



**Universitat**  
de les Illes Balears

**TESIS DOCTORAL**  
**2017**

**“CÁLCULO DEL  $pK_a$  A PARTIR DE LA  
REACCIÓN ISODÉSMICA. APLICACIÓN AL  
ESTUDIO DE LA GLICACIÓN PROTEICA”**

**Sebastián Sastre Tenreiro**





**Universitat**  
de les Illes Balears

**TESIS DOCTORAL**  
**2017**

**Programa de Doctorado en Ciencia y Tecnologia**  
**Química**

**“CÁLCULO DEL  $pK_a$  A PARTIR DE LA  
REACCIÓN ISODÉSMICA. APLICACIÓN AL  
ESTUDIO DE LA GLICACIÓN PROTEICA”**

**Sebastián Sastre Tenreiro**

**Director: Juan Frau Munar**

**Director: Rodrigo Casasnovas Perera**

**Doctor/a por la Universitat de les Illes Balears**



Los Directores de la tesis Dr. Juan Frau Munar, Catedrático del Área de Química Física del Departamento de Química de la Universitat de les Illes Balears, el Dr. Rodrigo Casanovas Perera, y el doctorando Sebastián Sastre Tenreiro.

DAN CONSTANCIA DE:

Que el presente trabajo titulado “Cálculo del  $pK_a$  a partir de la reacción Isodésmica. Aplicación al estudio de la glicación proteica” ha sido realizado bajo su dirección en el Departamento de Química (Área de Química Física) de la Universitat de les Illes Balears por Sebastián Sastre Tenreiro durante los años 2012 - 2017, constituyendo la Memoria que presenta para optar al grado de Doctor.

Fdo. Juan Frau Munar

Fdo. Rodrigo Casanovas Perera

Fdo. Sebastián Sastre Tenreiro

Palma de Mallorca, Septiembre de 2017.



## Artículos

### Artículo 1

“Isodesmic reaction for  $pK_a$  calculations of common organic molecules “.

Sastre, S., Casasnovas, R., Muñoz F., Frau, J. *Theor. Chem. Acc.*, **2013**, 132, 1310

### Artículo 2

“Isodesmic reaction for accurate theoretical  $pK_a$  calculations of amino acids and peptides”.

Sastre, S., Casasnovas, R., Muñoz, F., Frau, J. *Phys. Chem. Chem. Phys.*, **2016**, 18, 11202

### Artículo 3

“Computational prediction of the protonation sites of Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> peptides through conceptual DFT descriptors”.

Sastre, S., Frau, J., Glossman-Mitnik, D. *Molecules*, **2017**, 22, 458

### Artículo 4 (En preparación)

“Very reactive minor molar fraction protonation states account for macroscopic antioxidant activity of thiamine. A Density Functional Theory study”.

Sastre, S., Casasnovas, R., Glossman-Mitnik, D., Frau, J.

## Resumen

Se conocen con el nombre de reacciones de glicación, las reacciones que transcurren sin control enzimático entre los grupos amina de las biomoléculas y los grupos carbonilo de azúcares reductores, como por ejemplo la glucosa. Los procesos de glicación se han relacionado con diabetes, envejecimiento, Alzheimer y Parkinson. Estas reacciones modifican preferentemente los residuos lisina y arginina de las proteínas. Sin embargo, no todos estos residuos se glican homogéneamente y uno de los objetivos de esta tesis es investigar si el  $pK_a$  de estos residuos está relacionado con la tendencia a sufrir reacciones de glicación.

Se ha calculado el  $pK_a$  de moléculas orgánicas con grupos funcionales ácido base comunes empleando una reacción isodésmica y se ha determinado que dicha metodología ofrece ventajas significativas con respecto a los ciclos termodinámicos. Esta misma metodología se ha empleado para el cálculo de  $pK_a$ s de aminoácidos y péptidos. En este caso la reacción isodésmica proporciona resultados precisos para los  $pK_a$ s de los grupos  $\alpha$ -carboxílico,  $\alpha$ -amino y cadenas laterales de aminoácidos y péptidos, con desviaciones absolutas medias (MADs) de 1 unidad o inferior. Esta metodología es una alternativa viable a los ciclos termodinámicos ya que no considera estas especies en fase gas.

Se ha empleado una metodología alternativa al cálculo de  $pK_a$ s, DFT conceptual, para evaluar la acidez y reactividad frente a la glicación de diferentes residuos lisina en péptidos del tipo  $K(A)_nK$ . Se han empleado los funcionales de la familia Minnesota M11, M11L, MN12L, MN12SX, N12 y N12SX para el cálculo de diferentes descriptores de reactividad. En particular se ha observado que el descriptor dual condensado y las funciones de Parr permiten predecir cualitativamente los sitios reactivos preferentes de glicación y protonación y proporcionan una explicación sólida de la reactividad de estas moléculas.

Los productos de glicación pueden degradarse a los denominados AGEs (Advanced Glycation End-Products) por efecto de especies radicales reactivas de oxígeno (ROS), entre otros agentes. Una de las vías de inhibición de formación de AGEs se basa en la administración de moléculas con capacidad antioxidante que eliminan ROS del organismo. En esta tesis se estudia la capacidad antioxidante de la tiamina. Dado que esta molécula también presenta reactividad ácido-base, se ha calculado la capacidad antioxidante de cada uno de los estados de protonación posibles. Por otra parte, se ha empleado la metodología indicada anteriormente para determinar la abundancia relativa

de cada uno de dichos estados de protonación. A partir de la reactividad de cada estado de protonación y de su fracción molar, se ha podido estimar la capacidad antioxidante global en función del pH del medio. Los resultados indican que la tiamina elimina eficazmente el radical  $\bullet\text{OCH}_3$  independientemente del pH del medio según los mecanismos HAT (Hydrogen Atom Transfer) y RAF (Radical Adduct Formation). Sin embargo, la capacidad de eliminación de radicales peroxilo no es tan buena. Destaca que la actividad macroscópica de la tiamina se debe a especies no mayoritarias que presentan excepcionales constantes de reacción microscópicas.

## Resum

Es coneixen amb el nom de reaccions de glicació, aquelles reaccions que es donen sense control enzimàtic entre els grups amina de les biomolècules i els grups carbonil de sucres reductors, com per exemple la glucosa. Els processos de glicació s'han relacionat amb la diabetis, envelliment, Alzheimer i Parkinson. Les reaccions de glicació no enzimàtica modifiquen preferentment els residus lisina i arginina de les proteïnes. Però, no tots aquests residus es gliquen homogèniament i un dels objectius d'aquesta tesi és investigar si el  $pK_a$  d'aquests residus està relacionat amb la tendència a sofrir reaccions de glicació.

S'ha calculat el  $pK_a$  de molècules orgàniques amb grups funcionals àcid-base comuns fent ús de la reacció isodèsmica i s'ha determinat que l'esmentada metodologia mostra avantatges significatives respecte als cicles termodinàmics. Aquesta mateixa metodologia s'ha usat pel càlcul de  $pK_a$ s d'aminoàcids i pèptids. En aquest cas, la reacció isodèsmica proporciona resultats precisos per els  $pK_a$ s dels grups  $\alpha$ -carboxílic,  $\alpha$ -amino i cadenes laterals d'aminoàcids i pèptids, amb desviacions absolutes mitjanes (MADs) d'una unitat o inferior. Aquesta metodologia representa una alternativa viable als cicles termodinàmics ja que no es consideren aquestes espècies en fase gas.

També s'ha fet ús d'una metodologia alternativa al càlcul de  $pK_a$ s, DFT conceptual, per avaluar l'acidesa i la reactivitat en front de la glicació de diferents residus lisina en pèptids del tipus K(A)<sub>n</sub>K. A partir dels funcionals de la família Minnesota M11, M11L, MN12L, MN12SX, N12 i N12SX s'han calculat diferents descriptors de reactivitat. En particular s'ha observat que el descriptor dual condensat i les funcions de Parr prediuen qualitativament els llocs reactius preferencials de glicació i protonació, i proporcionen una explicació sòlida de la reactivitat d'aquestes molècules.

Els productes de glicació es poden degradar als denominats AGEs (Advanced Glycation End-Products) degut a espècies radicalàries reactives d'oxigen (ROS), entre d'altres agents. Una de les vies d'inhibició de la formació d'AGEs es basa en l'administració de molècules amb capacitat antioxidant que eliminen ROS de l'organisme. A aquesta tesi s'ha estudiat la capacitat antioxidant de la tiamina. Donat que aquesta molècula també presenta reactivitat àcid-base, s'ha calculat la capacitat antioxidant de cada un dels estats de protonació possibles. Per altra banda, s'ha fet ús de la metodologia indicada anteriorment per a determinar l'abundància relativa de cada un d'aquests estats de protonació. A partir de la reactivitat de cada estat de protonació i de la seva fracció molar, s'ha pogut estimar la capacitat antioxidant global en funció del pH del medi. Els resultats indiquen que la tiamina elimina eficaçment el radical  $\bullet\text{OCH}_3$ ,

independentment del pH del medi, amb els mecanismes HAT (Hydrogen Atom Transfer) i RAF (Radical Adduct Formation). Així i tot, la capacitat d'eliminació de radicals peroxil no és tan bona. Finalment, es destaca que l'activitat macroscòpica de la tiamina es deguda a espècies no majoritàries que presenten excepcionals constants de reacció microscòpiques..

## Summary

Non-enzymatic reactions between amino groups of biomolecules and carbonyl groups of reducing sugars, like glucose, are known as glycation reactions. Glycation processes have been linked with diabetes, ageing, Alzheimer's and Parkinson's diseases. These reactions preferentially modify lysine and arginine protein residues. However, the residues are not homogeneously glycated and one objective of this thesis is whether the  $pK_a$  of these residues is related with the inclination to undergo glycation reactions.

The  $pK_a$  of organic molecules with acid-base functional groups has been calculated by employing an isodesmic reaction. It has been determined that this methodology offers significant advantages with respect to thermodynamic cycles. The same methodology has been used for the  $pK_a$  calculation of amino acids and peptides. In this case, the isodesmic reaction yields precise results for the  $pK_a$  of the  $\alpha$ -carboxylic,  $\alpha$ -amino and sidechain groups of amino acids and peptides, with mean absolute deviations (MADs) of 1 unit or less. This methodology is a viable alternative to thermodynamic cycles because it does not consider these species in gas phase.

An alternative methodology to the  $pK_a$  calculations, conceptual DFT, has been used for the evaluation the acidity and reactivity against glycation of different lysine residues in  $K(A)_nK$  peptides. The functionals M11, M11L, MN12L, MN12SX, N12 and N12SX of the Minnesota family have been used for the calculation of different reactivity descriptors. In particular, it has been observed that the condensed dual descriptor and the Parr functions allow the qualitative prediction of the preferential reactive sites for glycation and protonation and give a solid explanation of the reactivity of these molecules.

Glycation products can degrade to AGEs (Advanced Glycation End-Products) by the effect of reactive oxygen species (ROS), among other agents. One of the inhibition routes to the AGEs formation is based on the administration of molecules with antioxidant capacity that scavenge ROS from the organism. In this thesis thiamine's antioxidant capacity is studied. Since this molecule also exhibits acid base reactivity, the antioxidant capacity of each possible protonation state has been studied. On the other hand, the methodology mentioned above has been used to determine the relative abundance of each such protonation states- From the reactivity and molar fraction of each protonation state, the global antioxidant capacity has been calculated as a function of the environment pH. The results indicate that thiamine scavenges the radical  $\bullet OCH_3$  efficiently and

independently of the environment's pH via the HAT (Hydrogen Atom Transfer) and RAF (Radical Adduct Formation) mechanisms. However, the capacity for scavenging peroxy radicals is not as good. It is worth mentioning that thiamine's macroscopic activity is due to non-majoritarian species that show exceptional microscopic reaction constants.

# Índice



<b>Índice</b>		1
<b>Abreviaturas</b>		3
<b>1</b>	<b>Introducción</b>	8
	<b>1.1. Glicación no enzimática de proteínas y formación de AGEs</b>	10
	Mecanismo general de la glicación proteica	10
	Productos finales de glicación avanzada (AGEs)	15
	<b>1.2. Vías de inhibición de la formación de AGEs</b>	16
	Especies reactivas de oxígeno (ROS) en la propagación del daño proteico: tipos y procedencia	18
	Estrés oxidativo procedente de glicación proteica	19
	Usos de métodos computacionales para el estudio de reacciones con radicales libres	21
	<b>1.3. Glicación selectiva de residuos de lisina</b>	23
<b>2.</b>	<b>Objetivos</b>	26
<b>3.</b>	<b>Metodología</b>	30
	<b>3.1. Cálculo teórico de pK<sub>a</sub>s</b>	32
	Modelos de disolvente continuo	33
	Uso de ciclos termodinámicos	35
	La reacción isodésmica	41
	Cálculo de pK <sub>a</sub> de aminoácidos y péptidos	44
	<b>3.2. DFT conceptual</b>	45
	<b>3.3. Detalles computacionales</b>	47
	Cálculo de pK <sub>a</sub> s de compuestos orgánicos	47
	Cálculo de pK <sub>a</sub> s de aminoácidos y péptidos	47
	Predicción de los sitios de protonación de péptidos tipo Ac-Lys-(Ala) <sub>n</sub> -Lys-NH <sub>2</sub> mediante descriptores de DFT conceptual	48
	Capacidad antioxidante de la tiamina	48
<b>4.</b>	<b>Resultados y discusión</b>	50
	<b>4.1. Cálculo de pK<sub>a</sub>s de compuestos orgánicos haciendo uso de la reacción isodésmica</b>	52
	<b>4.2. Cálculo de pK<sub>a</sub>s de aminoácidos y péptidos usando la reacción isodésmica</b>	59

	Cálculo del $pK_a$ del grupo $\alpha$ -carboxílico y del grupo $\alpha$ -amino e influencia del ácido de referencia	59
	Cálculo del $pK_a$ de las cadenas laterales de péptidos ( $pK_{a3}$ )	68
	Ciclos termodinámicos para el cálculo de $pK_{a3}$ de péptidos	71
	Muestreo conformacional y cálculos de $pK_a$	72
<b>4.3.</b>	<b>Predicción de los sitios de protonación de péptidos tipo Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> mediante descriptores de DFT conceptual</b>	76
<b>4.4.</b>	<b>Capacidad antioxidante de la tiamina</b>	80
	Equilibrios ácido-base de la tiamina en disolución acuosa	80
	Viabilidad termodinámica	82
	Viabilidad cinética	88
	Cálculo de las constantes cinéticas macroscópicas	91
<b>5</b>	<b>Conclusiones</b>	98
<b>6</b>	<b>Bibliografía</b>	102
<b>7</b>	<b>Apéndice</b>	124
	Apéndice A. Resultados adicionales	126
	Apéndice B. Métodos de Química Cuántica	141

## Abreviaturas

### A

ALT-711	Alagebrium
AG	Aminoguanidina
AGEs	Advanced Glycation End Products
ALEs	Advanced Lipoxidation end Products
AP	Producto de Amadori

### B

BS	Base de Schiff
----	----------------

### C

CBA	Crocin Bleaching Assay
CEL	Carboxietil lisina
CML	N-ε-(carboximetil)-lisina

### D

1-DG	Desoxiglucosona
3-DG	3-desoxiglucosona
DPPH	1,1difenil-2-picrilhidrazil
DOLD	3-deoxyglucosone Lysine Dimer

### F

FA	Formaldehido
----	--------------

### G

GLA	Glicolaldehido
GLO-I	Glioxalasa-I. Enzima
GO	Glioxal
GOLD	Glyoxal Lysine Dimer
GRD	Descriptorios globales de reactividad
GSH	Glutación (GluCysGly)

### H

HAT	Hydrogen Atom Transfer
4-HNE	4-Hidroxinonenal
HTRP	Hidroxitriptófano

### K

KYN	Quinureína
-----	------------

### L

LRD	Local Reactivity Descriptors
-----	------------------------------

### M

MDA	Malondialdehído
MGO	Metilglioxal
MOLD	Methylglyoxal Lysine Dimer

**N**  
NADPH Nicotínamida adenina dinucleótido fosfato  
NF-kB Factor pro-inflamatorio

**O**  
ORAC Oxygen Radical Absorbance Capacity

**P**  
PLP Fosfato de piridoxal  
PM Piridoxamina  
PTB Bromuro de fenaciltiazol

**R**  
RAF Radical Adduct Formation  
RAGE Receptor de AGE  
RCS Especies Reactivas de Carbono  
ROS Especies Reactivas de Oxígeno

**S**  
SET Sequential Electron Transfer

**T**  
TAGE Toxic AGES  
THPP Tiamina difosfato  
TPK Tiamina difosfoquinasa



# 1. Introducción



### **1.1. Glicación no enzimática de proteínas y formación de AGEs**

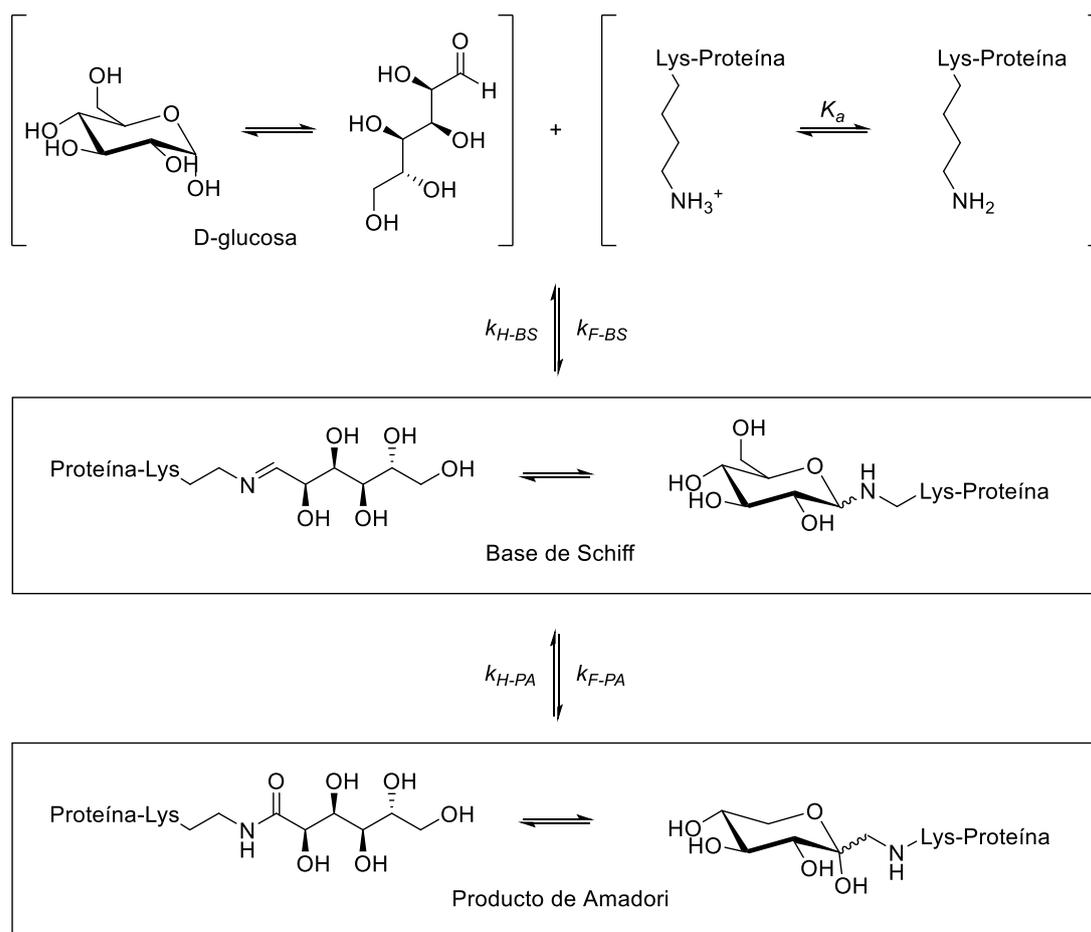
Se conoce como glicación no enzimática o reacción de Maillard, como reconocimiento a su descubridor [Maillard, 1912] a la reacción entre un grupo amino libre de un amino ácido o proteína y el grupo carbonilo de un azúcar reductor que tiene lugar sin control enzimático. Esta reacción puede provocar cambios patológicos a través de tres mecanismos generales: 1) por alteración de la estructura y función de las proteínas [Brownlee, 1988], 2) variando el nivel de señales solubles, como las citoquinas, hormonas y radicales libres que interactúan con receptores celulares específicos [Miyata, 1994] y 3) por la formación de especies reactivas de oxígeno [Baynes, 1999].

En los primeros años de investigación los esfuerzos se centraron en el estudio de esta reacción sobre los alimentos, observando que los productos mayoritarios son moléculas cíclicas y policíclicas que aportan sabor y aroma, aunque también pueden ser potencialmente cancerígenas [Skog, 1998]. La importancia de la glicación en la salud humana cobró relevancia cuando en 1971 se evidenció que este tipo de reacciones también ocurren in vivo [Trivelli, 1971] y no sólo en los alimentos elaborados.

La glicación de biomoléculas es un factor clave en muchas patologías asociadas a la diabetes (nefropatía, retinopatía, enfermedades cardiovasculares, etc.), así como en el envejecimiento y otras enfermedades neurodegenerativas [Thorpe, 1996], [Brownlee, 2001]. En individuos diabéticos no controlados, la falta o los niveles bajos de insulina aumentan considerablemente los niveles de glucosa en sangre, lo que afecta irreversiblemente a la integridad de las proteínas de larga vida [Fan, 2009], [Maillard-Lefebvre, 2009], [Zhang, 2011]. Las enfermedades neurodegenerativas, tales como el Alzheimer [Mimna, 2007] o el Parkinson [Xu, 2011], se relacionan con el plegamiento erróneo de proteínas y con los procesos de agregación y deposición de fibrillas de tipo amiloide, lo que podría ser consecuencia de la glicación no enzimática de las proteínas implicadas.

#### **Mecanismo general de la glicación proteica**

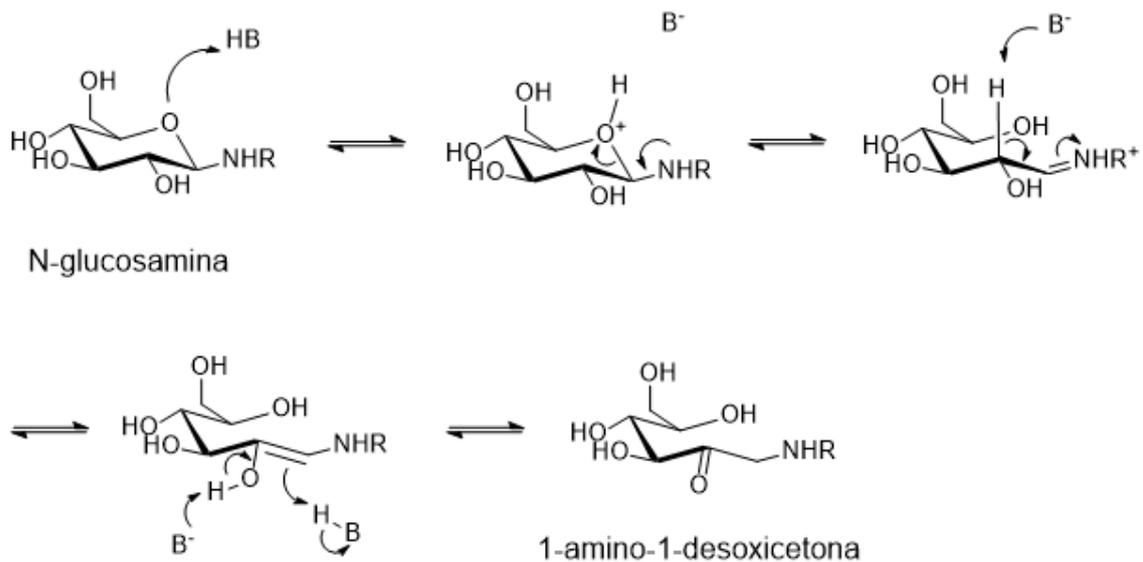
La reacción de Maillard no es una sola reacción, sino una serie de reacciones no enzimáticas en las que el primer paso es la adición de un grupo amino primario de una proteína al grupo carbonilo de un azúcar reductor, formando una base de Schiff (aldimina). Posteriormente, la base de Schiff sufre reacciones de reordenamiento intramolecular, formando un compuesto más estable (cetoamina) conocido como compuesto de Amadori (**Figura 1.1**).



**Figura 1.1.** Formación de la base de Schiff y del compuesto de Amadori. [Lee, 1992].

La reacción de glicación se inicia mediante la reacción de un grupo carbonilo de un azúcar reductor con un grupo amino libre de la proteína, principalmente con el amino terminal, con los de las cadenas laterales de las lisinas y con los grupos guanidinio de las cadenas laterales de argininas, en este orden de reactividad [Bunn, 1979]. Se forma una aldimina denominada base de Schiff (BS). Se trata de una reacción reversible que ocurre rápidamente, alcanzándose el equilibrio en unas horas [Brownlee, 1995], [Nimbalkar, 2012]. En el caso de azúcares reductores, la BS puede evolucionar rápidamente hacia un compuesto cíclico para formar la glicosilamina N-sustituida [Cho, 2007].

Una vez formada la Base de Schiff, el grupo hidroxilo que se halla en el carbono adyacente a la imina posibilita el llamado reordenamiento de Amadori, para dar lugar a un producto más estable (cetoamina), denominado Producto de Amadori (AP). Esta reacción también es reversible pero no tan rápida como la formación de la base de Schiff, alcanzándose el equilibrio termodinámico entre las 2 y las 4 semanas aproximadamente [Lee, 1992]. (**Figura 1.2**).



**Figura 1.2.** Reordenamiento de Amadori.

Smith y Thornalley [Smith, 1992] sugirieron que el reordenamiento de Amadori es un proceso mucho más lento que el de formación de la BS y que por ello ha de ser considerado como el paso limitante de la glicación, hipótesis que fue confirmada en los trabajos de Martins y van Boekel [Martins, 2005].

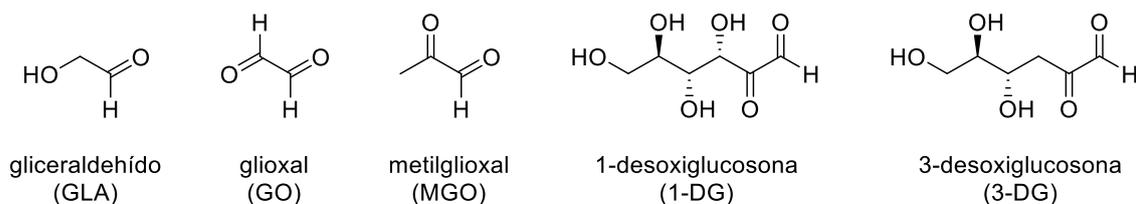
En conjunto, las Bases de Schiff y los Productos de Amadori reciben el nombre de productos de glicación temprana EGP “early glycation products”. Dado que, como se ha señalado, se trata de reacciones reversibles, estos productos de glicación temprana incrementan cuando los niveles de glucosa en sangre son elevados y disminuyen a niveles normales cuando la glucemia es regulada por tratamiento [Brownlee, 1988]. No obstante, si la exposición a la glucosa persiste, la glicación proteica puede continuar evolucionando hacia una tercera etapa.

Los productos de Amadori sufren una serie de reacciones de deshidratación, eliminaciones- $\beta$  sucesivas, condensación y oxidación, [Wautier, 2004], [Booth, 1997], [Ahmed, 2003] formando una variedad de productos carbonílicos (Glioxal (GO), Metilglioxal (MGO) y 3-desoxiglucosona (3-DG)) y un grupo heterogéneo de residuos o aductos unidos a proteínas conocidos como Productos de Glicación Avanzada o Advanced Glycation End-Products (AGEs) [Brownlee, 1995]. Los compuestos dicarbonílicos generados son altamente reactivos y, junto con otros compuestos monocarbonílicos como el formaldehído, reciben el nombre de RCS “reactive carbonyl species”. La generación de AGEs es una reacción lenta e irreversible que ocurre tras varios meses o incluso años de contacto con glucosa, principalmente en proteínas de vida

media larga como las que forman parte de la matriz extracelular como: colágeno, elastina, mioglobina, mielina y cristalina [Brownlee, 1988], [Vlassara, 1996], [Ahmed, 2005] y otras relacionadas con enfermedades neurodegenerativas como el Parkinson, como la alfa-sinucleína.

Los RCS actúan como propagadores de la reacción, siendo mucho más reactivos que los azúcares de los cuáles son derivados y volviendo a reaccionar con las biomoléculas produciendo más AGEs.

Otras vías derivadas de la reacción de Maillard llevan también a la producción de AGEs, en muchos casos a través de reacciones de oxidación. Así pues, en el proceso de glicación de proteínas, la formación de muchos de los AGEs incluye reacciones de oxidación, llamadas en conjunto glicoxidación. En las vías alternativas para la formación de AGEs, participan 1) la glucosa, 2) los  $\alpha$ -hidroxialdehídos (gliceraldehído (GLA) y glicolaldehído) y 3) los compuestos dicarbonilo intermediarios u oxoaldehídos altamente reactivos glioxal (GO), metilglioxal MGO) y 3-desoxiglucosona (3-DG). (**Figura 1.3**). Estas especies generan un total de seis diferentes clases de estructuras AGEs que se clasifican en dos grupos. Uno de ellos asociado a toxicidad celular que incluye a los AGEs derivados de gliceraldehído (AGE-2) y de glicolaldehído (AGE-3), denominados AGEs tóxicos (TAGE, por *Toxic AGEs*). El otro grupo formado por los AGEs no tóxicos, derivados de glucosa tales como AGE-1, N- $\epsilon$ -(carboximetil)-lisina (CML), carboxietil lisina (CEL), pentosidina, pirralina y crosslina, así como los derivados a partir de MGO, GO y 3-desoxiglucosona (AGE-4, AGE-5 y AGE-6, respectivamente), llamados hidroimidazolones. Estos últimos, a su vez, forman los correspondientes productos de entrecruzamiento o dímeros de lisina: GOLD (Glyoxal Lysine Dimer), MOLD (Methylglyoxal Lysine Dimer) y DOLD (3-Deoxyglucosone Lysine Dimer) [Thornalley, 2003A], [Sato, 2006].



**Figura 1.3.** Productos RCS: GLA: gliceraldehído. 1-DG: 1-desoxiglucosona. 3-DG: 3-desoxiglucosona. GO: glioxal. MGO: metilglioxal.

La oxidación de la glucosa, de las SB y de los APs produce compuestos reactivos de oxígeno (ROS) y compuestos carbonílicos y dicarbonílicos altamente reactivos (RCS). Los RCS se pueden generar 1) por fragmentación oxidativa de la BS (mecanismo de Namiki) [Fu, 1996], 2) por autooxidación de la glucosa mediante catálisis metálica (mecanismo de Wolff) y 3) por peroxidación lipídica [Wolff, 1987], [Wolff, 1991]. Por otro lado, el GO, MGO, 3-DG y GLA se forman, además, a partir de la BS y del AP [Thornalley, 1999]. El GO también se forma por peroxidación lipídica y a partir de la treonina, mientras que el GLA puede provenir de la oxidación de la L-serina [Grillo, 2008].

La autooxidación de la glucosa genera nuevos compuestos debido a las reacciones de degradación. A partir de la desprotonación del C(2) de la glucosa con la subsiguiente oxidación del 1,2-enediol formado se obtiene la desoxiglucosona (1-DG). La formación de la 3-DG se inicia también mediante la desprotonación del C(2) de la glucosa, seguido de deshidratación y formación de un 2,3-enol, el cual se reordena dando la 3-DG o bien condensa retroaldólicamente dando gliceraldehído y MGO [Thornalley, 1999]. La condensación retroaldólica puede darse también sobre la glucosa, formándose GO o bien una tetrosa y GLA que se oxida dando GO,  $\text{H}_2\text{O}_2$  y radicales  $\text{HO}\cdot$  [Thornalley, 1999].

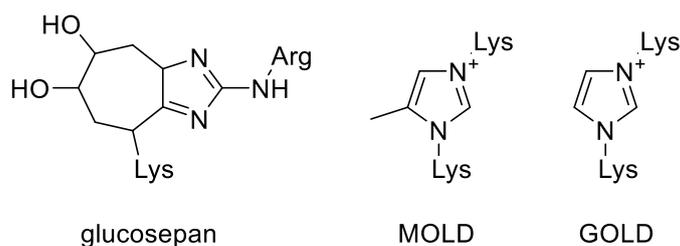
Se ha observado que la oxidación de la glucosa está afectada por la presencia del dianión fosfato, ya que éste cataliza la desprotonación del C(2). La concentración de  $\text{HPO}_4^{2-}$  es determinante en la formación de GO [Thornalley, 1999]. Los metales divalentes ( $\text{Fe}^{2+}$  y  $\text{Cu}^{2+}$ ) tienen un efecto similar, tal es así, que su complejación disminuye fuertemente la formación de GLA, GO y  $\text{HO}\cdot$  [Thornalley, 1984] y ligeramente los niveles de 3-DG, sin afectar a la formación de MGO a partir de 3-DG [Thornalley, 1999].

## Productos finales de glicación avanzada (AGES)

Los AGES son un conjunto complejo y heterogéneo de compuestos estables [Ikeda, 1996] que se forman fundamentalmente por dos vías. La primera es la deshidratación y reordenamiento del AP, tanto en condiciones oxidativas como en no oxidativas. Una segunda vía es por reacción de RCS con grupos funcionales amino, tiol y guanidinio.

Los AGES se agrupan en dos amplias clases: los AGES de no-entrecruzamiento (non crosslinking AGES) los cuales se forman mediante la reacción de la glucosa o de un grupo carbonílico o dicarbonílico con grupos de la cadena lateral de un residuo proteico y los AGES de entrecruzamiento (crosslinking AGES) que se forman como consecuencia de la reacción de un segundo grupo nucleófilo proteico sobre un AP o post-Amadori. La diversidad de los grupos carbonílicos y de las cadenas laterales proteicas hace que exista un gran número de AGES de entrecruzamiento.

Algunos AGES de entrecruzamiento son fluorescentes, por lo que se detectan usando la espectroscopia de fluorescencia. Sin embargo, la mayor parte de las estructuras de los AGES responsables del entrecruzamiento proteína-proteína in vivo no son fluorescentes (99 % del total). Algunos de estos compuestos, son el GOLD y el MOLD que se forman mediante la reacción de GO y MGO con los residuos de lisina, respectivamente [Ahmed, 2002], o el glucosepan, que se forma por el ataque del grupo guanidina sobre productos de la degradación del AP [Lederer, 1999], siendo el AGE que más se acumula en el colágeno humano [Sell, 2005]. (**Figura 1.4**).



**Figura.1.4.** Diferentes tipos de AGES.

Los AGES pueden interactuar con una proteína de la superficie celular específica, denominada receptor de los AGES (RAGE) [Schmidt, 1992]. El RAGE es un receptor multiligando perteneciente a la familia de las inmunoglobulinas. La unión de los AGES inicia la transcripción del factor pro inflamatorio (NF- $\kappa$ B) con la consiguiente respuesta inflamatoria. Adicionalmente, se generan especies oxigenadas radicalarias (ROS) que disminuyen la actividad de la enzima glioxalasa-I (GLO-I), interfiriendo de

esta manera en el sistema natural de eliminación de los precursores de los AGEs. Esta situación implica el aumento de la concentración de GO y MGO, con el consecuente incremento de la concentración de AGEs y de la activación de los RAGEs perpetuando la respuesta inflamatoria [Fleming, 2011], [Ramasamy, 2012].

## **1.2. Vías de inhibición de la formación de AGEs**

Dadas las implicaciones patológicas de la glicación proteica, de la formación de AGEs y de los productos secundarios que se producen, actualmente se dedican grandes esfuerzos a investigar sus modos de inhibición. La inhibición se presenta como una terapia para evitar o atenuar las complicaciones de la diabetes (nefropatías, retinopatías, problemas vasculares) y otras patologías asociadas al envejecimiento (osteoporosis, osteoartritis, neuropatías).

Los principales mecanismos de inhibición de la formación de AGEs son los siguientes [Wu, 2011]: 1) Secuestro de los Radicales libres para aliviar el estrés oxidativo y reducir la generación de nuevos grupos carbonílicos o RCS. 2) Secuestro y eliminación de los grupos carbonílicos o RCS. 3). Quelación de metales de transición con capacidad redox que participan en la oxidación de las BS y los AP formando AGEs, RCS y ROS 4). Inhibición de la formación de AGEs. 5). Ruptura de las estructuras de los AGEs de entrecruzamiento. 6) Bloqueo de los RAGEs, lo que puede reprimir el estrés oxidativo y la inflamación de los sistemas biológicos.

En base a estos requisitos, se han ensayado algunos inhibidores potenciales [Rahbar, 2000], [Rahbar, 2003] que han superado, en algunos casos, la fase de estudios clínicos [Williams, 2004], [Giannoukakis, 2005].

Los compuestos antioxidantes parecen ser buenos agentes de prevención de la formación de AGEs, ya que disminuyen el estrés oxidativo. Se ha estudiado un grupo numeroso de flavonoides naturales, demostrándose una buena correlación entre los efectos inhibitorios y el secuestro de radicales libres [Wu, 2005]. Por este motivo, en los últimos años se han estudiado las actividades inhibitoras de algunos extractos de plantas y de compuestos fenólicos [Peng, 2011], [Wu, 2011]. Cabe matizar que se ha señalado que varios polifenoles, moléculas con propiedades antioxidantes, del té pueden atrapar ciertos RCS en condiciones fisiológicas [Lo, 2006]. Estos resultados sugieren que no sólo las propiedades antioxidantes de los compuestos fenólicos son la razón de su capacidad inhibitoria.

Por otro lado, se han ensayado una gran variedad de compuestos sintéticos con propiedades inhibitorias de la formación de AGEs o con capacidad para romper los compuestos de entrecruzamiento generados durante la glicación. No obstante, en algunos casos se ha comprobado que pueden producir efectos secundarios severos como es el caso de la aminoguanidina (AG) [Thornalley, 2003A], [Thornalley, 2003B], [Williams, 2004].

Sólo unos pocos inhibidores son eficaces en las primeras fases de la glicación interfiriendo en la formación de la BS. El ácido acetilsalicílico (aspirina) bloquea los grupos amina libres de proteínas mediante su acetilación [Caballero, 2000] el diclofenaco interacciona de forma no covalente con las proteínas evitando la unión con los azúcares [van Boekel, 1992], mientras que el inositol atrapa las moléculas de glucosa [Ramakrishnan, 1999].

La mayoría de inhibidores sintéticos previenen la formación de AGEs secuestrando RCS o bloqueando la formación de intermediarios. Entre ellos, destaca la AG, compuesto que presenta dos centros de reacción, un grupo hidracina y un grupo guanidino [Koskinen, 1997]. Se ha observado que la AG reacciona con los  $\alpha$ -oxoaldehídos GO y MGO [Thornalley, 2000], con los precursores de ALEs [Thornalley, 2000], con la glucosa [Hirsch, 1995], con aldehídos  $\alpha,\beta$ -insaturados [Al-Abet, 1997].

La AG también presenta una capacidad complejante de iones metálicos [Price, 2001] e inhibe la formación de HO• [Giardino, 1998] y de la actividad de la enzima amino oxidasa, la cual contribuye a la generación de MGO y FA [Yu, 1997]. Se ha comprobado que la AG es capaz de inhibir la formación de AGEs y la modificación de lipoproteínas [Aldini, 2007], reduciendo la aparición de ciertas complicaciones patológicas de la diabetes [Thornalley, 2003A], [Nemoto, 2012]. Sin embargo, Khalifah [Khalifah, 1999] demostró que la capacidad inhibidora de la AG es muy inferior a la mostrada por la PM.

Otros compuestos que actúan mediante el secuestro de RCS son la metformina [Tanaka, 1997], [Ruggiero-López, 1999]; el OPB-9195 que reacciona con los RCS mediante un mecanismo similar al descrito para la AG [Miyata, 2000], [Miyata, 2003], aunque presenta efectos secundarios [Peyroux, 2006]; la D-penicilamina [Wondrak, 2002] y la hidralazina [Vindis, 2006], que también es un potente antioxidante que inhibe la activación de la oxidasa NADH con la consecuente generación de ROS [Munzel, 1996].

Existen algunos derivados de aminoácidos que presentan propiedades inhibitorias como la N-acetyl-Cys [Neely, 2000], la S-adenosilmetionina [Valentovik, 2004], el GSH (Glu-Cys-Gly) [Petersen, 2004] y la histidil hidrazida [Tang, 2007], los cuales inhiben la formación de 4-HNE. Por otro lado, la carnosina es un dipéptido que

posee habilidades de antioxidante, de secuestrador de carbonilos, de bloqueador de la formación de MDA y de MGO [Hipkins, 1997], y de protector del entrecruzamiento entre proteínas y ADN [Guiotto, 2005]. Trabajos recientes se han centrado en el estudio de derivados de carnosina que mantienen o mejoran la reactividad con los RCS [Vistoli, 2012].

Hay una serie de compuestos que rompen los AGEs de entrecruzamiento. Vasan y cols. [Vasan, 1996] demostraron que el bromuro de fenaciltiazol (PTB) era capaz de escindir los enlaces covalentes en AGEs de entrecruzamiento, aunque no llegó a utilizarse como fármaco ya que reduce los enlaces disulfuro proteicos [Thornalley, 1999]. En cambio, el alagebrium (ALT-711) es capaz de inhibir el entrecruzamiento proteico y la formación de AGEs en un 95% [Vasan, 2003].

### **Especies reactivas de oxígeno (ROS) en la propagación del daño proteico: tipos y procedencia**

Con el nombre ROS, se engloba a un conjunto de especies reactivas de oxígeno de diversa procedencia, siendo éste un término colectivo que incluye tanto radicales libres como el radical superóxido ( $O_2\bullet$ ), el radical anión superóxido ( $O_2\bullet^-$ ), el radical hidroxilo ( $\bullet OH$ ), los radicales peroxilo ( $RO_2\bullet$ ) y los radicales alcoxilo ( $RO\bullet$ ), y algunas especies derivadas del oxígeno, aunque no radicalarias, como el peróxido de hidrógeno,  $H_2O_2$ , o el anión superóxido,  $O_2^-$  [Halliwell, 1999].

Los ROS, son generados continuamente como parte del metabolismo aeróbico mitocondrial, donde actúan como intermedios necesarios en una gran variedad de reacciones bioquímicas normales. No obstante, si estos ROS se producen en exceso en las células como consecuencia de desórdenes metabólicos patológicos, generan estrés oxidativo, capaz de dañar múltiples macromoléculas, entre ellas las proteínas, lo cual está implicado en el desarrollo de las complicaciones asociadas a la diabetes mellitus [Matxain, 2009]. Los pacientes diabéticos son más propensos al daño producido por el estrés oxidativo. Esto se debe a que la hiperglucemia debilita las defensas antioxidantes, facilitando que los radicales libres dañen moléculas intra y extracelulares y también proteínas estructurales. En efecto, se ha comprobado una reducción de la concentración de alfa-tocoferol (vitamina E) y de ácido ascórbico (vitamina C) en pacientes diabéticos, además de una disminución eritroleucocitaria de enzimas antioxidantes como la superóxido dismutasa, la catalasa y la glutatión peroxidasa [Pennathur, 2004]. De hecho, ya en 1989, Taniguchi *et al.* determinaron que la superóxido dismutasa pierde gran parte

de su actividad enzimática cuando se halla glicada, lo que agrava, por tanto, el daño que los ROS pueden infligir en pacientes diabéticos [Taniguchi, 1989]. En el presente apartado nos centraremos en los tipos y procedencia de los ROS producidos por glicación proteica, que posteriormente llevarán a la formación de AGEs y ALEs.

### **Estrés oxidativo procedente de glicación proteica**

Las reacciones colaterales de oxidación de la glucosa, las Bases de Schiff y los Productos de Amadori dan lugar a ROS y a otros productos altamente reactivos como los RCS. Los ROS más habitualmente producidos a través de estas reacciones colaterales de oxidación son el radical superóxido ( $O_2\bullet$ ) producido principalmente a partir de la descomposición de los compuestos de Amadori [Mossine, 1999], [Chetyrkin, 2008B] o el radical anión superóxido ( $O_2\bullet^-$ ) producido a partir de la degradación oxidativa tanto de la base de Schiff como del producto de Amadori [Mullarkey, 1990], [Elgawish, 1996]. Estos radicales son precursores del peróxido de hidrógeno [Nakashima, 2010] el cual, a su vez, puede degradarse mediante la reacción de Fenton catalizada por Fe(II) a radical hidroxilo,  $\bullet OH$ , una de las especies más reactivas y causante de gran parte del daño sobre las proteínas [Chetyrkin, 2008B], [Chetyrkin, 2011]. Asimismo, el  $H_2O_2$  puede generarse también por degradación oxidativa de la glucosa y de algunos RCS, como el glioxal (GO) o el glicolaldehído (GLA) [Wolff, 1987], [Thornalley, 1999], [Mullarkey, 1990].

Según los estudios llevados a cabo por Mossine *et al.* en 1999 [Mossine, 1999], la generación de ROS por parte de las proteínas glicadas es una de las mayores causas de estrés oxidativo en la diabetes y la edad. En estos estudios, los autores hallaron que el  $Cu^{2+}$  aumentaba marcadamente la concentración de radical superóxido producido,  $O_2\bullet^-$ , mientras que el  $Fe^{3+}$  no aceleraba la reacción. Además, en adición a la formación de radicales libres de oxígeno catalizada por metales, también se produce la oxidación libre de uno de los enoles del producto de Amadori sin intervención de iones metálicos. Ambas vías contribuyen a la generación de ROS y su subsecuente daño oxidativo en las proteínas [Mossine, 1999].

El  $H_2O_2$  una vez formado se degrada inmediatamente para dar especies más reactivas, especialmente si el proceso es catalizado por el hierro mediante la reacción de Fenton para dar  $\bullet OH$  [Nakashima, 2010]. En 2010, Nakashima *et al.* usando un medio de cultivo incubado con un modelo in vitro de glicación proteica, hallaron que se producía anión superóxido,  $O_2\bullet^-$  como producto de la glicación proteica y que éste, posteriormente, se reducía a  $H_2O_2$  y ésta a  $\bullet OH$  [Nakashima, 2010]. Otro estudio in vitro del mismo grupo

llevado a cabo en 2012, mostró que la generación de  $O_2^-$  se veía incrementada cuando se agregaba un potente quelante de metales, con respecto a un experimento en el que había metales disponibles. Esto sugiere que la producción de  $O_2^-$  se sigue llevando a cabo, incluso, sin la implicación de metales de transición [Nakashima, 2012]. Por otra parte, además de la producción de radicales durante la propia secuencia de reacciones de glicación proteica, los productos finales de glicación avanzada que se forman pueden desencadenar una mayor producción de ROS mediante la unión a moléculasceptoras específicas, de la familia de las inmunoglobulinas, que reciben el nombre de receptores de AGEs, RAGEs [Schmidt, 1996].

En su conjunto, los ROS pueden causar daño a las cadenas laterales de los aminoácidos de las proteínas, como la oxidación de cisteínas a sulfonatos, la oxidación de metionina a sulfóxido y la hidroxilación de aminoácidos aromáticos, como el triptófano. La oxidación de los residuos de triptófano da lugar a productos como el hidroxitriptófano (HTRP), la quinureína (KYN) y la 3-hidroxiquinureína (3-OH-KYN) [Finley, 1998] y a la fragmentación del esqueleto proteico [Chetyrkin, 2008]. Además, según el mecanismo propuesto por Chetyrkin *et al.* en 2011 [Chetyrkin, 2011] para la fragmentación del esqueleto proteico causada por el radical hidroxilo, también se contempla la producción del radical hidroperoxilo,  $\bullet OOH$ .

Estudios recientes sugieren que la tiamina podría aliviar el estrés oxidativo en diferentes organismos. En microsomas de hígados de ratas, la tiamina inhibe la peroxidación lipídica y la oxidación por radicales libres del ácido oleico *in vitro* [Lukienko, 2000]. Adicionalmente, se disminuyó el estrés oxidativo en células de hígado cuando las mismas fueron suplementadas con una combinación de tiamina y ascorbato [Wang, 2007]. Sheline y Choi [Sheline, 2004] describieron que la tiamina es un potencial agente terapéutico para la enfermedad de Wilson, debido a la restauración de la actividad enzimática dependiente de tiamina, la cual es inhibida por las especies reactivas de oxígeno inducidas por  $Cu^{2+}$ . Mientras que los estudios *in vitro* sugieren que la tiamina puede actuar directamente como un antioxidante, la asociación entre las enzimas dependientes de tiamina y el estrés oxidativo puede indicar un papel vital para la tiamina en condiciones de estrés.

La tiamina puede actuar como un antioxidante indirecto, promoviendo la producción de equivalentes reductores (NADPH, a través de la ruta de las pentosas fosfato), o funcionando como un agente quelante de los metales de transición tales como hierro y cobre.

La tiamina inhibe la oxidación del ácido ascórbico [Gero, 1954] y, al igual que la cisteína y el ácido ascórbico, previene la oxidación de grupos tiol, como los de las cisteínas, en las células [Jones, 1983]. El complejo tiamina-ácido ascórbico inhibe la oxidación de los radicales libres de la dopamina [Florence, 1989]. Los altos niveles de radicales libres en el cerebro [Langlais, 1997] y productos de peroxidación de lípidos en el hígado [Sushko, 1981] se encontraron en ratas con deficiencia de tiamina. Al mismo tiempo, la tiamina se oxida rápidamente en medios que contienen  $\text{Cu}^{2+}$  y ácido ascórbico [Stepuro, 1997].

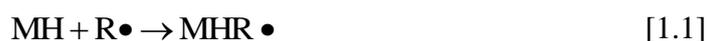
Un aspecto más importante de la actividad biológica de la vitamina B1 se relaciona con su efecto sobre la acumulación de productos de la glicación de proteínas, causando su envejecimiento [Shangari, 2007], [Booth, 1997]. Además del difosfato de tiamina y de la tiamina, otras sales de tiazolio son también inhibidores del envejecimiento de las proteínas. En este caso, los mecanismos de acción para la vitamina B1 y sus análogos estructurales de tiazolio pueden estar conectados tanto con la degradación de los restos  $\alpha$ -dicetona de las moléculas de proteína como con la inhibición de la autooxidación de los productos de glicosilación con la participación de especies reactivas de oxígeno. La evidencia a favor de esta última hipótesis proviene de datos que muestran que la tiamina y el difosfato de tiamina pueden actuar directamente como antioxidantes, reaccionando con especies activas de oxígeno y previniendo la peroxidación lipídica [Lukienko, 2000], [Jung, 2003].

**Usos de métodos computacionales para el estudio de reacciones con radicales libres.** Diferentes ensayos experimentales como CBA (Crocín Bleaching Assay), ORAC (Oxygen Radical Absorbance Capacity) y actividad frente al radical DPPH (1,1-difenil-2-picrilhidrazil), entre otros, asociados a determinaciones espectroscópicas por UV-Visible, fluorescencia o Resonancia Paramagnética electrónica pueden utilizarse para determinar la capacidad antioxidante de un compuesto. Sin embargo, estos ensayos están centrados en la capacidad antioxidante de compuestos sobre radicales no presentes en el organismo, como el DPPH, o bien determinan globalmente esta capacidad sin diferenciar entre distintos radicales. Por otra parte, no permiten explicar el mecanismo de reacción ni el sitio de la molécula que presenta una mayor capacidad antioxidante. Por ello, los métodos computacionales, y en concreto la teoría del funcional de la densidad (DFT), se posicionan como potentes herramientas para profundizar en el estudio de las reacciones con radicales libres, ya que permiten determinar tanto los mecanismos de reacción como

los valores cinéticos y termodinámicos correspondientes a las distintas posiciones reactivas. Así, a partir de un método teórico, no sólo se podrá determinar la capacidad antioxidante global de una molécula, sino también a través de qué mecanismo transcurre y en qué átomo se lleva a cabo preferentemente la reacción. Se han realizado numerosos trabajos como los estudios del grupo de Annia Galano y Raúl Álvarez-Idaboy, referidos a multitud de moléculas orgánicas con capacidad antioxidante, por ejemplo: capsaicina [Galano, 2012], cafeína [León-Carmona, 2011], canolol [Galano, 2011A] o melatonina [Galano, 2011B].

Los mecanismos más habituales son RAF (radical adduct formation), HAT (hydrogen transfer) y SET (single electron transfer). Se produce la reacción entre una molécula con propiedades antioxidantes (M) y un radical R·.

**Radical adduct formation (RAF):** El radical se enlaza a un átomo de la molécula antioxidante.



**Hydrogen atom transfer (HAT):** Corresponde a la abstracción de un átomo de hidrógeno, •H, por parte del radical, dejando la molécula antioxidante en forma radicalaria, T•, y neutralizándose el radical libre inicial.



**Single electron transfer (SET):** En este mecanismo se produce la transferencia de un electrón de la molécula antioxidante al radical R•, quedando el radical neutralizado como anión y el antioxidante en forma radicalaria y catiónica, MH•<sup>+</sup>



Esta reacción es muy frecuente con antioxidantes con enlaces  $\pi$  conjugados como los carotenoides.

### 1.3. Glicación selectiva de residuos lisina

Los distintos grupos amino de una proteína tienen diferente reactividad frente a un azúcar reductor dado. Es decir, existen sitios preferentes de glicación entre los diferentes residuos de una proteína. Esto se debe a que para que el grupo amino pueda actuar como nucleófilo en la formación de la base de Schiff, primero debe desprotonarse. De hecho, según la definición de acidez de Lewis, la formación de un enlace entre el nitrógeno amino de un residuo proteico con el carbono carbonílico de un azúcar reductor, es el mismo tipo de reacción que la protonación del grupo amino.

Las lisinas se desprotonan a valores de pH muy superiores al fisiológico. Esto podría dar la falsa apariencia de que las cadenas laterales de las lisinas de una proteína se mantienen protonadas a pH fisiológico y no son susceptibles a actuar como nucleófilos frente a los grupos carbonilos de azúcares reductores o RCS. Sin embargo, debe tenerse en cuenta que el  $pK_a$  está relacionado con energía libre de desprotonación frente al disolvente agua, pero pueden darse reacciones de intercambio protónico con otras moléculas de  $pK_a$  similar a las lisinas que permitan la existencia de cierto porcentaje de forma desprotonada a pH fisiológico. Por otra parte, los  $pK_a$ s de los residuos en una proteína pueden modificarse significativamente por la accesibilidad del disolvente y la interacción con otros grupos de la proteína. En cualquier caso, para un péptido o proteína con muchos grupos ácido-base, su desprotonación sigue un orden que viene determinado por sus valores de  $pK_a$ . Dicho orden es de interés porque, además de otros factores, es cualitativamente indicativo de la reactividad de cada grupo como nucleófilo en las reacciones de glicación.

Experimentalmente es posible determinar  $pK_a$ s de proteínas a partir de valoraciones ácido-base, pero las constantes determinadas de esta manera son

macroscópicas. Normalmente, las constantes macroscópicas no se corresponden a la desprotonación de un residuo o grupo en particular, sino que representan los equilibrios de protonación/desprotonación de los múltiples grupos presentes en la proteína. Por este motivo es difícil determinar los  $pK_a$ s individuales, o microscópicos, de cada residuo. Sin embargo, los métodos computacionales permiten estudiar los equilibrios microscópicos con resolución atómica. Por este motivo, el desarrollo de metodologías computacionales que permitan predecir  $pK_a$ s microscópicos es de interés general. En el caso particular de la glicación, el conocimiento de los  $pK_a$ s de los residuos permitiría comprender la relación entre la reactividad ácido-base y la preferencia por la glicación. De existir tal correlación, sería posible emplear métodos computacionales para identificar las dianas de glicación preferentes e investigar posibles mecanismos de inhibición de estos procesos.



## **2. Objetivos**



Se conocen con el nombre de reacciones de glicación, las reacciones que transcurren sin control enzimático entre los grupos amina de las biomoléculas y los grupos carbonilo de azúcares reductores, como por ejemplo la glucosa. Los procesos de glicación se han relacionado con diabetes, envejecimiento, Alzheimer y Parkinson. Las reacciones de glicación no enzimática modifican preferentemente los residuos lisina y arginina de las proteínas. El primer paso de estas reacciones es la formación de una base de Schiff. Para que se pueda iniciar esta reacción, los grupos amino y guanidinio deben desprotonarse. A pH fisiológico, la abundancia de la forma reactiva desprotonada de cada grupo depende de su  $pK_a$  y éste del entorno químico y de la accesibilidad del disolvente al grupo en cuestión. De hecho, algunos autores han indicado una posible relación directa entre el  $pK_a$  de los residuos de las proteínas y el proceso de la glicación [Jakas, 2004]. Por lo tanto, el conocimiento de los valores del  $pK_a$  de estos residuos es de vital importancia para comprender el proceso de glicación de una proteína.

La determinación computacional del  $pK_a$  de aminoácidos es un paradigma de especies cuyos tautómeros mayoritarios en disolución y en fase gas son diferentes. Por ello es necesario desarrollar un procedimiento de cálculo de  $pK_a$ s sin el uso de los ciclos termodinámicos y así evitar el cálculo de las energías libres en fase gaseosa. Recientemente nuestro grupo de investigación ha aplicado con éxito el esquema de la reacción isodésmica [Casasnovas, 2011], [Casasnovas, 2012], en el cálculo de  $pK_a$ s de ácidos de carbono en moléculas orgánicas. Por ello el primer objetivo será evaluar la robustez de la reacción isodésmica en el cálculo del  $pK_a$ :

- i) de moléculas orgánicas con grupos funcionales ácido-base comunes.
- ii) del grupo  $\alpha$ -carboxílico ( $pK_{a1}$ ), del grupo  $\alpha$ -amino ( $pK_{a2}$ ), y de la cadena lateral ( $pK_{a3}$ ) de los aminoácidos proteicos, así como de algunos péptidos.

La consecución de este objetivo permitirá la consecución de otros objetivos parciales:

- i) Evaluación de métodos de química cuántica que, utilizados junto a la reacción isodésmica, proporcionen las mejores predicciones de  $pK_a$ . En concreto se emplearán cálculos semiempíricos con el Hamiltoniano PM6, cálculos basados en la Teoría del Funcional de la Densidad (DFT) con los funcionales B3LYP, M05-2X y M06-2X y cálculos post-Hartree-Fock: Møller-Plesset y Coupled Cluster dentro del método compuesto CBS-4B3\*.

- ii) Evaluación de métodos empíricos para el cálculo de  $pK_a$ s: ChemOffice.

iii) Comparación de los valores de  $pK_a$  obtenidos con la reacción isodésmica con los obtenidos con ciclos termodinámicos.

iv) Importancia de la inclusión de múltiples conformaciones en la determinación de los  $pK_a$ s de moléculas flexibles como los péptidos.

v) Efecto del ácido de referencia en la determinación de  $pK_a$  con la reacción isodésmica y criterios básicos para su elección.

El segundo objetivo será estudiar la acidez de dos lisinas en una serie de péptidos del tipo Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> mediante un método alternativo a la reacción isodésmica y los ciclos termodinámicos: los descriptores de DFT conceptual.

Las reacciones colaterales de oxidación de la glucosa, de las Bases de Schiff y de los Productos de Amadori dan lugar a la formación de ROS y a otros productos altamente reactivos como los RCS que propagan el daño iniciado por la glucosa y así aumenta el número y tipo de AGEs producidos. Por lo tanto, uno de los mecanismos de inhibición en la formación de AGEs es el secuestro de las especies altamente reactivas ROS.

Se ha propuesto que la tiamina tiene propiedades antioxidantes y que puede actuar eliminando radicales libres [Okai, 2007]. Por ello el tercer objetivo será evaluar mediante cálculos basados en la Teoría del Funcional de la Densidad (DFT) la reactividad de la tiamina con diversos radicales libres ( $\cdot\text{OCH}_3$ ,  $\cdot\text{OOCH}_3$  y  $\cdot\text{OOH}$ ), así como la polaridad del medio y del pH.

# 3. Metodología



### 3.1. Cálculo teórico de pK<sub>a</sub>s

En los últimos años se han dedicado grandes esfuerzos en desarrollar métodos computacionales que permitan el cálculo teórico del pK<sub>a</sub> de un ácido con precisión [Ho, 2010A], [Ho, 2011], [Casasnovas, 2014], [Seybold, 2015].

Para calcular teóricamente el pK<sub>a</sub> de un ácido hace falta calcular la energía libre de la reacción de disociación del ácido en disolución. Un cálculo riguroso de las energías libres de desprotonación en disolución requeriría métodos de simulación que consideren explícitamente el disolvente a la temperatura y presión experimentales. Además, salvo en el caso de ácidos fuertes, la reacción de desprotonación puede ser endergónica y/o presentar barreras de energía superiores a la energía térmica  $k_B T$  por lo que se requeriría emplear algún método de cálculo de energía libre como thermodynamic integration o umbrella sampling. Por último, dado que las transferencias de protón implican ruptura y formación de enlaces, deberían emplearse simulaciones *ab initio*<sup>1</sup>. El inconveniente es que los métodos de simulación son más costosos computacionalmente que los métodos estáticos, especialmente en el caso de simulaciones *ab initio*.

Existen esencialmente tres formas de introducir los efectos de un disolvente en cálculos estáticos de química cuántica. En primer lugar, se puede construir una supermolécula o cluster formado por la molécula del soluto y moléculas de disolvente. Como el número de moléculas de disolvente que rodea al soluto incrementa rápidamente, únicamente se suelen considerar la primera o segunda capas de solvatación. Además, a medida que incrementa el número de moléculas de disolvente, el cálculo se hace más costoso computacionalmente y es más complicado optimizar las estructuras de mínima energía. La segunda aproximación consiste en reemplazar las moléculas del disolvente por un continuo caracterizado por una constante dieléctrica. El continuo y la densidad electrónica de la molécula se polarizan mutuamente para simular los efectos del disolvente. Esta aproximación es muy cruda pero es computacionalmente poco costosa y permite calcular energías de solvatación muy fácilmente. La tercera aproximación es un híbrido de las dos primeras en la que se emplea un número de moléculas de disolvente muy reducido y el cluster resultante se rodea por un modelo continuo. De esta forma, las moléculas explícitas de disolvente permiten describir las interacciones puntuales con el

---

<sup>1</sup> También sería posible calcular la energía libre de desprotonación empleando un Hamiltoniano clásico y el método *alchemical free energy perturbation* puesto que en este caso no es necesario que los estados intermedios de la desprotonación correspondan a estados reales del sistema.

soluto mientras que el modelo continuo permite describir el efecto del grueso del disolvente lejos del soluto.

Por estos motivos, la forma tradicional para el cálculo teórico del  $pK_a$  emplea cálculos estáticos formando ciclos termodinámicos que combinan la energía libre de reacción de desprotonación en fase gas ( $\Delta G_{gas}$ ) con las energías de solvatación de los ácidos y bases conjugadas ( $\Delta G_{solv}$ ) calculadas con modelos de disolvente continuo. El uso de ciclos termodinámicos se justifica debido a que no es rigurosamente correcto calcular energías libres del soluto en disolución simplemente considerando un modelo continuo y las aproximaciones armónica y del rotor rígido, válidas para moléculas en fase gaseosa [Ho, 2010B], [Casasnovas, 2011].

### **Modelos de disolvente continuo**

Como se ha indicado anteriormente, estos modelos consideran al disolvente como un medio continuo sin estructura interna caracterizado por su constante dieléctrica  $\epsilon$  que rodea una cavidad que contiene la molécula de soluto. Estos modelos pueden emplearse con Hamiltonianos clásicos en cálculos de mecánica molecular o dinámica molecular o se pueden emplear con Hamiltonianos cuánticos. En este último caso, la interacción entre la molécula, o cluster de moléculas, de soluto y el disolvente se calcula mediante un término  $V_{int}$  que se añade al Hamiltoniano electrónico. Así como el término  $V_{int}$  polariza la función de onda del soluto, el disolvente continuo también se polariza en la superficie de la cavidad que rodea dicho soluto. Por este motivo, la interacción-soluto disolvente se resuelve como parte del proceso iterativo de optimización de la función de onda y estos métodos se conocen como métodos de campo de reacción autoconsistente (Self-Consistent Reaction Field, SCRF).

En los métodos SCRF entran en juego, además de la energía de interacción electrostática y la energía de polarización del soluto, otros términos como las interacciones de dispersión-repulsión, que consideran todas las interacciones de van der Waals no clásicas, y la energía de cavitación, que considera el trabajo necesario para crear la cavidad que aloja el soluto en el medio dieléctrico. Normalmente, estos últimos términos se suponen proporcionales al área de la cavidad, ya sea con una única constante de proporcionalidad o con parámetros para cada tipo de átomo ajustados a valores experimentales.

$$G_{soln} = G_{SCRF} + G_{cav} + G_{disp-rep} \quad [3.1]$$

En la **Ecuación 3.1**, la energía calculada se considera una energía libre porque una forma habitual de optimizar los parámetros que determinan las cavidades y las energías de dispersión-repulsión y cavitación es ajustando la diferencia entre las energías calculadas en presencia del disolvente continuo y la energía en fase gas con la energía libre de solvatación experimental del soluto:

$$\Delta G_{solv} = G_{soln} - E_{gas} \quad [3.2]$$

En la **Ecuación 3.2** se asume que las contribuciones térmicas y entrópicas en fase gas y en disolución son iguales. Como en realidad esto no es necesariamente cierto, la diferencia entre estas contribuciones entrópicas y térmicas a la energía libre de solvatación se incluye en el término  $G_{soln}$  como efecto de la parametrización y por este motivo se considera como una energía libre.

Existen distintas formas de definir la cavidad que alberga el soluto, la representación de la distribución de carga del soluto y la forma de calcular la contribución de la dispersión a la energía de interacción [Rivail, 1976], [Rinaldi, 1983], [Dillet, 1994].

Quizá el más conocido de los métodos de disolvente continuo sea el llamado “modelo del continuo polarizable” (Polarizable Continuum Model, PCM) [Miertus, 1981], [Mennucci, 1996], [Olivares del Valle, 1991]. En este modelo la cavidad en la que se introduce el soluto se construye mediante una serie de esferas solapadas, centradas en los núcleos del soluto y con radios proporcionales a los correspondientes radios de Van der Waals. La presencia del soluto en la cavidad induce una polarización en el medio circundante, que se traduce en una densidad de carga aparente,  $\sigma$ , en la superficie de la cavidad. A efectos prácticos, esta superficie se divide en elementos de área  $S_i$  y la densidad de carga superficial se sustituye por unas cargas puntuales, evitándose el empleo de expansiones multipolares. Este método también se llama D-PCM (PCM dieléctrico). Con un fundamento similar al del modelo PCM existen algunas variaciones según la forma de definir las cargas en la superficie de la cavidad. Por ejemplo, el “modelo de

apantallamiento tipo conductor” (Conductor-like Screening Model, COSMO) impone la restricción de que el potencial electrostático ha de ser nulo en la superficie y resulta menos sensible a la forma y tamaño escogidos para la cavidad [Klamt, 1993], [Truong, 1995].

Una implementación particular de COSMO que permite una optimización eficiente de la geometría en disolución se llama C-PCM (PCM conductor).

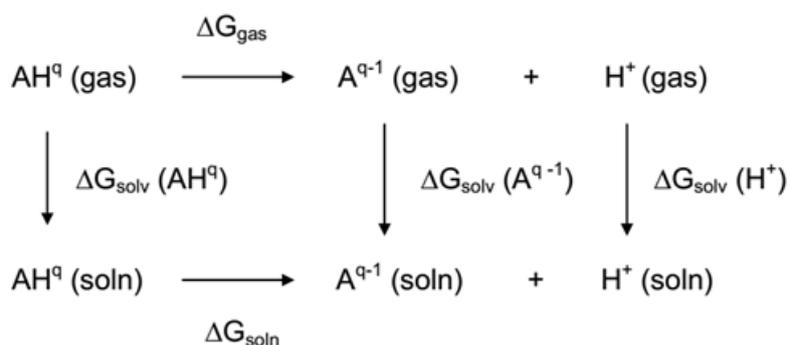
Los métodos IPCM (Isoelectrónico-PCM) y SCI-PCM (Isodensidad autoconsistente-PCM) definen la cavidad como una superficie de densidad electrónica constante del soluto. Así, la cavidad va cambiando durante la optimización de la función de onda del soluto.

### Uso de ciclos termodinámicos

En todos los ciclos termodinámicos considerados, la energía libre de la reacción de disociación del ácido en disolución ( $\Delta G_{soln}$ ) se obtiene como suma de la correspondiente energía libre de desprotonación en fase gaseosa del ácido ( $\Delta G_{gas}$ ) y el incremento de energías libres de solvatación de los reactivos y productos implicados ( $\Delta\Delta G_{solv}$ )

$$\Delta G_{soln} = \Delta G_{gas} + \Delta\Delta G_{solv} \quad [3.3]$$

Los términos de la **Ecuación 3.3** adoptan diferentes expresiones dependiendo del ciclo termodinámico. En el ciclo 1 o también llamado método directo (**Esquema 3.1**) se considera la desprotonación del ácido ( $AH^q$ ) en su base conjugada ( $A^{q-1}$ ) y el protón aislado ( $H^+$ ).



**Esquema 3.1.** Ciclo termodinámico 1 en el que interviene un ácido  $\text{AH}^q$  que se disocia en su base conjugada  $\text{A}^{q-1}$  y un protón.  $\Delta G_{\text{gas}}$  es la energía libre en fase gaseosa de la reacción de desprotonación,  $\Delta G_{\text{soln}}$  es la energía libre en disolución de la reacción de desprotonación y  $\Delta G_{\text{solv}}$  es la energía libre de solvatación. La carga del ácido  $\text{AH}^q$  se representa por  $q$ .

La energía libre en fase gas de la reacción de disociación ( $\Delta G_{\text{gas}}$ ) viene dada por

$$\Delta G_{\text{gas}} = G_{\text{gas}}(\text{H}^+) + G_{\text{gas}}(\text{A}^{q-1}) - G_{\text{gas}}(\text{AH}^q) + \Delta nRT \ln 24.46 \quad [3.4]$$

Las energías libres en fase gas se calculan respecto a un estado estándar de 1 atm pero la constante de acidez corresponde un proceso en disolución, para los que se suele emplear un estado estándar 1 M. El último término de la **Ecuación 3.4** corresponde a la variación de energía asociada al cambio del estado estándar de 1 atm a 1 M. En este término  $\Delta n$  indica la variación de moles o moléculas entre reactivos y productos.

El valor de la energía libre en fase gas del protón,  $G_{\text{gas}}(\text{H}^+)$ , puede calcularse con la ecuación de Sackur-Tetrode [McQuarrie, 1970], que a 298 K y 1 atm proporciona un valor de 7.76 kcal/mol al producto TS. Considerando la energía de traslación para una especie monoatómica en fase gas ( $1.5RT$ ), la ecuación de estado de un gas ideal y la definición de entalpía, se obtiene que el valor de la entalpía es de  $5/2RT$  (1.48 kcal/mol a 298 K y 1 atm). Finalmente, se obtiene que el valor de la energía libre en fase gas del protón a 298 K y 1 atm es de -6.28 kcal/mol.

El incremento de energía libre de solvatación ( $\Delta \Delta G_{\text{solv}}$ ) entre reactivos y productos puede calcularse como

$$\Delta\Delta G_{solv} = \Delta G_{solv}(H^+) + \Delta G_{solv}(A^{q-1}) - \Delta G_{solv}(AH^q) \quad [3.5]$$

En esta ecuación se requiere el valor de la energía libre de solvatación del protón,  $\Delta G_{solv}(H^+)$ , que debe determinarse experimentalmente. Los valores experimentales de  $\Delta G_{solv}(H^+)$  descritos en la literatura varían considerablemente entre -259 y -264 kcal/mol. Esta dispersión de valores ha originado diferencias en los valores del  $pK_a$  calculados de hasta 3 unidades. Además, el estado estándar en el que se han calculado estos valores ha originado cierta confusión. Camaioni y Schwerdtfeger [Camaioni, 2005] han propuesto el valor de -265.9 kcal/mol, valor determinado por Tissandier *et al.* [Tissandier, 1998] en 1998 y confirmado por Kelly *et al.* en 2006 [Kelly, 2006].

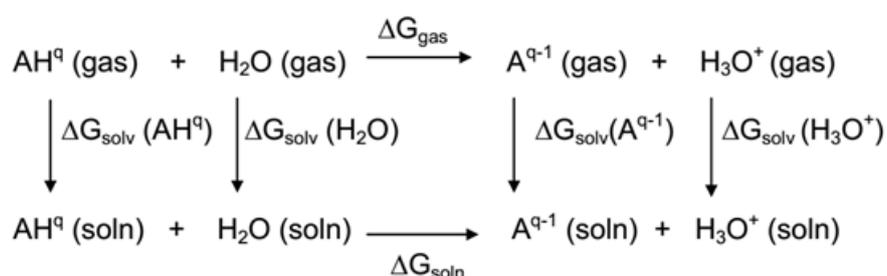
Finalmente, empleando las **Ecuaciones 3.4** y **3.5**, se obtiene la energía de desprotonación del ácido  $AH^q$  en disolución y el  $pK_a$  correspondiente se calcula como

$$pK_a = \frac{\Delta G_{aq}}{RT \ln 10} \quad [3.6]$$

Este ciclo ha sido muy utilizado en el cálculo de  $pK_a$ s. Liptak *et al.* han calculado el  $pK_a$  de fenoles [Liptak, 2002] y ácidos carboxílicos alifáticos [Liptak, 2001] con un error de 1 unidad de  $pK_a$ . También se ha empleado en el caso de ácidos inorgánicos [da Silva, 2006], aminas [da Silva, 2003] derivados de ácidos carboxílicos [Król, 2006], [Namazian, 2007], [Dong, 2008] y piridinas [Casasnovas, 2009]. Los resultados en el caso de ácidos carboxílicos, fenoles y piridinas han resultado buenos (errores del orden de 0.5 unidades de  $pK_a$ ) sin embargo en el caso de ácidos de carbono los errores han sido del orden de 7 unidades de  $pK_a$  [Gao, 2005], [Ho, 2009]. En un estudio realizado con 55 ácidos de diferentes tipos (ácidos carboxílicos, ácidos inorgánicos, fenoles, alcoholes y ácidos de carbono) y cinco modelos de solvatación (COSMO-RS, IPCM, SM6, CPCM-UAHF, CPCM-UAKS), Ho y Coote [Ho, 2010A] determinaron que la precisión de los resultados obtenidos con este ciclo depende del modelo de disolvente continuo usado para calcular las energías libres en disolución así como del tipo de ácido. Los modelos CPCM-UAHF y SM6 aplicados a fenoles, alcoholes y ácidos carboxílicos son los que producen errores promedio inferiores.

La incertidumbre en el valor de  $\Delta G_{solv}(H^+)$  y los errores asociados en el cálculo del  $pK_a$ , motivaron a diferentes grupos de investigación a considerar otros ciclos termodinámicos.

Uno de dichos ciclos, ciclo 2, consiste en la reacción de intercambio de protón entre un ácido y el agua (**Esquema 3.2**). Este ciclo requiere las energías libres en fase gaseosa y de solvatación del agua y del hidronio, que pueden ser calculadas teóricamente o determinadas experimentalmente. Al igual que en el ciclo 1, hay que considerar el cambio de estado estándar de 1 atm a 1 M para las energías de desprotonación en fase gaseosa, sin embargo, como el número de moléculas a ambos lados de la ecuación es el mismo, las energías libres de desprotonación a 1 atm y a 1 M son iguales.



**Esquema 3.2.** Ciclo termodinámico en el que interviene un ácido  $\text{AH}^q$  que cede un protón a una molécula de agua, obteniéndose la base conjugada  $\text{A}^{q-1}$  y el ion hidronio.  $\Delta G_{\text{gas}}$  es la energía libre en fase gaseosa y  $\Delta G_{\text{soln}}$  es la energía libre de la reacción de desprotonación en disolución.  $\Delta G_{\text{solv}}$  es la energía libre de solvatación. Las cargas del ácido y de su base conjugada se representan por  $q$  y  $q-1$  respectivamente.

En este ciclo, la energía libre en fase gas de la reacción de disociación,  $\Delta G_{\text{gas}}$ , y el incremento de energías libres de solvatación,  $\Delta \Delta G_{\text{solv}}$ , vienen dados por la **Ecuaciones 3.7** y **3.8** respectivamente.

$$\Delta G_{\text{gas}} = G_{\text{gas}}(\text{H}_3\text{O}^+) + G_{\text{gas}}(\text{A}^{q-1}) - G_{\text{gas}}(\text{AH}^q) - G_{\text{gas}}(\text{H}_2\text{O}) \quad [3.7]$$

$$\Delta \Delta G_{\text{solv}} = \Delta G_{\text{solv}}(\text{H}_3\text{O}^+) + \Delta G_{\text{solv}}(\text{A}^{q-1}) - \Delta G_{\text{solv}}(\text{AH}^q) - \Delta G_{\text{solv}}(\text{H}_2\text{O}) \quad [3.8]$$

En este caso la energía libre del proceso en disolución no corresponde directamente con la constante de acidez sino con la constante de equilibrio de la reacción de intercambio protónico con agua:

$$K_{eq} = \frac{[H_3O^+][A^{q-1}]}{[AH^q][H_2O]} = \frac{K_a}{[H_2O]} \quad [3.9]$$

Teniendo en cuenta las **Ecuaciones 3.9** y **3.6**, el  $pK_a$  del ácido  $AH^q$  puede calcularse como

$$pK_a = \frac{\Delta G_{aq}}{RT \ln 10} - \log[H_2O] \quad [3.10]$$

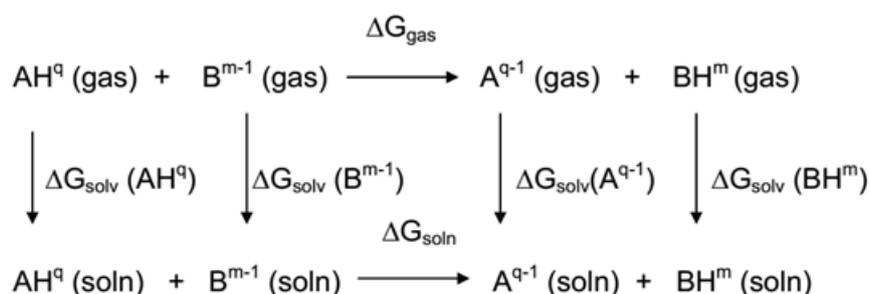
Este ciclo ha sido menos utilizado que el ciclo 1 en el cálculo de  $pK_a$ s en disolución [Casasnovas, 2009], [Liptak, 2001], [Pliego, 2002], [Brown, 2006].

Una variación del ciclo 2 se basa en la reacción entre un ácido y el hidroxilo para generar la base conjugada y agua [Pliego, 2002], [Brown, 2006].

A diferencia del ciclo 1, ni en el ciclo 2 mostrado en el **Esquema 3.2** ni en su variante se considera el protón aislado por lo que se evita la problemática asociada a la incertidumbre en la energía libre de solvatación de esta especie. Aquellos grupos que emplean valores experimentales consideran los valores de -110.3 kcal/mol para  $\Delta G_{solv}(H_3O^+)$ , -6.32 kcal/mol para  $\Delta G_{solv}(H_2O)$  y -104.7 kcal/mol para  $\Delta G_{solv}(OH^-)$  para un estado de referencia 1 M [Tissandier, 1998], [Kelly, 2006]. Por otra parte, este ciclo puede disminuir los errores asociados a las energías de solvatación de las especies iónicas ya que se puede combinar la disociación del ácido estudiado frente a una molécula de agua o un ion hidroxilo de forma que haya una especie iónica del mismo signo a cada lado de la ecuación química. De esta forma, si el error en las energías de solvatación de las especies iónicas es sistemático, se puede incrementar la precisión en la determinación del  $pK_a$  debido a la cancelación de errores.

Un ciclo alternativo a los ciclos 1 y 2 es el llamado de intercambio de protón, ciclo 3 (**Esquema 3.3**), que está basado en el concepto de reacción isodésmica. Una reacción isodésmica se caracteriza porque el número y tipo de enlaces que se rompen y se forman es el mismo. En este ciclo se evita el uso de las energías libres de solvatación

del protón, agua y del ion hidronio o hidroxilo. La principal razón por la que se plantea el tercer ciclo es para poder considerar especies con igual carga total en ambos lados de la reacción y así mejorar la cancelación de errores entre las energías de solvatación. Este ciclo proporciona valores relativos del  $pK_a$  y necesita de una especie de referencia ( $BH^m$ ) que actúa como base y cuyo  $pK_a$  es conocido.



**Esquema 3.3.** Ciclo termodinámico en el que un ácido  $AH^q$  cede un protón a una base  $B^{m-1}$ . Se obtiene su base conjugada  $A^{q-1}$  y el ácido de referencia  $BH^m$ .  $\Delta G_{gas}$  es la energía libre de la reacción ácido-base en fase gaseosa y  $\Delta G_{soln}$  es la energía libre en disolución de la reacción ácido-base.  $\Delta G_{solv}$  es la energía libre de solvatación. Las cargas del ácido y de su base conjugada se representan por  $q$  y  $q-1$  respectivamente. Las cargas del ácido de referencia y su base conjugada se representan por  $m$  y  $m-1$  respectivamente.

La constante de equilibrio de la reacción de intercambio protónico se puede expresar como el cociente de las constantes de acidez del ácido estudiado  $AH^q$  y del ácido de referencia  $BH^m$

$$K_{eq} = \frac{[BH^m][A^{q-1}]}{[AH^q][B^{m-1}]} = \frac{K_{a(AH)}}{K_{a(BH)}} \quad [3.11]$$

donde el  $pK_a$  del ácido  $AH^q$  se obtiene a partir de la **Ecuación 3.12** en la que  $pK_a(BH)$  es el  $pK_a$  del ácido de referencia.

$$pK_{a(AH)} = \frac{\Delta G_{aq}}{RT \ln 10} + pK_{a(BH)} \quad [3.12]$$

La energía libre de la reacción en disolución,  $\Delta G_{soln}$ , se calcula con la **Ecuación 3.3**, la energía de desprotonación en fase gas,  $\Delta G_{gas}$ , y la diferencia de energías de solvatación,  $\Delta \Delta G_{solv}$ , adoptan la forma

$$\Delta G_{gas} = G_{gas}(BH^m) + G_{gas}(A^{q-1}) - G_{gas}(AH^q) - G_{gas}(B^{m-1}) \quad [3.13]$$

$$\Delta \Delta G_{solv} = \Delta G_{solv}(BH^m) + \Delta G_{solv}(A^{q-1}) - \Delta G_{solv}(AH^q) - \Delta G_{solv}(B^{m-1}) \quad [3.14]$$

Como se ha indicado, este ciclo presenta la ventaja de que no emplea la energía de solvatación del protón aislado ( $\Delta G_{solv}(H^+)$ ) pero posee la desventaja de que hay que elegir un ácido de referencia ( $BH^m$ ). Otra ventaja respecto al ciclo 1 es que la presencia de especies de igual carga eléctrica a ambos lados de la ecuación química mejora la predicción de los  $pK_a$ s por cancelación de errores en las energías de solvatación calculadas [Kelly, 2006], [Cramer, 1999]. Además, en el ciclo 3 se emplea una molécula de referencia químicamente similar al ácido del que se desea calcular el  $pK_a$ , por lo que implica una mayor cancelación de errores en las energías libres en fase gaseosa.

Este ciclo ha sido utilizado en numerosas ocasiones [Dong, 2008], [Casasnovas, 2009], [Toth, 2001], [Namazian, 2006], [Caballero, 2006], [Ding, 2009], [Namazian, 2003] y generalmente con mejores resultados que el ciclo 1. Al igual que el ciclo 1, los resultados dependen del tipo de ácido del que se calcula el  $pK_a$  y del método de solvatación empleado. Ho y Coote [Ho, 2010A] han calculado el  $pK_a$  de cuatro grupos funcionales empleando como ácido de referencia el metanol, el fenol, el ácido fórmico y la metilamida respectivamente con cuatro modelos de disolvente continuo y, en el caso del modelo CPCM, dos formas de calcular la cavidad del soluto: CPCM-UAKS, CPCM-UAHF, SM6, IPCM y COSMO-RS. El menor error obtenido es de aproximadamente 1 unidad de  $pK_a$  en alcoholes con el modelo CPCM y, por otra parte, los errores máximos están en torno a las 4 unidades  $pK_a$ .

### **La reacción isodésmica**

Para aquellas especies químicas que no son estables en fase gas, o presentan grandes cambios conformacionales entre fase gas y disolución, el uso de ciclos termodinámicos presenta dificultades. Para evitar el cálculo de las energías libres en fase gaseosa nuestro grupo de investigación ha aplicado recientemente con éxito el esquema de la reacción

isodésmica [Casasnovas, 2011], [Casasnovas, 2012], [Casasnovas, 2013A], [Casasnovas, 2013B], caracterizada porque el número y el tipo de enlaces que se rompen y se forman es el mismo (**Esquema 3.4**).



**Esquema 3.4.** Reacción isodésmica empleada para el cálculo del  $pK_a$  de  $AH^q$ .  $\Delta G_{\text{soln}}$  es la energía libre en disolución de la reacción ácido-base. La carga del ácido y de su base conjugada se representan por  $q$  y  $q-1$  respectivamente. Las cargas del ácido de referencia y su base conjugada se representan por  $m$  y  $m-1$  respectivamente.

La reacción de intercambio protónico planteada en la reacción isodésmica es idéntica a la planteada en el ciclo 3, por lo que el  $pK_a$  del ácido  $AH^q$  se obtiene a partir del  $pK_a$  de un ácido de referencia,  $BH^m$ , según la **Ecuación 3.12**. La diferencia con el ciclo 3, es que la energía libre de la reacción en disolución,  $\Delta G_{\text{soln}}$ , se calcula a partir de las energías libres de cada soluto en disolución:

$$\Delta G_{\text{soln}} = G_{\text{soln}}(BH^m) + G_{\text{soln}}(A^{q-1}) - G_{\text{soln}}(AH^q) - G_{\text{soln}}(B^{m-1}) \quad [3.15]$$

Así hemos calculado el  $pK_a$  de piridinas mono y disustituidas, de ácidos de carbono [Casasnovas, 2011], [Casasnovas, 2012], [Casasnovas, 2013A] y de fenoles, piridinas y bencilaminas de complejos de piridoxamina y cobre [Casasnovas, 2013B]. En el caso de las piridinas, se observó que al emplear ciclos termodinámicos era necesario incluir una molécula de agua explícita para obtener  $pK_a$ s con buena precisión [Casasnovas, 2009]. Sin embargo, al emplear la reacción isodésmica, se obtuvo la misma precisión sin considerar moléculas de agua de forma explícita [Casasnovas, 2011]. Con los ácidos de carbono se han obtenido unos valores de la desviación media de unas 2 unidades de  $pK_a$  que son de la misma precisión que los mejores resultados obtenidos con ciclos termodinámicos y similares a la incertidumbre experimental para estas especies [Gao, 2005], [Ho, 2009], con la ventaja que la carga de la molécula de referencia es menos determinante [Casasnovas, 2011]. Posteriormente, en el caso de ácidos de carbono, se empleó un conjunto de moléculas de prueba en un rango de 30 unidades de  $pK_a$  para calcular las desviaciones de las energías de desprotonación en disolución debidas al modelo de disolvente continuo. Tras una regresión lineal se obtuvieron los parámetros para corregir las energías calculadas en estas especies y obtener  $pK_a$ s con mayor precisión

en todo el rango considerado [Casasnovas, 2012], [Casasnovas, 2013A]. La reacción isodésmica se empleó para calcular  $pK_a$ s de los complejos de piridoxamina y cobre, obteniendo una media de las desviaciones absolutas, MAD de 0.2 unidades de  $pK_a$  [Casasnovas, 2013B].

Sin embargo, las energías libres en disolución no se pueden calcular mediante las aproximaciones del oscilador armónico y del rotor rígido debido a que no representan correctamente los movimientos nucleares de moléculas en disolución. Nuestro grupo [Casasnovas, 2011] ha propuesto calcular las energías libres en disolución asumiendo que las funciones de partición vibracionales correspondientes al oscilador armónico son válidas en disolución y que las contribuciones entrópicas de los movimientos de rotación y translación correspondientes a reactivos y productos tienen magnitudes similares, por lo que, de acuerdo con el **Esquema 3.4** y la **Ecuación 3.15**, se anulan en gran medida y no contribuyen significativamente al valor de la energía de desprotonación,  $\Delta G_{soln}$ .

Por ello, la energía libre de cada soluto en disolución,  $G_{soln}$ , se aproxima sumando tres términos (**Ecuación 3.16**). El primero es la energía potencial del soluto incluyendo la polarización mutua con el disolvente continuo debido a la interacción electrostática soluto-dieléctrico,  $G_{SCRF}$ . En el segundo término,  $G_{nes}$ , se incluyen todas las interacciones no electrostáticas entre el soluto y el disolvente continuo: dispersión, repulsión y cavitación. El tercer término,  $\Delta G_{corr\_soln}$ , corresponde a la contribución vibracional de los núcleos a 298 K.

$$G_{soln} \approx G_{SCRF} + G_{nes} + \Delta G_{corr\_soln} \quad [3.16]$$

Hasta el momento, la reacción isodésmica ha sido empleada para calcular entalpías de formación [Wang, 2008], pero lo ha sido en menor medida que los ciclos termodinámicos en el cálculo de constantes de acidez. Govender y Cukrowski [Govender, 2009], [Govender, 2010] han empleado este método junto con el modelo de solvatación PCM con esferas UA0 para calcular las constantes de disociación sucesivas del ácido nitrilotripropanoico y del nitrilotriacético con un error de entre 0.2 y 3 unidades de  $pK_a$ .

## Cálculo del $pK_a$ de aminoácidos y péptidos

La determinación computacional del  $pK_a$  de aminoácidos es un paradigma de especies cuyos tautómeros mayoritarios en disolución y en fase gas son diferentes. Existen pocos trabajos del cálculo teórico de los  $pK_a$ s correspondientes a la ionización del grupo  $\alpha$ -carboxílico ( $pK_{a1}$ ), del grupo  $\alpha$ -amino ( $pK_{a2}$ ) y de la cadena lateral ( $pK_{a3}$ ) de los aminoácidos. Kiani *et al.* [Kiani, 2010] han determinado los valores del  $pK_{a1}$  y del  $pK_{a2}$  de la glicina, valina, fenilalanina, glicilvalina y glicilfenilalanina con una MAD de 0.32 y 0.40 unidades de  $pK_a$  respectivamente. Kiani *et al.* han utilizado el modelo de solvatación PCM, bases B3LYP/6-31+G(d) para calcular la variación de energía libre en disolución, y por lo tanto el  $pK_a$  de diversas reacciones del aminoácido solvatado con moléculas de agua. Gupta *et al.* [Gupta, 2013] han calculado el  $pK_{a1}$  y el  $pK_{a2}$  de 10 aminoácidos y el  $pK_{a3}$  de 2 aminoácidos. Han usado el ciclo termodinámico 1, diferentes modelos continuos de solvatación (PCM, SM8T y DivCon) y un modelo híbrido Explicit Solvation Shell Model, ESS) en el que se representa explícitamente la primera capa de solvatación con moléculas del disolvente y el clúster resultante se rodea con un modelo continuo. Los mejores resultados en el cálculo del  $pK_{a1}$ ,  $pK_{a2}$  y  $pK_{a3}$  los han obtenido con el modelo mixto ESS con unas MADs de 3.18, 1.78 y 1.65 unidades de  $pK_a$  respectivamente. Nuestro grupo de investigación ha usado la reacción isodésmica para calcular los valores del  $pK_a$  correspondientes a la ionización del grupo  $\alpha$ -carboxílico y del grupo  $\alpha$ -amino de algunos aminoácidos con una desviación media de 0.22 y 0.19 unidades de  $pK_a$  respectivamente [Casasnovas, 2011]. Estos valores son mucho menores que los obtenidos por Gupta *et al.* y sin la necesidad de usar moléculas de agua de forma explícita. Recientemente Ho [Ho, 2015] también ha usado la reacción isodésmica para calcular el  $pK_{a1}$  y el  $pK_{a2}$  de una serie de aminoácidos apolares, con resultados mejores que usando ciclos termodinámicos.

### 3.2. DFT conceptual

La teoría del funcional de la densidad permite definir rigurosamente conceptos de reactividad química según la teoría denominada DFT conceptual. Por ejemplo, el potencial químico se define como

$$\mu = \left( \frac{\partial E}{\partial N} \right)_v \quad [3.17]$$

Empleando diferencias finitas para diferenciar 3.17 se obtiene

$$\mu = \frac{\Delta E}{\Delta N} = \frac{E(N+1) - E(N-1)}{(N+1) - (N-1)} = \frac{(E(N) - A) - (E(N) + I)}{2} = -\frac{A+I}{2} = -\chi \quad [3.18]$$

donde  $A$  e  $I$  son respectivamente la primera afinidad electrónica y el primer potencial de ionización y  $\chi$  es la electronegatividad según la definición de Mulliken. Según la aproximación de Koopmans

$$\mu \approx -\chi \approx \frac{\varepsilon_{HOMO} + \varepsilon_{LUMO}}{2} \quad [3.19]$$

donde  $\varepsilon_{HOMO}$  y  $\varepsilon_{LUMO}$  son respectivamente las energías de los orbitales HOMO y LUMO.

Por otra parte, la dureza química se puede definir como

$$\eta = \left( \frac{\partial^2 E}{\partial N^2} \right)_v \quad [3.20]$$

Empleando diferencias finitas

$$\eta \approx \frac{\varepsilon_{LUMO} - \varepsilon_{HOMO}}{2} \quad [3.21]$$

La electrofilia  $\omega$  se define como

$$\omega = \frac{\mu^2}{2\eta} = \frac{(I+A)^2}{4(I-A)} \approx \frac{(\varepsilon_{LUMO} + \varepsilon_{HOMO})^2}{4(\varepsilon_{LUMO} - \varepsilon_{HOMO})} \quad [3.22]$$

La función de Fukui indica cómo varía la densidad electrónica por un cambio de electrones del sistema

$$f(\vec{r}) = \left( \frac{\partial \rho(\vec{r})}{\partial N} \right)_v \quad [3.23]$$

Otra interpretación de la función de Fukui es que refleja la capacidad de una molécula a aceptar o donar electrones. Altos valores de  $f(\vec{r})$  se asocian con alta reactividad en  $r$ .

Morell *et al.* [Toro-Labbé, 2007], [Morell, 2005], [Morell, 2006] propusieron un descriptor local de la reactividad denominado descriptor dual:

$$\Delta f(\vec{r}) = \left( \frac{\partial f(\vec{r})}{\partial N} \right)_v \quad [3.24]$$

El descriptor dual, que es una función espacial que toma el rango de la densidad electrónica molecular, puede condensarse sobre átomos de forma que valores positivos indican que el átomo es un centro electrófilo y valores negativos indican que es un centro nucleófilo. El procedimiento para calcular el descriptor dual condensado sobre un átomo en particular requiere la integración espacial de la densidad electrónica sobre dicho átomo. Dado que los límites de los átomos dentro de una molécula no están definidos, para realizar la integración debe emplearse un esquema de partición de cargas atómicas como el de Mulliken, Hirschfeld o Bader entre otros [Zielinski, 2012]. Martínez-Araya [Martínez-Araya, 2015] ha mostrado que el descriptor dual permite identificar las regiones electrófilas y nucleófilas de una molécula de manera inequívoca.

Domingo [Domingo, 2013], [Chamorro, 2013] definió las funciones de Parr como

$$P^+(\vec{r}) = \rho_s^{rc}(\vec{r}) \quad [3.24]$$

para ataques electrófilos. Y

$$P^-(\vec{r}) = \rho_s^{ra}(\vec{r}) \quad [3.25]$$

para ataques nucleófilos. Estas funciones están relacionadas con la densidad de espín atómica del átomo en  $\vec{r}$  del catión o anión de una molécula determinada. La densidad de espín atómica sobre cada átomo del catión radical y del anión radical de la molécula proporciona las funciones de Parr locales nucleófila  $P^-(\vec{r})$  y electrófila  $P^+(\vec{r})$  de la molécula neutra [Domingo, 2016].

### 3.3. Detalles computacionales

#### Cálculo de $pK_{as}$ de compuestos orgánicos

Todos los cálculos se han realizado con el programa Gaussian03 [Frisch, 2004], Gaussian09 [Frisch, 2009] usando tres métodos teóricos para obtener los valores de las energías libres: B3LYP [Becke, 1993], [Lee, 1988], M05-2X [Zhao, 2006] y CBS-4B3\* [Casasnovas, 2010]. Todas las estructuras se han optimizado y se han caracterizado como mínimos energéticos por la ausencia de frecuencias imaginarias empleando las bases 6-311++G(d,p). El método CBS-4B3\* [Casasnovas, 2010] es una simplificación del método CBS-QB3 en el que se han eliminado los cálculos CCSD(T), la corrección energética correspondiente a la contaminación de espín y las correcciones empíricas. Este método presenta ciertas ventajas frente al CBS-QB3 pues con un coste computacional menor, ya que se han eliminado los cálculos CCSD(T), se obtienen resultados de igual precisión de la energía libre de protonación en fase gas [Casasnovas, 2010].

Para considerar los efectos del disolvente se ha utilizado el modelo CPCM [Ho, 2009], [Barone, 1998A], [Barone, 1998B] junto con las cavidades UAKS implementado en el programa Gaussian 03 [Frisch, 2004] y el modelo SMD con el programa Gaussian 09 [Frisch, 2009]. El modelo SMD utiliza su propia definición de cavidades.

### **Cálculo de pK<sub>a</sub>s de aminoácidos y péptidos**

Todos los cálculos se han realizado con el programa Gaussian09 [Frisch, 2009] usando los funcionales de correlación intercambio DFT M05-2X [Zhao, 2006], M06-2X [Zhao, 2008A] y B3LYP [Becke, 1993], [Lee, 1988] con bases 6-31+G(d,p) y el Hamiltoniano semiempírico PM6 [Stewart, 2007]. En todos los casos se ha empleado el modelo de disolvente continuo SMD [Marenich, 2009]. Las estructuras se optimizaron y se caracterizaron como mínimos energéticos por la ausencia de frecuencias imaginarias.

En un trabajo anterior (cálculo de pK<sub>a</sub>s de compuestos orgánicos) se ha empleado el método CBS-4B3\* y en otros trabajos bibliográficos también se han empleado métodos como CBS-QB3 [Casasnovas, 2011] para el cálculo de pK<sub>a</sub>s. De esos trabajos se puede concluir que la precisión que se obtiene con métodos post-HF de alto nivel es similar a la precisión de métodos DFT. Por este motivo, y por eficiencia computacional, en este trabajo no se emplean métodos post-HF.

### **Predicción de los sitios de protonación de péptidos tipo Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> mediante descriptores de DFT conceptual**

Las estructuras de los seis péptidos con la fórmula Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> (0 ≤ n ≤ 5) se construyeron con el software Avogadro 1.2.0 [Avogadro, 2016], [Hanweel, 2012]. Las estructuras resultantes fueron pre-optimizadas realizando un muestreo aleatorio de los diedros para encontrar los conformeros más estables según el campo de fuerzas general de Amber (GAFF) [Wang, 2004]. Los conformeros resultantes fueron optimizados con cálculos DFT empleando los funcionales M11 [Peverati, 2011], M11L [Peverati, 2012A], MN12L [Peverati, 2012B], MN12SX [Peverati, 2012C], N12 [Peverati, 2012D] y N12X [Peverati, 2012C] con las bases Def2SVP y con el modelo de disolvente continuo SMD [Marenich, 2009] para simular la disolución acuosa. Las estructuras optimizadas se caracterizaron como mínimos energéticos por la ausencia de frecuencias de vibración imaginarias. Para el cálculo de los descriptores de DFT conceptual se hicieron cálculos puntuales con las bases Def2TZVP sobre las geometrías optimizadas con bases Def2SVP. Todos los cálculos DFT se realizaron con el software Gaussian09 [Frisch, 2009].

### **Capacidad antioxidante de la tiamina**

Todos los cálculos se han realizado con el software Gaussian 09 [Frisch, 2009]. Los cálculos basados en la Teoría del Funcional de la Densidad (DFT) se han realizado utilizando el funcional M05-2X [Zhao, 2006] con el conjunto de bases 6-311+G(d,p). El

funcional M05-2X ha sido recomendado para cálculos cinéticos [Zhao, 2006] y es también uno de los mejores funcionales para el cálculo de las energías de reacción en las que participan radicales libres [Zhao, 2008B]. En todos los casos han llevado a cabo cálculos no restringidos (unrestricted) para sistemas de capa abierta.

Los efectos del disolvente se han introducido con el modelo continuo SMD [Marenich, 2009] utilizando agua o etanoato de pentilo (PE) para imitar el medio lipídico/no polar de las membranas celulares.

Se han realizado optimizaciones de la geometría de todas las estructuras. Los mínimos energéticos se han caracterizado por la ausencia de frecuencias imaginarias y las estructuras del estado de transición lo han sido por la presencia de una única frecuencia imaginaria que corresponde al movimiento esperado a lo largo de la coordenada de reacción estudiada. Las energías del punto cero (ZPE), las correcciones térmicas y el estado estándar 1M se han utilizado para calcular la energía de Gibbs a 298.15 K.

## **4. Resultados y discusión**



#### 4.1. Cálculo de $pK_a$ s de compuestos orgánicos haciendo uso de la reacción isodésmica

Anteriormente se ha empleado la reacción isodésmica (**Esquema 3.4**) para calcular el  $pK_a$  de piridinas sustituidas, ácidos de carbono y el  $pK_{a1}$  y  $pK_{a2}$  de algunos aminoácidos neutros [Casasnovas, 2011]. En este trabajo se ha utilizado la reacción isodésmica (**Esquema 3.4**) para el cálculo de  $pK_a$ s de moléculas orgánicas con grupos ácido-base comunes en disolución acuosa. Se han elegido varios compuestos orgánicos pertenecientes a alcoholes alifáticos, ácidos carboxílicos, aminas alifáticas y sus equivalentes aromáticos: fenoles monosustituidos, ácidos benzoicos monosustituidos y piridinas sustituidas (**Tabla 4.1**). El  $pK_a$  de las moléculas de interés se ha calculado con las **Ecuaciones 3.15** y **3.12**.

**Tabla 4.1.** Piridinas, Alcoholes Alifáticos, Ácidos Carboxílicos, Aminas, Fenoles y Ácidos Benzoicos elegidos.

Piridinas	Alcoholes	Ácidos carboxílicos	Aminas	Fenoles	Ácidos benzoicos
2-metilpiridina	metanol	cloroacético	metilamina	p-cianofenol	o-clorobenzoico
3-metilpiridina	2-cloroetanol	fórmico	propilamina	o-clorofenol	m-clorobenzoico
4-metilpiridina	propanol	3-clorobutanoico	i-propilamina	m-cianofenol	p-clorobenzoico
2,3-dimetilpiridina	i-propanol	benzoico	butilamina	m-clorofenol	p-metilbenzoico
2,4-dimetilpiridina	2-butanol	4-clorobutanoico	2-butilamina	m-fluorofenol	m-metilbenzoico
3-bromopiridina	tert-butanol	hexanoico	tert-butilamina	p-clorofenol	p-fluorobenzoico
3-fluoropiridina		propanoico	trimetilamina	p-fluorofenol	
3-cianopiridina		pentanoico	dimetilamina	m-metilfenol	
3-cloropiridina		trimetilacético		p-metilfenol	
				o-metilfenol	

En lo que respecta a la elección del ácido de referencia, se ha elegido el criterio de que la especie  $AH^q$  y el compuesto de referencia  $BH^m$  tengan una estructura semejante. Así en el caso de los alcoholes, ácidos carboxílicos y aminas, se han elegido el etanol, el ácido etanoico y el ion etilamonio respectivamente. En el caso de los fenoles monosustituidos, ácidos benzoicos monosustituidos y piridinas el compuesto de referencia utilizado ha sido el fenol, el ácido benzoico y el ion piridinio, respectivamente.

En la **Tabla 4.2** se muestran las medias de las desviaciones absolutas (*Mean Absolute Deviation* MADs), desviaciones estándar (*Standard Deviation* SD) y

desviaciones absolutas máximas (*Absolute deviation* AD<sub>max</sub>) entre los valores de los pK<sub>a</sub>s teóricos obtenidos con la reacción isodésmica respecto a los valores experimentales, clasificados según el grupo funcional, el modelo de solvatación (CPCM y SMD) y el método de cálculo empleado (CBS-4B3\*, B3LYP y M05-2X).

**Tabla 4.2.** Media de las desviaciones absolutas (MADs), desviación estándar (SD) y máxima desviación absoluta (AD<sub>max</sub>) de los valores del pK<sub>a</sub> en disolución acuosa calculados usando diferentes métodos (CBS-4B3\*/6-311++G(d,p), B3LYP/6-311++G(d,p) y M05-2X/6-311++G(d,p)), y modelos de disolvente (CPCM y SMD).

		CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Piridinas</b>	MAD	0.57	0.75	0.83	0.62	0.80	0.78
	SD	0.48	0.74	0.55	0.36	0.39	0.44
	AD max	1.44	2.17	1.87	1.29	1.34	1.55
<b>Alcoholes</b>	MAD	0.85	0.68	0.62	1.20	1.01	1.29
	SD	0.49	0.46	0.51	0.96	0.75	0.97
	AD max	1.54	1.19	1.22	2.84	2.27	2.85
<b>Ácidos carboxílicos</b>	MAD	0.78	0.98	0.79	0.57	0.64	0.67
	SD	0.42	0.36	0.47	0.36	0.54	0.59
	AD max	1.32	1.44	1.51	1.13	1.54	1.89
<b>Aminas</b>	MAD	0.87	0.72	0.90	0.24	0.35	0.27
	SD	0.84	0.74	1.00	0.19	0.25	0.20
	AD max	2.51	2.20	2.89	0.63	0.83	0.71
<b>Fenoles</b>	MAD	0.90	1.08	1.02	0.70	0.89	0.87
	SD	0.57	0.68	0.68	0.48	0.57	0.56
	AD max	2.16	2.71	2.57	1.66	2.18	2.10
<b>Ácidos benzoicos</b>	MAD	0.45	0.57	0.41	0.41	0.50	0.34
	SD	0.29	0.41	0.30	0.35	0.48	0.29
	AD max	0.95	1.36	0.82	1.05	1.39	0.69

Los valores de las MADs son inferiores a 1 unidad de pK<sub>a</sub>, exceptuando los valores calculados para los alcoholes, que muestran valores inferiores a 1.3 unidades de pK<sub>a</sub>. Los mejores resultados corresponden a las aminas primarias y a los ácidos benzoicos con una MAD inferior a 0.35 y 0.50 unidades de pK<sub>a</sub> respectivamente (**Tabla 4.2**). En el caso de las aminas, si se consideran las aminas primarias junto a las secundarias y terciarias, los valores de las desviaciones medias con el modelo de solvatación CPCM son de unas 0.8 unidades de pK<sub>a</sub>. (**Tabla A3**).

Comparando entre sí los valores de las MADs obtenidos con el modelo de solvatación CPCM frente al modelo SMD, los valores más bajos corresponden al modelo SMD para los ácidos carboxílicos, las aminas, los fenoles y los ácidos benzoicos, sea cual sea el método teórico empleado. En el caso de las piridinas y de las aminas primarias,

apenas hay diferencias entre ambos modelos. Para los alcoholes, ocurre lo contrario ya que los valores de las MADs son menores con el modelo de solvatación CPCM.

Teniendo en cuenta el método de cálculo, los mejores resultados corresponden al método CBS-4B3\* para las piridinas (CPCM), ácidos carboxílicos, aminas, y fenoles (SMD), con unas MADs de 0.57, 0.57, 0.24 y 0.70 respectivamente (**Tabla 4.2**) y al M05-2X para los alcoholes (CPCM) y ácidos benzoicos (SMD), con unas desviaciones medias de 0.62 y 0.34 (**Tabla 4.2**).

Los resultados de la **Tabla 4.2** muestran que los  $pK_a$ s obtenidos con la reacción isodésmica se obtienen con errores bajos para todos los grupos funcionales independientemente de la combinación de método teórico y modelo de disolvente continuo.

Se ha evaluado la influencia de la elección del ácido de referencia en los valores de  $pK_a$  calculados para aminas alifáticas y ácidos carboxílicos. En las **Tablas 4.3** y **4.4**, se muestran los valores de MAD obtenidos empleando diferentes ácidos de referencia. En el caso de las aminas, la media de las desviaciones medias se encuentra entre 0.28 y 0.78 unidades de  $pK_a$ , con pequeñas diferencias dependiendo de si el compuesto de referencia es una amina primaria, secundaria o terciaria (**Tabla 4.3**). Para los ácidos carboxílicos, las MADs oscilan entre 0.5 y 1 unidades de  $pK_a$ , excepto si el compuesto de referencia es el ácido cloroacético y el ácido fórmico (**Tabla 4.4**).

**Tabla 4.3.** Efecto de la especie de referencia en la media de las desviaciones absolutas (MADs) de los  $pK_a$ s de aminas alifáticas.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Etilamina	0.87	0.72	0.90	0.24	0.35	0.27
Metilamina	0.61	0.93	0.80	0.61	0.93	0.80
Propilamina	0.35	0.32	0.17	0.35	0.32	0.17
i-propilamina	0.23	0.34	0.38	0.23	0.34	0.38
Butilamina	0.28	0.35	0.48	0.28	0.35	0.48
2-butilamina	0.21	0.35	0.22	0.21	0.35	0.22
tert-butilamina	0.31	0.63	0.30	0.31	0.63	0.30
trimetilamina	0.27	0.41	0.27	0.27	0.41	0.27
dimetilamina	0.36	0.58	0.33	0.36	0.58	0.33

**Tabla 4.4.** Efecto de la especie de referencia en la media de las desviaciones absolutas (MADs) de los  $pK_a$ s de ácidos carboxílicos.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Acético	0.78	0.98	0.79	0.57	0.64	0.67
Cloroacético	1.45	1.72	1.50	0.84	1.42	1.70
Fórmico	1.11	1.51	1.13	1.02	1.45	1.17
3-clorobutanoico	0.74	1.25	0.80	0.57	0.66	0.79
Benzoico	0.86	0.89	0.86	0.78	0.64	0.88
4-clorobutanoico	0.74	0.89	0.79	0.58	0.64	0.67
Hexanoico	0.88	1.09	0.98	0.69	0.82	0.72
Propanoico	0.84	1.00	0.89	0.91	0.96	0.82
Pentanoico	0.88	1.06	0.92	0.66	0.73	0.74
Trimetilacético	1.07	1.25	1.31	0.87	1.01	0.81

Aunque es cierto que la elección del compuesto de referencia afecta la precisión de los resultados obtenidos, la cancelación de errores de la reacción isodérmica permite seguir criterios más flexibles en la elección de dicha especie de referencia. En general si el ácido utilizado como referencia,  $BH^m$ , y el ácido de interés,  $AH^q$ , son estructuralmente similares, normalmente el mismo grupo funcional, los resultados obtenidos son también similares independientemente de la referencia elegida (**Tablas 4.3 y 4.4**).

Para evaluar mejor la influencia del ácido de referencia, se ha calculado el  $pK_a$  de algunos de los grupos funcionales usando una molécula estructuralmente diferente (**Tabla 4.5**). Así, se ha calculado el  $pK_a$  de los ácidos benzoicos, de los fenoles y de las piridinas, tomando como referencia el ácido acético, el etanol y la etilamina, respectivamente. En el primero de los casos, se obtienen unas desviaciones medias de 0.5 unidades de  $pK_a$ , del mismo orden de los que se obtienen empleando el ácido benzoico como compuesto de referencia (**Tabla 4.5**). Sin embargo, si se calcula el  $pK_a$  de los fenoles empleando el etanol como referencia, se obtienen unas desviaciones medias que varían entre 4 y 6.5 unidades de  $pK_a$  (**Tabla 4.5**). Esta diferencia de la influencia del compuesto de referencia en los resultados del  $pK_a$  de los fenoles y de los ácidos benzoicos, puede deberse a que el anión fenóxido deslocaliza carga negativa al sistema  $\pi$ , provocando que la interacción electrostática del oxígeno desprotonado con el modelo continuo sea considerablemente diferente a la de los alcóxidos alifáticos. En el caso del ácido benzoico, esta deslocalización de la carga negativa al anillo aromático del ion benzoato es menor y la distribución de carga en el grupo carboxilato es semejante a la de los ácidos carboxílicos alifáticos. Para las piridinas, los errores obtenidos al emplear la etilamina como referencia incrementan hasta 2 unidades de  $pK_a$  respecto a si se emplea

piridina como referencia para el modelo CPCM, pero la diferencia es de sólo 0.6 unidades si se emplea el modelo de solvatación SMD (**Tabla 4.5**). De estos resultados se deduce que el factor estructural no debe ser el único criterio a tener en cuenta para la elección de la referencia y que la cancelación de errores debida a las similitudes en la interacción soluto-continuo no siempre es trivial.

**Tabla 4.5.** Efecto de la especie de referencia en los valores del  $pK_a$ . Media de las desviaciones absolutas (MADs).

	Referencia	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Ácidos benzoicos</b>	benzoico	0.45	0.57	0.41	0.41	0.50	0.34
	acético	0.51	0.47	0.48	1.16	0.65	1.16
<b>Fenoles</b>	fenol	0.90	1.08	1.02	0.70	0.89	0.87
	etanol	4.01	4.31	5.01	5.21	5.59	6.46
<b>Piridinas</b>	piridina	0.57	0.75	0.83	0.62	0.80	0.78
	etilamina	1.86	2.77	1.92	1.22	0.82	1.42

Como se indica en la Metodología, las energías libres en disolución empleadas en la reacción isodérmica no incluyen las contribuciones traslacionales y rotacionales en disolución (**Ecuación 3.16**). Los  $pK_a$ s se han recalculado empleando energías libres que incluyen estas contribuciones, estimadas a partir de las funciones de partición traslacional y rotacional en fase gas. Estas contribuciones modifican los  $pK_a$ s en, únicamente, 0.01-0.1 unidades para los grupos funcionales estudiados. Sólo en el caso del metanol y la metilamina, estas contribuciones tienen un impacto significativo en los  $pK_a$ s calculados (i.e. 0.54 y 0.43 unidades respectivamente). Sin embargo, debe indicarse que este resultado únicamente pone de manifiesto que las energías rotacionales y traslacionales de ácidos y bases conjugados calculados con las funciones de partición en fase gas son muy similares y, por tanto, se cancelan casi completamente en la **Ecuación 3.15**. Además, este resultado no implica que las energías rotacionales y traslacionales reales de ácidos y bases conjugados tengan que ser necesariamente similares.

Los resultados obtenidos se pueden comparar con aquellos publicados anteriormente, en especial con los obtenidos haciendo uso de los ciclos termodinámicos.

El  $pK_a$  de alcoholes alifáticos ha sido calculado en diversas ocasiones [Pliego, 2002], [Namazian, 2003], [Silva, 2000], [Casey, 2006]. Entre los mejores resultados destacan los de Pliego y Riveros [Pliego, 2002], que calcularon el  $pK_a$  del metanol y del

etanol, con una desviación media de 0.5 unidades de  $pK_a$  con el modelo IPCM y 2 ó 3 moléculas de agua de forma explícita, y los de Namazian y Heidary [Namazian, 2003], que obtuvieron el  $pK_a$  del metanol, etanol, propanol y iso-propanol con un error máximo inferior a 0.5 unidades de  $pK_a$  empleando el modelo CPCM para calcular las energías libres de solvatación. Como se observa en la **Tabla 4.2**, con la reacción isodésmica y la combinación CPCM/M05-2X se obtienen unas desviaciones medias (MADs) de 0.62 unidades de  $pK_a$ , del mismo orden de las de los trabajos citados anteriormente.

Previamente a este estudio, el  $pK_a$  de ácidos carboxílicos se ha calculado con frecuencia. Por ejemplo, destaca el estudio de Namazian y Halvani [Namazian, 2006] que calcularon el  $pK_a$  de 66 ácidos carboxílicos con un error medio de 0.5 unidades de  $pK_a$ , usando el modelo de solvatación PCM y el funcional B3LYP. Por su parte, Toth *et al.* [Toth, 2001] calcularon el  $pK_a$  de cinco ácidos carboxílicos usando el modelo de solvatación CPCM y cálculos HF/6-31+G(d)//HF/6-31G(d) con una desviación media de 0.5 unidades de  $pK_a$ . Estos valores son semejantes a los obtenidos en nuestro estudio con el modelo de solvatación SMD (0.57, **Tabla 4.2**).

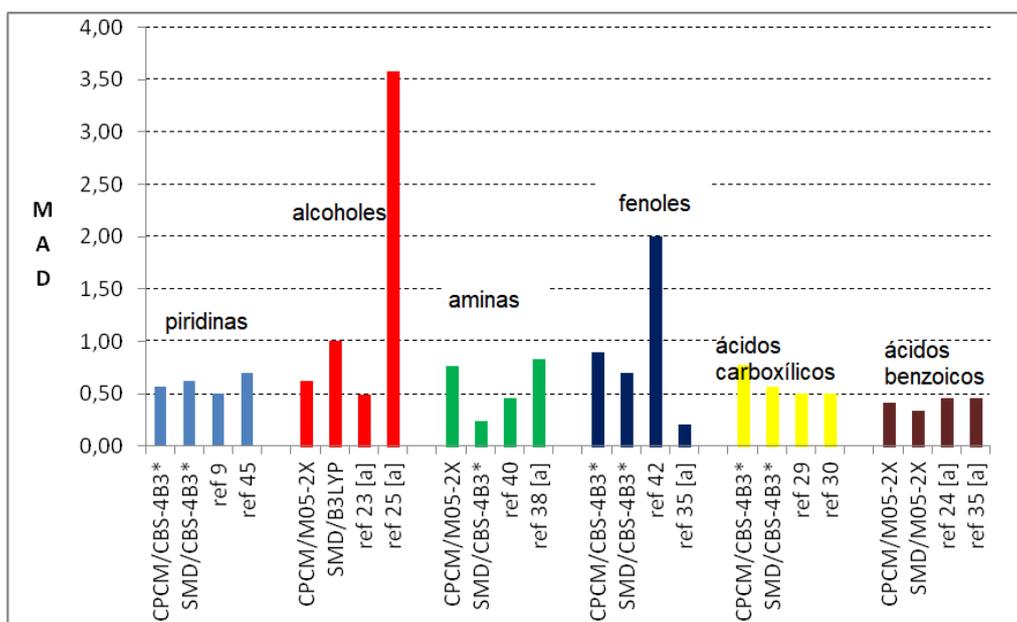
El  $pK_a$  de aminas ha sido calculado con muy buenos resultados. Así, Behjatmanesh-Ardakani *et al.* [Behjatmanesh-Ardakani, 2009] emplearon el modelo de solvatación PCM, cálculos B3LYP/6-31+G\*\*//MP2/6-31+G\* y una molécula de agua de forma explícita, y obtuvieron una desviación media de 0.46 unidades de  $pK_a$ . Nosotros, con el modelo de solvatación CPCM, considerando solo las aminas primarias, obtenemos MADs de orden de 0.3 unidades de  $pK_a$  y si se consideran todas las aminas se obtienen unas MADs de 0.9 unidades de  $pK_a$ . Sin embargo, con el modelo SMD se obtienen MADs de 0.3 unidades para todo el conjunto de aminas (**Tabla 4.2**).

Schmidt am Busch y Knapp [Schmidt am Busch, 2004] calcularon el  $pK_a$  del ácido benzoico, del ácido p-metilbenzoico, del m-metilbenzoico y del p-clorobenzoico con una desviación media de 0.5 unidades de  $pK_a$ . Como se puede apreciar en la **Tabla 4.2**, nuestros resultados son ligeramente mejores que los obtenidos usando ciclos termodinámicos, destacando los cálculos SMD/M05-2X con MAD de 0.34 unidades.

La obtención teórica del  $pK_a$  de fenoles sustituidos ha sido objeto de estudio en diversas ocasiones. Liptak *et al.* [Liptak, 2002] obtuvieron resultados con un error cuadrático medio (RMSD) de entre 0.4 y 3.9 unidades de  $pK_a$ , dependiendo del procedimiento de solvatación elegido. Nuestros resultados se encuentran en el mismo orden de magnitud que los mejores obtenidos con el uso de ciclos termodinámicos, siendo las MADs inferiores en todos los casos a 1 unidad de  $pK_a$  (**Tabla 4.2**).

El  $pK_a$  de piridinas ha sido calculado anteriormente, con una desviación media de una unidad de  $pK_a$ . Casasnovas *et al.* [Casasnovas, 2009] han estudiado la influencia de diferentes ciclos termodinámicos, del empleo de una molécula de agua de forma explícita, y de la importancia de optimizar las geometrías en fase gas o en disolución para el cálculo de energías de solvatación con el modelo CPCM. El  $pK_a$  de piridinas mono y disustituidas se obtiene con un error de 0.5 unidades. Los errores obtenidos en este trabajo son ligeramente mayores que los obtenidos con los ciclos termodinámicos, aunque en cualquier caso inferiores a una unidad de  $pK_a$ .

En la **Figura 4.1** se han representado los menores valores de MAD obtenidos con cada modelo de solvatación, así como las desviaciones medias correspondientes a otros trabajos que usan ciclos termodinámicos en el cálculo teórico del  $pK_a$ . Como se puede observar, nuestros resultados son del mismo orden que los mejores resultados publicados empleando ciclos termodinámicos para las piridinas, ácidos benzoicos y ácidos carboxílicos, independientemente del modelo de solvatación utilizado. En el caso de las aminas, el modelo SMD/CBS-4B3\* conduce a unos resultados significativamente mejores a los obtenidos con el uso de ciclos termodinámicos. De todos modos, en el peor de los casos, la diferencia entre los valores de la desviación media de nuestros resultados y los de los ciclos termodinámicos, no es mayor de 0.6 unidades de  $pK_a$ .



**Figura 4.1.** Comparación entre las medias de las desviaciones absolutas (MADs) de los  $pK_a$ s en disolución acuosa calculados usando diferentes modelos de solvatación (CPM y SMD) con los valores obtenidos empleando ciclos termodinámicos. [Casasnovas, 2009], [Pliego, 2002], [Namazian, 2003], [Silva, 2000], [Toth, 2001], [Namazian, 2006], [Schmidt am Busch, 2004], [Eckert, 2006], [Behjatmanesh-Ardakani, 2009], [Liptak, 2002], [Chen, 2000].

## 4.2. Cálculo de $pK_a$ s de aminoácidos y péptidos usando la reacción isodésmica

En el trabajo anterior se ha mostrado la robustez de la reacción isodésmica en la predicción de  $pK_a$ s de moléculas orgánicas con grupos ácido-base comunes. También se ha mostrado que la reacción isodésmica proporciona  $pK_a$ s con la misma precisión que los mejores valores obtenidos con ciclos termodinámicos.

Como se ha mencionado anteriormente, los ciclos termodinámicos que consideran reacciones en fase gas, como los empleados para el cálculo de  $pK_a$ s, presentan problemas para moléculas inestables en esta fase. Un ejemplo son los aminoácidos, cuyo estado de protonación más estable en disolución es zwitteriónico mientras que el estado de protonación más estable en fase gas no presenta separación de cargas. Como la reacción isodésmica no considera especies en fase gas, puede ser una buena alternativa en estos casos. De hecho, la reacción isodésmica se ha empleado con éxito para calcular los  $pK_a$ s de los grupos carboxilo y amino terminales,  $pK_{a1}$  y  $pK_{a2}$ , de un grupo reducido de aminoácidos [Casasnovas, 2011].

En este trabajo se evaluará en detalle si la reacción isodésmica puede emplearse para predecir con éxito  $pK_a$ s de los grupos carboxilo y amino terminales,  $pK_{a1}$  y  $pK_{a2}$ , y de las cadenas laterales que presenten grupos ácido-base,  $pK_{a3}$ , de aminoácidos y péptidos. También se compararán los resultados con predicciones realizadas con ciclos termodinámicos para aquellas especies que sea posible y con las predicciones realizadas por un método empírico de cálculo de  $pK_a$ s, en particular el implementado en el software ChemOffice [ChemOffice, 1998].

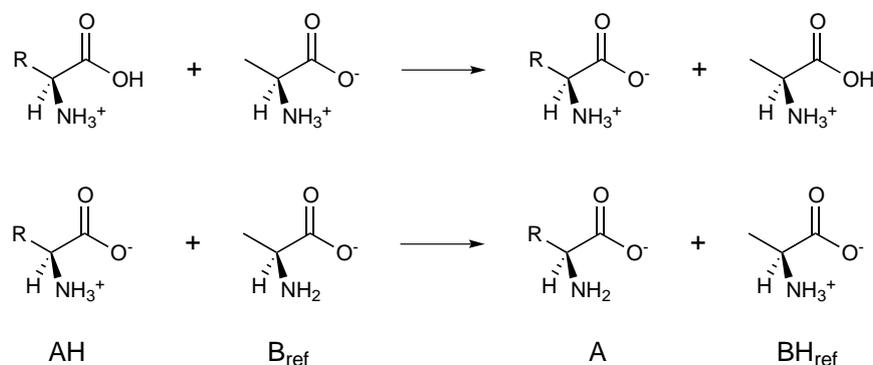
### Cálculo del $pK_a$ del grupo $\alpha$ -carboxílico y del grupo $\alpha$ -amino e influencia del ácido de referencia

Haciendo uso del esquema de la reacción isodésmica (**Esquema 3.4**) y las **Ecuaciones 3.15** y **3.12**, se han calculado los valores de  $pK_a$  del grupo  $\alpha$ -carboxílico,  $pK_{a1}$ , y del grupo  $\alpha$ -amino,  $pK_{a2}$ , de 19 de los aminoácidos proteicos y de una serie de péptidos.

En el caso de los aminoácidos, el ácido de referencia elegido para el  $pK_{a1}$  del grupo  $\alpha$ -carboxílico ha sido la alanina con carga positiva y para el  $pK_{a2}$  del grupo  $\alpha$ -amino la alanina zwitteriónica (**Figura 4.2**). Así se obtiene unas medias de las desviaciones absolutas (MADs) del orden 0.5 unidades de  $pK_a$  (**Tablas 4.6** y **4.7**).

Las **Tablas 4.6** y **4.7** muestran los errores absolutos, la MAD y la desviación estándar de los  $pK_{a1}$  y  $pK_{a2}$  con respecto a los valores experimentales. En términos de errores absolutos, la mayoría de las predicciones de  $pK_{a1}$  y  $pK_{a2}$  tienen errores menores

que 1.0 unidad para todos los funcionales DFT y el Hamiltoniano PM6. Los valores de las MAD son de aproximadamente 0.5 unidades con la excepción de las predicciones de  $pK_{a2}$  con PM6, que muestra un valor MAD más alto de 0.9 unidades.



**Figura 4.2.** Reacción isodésmica empleada para calcular el  $pK_a$  del grupo  $\alpha$ -carboxílico ( $pK_{a1}$ ). Reacción isodésmica empleada para calcular el  $pK_a$  del grupo  $\alpha$ -amino ( $pK_{a2}$ ).

El cálculo de  $pK_{a1}$  y  $pK_{a2}$  se repitió empleando sistemáticamente los otros aminoácidos como ácido de referencia. En este caso, los valores de las MADs obtenidos usando el funcional M052x se sitúan entre 0.5 y 1.0 unidades de  $pK_a$  (**Tabla 4.8**). Es importante destacar que la MAD de los  $pK_{a1}$  cuando se usa histidina como especie de referencia es significativamente mayor que los de los otros aminoácidos (2.09 unidades de  $pK_a$ ).

**Tabla 4.6.** Valores de los errores absolutos de los  $pK_{a1}$  calculados de aminoácidos empleando alanina como ácido de referencia.

	$pK_{a1}$ (exptl. <sup>a</sup> )	$\Delta pK_{a1}$ (M052X)	$\Delta pK_{a1}$ (M062X)	$\Delta pK_{a1}$ (B3LYP)	$\Delta pK_{a1}$ (PM6)	$\Delta pK_{a1}$ (ChemOffice)
Alanina (Ref)	2.34	-	-	-	-	0.08
Glicina	2.34	0.42	0.18	0.42	0.09	0.03
Valina	2.32	0.18	0.06	0.32	1.06	0.14
Iso-Leucina	2.35	0.28	0.11	0.13	0.36	0.06
Leucina	2.33	0.02	0.02	0.35	0.29	0.04
Metionina	2.28	0.37	0.13	0.04	0.15	0.04
Prolina	1.99	0.37	0.10	0.36	0.43	0.53
Fenilalanina	1.83	0.08	0.03	0.33	0.62	0.38
Triptófano	2.38	0.26	0.28	0.32	0.11	0.29
Serina	2.21	0.67	0.60	0.31	1.37	0.05
Treonina	2.09	0.57	0.49	0.87	1.35	0.09
Asparagina	2.01	0.76	0.97	0.16	1.70	0.09
Glutamina	2.17	0.44	0.28	0.19	0.27	0.02
Tirosina	2.18	0.06	0.08	0.12	0.09	0.16
Lisina	2.18	0.36	0.44	0.63	0.49	0.11
Arginina	2.17	0.92	1.06	0.46	0.04	0.02
Cisteína	1.96	0.28	0.70	0.21	0.02	0.14
Histidina	1.82	2.34	2.78	2.48	1.89	0.38
Ac. Aspártico	1.89	1.35	1.31	0.85	1.09	0.03
Ac. Glutámico	2.19	0.75	0.71	0.28	0.34	0.04
MAD		0.56	0.54	0.46	0.62	0.14
SD		0.55	0.66	0.54	0.60	0.14

<sup>a</sup>[Dean, 1999]

**Tabla 4.7.** Valores de los errores absolutos de los  $pK_{a2}$  calculados de aminoácidos empleando alanina como ácido de referencia.

	$pK_{a2}$ (exptl. <sup>a</sup> )	$\Delta pK_{a2}$ (M052X)	$\Delta pK_{a2}$ (M062X)	$\Delta pK_{a2}$ (B3LYP)	$\Delta pK_{a2}$ (PM6)	$\Delta pK_{a2}$ (ChemOffice)
Alanina (Ref)	9.69	-	-	-	-	0.5
Glicina	9.6	1.13	0.10	0.61	0.17	0.45
Valina	9.61	0.81	0.92	0.48	0.12	0.6
Iso-Leucina	9.68	0.64	0.69	0.56	0.63	0.54
Leucina	9.6	0.14	0.04	0.26	0.59	0.62
Metionina	9.21	0.03	0.87	0.00	1.20	0.85
Prolina	10.96	2.13	1.79	1.55	2.36	0.78
Fenilalanina	9.12	0.06	0.14	0.21	0.00	0.81
Triptófano	9.39	0.22	0.10	0.20	0.56	0.62
Serina	9.15	0.71	0.75	1.36	1.24	0.22
Treonina	9.1	0.90	0.99	0.89	0.84	0.35
Asparagina	8.8	0.22	0.50	0.29	0.40	1.35
Glutamina	9.13	0.30	0.41	0.19	1.40	0.38
Tirosina	9.11	0.06	0.18	0.48	0.84	1.56
Lisina	8.94	0.32	0.71	0.08	0.52	0.84
Arginina	9.04	0.36	0.55	0.05	0.84	1.14
Cisteína	10.29	0.69	0.82	0.65	1.28	0.44
Histidina	9.16	0.49	0.39	0.15	0.83	0.82
Ac. Aspártico	9.6	0.62	0.85	0.82	1.06	0.55
Ac. Glutámico	9.67	0.31	0.77	0.01	1.13	0.52
MAD		0.51	0.59	0.45	0.86	0.70
SD		0.44	0.37	0.40	0.61	0.34

<sup>a</sup>[Dean, 1999]

Así en el cálculo del  $pK_{a1}$  y del  $pK_{a2}$  de los aminoácidos, el empleo de la alanina como ácido de referencia proporciona una media de los errores absolutos (MAD) del orden 0.5 unidades de  $pK_a$  (Tablas 4.6 y 4.7). Con los otros aminoácidos como ácido de referencia, los valores de las MADs obtenidos usando el funcional M052x se sitúan entre 0.5 y 1.0 unidades de  $pK_a$  (Tabla 4.8). Es importante destacar que la MAD de los  $pK_{a1}$  cuando se usa histidina como especie de referencia es significativamente mayor que los de los otros aminoácidos (2.09 unidades de  $pK_a$ ).

**Tabla 4.8.** Media de los errores absolutos (MAD) de los  $pK_{a1}$  y  $pK_{a2}$  calculados de amino ácidos en función del ácido de referencia.

<b>Referencia</b>	<b>MAD <math>pK_{a1}</math></b>	<b>MAD <math>pK_{a2}</math></b>
Alanina	0.56	0.51
Glicina	0.84	1.40
Valina	0.66	0.76
Iso-Leucina	0.73	0.64
Leucina	0.58	0.47
Metionina	0.53	0.50
Prolina	0.80	1.68
Fenilalanina	0.55	0.52
Triptófano	0.52	1.27
Serina	0.62	1.52
Treonina	0.58	1.66
Asparagina	0.67	0.60
Glutamina	0.54	0.48
Tirosina	0.55	0.48
Lisina	0.52	0.48
Arginina	0.79	0.70
Cisteína	0.94	0.47
Histidina	2.09	0.56
Ac. Aspártico	1.16	0.62
Ac. Glutámico	0.66	0.48
<b>Media</b>	<b>0.74</b>	<b>0.79</b>

Los  $pK_{a1}$  y  $pK_{a2}$  de los péptidos considerados se calcularon considerando como ácido de referencia otro péptido, la glicilglicina (**Tablas 4.9 y 4.10**). En este caso se obtienen 46 valores de  $pK_{a1}$  con errores inferiores a 0.5 unidades y 29 con errores entre 0.5 y 1.0 unidades (**Tabla 4.9**). Para el  $pK_{a2}$ , 81 de los valores obtenidos tienen errores inferiores a 1.0 unidades (**Tabla 4.10**).

En el caso de la glicilvalina y la glicilfenilalanina existen valores calculados en la literatura. Se puede ver que nuestros cálculos muestran errores similares a los reportados Por Kiani *et al.* [Kiani, 2010].

También se ha utilizado el programa de predicción de  $pK_a$ s de ChemOffice para calcular el  $pK_{a1}$  y  $pK_{a2}$  de los mismos aminoácidos y péptidos. Los valores de la MAD y SD obtenidos con ChemOffice son 3-4 veces menores que los resultados de la reacción isodésmica  $pK_{a1}$  (**Tablas 4.6 y 4.9**). Para el  $pK_{a2}$ , la MAD y la SD obtenidos con ChemOffice son similares para los aminoácidos (**Tabla 4.7**), pero mejor para los péptidos (**Tabla 4.10**).

**Tabla 4.9.** Errores absolutos de los  $pK_{a1}$  calculados de péptidos empleando glicilglicina como ácido de referencia.

	$pK_{a1}$ (exptl.)	$\Delta pK_{a1}$ (M052X)	$\Delta pK_{a1}$ (M062X)	$\Delta pK_{a1}$ (B3LYP)	$\Delta pK_{a1}$ (PM6)	$\Delta pK_{a1}$ (ChemOffice)
GlyGly	3.13	-	-	-	-	0.17
GlyVal	3.18	0.13	0.40	0.29	0.20	0.12
GlyPhe	3.23	0.33	0.47	0.10	1.70	0.39
GlyAla	3.15	0.83	0.72	0.64	1.45	0.10
AlaGly	3.16 <sup>b</sup>	0.10	0.43	0.34	0.93	0.22
AlaAla	3.32	2.08	1.95	1.01	0.35	0.29
AsnGly	2.90 <sup>b</sup>	0.97	0.68	0.18	2.34	0.24
ValGly	3.23	1.37	1.56	1.22	1.12	0.24
SerGly	3.10 <sup>b</sup>	0.31	0.63	0.26	0.80	0.07
SerLeu	3.08	0.87	1.94	0.45	3.32	0.01
AlaHis	2.64 <sup>c</sup>	2.45	2.23	2.13	0.44	0.16
GlyTyr	2.93	2.39	2.39	0.25	1.31	0.28
HisGly	2.40	1.04	0.57	0.21	0.36	0.42
GlyAsp	2.81	0.78	0.76	0.55	1.54	0.27
CysAsn	2.97	0.60	0.63	1.68	0.83	0.37
PheArg	2.66	0.37	0.14	0.53	1.01	0.06
LysAla	3.22 <sup>b</sup>	2.26	1.87	1.89	0.91	0.20
LeuTyr	3.46	1.90	1.37	0.73	1.53	0.85
TyrGly	2.98 <sup>d</sup>	0.28	0.97	0.33	0.52	0.11
LysGlu	2.93	2.93	1.29	1.86	2.33	1.10
TyrArg	2.65	0.15	0.12	0.54	0.66	0.07
HisLys	2.50 <sup>e</sup>	2.13	2.12	1.08	1.21	0.28
GlyLys	2.96 <sup>e</sup>	0.01	0.07	0.31	0.47	0.04
AspGly	2.10 <sup>b</sup>	0.08	0.01	0.20	0.35	0.42
AlaGlyGly	3.19 <sup>b</sup>	0.12	0.18	0.07	4.26	0.42
GlyAlaAla	3.38	1.61	2.44	0.64	1.97	0.30
GlySerGly	3.32	0.41	0.35	0.96	0.21	0.36
GlyGlyGly	3.23	0.09	0.20	0.71	1.69	0.39
CysGlyGly	3.13 <sup>b</sup>	0.14	0.05	0.80	2.79	0.47
AlaLysAla	3.15 <sup>b</sup>	0.24	0.69	1.32	1.47	0.52
PheAlaArg	2.60 <sup>b</sup>	0.27	0.02	0.81	1.01	0.83
<b>MAD</b>		0.99	0.98	0.78	1.40	0.32
<b>SD</b>		0.85	0.78	0.53	0.96	0.25

<sup>a</sup>[Dean, 1999], <sup>b</sup>[Schmidt, 2001], <sup>c</sup>[Jenks, 1970], <sup>d</sup>[Tominaga, 1995], <sup>e</sup>[Remeli, 2000], <sup>f</sup>[Rabenstein, 1977]

Estos datos muestran que pueden obtenerse buenas predicciones de  $pK_a$  con métodos bien parametrizados, sin embargo, la reacción isodésmica no requiere parametrización explícita, aunque puede haber una parametrización implícita debida al funcional de correlación-intercambio, Hamiltoniano semiempírico y/o modelo continuo de disolvente. De hecho, en la reacción isodésmica, la parametrización tiene lugar *in situ* por la inclusión de la especie de referencia. Por lo tanto, cualquier  $pK_a$  de cualquier molécula se puede calcular, siempre que se elija una especie de referencia adecuada.

**Tabla 4.10.** Errores absolutos de los  $pK_{a2}$  calculados de péptidos empleando glicilglicina como ácido de referencia.

	$pK_{a2}$ (exptl.)	$\Delta pK_{a2}$ (M052X)	$\Delta pK_{a2}$ (M062X)	$\Delta pK_{a2}$ (B3LYP)	$\Delta pK_{a2}$ (PM6)	$\Delta pK_{a2}$ (ChemOffice)
GlyGly	8.25	-	-	-	-	0.71
GlyVal	8.18	0.34	0.22	0.59	1.68	0.61
GlyPhe	8.11 <sup>b</sup>	1.64	1.52	1.54	0.77	0.58
GlyAla	8.33	0.03	0.37	0.08	2.14	0.77
AlaGly	8.24 <sup>b</sup>	0.27	0.16	0.53	2.17	0.57
AlaAla	8.13	1.24	0.99	0.92	1.47	0.43
AsnGly	7.25 <sup>b</sup>	0.10	0.65	0.16	2.15	0.98
ValGly	8.00	0.91	1.17	1.61	0.49	0.31
SerGly	7.33 <sup>b</sup>	0.49	0.76	1.04	0.64	0.47
SerLeu	7.45	0.34	0.51	0.61	2.59	0.55
AlaHis	9.40 <sup>c</sup>	0.16	0.21	0.45	4.26	0.37
GlyTyr	8.45	0.72	0.11	0.17	0.18	0.91
HisGly	7.82	0.51	0.80	2.56	3.74	0.36
GlyAsp	8.60	0.68	0.64	0.39	1.34	1.04
CysAsn	8.47	2.43	3.02	3.31	0.66	1.25
PheArg	7.57	0.14	0.15	0.25	2.20	0.16
LysAla	8.47 <sup>b</sup>	0.41	0.69	0.47	2.70	1.18
LeuTyr	7.84	1.05	0.73	1.32	0.03	0.14
TyrGly	8.00 <sup>d</sup>	0.12	0.18	0.32	1.44	0.56
LysGlu	7.75	0.88	0.47	0.08	0.82	0.64
TyrArg	7.39	0.54	0.77	0.26	0.89	0.05
HisLys	7.41 <sup>e</sup>	0.23	0.45	0.70	2.01	0.07
AspHis	7.98 <sup>b</sup>	2.49	2.43	1.23	3.60	0.64
GlyHis	8.20	2.49	2.43	1.23	3.60	0.32
GlyLys	8.01 <sup>e</sup>	1.35	1.56	1.29	0.56	0.46
AspGly	9.07 <sup>b</sup>	1.69	1.84	1.93	4.28	1.43
AlaGlyGly	8.15 <sup>b</sup>	0.36	0.47	0.55	3.29	0.83
GlyAlaAla	8.10	0.33	1.03	1.29	1.24	0.89
GlySerGly	7.99	0.04	0.28	0.46	2.49	0.87
GlyGlyGly	8.09	0.29	0.18	0.13	3.31	0.9
CysGlyGly	6.95 <sup>b</sup>	4.43	4.71	4.58	0.46	0.04
AlaLysAla	7.65 <sup>b</sup>	0.50	0.62	0.49	1.53	0.32
PheAlaArg	7.54 <sup>b</sup>	0.12	0.49	0.20	2.25	0.46
GlyHisLys	8.06 <sup>f</sup>	0.84	0.09	0.23	0.96	0.85
MAD		0.91	0.98	0.92	1.71	0.61
SD		1.01	1.10	1.00	1.20	0.35

<sup>a</sup>[Dean, 1999], <sup>b</sup>[Schmidt, 2001], <sup>c</sup>[Jenks, 1970], <sup>d</sup>[Tominaga, 1995], <sup>e</sup>[Remeli, 2000], <sup>f</sup>[Rabenstein, 1977]

Como se ha mostrado en el caso de moléculas orgánicas, la elección del ácido de referencia es importante en la determinación de los  $pK_{as}$ , aunque la cancelación de errores de la reacción isodésmica permite considerar diversas especies de referencia, tanto en su estructura como en su carga global.

Cualquiera de los aminoácidos salvo la histidina es una buena referencia para los  $pK_{a1}$  y  $pK_{a2}$  de los otros aminoácidos (**Tabla 4.8**). Cuando se emplea la glicilglicina como referencia para los aminoácidos, la desviación absoluta media (MAD) de  $pK_{a1}$  aumenta levemente en 0.3 unidades mientras que la MAD del  $pK_{a2}$  aumenta notablemente en 1.61 unidades (**Tabla 4.11**). Por el contrario, cuando se emplea alanina como referencia para los péptidos, la MAD de  $pK_{a1}$  se mantiene y la de  $pK_{a2}$  disminuye 0.47 unidades (**Tabla 4.11**).

Para estudiar con más detalle el efecto del ácido de referencia en la precisión de los  $pK_a$ s, se emplearon el ácido acético y la etilamina respectivamente para el  $pK_{a1}$  y el  $pK_{a2}$  tanto de aminoácidos como de péptidos. En este caso, se obtienen unas MADs mucho más elevadas en el caso de  $pK_{a1}$  de aminoácidos y péptidos (**Tabla 4.11**). El valor de la MAD del  $pK_{a1}$  de los péptidos se duplica, pero en el caso de los aminoácidos se multiplica por 5. Este hecho se debe a que en un péptido el grupo  $\alpha$ -carboxílico y el  $\alpha$ -amino se encuentran más separados entre sí por lo que el ácido acético describe mejor las interacciones electrostáticas con el disolvente de un péptido que de un aminoácido. Por el mismo motivo, si se emplea la etilamina, la MAD del  $pK_{a2}$  de los aminoácidos aumenta 0.97 unidades para los aminoácidos, pero para los péptidos se obtiene una MAD parecida, aunque ligeramente inferior (i.e. 0.16 unidades) (**Tabla 4.11**).

**Tabla 4.11.** Media de las desviaciones absolutas (MAD) de los  $pK_{a1}$  y  $pK_{a2}$  de aminoácidos y péptidos según el ácido de referencia empleado.

Referencia	$pK_{a1}$		$pK_{a2}$		
	Aminoácidos	Péptidos	Referencia	Aminoácidos	Péptidos
Alanina	0.57	1.02	Alanina	0.49	1.44
GlyGly	0.87	1.01	GlyGly	2.10	0.97
Ácido Acético	2.63	2.06	Etilamina	1.46	0.81

Para explicar estos resultados se debe considerar que una buena especie de referencia debe tener interacciones soluto-disolvente similares al ácido estudiado. En realidad, y volviendo al esquema de la reacción isodésmica, la interacción de  $BH^m$  con el disolvente continuo debe ser lo más similar posible a la de  $AH^q$ , e igualmente para  $B^{m-1}$  y  $A^{q-1}$ . Por esta razón, al emplear un aminoácido como especie de referencia se obtienen MADs de 0.74 y 0.79 unidades de media para  $pK_{a1}$  y  $pK_{a2}$  de otros aminoácidos (**Tabla 4.8**). Por otro lado, si la especie de referencia contiene el mismo grupo funcional, pero muestra una distribución de la carga significativamente diferente,

los errores aumentan. Este es el caso cuando se utilizan ácido acético y etilamina como especie de referencia para el cálculo de  $pK_{a1}$  y  $pK_{a2}$  de aminoácidos (**Tabla 4.11**).

Cabe señalar que nos referimos a la distribución de la carga de la especie de referencia porque se ha demostrado que la interacción electrostática con el modelo continuo es la principal fuente de error en el cálculo de energías de solvatación [Cramer, 2008], [Casasnovas, 2014]. Sin embargo, el entorno químico también debe tenerse en cuenta. De hecho, los grandes errores obtenidos para el  $pK_{a1}$  de la histidina o el ácido aspártico con la alanina como referencia (**Tabla 4.6**) pueden ser atribuidos a ese factor. En ambos casos, la cadena lateral del aminoácido está lo suficientemente cerca como para interactuar con el  $\alpha$ -carboxilo y/o los grupos  $\alpha$ -amino. Sin embargo, estas interacciones no están presentes en la alanina y además se modifican tras la desprotonación del grupo  $\alpha$ -carboxilo, por eso la cancelación de errores con esta referencia no es buena.

También se ha evaluado la influencia de la carga de los aminoácidos y de los péptidos, así como del número de aminoácidos que forman los péptidos, es decir, de la separación entre los grupos  $\alpha$ -carboxilo y  $\alpha$ -amino, en los valores del  $pK_{a1}$  y  $pK_{a2}$  obtenidos (**Tabla 4.12**). Atendiendo al número de aminoácidos de los péptidos, los resultados son mejores para el  $pK_{a1}$  de los tripéptidos y al contrario son mejores para el  $pK_{a2}$  de los dipéptidos.

**Tabla 4.12.** Media de las desviaciones absolutas (MAD) de los  $pK_{a1}$  y  $pK_{a2}$  de aminoácidos y péptidos según su longitud y carga total,  $q_t$ .

Referencia	Aminoácidos	Dipéptidos	Tripéptidos			
$pK_{a1}$				$q_t = 1$	$q_t = 2$	
Alanina	0.57	1.10	0.72	0.76	1.10	
GlyGly	0.87	1.19	0.42	0.93	1.18	
$pK_{a2}$				$q_t = 1$	$q_t = 0$	$q_t = -1$
Alanina	0.49	1.35	1.76	1.26	1.10	0.90
GlyGly	2.10	0.98	0.91	0.50	0.82	2.23

Respecto a la carga tanto de los aminoácidos como de los péptidos, exceptuando el cálculo del  $pK_{a2}$  de los péptidos con carga negativa y empleando el péptido glicilglicina como referencia, no se observan diferencias en el  $pK_{a1}$  ni en el  $pK_{a2}$ , sea cual sea la referencia elegida, ya que la mayoría de las MADs son del orden de 1 unidad de  $pK_a$ . El elevado valor mencionado se debe a que sólo disponemos de cinco péptidos con carga negativa y en alguno de ellos se obtiene un error absoluto elevado.

La única tendencia clara que se observa es que la alanina es el mejor ácido de referencia para los aminoácidos y la glicilglicina para los péptidos. No se observan tendencias claras respecto a la longitud de la cadena o la carga global. Esto refuerza la hipótesis de que es más importante el entorno electrostático local del grupo ácido considerado que la carga total de la molécula en la elección del ácido de referencia.

### Cálculo del $pK_a$ de las cadenas laterales de péptidos ( $pK_{a3}$ )

Se ha calculado el valor del  $pK_{a3}$  de grupo  $\epsilon$ -amino de la lisina, guanidinio de la arginina, sulfhidrilo de la cisteína, fenol de la tirosina, imidazol de la histidina y carboxílico de los ácidos glutámico y aspártico de los péptidos considerados. El ácido de referencia empleado ha sido el grupo ácido de la cadena lateral de la lisina, de la arginina, de la cisteína, de la tirosina, de la histidina y del ácido glutámico respectivamente. Los errores absolutos, MADs y SDs entre los valores calculados y los experimentales se muestran en la **Tabla 4.13**.

**Tabla 4.13.** Errores absolutos de los  $pK_{a3}$  calculados de péptidos.

	$pK_{a3}$ (exptl.)	$\Delta pK_{a3}$ (M052X)	$\Delta pK_{a3}$ (M062X)	$\Delta pK_{a3}$ (B3LYP)	$\Delta pK_{a3}$ (PM6)	$\Delta pK_{a3}$ (ChemOffice)
<b>Lisina</b>						
Lys (ref.)	10.53 <sup>a</sup>	-	-	-	-	0.24
LysGlu	10.50 <sup>a</sup>	0.38	0.58	0.60	2.13	0.36
GlyLys	10.50 <sup>b</sup>	0.24	0.27	0.06	0.00	0.35
HisLys	10.49 <sup>b</sup>	0.02	0.28	0.58	0.50	0.34
LysAla	10.70 <sup>a</sup>	0.14	0.05	0.95	2.14	0.55
AlaLys	10.70 <sup>b</sup>	0.08	0.39	0.03	1.00	0.36
AlaLysAla	10.30 <sup>c</sup>	0.29	0.39	0.48	2.18	0.30
LysD-Ala	10.63 <sup>b</sup>	0.65	0.78	0.59	1.78	0.48
GlyHisLys	10.71 <sup>d</sup>	1.68	1.10	1.30	2.71	0.57
GlyGlyLysAla	11.10 <sup>e</sup>	2.29	1.08	0.38	3.54	1.10
<b>MAD</b>		<b>0.64</b>	<b>0.55</b>	<b>0.55</b>	<b>1.78</b>	<b>0.47</b>
<b>SD</b>		<b>0.80</b>	<b>0.37</b>	<b>0.40</b>	<b>1.11</b>	<b>0.25</b>
<b>Arginina</b>						
Arg (ref.)	12.47 <sup>a</sup>	-	-	-	-	3.88
PheArg	12.40 <sup>a</sup>	1.65	2.20	0.71	1.42	3.76
TyrArg	11.62 <sup>a</sup>	0.37	0.31	2.29	0.59	2.98
PheAlaArg	12.43 <sup>c</sup>	0.79	1.66	0.72	0.73	3.79
<b>MAD</b>		<b>0.94</b>	<b>1.39</b>	<b>1.24</b>	<b>0.91</b>	<b>3.60</b>
<b>SD</b>		<b>0.65</b>	<b>0.97</b>	<b>0.91</b>	<b>0.44</b>	<b>0.42</b>
<b>Histidina</b>						
His (ref.)	6.00 <sup>a</sup>	-	-	-	-	-
AlaHis	6.72 <sup>f</sup>	1.10	1.81	2.31	0.78	-
HisGly	5.80 <sup>a</sup>	1.38	2.51	4.58	2.84	-

HisLys	5.91 <sup>b</sup>	1.93	2.73	2.18	2.45	-
AspHis	6.82 <sup>a</sup>	0.52	0.98	1.01	0.02	-
GlyHis	6.77 <sup>a</sup>	0.48	0.99	0.64	0.70	-
GlyHisLys	6.60 <sup>d</sup>	2.00	2.23	2.54	0.89	-
GlyHisGly	6.62 <sup>d</sup>	1.93	2.24	2.69	0.13	-
GlyGlyHisAla	7.00 <sup>e</sup>	1.51	1.01	2.71	0.38	-
TyrHisOMe	6.41 <sup>f</sup>	2.17	2.59	3.05	1.95	-
GluHisOMe	6.44 <sup>f</sup>	2.34	2.94	2.76	1.81	-
<b>MAD</b>		<b>1.53</b>	<b>2.00</b>	<b>2.45</b>	<b>1.19</b>	-
<b>SD</b>		<b>0.66</b>	<b>0.76</b>	<b>1.08</b>	<b>1.00</b>	-
Cisteína						
Cys (ref.)	8.18 <sup>a</sup>	-	-	-	-	-
CysAsn	7.09 <sup>a</sup>	0.24	0.20	0.09	1.49	-
CysGlyGly	6.36 <sup>c</sup>	0.88	1.24	1.06	0.93	-
<b>MAD</b>		<b>0.56</b>	<b>0.72</b>	<b>0.57</b>	<b>1.21</b>	-
<b>SD</b>		<b>0.45</b>	<b>0.73</b>	<b>0.69</b>	<b>0.39</b>	-
Tirosina						
Tyr (ref.)	10.60 <sup>a</sup>	-	-	-	-	1.25
GlyTyr	10.49 <sup>a</sup>	2.45	1.97	0.23	0.63	1.08
TyrArg	9.36 <sup>a</sup>	1.10	1.32	0.81	0.82	0.17
LeuTyr	10.09 <sup>a</sup>	0.86	1.02	0.71	1.82	0.69
TyrGly	10.51 <sup>g</sup>	0.18	0.01	0.12	1.43	0.97
GlyGlyTyrAla	10.30 <sup>e</sup>	0.80	0.22	0.33	2.03	0.88
D-LeuTyr	10.35 <sup>a</sup>	0.62	0.30	1.00	1.01	0.95
TyrHisOMe	9.69 <sup>f</sup>	0.20	0.17	0.74	0.13	0.16
<b>MAD</b>		<b>0.89</b>	<b>0.72</b>	<b>0.56</b>	<b>1.13</b>	<b>0.77</b>
<b>SD</b>		<b>0.77</b>	<b>0.74</b>	<b>0.34</b>	<b>0.67</b>	<b>0.41</b>
Glu/Asp						
Glu (ref.)	4.25 <sup>a</sup>	-	-	-	-	1.20
Asp	3.65 <sup>a</sup>	1.07	1.76	0.42	0.60	1.79
LysGlu	4.47 <sup>a</sup>	0.14	0.74	0.16	0.42	0.44
GlyGlyGluAla	4.30 <sup>e</sup>	0.22	0.56	0.27	0.31	0.06
GlyAsp	4.45 <sup>a</sup>	1.58	2.15	1.33	0.14	0.80
GlyGlyAspAla	3.90 <sup>e</sup>	0.53	1.70	1.83	0.87	0.24
GluHisOMe	3.79 <sup>f</sup>	0.04	0.18	1.03	0.93	0.66
AspGly	4.53 <sup>c</sup>	0.84	0.27	0.93	0.21	2.76
<b>MAD</b>		<b>0.63</b>	<b>1.05</b>	<b>0.85</b>	<b>0.50</b>	<b>0.99</b>
<b>SD</b>		<b>0.56</b>	<b>0.80</b>	<b>0.61</b>	<b>0.31</b>	<b>0.90</b>
Todos						
<b>MAD</b>		<b>0.94</b>	<b>1.13</b>	<b>1.16</b>	<b>1.17</b>	<b>1.11</b>
<b>SD</b>		<b>0.75</b>	<b>0.87</b>	<b>1.05</b>	<b>0.89</b>	<b>1.14</b>

<sup>a</sup>[Dean, 1999], <sup>b</sup>[Remelli, 2000], <sup>c</sup>[Jenks, 1970], <sup>d</sup>[Rabenstein, 1977], <sup>e</sup>[Arnold,2002], <sup>f</sup>[Vistoli, 2012], <sup>g</sup>[Tominaga, 1995]

Con cálculos DFT se obtienen unos valores de MADs inferiores a 1.0 unidad de  $pK_a$  para todos los residuos, excepto las histidinas, cuyas MADs son de 1.53-2.45 unidades. Con el Hamiltoniano PM6 se obtienen MADs mayores para lisinas, cisteínas y tirosinas, pero menor para histidinas. Los errores absolutos son inferiores a 1.0 unidad de

$pK_a$  para 47 resultados independientemente del tipo de residuo. Los 15 casos restantes, en los que los errores son mayores de 1.0 unidad de  $pK_a$ , corresponden en su mayoría a las histidinas y luego a unas pocas lisinas, argininas, tirosinas y ácidos aspárticos. Por lo tanto, con la excepción de las histidinas, con la reacción isodésmica se obtienen valores del  $pK_{a3}$  muy satisfactorios (errores inferiores a 1.0 unidad de  $pK_a$ ).

El software ChemOffice también se empleó para calcular los  $pK_{a3}$  (**Tabla 4.13**). Como se observa, este software no permite el cálculo de  $pK_a$ s de grupos tiol e imidazol. Sin embargo, para el resto de grupos funcionales, los errores obtenidos son similares a los obtenidos empleando la reacción isodésmica (i.e. lisinas, tirosinas, ácidos aspárticos y glutámicos), pero peores para argininas (**Tabla 4.13**).

Para evaluar la influencia del ácido de referencia, se ha calculado el  $pK_{a3}$  usando un compuesto de referencia que no es aminoácido ni péptido para cada uno de los grupos estudiados: la etilamina, etilguanidinio, el 4-metilimidazol, el etanotiol, el fenol y el ácido acético respectivamente para la lisina, arginina, histidina, cisteína, tirosina y ácido glutámico/aspártico. (**Tabla 4.14**).

**Tabla 4.14.** Media de las desviaciones absolutas (MAD) y desviaciones estándar (SD) de los  $pK_{a3}$  s calculados empleando como referencia moléculas orgánicas simples <sup>a</sup>.

	$\Delta pK_{a3}$ (M052X)	$\Delta pK_{a3}$ (M062X)	$\Delta pK_{a3}$ (B3LYP)	$\Delta pK_{a3}$ (PM6)
Lisina				
MAD	0.59	0.48	0.48	1.54
SD	0.76	0.49	0.35	1.13
Arginina				
MAD	1.52	0.99	1.67	0.69
SD	0.89	0.84	1.28	0.54
Histidina				
MAD	0.63	0.77	0.92	1.90
SD	0.43	0.52	0.81	1.04
Cisteína				
MAD	2.34	2.46	1.94	2.58
SD	0.46	0.66	0.59	0.75
Tirosina				
MAD	1.46	0.71	0.77	2.17
SD	0.61	0.71	0.47	1.25
Glu/Asp				
MAD	0.70	1.22	0.89	2.24
SD	0.42	0.87	0.52	0.53
<b>Todos</b>				
<b>MAD</b>	<b>0.98</b>	<b>0.91</b>	<b>0.93</b>	<b>1.87</b>
<b>SD</b>	<b>0.77</b>	<b>0.79</b>	<b>0.75</b>	<b>1.05</b>

<sup>a</sup> Moléculas orgánicas usadas como especies de referencia (etilamina, etilguanidinio, 4-metilimidazol, etanotiol, fenol y ácido acético).

En general no se observan diferencias importantes en el valor de las MADs al emplear otra referencia con igual o diferente carga y/o estructura siendo como mucho de 1 unidad de  $pK_a$ . En el caso de los  $pK_{a3}$  de los péptidos que contienen histidinas, se obtienen mejores resultados empleando el 4-metilimidazol como compuesto de referencia que empleando el propio aminoácido histidina (**Tablas 4.13 y 4.14**). Esto se puede explicar con un razonamiento similar al empleado en la sección anterior, según el cual los errores de los  $pK_{a1}$  y  $pK_{a2}$  de la histidina al emplear la alanina como ácido de referencia se deben a que: 1) las interacciones intramoleculares de los grupos  $\alpha$ -carboxílico y  $\alpha$ -amino con el imidazol de la cadena lateral no están presentes en la referencia, y 2) dichas interacciones intramoleculares no se mantienen constantes sino que cambian cuando los grupos implicados se protonan/desprotonan. En el caso de los residuos histidina en péptidos, los grupos imidazol de las cadenas laterales están más alejados de los grupos carboxilo y amino terminales que en el aminoácido histidina y presentan menos interacciones intramoleculares. También por ese motivo, las interacciones de los imidazoles de las cadenas laterales con el disolvente son más similares a las presentes en el 4-metilimidazol, lo que lo convierte en un mejor ácido de referencia.

### **Ciclos termodinámicos para el cálculo de $pK_{a3}$ de péptidos**

El uso de ciclos termodinámicos implica inconvenientes para especies químicas que son inestables en la fase gaseosa. De hecho, Casasnovas *et al.* [Casasnovas 2011] han mostrado que los ciclos termodinámicos no son una aproximación útil para el cálculo de  $pK_{a1}$  y  $pK_{a2}$  porque las formas zwitteriónicas no son estables en fase gas y las transferencias de protón entre los grupos  $\alpha$ -amino y  $\alpha$ -carboxilo ocurren de forma espontánea. Otras especies que presentan complicaciones en el cálculo de  $pK_{a3}$  con ciclos termodinámicos son aquellas que sufren grandes cambios conformacionales como consecuencia de la solvatación/desolvatación.

Nuestro objetivo en este apartado es comparar la precisión de los valores de  $pK_{a3}$  calculados con la reacción isodésmica con los de los ciclos termodinámicos. El cálculo de  $pK_{a3}$  ha sido inviable para muchos péptidos debido a uno o más de los siguientes eventos: (a) transferencia de protones que implica al grupo funcional para el que se calcula el  $pK_{a3}$ . Típicamente, las transferencias de protones desde una lisina, histidina o

cadena lateral de arginina protonadas hacia el grupo  $\alpha$ -carboxilato desprotonado. (b) transferencia de protones que involucra a otros grupos distintos del objeto de estudio y cambia el estado de protonación del péptido considerado. (c) grandes cambios conformacionales del péptido entre la fase gaseosa y disolución acuosa.

La **Tabla 4.15** muestra los valores experimentales de  $pK_{as}$ , los errores absolutos, MAD y SD de los péptidos para los cuales el cálculo de  $pK_{as}$  fue posible con ciclos termodinámicos. Como se observa, únicamente la reacción isodésmica proporciona errores bajos de forma sistemática. Por el contrario, ninguno de los ciclos termodinámicos es una alternativa real para el cálculo de  $pK_{as}$  en péptidos (**Tabla 4.15**).

**Tabla 4.15.** Errores absolutos de los  $pK_{a3}$  calculados con ciclos termodinámicos y con la reacción isodésmica.

	$pK_{a3}$ (exptl.)	C1	C2	C3	Reacción Isodésmica
Tyr	10.60	5.15	6.09	-	-
His	6.00	9.82	10.76	-	-
Arg	12.47	5.04	4.09	-	-
GlyTyr	10.49	1.12	2.06	4.03	2.46
LeuTyr	10.09	3.81	2.87	8.96	0.86
TyrGly	10.12	5.09	6.03	0.06	0.21
GlyGlyTyrAla	10.30	3.32	4.26	1.83	0.80
D-LeuTyr	10.35	2.23	1.28	7.37	0.62
TyrHisOMe	9.69	0.37	0.57	5.52	0.19
PheArg	12.40	22.43	21.48	17.39	1.65
PheAlaArg	12.43	31.10	30.15	26.06	0.79
TyrArg	9.36	7.48	8.42	2.33	1.10
TyrArg	11.62	3.81	2.87	1.23	0.37
TyrHisOMe	6.41	0.98	1.93	8.84	2.18
<b>MAD</b>		<b>7.56</b>	<b>7.67</b>	<b>4.63</b>	<b>1.09</b>
<b>SD</b>		<b>10.46</b>	<b>9.95</b>	<b>3.35</b>	<b>0.82</b>

### Muestreo conformacional y cálculos de $pK_a$

Los péptidos pueden adoptar una enorme gama de conformaciones a temperatura ambiente. Esto implica que el grupo funcional ácido en la secuencia peptídica puede establecer potencialmente muchas interacciones intramoleculares. Por la misma razón, el estado de plegamiento del péptido también modula el grado de exposición al disolvente de los residuos que lo forman. Estos efectos pueden cambiar significativamente los  $pK_{as}$  de los residuos respecto a los  $pK_{as}$  de los aminoácidos correspondientes en disolución.

Es de esperar que la contribución conformacional a las modificaciones de los  $pK_{as}$  sea más importante para cadenas peptídicas largas, ya que pueden adquirir estructura

secundaria e incluso terciaria, que para péptidos pequeños. Se ha realizado un análisis conformacional de cada estado de protonación de los aminoácidos con cadena lateral con propiedades ácido-base y se han recalculado sus  $pK_{a1}$ ,  $pK_{a2}$  y  $pK_{a3}$  como

$$pK_{a(AH)} = pK_{a(BH)} - \log \frac{q_A q_{BH}}{q_{AH} q_B} \quad [4.1]$$

donde  $q_i$  es la función de partición configuracional molecular de la especie  $i$

$$q_i = \sum \exp(-E/k_B T) \quad [4.2]$$

Se generaron entre 10 y 18 conformaciones iniciales aleatorias para cada estado de protonación empleando el software OpenBabel [OpenBabel, 2011]. La búsqueda conformacional se hizo optimizando la diversidad de la desviación cuadrática media [OpenBabel, 2011]. Los conformeros resultantes se minimizaron con el Hamiltoniano PM6 y el modelo continuo SMD. En algunos casos, las geometrías iniciales no eran buenas y la optimización geométrica dio lugar a una quimera química o no convergió. El número final de conformaciones empleadas para cada aminoácido se indica en la **Tabla 4.16**. Como se puede ver, el uso de un ensemble conformacional, incluso si este es pequeño, también conduce a buenas predicciones, pero no implica una mejora sistemática de los  $pK_a$ s calculados con una única conformación (**Tabla 4.16**).

Las energías potenciales ( $E$ ) de las conformaciones de cada estado de protonación presentan desviaciones estándar de 0.5-1 kcal/mol en la mayoría de los casos. Esto muestra que, en general, no hay una conformación mucho más estable que domine sobre las otras. Por este motivo, la energía de una conformación cualquiera representa una buena aproximación a la energía promedio del ensemble, permitiendo obtener buenos resultados sólo con una conformación. Por otra parte, los bajos errores de los  $pK_a$ s calculados indican que, o las funciones de partición calculadas convergen rápidamente o que los errores se cancelan en el cociente de funciones de partición de la **Ecuación 4.2**. Se debe indicar que las variaciones de energía obtenidas se encuentran en los límites de precisión del Hamiltoniano PM6 y modelo continuo empleados. Un resultado positivo de estos resultados es que no es necesario realizar una búsqueda exhaustiva de las

configuraciones de mínima energía absolutas para moléculas pequeñas como los aminoácidos aquí considerados.

**Tabla 4.16.** Errores absolutos de los  $pK_{a1}$ ,  $pK_{a2}$  y  $pK_{a3}$  calculados con PM6 considerando el conjunto conformacional.

	$N^a$	$pK_{a1}$ (exptl.)	$\Delta pK_{a1}$ (PM6)	$pK_{a2}$ (exptl.)	$\Delta pK_{a2}$ (PM6)	$pK_{a3}$ (exptl.)	$\Delta pK_{a3}^b$ (PM6)
Alanina	25	2.34	-	9.69	-		
Ac. aspártico	38	1.89	1.17	9.6	1.21	3.65	1.94
Ac. glutámico	57	2.19	0.16	9.67	2.25	4.25	2.50
Histidina	67	1.82	0.18	9.16	0.81	6.0	0.93
Lisina	61	2.18	0.36	8.94	0.17	10.53	0.57
Arginina	60	2.17	0.03	9.04	0.18	12.47	0.12
Tirosina	46	2.18	0.30	9.11	0.84	10.6	1.71
Cisteína	40	1.71	0.07	10.78	2.49	8.18	2.65
MAD			0.33		1.14		1.49
SD			0.39		0.92		0.97
MAD <sup>c</sup>			0.59		0.9		1.87
SD <sup>c</sup>			0.59		0.62		1.05

<sup>a</sup>Número total de conformaciones empleadas. <sup>b</sup>Moléculas orgánicas usadas como especie de referencia en el cálculo del  $pK_{a3}$ . <sup>c</sup>Valores de MAD y SD calculados sin usar un ensemble conformacional (i.e. con una estructura de cada estado de protonación).

Dado que el espacio configuracional de un péptido crece rápidamente con el número de residuos que lo forman, se quiso evaluar el impacto en el cálculo de  $pK_a$  de emplear un número reducido de conformaciones. Los  $pK_{a1}$ ,  $pK_{a2}$  y  $pK_{a3}$  de los péptidos se calcularon empleando únicamente la conformación de menor energía obtenida en la búsqueda conformacional para la optimización geométrica con PM6/SMD. Las medias de las desviaciones absolutas de estos cálculos se muestran en la **Tabla 4.17**. Comparando los valores de  $pK_a$  obtenidos de esta manera con los obtenidos con PM6/SMD empleando geometrías iniciales generadas manualmente muestra que la búsqueda conformacional mejora las predicciones de  $pK_{a2}$  en  $\sim 0.2$  unidades, pero las predicciones del  $pK_{a1}$  y  $pK_{a3}$  empeoraron en  $\sim 1$  unidad de  $pK_a$  (**Tabla 4.17**). Los resultados de  $pK_{a3}$  empeoraron para todos los grupos funcionales, así que mientras que los errores para lisinas, cisteínas y tirosinas aumentaron moderadamente (i.e. 0.21, 0.25 y 0.7 unidades respectivamente), los errores de los ácidos glutámico/aspártico, histidinas y particularmente argininas aumentaron significativamente (i.e. 1.05, 1.42 y 2.49 unidades) (**Tabla 4.17**).

Un análisis más detallado de los valores individuales de los valores del  $pK_{a3}$  de cada péptido muestra que los errores aumentan para los péptidos más largos y flexibles (**Tabla A18**). Por ejemplo, el error del  $pK_{a3}$  del residuo de lisina en el dipéptido LysGlu

es de 0.44 unidades mientras que en el péptido GlyGlyLysAla es de 5.48 unidades. Se observan tendencias similares para histidinas, tirosinas, argininas y, en menor medida, aspártico y glutámico. Al examinar las estructuras de los péptidos, las conformaciones más estables generadas con OpenBabel tienden a ser compactas. De esta forma se maximizan las interacciones de van der Waals, Coulomb y enlaces de hidrógeno. Esto es así porque OpenBabel realiza la búsqueda conformacional en fase gas. Por lo tanto, las conformaciones generadas no pueden emplearse para validar las estructuras generadas manualmente. Sin embargo, lo que sí muestran claramente estos resultados es que emplear estructuras no representativas de las o conformaciones más pobladas en disolución conduce a grandes errores. Por el mismo motivo, como los errores obtenidos para los  $pK_a$ s de los péptidos en las **Tablas 4.9, 4.10 y 4.16** son bajos y similares a los errores obtenidos para los aminoácidos, se puede concluir que las geometrías generadas manualmente son aproximaciones decentes a las conformaciones más importantes en disolución acuosa.

**Tabla 4.17.** Media de las desviaciones absolutas y desviación estándar de los  $pK_{a1}$ ,  $pK_{a2}$  y  $pK_{a3}$  de péptidos calculados con PM6 usando el confórmero de mínima energía como estructura inicial.

	PM6	PM6 Confórmero de menor energía
$\Delta pK_{a1}^a$	$1.30 \pm 0.96$	$2.40 \pm 1.77$
$\Delta pK_{a2}$	$1.80 \pm 1.18$	$1.62 \pm 1.51$
$\Delta pK_{a3}$ total <sup>b</sup>	$1.17 \pm 0.89$	$2.13 \pm 1.51$
$\Delta pK_{a3}$ lys.	$1.78 \pm 1.11$	$1.99 \pm 1.75$
$\Delta pK_{a3}$ arg.	$0.92 \pm 0.44$	$3.41 \pm 2.97$
$\Delta pK_{a3}$ his.	$1.19 \pm 1.00$	$2.61 \pm 1.15$
$\Delta pK_{a3}$ cys.	$1.21 \pm 0.39$	$1.46 \pm 1.57$
$\Delta pK_{a3}$ tyr.	$1.13 \pm 0.67$	$1.83 \pm 1.09$
$\Delta pK_{a3}$ glu./asp.	$0.50 \pm 0.31$	$1.55 \pm 1.27$

<sup>a</sup>Errores expresados como  $MAD \pm SD$ . <sup>b</sup>Aminoácidos usados como especie de referencia en el cálculo de  $pK_{a3}$ .

### 4.3. Predicción de los sitios de protonación de péptidos tipo Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> mediante descriptores de DFT conceptual.

En los trabajos anteriores se ha mostrado que los pK<sub>as</sub> de moléculas orgánicas, aminoácidos y péptidos, se pueden calcular con métodos de química cuántica y el esquema de la reacción isodésmica. En este trabajo se evalúan descriptores de DFT para el estudio de los sitios de protonación y glicación de péptidos. En particular se estudian los péptidos Ac-Lys-(Ala)<sub>n</sub>-Lys-NH<sub>2</sub> (abreviado KA<sub>n</sub>K), estudiados experimentalmente por Makowska *et al.* [Makowska, 2012], donde dos residuos lisina se encuentran separados por un número variable ( $0 \leq n \leq 5$ ) de alaninas intermedias.

Las energías de los orbitales HOMO y LUMO, potenciales de ionización verticales  $I$ , afinidades electrónicas  $A$ , electronegatividades globales  $\chi$ , durezas químicas totales  $\eta$  y electrofilias globales  $\omega$  se obtuvieron a partir del teorema de Koopman o KID (“Koopman’s in DFT”) y a partir de cálculos de las energías verticales de ionización según el procedimiento  $\Delta$ SCF. Este procedimiento consiste en hacer cálculos puntuales de energía de las especies catión radical sobre las geometrías optimizadas de las especies neutras para obtener las energías verticales de ionización.

Para analizar la calidad de los resultados obtenidos con el procedimiento KID, se calcularon varios descriptores que relacionan los resultados KID con los  $\Delta$ SCF. Los tres primeros se relacionan con el cumplimiento del teorema de Koopman relacionando  $\varepsilon_H$  con  $-I$ ,  $\varepsilon_L$  con  $-A$ , y el comportamiento de éstos en la descripción de la banda prohibida:

$$J_I = \left| \varepsilon_H + E_{gs}(N-1) - E_{gs}(N) \right| \quad [4.3]$$

$$J_A = \left| \varepsilon_L + E_{gs}(N) - E_{gs}(N+1) \right| \quad [4.4]$$

$$J_{HL} = \sqrt{J_I^2 + J_A^2} \quad [4.5]$$

También consideramos otros cuatro descriptores que analizan cuán bien los funcionales empleados son útiles para la predicción de la electronegatividad  $\chi$ , dureza global  $\eta$ , electrofilia global  $\omega$  y una combinación de éstos considerando las energías HOMO y LUMO o las energías de ionización,  $I$ , y afinidades electrónicas,  $A$ , verticales

$$J_{\chi} = |\chi - \chi_K| \quad [4.6]$$

$$J_{\eta} = |\eta - \eta_K| \quad [4.7]$$

$$J_{\omega} = |\omega - \omega_K| \quad [4.8]$$

$$J_{D1} = \sqrt{J_{\chi}^2 + J_{\eta}^2 + J_{\omega}^2} \quad [4.9]$$

donde el subíndice  $K$  en las **Ecuaciones 4.7, 4.8 y 4.9** indica que el descriptor se ha obtenido según el procedimiento KID y el subíndice D1 indica el primer grupo de descriptores conceptuales DFT. Los resultados de  $J_I$ ,  $J_A$ ,  $J_{HL}$ ,  $J_{\chi}$ ,  $J_{\eta}$ ,  $J_{\omega}$  y  $J_{D1}$  de los seis péptidos se muestran en las **Tablas A19-A30**. Basados en estos resultados, se han recopilado los valores promedio obtenidos según cada funcional de correlación intercambio empleado para el conjunto de los péptidos en la **Tabla 4.18**.

Como puede observarse de la **Tabla 4.18**, el procedimiento KID se cumple con gran precisión para los funcionales MN12SX y N12SX que son respectivamente funcionales híbridos meta-NGA de rango separado e híbrido NGA de rango separado. De hecho, los valores de  $J_I$ ,  $J_A$  y  $J_{HL}$  no son exactamente cero. Sin embargo, los resultados son muy buenos, en particular para el funcional N12SX.

**Tabla 4.18.** Media de los descriptores  $J_\chi$ ,  $J_\eta$ ,  $J_\omega$ ,  $J_{D1}$ ,  $J_I$ ,  $J_A$  y  $J_{HL}$  de los péptidos  $K(A)_nK$  calculados con los funcionales M11, M11L, MN12L, MN12SX, N12 y N12SX.

	$J_\chi$	$J_\eta$	$J_\omega$	$J_{D1}$	$J_I$	$J_A$	$J_{HL}$
<b>M11</b>	0.20	5.37	0.37	5.19	2.79	2.38	3.67
<b>M11L</b>	0.80	1.73	0.17	1.91	1.66	0.07	1.67
<b>MN12L</b>	0.87	2.01	0.80	2.20	1.87	0.14	1.88
<b>MN12SX</b>	0.09	0.06	0.05	0.14	0.13	0.08	0.15
<b>N12</b>	0.59	1.62	0.10	1.74	1.40	0.22	1.44
<b>N12SX</b>	0.03	0.07	0.03	0.08	0.00	0.07	0.07

Es interesante ver que los mismos funcionales también cumplen el procedimiento KID para los otros descriptores, concretamente  $J_\chi$ ,  $J_\eta$ ,  $J_\omega$  y  $J_{D1}$ . Estos resultados son muy importantes porque muestran que no es suficiente confiar únicamente en  $J_I$ ,  $J_A$  y  $J_{HL}$ . Por ejemplo, si sólo se considera  $J_\chi$  para todos los funcionales, todos presentan valores cercanos a cero. En lo que respecta al resto de descriptores, sólo MN12SX y N12SX muestran este comportamiento. Esto quiere decir que los resultados para  $J_\chi$  pueden ser fortuitos debido a cancelación de errores.

Un hecho importante es que a pesar de que los funcionales híbridos NGA de rango separado e híbridos meta-NGA de rango separado pueden ser útiles para el cálculo de descriptores de DFT conceptual, no es lo mismo para el híbrido GGA de rango separado M11. Una inspección de la **Tabla A19** muestra que esto es debido a que este funcional describe erróneamente la energía del LUMO, conduciendo a valores negativos de  $A$  que están en contradicción con los resultados  $\Delta$ SCF.

Las funciones condensadas de Fukui,  $f$ , y el descriptor dual condensado,  $\Delta f_K$ , se calcularon empleando el programa AOMix [Gorelsky, 2011], [Gorelsky, 2001] a partir de cálculos de energía puntuales. En la **Tabla 4.19** se muestran los descriptores duales condensados  $\Delta f_K$  y las funciones de Parr de electrofilia  $P_{k(MPA)}^-$  y  $P_{k(HPA)}^-$  sobre los átomos de nitrógeno de los residuos lisina de los péptidos calculados con los funcionales MN12SX y N12SX con las bases Def2TZVP y agua como disolvente según el modelo SMD.

**Tabla 4.19.** Valores del descriptor dual  $\Delta f_K$  y de la función de Parr electrófila  $P_{k(MPA)}^-$  y  $P_{k(HPA)}^-$  para los péptidos  $K(A)_nK$  calculados con los funcionales MN12SX y N12SX.

	MN12SX			N12SX		
	$\Delta f_K$	$P_{k(MPA)}^-$	$P_{k(HPA)}^-$	$\Delta f_K$	$P_{k(MPA)}^-$	$P_{k(HPA)}^-$
<b>KK</b>	-0.63	0.76	0.74	-0.69	0.76	0.73
<b>KAK</b>	-0.62	0.76	0.73	-0.69	0.76	0.73
<b>KA2K</b>	-0.71	0.81	0.78	-0.74	0.78	0.73
<b>KA3K</b>	-0.63	0.76	0.74	-0.69	0.76	0.73
<b>KA4K</b>	-0.71	0.81	0.78	-0.74	0.81	0.78
<b>KA5K</b>	-0.71	0.81	0.78	-0.74	0.81	0.78

Los valores mostrados en la **Tabla 4.19** corresponden a los descriptores duales condensados y las funciones de Parr de los nitrógenos de las lisinas más próximas al extremo C-terminal (Lys1). Los valores de los descriptores correspondientes a la otra lisina (Lys2) no se muestran porque son cero en casi todos los casos. Esto implica que el sitio preferente de glicación será Lys1, aunque también será el sitio preferente de protonación. Por tanto, se puede predecir que el  $pK_a$  de Lys1 será mayor que el de Lys2, aunque no es posible asignar un valor absoluto a ninguno de ellos. Además, aunque no haya una tendencia clara, de los resultados de la **Tabla 4.19** se puede observar que los descriptores duales condensados y las funciones de Parr, tanto del análisis de población de Mulliken MPA como de Hirshfeld HPA, que los valores de los descriptores son mayores para los péptidos KA4K y KA5K que para los otros péptidos. Por tanto, se puede concluir que el efecto de los espaciadores alanina será el de incrementar la diferencia entre los  $pK_{as}$  de Lys1 y Lys2. Esto probablemente se puede explicar considerando que el incremento en el número de alaninas conduce a conformaciones en las que las lisinas están separadas una de otra y no compiten por la reacción con los protones.

#### 4.4. Capacidad antioxidante de la tiamina

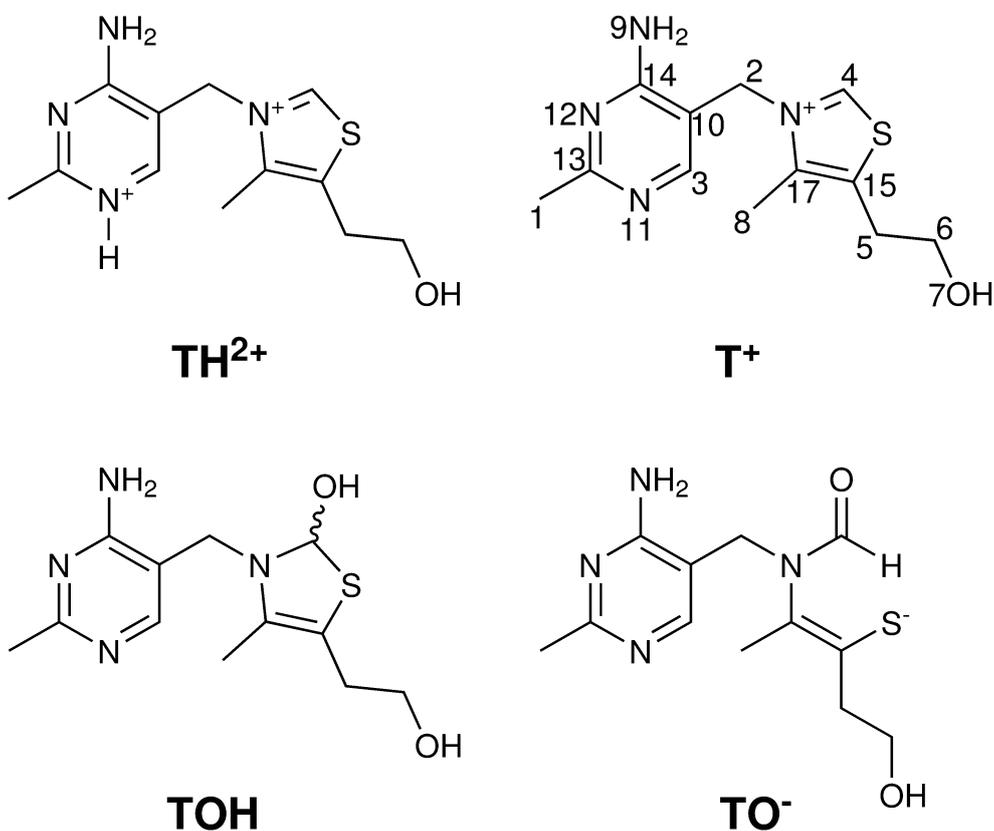
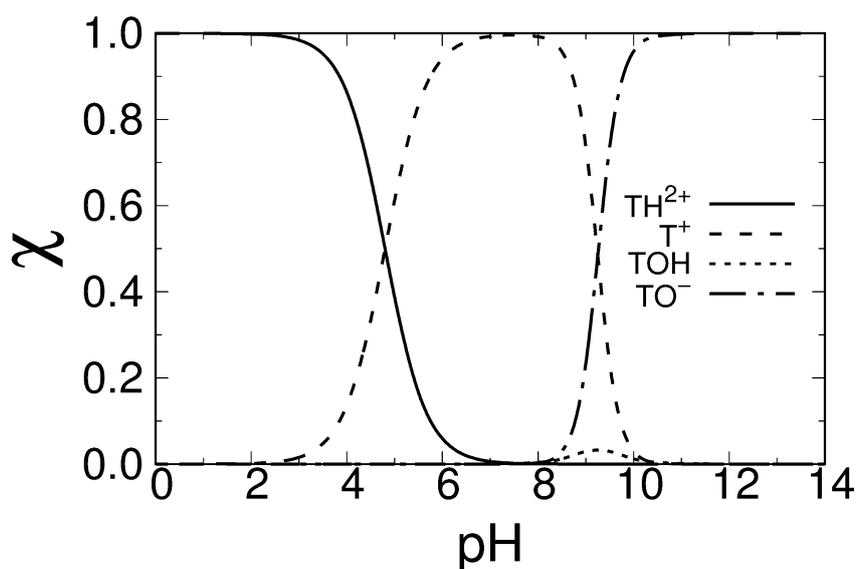
Se ha evaluado la reactividad de la Vitamina B1, o tiamina, frente a algunos de los radicales libres más usuales: metoxilo ( $\bullet\text{OCH}_3$ ), hidroperoxilo ( $\bullet\text{OOH}$ ) y metilperoxilo ( $\bullet\text{OOCH}_3$ ), teniendo en cuenta tres posibles mecanismos de reacción (HAT, RAF y SET) en disolución acuosa y en un disolvente orgánico para representar las condiciones de reacción en el interior de una membrana lipídica. El radical hidroperoxilo ( $\bullet\text{OOH}$ ) es una especie capaz de difundirse hasta localizaciones celulares remotas, [Pryor, 1986] con una vida media del orden de los segundos [Marnett, 1987]. Además, su comportamiento es probablemente similar al de los radicales peroxilo mayores,  $\bullet\text{OOR}$ , que son abundantes en sistemas biológicos. Por lo tanto, el empleo de  $\bullet\text{OOH}$  puede proporcionar una visión de las reacciones de otros radicales importantes.

Los tres posibles mecanismos de reacción mencionados corresponden con las reacciones químicas expresadas por las **ecuaciones 1.1, 1.2 y 1.3**.

#### Equilibrios ácido-base de la tiamina en disolución acuosa

La tiamina presenta diferentes estructuras dependiendo del pH puesto que es una molécula con propiedades ácido-base. A  $\text{pH} < 5$  está protonada ( $\text{TH}^{2+}$ ), existiendo tres posibles tautómeros dependiendo de si la protonación tiene lugar en el N11, N12 o N9 (**Figura 4.4**). Las energías libres de Gibbs calculadas para los tres tautómeros  $\text{TH}^{2+}$  indican que el más estable es el que está protonado en N11, con una diferencia de energías de 4.5, 19.8 kcal/mol respecto a los tautómeros protonados en N12 y N9 respectivamente. Lo mismo ocurre con los cálculos de la energía libre de Gibbs en etanoato de pentilo, en el que el tautómero protonado en N11 es 7.3 kcal/mol más estable que el protonado en N12. Esto implica que, según una distribución de Boltzmann, a 298.15 K y  $\text{pH} < 5$ , más del 99% de la tiamina existe en la forma del tautómero protonado en el N11 en ambos ambientes.

Se ha calculado el  $\text{p}K_a$  de los tres tautómeros  $\text{TH}^{2+}$  en disolución acuosa usando el método de la reacción isodésmica según una reacción de intercambio de protón [Casasnovas, 2011], [Casasnovas, 2014], [Sastre, 2013], [Sastre, 2016]. En nuestro caso hemos elegido la piridina como ácido de referencia para la  $\text{TH}^{2+}$  protonada en el N11 y en el N12 y la anilina para el N9. Los valores de  $\text{p}K_a$  obtenidos son 4.4, 1.1 y -7.7 respectivamente para el N11, N12 y N9, lo que indica que el N11 es más básico que N12 y N9. En consecuencia, la forma  $\text{TH}^{2+}$  considerada en el presente trabajo es la que está protonada en el N11 (**Figura 4.4**).



**Figura 4.4.** Diagrama de distribución (fracción molar) de las diferentes estructuras de la tiamina en función del pH. Diferentes estructuras de la tiamina: TH<sup>2+</sup>, T<sup>+</sup>, TOH y TO<sup>-</sup>. En la estructura T<sup>+</sup> se muestra la numeración de los átomos.

En medios básicos, es decir pH > 8, se ha propuesto que la tiamina existe en forma del anión tiolato con el anillo abierto TO<sup>-</sup> [Caballero, 2011]. El mecanismo de formación TO<sup>-</sup> propuesto en la literatura implica la existencia de una especie tetraédrica TOH (Figura 4.4) [Caballero, 2011].

Según los experimentos de Pérez-Caballero *et al.* [Caballero, 2011] la primera desprotonación ocurre en el anillo de piperidinio con un  $pK_a$  de 4.80, mientras que la segunda y tercera etapas de desprotonación corresponden a la formación de la especie tetraédrica TOH con un  $pK_a$  de 10.40 y a la desprotonación del hemitioacetal, lo que provoca la abertura del anillo de tiazol para dar la especie  $TO^-$ , con un  $pK_a$  de 8.10. A partir de estos valores, las fracciones molares indican que en solución acuosa la forma dominante de la tiamina a pH fisiológico es la del monocatión  $T^+$ , con una población superior al 99% (**Figura 4.4**). A pH ligeramente básico, las formas TOH y  $TO^-$  tienen fracciones molares bajas, pero no despreciables, de aproximadamente un 10%. Por lo tanto, estas formas de tiamina también se han considerado en el estudio de las reacciones de eliminación de radicales libres.

Se ha realizado un análisis conformacional de la tiamina generando múltiples conformeros iniciales variando sistemáticamente los diedros C3-C10-C2-N y C10-C2-N-C17 en incrementos de  $30^\circ$ . Este procedimiento generó conformeros que presentan átomos colisionando unos con otros, por lo que dichas estructuras fueron descartadas. Las estructuras restantes fueron sometidas a optimización geométrica manteniendo restricciones en los diedros mencionados. Al final se obtuvieron un total de 38 conformaciones, de las cuales 4 presentan diferencias de energía inferiores a  $k_B T$  (a 310 K) respecto a la conformación de mínima energía encontrada. Dado el número relativamente reducido de conformaciones accesibles a temperatura fisiológica, se empleó la conformación de mínima energía para el resto de cálculos de energías y barreras de reacción.

Se ha calculado que la energía libre de todas las formas de la tiamina es menor en disolución acuosa que en PE. Así la diferencia entre la  $\Delta G$  en agua y en PE es de -32 kcal/mol para  $TH^{2+}$ , -12 kcal/mol para la  $T^+$ , -6 kcal/mol para TOH y -13.6 kcal/mol para  $TO^-$ . Puede apreciarse que la diferencia es mayor si la tiamina está cargada. Por tanto los coeficientes de partición de cada forma serán -23, -9, -4, -10, indicando que la tiamina estará de forma muy minoritaria en entornos lipídicos o apolares.

### **Viabilidad termodinámica**

En primer lugar, se ha investigado la viabilidad termoquímica de los diferentes mecanismos de reacción y posibles canales.

#### *Single-electron transfer (SET)*

Las energías libres de Gibbs para las reacciones SET se presentan en la **Tabla 4.20**. Sus valores indican que la mayoría de las reacciones SET de la tiamina no son termodinámicamente espontáneas ( $\Delta G > 0$ ) en ninguno de los dos medios.

**Tabla 4.20.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones SET entre la tiamina y los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOH}$  y  $\bullet\text{OOCH}_3$  en disolución acuosa y en PE.

	Agua				PE			
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\bullet$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\bullet$
$\text{CH}_3\text{O}^\bullet$	55.28	42.33	-1.82	-7.21	126.23	98.55	42.18	16.41
$\text{HOO}^\bullet$	61.54	48.59	4.44	-0.95	135.04	107.36	50.99	25.22
$\text{CH}_3\text{OO}^\bullet$	63.44	50.49	6.34	0.96	136.58	108.90	52.53	26.76

Como se puede ver en la **Tabla 4.20**, las reacciones SET termodinámicamente más favorables tienen lugar con el radical  $\bullet\text{OCH}_3$  siendo las energías libres de reacción 6.3 y 8.2 kcal/mol más positivas para el radical  $\bullet\text{OOH}$  y  $\bullet\text{OOCH}_3$  respectivamente.

Los cambios en el estado de protonación de la tiamina tienen mayor impacto en la estabilidad de los productos formados, mostrando diferencias de 62.5 kcal/mol entre las energías libres de reacción de  $\text{TH}^{2+}$  y  $\text{TO}^\bullet$  en solución acuosa y de 109.8 kcal/mol en medio orgánico.

Los datos de la **Tabla 4.20** también muestran que las reacciones SET de la tiamina están desfavorecidas en ambientes lipídicos ya que las energías libres de reacción son ~ 23 a ~ 73 kcal/mol más positivas que en solución acuosa (**Tabla 4.20**). En este caso la razón es que los productos resultantes metilhidróxido, hidroperóxido e hidropéroxido de metilo aniones pequeños que se encuentra mejor estabilizados en medios altamente polares que en ambientes no polares.

#### *Hydrogen Atom Transfer (HAT)*

En la **Tablas 4.21**, **4.22** y **4.23** se presentan las energías libres de Gibbs de las reacciones HAT con los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$ , respectivamente, para todas las vías estudiadas en disolución acuosa y en medio lipídico (PE).

**Tabla 4.21.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones HAT entre la tiamina y el radical  $\bullet\text{OCH}_3$  en disolución acuosa y en PE.

	Agua				PE			
	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>•</sup>	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>•</sup>
<b>C1</b>	-11.18	-12.47	-12.20	-12.63	-8.85	-8.72	-8.87	-11.17
<b>C2</b>	-13.48	-12.76	-17.54	-20.14	-11.18	-12.61	-19.21	-18.15
<b>C3</b>	14.32	2.51	4.24	3.76	17.91	4.24	4.16	3.81
<b>C4</b>	22.93	16.58	-14.76	-9.61	18.68	16.96	-11.87	-8.09
<b>C5</b>	-16.66	-16.98	-20.49	-19.82	-11.25	-16.23	-19.18	-16.43
<b>C6</b>	-9.47	-9.29	-9.52	-11.73	-8.98	-5.11	-9.26	-10.63
<b>O7</b>	0.26	0.52	-6.55	-7.21	1.69	1.42	-1.01	-0.85
<b>C8</b>	-13.75	-13.25	-20.12	-27.01	-11.18	-11.74	-19.01	-26.97
<b>N9</b>	9.90	-0.47	-8.22	-10.10	12.64	2.16	-1.29	-8.91
<b>N11</b>	11.13				18.35			

**Tabla 4.22.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones HAT entre la tiamina y el radical  $\bullet\text{OOCH}_3$  en disolución acuosa y en PE.

	Agua				PE			
	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>•</sup>	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>•</sup>
<b>C1</b>	7.99	6.70	6.97	6.54	11.10	11.23	11.08	8.79
<b>C2</b>	5.69	6.41	1.64	-0.97	8.77	7.34	0.74	1.80
<b>C3</b>	33.49	21.68	23.41	22.93	37.86	24.19	24.11	23.76
<b>C4</b>	42.10	35.75	4.41	9.56	38.63	36.92	8.08	11.86
<b>C5</b>	2.51	2.19	-1.32	-0.65	8.70	3.72	0.77	3.52
<b>C6</b>	9.70	9.89	9.65	7.44	10.97	14.84	10.69	9.32
<b>O7</b>	19.43	19.69	12.62	11.96	21.64	21.37	18.94	19.10
<b>C8</b>	5.42	5.92	-0.95	-7.84	8.77	8.21	0.94	-7.02
<b>N9</b>	29.07	18.72	10.95	9.07	32.59	22.11	18.66	11.04
<b>N11</b>	30.30				38.30			

**Tabla 4.23.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones HAT entre la tiamina y el radical  $\bullet\text{OOH}$  en disolución acuosa y en PE.

	Agua				PE			
	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>-</sup>	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>-</sup>
<b>C1</b>	6.47	5.17	5.44	5.01	8.84	8.97	8.82	6.53
<b>C2</b>	4.16	4.88	0.11	-2.50	6.51	5.08	-1.52	-0.45
<b>C3</b>	31.96	20.15	21.88	21.40	35.60	21.94	21.85	21.51
<b>C4</b>	40.57	34.23	2.88	8.03	36.37	34.66	5.82	9.60
<b>C5</b>	0.98	0.67	-2.85	-2.17	6.45	1.47	-1.48	1.27
<b>C6</b>	8.17	8.36	8.13	5.91	8.71	12.59	8.43	7.06
<b>O7</b>	17.91	18.17	11.09	10.44	19.39	19.11	16.68	16.85
<b>C8</b>	3.90	4.39	-2.48	-9.37	6.51	5.95	-1.32	-9.28
<b>N9</b>	27.54	17.19	9.42	7.55	30.34	19.85	16.40	8.78
<b>N11</b>	28.78				36.05			

La carga eléctrica global de cada molécula se conserva en las reacciones HAT. Esto sugiere que la polaridad del medio no debería tener un gran impacto en los valores de la energía libre de la reacción. De hecho, en promedio estas son sólo 2 kcal/mol más negativas en disolución acuosa que en el medio lipídico (**Tablas 4.21, 4.22 y 4.23**).

La termodinámica de las reacciones HAT también está influenciada por los diferentes sitios de reacción de la tiamina. Además, la reactividad de cada sitio también puede ser modulada por el pH ya que los cambios en el estado de protonación afectan la polaridad local y la carga eléctrica de la molécula. Por ejemplo, la reacción HAT en el C1 muestra energías libres de reacción similares para TH<sup>2+</sup>, T<sup>+</sup>, TOH y TO<sup>-</sup> porque los radicales generados se estabilizan por deslocalización a través del anillo de aminopirimidina independientemente del estado de protonación. En contraposición, las energías libres de reacción de la reacción HAT en el C4 dependen fuertemente del estado de protonación. Para TH<sup>2+</sup> y T<sup>+</sup>, la abstracción de hidrógeno genera un radical completamente localizado en C4. Esto aumenta la electronegatividad del C4, pero no puede polarizar fácilmente el azufre del tioéter ni el nitrógeno positivo del grupo iminio ya que estos también son electronegativos. Como resultado, la reacción HAT en esta posición es endergónica. En cambio, para la forma TOH, la reacción HAT en el C4 es exergónica porque el radical generado puede ser deslocalizado a través del sistema  $\pi$  de anillo de tiazol. De manera similar, los radicales formados en el C3 no se pueden deslocalizar a otros centros y las reacciones son endergónicas.

Una tendencia que se puede observar para la reactividad entre diferentes sitios es que las energías libres de reacción se vuelven más negativas en el orden C6, C1, C2 y

C5-C8 (Tablas 4.21, 4.22 y 4.23). Dos factores deben ser considerados para explicar esta tendencia: 1) el aumento/disminución de la deslocalización del radical después de la reacción, que estabiliza/desestabiliza los productos finales y 2) la posible pérdida de aromaticidad de los anillos de aminopirimidina y/o tiazol, que desestabiliza los productos. Estos factores también ayudan a entender las importantes diferencias de energía libre observadas para la reacción HAT con los diferentes radicales. Los resultados muestran que las reacciones HAT son termodinámicamente más favorables con el radical  $\bullet\text{OCH}_3$  que con los radicales  $\bullet\text{OOH}$  y  $\bullet\text{OOCH}_3$  (17.6 y 19.2 kcal/mol respectivamente). La explicación es que el electrón desapareado en  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  está más deslocalizado que en  $\bullet\text{OCH}_3$ . Por lo tanto en la mayoría de las reacciones con  $\bullet\text{OCH}_3$ , dicho electrón está más deslocalizado en el producto radical de tiamina. Por otra parte, no todas las reacciones con  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  implican un aumento en la deslocalización del electrón desapareado y pueden alterar la aromaticidad de la tiamina. En realidad, la única reacción HAT exergónica para  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  tiene lugar en los C2, C5 y C8 de la forma  $\text{TO}^-$  y en el C5 y el C8 de la forma  $\text{TOH}$ , donde el anillo tiazol no es aromático.

#### *Radical Adduct Formation (RAF)*

En la Tablas 4.24, 4.25 y 4.26 se presentan los valores de la energía libre de las reacciones RAF con los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  para todas las vías estudiadas en disolución acuosa y en medio lipídico (PE).

**Tabla 4.24.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones RAF entre la tiamina y el radical  $\bullet\text{OCH}_3$  en disolución acuosa y en PE.

	Agua					PE				
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^-$		$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^-$	
<b>C3</b>	-3.15	4.71	3.77	4.12	7.40	1.25	6.27	4.58	6.47	8.04
<b>C4</b>	-19.88	-18.83				-17.65	-16.69			
<b>C10</b>	-0.75	-1.01	4.05	0.04	-0.34	-0.23	-1.25	6.26	0.60	2.56
<b>C13</b>	12.43	5.45	8.39	2.83	6.84	17.74	11.45	21.05	5.20	5.63
<b>C14</b>	17.90	14.77	14.30	9.29	8.64	22.08	14.26	14.48	15.83	13.53
<b>C15</b>	-3.19	-4.19	-10.67	-12.78	-8.07	-5.85	-5.34	-12.18	-12.61	-4.24
<b>C17</b>	-1.26	-1.87	-13.73	-12.85	-16.40	-1.99	-0.59	-16.10	-11.86	-17.55

**Tabla 4.25.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones RAF entre la tiamina y el radical  $\bullet\text{OOH}$  en disolución acuosa y en PE.

	Agua					PE				
	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>-</sup>		TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>-</sup>	
<b>C3</b>	16.11	20.29	19.10	18.86	18.10	19.67	23.59	18.69	21.00	16.87
<b>C4</b>	-0.23	0.60				3.54	4.35			
<b>C10</b>	17.49	14.92	18.31	14.44	15.18	19.94	17.85	21.44	16.89	16.41
<b>C13</b>	27.85	20.20	20.71	18.11	18.37	34.76	27.75	35.02	21.03	27.61
<b>C14</b>	36.52	28.94	29.49	21.98	19.01	42.69	32.10	30.24	30.98	21.55
<b>C15</b>	12.23	12.69	1.22	1.74	3.80	13.25	13.62	0.94	2.14	2.18
<b>C17</b>	16.35	16.63	-0.52	1.27	0.37	17.38	16.00	2.50	5.32	-2.54

**Tabla 4.26.** Energía libre de Gibbs a 298.15 K (kcal/mol) de las reacciones RAF entre la tiamina y el radical  $\bullet\text{OOCH}_3$  en disolución acuosa y en PE.

	Agua					PE				
	TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>-</sup>		TH <sup>2+</sup>	T <sup>+</sup>	TOH	TO <sup>-</sup>	
<b>C3</b>	19.82	21.48	21.82	22.22	21.38	22.47	26.37	23.35	23.43	23.76
<b>C4</b>	1.44	2.69				5.01	5.52			
<b>C10</b>	20.38	18.72	17.22	19.42	17.65	22.77	20.15	23.56	19.12	20.85
<b>C13</b>	30.20	23.19	20.20	22.71	20.31	36.21	31.26	38.28	25.07	30.65
<b>C14</b>	40.24	31.73	23.81	32.09	22.81	43.52	34.11	36.10	35.10	27.99
<b>C15</b>	12.15	14.49	3.50	3.72	3.54	21.39	18.27	5.02	5.54	6.68
<b>C17</b>	19.82	18.93	2.25	2.55	2.35	16.54	17.24	5.27	6.96	3.79

La adición del radical en los sitios C3, C4, C10, C13, C14, C15 y C17 genera nuevos centros quirales. Dado que la TOH ya tiene un centro quiral en C4, las reacciones RAF en la estructura TOH producen diastereómeros. Las dos estructuras posibles se han optimizado y se han obtenido dos valores de la energía libre de reacción (**Tablas 4.24, 4.25 y 4.26**).

La reacción RAF no modifica la carga eléctrica de los reactivos por lo que la polaridad del medio tiene poca influencia en la espontaneidad termodinámica. Los valores de la energía libre de Gibbs son solo de unas 5 kcal/mol menores en disolución acuosa que en el medio lipídico (**Tablas 4.24, 4.25 y 4.26**).

Se observan diferencias de energía mucho mayores para la reacción RAF dependiendo del tipo de radical. En particular, la reacción RAF con el radical  $\bullet\text{OCH}_3$  es 23.0 y 19.2 kcal/mol más favorable que con  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  respectivamente.

Las energías de la reacción RAF para diversos sitios también demuestran variaciones importantes. La siguiente discusión se centrará en las RAF con el radical  $\bullet\text{OCH}_3$ . Se puede observar que para cada forma de la tiamina, la formación del aducto es preferente en el anillo de tiazol (C4, C15 y C17) respecto al anillo de aminopirimidina (C3, C10, C13 y C14). El estado de protonación de la tiamina modula la reactividad de

estos sitios. Por ejemplo, la reacción RAF en el C3 y el C10 es termodinámicamente favorable solamente para las formas  $\text{TH}^{2+}$  y  $\text{T}^+$  debido a los átomos de nitrógeno vecinos positivos. Por otro lado, para  $\text{TH}^{2+}$  y  $\text{T}^+$ , la RAF en el C4 es claramente preferencial con energías libres de reacción de aproximadamente -20 kcal/mol, mientras que para TOH y  $\text{TO}^\cdot$  los sitios preferenciales son el C15 y el C17.

Se puede observar que las reacciones más exergónicas de la reacción RAF se obtienen para las formas  $\text{TH}^{2+}$  y  $\text{T}^+$ , mientras que en las reacciones SET y HAT las reacciones más exergónicas corresponden a las estructuras TOH y  $\text{TO}^\cdot$ .

### **Viabilidad cinética**

Después del estudio termoquímico de los diferentes mecanismos, se ha realizado un estudio cinético para las reacciones exergónicas. La cinética de las vías calculadas como endergónicas no se ha calculado porque, incluso si estas vías se producen a altas velocidades, al ser la reacción endergónica probablemente la velocidad inversa también es alta, haciendo la reacción reversible por lo que, después de un tiempo transitorio, los productos formados serían menores. Este enfoque se ha utilizado con éxito en las investigaciones de la capacidad antioxidante de otras moléculas [Galano, 2016], [Mazzone, 2016]. Por este motivo, el mecanismo SET no se ha considerado para el estudio cinético ya que la gran mayoría de energías libres de reacción calculadas son endergónicas. Únicamente se observa una reacción exergónica entre la forma  $\text{TO}^\cdot$  y el radical  $\cdot\text{OCH}_3$  en medio acuoso.

### *Hydrogen Atom Transfer (HAT)*

En la **Tablas 4.27 y 4.28** se presentan los valores de la energía libre de las reacciones HAT con los radicales  $\cdot\text{OCH}_3$ ,  $\cdot\text{OOCH}_3$  y  $\cdot\text{OOH}$  para todas las vías estudiadas en disolución acuosa y en medio lipídico (PE) respectivamente.

Como se ha visto antes, las reacciones termodinámicamente más favorables de las HAT tienen lugar frente al radical  $\cdot\text{OCH}_3$ . Sin embargo, mientras que las energías libres de reacción con  $\cdot\text{OOCH}_3$  y  $\cdot\text{OOH}$  eran 17-19 kcal/mol más altas, las barreras de energía libre son comparables. Una inspección de las distancias interatómicas de las estructuras del estado de transición HAT no muestra correlación entre la magnitud de las barreras de energía libre y las distancias X-O ó X-H, siendo X, H y O el átomo pesado en la tiamina, el hidrógeno transferido y el oxígeno del radical respectivamente (Figura 4.5, **Tablas A31, A32**).

**Tabla 4.27.** Energía libre de Gibbs de activación a 298.15 K (kcal/mol) de las reacciones HAT entre la tiamina y los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  en disolución acuosa.

	$\bullet\text{OCH}_3$				$\bullet\text{OOCH}_3$		$\bullet\text{OOH}$			
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}\cdot$	$\text{TOH}$	$\text{TO}\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}\cdot$
<b>C1</b>	17.56	16.37	16.07	10.17						
<b>C2</b>	18.11	16.12	14.79	9.08		21.60			23.28	19.42
<b>C4</b>			1.85							
<b>C5</b>	13.66	14.23	5.02	12.81	16.87	19.58	22.78	23.81	15.70	14.67
<b>C6</b>	14.40	15.00	13.59	11.69						
<b>O7</b>	17.36	18.35	6.79	12.13						
<b>C8</b>	15.84	16.36	4.22	14.32	16.68	11.91			13.94	19.55
<b>N9</b>		16.39	8.49	18.12						

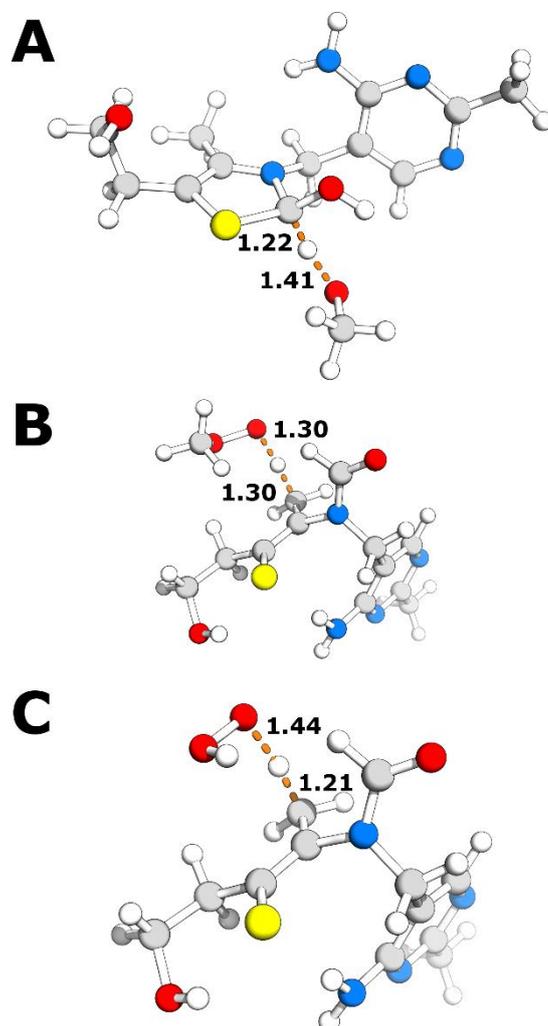
**Tabla 4.28.** Energía libre de Gibbs de activación  $\Delta G^\ddagger$  a 298.15 K (kcal/mol) de las reacciones HAT entre la tiamina y los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  en PE.

	$\bullet\text{OCH}_3$				$\bullet\text{OOCH}_3$		$\bullet\text{OOH}$	
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}\cdot$	$\text{TOH}$	$\text{TO}\cdot$	$\text{TOH}$	$\text{TO}\cdot$
<b>C1</b>	19.42	18.17	16.40	22.26				
<b>C2</b>	22.60	17.83	14.57	15.71	22.15		22.85	17.77
<b>C4</b>			9.65					
<b>C5</b>	17.83	18.34	14.86	13.95	23.18		22.08	
<b>C6</b>	18.83	15.70	13.11	13.59				
<b>O7</b>			11.96	24.22				
<b>C8</b>	18.95	15.98	11.83	17.03	26.71	11.11	21.58	11.01
<b>N9</b>			9.13					

La energía de estas reacciones está muy influenciada por el grado de deslocalización del electrón desapareado y la posible desaparición de la aromaticidad en la tiamina. Por lo tanto, la razón por la cual las barreras son similares para los tres radicales es que durante la transferencia de hidrógeno el electrón desapareado está deslocalizado entre tres centros  $\text{X} \bullet \bullet \text{H} \bullet \bullet \text{O}$  de manera similar para  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$ . También por la misma razón, en general la polaridad del entorno no implica modificaciones mayores de la barrera de energía libre (**Tablas 4.27 y 4.28**).

Las barreras más bajas se observan para las reacciones de  $\bullet\text{OCH}_3$  con la forma  $\text{TOH}$  en las posiciones C4, C8 y C5 en solución acuosa, 1.85, 4.22 y 5.02 kcal/mol respectivamente (**Tablas 4.27**). Por otro lado, las barreras más bajas de las reacciones con  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  también corresponden a reacciones en el C8 pero con la forma  $\text{TO}\cdot$  y  $\text{TOH}$  (**Tablas 4.27**). En todos estos casos, en el estado de transición de la extracción del hidrógeno, el electrón desapareado puede estabilizarse mediante una deslocalización

extensa a través del sistema  $\pi$  de la tiamina sin desestabilizar el sistema  $\pi$  aromático (Figura 4.5).



**Figura 4.5.** Estructuras de los estados de transición (TS) de las reacciones HAT con las barreras más bajas para la eliminación de  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$ . A) Estado de transición de la reacción entre  $\bullet\text{OCH}_3$  y TOH en C4 en disolución acuosa. B) Estado de transición entre  $\bullet\text{OOCH}_3$  y  $\text{TO}^-$  en C8 en PE. C) Estado de transición de la reacción entre  $\bullet\text{OOH}$  y  $\text{TO}^-$  en C8 en PE. Las distancias relevantes se indican en Å y se muestran con líneas segmentadas. Figura creada con Open-source PyMOL software [Schrödinger, 2011]

#### *Radical Adduct Formation (RAF)*

En las **Tablas 4.29** y **4.30** se presentan los valores de la energía libre de las reacciones RAF con los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  para todas las vías estudiadas en disolución acuosa y en medio lipídico (PE).

**Tabla 4.29.** Energía libre de Gibbs de activación,  $\Delta G^\ddagger$ , a 298.15 K (kcal/mol) de las reacciones RAF entre la tiamina y los radicales  $\bullet\text{OCH}_3$  y  $\bullet\text{OOH}$  en disolución acuosa.

	$\bullet\text{OCH}_3$				$\bullet\text{OOH}$			
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}\cdot$
<b>C3</b>	19.76							
<b>C4</b>	14.00	14.78			22.18	23.63		
<b>C10</b>	21.16	17.82	16.19		15.07			
<b>C15</b>	16.77	15.83	6.37	4.75	7.74			
<b>C17</b>	13.84	15.38	2.24	3.74	1.39		11.56	6.44

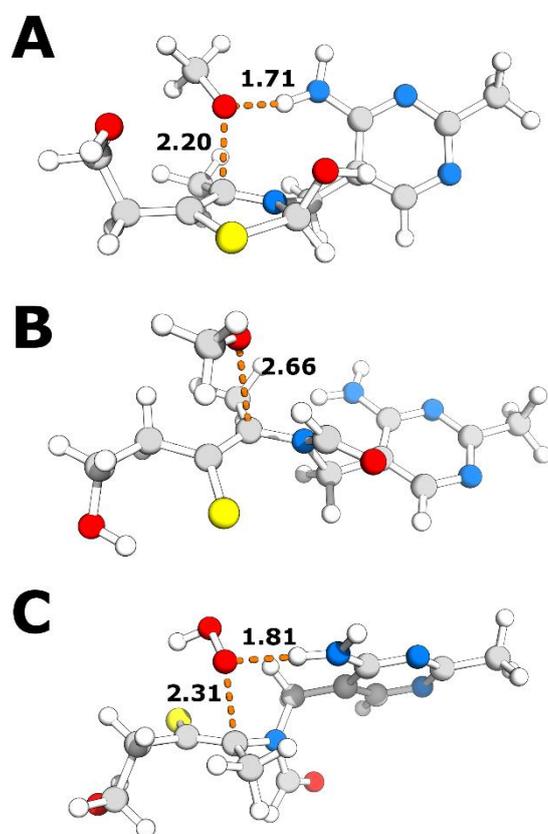
**Tabla 4.30.** Energía libre de Gibbs de activación  $\Delta G^\ddagger$  a 298.15 K (kcal/mol) de las reacciones RAF entre la tiamina y los radicales  $\bullet\text{OCH}_3$  y  $\bullet\text{OOH}$  en PE.

	$\bullet\text{OCH}_3$				$\bullet\text{OOH}$	
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}\cdot$	$\text{TOH}$	$\text{TO}\cdot$
<b>C4</b>	15.34	15.89				
<b>C10</b>	21.34	19.35	17.03			
<b>C15</b>	17.84	18.59	11.66	8.25	12.16	8.17
<b>C17</b>	13.89	15.03	6.55	6.98		3.97

De forma similar a las barreras de la reacción HAT, la polaridad del medio no cambia significativamente la magnitud de las barreras (**Tablas 4.29 y 4.30**). La menor barrera de energía libre de la reacción RAF corresponde a la del C17 para las formas TOH con  $\bullet\text{OCH}_3$  en disolución acuosa y la forma  $\text{TO}\cdot$  con  $\bullet\text{OCH}_3$  en medio lipídico (**Tablas 4.29 y 4.30**). Varias de estas reacciones muestran barreras del orden de pocas unidades de  $k_B T$ . La reacción RAF entre  $\bullet\text{OCH}_3$  y  $\text{TO}\cdot$  en C17 en medio acuoso presenta una barrera de sólo de  $2.4k_B T$ , la más baja de todas las calculadas aquí. En estos casos, las correcciones térmicas deberían calcularse a lo largo del camino de reacción de mínima energía para optimizar la posición de la barrera en el marco de la versión Variacional de la Teoría del Estado de Transición [Fernández-Ramos. 2007A]. Puesto que en cualquier caso estas barreras energéticas estarán dominadas por efectos térmicos y entrópicos, se asume que en este caso la reacción está limitada por la difusión como primera aproximación.

Las reacciones RAF en el C17 y en el C15 de las formas  $\text{TH}^{2+}$  y  $\text{T}^+$  muestran barreras de hasta  $\sim 14$  kcal/mol más grandes que las de las formas TOH y  $\text{TO}\cdot$ . La **Figura**

4.6 representa las estructuras de estado de transición de las reacciones RAF de TOH y  $\text{TO}^-$  en el C17.



**Figura 4.6.** Estructuras de los estados de transición (TS) de las reacciones RAF con las barreras más bajas para la eliminación de  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$ . A) Estado de transición de la reacción entre  $\bullet\text{OCH}_3$  y TOH en C17 en disolución acuosa. B) Estado de transición entre  $\bullet\text{OCH}_3$  y  $\text{TO}^-$  en C17 en disolución acuosa. C) Estado de transición de la reacción entre  $\bullet\text{OOH}$  y  $\text{TO}^-$  en C17 en PE. Las distancias relevantes se indican en Å y se muestran con líneas segmentadas. Figura creada con Open-source PyMOL software [Schrödinger, 2011]

Como se muestra, el grupo amino del anillo de la aminopirimidina puede participar en el estado de transición formando un enlace de hidrógeno con el oxígeno del radical. Sin embargo, esta característica también se ha observado en las estructuras de estado de transición de  $\text{TH}^{2+}$  y  $\text{T}^+$  y no puede explicar las diferencias en las barreras de energía libre. Por otro lado, las barreras para  $\text{TH}^{2+}$  y  $\text{T}^+$  en el C15 y el C17 son también similares a las calculadas para las posiciones C3, C10 y C4 (Tablas 4.29 y 4.30). Esto señala que la magnitud de las barreras de energía libre depende más de los cambios de la estructura electrónica de la tiamina entre los reactivos y el estado de transición. Específicamente, las barreras de  $\text{TH}^{2+}$  y  $\text{T}^+$  en el C17 y el C15 son mayores porque la

adición del radical rompe la aromaticidad del anillo de tiazol de una forma similar a la adición en C3 y C10, que hace perder la aromaticidad del anillo de la aminopirimidina.

### **Cálculo de las constantes cinéticas macroscópicas**

Los resultados indican que la tiamina experimenta reacciones exergónicas con los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$ , tanto en disolución acuosa como en medio lipídico a través de una variedad de vías de reacción, con energías libres de activación de entre  $\sim 1.4$  y  $\sim 27$  kcal/mol. Además, la tiamina muestra varios estados de protonación cuya concentración está modulada por el pH del medio (**Figura 4.4**). Dado que no siempre las formas de mayor concentración muestran las reacciones más exergónicas o más rápidas, es difícil proporcionar una imagen global que resuma y explique la actividad macroscópica de la tiamina. Galano y Álvarez-Idaboy han proporcionado un marco teórico unificado para estimar y comparar la actividad diferentes radicales [Galano. 2013], [Galano. 2014].

En primer lugar, se han calculado las constantes de velocidad microscópicas.  $k$ , de todas las reacciones exergónicas a partir de las correspondientes barreras de energía libre,  $\Delta G^\ddagger$ , considerando la Teoría del Estado de Transición (TST) [Fernández-Ramos, 2007A], [Eyring, 1935], [Evans, 1935]:

$$k = \sigma \frac{k_B T}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) \quad [4.10]$$

donde  $k_B$  y  $h$  son respectivamente las constantes de Boltzmann y Planck.  $R$  es la constante de los gases,  $T$  es la temperatura absoluta y  $\sigma$  es el número de simetría de la trayectoria de la reacción [Fernández-Ramos, 2007B].

Algunos de los valores de las constantes cinéticas calculados son superiores a  $10^8 \text{ M}^{-1}\text{s}^{-1}$ . Tales reacciones están dominadas por efectos entrópicos y pueden alcanzar el régimen de velocidad limitada por difusión. Este escenario es uno de los casos difíciles de la Teoría del Estado de Transición para los cuales las velocidades de reacción no se predicen correctamente. En los casos en que el valor de  $k$  esté próximo al límite de difusión,  $k_D$ , se calcula la constante aparente  $k_{app}$  que puede estimarse aplicando la aproximación en estado estacionario a la concentración del complejo reactivo [Evans, 1935].

$$k_{app} = \frac{k_D k}{k_{-D} + k} \quad [4.11]$$

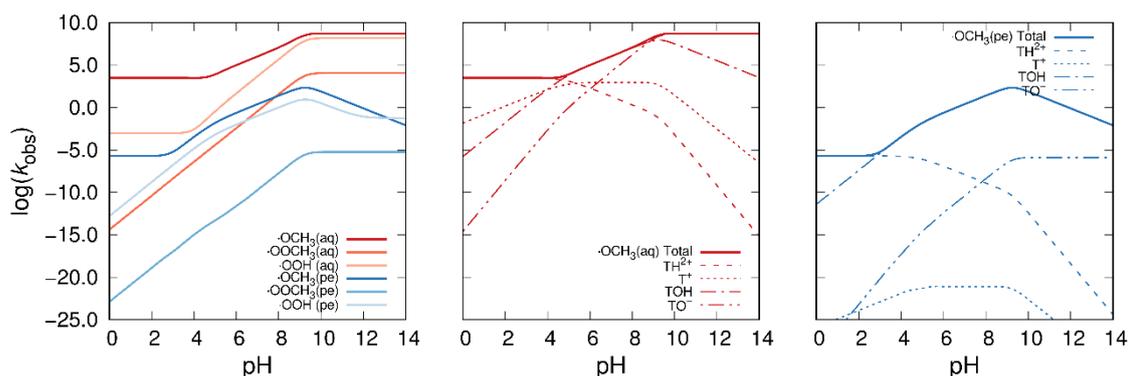
Donde  $k_{-D}$  es la velocidad de difusión de los reactivos desde el complejo de activación y se supone que es igual a  $k_D$ . Para cada radical y estado de protonación de la tiamina,  $k_D$  se ha aproximado mediante la consideración de las ecuaciones de Smoluchowski y Stokes-Einstein para la difusión [Galano. 2013].

Se han utilizado las constantes de velocidad microscópicas para calcular las constantes de velocidad macroscópicas,  $k_{obs}$ , por separado para cada radical  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  en disolución acuosa y en un entorno lipídico en función del pH como

$$k_{obs} = \sum_{ij} (k_i^{SET} + k_{ij}^{HAT} + k_{ij}^{RAF}) \chi_i \quad [4.12]$$

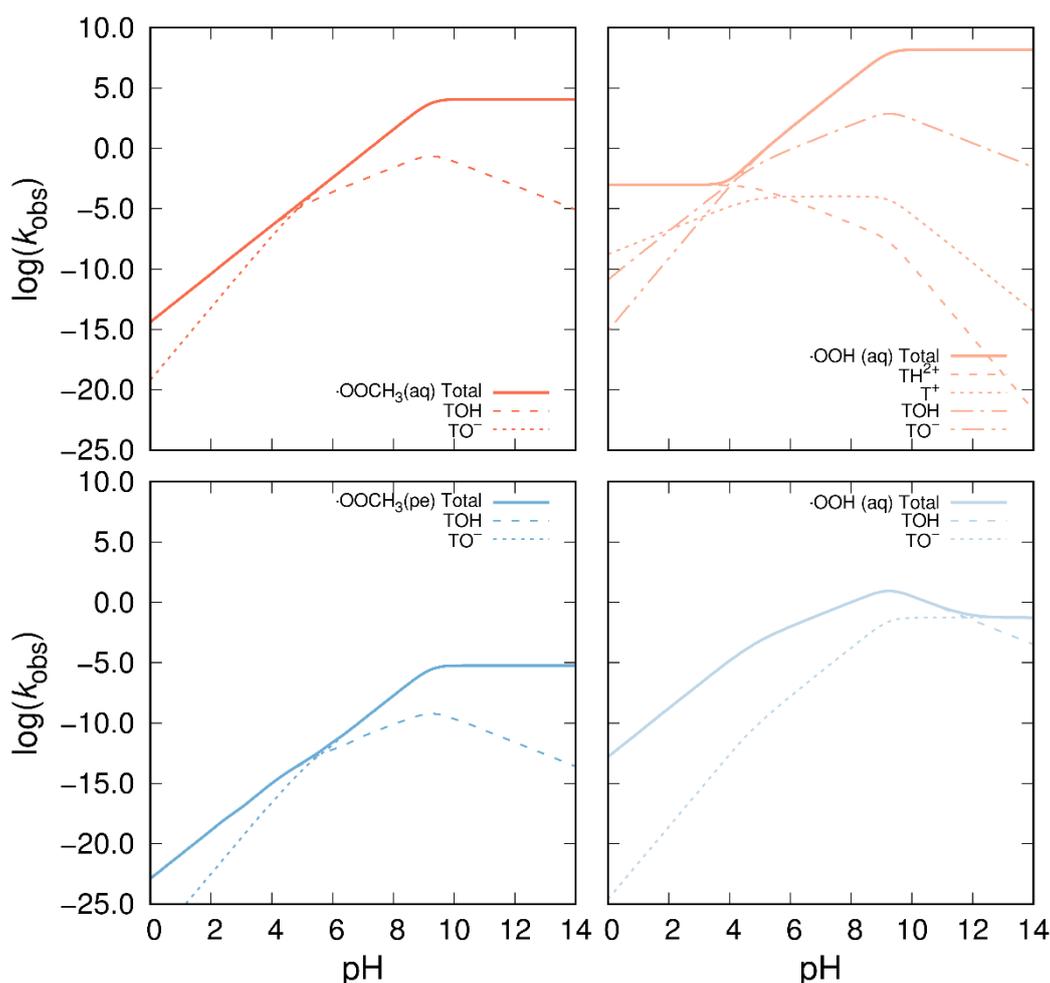
donde  $i, j$  representan el estado de protonación de la tiamina y el sitio de reacción.

En la **Figura 4.7** se han representado las constantes cinéticas macroscópicas,  $\log k_{obs}$ , de la eliminación de cada radical estudiado en disolución acuosa y en PE en función del pH. Las constantes cinéticas  $k_{obs}$  son mayores en disolución acuosa que en el medio lipídico, en parte debido a la diferente solubilidad de todas las formas de la tiamina, que es mayor en agua que en PE, pero también debido a que generalmente las barreras de energía libre en disolución acuosa son menores que en PE. Sin embargo, la actividad frente a los radicales  $\bullet\text{OCH}_3$  y  $\bullet\text{OOH}$  en medio lipídico es mayor que la actividad frente a  $\bullet\text{OOCH}_3$  en solución acuosa por debajo de  $\text{pH}=8$  (**Figura 4.7**). En ambos medios se observa la máxima reactividad frente a  $\bullet\text{OCH}_3$ , luego frente a  $\bullet\text{OOH}$  y finalmente frente a  $\bullet\text{OOCH}_3$ , siguiendo el orden de las barreras mínimas de energía libre obtenidas para las reacciones con cada una de ellas. Por último, la tendencia es que la actividad de eliminación es mayor a pH neutro/básico que a pH ácido (**Figura 4.7**).



**Figura 4.7.** (Izda.) Logaritmo de las constantes cinéticas macroscópicas de la eliminación de los radicales  $\bullet\text{OCH}_3$ ,  $\bullet\text{OOCH}_3$  y  $\bullet\text{OOH}$  en agua (líneas rojas) y en PE (líneas azules) en función del pH. (Centro) Logaritmo de las constantes cinéticas macroscópicas para la eliminación de  $\bullet\text{OCH}_3$  en agua deconvolucionada en las contribuciones de cada estado de protonación de la tiamina. (Dcha.) Logaritmo de las constantes cinéticas macroscópicas para la eliminación de  $\bullet\text{OCH}_3$  en PE deconvolucionada en las contribuciones de cada estado de protonación de la tiamina.

La **Figura 4.7** también muestra las contribuciones de cada estado de protonación de la tiamina a la reactividad total frente al radical  $\bullet\text{OCH}_3$  en disolución acuosa y entorno lipídico. En disolución acuosa por debajo de  $\text{pH} = 4$  la actividad macroscópica frente a  $\bullet\text{OCH}_3$  se debe principalmente a la forma  $\text{TH}^{2+}$  y a un pH superior a 10, la actividad macroscópica se puede atribuir principalmente a la forma  $\text{TO}^-$ . Entre pH 4 y 10, la mayor contribución a la actividad macroscópica se debe a la forma  $\text{TOH}$  y no a la forma  $\text{T}^+$ , que es la especie mayoritaria en dicho intervalo (**Figura 4.4**). El hecho es que, a pesar de que la fracción molar de  $\text{TOH}$  es de 1 a 5 órdenes de magnitud menor que la de  $\text{T}^+$ , la constante cinética total de la forma  $\text{TOH}$  es casi 7 órdenes de magnitud mayor. En los casos de  $\text{TOH}$  y  $\text{TO}^-$ , es posible atribuir la mayor parte de la actividad macroscópica a las reacciones de RAF en el C17, cuyas velocidades de reacción están casi limitadas por la difusión. En el medio lipídico, la forma ácida  $\text{TH}^{2+}$  sólo es relevante para la actividad macroscópica por debajo de  $\text{pH} = 3$ . Por encima, incluso a valores de pH muy básicos, la actividad macroscópica frente a  $\bullet\text{OCH}_3$  se debe principalmente a la forma  $\text{TOH}$ . Por un lado, la barrera de energía libre para la RAF en el C17 es significativamente menor que las barreras más bajas de las otras formas. De hecho, la constante cinética total de la especie  $\text{TOH}$  es de 4 a 6 órdenes de magnitud mayor que las de las otras formas. Por otro lado, la forma  $\text{TOH}$  presenta el mejor coeficiente de partición etanoato de pentilo/agua de todas las formas (es decir, los valores de  $\log P$  para  $\text{TH}^{2+}$ ,  $\text{T}^+$ ,  $\text{TOH}$  y  $\text{TO}^-$  son respectivamente -23, -9, -4 y -10) lo que la convierte en la forma más abundante en medio lipídico.



**Figura 4.8.** Logaritmo de las constantes cinéticas macroscópicas deconvolucionadas en las contribuciones de cada forma de la tiamina, en función del pH. (Arriba izda.) Actividad de eliminación frente a  $\bullet\text{OOCH}_3$  en agua. (Arriba dcha.) Actividad de eliminación frente a  $\bullet\text{OOH}$  en agua. (Abajo izda.) Actividad de eliminación frente a  $\bullet\text{OOCH}_3$  en PE. (Arriba dcha.) Actividad de eliminación frente a  $\bullet\text{OOH}$  en PE.

En otros trabajos se ha convertido en práctica habitual comparar las velocidades de secuestro de radicales con las del trolox (ácido 6-hidroxi-2.5.7.8-tetrametilcroman-2-carboxílico) para obtener una medida de la actividad antioxidante. La velocidad total para la eliminación de radicales aroxilo ( $\bullet\text{OR}$ ) por trolox en disolución acuosa con Triton X-100 (5.0% en peso) a 298.15 K es de  $5.01$ ,  $3.07$  y  $4.90 \cdot 10^4 \text{ M}^{-1}\text{s}^{-1}$  a pH 4.0, 7.0 y 12.0 respectivamente [Mitarai, 2008]. De acuerdo con nuestros resultados, la tiamina tendría una capacidad antioxidante reducida a pH bajo ( $k_{obs} = 3.1 \cdot 10^3 \text{ M}^{-1}\text{s}^{-1}$ ), pero a pH neutro y básico, la capacidad antioxidante de la tiamina sería respectivamente de 2 y 3 órdenes de magnitud mayor que el trolox ( $k_{obs} = 1.0 \cdot 10^6 \text{ M}^{-1}\text{s}^{-1}$  y  $k_{obs} = 5.1 \cdot 10^8 \text{ M}^{-1}\text{s}^{-1}$ ).

Por otro lado, el trolox muestra unas constantes cinéticas frente a radicales peroxilo de  $1.4, 4.1$  y  $1.5 \cdot 10^5 \text{ M}^{-1}\text{s}^{-1}$  respectivamente a pH 2.1, 7.4 y 12 [Amorati, 2016]. En este caso, la tiamina muestra una actividad de eliminación mucho menor a pH bajo ( $k_{obs} = 4.0 \cdot 10^{-11} \text{ M}^{-1}\text{s}^{-1}$  para  $\bullet\text{OOCH}_3$  y  $k_{obs} = 9.8 \cdot 10^{-4} \text{ M}^{-1}\text{s}^{-1}$  para  $\bullet\text{OOH}$ ). A pH neutro, la capacidad de la tiamina sigue siendo menor que la del trolox ( $k_{obs} = 3.7 \cdot 10^{-1} \text{ M}^{-1}\text{s}^{-1}$  frente a  $\bullet\text{OOCH}_3$  y  $k_{obs} = 4.9 \cdot 10^3 \text{ M}^{-1}\text{s}^{-1}$  frente a  $\bullet\text{OOH}$ ). A pH básico, la tiamina todavía no es tan eficaz como el trolox para  $\bullet\text{OOCH}_3$  ( $k_{obs} = 1.2 \cdot 10^4 \text{ M}^{-1}\text{s}^{-1}$ ), pero es mucho más reactiva frente a  $\bullet\text{OOH}$  ( $k_{obs} = 1.5 \cdot 10^8 \text{ M}^{-1}\text{s}^{-1}$ ).

## **5. Conclusiones**



Se ha comprobado la robustez de la reacción isodésmica en el cálculo de  $pK_a$  de compuestos pertenecientes a seis grupos funcionales comunes con valores de  $pK_a$ s comprendidos entre 1 y 19. Se han obtenido unos valores del  $pK_a$  con una precisión igual o mejor que la aportada con los ciclos termodinámicos empleando diversas combinaciones de métodos teóricos y modelos de disolvente y sin emplear moléculas de agua de forma explícita. La elección del ácido de referencia afecta a la precisión de los valores calculados, la cancelación de errores de la reacción isodésmica permite cierta flexibilidad en cuanto a la similitud estructural con el ácido de interés. Todo esto indica que este procedimiento tiene gran potencial para predicciones teóricas de  $pK_a$ , especialmente para los casos en que los ciclos termodinámicos muestran problemas relacionados con los cálculos en fase gaseosa de algunas especies.

Se ha mostrado que el esquema de la reacción isodésmica presenta ventajas significativas respecto a los ciclos termodinámicos para el cálculo de  $pK_a$ s de aminoácidos y péptidos, principalmente porque no requiere cálculos en fase gas. La reacción isodésmica proporciona resultados precisos para los  $pK_a$ s de los grupos  $\alpha$ -carboxílico,  $\alpha$ -amino y cadenas laterales de aminoácidos y péptidos, con medias de las desviaciones absolutas (MADs) de 1 unidad o inferior.

La precisión obtenida no es muy dependiente del método empleado. Los resultados para los diferentes funcionales de correlación intercambio de DFT empleados son muy similares entre sí. Además, la precisión obtenida con cálculos semiempíricos PM6 es similar a la de los cálculos DFT. Por último, la precisión obtenida con la reacción isodésmica es similar a la obtenida con métodos empíricos para la estimación de  $pK_a$ s. De forma que, a pesar de que los cálculos cuánticos son más costosos computacionalmente, no están limitados a la parametrización de los métodos empíricos.

La elección del ácido de referencia es importante para la precisión de los  $pK_a$ s calculados. Sin embargo, la cancelación de errores de la reacción isodésmica permite más flexibilidad en la elección de dichas especies. Como en el estudio mostrado anteriormente, confirmamos que es clave elegir un ácido de referencia para el que la distribución de carga vecina al grupo ácido es similar a la del grupo ácido del ácido estudiado. En muchos casos, esta condición se cumple eligiendo una molécula con el mismo grupo funcional.

Por otra parte, se ha empleado la teoría de DFT conceptual para estudiar los sitios preferentes de protonación y glicación de péptidos de lisina  $K(A)_nK$ . Se han probado los

funcionales de la familia Minnesota M11, M11L, MN12L, MN12SX, N12 y N12SX para el cumplimiento del procedimiento KID por comparación de los descriptores obtenidos con las energías HOMO y LUMO con los obtenidos con el procedimiento  $\Delta$ SCF. Se ha mostrado que el funcional MN12SX, híbrido meta-NGA de rango separado, y el funcional N12SX, híbrido NGA de rango separado, son los que cumplen mejor este objetivo. Como tal, son una buena alternativa a aquellos funcionales cuyo comportamiento se han puesto a punto por ajuste de la banda prohibida y son prometedores para la descripción de la reactividad química de sistemas moleculares de gran tamaño. De los resultados obtenidos, se observa que los sitios de preferentes de glicación y protonación de los péptidos estudiados se pueden predecir empleando descriptores de reactividad de DFT tales como el descriptor dual condensado y las funciones de Parr. Estos descriptores proporcionan una explicación sólida de la reactividad de estas moléculas. Además, los resultados presentados ayudan a entender el efecto de los espaciadores de alanina en propiedades como el  $pK_a$  de los residuos lisina. Como resultado, el  $pK_a$  de la lisina C-terminal será mayor que el de la N-terminal y la diferencia incrementará con el tamaño del espaciador de alanina.

En último lugar, se han considerado todas las posibles vías de reacción de la tiamina como secuestrador de ROS. La viabilidad termodinámica de cada vía de reacción se ha calculado, así como las constantes cinéticas de las reacciones termodinámicamente favorables con el fin de construir un modelo de la capacidad de eliminar los radicales. Los resultados indican que la tiamina elimina eficazmente el radical  $\bullet\text{OCH}_3$  independientemente del pH del medio, pero la actividad frente a los radicales peroxilo es fuertemente dependiente del pH. También se puede observar que la actividad macroscópica de la tiamina se debe a las especies con una fracción molar menor debido a sus excepcionales constantes de reacción microscópicas. Finalmente, se muestra que la variación de la deslocalización del electrón desapareado durante la reacción es el factor dominante en la determinación de la espontaneidad de la reacción y la magnitud de su barrera de energía libre, siempre y cuando no se afecte la aromaticidad de los anillos de la tiamina.

## **6. Bibliografía**



## A

### [Ahmed, 2002]

Ahmed, N., Argirov, O. K., Minhas H. S., Cordeiro, C. A., Thornalley, P. J., *Biochem. J.*, **2002**, 364, 1.

### [Ahmed, 2003]

Ahmed N., Thornalley P. J., *Biochem Soc Trans.*, **2003**, 31, 1417-22.

### [Ahmed, 2005]

Ahmed, N., Thornalley, P. J., *Ann. N.Y. Acad. Sci.*, **2005**, 1043, 260.

### [Albert, 1963]

Albert, A., Dipolar cycloaddition chemistry In: Katritzky AR 851 (ed) Physical methods in heterocyclic chemistry. vol 1. Academic Press. NY. **1963**

### [Al-Abet, 1997]

Al-Abet, Y., Bucala, R., *Chem. Res. Toxicol.*, **1997**, 10, 875.

### [Aldini, 2007]

Aldini, G., Dalle-donne, I., Facino, R. M., Milzani, A., Carini, M., *Med. Res. Rew.*, **2007**, 27, 817.

### [Amorati, 2016]

Amorati, R., Baschieri, A., Morroni, G., Gambino, R., Valgimigli, L., *Chem. Eur. J.*, **2016**, 22, 7924–7934.

### [Arnold, 2002]

Arnold, M. R., Kremer, W., Lüdemann, H-D., Kalbitzer, H. R., *Biophys. Chem.*, **2002**, 96, 129.

### [Avogadro, 2016]

Avogadro: An Open-Source Molecular Builder and Visualization Tool—Version 1.2.0. Available online: <http://avogadro.openmolecules.net> (accessed on 6 October 2016).

## B

### [Barone, 1998A]

Barone, V., Cossi, M., *J. Phys. Chem. A.*, **1998**, 102, 1995-2001.

### [Barone, 1998B]

Barone, V., Cossi, M., Tomassi, J., *J. Comput. Chem.*, **1998**, 19, 404-417.

### [Baynes, 1999]

Baynes, J. W., Thorpe, S. R., *Diabetes*, **1999**, 48, 1-9.

### [Becke, 1993]

Becke, A.D., *J. Chem. Phys.*, **1993**, 98, 5648–5652.

**[Behjatmanesh-Ardakani, 2009]**

Behjatmanesh-Ardakani, R., Karimi, M. A., Ebady, A., *J. Mol. Struct. Theochem*, **2009**, 910, 99-103.

**[Booth, 1997]**

Booth A. A., Khalifah R.G., Todd P, Hudson B.G., *J Biol Chem.*, **1997**, 272, 5430-7.

**[Brown, 2006]**

Brown, T. N., Mora-Diez, N., *J. Phys. Chem. B.*, **2006**, 110, 270-9279.

**[Brownlee, 1988]**

Brownlee M, Cerami A, Vlassara H., *N Engl J Med.*, **1988**, 318, 1315-21.

**[Brownlee, 1995]**

Brownlee M., *Annu Rev Med*, **1995**, 46, 223-234.

**[Brownlee, 2001]**

Brownlee, M., *Nature*, **2001**, 414, 813.

**[Bunn, 1979]**

Bunn, H.F., Shapiro, R., McManus, M., Garrick, L., McDonald, M. J., Gallop, P. M., Gabbay, K. H., *J. Biol. Chem.*, **1979**, 254, 3892–3898.

**C**

**[Caballero, 2000]**

Caballero, F., Gerez, E., Batlle, A., Vazquez, E. *Chem. Biol. Interact.*, **2000**, 126, 215-225.

**[Caballero, 2006]**

Caballero, N. A., Melendez, F. J., Muñoz-Caro, C., Niño. A., *Biophys. Chem.*, **2006**, 124, 155-160.

**[Caballero, 2011]**

Pérez-Caballero, G., Pérez-Arévalo, J. F., Morales-Hipólito, E. A., Arenas, M.E.C., Rojas-Hernández, A., *J. Mex. Chem. Soc.*, **2011**, 55, 126 – 131.

**[Camaioni, 2005]**

Camaioni, D. M., Schwerdtfeger, C. A., *J. Phys. Chem. A*, **2005**, 109, 10795-10797.

**[Casasnovas, 2009]**

Casasnovas, R., Frau, J., Ortega-Castro, J., Salvà, A., Donoso, J., Muñoz, F., *J. Mol. Struct. Theochem*, **2009**, 912, 5-12.

**[Casasnovas, 2010]**

Casasnovas, R., Frau, J., Ortega-Castro, J., Salvà, A., Donoso, J., Muñoz, F., *Int. J. Quantum Chem.*, **2010**, 110, 323-330.

**[Casasnovas, 2011]**

Casasnovas, R., Fernández, D., Ortega-Castro, J., Frau, J., Donoso, J., Muñoz, F., *Theor. Chem. Acc.*, **2011**, 130, 1-13.

**[Casasnovas, 2012]**

Casasnovas, R., Adrover, M., Ortega-Castro, J., Frau, J., Donoso, J., Muñoz, J., *Journal of Physical Chemistry B.*, **2012**, 116, 10665.

**[Casasnovas, 2013A]**

Casasnovas, R., Ortega-Castro, J., Donoso, J., Frau, J., Muñoz, F. *Phys. Chem. Chem. Phys.*, **2013**, 15, 16303.

**[Casasnovas, 2013B]**

Casasnovas, R., Frau, J., Ortega-Castro, J., Donoso, J., Muñoz, F., *J. Phys. Chem. B.*, **2013**, 117, 2339-2347.

**[Casasnovas, 2014]**

Casasnovas, R., Ortega-Castro, J., Frau, J., Donoso, J., Muñoz, F., *Int. J. Quantum Chem.*, **2014**, 114, 1350–1363.

**[Casey, 2006]**

Casey P. Kelly, Christopher J. Cramer, and Donald G. Truhlar. *J. Phys. Chem. A.*, **2006**, 110, 2493-2499.

**[Cramer, 1999]**

Cramer, C. J., Truhlar, D.G., *Chem. Rev.*, **1999**, 99, 2161-2200.

**[Cramer, 2008]**

Cramer, C. J., Truhlar, D. G., *Acc. Chem. Res.*, **2008**, 41, 760-768.

**[Curtiss, 1991]**

Curtiss, L. A., Raghavachari, K., Trucks, G. W., Pople, J. A., *J. Chem. Phys.*, **1991**, 94, 7221-7230.

**[Curtiss, 1998]**

Curtiss, L. A., K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, *J. Chem. Phys.*, **1998**, 109: 7764-7776.

## CH

**[Chamorro, 2013]**

Chamorro, E., Pérez, P., Domingo, L. R., *Chem. Phys. Lett.*, **2013**, 582, 141–143.

**[ChemOffice, 1998]**

ChemOffice, CambridgeSoft, Copyright 1998–2016 PerkinElmer Inc.

**[Chen, 2000]**

Chen, I. J., MacKerell, A. D., *Theo. Chem. Acc.*, **2000**, 103, 483-494.

**[Chetyrkin, 2008A]**

Chetyrkin, S., Zhang, W., Hudson, B. G., Serianni, A. S., Voziyan P. V., *Biochemistry*, **2008**, 47, 997.

**[Chetyrkin, 2008B]**

Chetyrkin, S.; Mathis, M. E., Ham, A-J., Hachey, D. L., Hudson, B. G., Voziyan, P. V., *Free Radic. Biol. Med.*, **2008**, 44, 1276.

**[Chetyrkin, 2011]**

Chetyrkin, S., Mathis, M., McDonald, W., Shackelford, X., Hudson, B., Voziyan, P. *Biochemical and Biophysical Research Communications*, **2011**, 411, 574–579.

**[Cho, 2007]**

Cho, S. J.; Roman, G.; Yeboah, F.; Konishi, Y., *Curr. Med. Chem.*, **2007**, 14, 1653.

**[Christen, 1988]**

Christen, H. R., F. Vögtle *Organische Chemie-Von den Grundlagen zur Forschung (Organic Chemistry-From Fundamentals to Research)* Otto Salle Verlag: Frankfurt, Germany, **1988**, Vol. 1, p 419.

**D**

**[da Silva, 2003]**

da Silva, E. F., Svendsen, H. F., *Ind. Eng. Chem. Res.*, **2003**, 42, 4414 - 4421.

**[da Silva, 2006]**

da Silva, G., Kennedy, E. M., Dlugogorski, B. Z., *J. Phys. Chem. A.*, **2006**, 110, 11371-11376.

**[Dean, 1999]**

J. A. Dean, *Lange's Handbook of Chemistry*, Mc Graw-Hill, inc., 14th ed., **1999**.

**[Dillet, 1994]**

Dillet, V., Rinaldi, D., Rivail, J-L., *J. Phys. Chem.*, **1994**, 98, 5034-5039.

**[Ding, 2009]**

Ding, F., Smith, J. M., Wang, H., *J. Org. Chem.*, **2009**, 74, 2679-2691.

**[Domingo, 2013]**

Domingo, L.R., Pérez, P., Sáez, J. *RSC Adv.*, **2013**, 3, 1486–1494.

**[Domingo, 2016]**

Domingo, L.R., *Molecules.*, **2016**, 21, 1–15.

**[Dong, 2008]**

Dong, H., Du, H., Qian, X., *J. Phys. Chem. A.*, **2008**, 112, 12687-12694.

**E**

**[Eckert, 2006]**

Eckert, F., Klamt, A., *J. Comput. Chem.*, **2006**, 27, 11-19.

**[Elgawish, 1996]**

Elgawish A., Glomb M., Friedlander M., Monnier V., *J. Biol. Chem.*, **1996**, 271, 22, 12964–12971.

**[Evans, 1935]**

Evans, M. G., Polanyi, M., *Trans. Faraday Soc.*, **1935**, 31, 875-894.

**[Eyring, 1935]**

H. Eyring, *J. Chem. Phys.*, **1935**, 3, 107–115.

**F**

**[Fan, 2009]**

Fan, X., Zhang, J., Theves, M., Strauch, C., Nemet, I., Liu, S., Qian, J., Giblin, F. J., Monnier, V. M. *J. Biol. Chem.*, **2009**, 284, 34618.

**[Fernandez-Ramos, 2007A]**

Fernandez-Ramos, B., Ellingson, A., Garrett, B.C., Truhlar, D.G., *Rev. Comp. Chem.* **2007**, 23, 125-232.

**[Fernandez-Ramos, 2007B]**

Fernandez-Ramos, B., Ellingson, A., Meana-Pañeda, R., Marques, J.M.C., Truhlar, D.G., *Theor. Chem. Acc.* **2007**, 118, 813-826

**[Fleming, 2011]**

Fleming, T. H.; Humpert, P. M.; Nawroth, P. P.; Bierhaus, A., *Gerontology*, **2011**, 57, 435.

**[Florence, 1989]**

Florence, T.M., Stanber, J.L., *Sci. Total Environ.*, **1989**, 78, 233- 240.

**[Finley, 1998]**

Finley, E., Dillon, L. J. et al., *Protein Sci.*, **1998**, 7, 11, 2391-7.

**[Frisch, 2004]**

Frisch M. J., G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A.

Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, revision C.02, Gaussian, Inc., Wallingford, CT, 2004.

**[Frisch, 2009]**

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, revision A. 02. Gaussian, Wallingford, CT, 2009.

**[Fu, 1996]**

Fu, M. X.; Requena, J. R.; Jenkins, A. J.; Lyons, T. J.; Baynes, J. W.; Thorpe, S. R.. *J. Biol. Chem.*, **1996**, *271*, 9982.

## G

**[Galano, 2011A]**

Galano A., Francisco-Márquez M., Alvarez-Idaboy J. R., *J. Phys. Chem. B.*, **2011**, *115*, 8590–8596.

**[Galano, 2011B]**

Galano, A., *Phys. Chem. Chem. Phys.*, **2011**, *13*, 7147–7157.

**[Galano, 2012]**

Galano, A., Martínez, A., *J. Phys. Chem.B.*, **2012**, *116*, 1200–1208.

**[Galano, 2013]**

Galano, A., Alvarez-Idaboy, J. R., *J. Comp. Chem.*, **2013**, *38*, 2430-2445.

**[Galano, 2014]**

Galano, A., Alvarez-Idaboy, J.R., *J. Comp. Chem.*, **2014**, *35*, 2019-2026.

**[Galano, 2016]**

Galano, A., Castañeda-Arriaga, R., Pérez-González, A., Tan, D. X., Reiter, R. J., *Molecules.*, **2016**, *21*, 1442.

**[Gao, 2005]**

Gao, D., Svoronos, P., Wong, P. K., Maddalena, D., Hwang, J., Walker, H., *J. Phys. Chem. A.*, **2005**, *109*, 10776-10785.

**[Gero, 1954]**

Gero, E., *Soc. Chim. Biol.*, **1954**, 36, 9, 1335.

**[Giannoukakis, 2005]**

Giannoukakis, N., *Curr. Opin. Investig. Drugs*, **2005**, 6, 410.

**[Giardino, 1998]**

Giardino, I., Fard, A. K., Hatchell, D. L., Brownlee, M., *Diabetes*, **1998**, 47, 1114.

**[Gorelsky, 2001]**

Gorelsky, S., Lever, A., *J. Organomet. Chem.*, **2001**, 635, 187–196.

**[Gorelsky, 2011]**

Gorelsky, S. AOMix Program for Molecular Orbital Analysis-Version 6.5, University of Ottawa: Ottawa, ON, Canada, **2011**.

**[Govender, 2009]**

Govender, K. K., Cukrowski, I., *J. Phys. Chem. A.*, **2009**, 113, 3639-3647.

**[Govender, 2010]**

Govender, K. K., Cukrowski, I., *J. Phys. Chem. A.*, **2010**, 114, 1868-1878.

**[Grillo, 2008]**

Grillo, M. A., Colombatto, S., *Amino Acids.*, **2008**, 35, 29.

**[Guiotto, 2005]**

Guiotto, A., Calderan, A., Ruzza, P., Borin, G., *Curr. Med. Chem.*, **2005**, 12, 2293.

**[Gupta, 2013]**

Gupta, M., da Silva, E. F., Svendsen, H. F., *J. Chem. Theory Comput.*, **2013**, 9, 5021.

## H

**[Halliwell, 1999]**

Halliwell B. and Gutteridge J. M. C. (1999) Free radicals in biology and medicine, 3rd ed. Oxford University Press, London.

**[Hanweel, 2012]**

Hanweel, M., Curtis, D.E., Lonie, D., Vandermeersch, T., Zurek, E., Hutchison, G., *J. Cheminf.*, **2012**, 4, 1–17.

**[Hipkins, 1997]**

Hipkiss, A. R., Preston, J. E., Himsworth, D. T., Worthington, V. C., Abbot, N. J., *Neurosci. Lett.*, **1997**, 238, 135.

**[Hirsch, 1995]**

Hirsch, J., Petrakova, E., Feather, M. S., Barnes, C. L., *Carbohydr. Res.*, **1995**, 267, 17.

**[Ho, 2009]**

Ho, J. M., Coote, M. L., *J. Chem. Theory Comput.*, **2009**, 5, 295-306.

**[Ho, 2010A]**

Ho, J. M., Coote, M. L., *Theor. Chem. Acc.*, **2010**, 125: 3-21.

**[Ho, 2010B]**

Ho, J., Klamt, A., Coote, M. L., *J. Phys. Chem. A.*, **2010**, 114, 13442–13444.

**[Ho, 2011]**

Ho, J., Coote, M. L., *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, **2011**, 1, 649–660.

**[Ho, 2015]**

Ho, J., *Phys. Chem. Chem. Phys.*, **2015**, 17, 2859–2868.

**I**

**[Ikeda, 1996]**

Ikeda, K.; Higashi, T.; Sano, H.; Jinnouchi, Y.; Yoshida, M.; Araki, T.; Ueda, S.; Horiuchi, S., *Biochemistry*, **1996**, 35, 8075.

**J**

**[Jakas, 2004]**

Jakas, H. S., *J. Pept. Sci.*, **2004**, 10, 119.

**[Jencks, 1970]**

Jencks, W. P., Regenstein, J., *Ionization Constants of Acids and Bases. Handbook of Biochemistry and Molecular Biology*, CRC Press, Cleveland, **1970**.

**[Jones, 1983]**

Jones, P. T., Anderson, R., *Int. Immunopharmacol.*, **1983**, 5, 377-389.

**[Jung, 2003]**

Jung, L., Kim, I. G., *Environ. Toxicol. Pharmacol.*, **2003**, 3, 19-26.

**K**

**[Kelly, 2006]**

Kelly, C. P., Cramer, C. J., Truhlar, D.G., *J. Phys Chem. B.*, **2006**, 110, 16066-16081.

**[Khalifah, 1999]**

Khalifah, R. G., Baynes, J. W., Hudson, B. G., *Biochem. Biophys. Res. Commun.*, **1999**, 257, 251.

**[Kiani, 2010]**

Kiani, F.; Rostami, A. A.; Sharifi, S.; Bahadori, A.; Chaichi, M. J. *J. Chem. Eng. Data* **2010**, 55, 2732.

**[Klamt, 1993]**

Klamt, A., Schüürmann, G., *J. Chem. Soc. Perkin Trans.*, **1993**, 2, 799-805.

**[Koskinen, 1997]**

Koskinen, M., Mutikainen, I., Tilus, T., Pelttari, E., Korvela, M., Elo, H., *Monatsh. Chem.*, **1997**, 128, 767.

**[Król, 2006]**

Król, M., Wrona, M., Page, C. S., Bates, P. A., *J. Chem. Theory Comput.*, **2006**, 2, 1520-1529.

**L**

**[Langlais, 1997]**

Langlais, P.J., Anderson, G., Guo, S. X., Bondy, S. C., *Metab. Brain Dis.*, **1997**, 12, 2, 137-143.

**[Lederer, 1999]**

Lederer, M. O., Bühler, H. P. *Bioorg. Med. Chem.*, **1999**, 7, 1081.

**[Lee, 1992]**

Lee A.T., Cerami A., *Ann NY Acad Sci.* **1992**, 663, 6370.

**[Lee, 1988]**

Lee, C., Yang, W., Parr, G., *Phys. Rev. B: Condens. Matter Mater. Phys.*, **1988**, 37, 785-789.

**[León-Carmona, 2011]**

Leon-Carmona, J. R., Galano., A., *J. Phys. Chem. B.*, **2011**, 115, 15, 4538-4546.

**[Lide, 2007]**

Lide. D.R. (ed). *CRC Handbook of chemistry and physics*. 87th edn. CRC Press Inc. Boca Raton. **2007**.

**[Liptak, 2001]**

Liptak, M. D., Shields, G. C., *J. Am. Chem. Soc.*, **2001**, 123, 7314-7319.

**[Liptak, 2002]**

Liptak, M. D., Gross, K. C., Seybold, P. G., Feldgus, S., Shields, G. C., *J. Am. Chem. Soc.*, **2002**, 124, 6421-6427.

**[Lo, 2006]**

Lo, C. Y., Li, S., Tan, D., Pan, M. H., Sang, S., Ho, C. T., *Mol. Nutr. Food Res.*, **2006**, 50, 1118.

**[Lukienko, 2000]**

Lukienko, P. I., Melnichenko, N. G., Zverinskii, I. V., Zabrodskaya, S. V., *Bull.Exp.Biol.Med.*, **2000**, 130, 9, 874-876.

## M

### [Maillard, 1912]

Maillard, L. C, *C. R. Acad. Sci. Ser.*, **1912**, 154, 66.

### [Makoowska, 2012]

Makowska, J., Liwo, A., Chmurzynski, L., Scheraga, H.A. *J. Solut. Chem.*, **2012**, 41, 1738–1746.

### [Marenich, 2009]

Marenich, A. V., Cramer, C. J., Truhlar, D. G., *J. Phys. Chem. B*, **2009**, 113, 6378–6396.

### [Maillard-Lefebvre, 2009]

Maillard-Lefebvre, H., Boulanger, E., Daroux, M.; Gaxatte, C., Hudson, B. I., Lambert, M., *Rheumatology*, **2009**, 48, 1190.

### [Marnett, 1987]

Marnett, L. J., *Carcinogenesis*, **1987**, 8, 1365-1373.

### [Martell, 2001]

Martell, A. E., Smith, R. M., Motekaitis, R. J., NIST Critically Selected Stability Constants of Metal Complexes. version 6.0; NIST Standard Reference Database 46; National Institute of Standards and Technology: Gaithersburg. MD. **2001**.

### [Martin, 1999]

Martin, J. M. L., de Oliveira, G. J., *Chem. Phys.*, **1999**, 111, 1843-1856.

### [Martínez-Araya, 2015]

Martínez-Araya, J.I., *J. Math. Chem.*, **2015**, 53, 451–465.

### [Martins, 2005]

Martins, S. I. F. S., van Boekel, M. A. J. S. A., *Food Chem.*, **2005**, 90, 257.

### [Mazzone, 2016]

Mazzone, G., Galano, A., Alvarez-Idaboy, J. R., Russo, N., *J. Chem. Inf. Model.*, **2016**, 56, 662-670.

### [McQuarrie, 1970]

McQuarrie, D.M. *Statistical Mechanics*, Harper and Row, New York, **1970**, p. 86.

### [Matxain, 2009]

Matxain J., Ristila M., Strid A. Eriksson L.A., *J. Phys. Chem. B.*, 2009, 113, 29, 9629–9632.

### [Mennucci, 1996]

Mennucci, B., Cossi, M., Tomasi, J., *J. Phys. Chem.*, **1996**, 100, 1807.

### [Miertus, 1981]

Miertus, S., Scrocco, E., Tomasi, J., *Chem, Phys.*, **1981**, 55, 117.

**[Mimna, 2007]**

Mimna, R., Camus, M. S., Schmid, A., Tuchscherer, G., Lashuel, H. A., Mutter, M., *Angew. Chem. Int. Ed.*, **2007**, 46, 2681.

**[Mitarai, 2008]**

Mitarai, A., Ouchi, A., Mukai, K., Tokunaga, A., Mukai, K., Abe, K., *J. Agric. Food Chem.* **2008**, 56, 84-91.

**[Miyata, 1994]**

Miyata T, Inagi R, Iida Y, Sato M, Yamada N, Oda O, et al., *J Clin Invest.* **1994**. 93:521-8.

**[Miyata, 2000]**

Miyata, T., *Bull. Mem. Acad. R. Med. Belg.*, **2002**, 157, 189.

**[Miyata, 2003]**

Miyata, T., van Ypersele de Strihou, C., *Arch. Biochem. Biophys.*, **2003**, 419, 50.

**[Montgomery, 1999]**

Montgomery, J. A., Frisch, M. J., Ochterki, J. W., Petersson, G. A., *J. Chem. Phys.*, **1999**, 110, 2822-2827.

**[Montgomery, 2000]**

Montgomery, J. A., Jr., Frisch, M. J., Ochterki, J. W., Petersson, G. A., *J. Chem. Phys.*, **2000**, 112, 6532-6542.

**[Morell, 2005]**

Morell, C., Grand, A., Toro-Labbé, A., *J. Phys. Chem. A.*, **2005**, 109, 205–212.

**[Morell, 2006]**

Morell, C., Grand, A., Toro-Labbé, A., *Chem. Phys. Lett.*, **2006**, 425, 342–346.

**[Mossine, 1999]**

Mossine V. V., Linetsky M., Glinsky G. V., Ortwerth B. J. And Feather M. S., *Chem. Res. Toxicol.*, **1999**, 12, 230–236.

**[Mullarkey, 1990]**

Mullarkey, C. J., Edelstein, D., Brownlee, M., *Biochem. Biophys. Res. Commun.* **1990**, 173, 932.

**[Munzel, 1996]**

Munzel, T., Kurz, S., Rajagopalan, S., Thoenes, M., Berrington, W. R., Thompson, J. A., Freeman, B. Harrison, D. G., *J. Clin. Invest.*, **1994**, 98, 1465.

N

**[Nakashima, 2010]**

Nakashima, T., Seki, T., Matsumoto, A., Miur, H., Sato, E., Niwano, Y., Kohno, M., Omura, S., and Takahashi, Y., *J. Biosci. Bioeng.*, **2010**, 110, 304-307

**[Nakashima, 2012]**

Nakashima, T., Ömura, S., Takahashi, Y., *J. Biosci. Bioeng.*, **2012**, 114, 275-280.

**[Namazian, 2003]**

Namazian, M., Heidary, H., *J. Molecular Structure-Theochem.*, **2003**, 620, 257-263.

**[Namazian, 2006]**

Namazian, M., Halvani, S., *J. Chem. Thermodynamics.*, **2006**, 38, 1495-1502.

**[Namazian, 2007]**

Namazian, M., Zakery, M., Noorbala, M. R., Coote, M. L., *Chem. Phys. Lett.*, **2007**, 45, 163-168.

**[Neely, 2000]**

Neely, M. D., Zimmerman, L., Picklo, M. J., Ou, J. J.; Morales, C. R., Montine, K. S., Amaranth, V., Montine, T. J., *Free Radic. Biol. Med.*, **2000**, 29, 1028.

**[Nemoto, 2012]**

Nemoto, S., Taguchi, K., Matsumoto, T., Kamata, K., Kobayashi, T., *Peptides.*, **2012**, 33, 109.

**[Nimbalkar, 2012]**

Nimbalkar V.V., *Biomed Aging Pathol.*, **2012**.

**[Nyden, 1981]**

Nyden, M. R., Petersson, G. A., *J. Chem. Phys.*, **1981**, 75, 1843-1862.

**O**

**[Ochterski, 1996]**

Ochterski, J. W., Petersson, G.A., Montgomery, J.A., Jr., *J. Chem. Phys.*, **1996**, 104, 2598-2619.

**[Okai, 2007]**

Okai, Y., Higashi-Okai, K., Sato, E. F., *J. Clin. Biochem. Nutr.*, **2007**, 40, 42-48.

**[Olivares del Valle, 1991]**

Olivares del Valle, F. J., Bonaccorsi, R., Cammi, R., Tomasi, J., *Theochem.*, **1991**, 76, 295.

**[OpenBabel, 2011]**

(a) N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminf.*, **2011**, 3, 33; (b) The Open Babel Package, version 2.3.1 <http://openbabel.org>.

## P

### [Peng, 2011]

Peng, X., Ma, J., Chen, F., Wang, M., *Food Funct.*, **2011**, 2, 289.

### [Pennathur, 2004]

Pennathur, S., Heinecke, J.W., *Front Biosci.*, **2004**, 9, 565-74.

### [Petersen, 2004]

Petersen, D. R., Doorn, J. A., *Free Radic. Biol. Med.*, **2004**, 37, 937.

### [Petersson, 1985]

Petersson, G. A., Yee, A. K., Bennett A., *J. Chem. Phys.*, **1985**, 83, 5105-5128.

### [Petersson, 1988]

Petersson, G. A., A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris., *J. Chem. Phys.*, **1988**, 89, 2193-2218.

### [Petersson, 1991]

Petersson, G. A., Al-Laham, M. A., *J. Chem. Phys.*, **1991**, 94, 6081-6090.

### [Petersson, 1994]

Petersson, G. A., Tenfeldt, T. G., Ochtanski, J. W., *J. Chem. Phys.*, **1994**, 101, 5900-5909.

### [Peverati, 2011]

Peverati, R., Truhlar, D.G., *J. Phys. Chem. Lett.*, **2011**, 2, 2810–2817.

### [Peverati, 2012A]

Peverati, R., Truhlar, D.G., *J. Phys. Chem. Lett.*, **2012**, 3, 117–124.

### [Peverati, 2012B]

Peverati, R., Truhlar, D.G., *Phys. Chem. Chem. Phys.*, **2012**, 14, 13171–13174.

### [Peverati, 2012C]

Peverati, R., Truhlar, D.G., *Phys. Chem. Chem. Phys.*, **2012**, 14, 16187–16191.

### [Peverati, 2012D]

Peverati, R., Truhlar, D.G., *J. Chem. Theory Comput.*, **2012**, 8, 2310–2319.

### [Peyroux, 2006]

Peyroux, J., Sternberg, M., *Pathol. Biol.*, **2006**, 54, 405.

### [Pliego, 2002]

Pliego, J. R., Riveros, J. M., *J. Phys. Chem. A.*, **2002**, 106, 7434-7439.

### [Pople, 1989]

Pople, J. A., Head-Gordon, M., Fox, D. J., Raghavachari, K., Curtiss, L.A., *J. Chem.*

*Phys.*, **1989**, 90, 5622-5629.

**[Price, 2001]**

Price, D. L., Rhett, P. M., Thorpe, S. R., Baynes, J. V., *J. Biol. Chem.*, **2001**, 276, 48967.

**[Pryor, 1986]**

Pryor, W. A., *Annu. Rev. Physiol.*, **1986**, 48, 657-667.

## **R**

**[Rabenstein, 1977]**

Rabenstein, D.L., Greenberg, M.S., Evans, C.A., *Biochemistry*, **1977**, 16, 977-981.

**[Rahbar, 2000]**

Rahbar, S., Natarajan, R., Yerneni, K., Scott, S., Gonzales, N., Nadler, J. L., *Clin. Chim. Acta*, **2000**, 301, 65.

**[Rahbar, 2003]**

Rahbar, S., Figarola, J. L., *Arch. Biochem. Biophys.*, **2003**, 419, 63.

**[Ramakrishnan, 1999]**

Ramakrishnan, S., Sulochana, K. N., Punitham, R., *Indian J. Biochem. Biophys.*, **1999**, 36, 129-133.

**[Ramasamy, 2012]**

Ramasamy, R.; Yan, S. F.; Schmidt, A. M., *Amino Acids*, **2012**, 42, 1151.

**[Remelli, 2000]**

Remelli, M., Conato, C., Agarossi, A., Pulidori, F., Mlynarz, P., Kozlowski, H., *Polyhedron.*, **2000**, 19, 2409-2419.

**[Rivail, 1976]**

Rivail, J. L., Rinaldi, D., *Chem. Phys.*, **1976**, 18, 233-242.

**[Rinaldi, 1983]**

Rinaldi, D., Ruiz-López, M. F., Rivail, J. L., *J. Chem. Phys.*, **1983**, 78, 834-838.

**[Ruggiero-López, 1999]**

Ruggiero-López, D., Lecomte, M., Moinet, G., Patereau, G., Lagarde, M., Wiernsperger, N., *Biochem. Pharmacol.*, **1999**, 58, 1765.

## **S**

**[Sastre, 2013]**

Sastre, S., Casasnovas, R., Muñoz, F., Frau, J., *Theor. Chem. Acc.*, **2013**, 132, 1310.

**[Sastre, 2016]**

Sastre, S., Casasnovas, R., Muñoz, F., Frau, J., *Phys Chem Chem Phys.*, **2016**, 18, 11202-11212.

**[Sato, 2006]**

Sato T, Shimogaito N, Wu X, Kikuchi S, Yamagishi S, Takeuchi M.. *Am J Alzheimers Dis Other Demen.*, **2006**, 21, 197-208.

**[Schmidt, 1992]**

Schmidt, A-M., Vianna, M., Gerlach, M., Brett, J., Ryan, J., Kao, J., Esposito, C., Hegarty, H., Hurley, W., Clauss, M., Wang, F., Pang, Y.E., Tsang, T. E., Stern, D., *J. Biol. Chem.*, **1992**, 267, 14987.

**[Schmidt, 1996]**

Schmidt A.M., Hori O., Cao R., Yan S.D., Brett J., Wautier J.L., Ogawa S., Kuwabara K., Matsumoto M., Stern D.. *Diabetes*, 1996, 45, S77-S80.

**[Schmidt, 2001]**

Schmidt, H., Andersson, I., Rehder, D., Pettersson, L., *Chemistry*, **2001**, 7, 251.

**[Schmidt am Busch, 2004]**

Schmidt am Busch, M., Knapp, E-W., *Chem. Phys. Chem*, **2004**, 5, 1513-1522.

**[Schrödinger, 2011]**

Schrödinger, 2011 The PyMOL molecular graphics system version 1.3. Schrödinger, LLC

**[Sell, 2005]**

Sell, D. R., Biemel, K. M., Reihl, O., Lederer, M. O., Strauch, C. M., Monnier, V. M. J., *Biol. Chem.*, **2005**, 280, 12310.

**[Seybold, 2015]**

Seybold, P. G., Shields, G. C., *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, **2015**, 5, 290–297.

**[Shangari, 2007]**

Shangari, N., Mehta, R., O'Brien, P. J., *Chem. Biol. Interact.*, **2007**, 165, 146-154.

**[Sheline, 2004]**

Sheline, C. T., Choi, D. W., *Ann Neurol*. **2004**, 55, 5, 645-53.

**[Silva, 2000]**

Silva, C. O., E. C. da Silva, and M. A. C. Nascimento., *J. Phys. Chem. A.*, **2000**, 104, 2402-2409.

**[Skog, 1998]**

Skog, K. I., Johansson, M. A., Jägerstad, M. I., *Food Chem. Toxicol.*, **1998**, 36, 879.

**[Smith, 1992]**

Smith, P. R., Thornalley, P. J., *Eur. J. Biochem.*, **1992**, 210, 729.

**[Stepuro, 1997]**

Stepuro, I.I., Piletskaya, T. P., Stepuro, V. I., Maskevich, S.D., *Biokhimiya*, **1997**, 62, 12, 1656-1662.

**[Stewart, 1985]**

Stewart, R. The Proton: Applications to Organic Chemistry, Wasserman. H. H., Ed., Vol. 46 of Organic Chemistry. A series of Monographs. Academic Press: New York. **1985**.

**[Stewart, 2007]**

Stewart, J. J. P., *J. Mol. Model.*, **2007**, 13, 1173–1213.

**[Sushko, 1981]**

Sushko, L.I., Lukienko, P.I., *Farmakol. Toksikol.*, **1981**, 2, 102-104.

**T**

**[Tanaka, 1997]**

Tanaka, Y., Iwamoto, H., Onuma, T., Kawamori. R., *Curr. Ther. Res.*, **1997**, 58, 693.

**[Tang, 2007]**

Tang, S. C., Arumugam, T. V., Cutler, R. G., Jo, D. G., Magnus, T., Chan, S. L., Mughal, M. R., Telljohann, R. S., Nassar, M., Ouyang, X., Calderan, A., Ruzza, P., Guiotto, A., Mattson, M. P., *J. Neurochem.*, **2007**, 101, 729.

**[Taniguchi, 1989]**

Taniguchi N., Kinoshita N., Arai K., Iizuka S., Usui M., Naito T., *Progr Clin Biol Res.*, **1989**, 304, 277-290.

**[[Thornalley, 1984]**

Thornalley, P. J.; Wolff, S. P.; Crabbe, J.; Stern, A., *Biochim. Biophys. Acta*, **1984**, 797, 276.

**[Thornalley, 1999]**

Thornalley, P.J., Langborg, A., Minhas, H. S., *Biochem. J.*, **1999**, 344, 109.

**[Thornalley, 2000]**

Thornalley, P. J., Yurek-George, A., Argirov, O. K., *Biochem. Pharmacol.*, **2000**, 60, 55.

**[Thornalley, 2003A]**

Thornalley P. J., Battah, S., Ahmed, N., Karachalias, N., Agalou, S., Babaei- Jadidi R., *Biochem J.*, **2003**. 375. 581-92.

**[Thornalley, 2003B]**

Thornalley, P. J., *Arch. Biochem. Biophys.*, **2003**, 419, 31.

**[Thornalley, 2005]**

Thornalley, P. J., *Ann. N.Y. Acad. Sci.*, **2005**, 1043, 111.

**[Thorpe, 1996]**

Thorpe, S. R., and Baynes, J. W. *Drugs & Aging*. **1996**, 9, 69–77.

**[Tissandier, 1998]**

Tissandier, M. D., Cowen, K. A., Feng, W. Y., Gundlach, E., Cohen, M. H., Earhart, A. D., Coe, J. V., *J. Phys. Chem. A.*, **1998**, 102, 7787-7794.

**[Tominaga, 1995]**

T. T. Tominaga, H. Imasato, O. R. Nascimento and M. Tabak, *Anal. Chim. Acta*, **1995**, 315, 217.

**[Toro-Labbé, 2007]**

Toro-Labbé, A. *Theoretical Aspects of Chemical Reactivity*, Elsevier Science: Amsterdam, The Netherlands, **2007**, 19.

**[Toth, 2001]**

Toth, A. M., Liptak, M. D., Phillips, D.L., Shields, G. C., *J. Chem. Phys.*, **2001**, 114, 4595-4606.

**[Trivelli, 1971]**

Trivelli, L. A., Ranney, H. M., Lai, H. T. *N. Engl. J. Med.*, **1971**, 284, 353.

**[Truong, 1995]**

Truong, T. N., Stefanovich, E. V., *Chem. Phys. Lett.*, **1995**, 240, 253-260.

**V**

**[Valentovik, 2004]**

Valentovic, M., Terneus, M., Harmon, R. C., Carpenter, A. B., *Toxicol. Lett.*, **2004**, 154, 165.

**[van Boekel, 1992]**

van Boekel, M. A., van der Bergh, P. J., Hoenders, H. J. *Biochim. Biophys. Acta*, **1992**, 1120, 201-204.

**[Vasan, 1996]**

Vasan, S., Zhang, X., Kapurniotu, A., Bernhagen, J., Teichberg, S., Basgen, J., Wagle, D., Shih, D., Terlecky, I., Bucala, R., Cerami, A., Egan, J., Ulrich, P., *Nature.*, **1996**, 382, 275.

**[Vasan, 2003]**

Vasan, S., Foiles, P., Founds, H., *Arch. Biochem. Biophys.*, **2003**, 419, 89.

**[Vindis, 2006]**

Vindis, C., Escargueil-Blanc, I., Elbaz, M., Marcheix, B., Grazide, M. H., Uchida, K., Salvayre, R., Negre-Salvayre, A., *Circ. Res.*, **2006**, 98, 785.

**[Vistoli, 2012]**

Vistoli, G., Straniero, V., Pedreti, A., Fumagalli, L., Bolchi, C., Pallavicini, M., Valoti, E., Testa, B., Chirality, **2012**, 24, 566–576.

**[Vlassara, 1996]**

Vlassara, H., Bucala, R., *Diabetes.*, **1996.**, 45 (Suppl. 3): S65-S66.

**W**

**[Wang, 2004]**

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A., *J. Comp. Chem.*, **2004**, 25, 1157–1174.

**[Wang, 2007]**

Wang, C., Liang J., Zhang C., Bi Y, Shi, X., Shi Q., *Ann Occup Hyg.* **2007**, 51, 563-569.

**[Wang, 2008]**

Wang, L., Heard, D. E., Pilling, M. J., Seakins, P., *J. Phys. Chem. A.*, **2008**, 112: 1832-1840.

**[Wautier, 2004]**

Wautier J. L., Schmidt, A. M., *Circ Res.*, **2004**, 95, 233-8.

**[Williams]**

Williams. R. pK<sub>a</sub> compilation.

[http://research.chem.psu.edu/brpgroup/pKa\\_compilation.pdf](http://research.chem.psu.edu/brpgroup/pKa_compilation.pdf)

**[Williams, 2004]**

Williams, M. E., *Curr. Diabetes Rep.*, **2004**, 4, 441.

**[Wolff, 1987]**

Wolff, S. P., Dean, R. T., *Biochem. J.*, **1987**, 245, 243.

**[Wolff, 1991]**

Wolff, S. P., Jiang, Z. Y., Hunt, J. V., *Free Radical Biol. Med.*, **1991**, 10, 339-52.

**[Wondrak, 2002]**

Wondrak, T. G., Cervantes-Laurean, D., Roberts, M. J., Qasem, J. G. Kim, M., Jacobson, E. L., Jacobson, M. K., *Biochem. Pharmacol.*, **2002**, 63, 361.

**[Wu, 2005]**

Wu, C. H., Yen, G. C., *Agric. Food Chem.*, **2005**, 53, 3167.

**[Wu, 2011]**

Wu, C.-H., Huang, S.-M., Lin, J.-A., Yen, G.-C., *Food Funct.*, **2011**, 2, 224.

**X**

**[Xu, 2011]**

Xu, Q., Park, Y., Huang, X., Hollenbeck, A., Blair, A., Schatzkin, A., Chen, H., *Diabetes Care*, **2011**, 34, 910.

## Y

### [Yu, 1997]

Yu, P. H., Zuo, D. M., *Diabetologia*, **1997**, *40*, 1243.

## Z

### [Zhang, 2011]

Zhang, Q., Monroe, M. E., Schepmoes, A. A., Clauss, T. R. W., Gritsenko, M. A., Meng, D., Petyuk, V. A., Smith, R. D., Metz, T. O., *J. Proteome Res.*, **2011**, *10*, 3076.

### [Zhao, 2006]

Zhao, Y., Schultz, N. E., Truhlar, D. G., *J. Chem. Theory Comput.*, **2006**, *2*, 364-382.

### [Zhao, 2008A]

Zhao, Y., Truhlar, D. G., *Theor. Chem. Acc.*, **2008**, *120*, 215–241.

### [Zhao, 2008B]

Zhao, Y., Truhlar, D. G., *J. Phys. Chem. A.*, **2008**, *112*, 1095–1099.

### [Zielinski, 2012]

Zielinski F., Tognetti V., Joubert L., *Chem. Phys. Lett.* **2012**, *527*, 67–72.



# 7. Apéndice



## Apéndice A. Resultados adicionales

Las referencias de esta sección están incluidas en el apartado 6. Bibliografía.

**Tabla A1.** Valores de  $pK_a$  experimentales y calculados de alcoholes alifáticos.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Etanol</b>	15.90						
<b>Metanol</b>	15.50	15.67	15.50	15.57	15.63	15.40	15.47
<b>2-cloroetanol</b>	14.30	13.40	13.11	13.08	13.51	13.25	13.18
<b>Propanol</b>	16.10	15.66	15.80	15.61	15.56	15.68	15.49
<b>i-propanol</b>	17.10	16.16	16.43	16.33	15.67	16.02	15.70
<b>2-butanol</b>	17.60	16.48	16.70	17.54	16.14	16.47	15.89
<b>tert-butanol</b>	19.20	17.66	18.15	18.06	16.36	16.93	16.35
<b>MAD</b>		0.85	0.68	0.62	1.20	1.01	1.29
<b>SD</b>		0.49	0.46	0.51	0.96	0.75	0.97
<b>AD max</b>		1.54	1.19	1.22	2.84	2.27	2.85

<sup>a</sup>[Stewart, 1985]

**Tabla A2.** Valores de  $pK_a$  experimentales y calculados de ácidos carboxílicos alifáticos.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Acético</b>	4.76						
<b>cloroacético</b>	2.86	1.91	1.63	1.84	1.94	1.36	0.97
<b>Fórmico</b>	3.75	3.19	2.76	3.14	2.62	2.21	2.46
<b>3-clorobutanoico</b>	4.05	4.20	3.44	3.98	3.90	3.74	3.40
<b>Benzoico</b>	4.20	4.00	4.62	4.00	3.37	3.98	3.35
<b>4-clorobutanoico</b>	4.52	5.10	5.09	5.02	4.57	4.52	4.38
<b>hexanoico</b>	4.82	5.93	6.08	5.95	5.11	5.24	5.04
<b>propanoico</b>	4.87	5.89	5.94	5.85	5.48	5.49	5.24
<b>pentanoico</b>	4.87	5.97	6.09	5.92	4.30	4.40	5.13
<b>trimetilacético</b>	5.05	6.37	6.49	6.56	5.62	5.72	5.42
<b>MAD</b>		0.78	0.98	0.79	0.57	0.64	0.67
<b>SD</b>		0.42	0.36	0.47	0.36	0.54	0.59
<b>AD máx</b>		1.32	1.44	1.51	1.13	1.54	1.89

<sup>a</sup>[Christen, 1998]

**Tabla A3.** Valores de  $pK_a$  experimentales y calculados de aminas alifáticas.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Etilamina</b>	10.63						
<b>Metilamina</b>	10.62	10.60	10.38	10.58	9.99	9.79	9.91
<b>propilamina</b>	10.53	10.93	10.95	10.96	10.87	10.91	10.72
<b>i-propilamina</b>	10.63	10.30	10.44	10.28	10.55	10.72	10.39
<b>Butilamina</b>	10.59	11.26	11.04	11.01	10.70	10.76	11.02
<b>2-butilamina</b>	10.56	10.40	10.67	10.55	10.47	10.74	10.65
<b>tert-butilamina</b>	10.55	10.02	10.42	10.10	10.70	11.11	10.77
<b>trimetilamina</b>	9.76	12.27	11.96	12.65	9.58	9.61	9.90
<b>dimetilamine</b>	10.64	12.23	11.99	12.39	10.30	10.20	10.48
MAD <sup>b</sup>		0.35	0.26	0.28	0.23	0.37	0.31
SD <sup>b</sup>		0.24	0.14	0.20	0.21	0.28	0.22
AD max <sup>b</sup>		0.67	0.45	0.45	0.63	0.83	0.71
MAD		0.87	0.72	0.90	0.24	0.35	0.27
SD		0.84	0.74	1.00	0.19	0.25	0.20
AD máx		2.51	2.20	2.89	0.63	0.83	0.71

<sup>a</sup>[Williams]<sup>b</sup>para aminas primarias**Tabla A4.** Valores de  $pK_a$  experimentales y calculados de fenoles.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Fenol</b>	9.98						
<b>p-cianofenol</b>	7.95	5.79	5.24	5.38	6.29	5.77	5.85
<b>o-clorofenol</b>	8.56	7.51	7.38	7.46	7.32	7.20	7.22
<b>m-cianofenol</b>	8.61	7.11	6.94	6.97	7.85	7.54	7.54
<b>m-clorofenol</b>	9.02	8.17	8.02	8.04	8.24	8.13	8.15
<b>m-fluorofenol</b>	9.28	8.34	8.14	8.07	8.41	8.25	8.17
<b>p-clorofenol</b>	9.38	8.77	8.70	8.68	8.91	8.87	8.86
<b>p-fluorofenol</b>	9.95	9.72	9.45	9.48	9.91	9.66	9.71
<b>m-metilfenol</b>	10.08	10.43	10.55	9.81	10.43	10.60	10.72
<b>p-metilfenol</b>	10.14	10.67	10.76	10.59	10.65	10.77	10.62
<b>o-metilfenol</b>	10.29	11.02	11.12	11.13	10.58	10.70	10.65
MAD		0.90	1.08	1.02	0.70	0.89	0.87
SD		0.57	0.68	0.68	0.48	0.57	0.56
AD max		2.16	2.71	2.57	1.66	2.18	2.10

<sup>a</sup>[Williams]

**Tabla A5.** Valores de  $pK_a$  experimentales y calculados de ácidos benzoicos.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>benzoico</b>	4.20						
<b>o-clorobenzoico</b>	2.93	15.67	15.50	15.57	15.63	15.40	15.47
<b>m-clorobenzoico</b>	3.83	13.40	13.11	13.08	13.51	13.25	13.18
<b>p-clorobenzoico</b>	3.99	15.66	15.80	15.61	15.56	15.68	15.49
<b>p-metilbenzoico</b>	4.37	16.16	16.43	16.33	15.67	16.02	15.70
<b>m-metilbenzoico</b>	4.27	16.48	16.70	17.54	16.14	16.47	15.89
<b>p-fluorobenzoico</b>	4.13	17.66	18.15	18.06	16.36	16.93	16.35
MAD		0.45	0.57	0.41	0.41	0.50	0.34
SD		0.29	0.41	0.30	0.35	0.48	0.29
AD máx		0.95	1.36	0.82	1.05	1.39	0.69

<sup>a</sup>[Martell, 2001]**Tabla A6.** Valores de  $pK_a$  experimentales y calculados de piridinas.

	Exptl	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Piridina</b>	5.23 <sup>a</sup>						
<b>2- metilpiridina</b>	5.97 <sup>b</sup>	6.77	6.96	6.90	6.45	6.65	5.78
<b>3-metilpiridina</b>	5.68 <sup>b</sup>	6.02	6.18	6.20	5.72	5.88	6.49
<b>4- metilpiridina</b>	6.02 <sup>b</sup>	6.23	6.44	6.89	6.24	6.42	6.15
<b>2,3-dimetilpiridina</b>	6.57 <sup>a</sup>	7.40	7.68	7.43	7.25	7.54	7.20
<b>2,4-dimetilpiridina</b>	6.99 <sup>a</sup>	8.07	8.44	7.73	7.78	7.54	7.73
<b>3-bromopiridina</b>	2.84 <sup>a</sup>	2.88	2.87	2.60	1.55	1.50	1.29
<b>3-fluoropiridina</b>	2.97 <sup>b</sup>	3.26	3.03	2.89	2.38	2.17	2.02
<b>3-cianopyridina</b>	1.45 <sup>c</sup>	0.01	-0.72	-0.42	0.78	0.51	0.41
<b>3-cloropiridina</b>	2.81 <sup>a</sup>	2.93	2.87	4.15	1.96	1.50	1.82
MAD		0.57	0.75	0.83	0.62	0.80	0.78
SD		0.48	0.74	0.55	0.36	0.39	0.44
AD max		1.44	2.17	1.87	1.29	1.34	1.55

<sup>a</sup>[Lide, 2007], <sup>b</sup>[Williams], <sup>c</sup>[Albert, 1963]

**Tabla A7.** Efecto del ácido de referencia en la media de las desviaciones absolutas (MADs) de  $pK_{as}$  calculados de alcoholes alfiáticos.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Etanol	0.85	0.68	0.62	1.20	1.01	1.29
Metanol	0.99	0.69	0.68	1.30	0.93	1.26
2-cloroetanol	0.55	0.71	0.82	0.89	0.66	0.89
Propanol	0.63	0.54	0.52	0.93	0.77	0.97
i-propanol	0.56	0.47	0.57	0.99	0.66	0.93
2-butanol	0.65	0.51	0.59	1.01	0.69	1.09
tert-butanol	1.00	0.59	0.76	2.16	1.64	2.04

**Tabla A8.** Efecto del ácido de referencia en la media de las desviaciones absolutas (MADs) de  $pK_{as}$  calculados de ácidos carboxílicos alfiáticos.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Acético	0.78	0.98	0.79	0.57	0.64	0.67
Cloroacético	1.45	1.72	1.50	0.84	1.42	1.70
Fórmico	1.11	1.51	1.13	1.02	1.45	1.17
3-clorobutanoico	0.74	1.25	0.80	0.57	0.66	0.79
Benzoico	0.86	0.89	0.86	0.78	0.64	0.88
4-clorobutanoico	0.74	0.89	0.79	0.58	0.64	0.67
Hexanoico	0.88	1.09	0.98	0.69	0.82	0.72
Propanoico	0.84	1.00	0.89	0.91	0.96	0.82
Pentanoico	0.88	1.06	0.92	0.66	0.73	0.74
Trimetilacético	1.07	1.25	1.31	0.87	1.01	0.81

**Tabla A9.** Efecto del ácido de referencia en la media de las desviaciones absolutas (MADs) de  $pK_{as}$  calculados de aminas alifáticas.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Etilamina	0.87	0.72	0.90	0.24	0.35	0.27
metilamina	0.61	0.93	0.80	0.61	0.93	0.80
propilamina	0.35	0.32	0.17	0.35	0.32	0.17
i-propilamina	0.23	0.34	0.38	0.23	0.34	0.38
butilamina	0.28	0.35	0.48	0.28	0.35	0.48
2-butilamina	0.21	0.35	0.22	0.21	0.35	0.22
tert-butilamina	0.31	0.63	0.30	0.31	0.63	0.30
trimetilamina	0.27	0.41	0.27	0.27	0.41	0.27
dimetilamina	0.36	0.58	0.33	0.36	0.58	0.33

**Tabla A10.** Efecto del ácido de referencia en la media de las desviaciones absolutas (MADs) de  $pK_a$ s calculados de fenoles.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Fenol	0.90	1.08	1.02	0.70	0.89	0.87
p-cianofenol	1.80	2.28	2.06	1.36	1.82	1.73
o-clorofenol	0.90	1.01	0.87	0.98	1.09	1.05
m-cianofenol	1.21	1.34	1.22	0.67	0.88	0.87
m-clorofenol	0.81	0.94	0.83	0.68	0.82	0.81
m-fluorofenol	0.84	0.98	0.93	0.72	0.86	0.89
p-clorofenol	0.79	0.91	0.80	0.64	0.78	0.77
p-fluorofenol	0.83	0.93	0.83	0.68	0.80	0.80
m-metilfenol	1.07	1.31	0.89	0.89	1.17	1.28
p-metilfenol	1.20	1.42	1.34	1.03	1.27	1.14
o-metilfenol	1.38	1.60	1.69	0.84	1.09	1.05

**Tabla A11.** Efecto del ácido de referencia en la media de las desviaciones absolutas (MADs) de  $pK_a$ s calculados de ácidos benzoicos.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
benzoico	0.47	0.46	0.44	0.41	0.50	0.34
o-clorobenzoico	1.41	1.71	1.42	1.42	1.68	1.17
m-clorobenzoico	2.06	2.01	1.84	2.12	2.05	1.94
p-clorobenzoico	1.12	1.19	1.30	1.21	1.21	1.12
p-metilbenzoico	1.12	1.22	1.15	1.12	1.15	1.07
m-metilbenzoico	1.21	1.27	1.18	1.15	1.19	1.16
p-fluorobenzoico	2.92	2.92	3.04	2.73	2.76	2.86

**Tabla A12.** Efecto del ácido de referencia en la media de las desviaciones absolutas (MADs) de  $pK_a$ s calculados de piridinas.

Referencia	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Piridina	0.57	0.75	0.83	0.62	0.80	0.78
2-metilpiridina	0.71	0.95	0.78	0.78	0.99	0.78
3-metilpiridina	0.50	0.74	0.71	0.62	0.82	1.17
4-metilpiridina	0.49	0.72	0.75	0.66	0.87	0.81
2,3-dimetilpiridina	0.73	1.03	0.74	0.92	1.19	1.03
2,4-dimetilpiridina	0.94	1.34	0.71	1.01	1.35	1.11
3-bromopiridina	0.55	0.73	0.97	1.30	1.43	1.45
3-fluoropiridina	0.49	0.72	0.87	0.76	1.00	0.95
3-cianopiridina	1.85	2.68	2.42	0.79	1.07	1.00
3-cloropiridina	0.51	0.72	1.15	0.91	1.06	0.97

**Tabla A15.**  $pK_a$ s calculados de fenoles empleando etanol como ácido de referencia.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Fenol	9.98	6.54	6.36	5.73	5.24	4.96	4.10
p-cianofenol	7.95	2.36	1.63	1.13	1.55	0.75	-0.03
o-clorofenol	8.56	4.07	3.76	3.21	2.58	2.18	1.34
m-cianofenol	8.61	3.67	3.33	2.73	3.11	2.52	1.66
m-clorofenol	9.02	4.73	4.40	3.79	3.50	3.11	2.27
m-fluorofenol	9.28	4.90	4.53	3.82	3.67	3.24	2.29
p-chlorofenol	9.38	5.33	5.09	4.44	4.17	3.86	2.98
p-fluorofenol	9.95	6.28	5.83	5.23	5.16	4.65	3.83
m-metilfenol	10.08	6.99	6.93	5.56	5.69	5.58	4.84
p-metilfenol	10.14	7.23	7.14	6.34	5.91	5.76	4.74
o-metilfenol	10.29	7.58	7.50	6.88	5.83	5.68	4.76
MAD		3.96	4.25	4.94	5.17	5.54	6.41
SD		0.89	1.07	0.96	0.70	0.88	0.86
AD max		5.59	6.32	6.82	6.40	7.20	7.98

<sup>a</sup>[Williams]**Tabla A16.**  $pK_a$ s calculados de ácidos benzoicos empleando ácido acético como ácido de referencia.

	Exptl <sup>a</sup>	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
benzoico	4.20	4.00	4.62	4.00	3.37	3.98	3.35
o-clorobenzoico	2.93	1.77	1.98	1.90	1.04	1.31	1.39
m-clorobenzoico	3.83	3.03	3.55	2.94	2.43	2.93	2.28
p-clorobenzoico	3.99	3.34	3.94	4.02	2.84	3.41	2.79
p-metilbenzoico	4.37	4.38	5.12	4.14	3.35	4.09	3.52
m-metilbenzoico	4.27	4.29	4.96	4.49	3.68	4.39	3.52
p-fluorobenzoico	4.13	3.67	4.23	3.65	3.18	3.71	3.07
MAD		0.47	0.46	0.44	1.12	0.59	1.12
SD		0.43	0.34	0.38	0.42	0.52	0.33
AD máx		1.15	0.94	1.03	1.88	1.62	1.55

<sup>a</sup>[Lide, 2007]

**Tabla A17.**  $pK_{a3}$ s calculados de piridinas empleando etilamina como ácido de referencia.

	Exptl	CPCM			SMD		
		CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
<b>Piridina</b>	5.23 <sup>a</sup>	7.73	6.73	4.14	5.21	4.08	4.14
<b>2- metilpiridina</b>	5.97 <sup>b</sup>	9.46	8.39	5.36	6.63	4.62	5.36
<b>3-metilpiridina</b>	5.68 <sup>b</sup>	8.68	7.69	4.63	5.86	5.33	4.63
<b>4- metilpiridina</b>	6.02 <sup>b</sup>	8.94	8.39	5.15	6.40	4.99	5.15
<b>2,3-dimetilpiridina</b>	6.57 <sup>a</sup>	10.18	8.93	6.16	7.52	6.05	6.16
<b>2,4-dimetilpiridina</b>	6.99 <sup>a</sup>	10.94	9.23	6.68	8.12	6.58	6.68
<b>3-bromopiridina</b>	2.84 <sup>a</sup>	5.37	4.10	0.45	1.48	0.14	0.45
<b>3-fluoropiridina</b>	2.97 <sup>b</sup>	5.53	4.38	1.29	2.15	0.87	1.29
<b>3-cianopyridina</b>	1.45 <sup>c</sup>	1.78	1.07	-0.31	0.49	-0.75	-0.31
<b>3-cloropiridina</b>	2.81 <sup>a</sup>	5.37	5.65	0.87	1.87	0.66	0.87
<b>MAD</b>		1.83	2.75	1.88	1.21	0.74	1.40
<b>SD</b>		0.69	0.99	0.73	0.70	0.43	0.85
<b>AD máx</b>		2.68	3.95	2.84	2.39	1.36	2.70

<sup>a</sup>[Lide, 2007], <sup>b</sup>[Williams], <sup>c</sup>[Albert, 1963]

**Tabla A18** Errores absolutos del  $pK_{a3}$  calculado de los péptidos con PM6/SMD usando únicamente el conformero de mínima energía como estructura inicial en la optimización geométrica.

	$pK_{a3}$ (exptl.)	$\Delta pK_{a3}$
<b>Lisina</b>		
Lys (ref.)	10.53 <sup>a</sup>	-
LysGlu	10.50 <sup>a</sup>	0.44
GlyLys	10.50 <sup>b</sup>	0.89
HisLys	10.49 <sup>b</sup>	0.88
LysAla	10.70 <sup>a</sup>	1.90
AlaLys	10.70 <sup>b</sup>	1.60
AlaLysAla	10.30 <sup>c</sup>	0.12
LysD-Ala	10.63 <sup>b</sup>	3.74
GlyHisLys	10.71 <sup>d</sup>	2.86
GlyGlyLysAla	11.10 <sup>e</sup>	5.48
<b>MAD</b>		1.99
<b>SD</b>		1.75
<b>Arginina</b>		
Arg (ref.)	12.47 <sup>a</sup>	-
PheArg	12.40 <sup>a</sup>	3.91
TyrArg	11.62 <sup>a</sup>	0.22
PheAlaArg	12.43 <sup>c</sup>	6.10
<b>MAD</b>		3.41
<b>SD</b>		2.97
<b>Histidina</b>		
His (ref.)	6.00 <sup>a</sup>	-
AlaHis	6.72 <sup>f</sup>	1.45
HisGly	5.80 <sup>a</sup>	3.85
HisLys	5.91 <sup>b</sup>	3.26

AspHis	6.82 <sup>a</sup>	3.93
GlyHis	6.77 <sup>a</sup>	1.80
GlyHisLys	6.60 <sup>d</sup>	1.73
GlyHisGly	6.62 <sup>d</sup>	2.74
GlyGlyHisAla	7.00 <sup>e</sup>	1.74
TyrHisOMe	6.41 <sup>f</sup>	1.28
GluHisOMe	6.44 <sup>f</sup>	4.29
<b>MAD</b>		2.61
<b>SD</b>		1.15
<b>Cisteina</b>		
Cys (ref.)	8.18 <sup>a</sup>	-
CysAsn	7.09 <sup>a</sup>	2.57
CysGlyGly	6.36 <sup>c</sup>	0.35
<b>MAD</b>		1.46
<b>SD</b>		1.57
<b>Tirosina</b>		
Tyr (ref.)	10.60 <sup>a</sup>	-
GlyTyr	10.49 <sup>a</sup>	0.74
TyrArg	9.36 <sup>a</sup>	2.20
LeuTyr	10.09 <sup>a</sup>	0.97
TyrGly	10.51 <sup>g</sup>	1.64
GlyGlyTyrAla	10.30 <sup>e</sup>	2.78
D-LeuTyr	10.35 <sup>a</sup>	0.85
TyrHisOMe	9.69 <sup>f</sup>	3.62
<b>MAD</b>		1.83
<b>SD</b>		1.09
<b>Glu/Asp</b>		
Glu (ref.)	4.25 <sup>a</sup>	-
Asp	3.65 <sup>a</sup>	0.60
LysGlu	4.47 <sup>a</sup>	1.09
GlyGlyGluAla	4.30 <sup>e</sup>	0.50
GlyAsp	4.45 <sup>a</sup>	3.62
GlyGlyAspAla	3.90 <sup>e</sup>	1.25
GluHisOMe	3.79 <sup>f</sup>	3.08
AspGly	4.53 <sup>c</sup>	0.70
<b>MAD</b>		1.55
<b>SD</b>		1.27
<b>Todos</b>		
<b>MAD</b>		<b>2.13</b>
<b>SD</b>		<b>1.51</b>

<sup>a</sup>[Dean, 1999], <sup>b</sup>[Remelli, 2000], <sup>c</sup>[Jencks, 1970], <sup>d</sup>[Rabenstein, 1997], <sup>e</sup>[Arnold, 2002], <sup>f</sup>[Vistoli, 2012], <sup>g</sup>[Tominaga, 1995]

**Tabla A19.** Energías de los orbitales HOMO y LUMO (eV), potenciales de ionización (I), afinidades electrónicas (A), electronegatividad global ( $\chi$ ), dureza ( $\eta$ ) y electrofilicidad global ( $\omega$ ) calculadas con el funcional M11

	HOMO	LUMO	$\chi_k$	$\eta_k$	$\omega_k$
<b>KK</b>	-9.450	1.922	3.764	11.372	0.623
<b>KAK</b>	-9.388	2.158	3.615	11.546	0.566
<b>KA2K</b>	-9.549	2.089	3.730	11.638	0.598
<b>KA3K</b>	-9.422	2.075	3.674	11.497	0.587
<b>KA4K</b>	-9.557	2.077	3.740	11.634	0.601
<b>KA5K</b>	-9.554	2.078	3.738	11.632	0.601
	<b>I</b>	<b>A</b>	<b>X</b>	<b><math>\eta</math></b>	<b><math>\omega</math></b>
<b>KK</b>	6.708	0.528	3.618	6.180	1.059
<b>KAK</b>	6.710	0.319	3.515	6.391	0.966
<b>KA2K</b>	6.695	0.270	3.483	6.425	0.944
<b>KA3K</b>	6.696	0.266	3.481	6.430	0.942
<b>KA4K</b>	6.702	0.248	3.475	6.454	0.936
<b>KA5K</b>	6.700	0.245	3.473	6.455	0.934

**Tabla A20:** Descriptores para los péptidos KK. KAK. KA2K. KA3K. KA4K y KA5K calculados con los resultados de la Tabla A19

	<b>J<sub>x</sub></b>	<b>J<sub><math>\eta</math></sub></b>	<b>J<sub><math>\omega</math></sub></b>	<b>J<sub>DI</sub></b>	<b>J<sub>I</sub></b>	<b>J<sub>A</sub></b>	<b>J<sub>HL</sub></b>
<b>KK</b>	0.15	5.19	0.44	5.21	2.74	2.45	3.68
<b>KAK</b>	0.10	5.16	0.40	5.17	2.68	2.48	3.65
<b>KA2K</b>	0.25	5.21	0.35	5.23	2.85	2.36	3.70
<b>KA3K</b>	0.19	5.07	0.36	5.08	2.73	2.34	3.59
<b>KA4K</b>	0.27	5.18	0.33	5.20	2.86	2.33	3.68
<b>KA5K</b>	0.27	5.18	0.33	5.19	2.85	2.32	3.68
<b>Media</b>	0.20	5.16	0.37	5.18	2.78	2.38	3.66

**Tabla A21:** Energías de los orbitales HOMO y LUMO (eV), potenciales de ionización (I), afinidades electrónicas (A) electronegatividad global ( $\chi$ ), dureza ( $\eta$ ) y electrofilicidad global ( $\omega$ ) calculadas con el funcional M11L

	HOMO	LUMO	$\chi_k$	$\eta_k$	$\omega_k$
<b>KK</b>	-6.135	-0.726	3.431	5.409	1.088
<b>KAK</b>	-6.135	-0.761	3.448	5.374	1.106
<b>KA2K</b>	-6.123	-0.811	3.467	5.312	1.131
<b>KA3K</b>	-6.126	-0.820	3.473	5.306	1.137
<b>KA4K</b>	-6.123	-0.828	3.476	5.295	1.141
<b>KA5K</b>	-6.123	-0.838	3.481	5.285	1.146
	I	A	X	$\eta$	$\omega$
<b>KK</b>	6.510	0.567	3.539	5.943	1.053
<b>KAK</b>	8.616	0.709	4.663	7.907	1.375
<b>KA2K</b>	8.240	0.770	4.505	7.470	1.358
<b>KA3K</b>	6.988	0.777	3.883	6.211	1.213
<b>KA4K</b>	8.203	0.782	4.493	7.421	1.360
<b>KA5K</b>	8.203	0.788	4.496	7.415	1.363

**Tabla A22:** Descriptores para los péptidos KK. KAK. KA2K. KA3K. KA4K y KA5K calculados con los resultados de la Tabla A21

	$J_x$	$J_\eta$	$J_\omega$	$J_{DI}$	$J_I$	$J_A$	$J_{HL}$
<b>KK</b>	0.11	0.53	0.03	0.55	0.38	0.16	0.41
<b>KAK</b>	1.21	2.53	0.27	2.82	2.48	0.05	2.48
<b>KA2K</b>	1.04	2.16	0.23	2.41	2.12	0.04	2.12
<b>KA3K</b>	0.41	0.91	0.08	1.00	0.86	0.04	0.86
<b>KA4K</b>	1.02	2.13	0.22	2.37	2.08	0.05	2.08
<b>KA5K</b>	1.02	2.13	0.22	2.37	2.08	0.05	2.08
<b>Media</b>	0.80	1.73	0.17	1.92	1.67	0.07	1.67

**Tabla A23:** Energías de los orbitales HOMO y LUMO (eV), potenciales de ionización (I), afinidades electrónicas (A), electronegatividad global ( $\chi$ ), dureza ( $\eta$ ) y electrofilicidad global ( $\omega$ ) calculadas con el funcional MN12L

	<b>HOMO</b>	<b>LUMO</b>	$\chi_k$	$\eta_k$	$\omega_k$
<b>KK</b>	-5.994	-0.486	3.240	5.508	0.953
<b>KAK</b>	-5.999	-0.335	3.167	5.664	0.885
<b>KA2K</b>	-6.030	-0.359	3.195	5.671	0.900
<b>KA3K</b>	-5.995	-0.357	3.176	5.638	0.895
<b>KA4K</b>	-6.029	-0.376	3.203	5.653	0.907
<b>KA5K</b>	-6.034	-0.369	3.202	5.665	0.905
	<b>I</b>	<b>A</b>	<b>X</b>	<b><math>\eta</math></b>	<b><math>\omega</math></b>
<b>KK</b>	8.294	0.280	4.287	8.014	1.147
<b>KAK</b>	7.557	0.185	3.871	7.372	1.016
<b>KA2K</b>	8.375	0.222	4.299	8.153	1.133
<b>KA3K</b>	8.280	0.231	4.256	8.049	1.125
<b>KA4K</b>	6.355	0.262	3.309	6.093	0.898
<b>KA5K</b>	8.448	0.261	4.355	8.187	1.158

**Tabla A24:** Descriptores para los péptidos KK. KAK. KA2K. KA3K. KA4K y KA5K calculados con los resultados de la Tabla A23

	<b>J<sub>x</sub></b>	<b>J<sub><math>\eta</math></sub></b>	<b>J<sub><math>\omega</math></sub></b>	<b>J<sub>DI</sub></b>	<b>J<sub>I</sub></b>	<b>J<sub>A</sub></b>	<b>J<sub>HL</sub></b>
<b>KK</b>	1.05	2.51	0.19	2.72	2.30	0.21	2.31
<b>KAK</b>	0.70	1.71	0.13	1.85	1.56	0.15	1.57
<b>KA2K</b>	1.10	2.48	0.23	2.73	2.35	0.14	2.35
<b>KA3K</b>	1.08	2.41	0.23	2.65	2.29	0.13	2.29
<b>KA4K</b>	0.11	0.44	0.01	0.45	0.33	0.11	0.35
<b>KA5K</b>	1.15	2.52	0.25	2.78	2.41	0.11	2.42
<b>Media</b>	0.87	2.01	0.18	2.20	1.87	0.14	1.88

**Tabla A25:** Energías de los orbitales HOMO y LUMO (eV), potenciales de ionización (I), afinidades electrónicas (A), electronegatividad global ( $\chi$ ), dureza ( $\eta$ ) y electrofilicidad global ( $\omega$ ) calculadas con el funcional MN12SX

	<b>HOMO</b>	<b>LUMO</b>	$\chi_k$	$\eta_k$	$\omega_k$
<b>KK</b>	-6.727	-0.291	3.509	6.436	0.957
<b>KAK</b>	-6.773	-0.389	3.581	6.384	1.004
<b>KA2K</b>	-6.782	-0.436	3.609	6.346	1.026
<b>KA3K</b>	-6.729	-0.434	3.582	6.295	1.019
<b>KA4K</b>	-6.783	-0.444	3.614	6.339	1.030
<b>KA5K</b>	-6.785	-0.454	3.620	6.331	1.035
	<b>I</b>	<b>A</b>	<b>X</b>	<b><math>\eta</math></b>	<b><math>\omega</math></b>
<b>KK</b>	6.622	0.355	3.489	6.267	0.971
<b>KAK</b>	6.628	0.338	3.483	6.290	0.964
<b>KA2K</b>	6.636	0.347	3.492	6.289	0.969
<b>KA3K</b>	6.624	0.333	3.479	6.291	0.962
<b>KA4K</b>	6.637	0.354	3.496	6.283	0.972
<b>KA5K</b>	6.639	0.354	3.497	6.285	0.973

**Tabla A26:** Descriptores para los péptidos KK, KAK, KA2K, KA3K, KA4K y KA5K calculados con los resultados de la Tabla A25

	<b>J<sub>x</sub></b>	<b>J<sub><math>\eta</math></sub></b>	<b>J<sub><math>\omega</math></sub></b>	<b>J<sub>DI</sub></b>	<b>J<sub>I</sub></b>	<b>J<sub>A</sub></b>	<b>J<sub>HL</sub></b>
<b>KK</b>	0.02	0.17	0.01	0.17	0.11	0.06	0.12
<b>KAK</b>	0.10	0.09	0.04	0.14	0.15	0.05	0.15
<b>KA2K</b>	0.12	0.06	0.06	0.14	0.15	0.09	0.17
<b>KA3K</b>	0.10	0.00	0.06	0.12	0.11	0.10	0.15
<b>KA4K</b>	0.12	0.06	0.06	0.14	0.15	0.09	0.17
<b>KA5K</b>	0.12	0.05	0.06	0.15	0.15	0.10	0.18
<b>Media</b>	0.10	0.07	0.05	0.14	0.13	0.08	0.16

**Tabla A27:** Energías de los orbitales HOMO y LUMO (eV), potenciales de ionización (I), afinidades electrónicas (A), electronegatividad global ( $\chi$ ), dureza ( $\eta$ ) y electrofilicidad global ( $\omega$ ) calculadas con el funcional M12

	<b>HOMO</b>	<b>LUMO</b>	$\chi_k$	$\eta_k$	$\omega_k$
<b>KK</b>	-5.287	-0.775	3.031	4.512	1.018
<b>KAK</b>	-5.291	-0.731	3.011	4.560	0.994
<b>KA2K</b>	-5.298	-0.766	3.032	4.532	1.014
<b>KA3K</b>	-5.289	-0.753	3.021	4.536	1.006
<b>KA4K</b>	-5.279	-0.771	3.025	4.508	1.015
<b>KA5K</b>	-5.298	-0.771	3.035	4.527	1.017
	<b>I</b>	<b>A</b>	<b>X</b>	<b><math>\eta</math></b>	<b><math>\omega</math></b>
<b>KK</b>	5.675	0.404	3.040	5.271	0.876
<b>KAK</b>	6.673	0.473	3.573	6.200	1.030
<b>KA2K</b>	7.166	0.577	3.872	6.589	1.137
<b>KA3K</b>	5.955	0.575	3.265	5.380	0.991
<b>KA4K</b>	7.323	0.611	3.967	6.712	1.172
<b>KA5K</b>	7.368	0.622	3.995	6.746	1.183

**Tabla A28:** Descriptores para los péptidos KK. KAK. KA2K. KA3K. KA4K y KA5K calculados con los resultados de la Tabla A27

	<b>J<sub>x</sub></b>	<b>J<sub><math>\eta</math></sub></b>	<b>J<sub><math>\omega</math></sub></b>	<b>J<sub>DI</sub></b>	<b>J<sub>I</sub></b>	<b>J<sub>A</sub></b>	<b>J<sub>HL</sub></b>
<b>KK</b>	0.01	0.76	0.14	0.77	0.39	0.37	0.54
<b>KAK</b>	0.56	1.64	0.04	1.73	1.38	0.26	1.41
<b>KA2K</b>	0.84	2.06	0.12	2.23	1.87	0.19	1.88
<b>KA3K</b>	0.24	0.84	0.02	0.88	0.67	0.18	0.69
<b>KA4K</b>	0.94	2.20	0.16	2.40	2.04	0.16	2.05
<b>KA5K</b>	0.96	2.22	0.17	2.42	2.07	0.15	2.08
<b>Media</b>	0.59	1.62	0.11	1.74	1.40	0.22	1.44

**Tabla A29:** Energías de los orbitales HOMO y LUMO (eV), potenciales de ionización (I), afinidades electrónicas (A), electronegatividad global ( $\chi$ ), dureza ( $\eta$ ) y electrofilicidad global ( $\omega$ ) calculadas con el funcional N12SX

	<b>HOMO</b>	<b>LUMO</b>	$\chi_k$	$\eta_k$	$\omega_k$
<b>KK</b>	-6.361	-0.201	3.281	6.160	0.874
<b>KAK</b>	-6.365	-0.185	3.275	6.180	0.868
<b>KA2K</b>	-6.394	-0.182	3.288	6.212	0.870
<b>KA3K</b>	-6.362	-0.188	3.275	6.174	0.869
<b>KA4K</b>	-6.389	-0.198	3.294	6.191	0.876
<b>KA5K</b>	-6.393	-0.177	3.285	6.216	0.868
	<b>I</b>	<b>A</b>	<b>X</b>	<b><math>\eta</math></b>	<b><math>\omega</math></b>
	6.359	0.280	3.320	6.079	0.906
<b>KAK</b>	6.364	0.258	3.311	6.106	0.898
<b>KA2K</b>	6.395	0.245	3.320	6.150	0.896
<b>KA3K</b>	6.361	0.246	3.304	6.115	0.892
<b>KA4K</b>	6.383	0.241	3.312	6.142	0.893
<b>KA5K</b>	6.387	0.241	3.314	6.146	0.893

**Tabla A30:** Descriptores para los péptidos KK, KAK, KA2K, KA3K, KA4K y KA5K calculados con los resultados de la Tabla A29

	<b>J<sub>x</sub></b>	<b>J<sub><math>\eta</math></sub></b>	<b>J<sub><math>\omega</math></sub></b>	<b>J<sub>DI</sub></b>	<b>J<sub>I</sub></b>	<b>J<sub>A</sub></b>	<b>J<sub>HL</sub></b>
<b>KK</b>	0.04	0.08	0.03	0.10	0.00	0.08	0.08
<b>KAK</b>	0.04	0.07	0.03	0.09	0.00	0.07	0.07
<b>KA2K</b>	0.03	0.06	0.03	0.07	0.00	0.06	0.06
<b>KA3K</b>	0.03	0.06	0.02	0.07	0.00	0.06	0.06
<b>KA4K</b>	0.02	0.05	0.02	0.06	0.01	0.04	0.04
<b>KA5K</b>	0.03	0.07	0.03	0.08	0.01	0.06	0.06
<b>Media</b>	0.03	0.07	0.03	0.08	0.00	0.06	0.06

**Tabla A31.** Distancias interatómicas X-O (Å) de las estructuras de los estados de transición de las reacciones HAT exergónicas con  $\cdot\text{OCH}_3$ ,  $\cdot\text{OOH}$  y  $\cdot\text{OOCH}_3$ .

	$\cdot\text{OCH}_3$				$\cdot\text{OOCH}_3$				$\cdot\text{OOH}$			
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$
<b>HAT (aq)</b>												
C1	2.523	2.549	2.556	2.626								
C2	2.509	2.528	2.603	2.614				2.614			2.565	2.593
C4			2.621									
C5	2.581	2.582	2.674	2.645			2.631	2.637	2.534	2.534	2.634	2.638
C6	2.589	2.588	2.601	2.489								
O7	2.303	2.304	2.216	2.102								
C8	2.542	2.542	2.685	2.672			2.646	2.673			2.645	2.621
N9		2.163	2.588	2.478								
<b>HAT (pe)</b>												
C1	2.497	2.516	2.534	2.662								
C2	2.486	2.507	2.576	2.608			2.557				2.554	2.530
C4			2.726									
C5	2.514	2.520	2.656	2.640			2.594				2.585	
C6	2.558	2.537	2.572	2.534								
O7			2.496	2.442								
C8	2.508	2.533	2.630	2.375			2.579	2.596			2.638	2.624
N9			2.400	2.481								

**Tabla A32.** Distancias interatómicas X-H (Å) de las estructuras de los estados de transición de las reacciones HAT exergónicas con  $\cdot\text{OCH}_3$ ,  $\cdot\text{OOH}$  y  $\cdot\text{OOCH}_3$ .

	$\cdot\text{OCH}_3$				$\cdot\text{OOCH}_3$				$\cdot\text{OOH}$			
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$
<b>HAT (aq)</b>												
C1	1.231	1.208	1.203	1.405								
C2	1.243	1.223	1.184	1.356				1.389			1.249	1.330
C4			1.225									
C5	1.180	1.179	1.225	1.163			1.274	1.315	1.268	1.268	1.266	1.301
C6	1.183	1.184	1.173	1.210								
O7	1.169	1.169	0.980	1.102								
C8	1.208	1.207	1.221	1.266			1.275	1.251			1.270	1.346
N9		1.100	1.205	1.450								
<b>HAT (pe)</b>												
C1	1.255	1.222	1.216	1.262								
C2	1.267	1.242	1.198	1.196			1.263				1.283	1.295
C4			1.186									
C5	1.236	1.230	1.168	1.163			1.235				1.260	
C6	1.193	1.203	1.188	1.153								
O7			1.250	1.356								
C8	1.236	1.210	1.176	1.175			1.232	1.299			1.275	1.211
N9			1.200	1.450								

**Tabla A33.** Distancias interatómicas X-O (Å) de las estructuras de los estados de transición de las reacciones RAF exergónicas con  $\cdot\text{OCH}_3$ ,  $\cdot\text{OOH}$  y  $\cdot\text{OOCH}_3$ .

	$\cdot\text{OCH}_3$				$\cdot\text{OOH}$			
	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$	$\text{TH}^{2+}$	$\text{T}^+$	$\text{TOH}$	$\text{TO}^\cdot$
<b>RAF (aq)</b>								
C3		1.928						
C4		2.034	2.033				1.925	1.928
C10		1.954	2.011	2.049	1.998			
C15		1.987	1.991	2.127/2.000	2.193			
C17		1.973	1.971	2.201/2.000	2.000		2.000	2.294
<b>RAF (pe)</b>								
C4		2.023	1.996					
C10		2.002	1.942	2.013				
C15		1.957	1.957	2.191/2.204	2.201		2.466	
C17		1.955	1.963	2.184/2.199				2.309

## Apéndice B. Métodos de Química Cuántica

### 1. La ecuación de Schrödinger

El estado de un sistema cuántico se describe por su función de onda  $\Phi$ . La ecuación de Schrödinger es una ecuación de valores propios que relaciona el operador Hamiltoniano cuántico  $\hat{H}$ , i.e. operador energía, con sus funciones propias  $\Phi$  y valores propios  $E$ . Donde  $\Phi$  y  $E$  corresponden a las funciones de onda de los estados del sistema y las energías de cada estado respectivamente. La forma no relativista e independiente del tiempo de la ecuación de Schrödinger tiene la forma

$$\hat{H} \Phi = E \Phi \quad [1]$$

Para un Sistema de  $N$  electrones y  $M$  núcleos y considerando unidades atómicas, en las que la carga del electrón  $e$  y su masa  $m$ , la constante reducida de Planck  $\hbar$  y la constante de Coulomb  $1/4\pi\epsilon_0$  adoptan el valor unidad por definición, el operador Hamiltoniano es

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>1}^M \frac{Z_A Z_B}{R_{AB}} \quad [2]$$

Siendo  $r_{iA}$  la distancia entre el electrón  $i$  y el núcleo  $A$ ,  $r_{ij}$  la distancia entre los electrones  $i$  y  $j$ ,  $R_{AB}$  la distancia entre los núcleos  $A$  y  $B$ ,  $Z_A$  la carga del núcleo  $A$  y  $M_A$  la masa del núcleo  $A$ . Los primeros dos términos de la ecuación [2] son los operadores energía cinética de los electrones y núcleos respectivamente. Mientras que los términos tercero a quinto son respectivamente los operadores de energía potencial electrostática núcleo-electrón, electrón-electrón y núcleo-núcleo.

### 2. La aproximación de Born-Oppenheimer

Las masas de los núcleos son mucho mayores que las de los electrones, i.e.  $\sim 1836$  veces mayor en el caso del protón, que constituye el núcleo del átomo más ligero  $^1\text{H}$ . Por este motivo, es razonable considerar que los electrones se mueven mucho más rápido que los núcleos. La aproximación de Born-Oppenheimer considera que los movimientos de electrones y núcleos están desacoplados y por tanto la función de onda total del sistema  $\Phi(\{r_i, R_A\})$  se puede expresar como un producto de una función de onda electrónica que describe el movimiento e interacciones de los electrones en un potencial determinado por las posiciones fijas de los núcleos,  $\Phi_{el}(\{r_i\}, \{R_A\})$ , y de una función de onda nuclear que describe el movimiento e interacciones de los núcleos en el campo de potencial promedio de los electrones,  $\Phi_{nuc}(\{R_A\})$ :

$$\Phi(\{r_i, R_A\}) = \Phi_{el}(\{r_i\}, \{R_A\}) \Phi_{nuc}(\{R_A\}) \quad [3]$$

De esta forma, se puede escribir la ecuación de Schrödinger electrónica como

$$\hat{H}_{el} \Phi_{el} = E_{el} \Phi_{el} \quad [4]$$

donde  $\hat{H}_{el}$  es el Hamiltoniano electrónico

$$\hat{H}_{el} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>1}^M \frac{Z_A Z_B}{R_{AB}} \quad [5]$$

y los valores propios de  $\hat{H}_{el}$ ,  $E_{el}$ , se denominan energías electrónicas. Así, la energía total del sistema para unas coordenadas nucleares determinadas es la suma de la energía electrónica y de la repulsión internuclear para dichas coordenadas

$$E_{tot} = E_{el} + \sum_{A=1}^M \sum_{B>1}^M \frac{Z_A Z_B}{R_{AB}} \quad [6]$$

Las ecuaciones [5] y [6] muestran que la energía y función de onda electrónicas dependen explícitamente de las coordenadas electrónicas y paraméricamente de las coordenadas nucleares. Por tanto, en la aproximación de Born-Oppenheimer debe resolverse la ecuación de Schrödinger electrónica para obtener la función de onda,  $\Phi_{el}$ , y energía electrónicas,  $E_{el}$ , que mejor se adaptan a las coordenadas internucleares especificadas.

### 3. Principio de antisimetría

Los electrones son partículas elementales, por lo que poseen un momento angular intrínseco denominado espín. En particular, los electrones son fermiones cuyo espín tiene valor  $\frac{1}{2}$ . Además, al ser los electrones indistinguibles entre sí, según el Teorema espín-estadística, la función de onda electrónica total debe ser antisimétrica respecto al intercambio de las coordenadas espaciales-espín  $x_i$  de dos electrones  $i, j$  cualesquiera:

$$\Phi_{el}(x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -\Phi_{el}(x_1, \dots, x_j, \dots, x_i, \dots, x_N) \quad [7]$$

Como la ecuación [5] muestra, el Hamiltoniano electrónico no considera las coordenadas de espín  $w$  y en consecuencia sus funciones propias,  $\Phi_{el}$ , tampoco. Por tanto, en la formulación no relativista de la mecánica cuántica, las funciones de espín  $\alpha(w)$  y  $\beta(w)$  deben introducirse *a posteriori* para describir correctamente el comportamiento electrónico. Las funciones de espín  $\alpha(w)$  y  $\beta(w)$  cumplen la propiedad de ortonormalidad:

$$\int dw \alpha^*(w) \alpha(w) = \int dw \beta^*(w) \beta(w) = 1 \quad [8]$$

$$\int dw \alpha^*(w) \beta(w) = \int dw \beta^*(w) \alpha(w) = 0 \quad [9]$$

A las funciones de onda electrónicas que describen las coordenadas espaciales,  $r_i$ , y de espín,  $w_i$ , de un electrón  $i$  se conocen como funciones espín-orbitales  $\chi_i(x_i)$ . Una manera de construir funciones de onda espín-orbital es multiplicar funciones de onda espaciales,  $\psi_i(r_i)$ , por una de las funciones de espín  $\alpha(w)$  o  $\beta(w)$ .

#### 4. Determinantes de Slater

La ecuación de Schrödinger electrónica [4] es un problema de  $N$  cuerpos que no tiene solución analítica general para sistemas de más de un electrón. En la práctica, las funciones de onda polielectrónicas se expresan en términos de funciones de onda espín-orbital monoeléctricas  $\chi_i(x_i)$ . Sin embargo, el principio de antisimetría implica que únicamente aquellas funciones de onda polielectrónicas que satisfacen la ecuación [7] son aceptables. Una forma de asegurar que esta condición se cumple es combinar las funciones de onda espín-orbital monoeléctricas  $\chi_i(x_i)$  según determinantes de Slater:

$$\Psi(x_1, x_2, \dots, x_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \dots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \dots & \chi_k(x_2) \\ \vdots & \vdots & & \vdots \\ \chi_i(x_N) & \chi_j(x_N) & \dots & \chi_k(x_N) \end{vmatrix} = |\Psi\rangle \quad [10]$$

donde  $(N!)^{-1/2}$  es una constante de normalización. En un determinante de Slater, los elementos de una fila describen al mismo electrón según todos los posibles espín-orbitales empleados mientras que los elementos de una columna corresponden a todos los electrones del sistema descritos por el mismo espín-orbital. De esta manera, el intercambio de coordenadas de dos electrones, que corresponde al intercambio de dos filas del determinante, cambia el signo de la función de onda total. Además, si dos electrones estuvieran descritos por el mismo espín orbital, dos filas del determinante serían iguales

y el determinante se anularía, indicando que no existen tales funciones de onda y cumpliendo el principio de exclusión de Pauli. Estas dos propiedades de los determinantes implican que las funciones de onda polieletrónicas construidas de esta manera incluyen cierto tipo de correlación, denominada de intercambio. Sin embargo, es posible construir determinantes de Slater en los que dos columnas estén constituidas por espín-orbitales que únicamente difieran en las coordenadas de espín. Esto implica que las funciones de onda totales resultantes describen estados en los que hay probabilidad no nula de que dos electrones coexistan en la misma región espacial o, lo que es lo mismo, que los electrones de espín antiparalelo no están correlacionados.

## 5. El método de Hartree-Fock

El método Hartree-Fock, HF, permite obtener una función de onda polieletrónica de forma aproximada. En primer lugar, HF considera que la función de onda electrónica total se describe mediante un único determinante de Slater, es decir, HF es un método monodeterminantal.

$$\Phi \approx |\Psi_0\rangle \quad [11]$$

Debe recordarse, que al emplear un único determinante de Slater, HF no describe correctamente la correlación electrónica entre electrones de espín opuesto. Por otra parte, HF no es apropiado para estados degenerados que no pueden describirse por un único determinante de Slater. A pesar de todo, HF es una aproximación central en los métodos de química cuántica ya que supone el punto de partida de otros métodos más precisos. El método HF emplea el principio variacional para resolver la ecuación de valores propios [4] y proporciona un conjunto de funciones propias ( $\Psi_0, \Psi_1, \Psi_2\dots$ ) con valores propios ( $E_0, E_1, E_2\dots$ ). La función propia  $\Psi_0$  asociada al menor valor propio  $E_0$  corresponden a la función de onda del estado fundamental y su energía dentro de la aproximación del método HF.

$$H|\Psi_0\rangle = E_0|\Psi_0\rangle \quad [12]$$

$$\langle\Psi_0|H|\Psi_0\rangle = E_0 \quad [13]$$

En la siguiente discusión se denotarán con los índices  $a, b, c\dots$  a los espín-orbitales ocupados y los índices  $r, s, t\dots$  designarán los espín-orbitales vacíos o virtuales. Considerando la forma de  $\hat{H}_{el}$  dado por la ecuación [5], la ecuación [13] se puede expresar como

$$E_0 = \langle \Psi_0 | H | \Psi_0 \rangle = \sum_a^N \langle a | h | a \rangle + \frac{1}{2} \sum_a^N \sum_b^N \langle ab || ab \rangle \quad [14]$$

Donde  $h$  es un operador,  $\hat{h}$ , formado por los operadores energía cinética del electrón y potencial electrostática de interacción electrón-núcleos de [5]

$$\hat{h} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} \quad [15]$$

$\langle i|h|j \rangle$  es la notación para las integrales monoeléctricas

$$\langle i|h|j \rangle = \langle \chi_i | h | \chi_j \rangle = \int dx_1 \chi_i^*(x_1) h(r_1) \chi_j(x_1) \quad [16]$$

y  $\langle ij||kl \rangle$  indica la diferencia  $\langle ij|kl \rangle - \langle ij|lk \rangle$ , donde  $\langle ij|kl \rangle$  es la notación de las integrales bielectrónicas:

$$\langle ij|kl \rangle = \langle \chi_i \chi_j | \chi_k \chi_l \rangle = \int dx_1 dx_2 \chi_i^*(x_1) \chi_j^*(x_2) r_{12}^{-1} \chi_k(x_1) \chi_l(x_2) \quad [17]$$

En la ecuación [14] se pueden distinguir dos tipos de integrals bielectrónicas, las de tipo  $\langle ab|ab \rangle$  y las de tipo  $\langle ab|ba \rangle$ . Las integrales  $\langle ab|ab \rangle$  se denominan integrales de Coulomb y corresponden a la repulsión electrostática entre la densidad de carga del electrón 1,  $\chi_a^*(x_1)\chi_a(x_1)$ , y la del electrón 2,  $\chi_b^*(x_2)\chi_b(x_2)$ .

$$\langle ab|ab \rangle = \int dx_1 dx_2 \chi_a^*(x_1) \chi_b^*(x_2) r_{12}^{-1} \chi_a(x_1) \chi_b(x_2) \quad [18]$$

Las integrales  $\langle ab|ba \rangle$  se denominan integrales de intercambio. No tienen un equivalente en mecánica clásica y surgen del producto antisimétrico de los determinantes de Slater.

$$\langle ab|ba \rangle = \int dx_1 dx_2 \chi_a^*(x_1) \chi_b^*(x_2) r_{12}^{-1} \chi_b(x_1) \chi_a(x_2) \quad [19]$$

Tras integrar las coordenadas de espín, las integrales  $\langle ab|ab \rangle$  y  $\langle ab|ba \rangle$  se denotan por  $J_{ab}$  y  $K_{ab}$  respectivamente. El operador de Coulomb  $\hat{J}_b$  actuando sobre  $\chi_a$  representa la repulsión promedio del electrón en  $\chi_a$  causada por otro electrón en  $\chi_b$ .

$$\hat{J}_b \chi_a(x_1) = \left[ \int dx_2 \chi_b^*(x_2) r_{12}^{-1} \chi_b(x_2) \right] \chi_a(x_1) \quad [20]$$

El operador de intercambio  $\hat{K}_b$  no tiene una interpretación simple en términos clásicos, representa la energía de intercambio debido al principio de antisimetría y se define por su efecto al actuar sobre un espín-orbital  $\chi_a$ .

$$\hat{K}_b \chi_a(x_1) = \left[ \int dx_2 \chi_b^*(x_2) r_{12}^{-1} \chi_a(x_2) \right] \chi_b(x_1) \quad [21]$$

Según la ecuación [13], la energía  $E_0$  es un funcional de los espín-orbitales. Las ecuaciones de Hartree-Fock surgen de la minimización de  $E_0$  respecto al conjunto de espín-orbitales  $\{\chi_a\}$  bajo la restricción de que éstos sean ortonormales

$$\langle a|b \rangle = \int dx_1 \chi_a^*(x_1) \chi_b(x_1) = \delta_{ab} \quad [22]$$

Empleando el método de los multiplicadores indeterminados de Lagrange y la restricción [22], se obtiene el siguiente funcional de los espín-orbitales

$$L[\{\chi_a\}] = E_0[\{\chi_a\}] - \sum_a \sum_b \varepsilon_{ab} (\langle a|b \rangle - \delta_{ab}) \quad [23]$$

donde el funcional de la energía  $E_0[\{\chi_a\}]$  viene determinado por la ecuación [14] y  $\varepsilon_{ab}$  son los multiplicadores de Lagrange. La minimización de  $E_0$  se consigue derivando e igualando a cero  $L[\{\chi_a\}]$

$$\delta L = 2 \sum_a^N \int dx_1 \delta \chi_a(1) \left[ \hat{h}(1) \chi_a(1) + \sum_b^N (\hat{J}_b(1) - \hat{K}_b(1)) \chi_a(1) - \sum_b^N \varepsilon_{ba} \chi_b(1) \right] = 0 \quad [24]$$

Como la variación de los espín-orbitales,  $\delta \chi_a(1)$ , es arbitraria y diferente de cero, el término entre corchetes de [24] debe ser cero

$$\left[ \hat{h}(1) + \sum_b^N \hat{J}_b(1) - \hat{K}_b(1) \right] \chi_a(1) = \sum_b^N \varepsilon_{ba} \chi_b(1) \quad a = 1, 2, \dots, N \quad [25]$$

resultando un sistema de  $N$  ecuaciones denominado ecuaciones de Hartree-Fock. El conjunto de operadores de la ecuación [25] se conoce como operador de Fock,  $\hat{f}$ , cuyas funciones y valores propios son los espín-orbitales y las energías orbitales.

$$\hat{f} = \hat{h}(1) + \sum_b^N \hat{J}_b(1) - \hat{K}_b(1) \quad [26]$$

Expresando las ecuaciones de Hartree-Fock en forma matricial y tras transformaciones unitarias se obtiene la forma canónica de las ecuaciones de Hartree-Fock

$$f | \chi_a \rangle = \varepsilon_a | \chi_a \rangle \quad [27]$$

Las funciones propias de las ecuaciones canónicas de Hartree-Fock [27] son los espín-orbitales canónicos, que se deslocalizan sobre todos los núcleos. Por esta razón también se conocen como espín-orbitales moleculares canónicos. Como indica la ecuación [26], el operador de Fock depende de sus funciones propias a través de los operadores de Coulomb [20] e intercambio [21]. Esto implica que las ecuaciones de Hartree-Fock no son lineales y deben resolverse de forma iterativa. Empezando desde un conjunto de espín-orbitales de prueba se define el operador de Fock, se resuelve la ecuación de valores propios [27], se obtienen nuevos espín-orbitales y se repite el ciclo iterativamente hasta alcanzar la convergencia. Por este motivo, el método de Hartree-Fock también se conoce como método de campo autoconsistente o, en inglés, Self-Consistent Field (SCF) method.

## 6. Las ecuaciones de Roothan-Hall

El problema de las ecuaciones de Hartree-Fock es que no existe una expresión analítica de los espín-orbitales que resuelven [27] y debe buscarse una solución numérica. Roothan y Hall mostraron que las ecuaciones integro-diferenciales de Hartree-Fock pueden ser transformadas en un conjunto de ecuaciones lineales y ser resueltas con técnicas algebraicas. En este método, conocido como Combinación Lineal de Orbitales Atómicos CLOA, se introduce un conjunto de  $K$  funciones espaciales de base conocidas  $\{\phi_\mu\}$  de forma que la parte espacial de los orbitales moleculares se exprese como combinación lineal de las funciones de base.

$$\psi_i = \sum_{\mu}^K C_{\mu i} \phi_{\mu} \quad i = 1, 2, \dots, K \quad [28]$$

Así el problema de optimizar  $N$  funciones (orbitales moleculares) de expresión matemática desconocida se transforma en optimizar los  $K$  coeficientes de las combinaciones lineales de los  $N$  orbitales moleculares. En la siguiente discusión se obtendrán las ecuaciones de Roothan-Hall para el método de “restricted Hartree-Fock” para sistemas de capa cerrada. En esta formulación del método HF, los espín-orbitales  $\alpha$  y  $\beta$  se construyen a partir de los mismos orbitales espaciales. Además, como los sistemas de capa cerrada están formados por un número par de electrones, en un cálculo “restricted HF”, los electrones  $\alpha$  y  $\beta$  están apareados de forma que únicamente  $N/2$  orbitales espaciales están ocupados. Integrando los espines de [27], las ecuaciones de Hartree-Fock se transforman en

$$f|\psi_i\rangle = \varepsilon_i|\psi_i\rangle \quad [29]$$

Sustituyendo los orbitales moleculares  $\psi_i$  por las combinaciones lineales de funciones de base [28], multiplicando por la izquierda por  $\langle\phi_\mu|$  e integrando, la ecuación [29] se transforma en la forma matricial de las ecuaciones de Roothan-Hall

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \quad [30]$$

donde  $\mathbf{C}$  es una matriz cuyas  $K$  columnas contienen los  $K$  coeficientes de las combinaciones lineales de cada orbital  $\psi_i$ ,  $\boldsymbol{\varepsilon}$  es una matriz diagonal de las energías de los orbitales  $\varepsilon_i$ ,  $\mathbf{F}$  es la representación matricial del operador de Fock en la base  $\{\phi_\mu\}$  y  $\mathbf{S}$  es la matriz de solapamiento. Los elementos de la matriz de solapamiento,  $S_{\mu\nu}$ , vienen determinados por las integrales

$$S_{\mu\nu} = \int d\mathbf{r}_1 \phi_{\mu}^*(1) \phi_{\nu}(1) = (\mu|\nu) \quad [31]$$

y los elementos de la matriz de Fock,  $F_{\mu\nu}$ , corresponden a

$$F_{\mu\nu} = (\mu|h|\nu) + \sum_a \sum_{\lambda\sigma}^{N/2} C_{\lambda a} C_{\sigma a}^* [2(\mu\nu|\sigma\lambda) - (\mu\lambda|\sigma\nu)] \quad [32]$$

donde las integrales  $(\mu|h|\nu)$  son el equivalente a las integrales monoeléctricas  $\langle\mu|h|\nu\rangle$  [16] para orbitales espaciales mientras que  $(\mu\mu|h|\nu\nu)$  y  $(\mu\nu|h|\mu\nu)$  son el equivalente a las integrales bielectrónicas de Coulomb  $\langle\mu\nu|\mu\nu\rangle$  [18] e Intercambio  $\langle\mu\nu|\nu\mu\rangle$  [19].

A partir de la expresión de los elementos del operador de Fock [32] se puede definir una nueva matriz  $\mathbf{P}$ , denominada matriz densidad, cuyos elementos  $\mu\nu$  vienen determinados por el producto de los coeficientes de la combinación lineal de los orbitales espaciales  $\mu$  y  $\nu$ .

$$P_{\mu\nu} = 2 \sum_a^{N/2} C_{\mu a} C_{\nu a}^* \quad [34]$$

Las ecuaciones [34], [32] y [28] ponen de manifiesto que la definición del operador de Fock depende de sus funciones propias.

$$\mathbf{F}(\mathbf{C})\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\varepsilon} \quad [35]$$

Por lo que las ecuaciones de Roothan-Hall no son lineales y deben resolverse de forma iterativa. Si la matriz de solapamiento fuera la matriz identidad, la ecuación [30] se simplificaría a

$$\mathbf{F}\mathbf{C} = \mathbf{C}\boldsymbol{\varepsilon} \quad [36]$$

y sus funciones y valores propios se podrían obtener por diagonalización de la matriz de Fock  $\mathbf{F}$ . Para que  $\mathbf{S}$  fuera una matriz identidad, las funciones de base  $\phi_{\mu}$  deberían ser ortonormales. Aunque en principio es posible elegir funciones de base ortonormales otra estrategia es emplear las funciones de base más convenientes y emplear la

transformación de Löwdin para ortogonalizar el conjunto de funciones de base  $\{\phi_\mu\}$ . En la base de las funciones ortonormales,  $\mathbf{F}'$  y  $\mathbf{C}'$  se definen como

$$\mathbf{F}' = \mathbf{S}^{-1/2} \mathbf{F} \mathbf{S}^{-1/2} \quad [37]$$

$$\mathbf{C} = \mathbf{S}^{-1/2} \mathbf{C}' \quad [38]$$

Empleando las transformaciones [37] y [38], las ecuaciones de Roothan-Hall resultantes se pueden resolver diagonalizando  $\mathbf{F}'$ .

$$\mathbf{F}' \mathbf{C}' = \mathbf{C}' \boldsymbol{\varepsilon} \quad [39]$$

La solución iterativa de [39] proporciona la matriz  $\mathbf{C}'$  a partir de la cual calcular  $\mathbf{C}$  a partir [38] y, a su vez,  $\mathbf{C}$  se emplea para definir la matriz de Fock  $\mathbf{F}$  según las ecuaciones [34] y [32]. De esta forma se pueden resolver las ecuaciones de Roothan-Hall [30] para obtener las nuevas matrices  $\mathbf{C}$  y  $\boldsymbol{\varepsilon}$  formadas por las funciones propias y valores propios de  $\mathbf{F}$ .

## 7. Correlación Electrónica

Como se ha mostrado, el método de Hartree-Fock aproxima una función  $N$ -electrónica como un producto antisimetrizado de  $N$  funciones monoeléctricas según un único determinante de Slater. Las propiedades de los determinantes aseguran que dos electrones no pueden estar descritos simultáneamente por el mismo espín-orbital. Sin embargo, el método HF no tiene en cuenta la correlación instantánea entre electrones de espín opuesto porque únicamente se considera la interacción promedio de sus densidades electrónicas. Aunque HF es una aproximación inicial aceptable, un cálculo preciso de propiedades como energías de disociación requieren un tratamiento más preciso de la correlación electrónica.

Se puede demostrar que la forma exacta de una función de onda  $N$  electrónica se puede expresar como combinación lineal de los  $N$  determinantes de Slater formados por una base completa de espín orbitales  $\{\chi_i\}$ . Tomando el determinante de Slater del estado fundamental  $|\Psi_0\rangle$  como referencia, la expansión exacta de la función de onda es

$$\Phi = c_0 |\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c \\ r<s<t}} c_{abc}^{rst} |\Psi_{abc}^{rst}\rangle + \dots \quad [39]$$

donde  $|\Psi_a^r\rangle$  es una función de onda construida a partir del determinante de Slater del estado fundamental  $|\Psi_0\rangle$  en el que se ha reemplazado el espín-orbital ocupado  $a$ , por el espín-orbital virtual  $r$  y  $c_a^r$  es su coeficiente en la combinación lineal de determinantes de Slater. Las sumas que implican desigualdades como  $a < b$  implican sumas sobre todos los espín-orbitales ocupados  $a$  y sobre todos los espín-orbitales ocupados  $b$  de mayor energía a  $a$ . Las desigualdades  $r < s$  son análogas a  $a < b$  para los espín-orbitales virtuales.

Así, se puede considerar que los determinantes del segundo término de [39] describen estados monoexcitados, los del tercer término estados diexcitados y así sucesivamente para los términos posteriores. Como cada determinante representa una configuración electrónica, este método se denomina Interacción de Configuraciones (Configuration Interaction, CI, en inglés).

La energía de correlación se define como la diferencia de energía del estado fundamental obtenida con los métodos de Interacción de Configuraciones,  $E_{CI}$ , y Hartree-Fock,  $E_{HF}$ .

$$E_{corr} = E_{HF} - E_{CI} \quad [40]$$

En el límite en el que se emplea una base infinita de espín-orbitales, el valor EHF alcanza el denominado límite de Hartree-Fock y ECI alcanza el valor exacto de la energía del sistema sin tener en cuenta efectos relativistas y dentro de la aproximación de Born-Oppenheimer.

Para un sistema finito de  $K$  funciones de base y  $N$  electrones, se pueden formar  $2K$  espín-orbitales y  $\binom{2K}{N}$  determinantes. Cuando se emplean la totalidad de los determinantes, el método se denomina “full-CI”. Dado que el número de determinantes posibles es astronómico salvo para sistemas con pocos electrones y conjuntos de base pequeños, en la práctica se suele restringir el cálculo CI a un número limitado de determinantes.

El inconveniente de restringir el cálculo CI a un orden determinado de excitaciones es que el método deja de ser consistente con el tamaño: la energía calculada para un sistema formado por dos fragmentos no interaccionantes no es igual a la energía calculada para los fragmentos por separado. Existen otros métodos aproximados que calculan ciertos elementos de la matriz full-CI, recuperando parte de la correlación electrónica, y sí son consistentes con el tamaño. Entre los más empleados se encuentran los métodos Coupled Cluster (CC) y Møller-Plesset (MP). Partiendo de una función de onda calculada con Hartree-Fock, estos métodos corrigen la energía del estado fundamental empleando teoría de perturbaciones.

Como estos métodos parten de una función de onda monodeterminantal, no son apropiados cuando el estado fundamental está descrito por estados degenerados o quasidegenerados. En estos casos, se requiere emplear varios determinantes de Slater para

describir la función de onda. Estos métodos se conocen como métodos multiconfiguracionales y los más comunes son el “Multiconfigurational Self Consistent Filed” (MCSCF), el “Complete Active Site Self Consistent Field” (CASSCF) y su variante perturbativa CASPT2.

### 8. Teoría del Funcional de la Densidad, Teoremas de Hohenberg-Kohn

La Teoría del Funcional de la Densidad, DFT en inglés, es una formulación alternativa al problema electrónico, en la que el objeto físico fundamental es la densidad electrónica en lugar de la función de onda. Los métodos DFT son de gran interés puesto que permiten incluir correlación electrónica con un coste computacional bajo.

Los fundamentos rigurosos de DFT fueron propuestos por Hohenberg y Kohn en dos teoremas. El primero indica que: *cualquier observable de un estado estacionario fundamental no degenerado puede ser calculado, en principio de forma exacta, a partir de la densidad electrónica de este estado fundamental. Esto es, cualquier observable puede describirse como un funcional de la densidad electrónica del estado fundamental.*

Como indica el propio teorema, éste es válido únicamente para estados no degenerados y también requiere que 1) la densidad electrónica  $\rho(r)$  sea una función positiva definida en todo el espacio y 2) la integral sobre todo el espacio de la densidad debe ser igual al número de electrones del sistema.

$$\rho(r) \geq 0 \quad \text{and} \quad \int dr \rho(r) = N \quad [41]$$

La demostración del primer teorema proporciona una conclusión fundamental por la cual la densidad del estado fundamental  $\rho(r)$  está directamente conectada con el denominado potencial externo  $v(r)$  definido por los núcleos de la molécula. Además, puesto que tanto el número de electrones como su interacción con los núcleos están determinados por  $\rho(r)$ , también lo están el Hamiltoniano y la función de onda del estado fundamental. Por tanto, la densidad electrónica contiene toda la información necesaria para calcular cualquier observable del estado fundamental. La expresión del funcional de la energía  $E[\rho]$  es de especial interés:

$$E[\rho] = T[\rho] + V_{Ne}[\rho] + V_{ee}[\rho] + (V_{NN}) \quad [42]$$

$T[\rho]$  y  $V_{ee}[\rho]$  son los funcionales de la energía cinética electrónica y del potencial de repulsión interelectrónico, y son funcionales universales puesto que no dependen del potencial externo.  $V_{Ne}[\rho]$  es el funcional del potencial de atracción núcleo-electrón y  $V_{NN}$  es el potencial de repulsión internuclear, que se añade como una constante.

El segundo teorema indica que: *la densidad electrónica de un estado fundamental no degenerado puede ser calculada, en principio de forma exacta,*

determinando aquella densidad que minimiza la energía del estado fundamental. Este teorema implica

$$E_0 \leq E_v[\tilde{\rho}] \quad [43]$$

donde  $E_0$  es la energía exacta del estado fundamental y  $E_v$  es la energía dada por el potencial externo  $v(r)$  determinado por la densidad de prueba  $\tilde{\rho}(r)$ . Este teorema proporciona un principio variacional que permite determinar la densidad exacta del estado fundamental mediante el cálculo de la densidad que minimiza el funcional de la energía.

$$\frac{\delta E_v[\rho]}{\delta \rho(r)} = 0 \quad [44]$$

Sin embargo, esta minimización requiere que la condición impuesta por la ecuación [41] se mantenga. Así, se puede emplear el método de multiplicadores indeterminados de Lagrange con la restricción [41] para obtener el funcional  $L[\rho]$

$$L[\rho] = E_0[\rho] - \mu \left( \int dr \rho(r) - N \right) \quad [45]$$

donde  $\mu$  es un multiplicador de Lagrange y tiene el significado físico de un potencial químico. Los mínimos de los funcionales  $L[\rho]$  y  $E_0[\rho]$  coinciden ya que el segundo término de [45] es cero. Esto permite minimizar  $E_0[\rho]$  derivando e igualando a cero  $L[\rho]$ .

$$\int \frac{\delta E_v[\rho]}{\delta \rho(r)} \delta \rho(r) dr - \mu \int \delta \rho(r) dr = 0 \quad [46]$$

Donde la definición de derivada funcional se ha tenido en cuenta. La condición de mínimo bajo la restricción impuesta se obtiene cuando

$$\mu = \frac{\delta E_v[\rho]}{\delta \rho(r)} \quad [47]$$

Recuperando la definición de  $E[\rho]$  dada por la ecuación [42], [47] se transforma en la ecuación Fundamental de la Teoría del Funcional de la Densidad.

$$\mu = \frac{\delta E_v[\rho]}{\delta \rho(r)} = v_n(r) + \frac{\delta(T[\rho] + V_{ee}[\rho])}{\delta \rho(r)} \quad [48]$$

### 9. El método de Kohn-Sham

Los teoremas de Hohenberg-Kohn demuestran que la densidad electrónica determina el Hamiltoniano y la función de onda de un sistema  $N$ -electrónico. Esta interpretación conduce de nuevo a la formulación ondulatoria de la mecánica cuántica y, hasta aquí, la Teoría del Funcional de la Densidad no implica ninguna ventaja respecto a los métodos basados en la función de onda. Sin embargo, de conocerse la expresión exacta del funcional de la densidad, se podría resolver la ecuación [48] con menor complejidad que los métodos basados en la función de onda.

Kohn y Sham propusieron emplear un sistema de referencia de electrones no interaccionantes bajo el efecto de un potencial ficticio,  $v_s(r)$ , cuya densidad electrónica sea la misma que la de un sistema real de electrones interaccionantes. La ventaja de esta estrategia es que el método de Hartree-Fock proporciona las soluciones exactas para sistemas de electrones no interaccionantes. En el sistema de referencia de electrones no interaccionantes, la energía puede expresarse como

$$E_v[\rho] = T_s[\rho] + \int dr \rho(r) v_s(r) \quad [49]$$

donde  $T_s(\rho)$  es la energía cinética de los electrones en el sistema no interaccionante. En el caso del sistema real, la energía puede descomponerse en las siguientes contribuciones

$$E_v[\rho] = T[\rho] + \int dr \rho(r) v_n(r) + V_{ee}(\rho) \quad [50]$$

El término  $T(\rho)$  es la energía cinética exacta, que difiere de  $T_s(\rho)$  del sistema de referencia porque en el sistema real, el movimiento de cada electrón está correlacionado con el movimiento de los otros electrones. Añadiendo y sustrayendo  $T_s(\rho)$  y  $J(\rho)$  a [50] se obtiene

$$E_v[\rho] = T_s[\rho] + \int dr \rho(r) v_n(r) + J[\rho] + (T[\rho] - T_s[\rho]) + (V_{ee}(\rho) - J[\rho]) \quad [51]$$

El término  $J(\rho)$  es la repulsión electrostática de Coulomb electrón-electrón entre los electrones que definen la densidad electrónica  $\rho(r)$ :

$$J[\rho] = \frac{1}{2} \int dr_1 dr_2 \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} \quad [52]$$

Debe indicarse que, de acuerdo con la ecuación [52],  $J(\rho)$  tiene en cuenta la repulsión de los  $N$  electrones de  $\rho(r)$  consigo mismos. Esto implica que se hay un error resultante de sobreestimar la repulsión interelectrónica.

El término  $T(\rho) - T_s(\rho)$  de [51] es la energía cinética de correlación puesto que mide la diferencia entre la energía cinética de los electrones correlacionados en el sistema real con la energía cinética de los electrones no correlacionados del sistema de referencia. El término  $V_{ee}(\rho) - J(\rho)$  es la diferencia entre la interacción clásica electrón-electrón, incluyendo los errores de auto-interacción, y la interacción electrón-electrón cuántica exacta y corresponde con la energía de correlación-intercambio. Estos dos últimos términos de [51] se agrupan como la energía total de correlación-intercambio.

$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (V_{ee}(\rho) - J[\rho]) \quad [53]$$

Esta expresión conduce a la forma del funcional de la energía

$$E_v[\rho] = T_s[\rho] + \int dr \rho(r) v_n(r) + \frac{1}{2} \int dr_1 dr_2 \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} + E_{xc}(\rho) \quad [54]$$

Aplicando la ecuación Fundamental de DFT, introducida en [48], al funcional [54] se obtiene

$$\mu = \frac{\delta T_s[\rho]}{\delta \rho(r)} + v_n(r) + \int dr_2 \frac{\rho(r_2)}{|r_1 - r_2|} + v_{xc}(r) = \frac{\delta T_s[\rho]}{\delta \rho(r)} + v_{eff}(r) \quad [55]$$

La expresión de la energía para el sistema de referencia dada por la ecuación [50] es análoga a la del sistema real [55] con la excepción de la forma explícita del

potencial. Por tanto, la solución de [55] para el sistema real es la misma que para el sistema de referencia de electrones no interaccionantes. Esto implica que dicha solución puede obtenerse por un procedimiento análogo al método de Hartree-Fock. Las ecuaciones de Kohn-Sham resultantes son análogas a las ecuaciones de Hartree-Fock.

$$\hat{h}_{KS}\chi_i = \varepsilon_i\chi_i \quad [56]$$

En este caso, los orbitales de la base  $\{\chi_i\}$  son los orbitales de Kohn-Sham, que permiten calcular la densidad electrónica según

$$\rho(r) = \sum_i^N |\chi_i(r)|^2 \quad [57]$$

El operador Hamiltoniano de Kohn-Sham  $\hat{h}_{KS}$  se define como la suma del operador monoeléctrico energía cinética y del operador potencial externo  $v_{eff}(r)$

$$\hat{h}_{KS} = -\frac{1}{2}\nabla^2 + v_{eff}(r) \quad [58]$$

$$v_{eff}(r) = v_n(r) + \int dr_2 \frac{\rho(r_2)}{|r_1 - r_2|} + v_{XC}(r) \quad [59]$$

De forma análoga al método HF, partiendo de unos orbitales aproximados de Kohn-Sham  $\{\chi_i\}$  se define la densidad de prueba según [57], que a su vez define el potencial externo  $v_{eff}(r)$  [59] y permite la resolución de la ecuación de autovalores [56] para obtener nuevos orbitales de Kohn-Sham  $\{\chi_i^3\}$  que se emplearán en la siguiente iteración hasta que la convergencia se alcance. En la práctica, las ecuaciones de Kohn-Sham se resuelven empleando el método de combinación lineal de orbitales atómicos LCAO y las ecuaciones de Roothan-Hall correspondientes en las que el operador de Fock se denomina operador de Kohn-Sham.

## 10. Funcionales de correlación-intercambio

El método de Kohn-Sham proporciona un protocolo para minimizar la energía del sistema pero no resuelve el hecho de que el funcional exacto de repulsión interelectrónica,  $V_{ee}(\rho)$ , y la energía cinética electrónica exacta,  $T(\rho)$ , o alternativamente el potencial de correlación-intercambio,  $v_{XC}(r)$ , sean desconocidos. Sin embargo, el potencial de DFT

frente a HF es que es posible incorporar toda la energía de correlación en tanto que el potencial de correlación-intercambio se aproxime a la forma exacta.

A partir de la ecuación [53] de la energía total de correlación-intercambio, se puede definir la energía de correlación-intercambio por partícula,  $\varepsilon_{xc}[\rho]$ , y relacionarse con el potencial de correlación-intercambio

$$E_{xc}[\rho] = \int dr \rho(r) \varepsilon_{xc}[\rho] \quad [60]$$

$$v_{xc}(r) = \frac{\delta(\rho(r) \varepsilon_{xc}[\rho])}{\delta \rho(r)} \quad [61]$$

Es común separar las contribuciones de correlación  $\varepsilon_c[\rho]$  e intercambio  $\varepsilon_x[\rho]$  según

$$\varepsilon_{xc}[\rho] = \varepsilon_x[\rho] + \varepsilon_c[\rho] \quad [62]$$

Según la aproximación “Local Density Approximation”, LDA, la contribución de intercambio  $\varepsilon_x[\rho]$  se calcula a partir de la densidad considerando que es constante como en un gas homogéneo de electrones y la contribución de correlación  $\varepsilon_c[\rho]$  se desprecia o se considera constante. La aproximación “Local Spin Density Approximation”, LSDA, es la versión no restringida de LDA en la que la contribución del intercambio se calcula de forma separada para los electrones  $\alpha$  y  $\beta$  teniendo en cuenta sus densidades respectivas. Estas aproximaciones consideran que los efectos de correlación e intercambio son locales y sólo dependen de los valores de la densidad en cada punto del espacio. A pesar de lo drástico de esta simplificación, LDA proporciona buenos resultados para cálculos de estado sólido. Sin embargo, la densidad electrónica en moléculas es mucho menos homogénea e irregular que en los sólidos así que la aproximación LDA no suele describir correctamente propiedades moleculares.

Para mejorar la aproximación local, se puede considerar que las energías de correlación e intercambio dependen tanto de la densidad como de su gradiente  $\nabla \rho(r)$ . Esta aproximación se conoce como “Generalized Gradient Approximation”, GGA, y se considera como semi-local en lugar de completamente no-local. Los cálculos DFT con funcionales de tipo GGA mejoran la predicción de geometrías y frecuencias moleculares, así como los cálculos de propiedades termoquímicas de los funcionales LDA. Sin embargo, la descripción de las interacciones de van der Waals no es precisa debido al carácter no-local de éstas.

En la aproximación Meta-GGA, las contribuciones de correlación e intercambio se calculan con la densidad y el gradiente de la densidad como en la aproximación GGA, pero además se emplea la densidad de energía cinética  $\tau(r)$  y el laplaciano de la densidad,  $\nabla^2\rho(r)$ .

Dado que el método HF proporciona la energía de intercambio exacta, una solución alternativa para mejorar la contribución de intercambio es incluir parte de la energía de intercambio calculada con HF. En estos casos se emplea la conexión adiabática para calcular la cantidad de energía de intercambio HF que debe introducirse en el funcional. Los funcionales que incluyen esta corrección se denominan híbridos y pueden ser funcionales tanto GGA como meta-GGA. En general, los funcionales híbridos-GGA e híbridos-meta-GGA proporcionan resultados mejores respecto a los funcionales GGA y meta-GGA respecto al cálculo de propiedades termodinámicas, cinéticas de reacción e interacciones no covalentes.

### 11. Método CBS-4B3 y el CBS-4B3\*

Los métodos denominados compuestos emplean cálculos HF o DFT junto con diversos métodos post-HF con el objetivo de calcular energías con la mayor precisión posible. Fueron originalmente propuestos por Pople y colaboradores. El método G1 [Pople, 1989] fue rápidamente desbancado por los métodos G2 [Curtiss, 1991] y G3 [Curtiss, 1998]. En estos métodos se combina la optimización de la geometría y el cálculo de frecuencias a nivel HF o B3LYP con posteriores cálculos de la energía a nivel post-HF y una o más correcciones empíricas finales. Otros métodos son los Weizmann W1 [Martín, 1999], W2 y W3 que son parecidos a los  $G_n$  pero sin ninguna corrección empírica. En estos métodos, las geometrías se optimizan y se calculan las frecuencias con B3LYP y luego se realizan cálculos de la energía a nivel CCSD(T) con bases grandes.

Los métodos CBS (Complete Basis Set) de Petersson y colaboradores [Peterson, 1994], [Nyden 1981], [Petersson, 1988], [Petersson, 1991]. [Petersson, 1985] emplean métodos a nivel bajo, HF o DFT, con bases grandes para las optimizaciones geométricas y cálculos de frecuencias. Como la mayor parte de la energía de correlación se obtiene de perturbaciones de bajo orden, se emplean bases de tamaño medio junto con una extrapolación a set de bases infinito para las energías de segundo orden. Posteriormente, se emplean cálculos a niveles de teoría superiores con bases pequeñas y, finalmente, se calculan correcciones de contaminación de espín y empíricas. Existen diferentes métodos de esta familia como el CBS-4, CBS-QB3 y CBS-APNO, pero el más empleado es el CBS-QB3 [Ochterski, 1996] [Montgomery, 1999], [Montgomery, 2000]. En particular, el método CBS-QB3 implica los siguientes cálculos:

1. Optimización de la geometría a nivel B3LYP/6-311G(2d,d,p)
2. Cálculo de frecuencias a nivel B3LYP/6-311G(2d,d,p)
3. Cálculo de la energía a nivel MP2/6-311+G(3d2f,2df,2p)
4. Extrapolación a base infinita sobre las energías MP2/6-311++G(2d2p)
5. Cálculo de la energía a nivel MP4(SDQ)/6-31+G(d(f),p)
6. Cálculo de energía a nivel CCSD(T)/6-31+G<sup>†</sup>
7. Correcciones empíricas y de contaminación de espín.

La energía electrónica según el método CBS-QB3 se calcula como:

$$E_{CBSQB3} = E_{MP2} + \Delta E_{MP4} + \Delta E_{CCSD(T)} + \Delta E_{CBS} + \Delta E_{emp} + \Delta E_{int} \quad [63]$$

donde  $\Delta E_{CBS}$  es el término que corrige el error debido al truncamiento de las bases en la energía de segundo orden y los términos  $\Delta E_{MP4}$ ,  $\Delta E_{CCSD(T)}$ ,  $\Delta E_{emp}$  y  $\Delta E_{int}$  se calculan como

$$\Delta E_{MP4} = E_{MP4(SDQ)/6-31+G(d,f),p} - E_{MP4(SDQ)/6-31+G(d,f),p} \quad [64]$$

$$\Delta E_{CCSD(T)} = E_{CCSD(T)/6-31+G^\dagger} - E_{MP4(SDQ)/6-31+G^\dagger} \quad [65]$$

$$\Delta E_{emp} = -0.00579 \sum_{i=1}^{n_\beta} \left( \sum_{\mu=1}^{N_{virt}+1} C_{\mu i} \right)^2 |S_{ii}|^2 \quad [66]$$

$$\Delta E_{int} = -0.00954 \langle S^2 \rangle - S_z(S_z - 1) \quad [67]$$

El método CBS-4B3 [Casasnovas, 2010] es una simplificación del método CBS-QB3 en el que se han eliminado los cálculos CCSD(T), la corrección energética correspondiente a la contaminación de espín y las correcciones empíricas. Este método presenta ciertas ventajas frente al CBS-QB3 pues con un coste computacional menor, ya que se han eliminado los cálculos CCSD(T), se obtienen los mismos resultados de energías libres de protonación en fase gas [Casasnovas, 2010]. En este caso la energía se calcula como

$$E_{CBS4B3} = E_{MP2} + \Delta E_{MP4} + \Delta E_{CBS} \quad [68]$$

Los métodos CBS-QB3 y CBS-4B3 usan el funcional B3LYP con las bases 6-311G(2d,d,p) para la optimización de la estructura y el cálculo de las frecuencias. Dado que estas bases no contienen funciones difusas, las interacciones por puente de hidrógeno pueden no describirse con propiedad. Por este motivo en el método CBS-4B3\* [Casasnovas, 2011] se han sustituido las funciones 6-311G(2d,d,p) por las funciones 6-311++G(d,p) en la optimización de la estructura y el cálculo de las frecuencias.