a hydride transfer from the cofactor to the carbonyl carbon atom of the substrate (see Scheme I).⁵⁹ LDH presents at least two unique features in regard with our previous discussion on the theoretical description of enzymatic reactions. By one side, the chemical step involves a concerted proton and hydride transfer to the substrate (pyruvate) and thus the coordinate governing the chemical step can not be simply described as a single antisymmetric combination of distances.³⁸ Second, conformational changes of the enzyme are intrinsically coupled to the chemical process as far as this change is required for the reaction to occur.⁶⁰ Single-molecule kinetics have been followed by Hue and Yeung for LDH-1 (one of the five human isozymes and present in heart tissue). These authors described differences in the single enzyme activity by up to a factor of four, suggesting that different conformations in enzyme molecules may have different reactivity.^{41,42}



One of the most studied LDH enzymes is BsLDH (the lactate dehydrogenase of *Bacillus starothermophilus*). BsLDH exists in dimeric and tetrameric forms, the latter favoured by the presence of fructose-1,6-bisphosphate.^{60,61} Although the kinetic behaviour of the two forms is different, the chemical step takes place at similar rates in both forms.⁶⁰ The rate constant at 25°C is 250 s⁻¹ and it is assigned to the active site loop closure (in this case, it is formed by residues 98-110). In this movement the invariant arginine, Arg109, moves 8 Å from a position in the solvent to one in the active site.⁵⁰ A combination of site-directed mutagenesis and kinetic isotopic effects allowed concluding that the rate constant of the chemical step should be of at least 750 s⁻¹ at 25°C,⁵⁰ which means that the phenomenological activation free energy should be lower than 13.5 kcal/mol. Primary kinetic isotope effects (KIEs) were measured using NAD²H instead of NADH on the Arg109Gln mutant enzyme, resulting in a 3-fold decrease of the rate constant at 25°C.⁵⁰ Schowen and coworkers determined a primary KIE of 5.5 at 55°C for the same LDH.⁶² Primary isotope effects of 3-6 are typical for dehydrogenases.⁵⁰ An experimental value of 2.70 was determined for mutant LDH in which Arg109 was replaced by lysine.

Theoretical calculations have revealed the difficulties associated to the study of this enzyme, in particular for determining the relative timing of the hydride and the proton transfer. Ranganathan and Gready, using hybrid quantum mechanics/molecular mechanics (QM/MM) methods, found a mechanism in which the hydride transfer preceded the proton transfer in a stepwise manner, in agreement with an earlier empirical valence bond study and with QM/MM results of an analogous malate dehydrogenase.^{11,60,63,64} One of us found a family of transition structures, obtained from different starting configurations, corresponding to a concerted mechanism where the proton transfer was considerably more advanced than hydride transfer.⁶⁵ The analysis of these transition structures revealed subtle differences in the interactions established between the substrate and the invariant Arg109. In particular it was found that the TS-stabilising role of this residue could also be alternatively played by Asn140.⁶⁵ It was also described that flexible

QM/MM modelling, allowing the enzymatic environment to change while the system advances along the reaction coordinate, could embrace the alternative mechanisms previously described.⁶⁶ While all these QM/MM calculations were made using a semiempirical AM1 Hamiltonian, we recently tackled the problem of improving the quality of the potential energy surface including corrections to the quantum description.⁶⁷ When these corrections were included, together with a systematic assignment of the protonation state of the different residues according to electrostatic calculations, a concerted mechanism was obtained where the hydride transfer was more advanced than the proton transfer.⁶⁷ It is worth mentioning that relative ordering of the proton and hydride transfer and the possible existence of shallow reaction intermediates are very elusive questions. Slightly different theoretical treatments of the system could lead to qualitatively different descriptions of the reaction. Different orderings between hydride and proton transfers and different degrees of concertedness between them could be plausible. In a recent transition path sampling study on the human heart isozyme, h-H₄LDH, several reaction trajectories were described differing in the time lag between the hydride and the proton transfer.⁶⁸

In this paper we present a theoretical study on the pyruvate to lactate transformation catalyzed by BsLDH. We specifically afford the problem of the proper sampling of the relevant configurations of the system, taking into account that the global process involves not only two chemical events (the proton and the hydride transfer) but also configurational changes of the enzyme that can determine the differences in the reactivity. Thus, our study considers the contribution of different reactant complexes by means of the calculation of several free energy profiles, each of them obtained from different starting configurations. We suggest that these different configurations can be obtained through analysis of intrinsic reaction coordinate (IRC)⁶⁹ calculations traced from transition structures localized from a sampling in the transition state region. Moreover, because of the complexity of the chemical step we suggest that the intrinsic reaction coordinate should be a more appropriate reaction coordinate than any

simple combination of distances. Thus, each of the free energy profiles is obtained sampling configurations of the system at different fixed positions of the IRC, by means of an improved free energy perturbation technique. Finally, tunnelling contributions to the rate constant are incorporated through an adaptation of the Ensemble Averaged-Variational Transition State Theory-Optimized Multidimensional Tunneling (EA-VTST-OMT) methodology.^{33,70,71}

2. Methodology

Model of the Enzyme-Substrate-Coenzyme Complex.

The X-ray structure of Lactate Dehydrogenase from *Bacillus stearothermophilus* comes from the PDB code 1LDN⁷² which is an octamer, although in this work only the tetramer is used. X-Ray studies have proposed that the tetramer is the functional form of Lactate Dehydrogenase from *Bacillus stearothermophilus*.⁷² In each monomer the crystal structure contains a total of 316 amino acid residues, the cofactor NADH, and the inhibitor oxamate (OXM), which is replaced by pyruvate in our study, in each active site.

The coordinates of the hydrogen atoms of the protein and coenzyme were determined using the HBUILD facility of the CHARMM package.⁷³ The pK_a of titrable aminoacid groups within the protein were calculated using the “cluster method” by Gilson and co-workers^{74,75} as implemented by Field and co-workers.⁷⁶ According to this method, each titrable residue in the protein is perturbed by the electrostatic effect of the protein environment. As demonstrated in our previous study on LDH,³⁸ a non-accurate protonation state of only a few residues of the protein can render erroneous PESs. The most significant change with respect to the use of aqueous solution pK_a values is the fact that Glu199 is found in its protonated form. Glu199 is close to Asp168, a residue that interacts directly to maintain His195 in a protonated state in preparation for substrate binding and turnover.³⁸

Then, the system was partitioned into a quantum mechanical region consisting of 78 atoms and a molecular mechanical region containing the rest of the system. The QM subsystem includes 9 atoms of the pyruvate, 13 atoms of the His195 residue -including the C_α as a boundary atom-, 30 atoms of the NADH -which include the dihydronicotinamide and ribose rings, and the C5' ribose atom as a boundary atom- and 13 atoms -including the C_γ as boundary atoms- from each of the two arginines present in the active

site –Arg109 and Arg171-. The QM subsystem is described using the AM1 Hamiltonian. The boundary atoms are represented by the GHO method. A picture of the active site showing the partition into QM and MM subsystems is presented in Figure 1. We carried out a protocol to solvate the active site and relax the energetically unfavorable contacts.⁶⁷ The QM/MM van der Waals interactions were recalibrated. The CHARMM22 van der Waals parameters are expected to work between MM atoms belonging to this force field, but they may not describe correctly a QM-MM(AM1) interaction. The new parameters were obtained following the procedure described by Freindorf and Gao⁷⁷ and applied recently by our group.⁶⁷

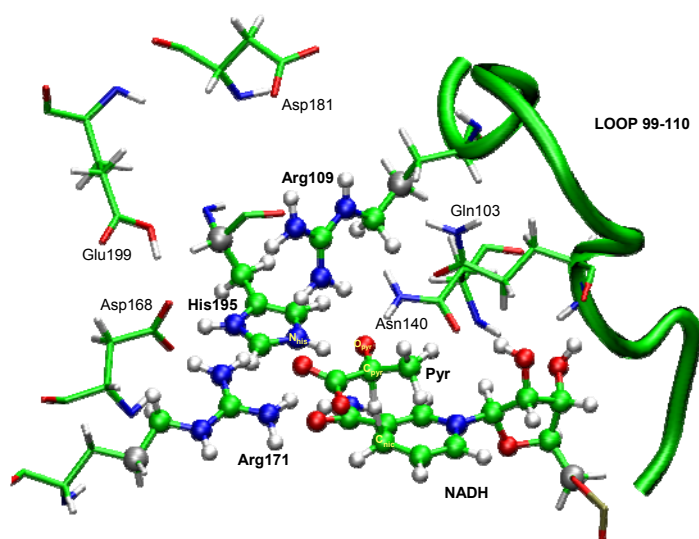


Figure 1. Transition structure for the pyruvate to lactate transformation in the LDH active center with the cofactor (NADH): The QM region is represented using balls and sticks. The position of the GHO atoms used to define the boundary with the MM region is also indicated as grey balls. Bold characters are used to indicate that the residue is described, at least partially, quantum mechanically.

The chemical reaction catalyzed by LDH involves two elementary processes: a hydride transfer from a carbon atom of dihydronicotinamide, C_{nic} (donor atom), to the carbonyl carbon atom of pyruvate, C_{pyr} (acceptor atom), and a proton transfer, from the N atom of His195, N_{his} (donor atom), to the carbonyl oxygen atom of pyruvate, O_{pyr} (the acceptor atom). Both processes can be described on a single PES obtained as a function of two reaction coordinates. The reaction coordinate R_1 for the hydride transfer is defined in this work as the difference in the distances of the bonds between the transferring hydride-ion and the donor and the acceptor atoms.:

$$R_1 = r_{C_{\text{nic}}H'} - r_{C_{\text{pyr}}H'} \quad (2)$$

For the proton transfer a reaction coordinate R_2 is defined as the difference in the distance of the bonds between the transferring proton and the donor (N_{his}) and the acceptor (O_{pyr}) atoms:

$$R_2 = r_{N_{\text{his}}H''} - r_{O_{\text{pyr}}H''} \quad (3)$$

The exploration of the PES was then carried out using the R_1 and R_2 coordinates defined before. CHARMM program was employed to carry out this exploration by means of the use of the RESDISTANCE keyword to define the reaction coordinates (R_i).

In order to represent as accurately as possible the energetics of the reaction and taking into account the limitations of the AM1 method we decided to use a modified QM/MM energy function in our simulations:

$$E = E_{\text{QM}}^{0,\text{AM1}} + E_{\text{QM/MM}}^{\text{AM1}} + E_{\text{MM}} + 2D_spline\{E_{\text{QM}}^{0,\text{MP2}}(R_1, R_2) - E_{\text{QM}}^{0,\text{AM1}}(R_1, R_2)\} \quad (4)$$

where the superscript 0 indicates the gas phase energy. In this expression we have included a correction term to the gas phase quantum energy as a function of the R_1 and R_2 coordinates. This correction, calculated as the difference between MP2/6-31G* and AM1 calculations for a subset of the QM system on structures picked up from the AM1/MM PES,⁶⁷ is fitted through the use of a bidimensional cubic spline function, ensuring then the

continuity of the energy and the first and second derivatives. During the simulations, those atoms 24 Å away from the active site, 14996 atoms, were kept frozen in order to reduce the computational cost while 7143 atoms were allowed to move. In all the cases, the total energy of system was minimized with the ABNR⁷³ method until the norm of the gradient was less than 0.001 kcal mol⁻¹·Å⁻¹. The GRACEFUL algorithm⁷⁸ was used to optimize and characterize an initial transition structure using a Hessian matrix containing all the coordinates of the QM subsystem.³⁸

In order to obtain additional transition structures, we carried out molecular dynamic simulations in the transition state region, restraining the R₁ and R₂ coordinates using Langevin dynamics with the atoms placed more than 24 Å of the active site kept fixed, while those placed between 20 Å and 24 Å define the buffer region. The simulation protocol is described in detail elsewhere⁶⁷ and the time step was of 1 fs. Configurations generated during the dynamics were used as starting point for the localization and characterization of new transition structures. From these structures, the intrinsic reaction coordinates (IRC)⁶⁹ were traced down up to reactants and products valleys. With this purpose new structures were generated following the gradient vector, except in the vicinity of the transition structure where the transition vector was used instead. To obtain the free energy profiles associated to some of the reaction paths using a free energy perturbation scheme (see below), MD Langevin simulations were carried out for all the structures appearing between the reactant and the product structures keeping fixed the coordinates of the QM subsystem. In addition, long simulations (500 ps) were carried out for some of the reactant complexes without fixing any coordinate of the QM region in order to characterize the average properties of these complexes.

3. Results

3.1. Exploring the Transition State Region

Our computational study begins with the exploration of the TS region. For this purpose we selected as initial configuration of the system a transition structure located and characterized after exploration of a PES defined through the use of the R_1 and R_2 coordinates.⁶⁷ This transition structure corresponds to a concerted hydride and proton transfer where the former is considerably more advanced than the latter. In this structure the pyruvate is fixed in the active site through interactions of the carboxylate group with Arg171. The carbonyl oxygen atom of pyruvate establishes hydrogen bond interactions with Arg109, although this role can also be played by Asn140. Other key residues, as the proton donor His195, and the NADH cofactor are also shown in **Figure 1**. In order to explore the TS region we run molecular dynamic simulations of the system restraining the values of the R_1 and R_2 coordinates. After heating at 298K we performed two different simulations at slightly different target temperatures (298 and 300K) in order to enhance the exploration over a wider range of configurations of the system. These two simulations in fact explored regions of the configurational space with important differences from the point of view of catalysis. **Figure 2** shows the evolution of the CA-CB-CD-CG and CA-CB-CD-N dihedral angles of Arg109 and Asn140 in both MD simulations. The different orientation reached in the 300K MD after some picoseconds corresponds to a different pattern of hydrogen bonds with the carbonyl oxygen atom of pyruvate. In the 298K MD Arg109 establishes a hydrogen bond with the carbonyl oxygen atom of the substrate (O_{pyr}) but in the 300K MD this residue is displaced by Asn140 after some picoseconds. This competition between these two residues was already reported in our previous computational studies and it was proposed as a redundant stabilization mechanism that could explain the fact that mutation of Arg109 does not completely destroy the catalytic activity of the enzyme.⁵⁸

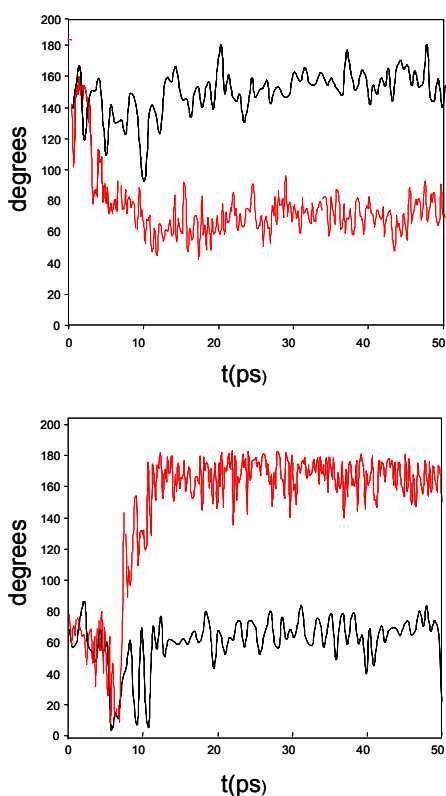


Figure 2. Time evolution of the CA-CB-CD-CG angle of Arg109 (top) and CA-CB-CD-N angle of Asn140 (bottom) during 298 (black) and 300 (red) K MD simulations in the transition state region.

Different snapshots of the two MD simulations were selected as starting points to locate and characterize transition structures of the process. The transition structure search is carried out by means of a combination of GRACEFUL and CHARMM programs, defining as active space the coordinates of all the QM atoms, while the rest of coordinates of the mobile atoms form the complementary space. The stationary structures obtained in this way (with zero or nearly zero gradients in both the active and complementary spaces) present one and only one imaginary frequency in the active space. We obtained up to 24 transition structures that are gathered in **Table I**. These transition structures are characterized from the bond breaking and bond forming distances corresponding to the proton and

hydride transfer, the hydrogen bond distances to either Arg109 or Asn140, the total potential energy as well as by the value of the imaginary frequency. The structures localized show an important dispersion in these properties, what clearly shows the dependence of the transition structure properties on the selected initial configuration. In all cases the transition structure obtained describes a concerted (but asynchronous) process where the local curvature seems extremely sensitive to fluctuations of the system, the imaginary frequency changes by a factor of ten: from 115i to 1188i cm^{-1} . The maximum of the imaginary frequencies distribution function appears around 300i cm^{-1} , but it displays a very long tail up to very large values (about 1200i cm^{-1}). With respect to the hydride transfer the bond breaking distance ($C_{\text{nic}}\text{H}'$) ranges between 1.39 and 1.69 Å while the bond forming distance ($C_{\text{pyr}}\text{H}'$) is found in a narrower interval, between 1.24 and 1.36 Å. In the case of the proton transfer, the $N_{\text{his}}\text{H}''$ distance varies between 1.08 and 1.20 Å and the pyruvate carbonyl oxygen-proton ($O_{\text{pyr}}\text{H}''$) distance between 1.37 and 1.66 Å. In terms of the hydride and proton transfer coordinates, the dispersion of the transition structures presented in **Table I** is graphically depicted in **Figure 3** where the localized structures are represented as a function of R_1 and R_2 transfer coordinates. For all of these structures we traced the IRC up to the products and reactants valleys. The energy barriers obtained as the potential energy difference between the transition and the reactant structures ranges between 17 and 83 Kcal/mol. Such a wide distribution of energy barriers is not completely surprising. Very similar results were obtained for the energy barrier of the hydride transfer in Dihydrofolate Reductase (DHFR), an enzyme that also has a flexible loop in the active site.^{43,46}

Table I. Selected distances (in Å) of the localized transition structures for the chemical reaction catalyzed by LDH. Energies and potential energy barriers are given in Kcal/mol and imaginary frequencies in cm^{-1} .

$d(\text{C}_{\text{nic}}\text{H}')$	$d(\text{H}'\text{C}_{\text{pyr}})$	$d(\text{N}_{\text{his}}\text{H}'')$	$d(\text{O}_{\text{pyr}}\text{H}'')$	Energy	Freq
1.43	1.31	1.09	1.57	-17907	776i
1.39	1.36	1.10	1.59	-18213	1186i
1.58	1.24	1.20	1.35	-18382	630i
1.39	1.35	1.08	1.66	-18165	1095i
1.63	1.24	1.16	1.43	-18305	462i
1.69	1.23	1.15	1.43	-18329	356i
1.61	1.25	1.14	1.45	-18354	556i
1.58	1.25	1.17	1.41	-18383	352i
1.57	1.25	1.15	1.43	-18410	619i
1.57	1.24	1.18	1.38	-18419	222i
1.54	1.26	1.12	1.52	-18419	193i
1.52	1.27	1.11	1.55	-18418	257i
1.55	1.25	1.12	1.52	-18470	215i
1.53	1.26	1.12	1.53	-18446	284i
1.55	1.25	1.14	1.45	-18421	352i
1.57	1.25	1.15	1.43	-18426	501i
1.56	1.24	1.18	1.37	-18410	541i
1.54	1.25	1.18	1.37	-18429	284i
1.55	1.25	1.14	1.45	-18379	317i
1.53	1.26	1.13	1.48	-18426	476i
1.55	1.26	1.16	1.41	-18402	720i
1.40	1.34	1.08	1.65	-18183	1058i
1.41	1.32	1.08	1.62	-18187	850i
1.55	1.24	1.11	1.52	-18326	115i

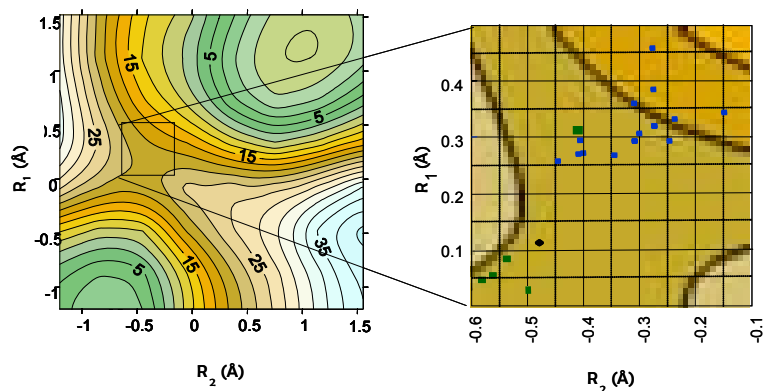


Figure 3. Representation of the localized transition structures on the Potential Energy Surface. The initial transition structure is represented as a black dot, those coming from the 298 K simulation as green dots while blue dots represent those coming from the 300 K simulation. The isoenergetical lines correspond to the transition structure represented in black.

3.2. Reactant State Analysis

The basic idea exploited in this paper is that reactant enzyme-substrate complexes displaying important differences can be obtained starting from different transition structures. Effectively, downhill IRC trajectories started at slightly different starting points (the transition structures) and following slightly different directions (the transition vectors) can lead to quite different reactant valleys. This procedure may be assimilated to a kind of ‘chemical simulated annealing’ to find those system minima relevant for the chemical process under study. To check this point we selected four reactant structures obtained from the IRCs of the transition structures appearing in Table I. In particular we selected the four first transition structures appearing in Table I (hereafter denoted as TS1, TS2, TS3 and TS4). We believe that with this selection we cover a significant interval of possible reaction barriers (from 17 to 36 kcal/mol) while the highest energy barrier values are expected to

contribute much less to the experimentally observed free energy barrier considering the next connection between these magnitudes:⁴⁶

$$\Delta G^\ddagger = -kT \ln \left\langle \exp \left(-\frac{\Delta E^\ddagger}{kT} \right) \right\rangle \quad (5)$$

Relatively long (500 ps) of NVT molecular dynamic simulations were run for each of these reactant complexes. The analysis of these trajectories effectively showed significant differences during the simulated time. Of course, one may wonder if longer simulation times will show significant overlap of the four trajectories in the configurational space. However, from a practical point of view, and considering the typical simulation times used to obtain free energy profiles, we can consider these complexes as different valleys and contributing each of them to the macroscopic reaction flux. **Table II** provides the averaged values and fluctuations (measured as standard deviations of the averaged property) of some geometrical parameters related to pyruvate obtained during the dynamics of these four reactant complexes.

Table II. Averaged geometries of the four reactant complexes analyzed. Distances are given in Å and the standard deviation is provided in parenthesis.

	REAC1	REAC2	REAC3	REAC4
d(C _{nic} H')	1.13 (0.03)	1.13 (0.02)	1.13 (0.02)	1.13 (0.03)
d(C _{pyr} H')	4.3 (0.7)	3.3 (0.3)	3.5 (0.4)	3.20 (0.2)
d(N _{his} H'')	1.00 (0.02)	1.01 (0.02)	1.01 (0.02)	1.01 (0.02)
d(O _{pyr} H'')	3.0 (0.2)	3.4 (0.4)	3.1 (0.2)	3.0 (0.2)
d(O _{pyr} H _{nic})	3.9 (1.0)	2.5 (0.3)	3.2 (0.7)	2.7 (0.4)
d(O _{pyr} HH _{Arg109})	7.1 (0.4)	6.9 (0.6)	6.0 (0.6)	6.7 (0.6)
d(O _{pyr} HD1 _{Asn140})	6.3 (1.4)	2.6 (0.3)	4.8 (0.7)	6.1 (0.9)
d(O _{pyr} HH _{Arg171})	5.4 (0.2)	5.5 (0.5)	4.0 (0.2)	5.2 (0.5)
d(O _{pyr} H _w)	5.2 (0.7)	6.0 (0.7)	5.8 (0.3)	4.6 (0.6)

With respect to the hydrogen transfer parameters, the donor-hydrogen distances (C_{nic}H' and N_{his}H'') are nearly identical in the four simulations. Obviously, and because of the significantly lower force constants associated to these parameters, the fluctuations and the differences are much larger for the acceptor-hydrogen distances. In particular, the reactant state obtained

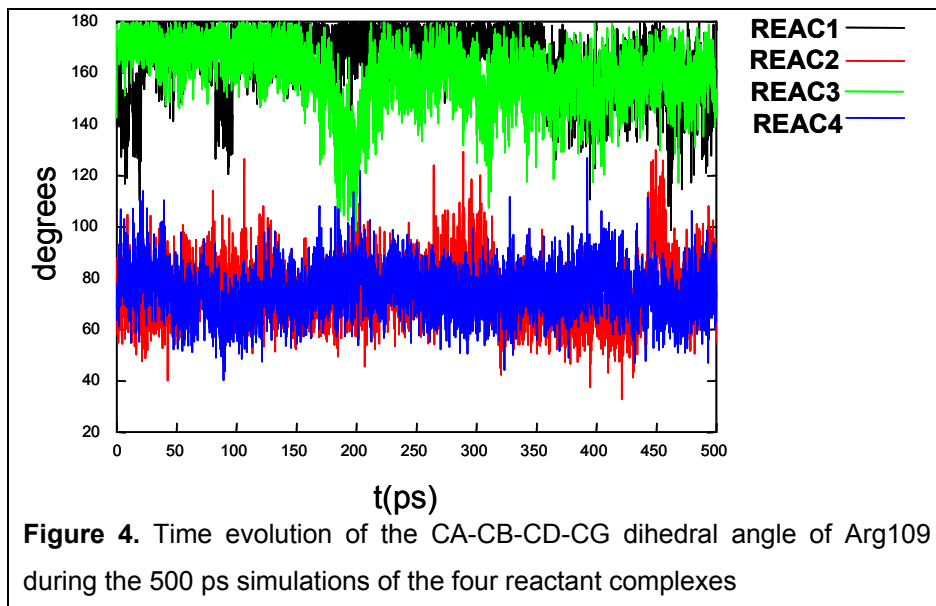
from TS1 (REAC1) displays a $C_{\text{pyr}}\text{H}'$ distance larger than in the other reactant complexes. The standard deviation is also significantly larger reflecting the fact that wider fluctuations are observed for REAC1. Part of this difference could be attributed to a compression effect due to two hydrophobic residues (Ile253 and Val33) placed at the other side of the nicotinamide ring of NADH with respect to pyruvate. In a recent study on the human heart isozyme, h-H₄LDH, it was suggested that a residue occupying a similar position (Val31) could play an active role favouring the hydride transfer.⁴⁹ Analysis of the averaged minimum distance between C_{nic} and carbon atoms of these two residues showed that REAC2 and REAC3 display the shorter distances to Ile253 (3.7 and 3.8 Å respectively), while REAC4 presents the shortest distance to Val33 (4.9 Å). It then seems that in REAC1 there is somewhat more room for nicotinamide ring and this could contribute to the larger value observed in the standard deviation of the $C_{\text{pyr}}\text{H}'$ distance. More significant differences, that could also explain this characteristic found in REAC1 simulation, are found in the analysis of pyruvate-enzyme interactions. In fact, REAC1 trajectory presents the largest value for the averaged RMS displacement of pyruvate (1.04, 1.00, 0.84 and 0.83 Å for REAC1 to REAC4, respectively). In the four reactant trajectories analyzed here the substrate is anchored by means of hydrogen bond interactions between pyruvate carboxylate group and Arg171. This interaction is well conserved during the four simulations. The differences appear in the interactions of the carbonyl oxygen (O_{pyr}). The pattern of interactions established by this atom in each of the simulation is quite different. In all cases there is a strong hydrogen bond interaction with the protonated His195, the residue responsible of the proton transfer to the substrate. The function and positioning of this residue is modulated by a negatively charged Asp168 which, in turns, establishes strong contacts with Arg171 and a protonated Glu199. The pKa of this last residue is considerably shifted from its standard value in aqueous solution due precisely to the proximity of Asp168, an essential feature for the energetics of the reaction and that was addressed in a previous work³⁸. Coming back to the description of the pyruvate carbonyl oxygen interactions it is interesting to note that in none of

the simulation Arg109 establishes hydrogen bond interactions with this atom. Considering that these interactions may appear in the transition structures (at least in some of them, as discussed in the previous section) one may conclude that the coordinates of this residue should be included in the correct definition of the reaction coordinate, at least for some of the possible reaction paths. The $O_{\text{pyr}}-\text{Arg109}$ hydrogen bond observed in some of the transition structures must be established during the reaction progress. Considering that the hydride transfer precedes the proton transfer in the concerted asynchronous transfer, it could be argued that an additional negative charge (or fraction of charge) on the pyruvate could be the driving force needed to approach this residue. Otherwise, Asn140 a residue candidate to explain the conserved, although reduced, activity of Arg109 mutants⁵⁸ forms well-conserved hydrogen bond with O_{pyr} during REAC2 simulations. This hydrogen bond is kept during the reaction path and then it is also observed in the corresponding transition structure. The existence of this hydrogen bond prevents the formation of an additional hydrogen bond with Arg109. As discussed in the analysis of transition states trajectories the positioning of these two residues is highly correlated and they seem to alternate in the transition state stabilization. Finally, it is interesting to comment on a particular feature observed in REAC2 simulation. In this case the substrate adopts a slightly different orientation in the active site. Effectively, in this simulation the pyruvate molecule is slightly rotated around a perpendicular axes to the molecular plane in such a way that a NH_2 moiety of Arg171 is placed in between a carboxylate oxygen and the carbonyl oxygen. In this case we can then occasionally observe a hydrogen bond interaction between O_{pyr} and Arg171, as reflected in the reduced averaged distance provided in Table II. As a result of all the interactions described between the pyruvate molecule and the enzyme, we can conclude than in the first reactant state, REAC1, the substrate molecule is less constrained and can displace more freely in the active site. Apart of the hydrogen bond interaction with His195, the carbonyl oxygen atom of the rest of complexes analyzed also display interactions with Asn140 or Arg171.

In addition to the reported enzyme-substrate interactions, we have found differences in the interactions with water molecules. Using the CONTACT keyword of CHARMM we computed the average number of water molecules hydrogen bonded to the QM atoms of our system. This averaged number was close to 8 for the simulations corresponding to REAC1, REAC2 and REAC4, but it was 10.5 for REAC3. The standard deviation in all cases was of about ± 1 . One of the additional water molecules in REAC3 is hydrogen bonded to the O_{pyr} atom of the substrate. This interaction is found in 32% of the configurations analyzed and for this subset the $O_{\text{pyr}}H_w$ averaged distance was 2.1 Å (± 0.2). An interaction of this type between the substrate and water molecules was found for only a 3% of structures corresponding to the simulation of REAC1 and much less frequently in REAC2 and REAC4 simulations. The $O_{\text{pyr}}H_w$ distance averaged for the whole simulation is also provided in Table II. The other additional water molecule found in REAC3 simulation is placed in the solvation shell of the ribose ring of NADH.

We also found significant differences among the reactants simulations analyzing the behaviour of key enzyme residues, in particular Arg109. As explained before this residue is part of the flexible loop closing the active site. **Figure 4** shows the time evolution of the CA-CB-CD-CG dihedral angle of Arg109 during the 500 ps simulations of the four reactant complexes. While in REAC2 and REAC4 this dihedral angle fluctuates around 90 degrees, in the case of REAC1 and REAC3 it is close to 180. This orientation of the Arg109 side-chain is kept during all the simulation. Accordingly, this residue is able to establish different interactions in each case. **Table III** provides the averaged hydrogen bond distances and standard deviation between Arg109 and some residues and water molecules. Both in REAC2 and REAC4 we have observed a water molecule hydrogen bonded to Arg109 during the simulation. This feature was not observed in REAC1 and REAC3 simulations, what can be related to the different disposition of the side-chain. In addition, Arg109 also forms hydrogen bonds with Gln103 and Asp181 in REAC2. In REAC3 we observed a more frequent contact with Tyr224. In the case of REAC1 Arg109 forms a

strong hydrogen bond with Asp181, which is conserved during the whole simulation. In the simulation of REAC3 we observe occasional hydrogen bonds of Arg109 either with Asp181 and Tyr224.



The Arg109-Gln103 interaction, observed in the simulation corresponding to the REAC2 complex deserves some especial attention. This interaction appears between two residues of the flexible loop closing the active site. The existence of this interaction leads to a slightly different disposition of the loop, which can have dramatic consequences on the reaction energetics. The positioning of the flexible loop at the end of the molecular dynamic simulations of the four reactant complexes is represented in **Figure 5. Table III** also provides the energy effect on the potential energy barrier, obtained as a snapshot difference between the gas phase energy barrier calculated with and without the presence of Arg109 at the AM1 level. Calculations were carried out on the optimized transition structures and the corresponding reactant structures localized after IRC following. In all cases Arg109 plays a stabilizing role of the transition structure relative to the reactants. However, there are significant differences between the four reaction paths. While the first attains the maximum stabilization energy (9.7 Kcal/mol), in the second

(the one corresponding to REAC2) Arg109 plays a minor role (just 1.0 Kcal/mol). REAC3 and REAC4 display intermediate values (6.7 and 8.2 Kcal/mol respectively). Comparing this Arg109 effect with the total potential energy barrier (see Table I) we can observe a good correspondence between the two sets of values. The maximum stabilization effect corresponds to the minimum energy barrier, while the minimum Arg109 effect corresponds to the higher energy barrier. Thus, the positioning of Arg109 is a key element to determine the reaction rate. This residue belongs to an enzymatic flexible loop and it seems that different dispositions can be attained, determining the existence of different reactant valleys from which different reaction paths can be followed.

Table III. Averaged distances (in Å) between Arg109 and other residues and water molecules. The stabilizing effect (see text) of this residue on the potential energy barrier (in Kcal/mol) is also provided.

	REAC1	REAC2	REAC3	REAC4
d(HH11 _{Arg109} -OD1 _{Asp181})	5.2 (0.2)	3.5 (1.3)	7.1 (0.8)	8.2 (0.8)
d(HH12 _{Arg109} -OD1)	5.4 (0.3)	3.2 (1.6)	6.5 (1.5)	8.1 (1.1)
d(HH21 _{Arg109} -OD1)	2.6 (0.7)	5.5 (0.6)	4.2 (0.6)	6.0 (0.5)
d(HH22 _{Arg109} -OD1)	4.2 (0.6)	4.8 (1.0)	4.8 (1.3)	6.9 (0.6)
d(HE _{Arg109} -OE1 _{Gln103})	5.1 (0.3)	3.7 (1.7)	5.4 (0.3)	6.2 (0.9)
d(HH _{Arg109} -OH _{Tyr239})	3.2 (0.4)	3.3 (0.8)	2.8 (0.8)	3.0 (0.4)
d(HH _{Arg109} -O _w)	6.5 (0.5)	2.5 (0.7)	4.7 (0.5)	2.7 (0.5)
ΔE_{stab}	9.7	1.0	6.7	8.2

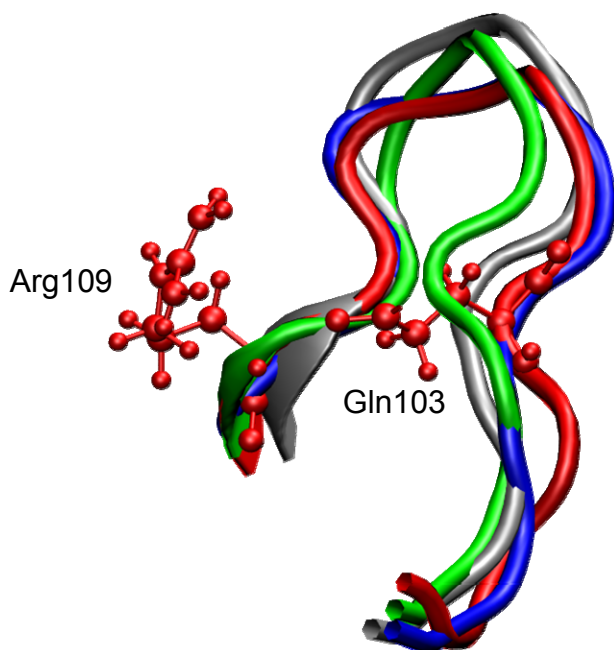


Figure 5. Overlapping of the flexible loop at the end of the four reactant complexes simulations. The colour code is the same as in Figure 4. The Arg109-Gln103 interaction in REAC2 is also shown.

3.3. Free Energy Calculations

As discussed in the introduction, the fact that different reactant complexes may contribute to the observed reaction flux could be incorporated by means of sufficiently long sampling trajectories. Considering that the conversion rate between two different reactant complexes could be even slower than the reaction rate this is not an efficient way to explore the different reaction channels except in very simple cases. Another possibility is to identify some of the reactant complexes contributing to the reaction and then obtain the free energy profile corresponding to each case by means of much shorter simulations. Finally the obtained free energy barrier should be conveniently

combined depending on the relative values between the interconversion and reaction rates.

In order to obtain free energy profiles corresponding to the pyruvate to lactate transformation for each of the reactant valleys analyzed above we first tried to use the umbrella sampling technique. This technique allows, in principle, for a correct sampling of all the degrees of freedom of the system except, obviously, the reaction coordinate. We tried to trace the free energy profiles using umbrella sampling with different definitions of the reaction coordinate of the type:

$$RC = \alpha_1 r_{CnicH} - \alpha_2 r_{CpyrH} + \alpha_3 r_{NhisH} - \alpha_4 r_{OpyrH} \quad (6)$$

However, the results were quite discouraging. While we were able to control the behaviour of the hydride transfer (monitoring the corresponding transfer coordinate we observed a soft and continuous variation as the reaction coordinate was changed), the proton transfer invariably took place suddenly, being transferred from His195 to pyruvate in a short time interval of a single window. Attempts to control the variation of the proton transfer coordinate changing the values of the α_i weights appearing in the definition of the global reaction coordinate were always failed. The free energy profiles obtained in this way displayed unexpectedly high reaction free energy barriers (corresponding to exploration of high energy regions of the PES) and discontinuous variations along the reaction coordinate. We then decided to employ a different strategy based on the use of the Free Energy Perturbation (FEP) technique. The configurations for which the free energy difference is estimated correspond to those structures obtained along the IRC calculation and are thus characterized by a single reaction coordinate:

$$s = \sqrt{\sum_{i \in QM} m_i [(x_i - x_{0i})^2 + (y_i - y_{0i})^2 + (z_i - z_{0i})^2]} \quad (7)$$

where x_0, y_0, z_0 are the Cartesian coordinates of the QM atoms in the transition structure while x_i, y_i and z_i are the coordinates of a structure belonging to the IRC traced from this transition structure and m_i are the masses of the atoms. Within this treatment the free energy relative to the

reactant structure can be approximately expressed as a function of the s coordinate as:

$$\begin{aligned} \Delta G_{\text{FEP}}(s^R \rightarrow s^j) &\approx \Delta E_{\text{QM}}^{0,R \rightarrow j} + \Delta G_{\text{QM/MM}}^{R \rightarrow j} = \\ &= \left(E_{\text{QM}}^0(s^j) - E_{\text{QM}}^0(s^R) \right) - k_B T \sum_{i=R}^{i=j-1} \ln \left\langle \exp \left(E_{\text{QM/MM}}(s^{i+1}) - E_{\text{QM/MM}}(s^i) \right) \right\rangle_{\text{MM},i} \end{aligned} \quad (8)$$

where E_{QM}^0 is the gas phase energy of the QM subsystem, including the bidimensional correction depending on the R_1 and R_2 coordinates. The QM/MM interaction contribution to the free energy difference between two different configurations of the system is obtained by averaging the QM/MM interaction energy (including the polarization energy) over all the MM coordinates of the system obtained for a particular value of the s coordinate. For this reason we added a term to include the contribution of the fluctuations of all these coordinates (except s) to the free energy profile. This contribution is estimated assuming a quantum harmonic treatment for the vibrational modes of the QM subsystem in the field created by the rest of the enzyme. Then, the quasiclassical free energy profile can be obtained adding the contribution of the QM vibrational degrees of freedom as:

$$\begin{aligned} G(s^j) - G(s^R) &= \Delta E_{\text{QM}}^{0,R \rightarrow j} + \Delta G_{\text{QM/MM}}^{R \rightarrow j} + \Delta G_{\text{fluc,QM}} = \\ &= \Delta E_{\text{QM}}^{0,R \rightarrow j} + \Delta G_{\text{QM/MM}}^{R \rightarrow j} + \sum_{x=1}^{3N-7} \frac{1}{2} h \nu_x^j + k_B T \ln \left(1 - e^{-\frac{h \nu_x^j}{k_B T}} \right) - G_{\text{fluc,QM}}(s^R) \end{aligned} \quad (9)$$

Pseudotraslations and pseudorotations of the QM subsystem are not expected to contribute significantly to the free energy differences and then are not considered. In addition, the frequencies are obtained projecting out the contribution of the reaction coordinate in the Hessian for all the structures but the reactant state. Thus, the vibrational term is obtained after summation over $3N-6$ vibrational modes (being N the number of QM atoms) in the reactant structure and $3N-7$ for the rest.⁷⁹ The free energy profile obtained in this way corresponds to a Potential of Mean Force because, within the approximations described, it can be directly translated to the

probability of finding the system at a particular value of the s coordinate. The quasiclassical activation free energy is then obtained as:

$$\Delta G^\ddagger = G(s^\ddagger) - G(s^R) \quad (10)$$

Where s^\ddagger indicates the maximum of the free energy profile obtained according to equation 9. To obtain the free energy profile we run MD simulations for the structures appearing along each reaction path. Each simulation window consisted in 2ps of equilibration and 10 ps of production. We used up to 50 windows for each reaction path and employed the double wide sampling technique. The free energy profiles obtained in this way for the four reaction paths analyzed here are presented in **Figure 6** while the activation free energies are given in **Table IV**. The activation free energies obtained in this quasiclassical approximation do not exactly correlate with the potential energies obtained from the reaction path following. While the potential energy barrier ordering was $TS4 > TS3 > TS2 > TS1$, the free energy ordering is $FEP2 > FEP4 > FEP3 > FEP1$. In addition, to this qualitative difference, the quantitative variation is quite significant for $TS3$ and $TS4$ (more than 15 Kcal/mol in the last case). This fact clearly underlines the importance of considering the fluctuations of the environment to evaluate the energetics of the reaction. The term related to the fluctuations of the QM subsystem does not introduce additional changes in the relative ordering among the reaction paths but, nonetheless, the quantitative importance is out of discussion. This term can amount up to more than 3 Kcal/mol.

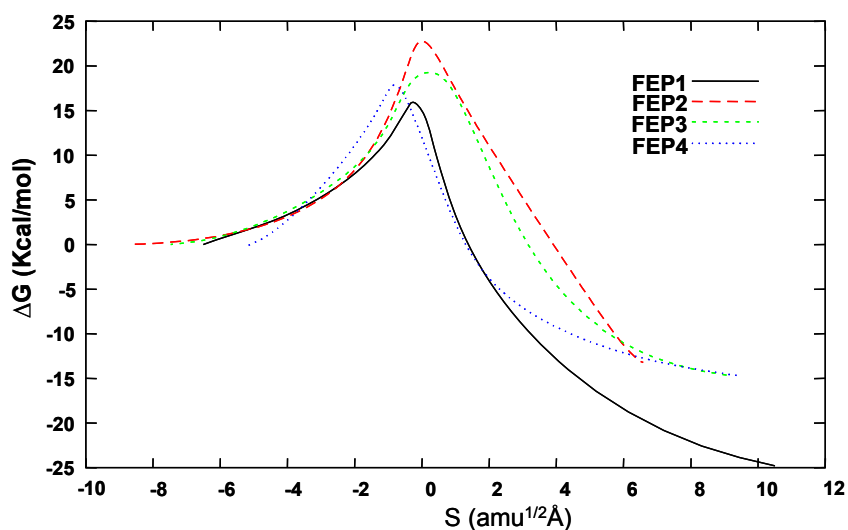


Figure 6. Quasiclassical free energy profiles obtained for the four reaction paths.

Table IV. Contributions to the activation free energy (in Kcal/mol), transmission coefficients and catalytic rate constants (in s^{-1}) obtained for each of the four reaction paths analyzed.

	FEP1	FEP2	FEP3	FEP4
ΔG_{FEP}	16.7	23.5	19.4	20.6
$\Delta G_{\text{fluc,QM}}$	-3.1	-2.5	-2.6	-3.2
ΔG^\ddagger	13.6	21.0	16.8	17.4
κ	4.116	4.051	3.879	3.762
ΔG_{tun}	-0.84	-0.83	-0.80	-0.78
$\Delta G_{\text{eff}}^\ddagger$	12.8	20.2	16.0	16.6
k_{cat}	$2.77 \cdot 10^3$	$9.50 \cdot 10^{-3}$	$1.01 \cdot 10^1$	$4.02 \cdot 10^0$

3.4. Tunnelling Transmission Coefficient

As explained before we used EA-VTST-OMT methodology to obtain the tunnelling transmission coefficient. In this way we try to account for quantum tunnelling and non-classical diffractive reflections. Contributions of recrossing are not considered here. As far as we follow the reaction paths using a quite complete reaction coordinate (defined by all the atoms of the QM subsystem) the corresponding transmission factor should be maximized and then it should be close to unity.

We finally evaluated the contribution of tunnelling to the free energy barrier

$$\text{as } \Delta G_{\text{tun}} = -RT \ln \kappa \quad (11)$$

and the final estimation of the effective free energy barrier is obtained as:

$$\Delta G_{\text{eff}}^{\ddagger} = -RT \ln k_{\text{cat}} = \Delta G_{\text{tun}} + \Delta G^{\ddagger} \quad (12)$$

The value of the transmission coefficient and its free energy equivalency for each of the four reaction path are given in Table IV. In free energy terms tunnelling contributes by about 0.8 Kcal/mol to the free energy barrier reduction and there are no significant differences between the four reaction paths analyzed. The maximum difference amounts to only 0.06 Kcal/mol.

3.5. Evaluation of the Rate Constant

A simple arithmetic average of the four reaction path rate constants presented in Table IV gives $6.96 \cdot 10^2 \text{ s}^{-1}$, a value very close to the experimental lower limit estimation. However, this must be considered only as a rough indication of the goodness of the description provided by the computational model. Weighting among different reactant complexes to obtain the macroscopic rate constant is in fact a very delicate question. If the interconversion rate among the different complexes is much lower than the reaction rate, then Otherwise, when the interconversion rate is much faster than the reaction rate then equilibrium may be assumed and the individual rate constant must then be weighted using the free energy differences

through the Boltzmann factors. Obviously, any intermediate behaviour is also possible.^{39,40}

Evaluation of the relative free energy differences among the four reactant complexes studied here is a difficult task because of the subtle differences appearing among them. We have been not able to find a single geometrical parameter (or combination of some of them) to follow the transformations among these reactant complexes and the direct evaluation of the free energy differences is, by the moment, out of our possibilities. As an indication of their relative stabilities we have computed the averaged potential energies as well as the interaction energies of the enzyme with NADH plus pyruvate, this is the interaction energy of the enzyme with the cofactor and the substrate. With this purpose we selected 5000 equally distributed snapshots from the 500 ps MD simulations of the reactant complexes. For these structures we redefined NADH and pyruvate as the QM subsystem and then we computed the potential energy of the full QM/MM system and the energy of non-interacting QM and MM subsystems. The interaction energy was then calculated as the difference between these two quantities. Averaged values and the standard deviation are given in **Table V**. It is a well-known fact that computation of internal energies and entropies is accompanied by much larger uncertainties than free energies, as reflected by the standard deviations given in Table V.

Table V. Averaged potential energies and interaction energies between the enzyme and NADH plus pyruvate (in Kcal/mol) and their respective standard deviation (in parenthesis) for the four reactant complexes analyzed.

	REAC1	REAC2	REAC3	REAC4
E	-12181 (34)	-12241 (13)	-12265 (37)	-12261 (44)
$E_{\text{QM/MM}}$	-211 (51)	-329 (52)	-411 (44)	-322 (48)

As a further evidence of the different characteristics of the reactant complexes analyzed we have obtained quite large differences in the averaged interaction energies between the enzyme and the cofactor plus the

substrate. These differences amount up to 200 Kcal/mol and are partly compensated by the enzyme self-energy, as reflected in the much smaller differences in the total potential energy (up to 80 Kcal/mol). It thus seems that in order to gain interaction energy with the cofactor and the substrate the enzyme must be deformed increasing the enzymatic energy term.

According to these results, the most reactive reactant complex (REAC1) also presents the lower stability (in potential energy terms). It is thus tempting to conclude that LDH does not work under the equilibrium regime. Obviously, relative free energies (and also free activation free energies) would be needed in order to conclude definitively this question but, in addition to the potential energy analysis, there are other facts supporting this idea. We have shown that the differences among the four reactant complexes can be related to the disposition of the flexible loop closing the active site. The way in which this loop is folded may determine the interactions established between the substrate and some key residues (such as Arg109 or Asn140) or the number of water molecules found in the active site. Thus, we can imagine a picture of LDH catalysis where the closure of the flexible loop over the active site can lead to different reactant complexes each of them presenting different inherent reactivity. In any case, there has probably not been any evolutive pressure to reduce the rate constant of the chemical step (or the macroscopic average) under the value of the rate-limiting step (the loop closure over the active site). The stability of these complexes can also be very different depending in the way in which the loop has been folded. In any case interconversion among them could not be an easy task. One way to go from a reactant complex to another is obviously through the unfolding and refolding of this loop, but we must take into account that this movement is in fact the rate determining step in the catalytic process and then the process could be slower than the chemical step.

3.6 *Kinetic Isotope Effects*

We have calculated the kinetic isotope effects corresponding to the use of NAD²H instead of NAD¹H as cofactor. This is, the transferred hydride is changed from ¹H to ²H. The kinetic isotope effects are obtained as the ratio between the rate constants obtained using NAD²H and NAD¹H:

$$\text{KIE} = \frac{k_{\text{cat}}^{\text{H}}}{k_{\text{cat}}^{\text{D}}} \quad (13)$$

To obtain the rate constant corresponding to NAD²H we just needed to recalculate the contribution of the QM subsystem vibrations to the free energy profile and the tunnelling transmission coefficient. The classical free energy profile is independent on the nuclear masses and then it takes the same value when using either NAD¹H or NAD²H. The calculations were carried out for the four reaction paths and the results are summarized in **Table VI**.

Table VI. Free energy contributions (in Kcal/mol), tunnelling transmission coefficients and rate constants (in s⁻¹) for four possible reaction paths corresponding to ²H hydride transfer. Kinetic Isotope Effects are also provided.

	FEP1	FEP2	FEP3	FEP4
$\Delta G_{\text{fluc,QM}}^{\text{D}}$	-2.5	-1.7	-1.8	-2.1
ΔG^{\ddagger}	14.2	21.8	17.6	18.0
κ^{D}	2.706	2.963	2.761	2.879
$k_{\text{cat}}^{\text{D}}$	$6.62 \cdot 10^2$	$1.88 \cdot 10^{-3}$	$2.04 \cdot 10^0$	$1.16 \cdot 10^0$
KIE	4.19	5.05	4.98	3.47

In all the cases, substitution of NAD¹H by NAD²H provokes an increase of the quasiclassical activation free energy and a reduction of the tunnelling transmission coefficient. The final result is a reduction of the catalytic rate constant. The obtained kinetic isotope effects range between 3.47 and 5.05, in good agreement with the experimental estimations.

4. Conclusions

Lactate Dehydrogenase catalyzes the reversible transformation of pyruvate into lactate. The chemical step consists in a hydride and a proton transfer from the cofactor (NADH) and a protonated histidine (His195). Our results indicate that the two transfers take place in a single concerted process although the degree of asynchronicity may depend on the particular reaction path followed. We have been able to locate different transition structures which have been connected through reaction path following to their respective reactant structures. Four of these reactant complexes have been analyzed in detail by means of relatively long molecular dynamic simulations. We have found some differences among these complexes related to the positioning of the flexible loop closing the active site. This is translated to a different interaction pattern of the enzyme with the substrate, a different number of water molecules in the active site and other differences related to the key residue Arg109.

A practical consequence of the existence of these different reactant conformations is that very long simulation times could be required to obtain significant free energy profiles where the most contributing energy paths were correctly sampled. We then decided to obtain different free energy profiles for each of the four reactant complexes analyzed in detail. In addition, because the concerted hydride and proton transfer involves a complicated geometrical coordinate we employed a free energy perturbation technique adapted to follow intrinsic reaction coordinates. Fluctuations of the quantum subsystem are incorporated through a harmonic normal mode approximation while tunnelling contributions are included using an adaptation of the VTST-OMT methodology. The four reaction paths analyzed display significant differences in the calculated rate constants. A simple arithmetic average of the four reaction path rate constants gives $6.96 \cdot 10^2 \text{ s}^{-1}$, a value very close to the experimental lower limit estimation. However, averaging among different reactant complexes to obtain the macroscopic rate constant is not an obvious question. To obtain a definitive answer one

would need to trace the free energy profiles connecting them. Our estimated the potential energies display significant differences but can not be taken as conclusive. Anyway, taking into account that the movement of the loop is in fact the rate-limiting step in the global reaction we can imagine an scenario where the reactant complexes interconversion is slower that the chemical reaction and then equilibrium can not be assumed. We have also calculated kinetic isotope effects for each of the four reaction paths. Our calculated values range between 3.47 and 5.05 in good agreement with experimental determinations.

5. Acknowledgements

We thank the DGI for projects DGI BQU2003-04168-C03 and BQU-2002-00301, BANCAIXA for project P1A99-03 and Generalitat Valenciana for project GV01-324, GV04B-21, GV04B-131 and GRUPOS04/28. S. F. acknowledges a doctoral fellowship of the Ministerio de Ciencia y Tecnología. S. F. also acknowledges the Ministerio for financial support during her visit to the UAB and the warm hospitality of all the people of the Departament de Química of the Universitat Autònoma de Barcelona.

References

- (1) Wolfenden, R.; Snider, M. J. *Accounts of Chemical Research* **2001**, *34*, 938.
- (2) Gao, J. L.; Truhlar, D. G. *Annual Review of Physical Chemistry* **2002**, *53*, 467.
- (3) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186.
- (4) Marti, S.; Roca, M.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *Chemical Society Reviews* **2004**, *33*, 98.
- (5) Shurki, A.; Villa, J.; Warshel, A. *Abstr. Pap. Am. Chem. S.* **2002**, *223*, C59.
- (6) Villa, J.; Strajbl, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. *P. Natl. Acad. Sci. USA* **2000**, *97*, 11899.
- (7) Warshel, A. *Abstracts of Papers of the American Chemical Society* **1989**, *197*, 101.
- (8) Warshel, A.; Villa, J.; Strajbl, M.; Florian, J. *Abstracts of Papers of the American Chemical Society* **2001**, *221*, U399.
- (9) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *P. Natl. Acad. Sci. USA* **2001**, *98*, 9032.
- (10) Ma, J. P.; Zheng, X. F.; Schnappauf, G.; Braus, G.; Karplus, M.; Lipscomb, W. N. *P. Natl. Acad. Sci. USA* **1998**, *95*, 14640.
- (11) Almarsson, O.; Bruice, T. C. *Journal of the American Chemical Society* **1993**, *115*, 2125.
- (12) Hur, S.; Bruice, T. C. *Journal of the American Chemical Society* **2003**, *125*, 10540.
- (13) Hur, S.; Bruice, T. C. *Journal of the American Chemical Society* **2003**, *125*, 1472.
- (14) Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Mol. Struc.-Theochem.* **2003**, *632*, 297.
- (15) Mulholland, A. J. *Drug Discov. Today* **2005**, *10*, 1393.
- (16) Roca, M.; Moliner, V.; Ruiz-Pernia, J. J.; Silla, E.; Tunon, I. *Journal of Physical Chemistry A* **2006**, *110*, 503.
- (17) Soriano, A.; Silla, E.; Tunon, I.; Marti, S.; Moliner, V.; Bertran, J. *Theoretical Chemistry Accounts* **2004**, *112*, 327.
- (18) Garcia-Viloca, M.; Alhambra, C.; Truhlar, D. G.; Gao, J. L. *Journal of Computational Chemistry* **2003**, *24*, 177.
- (19) Truhlar, D. G.; Gao, J. L.; Garcia-Viloca, M.; Poulsen, T.; Pu, J. Z.; Alhambra, C.; Corchado, J. C.; Sanchez, M. L. *Abstracts of Papers of the American Chemical Society* **2002**, *223*, C54.
- (20) Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Phys. Chem. B* **2005**, *109*, 21089.
- (21) Rosta, E.; Klahn, M.; Warshel, A. *J. Phys. Chem. B* **2006**, *110*, 2934.
- (22) Olsson, M. H. M.; Warshel, A. *Journal of the American Chemical Society* **2004**, *126*, 15167.

- (23) Garrett, B. C.; Truhlar, D. G. *Abstracts of Papers of the American Chemical Society* **1980**, *180*, 233.
- (24) Garrett, B. C.; Truhlar, D. G.; Grev, R. S.; Magnuson, A. W.; Connor, J. N. L. *Journal of Chemical Physics* **1980**, *73*, 1721.
- (25) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *Journal of Chemical Physics* **2003**, *119*, 5828.
- (26) Roca, M.; Andres, J.; Moliner, V.; Tunon, I.; Bertran, J. *Journal of the American Chemical Society* **2005**, *127*, 10648.
- (27) Soriano, A.; Silla, E.; Tunon, I.; Ruiz-Lopez, M. F. *Journal of the American Chemical Society* **2005**, *127*, 1946.
- (28) Nam, K.; Prat-Resina, X.; Garcia-Viloca, M.; Devi-Kesavan, L. S.; Gao, J. L. *Journal of the American Chemical Society* **2004**, *126*, 1369.
- (29) Truhlar, D. G.; Gao, J. *Acc. Chem. Res.* **2006**, *en imprenta*.
- (30) Roca, M.; Moliner, V.; Tuñón, I.; Hynes, J. T. *J. Am. Chem. Soc.* **2006**, *ASAP*.
- (31) Neria, E.; Karplus, M. *Chemical Physics Letters* **1997**, *267*, 23.
- (32) Staib, A.; Borgis, D.; Hynes, J. T. *Journal of Chemical Physics* **1995**, *102*, 2487.
- (33) Alhambra, C.; Corchado, J.; Sanchez, M. L.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2001**, *105*, 11326.
- (34) Agarwal, P. K.; Billeter, S. R.; Hammes-Schiffer, S. *Journal of Physical Chemistry B* **2002**, *106*, 3283.
- (35) Garcia-Viloca, M.; Truhlar, D. G.; Gao, J. L. *Biochemistry* **2003**, *42*, 13558.
- (36) Alhambra, C.; Sanchez, M. L.; Corchado, J. C.; Gao, J.; Truhlar, D. G. *Chemical Physics Letters* **2002**, *355*, 388.
- (37) Hammes-Schiffer, S. *Biochemistry* **2002**, *41*, 13335.
- (38) Ferrer, S.; Silla, E.; Tunon, I.; Oliva, M.; Moliner, V.; Williams, I. H. *Chemical Communications* **2005**, 5873.
- (39) Min, W.; English, B. P.; Luo, G. B.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. *Accounts Chem. Res.* **2005**, *38*, 923.
- (40) Antikainen, N. M.; Smiley, R. D.; Benkovic, S. J.; Hammes, G. G. *Biochemistry* **2005**, *44*, 16835.
- (41) Tan, W. H.; Yeung, E. S. *Anal. Chem.* **1997**, *69*, 4242.
- (42) Xue, Q. F.; Yeung, E. S. *Nature* **1995**, *373*, 681.
- (43) Thorpe, I. F.; Brooks, C. L. *Journal of the American Chemical Society* **2005**, *127*, 12997.
- (44) English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G. B.; Sun, H. Y.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. *Nat. Chem. Biol.* **2006**, *2*, 87.
- (45) Ansari, A.; Berendzen, J.; Bowne, S. F.; Frauenfelder, H.; Iben, I. E. T.; Sauke, T. B.; Shyamsunder, E.; Young, R. D. *P. Natl. Acad. Sci. USA* **1985**, *82*, 5000.
- (46) Thorpe, I. F.; Brooks, C. L. *J. Phys. Chem. B* **2003**, *107*, 14042.

- (47) Zhang, Y. K.; Kua, J.; McCammon, J. A. *J. Phys. Chem. B* **2003**, *107*, 4459.
- (48) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645.
- (49) Basner, J. E.; Schwartz, S. D. *J. Am. Chem. Soc.* **2005**, *127*, 13822.
- (50) Clarke, A. R.; Wigley, D. B.; Barstow, D. A.; Chia, W. N.; Waldman, A. D. B.; Hart, K. W.; Atkinson, T.; Holbrook, J. J. *Biochemical Society Transactions* **1987**, *15*, 152.
- (51) Badcoe, I. G.; Smith, C. J.; Wood, S.; Halsall, D. J.; Holbrook, J. J.; Lund, P.; Clarke, A. R. *Biochemistry* **1991**, *30*, 9195.
- (52) Clarke, A. R.; Wilks, H. M.; Barstow, D. A.; Atkinson, T.; Chia, W. N.; Holbrook, J. J. *Biochemistry* **1988**, *27*, 1617.
- (53) Deng, H.; Zheng, J.; Clarke, A.; Holbrook, J. J.; Callender, R.; Burgner, J. W. *Biochemistry* **1994**, *33*, 2297.
- (54) Leach, S. J. *Physical principles and techniques of protein chemistry*; Academic P: New York ; London, 1969.
- (55) Brown, W. M.; Yowell, C. A.; Hoard, A.; Jagt, T. A. V.; Hunsaker, L. A.; Deck, L. M.; Royer, R. E.; Piper, R. C.; Dame, J. B.; Makler, M. T.; Jagt, D. L. V. *Biochemistry* **2004**, *43*, 6219.
- (56) Kavanagh, K. L.; Elling, R. A.; Wilson, D. K. *Biochemistry* **2004**, *43*, 879.
- (57) Chaikuad, A.; Fairweather, V.; Conners, R.; Joseph-Horne, T.; Turgut-Balik, D.; Brady, R. L. *Biochemistry* **2005**, *44*, 16221.
- (58) Clarke, A. R.; Wigley, D. B.; Chia, W. N.; Barstow, D.; Atkinson, T.; Holbrook, J. J. *Nature* **1986**, *324*, 699.
- (59) Clarke, A. R.; Waldman, A. D. B.; Hart, K. W.; Holbrook, J. J. *Biochimica Et Biophysica Acta* **1985**, *829*, 397.
- (60) Jackson, R. M.; Gelpi, J. L.; Cortes, A.; Emery, D. C.; Wilks, H. M.; Moreton, K. M.; Halsall, D. J.; Sleigh, R. N.; Behanmartin, M.; Jones, G. R.; Clarke, A. R.; Holbrook, J. J. *Biochemistry* **1992**, *31*, 8307.
- (61) Kotik, M.; Zuber, H. *Biochemistry* **1992**, *31*, 7787.
- (62) Seravalli, J.; Huskey, W. P.; Schowen, K. B.; Schowen, R. L. *Pure and Applied Chemistry* **1994**, *66*, 695.
- (63) Madern, D.; Ebel, C.; Mevarech, M.; Richard, S. B.; Pfister, C.; Zaccari, G. *Biochemistry* **2000**, *39*, 1001.
- (64) Read, J. A.; Winter, V. J.; Eszes, C. M.; Sessions, R. B.; Brady, R. L. *Proteins-Structure Function and Genetics* **2001**, *43*, 175.
- (65) Turner, A. J.; Moliner, V.; Williams, I. H. *Physical Chemistry Chemical Physics* **1999**, *1*, 1323.
- (66) Moliner, V.; Turner, A. J.; Williams, I. H. *Chemical Communications* **1997**, 1271.
- (67) Ferrer, S.; Ruiz-Pernia, J. J.; Tunon, I.; Moliner, V.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. *Journal of Chemical Theory and Computation* **2005**, *1*, 750.
- (68) Basner, J. E.; Schwartz, S. D. *J. Phys. Chem. B* **2004**, *108*, 444.

- (69) Fukui, K. *Accounts of Chemical Research* **1981**, *14*, 363.
- (70) Truhlar, D. G.; Gao, J. L.; Alhambra, C.; Corchado, J.; Garcia-Viloca, M.; Sanchez, M. L.; Kesavan, L. D.; Poulsen, T.; Villa, J. *Abstracts of Papers of the American Chemical Society* **2002**, *224*, U480.
- (71) Truhlar, D. G.; Gao, J. L.; Garcia-Viloca, M.; Alhambra, C.; Corchado, J.; Sanchez, M. L.; Poulsen, T. D. *International Journal of Quantum Chemistry* **2004**, *100*, 1136.
- (72) Wigley, D. B.; Gamblin, S. J.; Turkenburg, J. P.; Dodson, E. J.; Piontek, K.; Muirhead, H.; Holbrook, J. J. *J. Mol. Biol.* **1992**, *223*, 317.
- (73) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J. *J. Comput. Chem* **1983**, *4*, 187.
- (74) Gilson, M. K. *Proteins-Structure Function and Genetics* **1993**, *15*, 266.
- (75) Antosiewicz, J.; Mccammon, J. A.; Gilson, M. K. *Journal of Molecular Biology* **1994**, *238*, 415.
- (76) Field, M. J.; David, L.; Rinaldo, D., personal communication
- (77) Freindorf, M.; Gao, J. *J. Comp. Chem* **1996**, *17*, 386.
- (78) Martí, S.; Moliner, V.; Tuñón, I. *Journal of Chemical Theory and Computation* **2005**, *1*, 1008.
- (79) The vibrational contribution of the QM atoms to the free energy profile was calculated for all the structures of each reaction path and afterwards fitted to a polynomial as a function of the s coordinate. In this process we took into account that the first steps in the IRC were traced following the transition vector and then the structures closer to the transition structure were not really stationary after projecting out the component of the reaction path coordinate.