



Transformaciones alquímicas en *clusters* y moléculas.

Tesis
entregada a la
Universidad de Chile
en cumplimiento parcial de los requisitos
para optar al grado de
Doctor en Ciencias con mención en Física
Facultad de Ciencias

por

Macarena del Pilar Muñoz González

Junio, 2018

Director de Tesis: **Dr. Carlos Cárdenas Valencia.**

FACULTAD DE CIENCIAS
UNIVERSIDAD DE CHILE

INFORME DE APROBACIÓN
TESIS DE DOCTORADO

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato

Macarena del Pilar Muñoz González

ha sido aprobada por la Comisión de Evaluación de la Tesis como requisito para optar al grado de Doctor en Ciencias con mención en Física, en el examen de Defensa de Tesis rendido el día 25 de Junio de 2018.

Director de Tesis

Dr. Carlos Cárdenas Valencia

Comisión de Evaluación de la Tesis

Dr. Eduardo Menéndez (Presidente)

Dr. Patricio Fuentealba

Dr. Ricardo Ramírez

Con amor para Luciano

BIOGRAFÍA



Macarena del Pilar Muñoz González nace el 28 de julio de 1983 en Santiago de Chile, siendo la primera hija de sus padres Carmen y Sergio. Tres años más tarde nacería su hermana y compañera de vida, Carmen Gloria, con quien hasta el día de hoy mantienen una hermosa relación. Con cinco años, Macarena ingresa al Colegio San Marcos en la comuna de Macul, colegio en el cual estaría desde kinder hasta cuarto medio, y donde descubre sus gustos y habilidades para las matemáticas. Luego de egresar de la enseñanza media, debido a su interés en la astronomía, ingresa a la carrera de Ciencias Físicas y Astronómicas en la Universidad de Concepción el año 2004, donde conoce a muy buenas amigas y además descubre que sus intereses académicos estaban realmente ligados a la física. Macarena realiza los primeros años de su carrera en Concepción y termina el último año en la Facultad de Ciencias de la Universidad de Chile. Terminada su licenciatura en física, continua sus estudios de postgrado en esta misma casa de estudios, realizando el Magíster en Ciencias Físicas el cual finaliza exitosamente el año 2013. Debido a sus inquietudes académicas y curiosidad científica, ingresa al programa de doctorado en agosto 2013, trabajo que está desarrollado en esta tesis.

AGRADECIMIENTOS

Deseo agradecer a todas las personas quienes me brindaron apoyo en esta tesis. Primero que todo a mi familia, a mis padres Carmen y Sergio, que con todo su amor y dedicación siempre me entregaron la mejor educación posible. A Carmen Gloria, por darme siempre los momentos más lindos que una hermana puede tener. También quisiera agradecer a mis grandes amigas de toda la vida, gracias por todos los momentos de ánimo que me brindaron en este difícil camino. A todos mis compañeros en el trascurso del doctorado, a mis compañeros de oficina y por supuesto a mis compañeros y profesores de grupo de investigación “*Group of Theoretical Chemical Physics*”.

A todos los profesores de los cuales aprendí en esta etapa, a los miembros de la comisión por sus valiosos comentarios y especialmente a mi tutor Dr. Carlos Cárdenas, gracias por todo el apoyo brindado en este trabajo y sobre todo, por las enseñanzas tanto en el ámbito científico como personal.

Por último, quisiera agradecer a la persona que siempre confió en mí, que siempre me ha apoyado en los momentos más difíciles, a quien me ha amado incondicionalmente, a aquella persona que le dá completo sentido a lo que hago, Alejandro.

Mi permanencia en el programa de doctorado fue posible gracias a la Beca de Doctorado CONICYT #21130691.

Además, este trabajo ha sido parcialmente financiado por FONDECYT, proyecto #1140313.

Índice

1. Introducción	1
2. Antecedentes bibliográficos	6
2.1. Teoría del Funcional de la Densidad	6
2.1.1. Funciones de respuesta química en la Teoría del Funcional de la Densidad	9
2.2. Transformaciones alquímicas	12
2.3. Cálculo de las funciones de respuesta	16
2.4. Función de respuesta lineal $\chi(\mathbf{r}, \mathbf{r}')$	18
3. Evaluación de las derivadas alquímicas	23
3.1. Aproximaciones IPA y RPA para la función de respuesta lineal.	24
3.1.1. Moléculas diatómicas	24
3.2. Métodos CPHF y CPKS para la matriz hessiana de segundas derivadas alquímicas	30
3.2.1. Moléculas tetraédricas	31
4. Estabilidad frente al dopaje sustitucional de un <i>cluster</i> semimetálico	37
4.1. Transformación alquímica tipo I: dopaje en <i>clusters</i> heteroatómicos	39

4.2. Transformación en la cual el número de electrones permanece constante.	41
4.3. Transformación en la cual el número de electrones cambia.	46
4.4. Análisis de correlación de las derivadas químicas	51
5. Prediciendo sitios de deprotonación en moléculas mono y polipróticas	54
5.1. Transformación química tipo II: Deprotonación	55
5.1.1. Cálculos de acidez en fase gaseosa y afinidad protónica	56
5.2. Moléculas monopróticas	58
5.3. Moléculas polipróticas	65
6. Conclusiones	74

Índice de figuras

2.1. Representación de la respuesta de una molécula a una perturbación inducida por cualquier modificación en el número de electrones N y en el potencial externo $\delta v(\mathbf{r})$	11
2.2. Transformación en la cual un electrón y un protón “escapan” de uno de los átomos de hidrógeno del H_2O para formar un átomo de F. . . .	12
2.3. Transformación alquímica que corresponde al dopaje de un <i>cluster</i> de Al_{13} con un átomo de Si.	13
2.4. El proceso de deprotonación puede ser visto como una transformación alquímica en la cual el protón es removido mientras el número de electrones permanece constante.	15
3.1. Comparación de los términos <i>diagonales</i> de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\alpha} = \frac{\partial^2 E}{\partial Z_\alpha^2}$) para un grupo de veinte moléculas diatómicas en la aproximación IPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).	27

- 3.2. Comparación de los términos *no diagonales* de la de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\beta} = \frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta}$) para un grupo de veinte moléculas diatómicas en la aproximación IPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. Los elementos de χ está en E_h ($1E_h = 27.2114$ eV). 28
- 3.3. Comparación de los términos *diagonales* de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\alpha} = \frac{\partial^2 E}{\partial Z_\alpha^2}$) para un grupo de veinte moléculas diatómicas en la aproximación RPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ está en E_h ($1E_h = 27.2114$ eV). 30
- 3.4. Comparación de los términos *no diagonales* de la de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\beta} = \frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta}$) para un grupo de veinte moléculas diatómicas en la aproximación RPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ está en E_h ($1E_h = 27.2114$ eV). 30
- 3.5. Comparación de los términos diagonales $\chi_{\alpha\alpha}$ para un grupo de moléculas tetraédricas para las distintas aproximaciones IPA (azul), RPA (rojo) y CPHF (verde) usando HF versus CCSD(CFD). La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV). . . . 33

- 3.6. Comparación de los términos diagonales $\chi_{\alpha\alpha}$ para un grupo de moléculas tetraédricas para las distintas aproximaciones IPA (azul), RPA (rojo) y CPKS (verde) usando a) BLYP vs. CCSD y b) B3LYP vs. CCSD. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV). 35
- 4.1. *Cluster* de Al_{13} 40
- 4.2. Energía alquímica vs. energía *ab initio* (ΔE_{rel}^{alchem} vs. $\Delta E_{rel}^{ab initio}$) de los *clusters* de aluminio dopados con átomos de silicio isoelectrónicamente $Al_{13-n}Si_n^{+n}$. Los puntos azules corresponden a derivadas alquímicas que incluyen hasta el término de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) $Al_{12}Si^+$, b) $Al_{11}Si_2^{+2}$, c) $Al_{10}Si_3^{+3}$ y d) $Al_9Si_4^{+4}$ 43
- 4.3. Energía alquímica vs. energía *ab initio* (ΔE_{rel}^{alchem} vs. $\Delta E_{rel}^{ab initio}$) de los *clusters* de aluminio dopados con átomos de silicio isoelectrónicamente $Al_{13-n}Si_n^{+n}$. Los puntos azules corresponden a derivadas alquímicas que incluyen hasta el término de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) $Al_{12}Si^+$ and b) $Al_{11}Si_2^{+2}$ 45

- 4.4. Energía alquímica vs. energía *ab initio* ($\Delta E_{\text{rel}}^{\text{alchem}}$ vs. $\Delta E_{\text{rel}}^{\text{ab initio}}$) de los *clusters* de aluminio dopados con átomos de silicio isoelectrónicamente $\text{Al}_{13-n}\text{Si}_n^{+n}$. Los puntos azules corresponden a derivadas alquímicas que incluyen hasta el término de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) $\text{Al}_{10}\text{Si}_3^{+3}$ and b) $\text{Al}_9\text{Si}_4^{+4}$ 47
- 4.5. Energía alquímica vs. energía *ab initio* ($\Delta E_{\text{rel}}^{\text{alchem}}$ vs. $\Delta E_{\text{rel}}^{\text{ab initio}}$) de los *clusters* de aluminio dopados no isoelectrónicamente con átomos de silicio $\text{Al}_{13-n}\text{Si}_n$. Los puntos azules corresponden a derivadas alquímicas que incluyen los términos electrostáticos, de transferencia de electrones y de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática y de polarización. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) Al_{12}Si , b) $\text{Al}_{11}\text{Si}_2$, c) $\text{Al}_{10}\text{Si}_3$ and d) Al_9Si_4 . 49
- 4.6. Coeficiente de correlación de Kendall τ para sustituciones desde uno a cuatro átomos de aluminio. En el cuadro a) se muestra el caso isoelectrónico donde los puntos rojos corresponde a mutaciones que sólo incluye el término electrostático y los puntos azules corresponden a mutaciones que incluyen tanto los efectos electrostáticos como los de polarización. En el cuadro b) se muestra el caso no-isoelectrónico, donde los puntos rojos corresponden a mutaciones que incluyen las contribuciones electrostática y de polarización, y los puntos azules también incluyen los efectos de transferencia electrónica. 52

5.1. Representación de la transformación alquímica adiabática del metanol ($\lambda = 0$) al anión metóxido ($\lambda = 1$).	55
5.2. La N-metiletanamina es una amina secundaria donde el nitrógeno es capaz de aceptar un protón.	59
5.3. El ácido propanoico es un ácido carboxílico monoprótico capaz de donar el protón hidrógeno unido al oxígeno.	59
5.4. El pirrol es un base débil en la posición N-H.	60
5.5. Comparación entre los valores de PA (en kcal/mol) <i>ab initio</i> versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para el grupo de las aminas.	63
5.6. Comparación entre los valores de PA (en kcal/mol) <i>ab initio</i> versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para el grupo de los ácidos carboxílicos.	64
5.7. Comparación entre los valores de PA (en kcal/mol) <i>ab initio</i> versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para el grupo de las moléculas aromáticas.	64
5.8. Comparación entre los valores de PA (en kcal/mol) <i>ab initio</i> versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para todo el conjunto de moléculas monopróticas: aminas, ácidos carboxílicos y aromáticas.	65
5.9. Uracilos con dos sitios de deprotonación: N ₁ y N ₃ . En la posición 5 los sustituyentes son X=CH ₃ , H, F, CHO y NO ₂	66
5.10. Moléculas polipróticas con tres o más sitios de deprotonación: a) adenina, b) guanina y c) citosina.	66

- 5.11. Comparación de los valores de afinidad protónica *ab initio* (PA_{B3LYP}) versus alquímicos ($PA_{\text{alquímico}}$) (en kcal/mol) para un conjunto de moléculas polipróticas: uracilos, adenina, guanina y citosina. 69
- 5.12. Comparación de los valores de afinidad protónica relativa *ab initio* ($PA_{\text{B3LYP}}^{\text{relativa}}$) versus alquímicos ($PA_{\text{alquímico}}^{\text{relativa}}$) (en kcal/mol) para un conjunto de moléculas polipróticas: uracilos, adenina, guanina y citosina. 70
- 5.13. Cambio en la energía de la transformación adiabática del metanol al anión metóxido. La curva negra (verde) corresponde al ajuste cuadrático (cúbico) de los puntos de la transformación alquímica adiabática. La curva azul corresponde a los valores de las derivadas alquímicas evaluadas en $Z = 1$. La línea roja punteada corresponde a la energía del anión metóxido relajado. 73

Índice de tablas

3.1. Términos diagonales de la matriz hessiana χ calculados con CCSD(CFD), HF(CFD), HF(IPA), HF(RPA) y CPHF. Los elementos de χ está en E_h ($1E_h = 27.2114$ eV).	33
3.2. Términos diagonales de la matriz hessiana χ calculados con CCSD(CFD), BLYP(CFD), BLYP(IPA), BLYP(RPA) y BLYP(CPKS). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).	34
3.3. Términos diagonales de la matriz hessiana χ calculados con CCSD(CFD), B3LYP(CFD), B3LYP(IPA), B3LYP(RPA) y B3LYP(CPKS). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).	35
5.1. Comparación entre las entalpías de disociación del protón, PA, (en kcal/mol, $1\text{eV}=23.06$ kcal/mol) obtenidas con derivadas alquímicas, B3LYP/def2-TZVPPD y valores experimentales para un conjunto de aminas, moléculas aromáticas y ácidos carboxílicos.	62
5.2. Comparación entre las energías de disociación del protón (en kcal/mol) obtenidas con derivadas alquímicas, B3LYP/def2-TZVPPD y experimentales para un conjunto de pirimidinas y purinas.	68

RESUMEN

Uno de los grandes retos actuales en ciencia de materiales es el desarrollo racional de compuestos, es decir, establecer protocolos experimentales y teóricos que permitan el diseño de materiales con propiedades optimizadas para aplicaciones específicas. Desde el punto de vista teórico, y en particular de la estructura electrónica, el reto es enorme. Para evidenciar esto es suficiente notar la vastedad del “espacio químico”, es decir, el conjunto de compuestos estables que hipotéticamente pueden ser creados a partir de los elementos de la tabla periódica. Estimaciones conservadoras de tan sólo un subconjunto de posibles moléculas orgánicas pequeñas, llevan a concluir que este espacio contiene muchísimo más que 10^{60} compuestos. Este proyecto propone una alternativa eficiente para explorar el espacio químico. En particular, se implementará un método perturbativo de la energía de un sistema electrónico, con el cual poder estudiar y optimizar propiedades de *clusters* atómicos y moléculas. Mediante el uso de transformaciones alquímicas, las cuales se definen como cualquier proceso real o ficticio en el que un sistema cambia su estequiometría, se estudiará la estabilidad de *clusters* heteroatómicos, como también comprender los mecanismos que promuevan la deprotonación en moléculas orgánicas.

ABSTRACT

One of the current major challenges in materials science is the rational design of compounds, that is, to establish experimental and theoretical protocols that allow the design of materials with properties optimized for specific applications. From a theoretical point of view, and in particular from the electronic structure, the challenge is enormous. To demonstrate this it is enough to note the vastness of the chemical space, that is, the set of stable compounds that hypothetically can be created from the elements of the periodic table. Conservative estimates of only a subset of possible small organic molecules, lead to conclude that this space contains much more than 10^{60} compounds. This work aims to provide an efficient alternative to explore the chemical space. In particular, we will introduce a perturbative method to capture changes of the electronic energy of the system associated to changes in its chemical composition. This method will be used to understand and optimize specific properties of atomic clusters and organic molecules. Using alchemical transformations, which are defined as any real or fictitious process in which a system changes its stoichiometry, we will explore the stability of heteroatomic clusters, as well as possible mechanisms that promote deprotonation in organic molecules.

Capítulo 1

Introducción

Los métodos *ab initio* en química cuántica y en ciencia computacional de materiales pueden ser notablemente precisos para reproducir propiedades de compuestos tales como energías relativas, reactividad, parámetros espectroscópicos, propiedades magnéticas y ópticas, etc. Esta precisión es crucial en el uso de métodos de química cuántica para diseñar compuestos con propiedades específicas. Tradicionalmente, el diseño racional de compuestos comienza con diferentes niveles de selección de grandes bibliotecas, reales o virtuales, de compuestos. Normalmente esta selección está guiada por reglas de propiedades estructurales cuyo objetivo es reducir el número de compuestos candidatos a optimizar la propiedad deseada. Para el éxito de este enfoque, es fundamental que la biblioteca inicial sea lo suficientemente diversa para contener compuestos con las propiedades deseadas y que la selección sea de alto rendimiento [1]. Debido a que las bibliotecas pueden contener millones de compuestos, los métodos *ab initio* se usan normalmente en la última etapa de refinamiento de la búsqueda. Sería ideal que toda la biblioteca se pudiera evaluar desde el principio con métodos *ab initio* en lugar de reglas empíricas. Esto, por ejemplo, evitaría que buenos candidatos sean descartados en los pasos iniciales de la selección. Sin embargo, el costo computacional de los métodos *ab initio* y el tamaño de las bibliotecas

descartan esta posibilidad. La biblioteca *ab initio* más grande de la que tenemos conocimiento tiene 134000 moléculas con hasta nueve átomos pesados [2]; un logro sorprendente, pero lejos de los millones de compuestos que requiere una biblioteca diversa [3].

El diseño inverso de compuestos aborda el problema de manera diferente haciendo la siguiente pregunta: dada la propiedad objetivo, ¿cuál es el compuesto que la optimiza? El punto de partida de la búsqueda es un conjunto de pocos compuestos con propiedades similares a las deseadas [4]. El proceso de optimización debería ser capaz de introducir modificaciones que den como resultado nuevos compuestos con propiedades más cercanas a las objetivo [5–11]. Esto requiere la habilidad de establecer de alguna manera una relación inversa estructura-propiedad, es decir, es necesario saber cómo una variación estructural o estequiométrica cambia la propiedad objetivo. Los métodos *ab initio* proporcionan un escenario natural para calcular estos cambios a través de teoría de perturbaciones. Varias estrategias en esta dirección han demostrado que el diseño inverso *ab initio* es posible con capacidades computacionales limitadas [12–15]. En este trabajo de investigación nos centraremos en la capacidad de las derivadas alquímicas para navegar el aspecto energético del llamado espacio químico (CS, por su sigla en inglés *Chemical Space*). Aunque existen varias definiciones del espacio químico, adoptamos una que parece más apropiada para su exploración con métodos *ab initio* de compuestos no solamente orgánicos [8, 16, 17]. En esta tesis, el CS se define como el espacio de los compuestos electrónicamente estables que se pueden crear a partir de los elementos de la tabla periódica. Una advertencia es necesaria, esta definición no implica que el espacio químico sea un espacio vectorial, sin embargo, cierta matematización es posible. La ecuación de Schrödinger de un sistema con N electrones y N_Z protones distribuidos en átomos con núme-

ros atómicos Z_1, Z_2, \dots, Z_M ubicados en las posiciones $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M$ está dada (en unidades atómicas, $\hbar = e = m_e = 4\pi\epsilon_0 = 1$) por

$$\left(-\sum_{i=1}^N \frac{\nabla_i^2}{2} - \sum_i^N \sum_{\beta}^M \frac{Z_{\beta}}{|\mathbf{r}_i - \mathbf{R}_{\beta}|} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \Psi = E\Psi. \quad (1.1)$$

Entonces, el CS está dado por el conjunto de todas las estructuras moleculares en las cuales los electrones están ligados al potencial de los núcleos atómicos,

$$\text{CS} = \{ \{ \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M \}, \{ Z_1, Z_2, \dots, Z_M \}, N, N_Z : \Psi \text{ es un estado ligado} \}. \quad (1.2)$$

Cualquier transformación que conecte dos puntos en el CS se llama transformación alquímica ya que lo que cambia en el proceso es la estequiometría del compuesto inicial [8, 18, 19]. Las transformaciones alquímicas pueden corresponder a procesos físicos, como la deprotonación, o fenómenos ficticios, como el cambio adiabático del número atómico de un átomo dentro de una molécula. La inmensidad del CS es, sin embargo, abrumadora. Estimaciones conservadoras de un sector del CS que comprende solo moléculas pequeñas (menos de veinte átomos) establece un límite inferior de 10^{60} compuestos [3, 20]. Los números se vuelven más y más grandes si se consideran otros sectores del espacio químico. Esto es especialmente cierto si se consideran sectores de compuestos donde las reglas de la química no pueden ser usadas como guías constructivas, este es el caso, por ejemplo, de los *clusters* metálicos [21, 22]. Por lo tanto, el desarrollo de herramientas para navegar eficientemente el CS es una necesidad. En esta tesis nos centramos en una aproximación perturbativa para realizar transformaciones alquímicas basadas en “derivadas alquímicas”, es decir, derivadas de la energía con respecto a cambios en la composición. Las derivadas/transformaciones alquímicas ya han sido usadas para explicar el enlace químico, reactividad, información de difracción de rayos X, estabilidad de las fases cristalinas y más [19, 23–30].

Aquí nos enfocaremos en cómo las derivadas alquímicas son capaces, o no, de predecir las potenciales estructuras estables de *clusters* semimetálicos dopados por sustitución. A pesar del interés actual en el estudio experimental y teórico de *clusters* y sus aplicaciones en áreas tan diversas como la nanociencia, la química, el diseño de fármacos, ciencia de materiales, medicina, etc. [31–35], elegimos este tipo de sistemas porque las características de su estructura electrónica plantean un desafío para las derivadas alquímicas.

Otro tipo de transformaciones en donde existe interés en estudiar la capacidad predictiva de las derivadas alquímicas es el proceso de deprotonación en moléculas orgánicas. La deprotonación de moléculas orgánicas es un proceso clave involucrado en muchos procesos químicos y biológicos [36,37]. Su mecanismo de reacción depende de varios factores tales como la naturaleza del grupo funcional ácido (carboxilo, hidroxilo, etc.), los sustituyentes de la molécula y el tipo de solvente [38]. Sin embargo, a pesar de su importancia, el mecanismo de estos procesos aún no se comprende bien. En este trabajo de tesis analizaremos si las derivadas alquímicas son capaces de predecir el sitio más fácil de deprotonación de un conjunto de moléculas polipróticas, es decir, moléculas que son capaces de ceder más de un protón desde diferentes sitios.

Esta tesis se desarrolla de la siguiente manera. En el capítulo 2 se presenta una breve reseña de los conceptos teóricos necesarios para el desarrollo de los métodos que se usan en esta investigación. En el capítulo 3 se presenta el formalismo de diferentes estrategias para calcular las derivadas alquímicas, en particular, mostramos como estas dependen del modelo de la densidad electrónica y del tipo de aproximación de la función de respuesta lineal utilizada en su evaluación. En el capítulo 4 se evalúa cómo las derivadas alquímicas son capaces de predecir estructuras estables de *clusters* semi-metálicos de aluminio dopados por sustitución con silicio. En el capítulo

5 se muestra la capacidad predictiva de las derivadas alquímicas en el proceso de deprotonación de un conjunto de moléculas polipróticas. Finalmente en el capítulo 6 se discuten los resultados principales de esta tesis, que se pueden resumir en que las derivadas alquímicas *ab initio* son una herramienta poderosa y libre de empirismo para explorar el espacio químico y procesos complejos que suceden en él.

Capítulo 2

Antecedentes bibliográficos

A continuación se revisan algunos conceptos básicos necesarios para el desarrollo de los métodos que se usan en esta tesis. Primero, se realiza una breve reseña de la Teoría del Funcional de la Densidad y luego se define en más detalle las transformaciones y derivadas alquímicas.

2.1. Teoría del Funcional de la Densidad

La Teoría del Funcional de la Densidad (DFT) [39] se basa en los teoremas de Hohenberg y Kohn (HK) [40]. El primer teorema establece que la densidad electrónica, $\rho(\mathbf{r})$, es suficiente para determinar la energía del estado fundamental de cualquier sistema electrónico

$$E[\rho(\mathbf{r})] = F[\rho(\mathbf{r})] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} , \quad (2.1)$$

donde $F[\rho(\mathbf{r})]$ es un funcional universal, en el sentido que es el mismo para todos los sistemas, $v(\mathbf{r})$ es el potencial externo debido a los núcleos y $\rho(\mathbf{r})$ es la densidad electrónica definida como

$$\rho(\mathbf{r}) = N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N) d\mathbf{r}_2 \dots d\mathbf{r}_N, \quad (2.2)$$

donde N es el número total de electrones y Ψ la función de onda del sistema. El funcional universal $F[\rho(\mathbf{r})]$ en la ecuación (2.1) es convenientemente dividido en diferentes contribuciones

$$F[\rho(\mathbf{r})] \equiv T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})], \quad (2.3)$$

donde $T[\rho(\mathbf{r})]$ es el funcional de energía cinética y $V_{ee}[\rho(\mathbf{r})]$ es la repulsión electrón-electrón, la cual incluye tanto efectos clásicos como cuánticos.

El segundo teorema de HK establece que para cualquier densidad $\tilde{\rho}(\mathbf{r})$ que proviene de un potencial externo físico y que es N -representable [41], es decir, que satisface las condiciones $\tilde{\rho}(\mathbf{r}) \geq 0$, $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$ y $\int |\nabla \tilde{\rho}(\mathbf{r})| d\mathbf{r} < \infty$, se cumple el siguiente principio variacional

$$E_0[\rho_0(\mathbf{r})] \leq E_v[\tilde{\rho}(\mathbf{r})], \quad (2.4)$$

donde $\rho_0(\mathbf{r})$ y E_0 son la densidad electrónica y la energía exacta del estado fundamental respectivamente, de un sistema con N electrones y potencial externo $v(\mathbf{r})$.

Los teoremas de HK no indican como construir explícitamente el funcional universal $F[\rho]$ que es desconocido. Muchos intentos se han hecho en construir un funcional explícito de la densidad [42, 43], sin embargo, todas las aproximaciones al funcional de energía cinética están lejos de ser satisfactorias. Luego, Kohn y Sham [44] idearon un procedimiento indirecto para encontrar la densidad electrónica que consiste en acudir a un sistema de referencia, el sistema s de Kohn y Sham (KS), de electrones que no interactúan, sujetos a un potencial efectivo, $v_{KS}(\mathbf{r})$ y que tiene la misma den-

sidad electrónica que el sistema real. En tal caso, el funcional universal se divide en varias componentes

$$F[\rho(\mathbf{r})] \equiv T_s[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] + E_{xc}[\rho(\mathbf{r})], \quad (2.5)$$

donde $T_s[\rho(\mathbf{r})]$ es la energía cinética de los electrones del sistema de KS, $J[\rho(\mathbf{r})]$ es la repulsión clásica entre electrones (Hartree), y $E_{xc}[\rho(\mathbf{r})]$ es un funcional desconocido de intercambio y correlación encargado de dar cuenta del principio de exclusión de Pauli, las correlaciones electrónicas y la inexactitud asociada a la energía cinética del sistema no interactuante. El resultado del procedimiento de KS es cambiar el problema de muchos cuerpos (N) en N ecuaciones monoeléctricas conocidas como las ecuaciones de Kohn y Sham

$$\left(-\frac{\nabla^2}{2} + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}') d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} + v_{xc}(\mathbf{r}) \right) \varphi_i(\mathbf{r}) = \varepsilon_i \varphi_i(\mathbf{r}) ,$$

de cuyos orbitales, $\varphi_i(\mathbf{r})$, se puede construir la densidad electrónica

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\varphi_i(\mathbf{r})|^2 ,$$

y donde $v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho}$ es el potencial desconocido de intercambio y correlación para el cual existen diversas aproximaciones [45]. Note que los orbitales de KS son en principio solo un sistema de referencia, y por tanto, en términos estrictos, ellos no constituyen una aproximación a la función de onda. Sin embargo, la práctica muestra que los orbitales de KS son muy parecidos a los orbitales canónicos de Hartree-Fock (HF) y que las diferencias entre autovalores de KS son aproximaciones a las energías de excitación.

2.1.1. Funciones de respuesta química en la Teoría del Funcional de la Densidad

En muchos procesos químicos y físicos uno está interesado en cuantificar la respuesta del sistema frente a perturbaciones externas. Si esa perturbación es una molécula que se presta a reaccionar con un sustrato, entonces habrán funciones de respuesta que cuantifican la reactividad del sustrato. El uso de estas funciones de respuesta en la comunidad de químicos constituye la llamada Teoría del Funcional de la Densidad Conceptual (DFT-Conceptual). Dado que el número de electrones N y el potencial externo $v(\mathbf{r})$ [39] determinan completamente el Hamiltoniano del sistema, una perturbación externa cualquiera altera estas dos variables. Frente a una variación de N y $v(\mathbf{r})$, el cambio en la energía del sistema puede ser evaluada mediante una expansión en series de Taylor de la energía electrónica $E[N, v(\mathbf{r})]$:

$$\begin{aligned}
\Delta E[N, v(\mathbf{r})] &= E[N + \Delta N, v(\mathbf{r}) + \Delta v(\mathbf{r})] - E[N, v(\mathbf{r})] \\
&= \left(\frac{\partial E}{\partial N} \right)_v \Delta N + \int d\mathbf{r} \left(\frac{\delta E}{\delta v(\mathbf{r})} \right)_N \delta v(\mathbf{r}) + \int d\mathbf{r} \frac{\delta^2 E}{\partial N \delta v(\mathbf{r})} \delta v(\mathbf{r}) \Delta N + \\
&\quad \frac{1}{2} \left(\frac{\partial^2 E}{\partial N^2} \right)_v (\Delta N)^2 + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \left(\frac{\delta^2 E}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} \right)_N \delta v(\mathbf{r}) \delta v(\mathbf{r}') + \\
&\quad \Delta Z_\alpha \sum_{\beta \neq \alpha} \frac{Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_\alpha|} + \sum_{l+n=3} O((\Delta N)^l (\delta v)^n). \tag{2.6}
\end{aligned}$$

La convergencia de esta serie ha sido discutida por Ayers *et al.* [46]. Los resultados empíricos muestran que una aproximación hasta segundo orden suele ser adecuada incluso para perturbaciones grandes. Como veremos más adelante, esta aproximación permite recuperar sobre el 99 % del cambio de la energía asociada a transmutación de átomos.

En la ecuación (2.6) se puede identificar cada coeficiente de la expansión como

funciones de respuesta del sistema frente a una perturbación externa. En el caso de una expansión a segundo orden dichas funciones de respuesta son: el potencial químico μ [47],

$$\left(\frac{\partial E}{\partial N}\right)_v = \mu, \quad (2.7)$$

que mide la tendencia de los electrones a escapar del sistema, la densidad electrónica $\rho(\mathbf{r})$

$$\left(\frac{\delta E}{\delta v(\mathbf{r})}\right)_N = \rho(\mathbf{r}), \quad (2.8)$$

la función de Fukui $f(\mathbf{r})$ [48]

$$\frac{\delta^2 E}{\delta N \delta v(\mathbf{r})} = \left(\frac{\partial \rho(\mathbf{r})}{\partial N}\right)_v = f(\mathbf{r}), \quad (2.9)$$

la cual indica en que regiones el sistema es propenso a aceptar o donar electrones. Por otro lado están la dureza química η [49]

$$\left(\frac{\partial^2 E}{\partial N^2}\right)_v = \eta, \quad (2.10)$$

que mide la resistencia de un sistema a transferir electrones, y la función de respuesta lineal $\chi(\mathbf{r}, \mathbf{r}')$ que cuantifica la polarización de la densidad electrónica

$$\left(\frac{\delta^2 E}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')}\right)_N = \left(\frac{\delta \rho(\mathbf{r})}{\delta v(\mathbf{r}')}\right)_N = \chi(\mathbf{r}, \mathbf{r}'). \quad (2.11)$$

Estas funciones de respuesta representan cuan sensible es un sistema frente a perturbaciones externas en el número de electrones N y en el potencial externo $v(\mathbf{r})$ (ver figura 2.1).

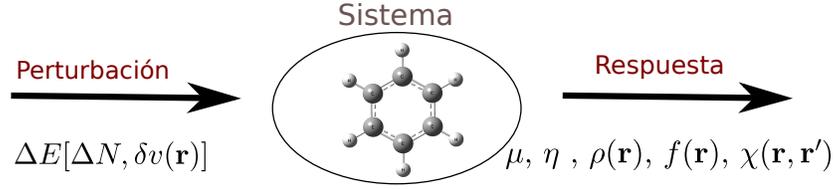


Figura 2.1: Representación de la respuesta de una molécula a una perturbación inducida por cualquier modificación en el número de electrones N y en el potencial externo $\delta v(\mathbf{r})$.

Si la perturbación es tal que representa un cambio en la naturaleza de alguno de los átomos del sistema, es decir, cambios en el N y el número atómico, entonces estas derivadas de la energía constituyen *derivadas alquímicas*, las cuales se pueden definir como cualquier función de respuesta que mide la sensibilidad de alguna propiedad del sistema con respecto a variaciones en su estequiometría. En este trabajo, consideramos de particular importancia las derivadas alquímicas que miden el cambio de la energía con respecto al número y tipo de átomos.

Introduciendo las ecuaciones (2.7-2.11) en (2.6), podemos reescribir la expresión para la expansión de la energía electrónica como

$$\begin{aligned} \Delta E[N, v(\mathbf{r})] = & \mu \Delta N + \int d\mathbf{r} \rho(\mathbf{r}) \delta v(\mathbf{r}) + \int d\mathbf{r} f(\mathbf{r}) \delta v(\mathbf{r}) \Delta N + \\ & \frac{1}{2} \eta (\Delta N)^2 + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \chi(\mathbf{r}, \mathbf{r}') \delta v(\mathbf{r}) \delta v(\mathbf{r}') + \\ & \Delta Z_\alpha \sum_{\beta \neq \alpha} \frac{Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_\alpha|} + \sum_{l+n=3} O((\delta N)^l (\delta v)^n). \end{aligned} \quad (2.12)$$

En la ecuación (2.12), los términos en azul (que sólo acompañan a ΔN) son los que determinan los efectos de transferencia de electrones. Los términos en rojo (que sólo acompañan a $\delta v(\mathbf{r})$) son los que determinan los efectos electrostáticos y

de polarización. Y por último, los términos cruzados, (que acompañan a ΔN y a $\delta v(\mathbf{r})$), determinan el acoplamiento entre la transferencia de electrones y los efectos electrostáticos.

2.2. Transformaciones alquímicas

Una transformación alquímica es cualquier proceso real o ficticio, mediante el cual un sistema cambia su estequiometría a través de cambios en el número y tipo de átomos y en el número de electrones. Por ejemplo, si uno se limitara a compuestos con 10 electrones y 10 protones, entre ellos se encontrarían la molécula de agua H_2O y el ácido fluorhídrico HF. El HF puede pensarse como una transformación en la cual un electrón y un protón del H_2O son transferidos al O. Sin embargo, una transformación en la cual un electrón y un protón “escapan” de uno de los átomos de hidrógeno del H_2O para formar un átomo de F no ocurre en el laboratorio [50]. Aún así, si la carga nuclear y el número de electrones se consideran parámetros adiabáticos del hamiltoniano, nada impide que el cambio de energía de esta transformación ficticia se pueda estimar con teoría de perturbaciones (ecuación 2.12).

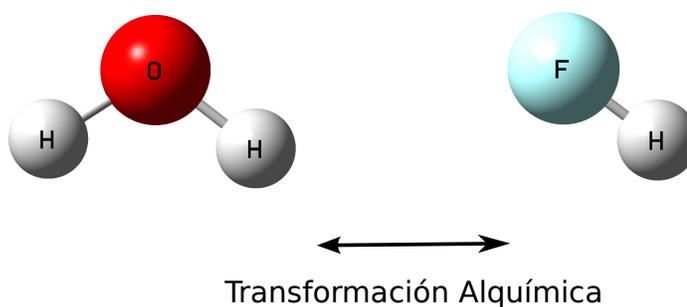


Figura 2.2: Transformación en la cual un electrón y un protón “escapan” de uno de los átomos de hidrógeno del H_2O para formar un átomo de F.

Existen diversos tipos de transformaciones alquímicas, y aunque su clasificación es arbitraria, en esta tesis serán importantes dos tipos de ellas:

1. **Transformaciones tipo I:** en la que cambia la naturaleza del átomo o grupo de átomos en un *cluster* o molécula de referencia. Un proceso al cual corresponde una transformación de este tipo es el dopaje sustitucional de *clusters* heteroatómicos a partir de *clusters* homoatómicos. A partir del *cluster* homoatómico estable, la mejor posición por sustitución puede ser obtenida optimizando la ecuación (2.15) sobre todas las posiciones atómicas con las restricciones adecuadas en $\{\Delta Z_\alpha\}$ y $\{\Delta N\}$. En la Figura 2.3, se muestra una representación del dopaje sustitucional de un *cluster* de Al_{13} con un átomo de Si. Note que en este caso la perturbación es tal que $\Delta N = 1$ y $\Delta Z_\alpha = 1$. Luego, la expresión dada en (2.15) para esta transformación, corresponde al cambio en la energía al agregar un electrón y un protón en la posición ocupada por algún átomo de Si.

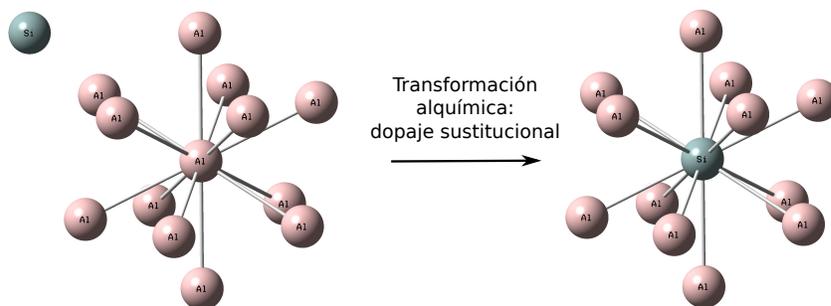


Figura 2.3: Transformación alquímica que corresponde al dopaje de un *cluster* de Al_{13} con un átomo de Si.

Como se mencionó anteriormente, los *clusters* de particular interés en esta tesis estarán compuestos de elementos del grupo principal de la tabla periódica.

No se considerarán metales de transición, ya que la descripción *all electron* de éstos es técnicamente difícil de abordar, pues suelen existir problemas de convergencia con el cálculo autoconsistente de la energía (*self-consistent field method*, SCF) para estados de carga nuclear fraccionaria.

2. **Transformaciones tipo II:** en la que se agrega o remueve un átomo o ion (con número atómico Z_α) en cierta posición \mathbf{R}_α . Este tipo de transformaciones pueden ser utilizadas para comprender, por ejemplo, posibles mecanismos de crecimiento de *clusters* [51]. Si un átomo es agregado a un *cluster*, el lugar más favorable \mathbf{R}_α^* entre todas las posiciones posibles, deberá ser el que minimiza el cambio de la energía, es decir:

$$\mathbf{R}_\alpha^* = \arg \min_{\mathbf{R}_\alpha} \Delta E(\mathbf{R}_\alpha). \quad (2.13)$$

Otra transformación alquímica de este tipo es la deprotonación en moléculas orgánicas [52]. Notar que si $\Delta N = 0$, $\Delta Z_\alpha = -1$ y \mathbf{R}_α es la posición del protón, entonces, la transformación dada por (2.15) corresponde a la eliminación del protón de un hidrógeno y representa una aproximación a la energía de deprotonación. Una aproximación porque dicha transmutación ignora el reordenamiento de la molécula y los efectos térmicos. Estas transformaciones están representadas en la figura 2.4.

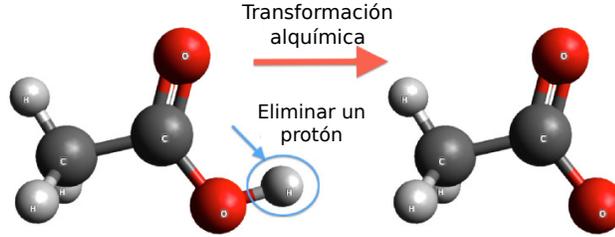


Figura 2.4: El proceso de deprotonación puede ser visto como una transformación alquímica en la cual el protón es removido mientras el número de electrones permanece constante.

Para ambos tipos de transformaciones alquímicas, el cambio de la energía hasta segundo orden en la perturbación cuando se cambia un átomo en la posición \mathbf{R}_α se obtiene de la ecuación (2.12). En este caso el cambio en el potencial externo es conocido

$$\delta v(\mathbf{r}) = -\frac{\Delta Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}|}, \quad (2.14)$$

donde $\Delta Z_\alpha = Z_\alpha^* - Z_\alpha$. Z_α^* es el número atómico del sistema perturbado y Z_α es el número atómico del sistema sin perturbar. Luego podemos escribir la expansión de la energía como

$$\begin{aligned} \Delta E[\mathbf{R}_\alpha] = & \mu \Delta N + \frac{1}{2} \eta (\Delta N)^2 - \Delta Z_\alpha \int d\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} - \Delta Z_\alpha \Delta N \int d\mathbf{r} \frac{f(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} + \\ & \frac{1}{2} (\Delta Z_\alpha)^2 \int d\mathbf{r} d\mathbf{r}' \frac{\chi(\mathbf{r}, \mathbf{r}')}{|\mathbf{R}_\alpha - \mathbf{r}| |\mathbf{R}_\alpha - \mathbf{r}'|} + \Delta Z_\alpha \sum_{\beta \neq \alpha} \frac{Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_\alpha|}. \end{aligned} \quad (2.15)$$

Por simplicidad, en la ecuación (2.15) se ha escrito la expresión para la energía al cambiar un solo átomo, pero la extensión para un grupo de átomos es directa.

Por ejemplo, al cambiar dos átomos i y j se tiene

$$\begin{aligned}
\Delta E[\mathbf{R}_\alpha] &= \mu\Delta N + \frac{1}{2}\eta(\Delta N)^2 - \int d\mathbf{r}\rho(\mathbf{r}) \left[\frac{\Delta Z_i}{|\mathbf{R}_i - \mathbf{r}|} + \frac{\Delta Z_j}{|\mathbf{R}_j - \mathbf{r}|} \right] - \\
&\Delta N \int d\mathbf{r}f(\mathbf{r}) \left[\frac{\Delta Z_i}{|\mathbf{R}_i - \mathbf{r}|} + \frac{\Delta Z_j}{|\mathbf{R}_j - \mathbf{r}|} \right] + \\
&\frac{1}{2} \int d\mathbf{r}d\mathbf{r}'\chi(\mathbf{r}, \mathbf{r}') \left[\frac{\Delta Z_i}{|\mathbf{R}_i - \mathbf{r}|} + \frac{\Delta Z_j}{|\mathbf{R}_j - \mathbf{r}|} \right] \left[\frac{\Delta Z_i}{|\mathbf{R}_i - \mathbf{r}'|} + \frac{\Delta Z_j}{|\mathbf{R}_j - \mathbf{r}'|} \right] + \\
&\Delta Z_i \sum_{\beta \neq i} \frac{Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_i|} + \Delta Z_j \sum_{\beta \neq j} \frac{Z_\beta}{|\mathbf{R}_\beta - \mathbf{R}_j|} + \frac{\Delta Z_\alpha \Delta Z_\beta}{|\mathbf{R}_i - \mathbf{R}_j|}. \quad (2.16)
\end{aligned}$$

2.3. Cálculo de las funciones de respuesta

Los términos globales en la ecuación (2.15) ($\mu\Delta N + \frac{1}{2}\eta(\Delta N)^2$) no son necesarios incluirlos si uno está interesado en comparar *clusters* con diferentes sitios de sustitución; sin embargo, se han incluido para evaluar cuan rápido converge la serie. No hay expresiones únicas para μ y η , éstas dependen del modelo elegido para la energía E en función del número de electrones N de un sistema abierto. El modelo más simple es la aproximación cuadrática de Parr [53], que establece las siguientes expresiones para μ y η :

$$\mu = -\frac{1}{2}(I + A) \quad (2.17)$$

y

$$\eta = I - A, \quad (2.18)$$

donde I y A son el potencial de ionización y electroafinidad verticales, respectivamente:

$$\begin{aligned}
I &= E_{N-1} - E_N \\
A &= E_N - E_{N+1},
\end{aligned}
\tag{2.19}$$

donde E_N es la energía del sistema neutro, E_{N-1} es la energía del catión y E_{N+1} la energía del anión, todos en la geometría del sistema con N electrones. Según el teorema de Hellmann-Feynman [54], la derivada de la energía E con respecto al número atómico Z_α es justamente la parte electrónica del potencial electrostático molecular, evaluado en la posición del núcleo α

$$\frac{\partial E}{\partial Z_\alpha} = - \int d\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} = \Phi_e(\mathbf{R}_\alpha).
\tag{2.20}$$

Similarmente, el término de transferencia electrónica es el negativo del potencial de Fukui [23, 55] en la posición del núcleo α

$$\frac{\partial^2 E}{\partial N \partial Z_\alpha} = - \int d\mathbf{r} \frac{f(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} = -v_f(\mathbf{R}_\alpha).
\tag{2.21}$$

Sin embargo, debido a que las funciones de Fukui son diferencias de densidad entre el sistema de interés y sus iones ($N \pm 1$), entonces la ecuación (2.21) también puede ser escrita en términos de la parte electrónica del potencial electrostático molecular. Por ejemplo, si la función de Fukui para incremento del número de electrones es $f^+(\mathbf{r}) = \rho_{N+1}(\mathbf{r}) - \rho_N(\mathbf{r})$, [56] entonces

$$\frac{\partial^2 E}{\partial N \partial Z_\alpha} = \Phi_e^{N+1}(\mathbf{R}_\alpha) - \Phi_e^N(\mathbf{R}_\alpha).
\tag{2.22}$$

El término de polarización es el más difícil de evaluar en la ecuación (2.15) ya que requiere la función de respuesta lineal exacta, pero en un cálculo DFT sólo se tiene acceso a la función de respuesta del sistema de KS. Sin embargo, es posible evaluar la función de respuesta exacta en términos de los orbitales de KS usando la teoría del funcional de la densidad tiempo dependiente (TD-DFT) [57, 58]. Sin embargo en este trabajo hemos visto que a menos que se ponga un gran esfuerzo en la implementación, el escalamiento de la evaluación de todas las integrales del tipo “potencial externo” que resultan cuando la ecuación (2.15) es escrita en la base de los orbitales de KS escala como $O(N^3)$ a $O(N^5)$, dependiendo del grado de aproximación en el *kernel* de intercambio y correlación. Por lo tanto, aquí hemos adoptado una forma alternativa que es evaluar $\frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta}$ como la primera derivada de la ecuación (2.20)

$$\frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta} = \int d\mathbf{r} d\mathbf{r}' \frac{\chi(\mathbf{r}, \mathbf{r}')}{|\mathbf{R}_\alpha - \mathbf{r}| |\mathbf{R}_\beta - \mathbf{r}'|} = \frac{\partial \Phi_e(\mathbf{R}_\beta)}{\partial Z_\alpha} \equiv \frac{\partial \Phi_\beta}{\partial Z_\alpha}. \quad (2.23)$$

2.4. Función de respuesta lineal $\chi(\mathbf{r}, \mathbf{r}')$

La función de respuesta lineal, o propagador de la densidad, definida como la segunda derivada funcional de la energía con respecto al potencial externo $v(\mathbf{r})$ y $v(\mathbf{r}')$, se puede entender también como la respuesta de la densidad electrónica (polarización) en un punto \mathbf{r} al cambiar el potencial externo en otro punto \mathbf{r}' [59]:

$$\chi(\mathbf{r}, \mathbf{r}') = \left(\frac{\delta^2 E}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} \right)_N = \left(\frac{\delta \rho(\mathbf{r})}{\delta v(\mathbf{r}')} \right)_N. \quad (2.24)$$

En un cálculo DFT del estado fundamental no se tiene acceso directo a $\chi(\mathbf{r}, \mathbf{r}')$ sino a la función de respuesta lineal del sistema no interactuante de Kohn-Sham $\chi_{\text{KS}}(\mathbf{r}, \mathbf{r}')$:

$$\chi_{\text{KS}}(\mathbf{r}, \mathbf{r}') = -2 \sum_i^{\text{occ}} \sum_a^{\text{virt}} \sum_\sigma (\varepsilon_{a\sigma} - \varepsilon_{i\sigma})^{-1} \phi_{i\sigma}^*(\mathbf{r}') \phi_{a\sigma}(\mathbf{r}') \phi_{i\sigma}(\mathbf{r}) \phi_{a\sigma}^*(\mathbf{r}), \quad (2.25)$$

donde ϕ son los orbitales de KS, los índices i, j, \dots hacen referencia a los orbitales moleculares ocupados (huecos), mientras que los índices a, b, \dots hacen referencia a los orbitales moleculares virtuales (partículas). Las letras griegas σ, τ , especifican el espín de cada electrón (α or β).

Sin embargo, la TD-DFT en su régimen lineal provee una ecuación tipo Dyson [60] que permite encontrar la función de respuesta lineal exacta en términos de $\chi_{\text{KS}}(\mathbf{r}, \mathbf{r}')$

$$\chi(\mathbf{r}, \mathbf{r}')^{-1} = \chi_{\text{KS}}(\mathbf{r}, \mathbf{r}')^{-1} - \frac{1}{|\mathbf{r} - \mathbf{r}'|} - f_{\text{xc}}(\mathbf{r}, \mathbf{r}'), \quad (2.26)$$

donde $f_{\text{xc}}(\mathbf{r}, \mathbf{r}') = \delta v_{\text{xc}}(\mathbf{r}) / \delta \rho(\mathbf{r}')$ es el *kernel*¹ de intercambio y correlación. Note que en esta ecuación la inversa es una inversa funcional, es decir,

$$\int \chi(\mathbf{r}, \mathbf{r}')^{-1} \chi(\mathbf{r}', \mathbf{r}'') d\mathbf{r}' = \delta(\mathbf{r} - \mathbf{r}''). \quad (2.27)$$

Otra forma conveniente de expresar la función de respuesta lineal $\chi(\mathbf{r}, \mathbf{r}')$ es expandirla en productos de excitaciones partícula-hueco (orbitales virtuales - orbitales ocupados)

$$\chi(\mathbf{r}, \mathbf{r}') = -2 \sum_{ia\sigma, jb\tau} (\mathbb{M}^{-1})_{ia\sigma, jb\tau} (\phi_{i\sigma}(\mathbf{r}) \phi_{a\sigma}^*(\mathbf{r})) (\phi_{j\tau}^*(\mathbf{r}') \phi_{b\tau}(\mathbf{r}')), \quad (2.28)$$

¹En este trabajo de tesis entenderemos el anglicismo *kernel* como una función generadora y como el integrando de la representación integral de la acción de un operador.

donde \mathbb{M}^{-1} es una matriz de respuesta de la matriz de densidad del sistema. Es decir, en la base de orbitales de KS, un cambio en la matriz de densidad \mathbb{P} , debido a una perturbación en el potencial $\delta\mathbb{V}$ está dada por

$$\delta\mathbb{P} = -\mathbb{M}^{-1}\delta\mathbb{V}. \quad (2.29)$$

La estructura de la matriz \mathbb{M} dependerá del nivel de aproximación que se use para los términos del lado derecho de la ecuación (2.26). Algunas de esas aproximaciones para $\chi(\mathbf{r}, \mathbf{r}')$ [61] se revisan a continuación.

- *Independent Particle Approximation (IPA)*: Consiste en aproximar la función de respuesta a la función de respuesta no interactuante de KS [62, 63]. Esta aproximación es equivalente a igualar la variación del potencial externo a la del potencial de KS

$$\delta v_{\text{KS}} = \delta v_{\text{ext}} + \cancel{\delta v_{\text{J}}} + \cancel{\delta v_{\text{xc}}}, \quad (2.30)$$

es decir, despreciar los efectos de intercambio y correlación y la interacción electrostática entre partícula y hueco. En el lenguaje de materia condensada es asumir que la energía de enlace partícula-hueco es nula.

Luego, en la aproximación IPA las derivadas de segundo orden en el potencial externo son

$$\begin{aligned} \frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta} &= \int d\mathbf{r}d\mathbf{r}' \left(\frac{\delta^2 E}{\delta v(\mathbf{r})\delta v(\mathbf{r}')} \right) \frac{\partial v(\mathbf{r})}{\partial Z_\alpha} \frac{\partial v(\mathbf{r}')}{\partial Z_\beta} \\ &\approx \int d\mathbf{r}d\mathbf{r}' \chi_{\text{KS}}(\mathbf{r}, \mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{R}_\alpha|} \frac{1}{|\mathbf{r}' - \mathbf{R}_\beta|}, \end{aligned} \quad (2.31)$$

que escritas en términos de los orbitales moleculares de KS toman la siguiente forma

$$\frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta} = -2 \sum_i^{occ} \sum_a^{virt} \sum_\sigma (\epsilon_{a\sigma} - \epsilon_{i\sigma})^{-1} \left(\int d\mathbf{r} \frac{\phi_{i\sigma}(\mathbf{r}) \phi_{a\sigma}^*(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_\alpha|} \right) \left(\int d\mathbf{r}' \frac{\phi_{i\sigma}^*(\mathbf{r}') \phi_{a\sigma}(\mathbf{r}')}{|\mathbf{r}' - \mathbf{R}_\beta|} \right). \quad (2.32)$$

Note que la ecuación (2.32) sólo contienen integrales del tipo “potencial externo” sobre los orbitales moleculares, las cuales tienen solución analítica en una base gaussiana [64–66].

- *Random Phase Approximation* (RPA): es importante notar que no hay una sola convención en la literatura para lo que se nombra como aproximación RPA. Aquí llamaremos aproximación RPA de la función de respuesta lineal $\chi(\mathbf{r}, \mathbf{r}')$ [63, 67] al caso en que se desprecian los efectos de intercambio y correlación pero se incluye el potencial electrostático partícula-hueco, es decir

$$\delta v_{KS} = \delta v_{\text{ext}} + \delta v_J + \cancel{\delta v_{xc}}.$$

En tal caso los elementos de la matriz \mathbb{M} son

$$M_{ia\sigma, jb\tau} = (\epsilon_{a\sigma} - \epsilon_{i\sigma}) \delta_{ij} \delta_{ab} \delta_{\sigma\tau} + 2 \langle ij|ab \rangle, \quad (2.33)$$

donde $\langle ij|ab \rangle$ son las integrales de repulsión electrónica, dadas por

$$\langle ij|ab \rangle = \int d\mathbf{r} d\mathbf{r}' \frac{\phi_i^*(\mathbf{r}) \phi_a(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_b(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.34)$$

Es importante anotar que en el caso de una función de onda HF, la aproximación RPA es equivalente a resolver las de HF dependientes del tiempo en el régimen lineal de la perturbación.

- *Coupled Perturbed Hartree-Fock* (CPHF): incluye la variación en el potencial externo de Coulomb y además incluye la variación del potencial de intercambio exacto de la aproximación de Hartree Fock. Esta aproximación es equivalente a calcular la función de respuesta lineal en el régimen de teoría de perturbaciones de Moller-Plesset de segundo orden [68]. En esta aproximación los elementos de la matriz \mathbb{M} son

$$M_{ia\sigma,jb\tau} = (\epsilon_{a\sigma} - \epsilon_{i\sigma})\delta_{ij}\delta_{ab}\delta_{\sigma\tau} + 2\langle ij|ab\rangle - \langle ij|ba\rangle - \langle ia|jb\rangle. \quad (2.35)$$

- *Coupled Perturbed Kohn-Sham* (CPKS): incluye el efecto de los cambios en el potencial externo sobre el potencial Coulombiano y el potencial de intercambio y correlación en la aproximación de KS. Este método es equivalente a calcular la función de respuesta lineal en el régimen de respuesta lineal de la TD-DFT. En principio es un método exacto y el único error es aquel cometido en la aproximación del *kernel* de intercambio y correlación. Los elementos de la matriz \mathbb{M} en este caso son

$$M_{ia\sigma,jb\tau} = (\epsilon_{a\sigma} - \epsilon_{i\sigma})\delta_{ij}\delta_{ab}\delta_{\sigma\tau} + 2\langle ij|ab\rangle + \langle ij|f_{xc}|ba\rangle. \quad (2.36)$$

Capítulo 3

Evaluación de las derivadas alquímicas

En el capítulo anterior hemos mencionado que el término más complejo de evaluar en la ecuación (2.15) es el término de polarización o función de respuesta lineal $\chi(\mathbf{r}, \mathbf{r}')$ ya que requiere la función de respuesta lineal exacta, pero en un cálculo DFT sólo se tiene acceso a la función de respuesta del sistema de KS, $\chi_{\text{KS}}(\mathbf{r}, \mathbf{r}')$. También hemos visto que existen distintos tipos de aproximaciones para la función de respuesta lineal: IPA, RPA, CPHF y CPKS, en donde cada uno de ellos incluyen distintos grados de correlación electrónica en la función de respuesta del sistema no interactuante. Una de las principales interrogantes que surge a partir de esto es cómo medir el efecto del método de cálculo en la evaluación de las derivadas alquímicas. Una utilidad interesante sería obtener una parametrización simple de uno de los métodos que ofrezca la precisión necesaria para alguna aplicación práctica de diseño racional de compuestos con un número considerable de átomos que excluya la posibilidad de usar métodos altamente correlacionados como lo son los métodos post-HF (CC, CI, CASSCF, etc.)

En este capítulo analizaremos cómo las derivadas alquímicas dependen del grado de aproximación de la función de respuesta lineal utilizada en su evaluación.

3.1. Aproximaciones IPA y RPA para la función de respuesta lineal.

La evaluación de la función de respuesta lineal se centra principalmente en la complejidad de las aproximaciones para calcularla. La primera aproximación para evaluar $\chi(\mathbf{r}, \mathbf{r}')$ es conocida como aproximación de partícula independiente (IPA) donde la variación del potencial de KS se iguala a la variación del potencial externo (ecuación 2.30), es decir, en esta aproximación se desprecian la interacción de Coulomb, el intercambio y la correlación entre los pares partícula-hueco. La aproximación IPA constituye la versión más simple de la función de respuesta lineal que se pueda tener dentro de un método monodeterminantal como DFT y HF. Un siguiente nivel de aproximación para $\chi(\mathbf{r}, \mathbf{r}')$ es la aproximación RPA de la función de respuesta lineal, $\chi_{\text{RPA}}(\mathbf{r}, \mathbf{r}')$. Es importante aclarar en este punto que no hay un acuerdo en la literatura de lo que se considera aproximación RPA [61, 63, 67]. En esta tesis llamaremos aproximación RPA a aquella que incluye la interacción de Coulomb entre el par partícula-hueco pero que desprecia los efectos de intercambio y correlación electrónica. En este capítulo analizaremos estas dos aproximaciones en distintos sistemas que cubren una gran diversidad en el tipo de enlace.

3.1.1. Moléculas diatómicas

Como primera instancia, se han calculado los términos diagonales y no diagonales de la matriz χ (hessiana) de segundas derivadas alquímicas, esto es $\chi_{\alpha\alpha} = \frac{\partial^2 E}{\partial Z_\alpha^2}$ y $\chi_{\alpha\beta} = \frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta}$ para un grupo de veinte moléculas diatómicas: LiH, BeH, CH, NH, OH, FH, SH, ClH, ClF, ClO, CN, CO, CS, LiF, LiN, NaCl, NO, OF, OS y SiO. Este grupo comprende un espectro amplio de tipos de enlace, desde muy iónico (NaCl) a covalente levemente polar (SiO). Primero se realizó la optimización de geometría de

estas moléculas utilizando el método post-HF conocido como *Coupled Cluster* [69,70] para incluir gran parte de la correlación electrónica dinámica. En breve, este método toma la función de onda de referencia (HF), y construye funciones de onda multi-electrónicas utilizando un operador de excitación para dar cuenta de la correlación entre los electrones. Dependiendo del orden de los operadores de excitación considerados, se da nombre al modelo *Coupled Cluster* (CC). En este trabajo hemos considerado operadores de excitaciones simples y dobles (CCSD). CCSD es conocido como el método *gold standard* de comparación para DFT y otros métodos *ab initio* que dan cuenta de la correlación dinámica.

La evaluación de la función de respuesta lineal a partir de los orbitales de KS se ha realizado con tres funcionales de intercambio y correlación distintos: un funcional de tipo GGA (*Generalized Gradient Approximation*) que combina el funcional de intercambio de Becke [71] con el funcional de correlación de Lee, Yang y Parr [72] conocido como BLYP; el funcional híbrido B3LYP [73,74] y su versión con correcciones de largo alcance, CAM-B3LYP [75]. Además hemos calculado las derivadas de segundo orden mediante diferencias finitas centrales (que llamaremos CFD, por su sigla en inglés *Central Finite Difference*) con un nivel de teoría CCSD. Para todos estos cálculos hemos utilizado el conjunto de funciones base def2-TZVPPD desarrollado por D. Rappoport y F. Furche específicamente para el cálculo de polarizabilidades [76]. Los cálculos de optimización de geometría se han realizado con el software de química computacional Gaussian09 [77] y los cálculos de la función de respuesta lineal tanto para la aproximación IPA como RPA se realizaron con ChemTools [78], que es un módulo dedicado al cálculo de funciones de respuesta química implementado en HORTON [79].

Evaluación de la matriz hessiana de segundas derivadas alquímicas, en la aproximación IPA

En las figuras 3.1 y 3.2 se han graficado los términos diagonales y no diagonales, respectivamente, de χ utilizando la aproximación IPA con los funcionales BLYP, B3LYP y CAM-B3LYP versus los valores encontrados mediante diferencias finitas (ecuación 4.1) usando el método CCSD, que son los valores de referencia. Para el caso de los términos diagonales $\chi_{\alpha\alpha}$ (ver Fig. 3.1), vemos que el funcional CAM-B3LYP provee la mejor descripción de la función de respuesta lineal no interactuante no solo en la dispersión de los datos sino en la subestimación de la magnitud de $\chi_{\alpha\alpha}$: hasta $\approx 80\%$ para BLYP, $\approx 57\%$ para B3LYP y $\approx 41\%$ para CAM-B3LYP. Esto se puede entender ya que mientras en los funcionales GGA, como BLYP, el potencial de intercambio decae exponencialmente en el régimen de largo alcance, el funcional CAM-B3LYP posee el decaimiento correcto que es proporcional a $-\frac{1}{r}$. Por tanto, los autovalores, $\{\epsilon_i\}$, de los orbitales CAM-B3LYP se esperan que sean más cercanos a los autovalores exactos que los autovalores de BLYP. De hecho, todos los funcionales GGA tienden a sobrestimar los niveles de energía de KS. Ahora, dado que en la aproximación IPA las verdaderas energías de excitación (los polos de $\chi(\mathbf{r}, \mathbf{r}')$) son aproximadas por diferencias entre las energías de los orbitales ocupados y desocupados de KS, es de esperar que la respuesta lineal en la aproximación IPA con un funcional como BLYP no solo esté muy subestimada sino que sea de peor calidad que la de un funcional con el decaimiento correcto del potencial de intercambio, como CAM-B3LYP. Por su parte, el funcional B3LYP aunque no tiene el decaimiento correcto del potencial de intercambio, posee una fracción (20%) global de intercambio de HF, de manera que sus autovalores se sitúan entre aquellos de BLYP y CAM-B3LYP. Cabe notar que dada la pobre correlación en los gráficos de la figura 3.1, los

resultados para BLYP y CAM-B3LYP pueden no tener una diferencia estadísticamente significativa.

En el caso de los términos no diagonales $\chi_{\alpha\beta}$ (ver Fig. 3.2), la aproximación IPA fracasa completamente, como puede verse de la absoluta falta de correlación entre los valores de referencia (CCSD) de $\chi_{\alpha\beta}$ y los obtenidos con todos los funcionales.

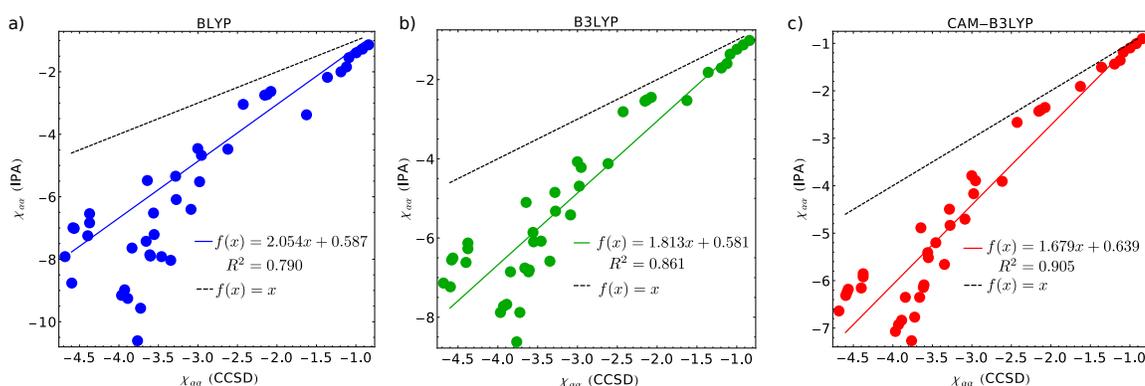


Figura 3.1: Comparación de los términos *diagonales* de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\alpha} = \frac{\partial^2 E}{\partial Z_\alpha^2}$) para un grupo de veinte moléculas diatómicas en la aproximación IPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).

Evaluación de la matriz hessiana de segundas derivadas alquímicas, en la aproximación RPA

Es sorprendente observar como la aproximación RPA permite una corrección excelente de los términos diagonales $\chi_{\alpha\alpha}$ (ver figura 3.3) para todos los funcionales con un desempeño levemente mejor para BLYP ($R^2 = 0.999$) que para B3LYP ($R^2 = 0.998$) y CAM-B3LYP ($R^2 = 0.997$). Es de notar que RPA tiene una leve tendencia a sobrestimar los valores de $\chi_{\alpha\alpha}$ como se nota de la pendiente del ajuste lineal en la figura 3.3. Esta sobrestimación es pequeña sin embargo, 7% para BLYP, 11%

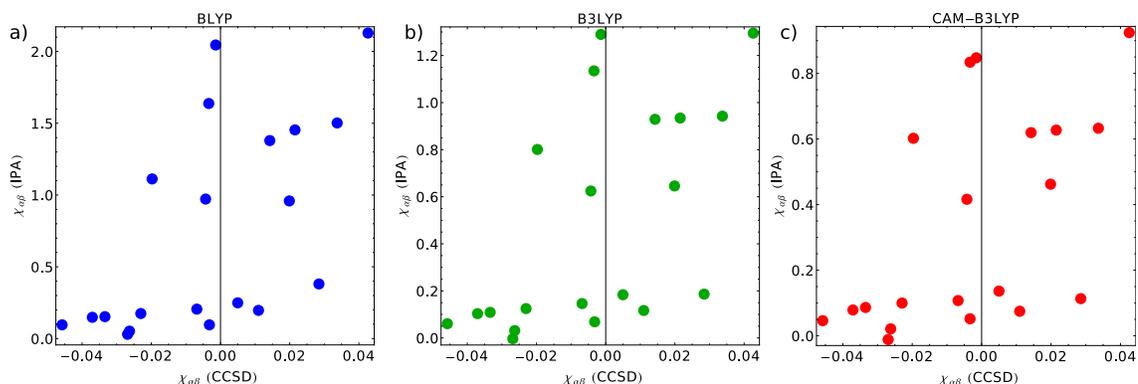


Figura 3.2: Comparación de los términos *no diagonales* de la de la matriz χ (hesiana) de segundas derivadas alquímicas ($\chi_{\alpha\beta} = \frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta}$) para un grupo de veinte moléculas diatómicas en la aproximación IPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. Los elementos de χ está en E_h ($1E_h = 27.2114$ eV).

para B3LYP y 13% para CAM-B3LYP. Las leves diferencias entre BLYP, B3LYP y CAM-B3LYP son quizás atribuibles a una cancelación de errores debido a la inclusión indirecta de integrales de intercambio en los funcionales híbridos. Las ecuaciones RPA, como las definimos acá, solo incluyen la interacción clásica de Coulomb entre el par partícula-hueco, sin embargo, en los funcionales híbridos siempre existe una fracción de intercambio de HF. En términos estrictos, para que la aproximación RPA sea justamente comparable entre funcionales GGA e híbridos, la fracción de intercambio en estos últimos debería ser descontada. Sin embargo, en la práctica, RPA funciona bastante bien con todos los funcionales para calcular los términos diagonales $\chi_{\alpha\alpha}$.

Es interesante notar que las figuras 3.3 a-c) hay un grupo de datos para valores más positivos de $\chi_{\alpha\alpha}$ que parecieran tener su propia tendencia. Una inspección revela que se trata de moléculas en las que uno de los átomos es H. Nuestra mejor explicación para esto es que, aunque hemos usado una base grande, esta está aún lejos del

límite de base infinita para el H, en especial cuando este está unido a átomos muy electronegativos como el F y el Cl.

Por su parte, los términos no diagonales $\chi_{\alpha\beta}$ son tan bien aproximados por RPA como los diagonales (ver figura 3.4a-c). Esto se esperaba para todos los métodos ya que los términos no diagonales miden respuestas no locales a una perturbación, por tanto demandan una función de respuesta lineal que se comporte correctamente en el régimen de grandes distancias $|\mathbf{r} - \mathbf{r}'| \gg 1$ ya que $\chi(\mathbf{r}, \mathbf{r}')$ no decae rápidamente con $|\mathbf{r} - \mathbf{r}'|$ [80]. Bajo estas consideraciones RPA es una aproximación satisfactoria para los elementos no diagonales de χ . Vemos de las figuras 3.4 a-c) que con CAM-B3LYP la correlación es menor ($R^2 = 0.933$) que con los otros dos funcionales, BLYP ($R^2 = 0.955$) y B3LYP ($R^2 = 0.958$). Además, de la regresión lineal realizada se encuentra que la pendiente para el caso de CAM-B3LYP es la que más se aleja de la unidad.

En resumen, la aproximación RPA con los funcionales BLYP y B3LYP reproducen bastante bien, comparadas con valores CCSD, las segundas derivadas alquímicas. Esto representa una tremenda reducción en el costo computacional de las derivadas alquímicas de segundo orden ya que la evaluación de la energía con el método CCSD escala como $O(N^6)$ mientras que RPA escala tan solo como $O(N^4)$. Además, la única forma práctica de evaluar derivadas alquímicas con métodos correlacionados basados en la función de onda (como CCSD) es mediante diferencias finitas, mientras que en la aproximación RPA todo el procedimiento es analítico y toda la matriz hessiana puede ser evaluada en un solo paso.

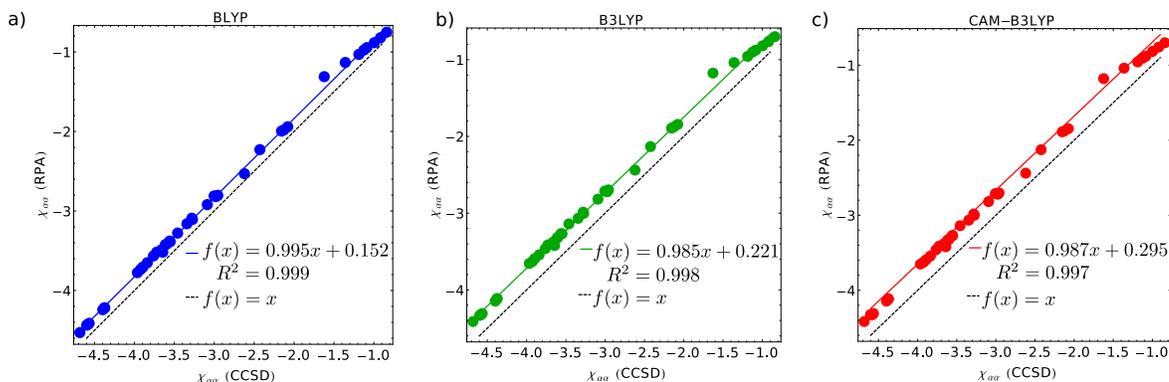


Figura 3.3: Comparación de los términos *diagonales* de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\alpha} = \frac{\partial^2 E}{\partial Z_\alpha^2}$) para un grupo de veinte moléculas diatómicas en la aproximación RPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ está en E_h ($1E_h = 27.2114$ eV).

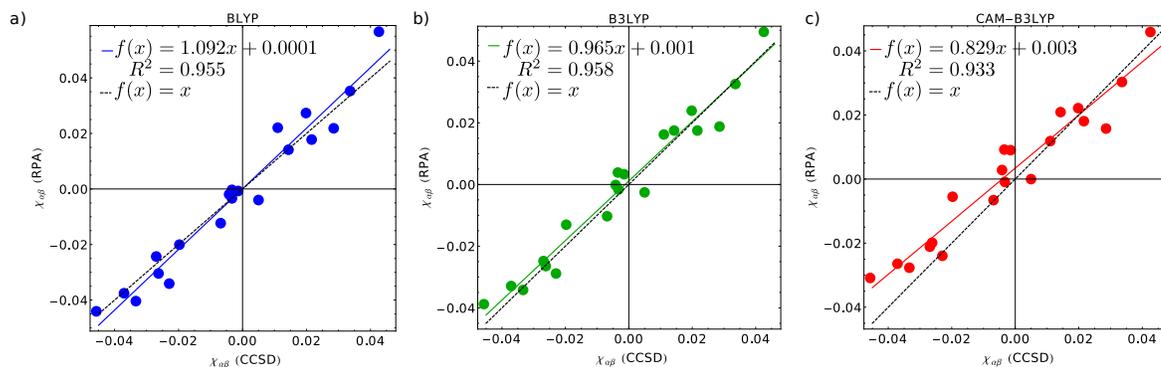


Figura 3.4: Comparación de los términos *no diagonales* de la de la matriz χ (hessiana) de segundas derivadas alquímicas ($\chi_{\alpha\beta} = \frac{\partial^2 E}{\partial Z_\alpha \partial Z_\beta}$) para un grupo de veinte moléculas diatómicas en la aproximación RPA versus CCSD(CFD) usando los funcionales a)BLYP, b)B3LYP y c)CAM-B3LYP. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ está en E_h ($1E_h = 27.2114$ eV).

3.2. Métodos CPHF y CPKS para la matriz hessiana de segundas derivadas alquímicas

Los siguientes niveles de aproximación para $\chi(\mathbf{r}, \mathbf{r}')$ están desarrollados dentro del marco de teoría de perturbaciones aplicadas a las ecuaciones de Hartree-Fock-

Rootan(CPHF) y de Kohn-Sham (CPKS) [81–83]. Como hemos visto en la sección 2.4, la aproximación CPHF incluye la variación en el potencial externo de Coulomb y además incluye la variación del potencial de intercambio exacto de la aproximación de Hartree-Fock. La misma teoría de perturbaciones aplicada a las ecuaciones de KS constituye el método CPKS, que incluye los efectos de correlación electrónica e intercambio según la DFT. En esta sección evaluaremos la matriz Hessiana de segundas derivadas químicas con las aproximaciones IPA, RPA, y los métodos CPHF y CPKS. Los métodos CPHF y CPKS también permiten una evaluación analítica de las derivadas de segundo orden mediante las ecuaciones (2.28), (2.35) y (2.36) y escalan formalmente como $O(N^5)$. De nuevo, el método con el cual se comparan las aproximaciones para $\chi_{\alpha\alpha}$ es CCSD con diferencias finitas.

3.2.1. Moléculas tetraédricas

En este caso hemos aumentado la complejidad de los sistemas y hemos escogido un conjunto de moléculas tetraédricas estudiadas previamente por Lesiuk *et. al* [57]. En ese estudio, Lesiuk *et. al* solo reportaron las derivadas de segundo orden con respecto a la carga nuclear del átomo central, es decir, el término diagonal $\chi_{\alpha\alpha}$. Nosotros nos limitaremos a un subconjunto de las moléculas estudiadas por Lesiuk (BH_4^- , CH_4 , NH_4^+ , BF_4^- , CF_4 , NF_4^+), ya que para el resto de moléculas fue imposible converger el ciclo autoconsistente en algunos casos (PF_6^- , SF_6 , ClF_6^+), y en otros varios de los estados ocupados corresponden a estados no ligados que preferimos dejar afuera (PO_4^{3-} , SO_4^{2-} , ClO_4^-). La segunda derivada química se ha evaluado entonces con las aproximaciones IPA y RPA y los métodos CPHF y CPKS utilizando HF y los funcionales BLYP y B3LYP, dado que en la sección anterior fueron los dos mejores para representar la función de respuesta lineal.

En la tabla 3.1 se muestran los resultados obtenidos con HF a través de diferencias finitas HF(CFD), y con las distintas aproximaciones HF(IPA), HF(RPA) y CPHF además de los resultados obtenidos con CCSD mediante diferencias finitas CCSD(CFD). El cálculo HF(CFD) ha sido realizado para verificar una vez más la calidad de las diferencias finitas. En el régimen cuadrático HF(CFD) debe coincidir con el método analítico CPHF; toda diferencia debe ser adjudicada al error cometido en usar diferencias finitas. Como es claro de la tabla 3.1, HF(CFD) coincide perfectamente con CPHF. Esto entonces nos da tranquilidad para confiar en CCSD(CFD) que es el método que hemos usado como patrón de comparación.

De la figura 3.5 se puede apreciar que la correlación lineal entre los valores de $\chi_{\alpha\alpha}$ encontrados con CPHF y CCSD(CFD) (puntos verdes) es sorprendentemente buena ($R^2 = 0.996$), con una mínima tendencia a sobrestimar $\chi_{\alpha\alpha}$ en un 0.04 %, Este buen acuerdo no es de extrañar ya que la teoría de perturbaciones de segundo orden aplicada a HF incluye efectos de correlación más los efectos de intercambio ya incluidos dentro de HF. Por su parte, la aproximación HF(IPA) (puntos azules) subestima $\chi_{\alpha\alpha}$ hasta en un 13 %, tal como en el caso de las moléculas diatómicas. Por el contrario, la aproximación HF(RPA) (puntos rojos) sobrestima $\chi_{\alpha\alpha}$ hasta en un 19 %. Note sin embargo que la subestimación de HF(IPA) no es homogénea (pendiente muy lejana de uno) en todo el rango de valores, como si lo es HF(RPA). A pesar de la (sub-)sobrestimación de CPHF, HF(RPA) y HF(IPA), los buenos coeficientes de correlación ($R^2 = 0.996$, $R^2 = 0.990$ y $R^2 = 0.988$ respectivamente) sugieren que es posible parametrizar aproximaciones computacionalmente “simples” como RPA para la exploración del espacio químico de compuestos realmente grandes, del orden de más de cien átomos.

Tabla 3.1: Términos diagonales de la matriz hessiana χ calculados con CCSD(CFD), HF(CFD), HF(IPA), HF(RPA) y CPHF. Los elementos de χ está en E_h ($1E_h = 27.2114$ eV).

Molécula	CCSD(CFD)	HF(CFD)	HF (IPA)	HF(RPA)	CPHF
BH_4^-	-2.775	-2.769	-3.016	-2.199	-2.769
CH_4	-3.006	-3.025	-3.480	-2.449	-3.025
NH_4^+	-3.246	-3.285	-3.967	-2.688	-3.285
BF_4^-	-2.590	-2.563	-2.628	-2.075	-2.563
CF_4	-2.934	-2.904	-3.187	-2.368	-2.904
NF_4^+	-3.330	-3.337	-4.055	-2.703	-3.337

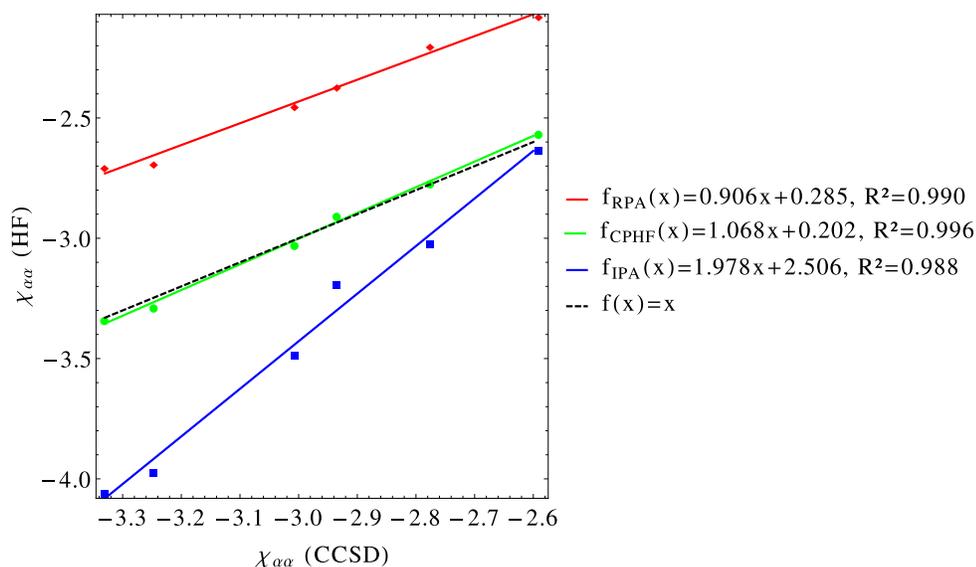


Figura 3.5: Comparación de los términos diagonales $\chi_{\alpha\alpha}$ para un grupo de moléculas tetraédricas para las distintas aproximaciones IPA (azul), RPA (rojo) y CPHF (verde) usando HF versus CCSD(CFD). La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).

Similarmente al caso de HF, las tablas 3.2 y 3.3 muestran los valores de $\chi_{\alpha\alpha}$ en las distintas aproximaciones para los funcionales BLYP y B3LYP. Nuevamente se corrobora que los valores obtenidos con el método analítico BLYP(CPKS) coinciden perfectamente con los encontrados mediante diferencias finitas BLYP(CFD). Por

su parte, de la figura 3.6 se aprecia que al igual que en HF, en DFT la aproximación IPA (puntos azules) subestima $\chi_{\alpha\alpha}$ pero ahora mucho más, 67% en BLYP(IPA) y 50% en B3LYP(IPA). Esto sugiere una vez más la importancia que tiene en la aproximación IPA el decaimiento correcto del potencial de intercambio. Por otro lado la aproximación RPA en DFT (puntos rojos) también sobrestima $\chi_{\alpha\alpha}$ pero en menor medida que HF, alrededor de 5% para BLYP(RPA) y de 8% para B3LYP(RPA). De nuevo, los coeficientes de correlación lineal de todas las aproximaciones son satisfactorios: BLYP(CPKS) ($R^2 = 0.999$), BLYP(IPA) ($R^2 = 0.956$), BLYP(RPA) ($R^2 = 0.995$), B3LYP(CPKS) ($R^2 = 0.999$), B3LYP(IPA) ($R^2 = 0.979$), B3LYP(RPA) ($R^2 = 0.996$). En particular la aproximación BLYP(RPA) introduce un pequeño error como había sido ya anticipado en la moléculas diatómicas.

Tabla 3.2: Términos diagonales de la matriz hessiana χ calculados con CCSD(CFD), BLYP(CFD), BLYP(IPA), BLYP(RPA) y BLYP(CPKS). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).

Molécula	CCSD(CFD)	BLYP(CFD)	BLYP (IPA)	BLYP(RPA)	BLYP(CPKS)
BH ₄ ⁻	-2.775	-2.795	-4.431	-2.622	-2.795
CH ₄	-3.006	-3.036	-5.019	-2.860	-3.036
NH ₄ ⁺	-3.246	-3.283	-5.712	-3.100	-3.283
BF ₄ ⁻	-2.590	-2.637	-3.608	-2.493	-2.638
CF ₄	-2.934	-2.978	-4.675	-2.818	-2.978
NF ₄ ⁺	-3.330	-3.358	-6.607	-3.182	-3.358

De este capítulo podemos concluir que la aproximación RPA es mejor que IPA para estimar la función de respuesta lineal, y por tanto χ , en el conjunto de moléculas estudiadas, independiente del método utilizado (HF, BLYP ó B3LYP). También se observa que en la aproximación IPA el método HF sobrestima menos los elementos

Tabla 3.3: Términos diagonales de la matriz hessiana χ calculados con CCSD(CFD), B3LYP(CFD), B3LYP(IPA), B3LYP(RPA) y B3LYP(CPKS). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).

Molécula	CCSD(CFD)	B3LYP(CFD)	B3LYP (IPA)	B3LYP(RPA)	B3LYP(CPKS)
BH_4^-	-2.775	-2.775	-4.000	-2.524	-2.789
CH_4	-3.006	-3.009	-4.561	-2.765	-3.032
NH_4^+	-3.246	-3.247	-5.194	-3.006	-3.282
BF_4^-	-2.590	-2.600	-3.321	-2.391	-2.617
CF_4	-2.934	-2.930	-4.200	-2.708	-2.955
NF_4^+	-3.330	-3.307	-5.712	-3.065	-3.344

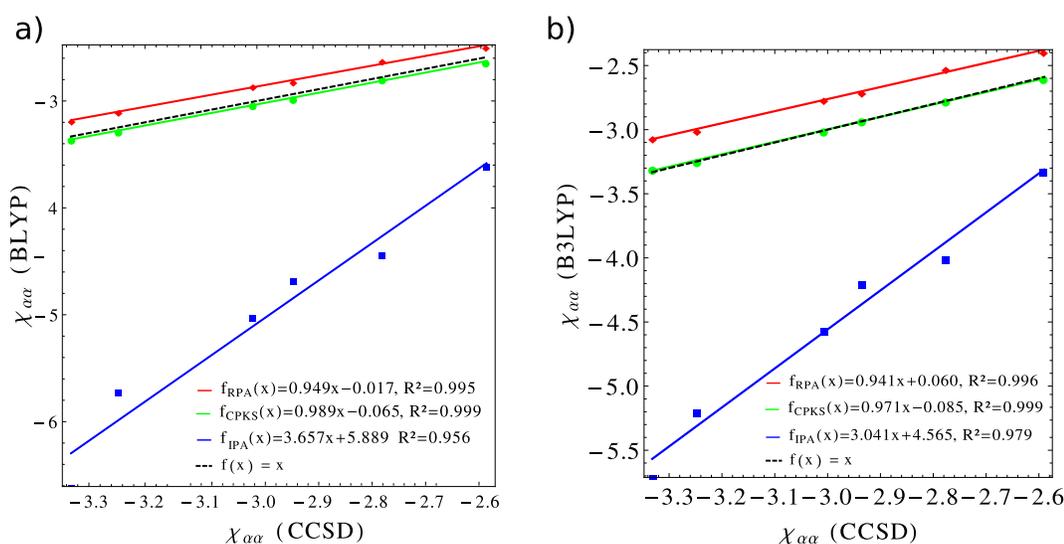


Figura 3.6: Comparación de los términos diagonales $\chi_{\alpha\alpha}$ para un grupo de moléculas tetraédricas para las distintas aproximaciones IPA (azul), RPA (rojo) y CPKS (verde) usando a) BLYP vs. CCSD y b) B3LYP vs. CCSD. La línea negra es una referencia visual de la recta hipotética de correlación perfecta ($f(x) = x$). Los elementos de χ están en E_h ($1E_h = 27.2114$ eV).

diagonales de χ que la misma aproximación dentro DFT (BLYP y B3LYP). Creemos que esto se debe a que en el método HF el potencial de Coulomb decae exactamente como $-1/r$ para largas distancias, mientras que en DFT los potenciales de la mayoría de los funcionales decaen exponencialmente. Parte de este problema se ve solucionado en el caso de los funcionales híbridos, como B3LYP, ya que incorpora una

fracción de intercambio exacto de HF en el funcional de intercambio. Finalmente, la aproximación RPA es la que mejor correlaciona con los valores encontrados con CCSD de referencia independiente de si se usa HF o DFT, aunque el mejor resultado se obtiene con RPA (BLYP) que formalmente escala solo como $O(N^4)$ mientras que CCSD escala como $O(N^6)$.

Capítulo 4

Estabilidad frente al dopaje sustitucional de un *cluster* semimetálico

El interés en el estudio experimental y teórico de las propiedades de *clusters* ha aumentado durante las últimas décadas en distintas áreas como la nanociencia, la química, el diseño de medicamentos, las ciencia de materiales, etc.; ello debido al potencial de aplicaciones tecnológicas y médicas que estos materiales poseen [84–88].

Determinar propiedades físicas y químicas de *clusters* es un problema complejo en la física de muchos cuerpos (electrones y núcleos). Es sabido que en *clusters*, las propiedades no escalan homogéneamente con el tamaño de éstos, sino que dependen fuertemente del número y tipo de átomos y de su estructura, la cual usualmente no puede ser determinada directamente con experimentos. Por lo tanto, la determinación de la estructura de un *cluster* es de gran importancia para comprender las relaciones entre estructura y propiedades [89, 90].

En el caso de *clusters* homoatómicos, sus propiedades son determinadas básicamente por el número y tipo de átomos que lo conforman. En el caso de *clusters* formados por dos o más tipos de átomos, llamados *clusters* heteroatómicos, la este-

quiometría agrega un grado de libertad adicional al estructural, haciendo aún más complejo el estudio de su estructura.

Se han desarrollado varios métodos para encontrar las estructuras estables de *clusters* tanto homoatómicos como heteroatómicos, entre ellos cabe destacar: algoritmo genético [5], *simulated annealing* [91], optimizaciones basadas en gradientes [8, 9, 92, 93], *basing hopping* [94], metadinámica [95] y Montecarlo [96, 97], por nombrar los más comunes. Estos métodos han contribuido de manera eficiente a la búsqueda de estructuras estables. Sin embargo, para predecir alguna propiedad del *cluster* o molécula, se debe realizar al menos un cálculo *single-point* por cada configuración de éstos, lo que generalmente exige una alta demanda computacional debido al costo de los de métodos *ab initio*.

Durante los últimos años, el paradigma en el estudio teórico de propiedades de compuestos ha cambiado. Ya no basta con poder calcular propiedades sino que también necesitan ser optimizadas: dada una cierta propiedad objetivo, ¿cuál es el compuesto que optimiza esta propiedad? [98]. Particularmente, la necesidad de poder crear nuevos materiales y compuestos con propiedades específicas, ha llevado al desarrollo de métodos y técnicas para explorar el espacio químico (CS) [99, 100], que en esta tesis definimos como el conjunto de todos los compuestos que pueden ser creados de los elementos de la tabla periódica. Debido a la vastedad del CS (mayor a 10^{60} compuestos [101]) el propósito de explorarlo y encontrar compuestos con propiedades optimizadas no es trivial. Luego, es claro que el desarrollo de herramientas que guíen la exploración de éste es importante.

Entender las transformaciones a lo largo del CS que implican cambios en la composición y la geometría de un sistema es posible utilizando funciones de respuesta química dentro de la DFT (sección 2.1.1). En esta tesis nos enfocaremos en la ca-

pacidad de las derivadas alquímicas para explorar el aspecto energético del espacio químico. Específicamente, en este capítulo nos centraremos en estudiar cómo las derivadas alquímicas son capaces de predecir estructuras estables de *clusters* semi-metálicos dopados por sustitución.

4.1. Transformación alquímica tipo I: dopaje en *clusters* heteroatómicos

Hemos descrito las transformaciones tipo I (sección 2.2) como una transformación en la cual cambia la naturaleza del átomo o grupo de átomos en un *cluster* o molécula de referencia. Esta transformación, en un *cluster*, puede explicar los cambios de estabilidad a través del dopaje sustitucional. Un sistema interesante de estudiar es el caso del *cluster* de Al_{13} dopado con átomos de Si (ver figura 4.1). Elegimos este sistema porque sus características de estructura electrónica representan un desafío para las derivadas alquímicas. En particular, Al_{13} es un *cluster* semimetálico con cierto carácter covalente y una direccionalidad de enlace considerable. Por lo tanto, cuando un átomo de Al se sustituye por uno de Si, no solo son importantes los efectos electrostáticos debidos al cambio en el número atómico, sino que se deben tener en cuenta los efectos de polarización y la transferencia de electrones. Además, debido al número atómico relativamente pequeño de Al ($Z = 13$), transmutarlo en Si ($Z = 14$) no es de ninguna manera una perturbación pequeña.

El *cluster* elegido presenta un reto adicional ya que es de capa abierta (grupo puntual S_6) y no el icosaedro más estable que corresponde el anión de capa cerrada Al_{13}^- [102–104]. La mayoría de los trabajos sobre transformaciones alquímicas se han realizado en sistemas sin polarización de espín. El *cluster* elegido es claramente un caso desafiante para las derivadas alquímicas ya que las exige al límite.

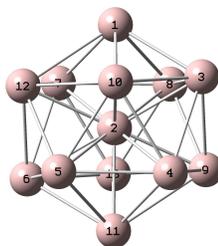


Figura 4.1: *Cluster* de Al_{13}

Cálculo de derivadas alquímicas

Para el cálculo de relajación de geometría del Al_{13} se utilizó el funcional de intercambio y correlación PBE [105] y un conjunto base del tipo Pople 6-311+g(d) [106, 107]. El cálculo de frecuencias vibracionales muestra que la estructura corresponde a un mínimo sobre la superficie de energía potencial. La estructura resultó ser un doblete con simetría S_6 . La energía de las estructuras dopadas de $\text{Al}_{13-n}\text{Si}_n$ se evaluó con el mismo método utilizado para optimizar el *cluster* de Al_{13} . Estos cálculos se realizaron con el software de química computacional Gaussian09 [77]. Las derivadas alquímicas fueron evaluadas usando las ecuaciones (2.20), (2.22) y (2.23). En estas ecuaciones, el MEP no se calculó a partir del teorema de Hellmann-Feynman, sino que se usaron gradientes analíticos de la energía para evitar el error de Pulay [108] asociado al conjunto de bases gaussianas centradas en los átomos. La derivada de la ecuación (2.23) es evaluada usando la aproximación de diferencias finitas centrales con precisión a cuarto orden, de la siguiente manera

$$\frac{\partial \Phi_\beta}{\partial Z_\alpha} = \frac{\frac{1}{12}\Phi_\beta(Z_\alpha - 2\Delta Z_\alpha) - \frac{2}{3}\Phi_\beta(Z_\alpha - \Delta Z_\alpha) + \frac{2}{3}\Phi_\beta(Z_\alpha + \Delta Z_\alpha) - \frac{1}{12}\Phi_\beta(Z_\alpha + 2\Delta Z_\alpha)}{\Delta Z_\alpha}. \quad (4.1)$$

Las derivadas alquímicas se calcularon en nuestro propio código y se verificaron usando la mayoría de las opciones internas (IOps) de Gaussian09 y una implementa-

ción reciente del módulo ChemTools [78] de HORTON [79], que está especialmente dedicado a calcular las funciones de respuesta química.

4.2. Transformación en la cual el número de electrones permanece constante.

Primero analizamos el caso más simple en el cual el número de electrones permanece constante durante la sustitución de átomos. Es decir, transformaciones en las que el *cluster* Al_{13} cambia a $\text{Al}_{13-n}\text{Si}_n^{+n}$. En estudios previos de Weigend *et al.* [109–112] en transformaciones de *clusters* de metales de transición, se argumentó que el término electrostático es suficiente para explicar el cambio de energía en la transformación. Encontraron buenas correlaciones entre el cambio de energía *ab initio* real y el cambio de energía alquímica en, por ejemplo, un *cluster* de Pt dopado con Ir. Notemos que esto se puede entender en términos de la magnitud de la perturbación y la naturaleza metálica del *cluster* de Pt. Si un átomo de Pt se muta en un átomo de Ir, la magnitud relativa de la perturbación es $\Delta Z/Z = 1/78$, mientras que en nuestro caso la perturbación es seis veces más fuerte $\Delta Z/Z = 1/13$. Si la magnitud relativa de la perturbación es pequeña, pueden despreciarse los términos de segundo orden. Además, si el sistema principal es genuinamente metálico, el *screening* electrónico (término de polarización) de todos los sitios del *cluster* será similar, sin diferenciar entre sustituciones en diferentes lugares. Claramente, este no es el caso de los *clusters* de Al sustituidos con Si porque la magnitud relativa de la perturbación no es pequeña y el *cluster* homoatómico original no es completamente metálico. Luego, nuestro cambio de energía alquímica incluye todos los términos de la ecuación (2.12) hasta segundo orden con $\Delta N = 0$.

En la figura 4.2 se muestra el cambio de energía alquímica vs. energía *ab initio*

para la sustitución de 1, 2, 3 y 4 átomos. Para evaluar la importancia del término de polarización, se ha graficado la energía alquímica con y sin este término. La energía en estos gráficos es relativa al *cluster* sustituido de menor energía. En el caso de la sustitución de un átomo, solo hay tres sitios no equivalentes. Del cuadro (a) se desprende que ambas transformaciones alquímicas, con y sin el término de polarización, son capaz de predecir la estructura de mínima energía. Sin embargo, si no se incluye el término de polarización (círculos rojos), las derivadas alquímicas no pueden predecir el orden correcto de las dos estructuras restantes. Cuando uno pasa a los casos de 2, 3 y 4 sustituciones parece que hay una falla completa de las derivadas alquímicas para ordenar las estructuras por sus energías. Esto es sorprendente ya que las derivadas alquímicas recuperan no menos del 99.8% del cambio absoluto de energía. Sin embargo, el restante 0.2% podría ser tan grande como tres veces la diferencia de energía máxima entre estructuras sustituidas. Sin embargo, un examen detallado de los términos de la ecuación (2.12) revela que la falta de correlación y la dispersión de la energía alquímica se explica por el cambio en la energía de repulsión nuclear. Cuando más de un átomo es sustituido, el cambio en la energía de repulsión nuclear tiene dos contribuciones, una lineal y una de segundo orden. Por ejemplo, cuando dos átomos son sustituidos, este cambio es

$$\Delta V_{NN} = \Delta Z_{\alpha} \sum_{\beta \neq \alpha} \frac{Z_{\beta}}{|\mathbf{R}_{\beta} - \mathbf{R}_{\alpha}|} + \frac{\Delta Z_{\alpha} \Delta Z_{\beta}}{|\mathbf{R}_{\beta} - \mathbf{R}_{\alpha}|} \quad (4.2)$$

El término lineal en la ecuación (4.2) podría ser tan grande como 100 a 300 veces el término no lineal para cuatro y dos sustituciones, respectivamente. Pero, aunque el término lineal hace la mayor contribución a ΔV_{NN} , los valores que toma para las sustituciones en diferentes sitios muestran una dispersión significativamente menor

que los valores del término de segundo orden. Especulamos que la única corrección formal a esto es incluir derivadas electrónicas de orden superior. Sin embargo, si se revisa la literatura donde se han usado series de perturbaciones similares, se encuentra que este término siempre ha sido despreciado sin más explicaciones [109–112]. A partir de ahora la contribución de segundo orden de la energía de repulsión nuclear no será considerado.

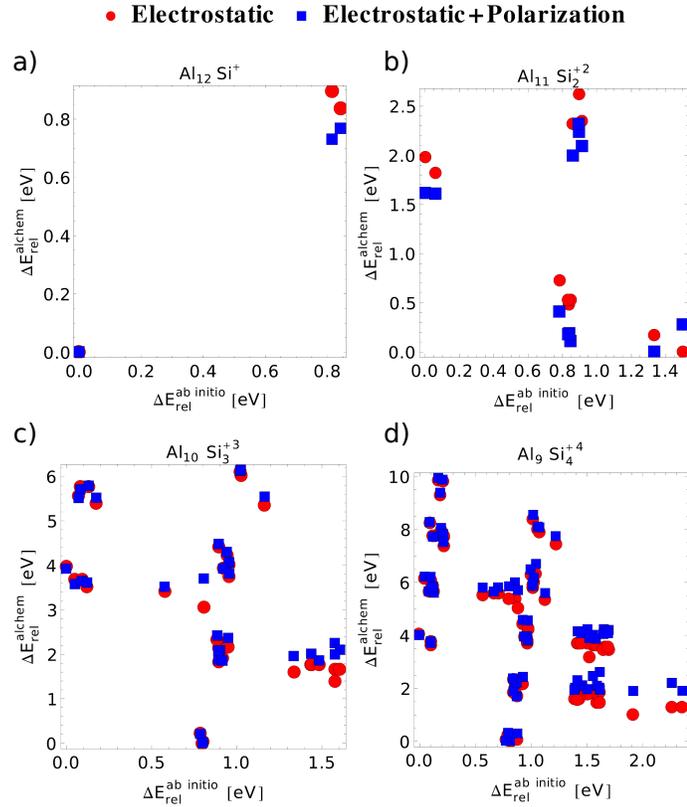


Figura 4.2: Energía alquímica vs. energía *ab initio* ($\Delta E_{\text{rel}}^{\text{alchem}}$ vs. $\Delta E_{\text{rel}}^{\text{ab initio}}$) de los *clusters* de aluminio dopados con átomos de silicio isoelectrónicamente $\text{Al}_{13-n}\text{Si}_n^{+n}$. Los puntos azules corresponden a derivadas alquímicas que incluyen hasta el término de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) $\text{Al}_{12}\text{Si}^+$, b) $\text{Al}_{11}\text{Si}_2^{+2}$, c) $\text{Al}_{10}\text{Si}_3^{+3}$ y d) $\text{Al}_9\text{Si}_4^{+4}$.

Las figuras 4.3 y 4.4 son completamente equivalentes a la figura 4.2 con la única diferencia que no se ha incluido el término nuclear de segundo orden. La figura 4.3 muestra como las derivadas alquímicas ordenan muy bien las estructuras de los *clusters* de $\text{Al}_{12}\text{Si}^+$ y $\text{Al}_{11}\text{Si}_2^{+2}$ según sus energías *ab initio*. De hecho, se encuentra una buena correlación lineal entre las energías alquímica y *ab initio*: $R^2 = 0.99$ y $R^2 = 0.94$ para las sustituciones simples y dobles, respectivamente. Nuevamente, se observa que la polarización es necesaria para ordenar correctamente las estructuras de mayor energía. La importancia de la polarización se vuelve más obvia cuando tres y cuatro átomos son mutados (figura 4.4 (a) y (b)). Se observa que la polarización es necesaria para que las derivadas alquímicas predigan el ordenamiento correcto de la mayoría de las estructuras de mayor energía y algunas de energías más bajas. La importancia de la polarización es doble. i) A medida que aumenta el número de átomos sustituidos, la perturbación se hace más grande y es necesario incluir al menos el término de segundo orden en la serie de perturbaciones. De hecho, si no se incluye la polarización, la energía relativa se subestima hasta en un 50%. Por otra parte, ii) el término electrostático es incapaz de distinguir las sustituciones múltiples de átomos equivalentes. Esto es fácil de entender. Supongamos que hay tres átomos no equivalentes con el mismo o casi el mismo potencial electrostático en el núcleo. Si se sustituyen dos de estos átomos, la contribución electrostática será la misma para los tres pares únicos de sustituciones. Los términos de polarización “diagonal” $\partial^2 E / \partial Z_\alpha^2$ también contribuirán en la misma cantidad. El único término que hace la diferencia es el término de polarización “no diagonal” $\partial^2 E / \partial Z_\alpha \partial Z_\beta$. La falta de este término es lo que hace que las energías alquímicas puramente electrostáticas en la figura 4.4 (círculos rojos) muestren un agrupamiento de estructuras de energía casi constante. Este efecto se vuelve más relevante en los *clusters* de alta simetría con muchos sitios

equivalentes.

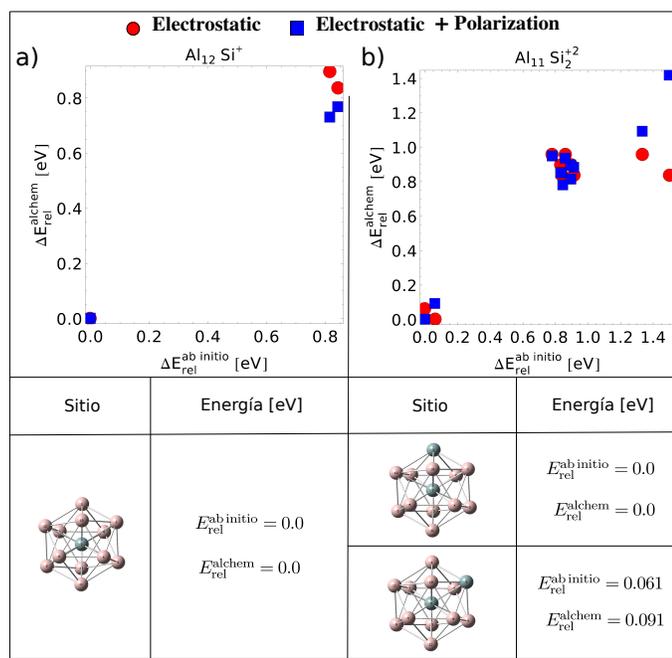


Figura 4.3: Energía alquímica vs. energía *ab initio* (ΔE_{rel}^{alchem} vs. $\Delta E_{rel}^{ab\ initio}$) de los *clusters* de aluminio dopados con átomos de silicio isoelectrónicamente $Al_{13-n}Si_n^{+n}$. Los puntos azules corresponden a derivadas alquímicas que incluyen hasta el término de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) $Al_{12}Si^+$ and b) $Al_{11}Si_2^{+2}$.

Vale la pena destacar que esperamos que las derivadas alquímicas sean una herramienta eficiente para filtrar las estructuras que no son buenas candidatas para los isómeros de baja energía. No esperamos que las derivadas alquímicas sean una alternativa a la evaluación real de la energía *ab initio*. Luego, para evaluar la eficiencia de las derivadas alquímicas, lo que importa es su capacidad de agrupar isómeros de acuerdo a su energía y no si existe una correlación estadísticamente significativa entre las energías alquímica y *ab initio*. En otras palabras, si se eligiera el 10% de los isómeros con menor energía alquímica para un refinamiento *ab initio* adicional,

a uno le gustaría maximizar la probabilidad de que el isómero de mínima energía real esté dentro de ese conjunto. En nuestro caso, esa probabilidad sería del 100 % para sustituciones de hasta cuatro átomos. Aun así, la correlación entre la energía alquímica y la energía *ab initio* para tres y cuatro sustituciones de átomos es sorprendentemente buena: $R^2 = 0.91$ y $R^2 = 0.89$ respectivamente.

Las figuras 4.3 y 4.4 también muestran las estructuras dopadas más estables y sus energías alquímicas y *ab initio* relativas. Podemos observar una característica atractiva, a medida que aumenta el orden de sustitución, las estructuras más estables son decoraciones de estructuras de baja energía de menor orden de sustitución. Por ejemplo, en la figura 4.3, las dos estructuras de energía más baja de $\text{Al}_{11}\text{Si}_2^{+2}$ corresponden a sustituciones del átomo central más uno externo; mientras que la sustitución favorable en $\text{Al}_{12}\text{Si}^+$ está en el átomo central. De manera similar, las seis estructuras de más baja energía de $\text{Al}_9\text{Si}_4^{+4}$ corresponden a una sustitución adicional en las seis estructuras de más baja energía de $\text{Al}_{10}\text{Si}_3^{+3}$. Esta característica es muy importante para aplicaciones más robustas de derivadas alquímicas dentro de, por ejemplo, algoritmos de búsqueda de estructuras estables de estequiometría diversa. Debido a que el poder de predicción de las derivadas alquímicas se degrada con el orden de sustitución (fuerza de la perturbación), sería más seguro filtrar estructuras al mutar un átomo a la vez con refinamientos *ab initio* intermedios.

4.3. Transformación en la cual el número de electrones cambia.

Ahora analizamos el caso en el que el número de electrones también cambia durante la sustitución de átomos. Es decir, mutaciones en las que el *cluster* Al_{13} cambia

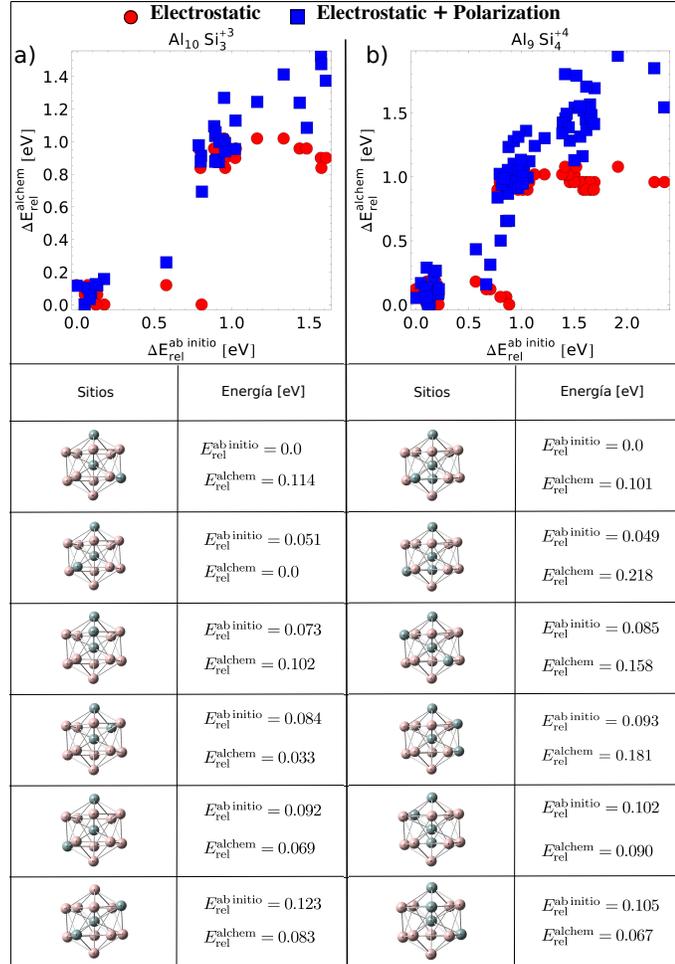


Figura 4.4: Energía alquímica vs. energía *ab initio* ($\Delta E_{\text{rel}}^{\text{alchem}}$ vs. $\Delta E_{\text{rel}}^{\text{ab initio}}$) de los *clusters* de aluminio dopados con átomos de silicio isoelectrónicamente $\text{Al}_{13-n}\text{Si}_n^{+n}$. Los puntos azules corresponden a derivadas alquímicas que incluyen hasta el término de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) $\text{Al}_{10}\text{Si}_3^{+3}$ and b) $\text{Al}_9\text{Si}_4^{+4}$.

a $\text{Al}_{13-n}\text{Si}_n$. En términos de derivadas alquímicas, la única diferencia con el caso isoelectrónico es que todos los términos que contienen ΔN en la ecuación (2.12) ahora están incluidos. En particular, el término que acopla cambios en el número atómico con transferencia de electrones es el asociado al potencial de Fukui (ecuación 2.22)

o la derivada mixta $\partial^2 E / \partial N \partial Z$. Hemos observado que con el cambio del número de electrones, la multiplicidad mínima de espín puede cambiar. Luego, en todos los cálculos *ab initio* hemos probado diferentes multiplicidades de espín, y se ha elegido el estado de mínima energía para cada estructura sustituida. La figura 4.5 compara el cambio alquímico relativo de la energía con el valor *ab initio* para la sustitución de 1, 2, 3 y 4 átomos. La energía alquímica sin incluir el término de transferencia de electrones (transformación isoelectrónica) también se muestra en la figura para evidenciar la relevancia de la misma. La importancia de incluir la transferencia de electrones es mixta. Para sustituciones de un átomo ($\Delta N = 1$) está claro que incluirla es una mejora sobre la transformación isoelectrónica. Sin embargo, para sustituciones de dos átomos ($\Delta N = 2$) a la vez, no es claro que el término de transferencia electrónica mejore el poder predictivo de las derivadas alquímicas. Aunque la transformación isoelectrónica rompe la tendencia de la transformación no isoelectrónica a agrupar incorrectamente estructuras de energía constante en el rango de 0.5 a 0.7 eV, falla para predecir la estructura de energía mínima. Cuando tres átomos son mutados ($\Delta N = 3$), al incluir los efectos de transferencia de electrones la energía alquímica empeora. Esto se puede ver en la figura 4.5(c) donde la correlación con la energía *ab initio* es mejor para la transformación isoelectrónica. En el caso en que cuatro átomos son sustituidos, figura 4.5(d), incluir la transferencia de electrones no representa ninguna mejora. Sin embargo, incluso para una perturbación tan grande ($\Delta N = 4$), las derivadas alquímicas son capaces de distinguir correctamente entre estructuras de alta y baja energía, que se manifiesta en los dos grupos de puntos en la figura 4.5(d). El grupo de puntos de menor energía corresponde a estructuras con un átomo de Si en el centro, mientras que el grupo de puntos de mayor energía corresponde a las estructuras que tienen todos los átomos de Si en el capa externa.

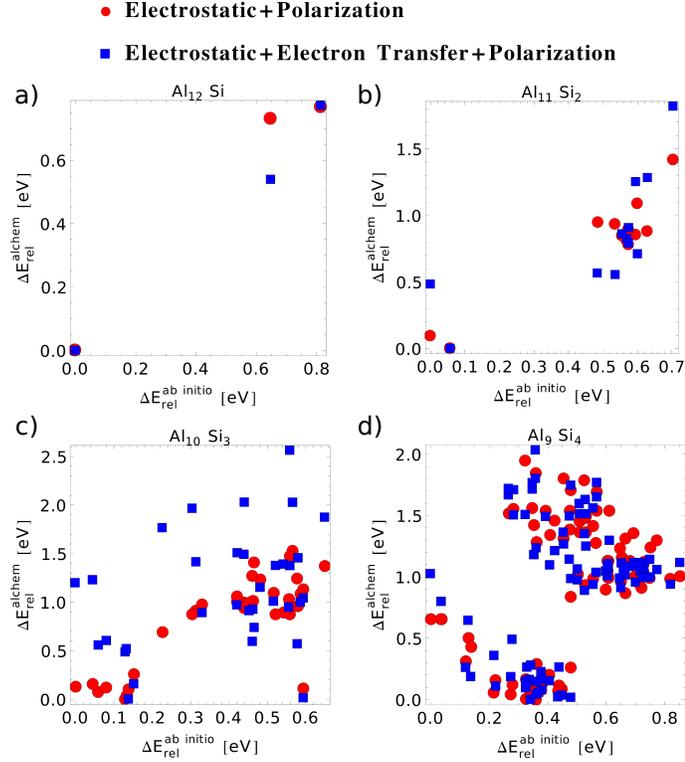


Figura 4.5: Energía alquímica vs. energía *ab initio* ($\Delta E_{\text{rel}}^{\text{alchem}}$ vs. $\Delta E_{\text{rel}}^{\text{ab initio}}$) de los *clusters* de aluminio dopados no isoelectrónicamente con átomos de silicio $\text{Al}_{13-n}\text{Si}_n$. Los puntos azules corresponden a derivadas alquímicas que incluyen los términos electrostáticos, de transferencia de electrones y de polarización, mientras que los puntos rojos corresponden a derivadas alquímicas que sólo incluye la contribución electrostática y de polarización. En todos los casos la energía es relativa al *cluster* dopado de mínima energía. a) Al_{12}Si , b) $\text{Al}_{11}\text{Si}_2$, c) $\text{Al}_{10}\text{Si}_3$ and d) Al_9Si_4 .

La conclusión que se puede extraer de la figura 4.5 es que para las transformaciones donde el número de electrones cambia, la serie de perturbaciones converge más lentamente con el número de electrones que con el potencial externo. Por lo tanto, parece razonable sugerir que cada vez que cambie el número de electrones, se deben evitar las transformaciones alquímicas que cambian más de un átomo a la vez, a menos que se tengan en cuenta los términos de orden superior en ΔN . El siguiente término en la serie de perturbaciones es la tercera derivada de orden mixto, $\frac{\partial^3 E}{\partial N \partial N \partial Z_\alpha}$,

la cual está relacionada con el potencial eléctrico del descriptor dual $f^{(2)}(\mathbf{r})$ [113,114]

$$\frac{\partial^3 E}{\partial N \partial N \partial Z_\alpha} = - \int d\mathbf{r} \frac{f^{(2)}(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} \quad (4.3)$$

La inclusión de este término en la serie de perturbaciones no aumenta drásticamente el costo computacional de las derivadas alquímicas debido a que la integral en la ecuación (4.3) también se puede expresar en términos de potenciales electrostáticos moleculares. Pero, el descriptor dual generalmente tiene una intrincada estructura nodal que podría acarrear una inestabilidad numérica en su evaluación. El cálculo de las funciones de respuesta de orden superior se vuelve más complicado y exigente computacionalmente. Sin embargo, una estrategia aproximada para calcular funciones de respuesta de alto orden en términos de órdenes de menor nivel podría ser útil. Una estrategia como esa se trata en el apéndice de la referencia [114]. Además, los modelos analíticos de la curva E vs. N podrían usarse para incluir términos de orden superior [78, 115]. Finalmente, si uno compara los cuadros en la figura 4.5, parece paradójico que las derivadas alquímicas son peores para la sustitución de tres átomos que cuatro. Sin embargo, hemos observado que las estructuras de los *clusters* Al_{12}Si , $\text{Al}_{11}\text{Si}_2$ y Al_9Si_4 prefieren estados de baja multiplicidad de spin; singlete, doblete y doblete respectivamente. Contrariamente, los *clusters* de $\text{Al}_{10}\text{Si}_3$ tienden a ser triplete en lugar de singlete. Las derivadas alquímicas que hemos usado aquí no toman en cuenta el hecho de que hay más de una multiplicidad de espín posible para un cambio dado en el número de electrones. Para incluir estos efectos, se deben usar las funciones de respuesta *spin resolved* [116–119].

4.4. Análisis de correlación de las derivadas alquímicas

Es importante enfatizar que lo que se desea en una herramienta para explorar espacios configuracionales y estequiométricos es que sea capaz de filtrar estructuras de alta energía. Por lo tanto, lo que es importante es la capacidad de las derivadas alquímicas para clasificar las estructuras de acuerdo con su energía y no cuán correlacionadas están las energías calculadas con química con el valor *ab initio* real. De las figuras 4.4 y 4.5 no es obvio juzgar qué tan buenas son las derivadas alquímicas para clasificar las estructuras por energía. Luego, hemos calculado el coeficiente de correlación Kendall τ [120–122], el cual mide la correspondencia entre dos sistemas de *rankings* diferentes. En nuestro caso, se desea medir la correspondencia entre las energías alquímica y *ab initio*. La definición para el coeficiente de Kendall τ que hemos usado es

$$\tau = \frac{P - Q}{\sqrt{(P + Q + T)(P + Q + U)}}, \quad (4.4)$$

donde P es el número de pares concordantes, Q es el número de pares discordantes, T es el número de vínculos solo en x , y U el número de vínculos solo en y , donde x e y son dos sistemas de *rankings* diferentes:

$$\begin{aligned} x &: x_1, x_2, x_3, \dots, x_n \\ y &: y_1, y_2, y_3, \dots, y_n. \end{aligned} \quad (4.5)$$

Un Kendall τ positivo (correlación positiva) indica que la correlación de ambas variables (energías alquímica y *ab initio*) es alto. Cuanto más cerca está el coeficiente de Kendall τ a 1, más correlacionados están los pares de elementos de ambas listas. Entonces, $\tau = 1$ significaría que las derivadas alquímicas clasifican la energía de las

estructuras exactamente igual a los cálculos *ab initio*.

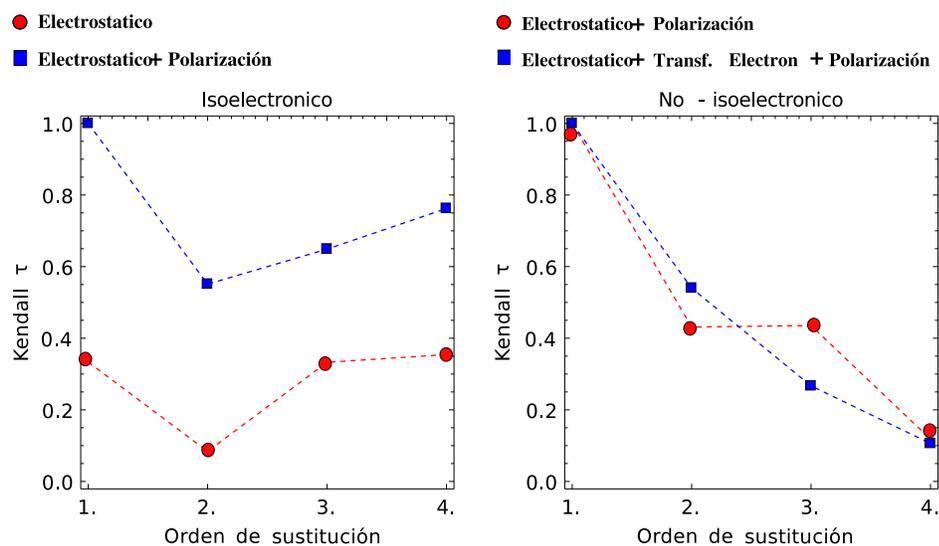


Figura 4.6: Coeficiente de correlación de Kendall τ para sustituciones desde uno a cuatro átomos de aluminio. En el cuadro a) se muestra el caso isoelectrónico donde los puntos rojos corresponde a mutaciones que solo incluye el término electrostático y los puntos azules corresponden a mutaciones que incluyen tanto los efectos electrostáticos como los de polarización. En el cuadro b) se muestra el caso no-isoelectrónico, donde los puntos rojos corresponden a mutaciones que incluyen las contribuciones electrostática y de polarización, y los puntos azules también incluyen los efectos de transferencia electrónica.

La figura 4.6 muestra el coeficiente de Kendall τ para ambas mutaciones: isoelectrónicas (cuadro a)) y no isoelectrónicas (cuadro b)) para sustituciones de uno a cuatro átomos de Al. Se puede apreciar que el comportamiento del coeficiente de Kendall τ está en acuerdo con la discusión de las dos secciones anteriores: i) al incluir la polarización en mutaciones isoelectrónicas tiene un mayor impacto en el poder de correspondencia de las derivadas alquímicas para todos los órdenes de sustitución; y ii) para sustituciones de más de dos átomos en mutaciones no isoelectrónicas, la

inclusión del término de transferencia de electrones no tiene un efecto claro en el poder de correspondencia de las derivadas alquímicas.

Este capítulo ha sido publicado en la referencia [123].

Capítulo 5

Prediciendo sitios de deprotonación en moléculas mono y polipróticas

La protonación y la deprotonación son quizás unos de los procesos más importantes en las reacciones químicas [124, 125]. La facilidad de deprotonación, por ejemplo, determina la acidez de las soluciones acuosas, que es de suma importancia en la síntesis química y el funcionamiento de la maquinaria molecular de sistemas biológicos [36, 126, 127]. En este capítulo, analizaremos la capacidad de las derivadas químicas para predecir la energía de deprotonación (acidez) de una serie de moléculas orgánicas. La deprotonación puede verse como una transformación química en la que el protón es eliminado mientras que el número de electrones permanece constante. En particular nos centraremos en predecir la energía de deprotonación de grupos de moléculas orgánicas monopróticas y también evaluaremos la capacidad de las derivadas químicas para predecir el sitio de más fácil deprotonación en moléculas polipróticas.

5.1. Transformación alquímica tipo II: Deprotonación

Hemos definido las transformaciones alquímicas tipo II (sección 2.2) como una transformación en la que se agrega o remueve un átomo o ion con número atómico Z_α en cierta posición \mathbf{R}_α . Si consideramos $\Delta N = 0$, $\Delta Z_\alpha = -1$ y \mathbf{R}_α como la posición del protón, entonces, la transformación dada por la ecuación (2.12) corresponde a la eliminación del protón del átomo de hidrógeno ácido y representa una aproximación a la energía de deprotonación. A una molécula que pierde espontáneamente un protón en solución acuosa, e incluso en fase gaseosa, se le llama un ácido de Brønsted-Lowry, mientras que a una que lo acepta se le llama una base [128, 129]. La acidez es quizás una de las propiedades más relevantes para la síntesis de moléculas en presencia de solventes acuosos. Otra manera de entender la deprotonación es como un proceso adiabático entre dos puntos del CS con diferentes estequiometrías. Por ejemplo la deprotonación del metanol puede ser vista como una transformación alquímica (adiabática) donde el estado inicial ($\lambda = 0$) es la especie protonada y el estado final ($\lambda = 1$) es el anión deprotonado (ver figura 5.1). En este caso el parámetro de transformación adiabática no es más que el número atómico del protón ácido.

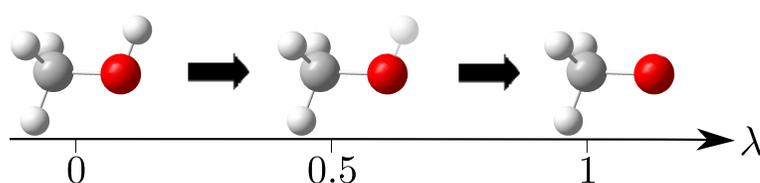


Figura 5.1: Representación de la transformación alquímica adiabática del metanol ($\lambda = 0$) al anión metóxido ($\lambda = 1$).

5.1.1. Cálculos de acidez en fase gaseosa y afinidad protónica

Aunque la cantidad física que nos interesa predecir es la energía necesaria para eliminar un protón de una molécula, conviene poner esto en el contexto del lenguaje químico y de cantidades susceptibles de determinar con experimentos. La afinidad protónica, PA, de una molécula, que puede ser neutra (B) o un anión (A^-), en fase gaseosa se define como el negativo del cambio de entalpía de la reacción entre la molécula y un protón:



Estas reacciones son siempre exotérmicas ($\Delta H < 0$) y por tanto la PA es positiva y también puede ser pensada como el cambio de entalpía de la reacción inversa de (5.1). Note que en la reacciones (5.1) las moléculas A^- y B aceptan un protón, por tanto son una base de Brønsted-Lowry. Pero si uno observa la reacción inversa de (5.1), las moléculas AH y BH^+ donan un protón, por tanto son ácidos de Brønsted-Lowry. Al par (AH, A^-) se le conoce como un par conjugado ácido/base, es decir, AH es el ácido conjugado de la base A^- y lo mismo para el par (BH^+ , B). Si la PA es grande significa que la base pierde fácilmente el protón y por tanto es una base fuerte, mientras que su ácido conjugado es un ácido débil. En resumen, si uno es ligero con el lenguaje químico, una molécula es un “buen” ácido si la PA de su base conjugada es pequeña, o si la entalpía de deprotonación es pequeña. Finalmente, la PA en fase gaseosa es importante porque proporciona información sobre la capacidad intrínseca de una molécula para aceptar/donar un protón sin considerar las complicaciones del solvente.

En resumen, la PA es la entalpía de disociación del protón desde el ácido conjugado:

$$\text{PA} = \Delta E + RT \quad (5.2)$$

Donde ΔE es el cambio de energía interna en la **deprotonación del ácido conjugado**. Este ΔE incluye las contribuciones translacional, rotacional, vibracional y electrónica de la energía

$$\Delta E = \Delta E_{\text{trans}} + \Delta E_{\text{rot}} + \Delta E_{\text{vib}} + \Delta E_{\text{elect}}. \quad (5.3)$$

Como la PA es medida en condiciones termodinámicas estándar, la temperatura ambiente es mucho mayor que la temperatura rotacional de las moléculas, por tanto, el cambio en la energía rotacional es despreciable ($\Delta E_{\text{rot}} \approx 0$). El cambio de energía por mol ($R = N_A k_B$) asociado a los grados de libertad traslacional extra del protón es $\Delta E_{\text{trans}} = \frac{3}{2}RT$. Luego, podemos escribir la PA como

$$\text{PA} = \Delta E_{\text{elect}} + \Delta E_{\text{vib}} + \frac{5}{2}RT. \quad (5.4)$$

Note que en la expresión usual para la PA los términos ΔE_{elect} y ΔE_{vib} van acompañados de un signo negativo en lugar de uno positivo. Debe entonces recordarse que en la ecuación (5.4) los cambios de energía son para la reacción inversa de (5.1). En términos de las transformaciones alquímicas es más sencillo calcular la energía para eliminar un protón desde el ácido conjugado que agregar un protón a la base dado que en el ácido se conoce la posición de dicho protón. Entonces, el cambio de energía electrónica para remover un proton de la posición \mathbf{R}_α , $\Delta E_{\text{elect}}[Z(\mathbf{R}_\alpha)]$, puede ser estimado con derivadas alquímicas:

$$\begin{aligned}\Delta E_{\text{elec}}[Z(\mathbf{R}_\alpha)] &= \left(\frac{\partial E}{\partial Z_\alpha}\right)_N \Delta Z_\alpha + \frac{1}{2} \left(\frac{\partial^2 E}{\partial Z_\alpha^2}\right)_N \Delta Z_\alpha^2 \\ &= \Phi_e(\mathbf{R}_\alpha) \Delta Z_\alpha + \frac{1}{2} \frac{\partial \Phi_e(\mathbf{R}_\alpha)}{\partial Z_\alpha} \Delta Z_\alpha^2.\end{aligned}\quad (5.5)$$

La derivada de $\Phi_e(\mathbf{R}_\alpha)$ con respecto a Z_α en la ecuación (5.5) es evaluada usando la aproximación de diferencias finitas hacia atrás con precisión de orden tres con un paso $\Delta Z_\alpha = 0.05$:

$$\frac{\partial \Phi_e(\mathbf{R}_\alpha)}{\partial Z_\alpha} \approx \frac{\frac{11}{6} \Phi_e(Z_\alpha) - 3 \Phi_e(Z_\alpha - \Delta Z_\alpha) + \frac{3}{2} \Phi_e(Z_\alpha - 2\Delta Z_\alpha) - \frac{1}{3} \Phi_e(Z_\alpha - 3\Delta Z_\alpha)}{\Delta Z_\alpha}\quad (5.6)$$

5.2. Moléculas monopróticas

Para evaluar la capacidad de las derivadas alquímicas para predecir la afinidad protónica, se calcularon las entalpías de deprotonación para un grupo de moléculas orgánicas monopróticas (ácidos y bases comunes). En este grupo encontramos principalmente tres tipos: aminas, ácidos carboxílicos y moléculas aromáticas. Las aminas se pueden considerar como compuestos nitrogenados derivados del amoníaco (NH_3) y resultan de la sustitución de uno o varios de los hidrógenos de la molécula de amoníaco por otros sustituyentes. Según la teoría ácido-base de Brønsted-Lowry estos compuestos son considerados como bases ya que son aceptores de protones, como por ejemplo la molécula N-metiletanamina (ver figura 5.2).

Los ácidos carboxílicos se caracterizan por poseer un grupo funcional llamado grupo carboxilo ($-\text{COOH}$). Estos compuestos son donantes de protones debido a que

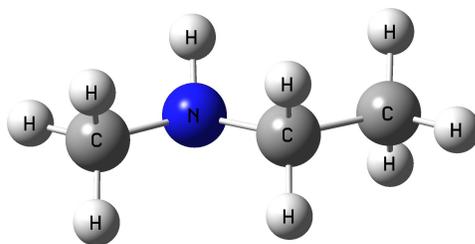


Figura 5.2: La N-metiletanamina es una amina secundaria donde el nitrógeno es capaz de aceptar un protón.

en el grupo carboxilo los dos átomos de oxígeno son electronegativos y tienden a atraer a los electrones del átomo de hidrógeno del grupo hidroxilo. Ello debilita el enlace O-H, provocando en ciertas condiciones una ruptura heterolítica y donando el correspondiente protón. La molécula deprotonada queda con carga negativa al alojar el electrón del átomo de hidrógeno. Un ejemplo de este tipo de compuestos es el ácido propanoico, mostrado en la figura 5.3

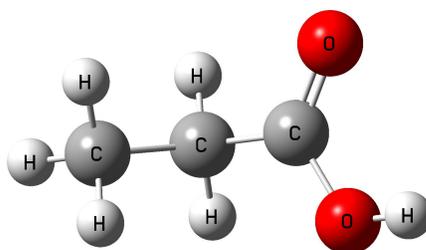


Figura 5.3: El ácido propanoico es un ácido carboxílico monoprótico capaz de donar el protón hidrógeno unido al oxígeno.

El último grupo de moléculas monopróticas que se ha considerado son ácidos y bases de Brønsted de compuestos aromáticos. Estas moléculas son compuestos orgánicos cíclicos que poseen gran estabilidad debido a la deslocalización electrónica en los enlaces π . Uno de los compuestos de este grupo que hemos estudiado es

el pirrol (ver figura 5.4). El pirrol es un compuesto aromático y heterocíclico de cinco miembros con un átomo de nitrógeno y cuatro átomos de carbono. Cuando el nitrógeno del pirrol es protonado, este pierde su estabilización aromática, por lo tanto, la protonación en el nitrógeno es desfavorable, haciendo del pirrol una base muy débil en comparación con las aminas convencionales y algunos otros compuestos aromáticos como la piridina.

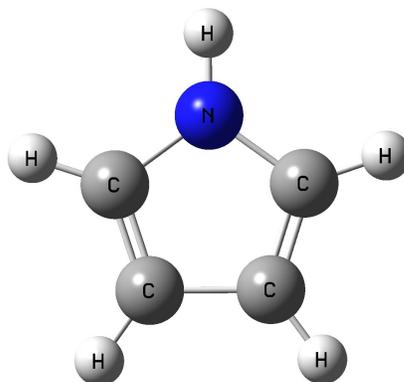


Figura 5.4: El pirrol es un base débil en la posición N-H.

Los cálculos de relajación de geometría y de frecuencias vibracionales de cada molécula se realizaron utilizando dos tipos de funcionales: el funcional híbrido B3LYP [73] y el funcional con correcciones de largo alcance ω B97XD [130]. Además se utilizaron los conjuntos de bases def2-TZVPPD [76] y cc-pVTZ [131]. Al comparar los cálculos de PA con valores experimentales [132–135], encontramos que con B3LYP/def2-TZVPPD se obtienen mejores resultados para los sistemas de interés. Todos los cálculos se realizaron con el software de química computacional Gaussian09 [77].

En la tabla 5.1 se muestran los valores de afinidad protónica calculados usando derivadas alquímicas (usando las ecuaciones (5.4), (5.5) y (5.6)), diferencias de entalpías *ab initio* (B3LYP/def2-TZVPPD) de las especies con y sin protón y valores experimentales para todos los miembros del conjunto de moléculas mencionadas. Se observa que existe una diferencia de energía de 70-80 (kcal/mol) entre los valores alquímicos y los *ab initio*. El origen de estas diferencias se debe en parte a que las transformaciones alquímicas no consideran la relajación estructural de la geometría inicial; pero tampoco se puede descartar que también contribuyan las derivadas de tercer orden y mayores no consideradas en la aproximación de la ecuación (5.5).

Tabla 5.1: Comparación entre las entalpías de disociación del protón, PA, (en kcal/mol, $1\text{eV}=23.06\text{ kcal/mol}$) obtenidas con derivadas alquímicas, B3LYP/def2-TZVPPD y valores experimentales para un conjunto de aminas, moléculas aromáticas y ácidos carboxílicos.

Base conjugada o ácido	Entalpía de disociación del protón (PA)		
	Alquímico	B3LYP/def2-TZVPPD	Experimental ^a
Aminas			
N,N-dimetilbutanamina	151.2	230.7	231.6
N-metiletanamina	145.6	224.9	225.2
N,N-dimetilpropanamina	150.6	230.0	230.1
N,N-dimetiletanamina	149.6	228.9	229.4
N,N-dietilpropanamina	156.0	235.4	233.9
N-etil-N-metilpropanamina	152.6	231.8	232.1
Aromáticas			
Anilino	138.5	213.0	210.9
Ácido Benzoico	135.6	206.5	196.2
Indol	275.9	351.3	351.8
Fenol	279.2	348.5	349.4
Piridinio	144.7	224.3	222.3
Pirrol	285.6	361.0	359.5
Ácidos carboxílicos			
Ácido butanoico	276.5	345.2	346.6
Ácido etanoico	278.1	347.0	348.4
Ácido propanoico	277.0	345.4	347.5
Ácido fluoracético	267.6	336.3	339.2
Ácido difluoracético	258.9	328.4	330.2
Ácido trifluoracético	254.0	322.6	322.6

^a Referencias [132–135]

Respecto a la precisión de las derivadas alquímicas, se ha evaluado la correlación entre los resultados alquímicos y valores experimentales (puntos púrpuras), además de compararlas también con los cálculos realizados con DFT(B3LYP) (puntos azules). Podemos ver de las figuras 5.5, 5.6 y 5.7 que esta correlación es sorprendentemente buena para los tres grupos de moléculas estudiados. Particularmente, para el grupo de las aminas (figura 5.5), la correlación entre las derivadas alquímicas y los

valores experimentales es menor que para los otros dos grupos, siendo el coeficiente de correlación $R^2 = 0.952$. Para el grupo de ácidos carboxílicos (figura 5.6) y moléculas orgánicas (figura 5.7), el coeficiente es $R^2 = 0.991$ y $R^2 = 0.995$ respectivamente. La correlación encontrada entre los valores de acidez calculados con B3LYP y con derivadas alquímicas resultó en una excelente concordancia para los tres grupos, con un coeficiente $R^2 = 0.999$.

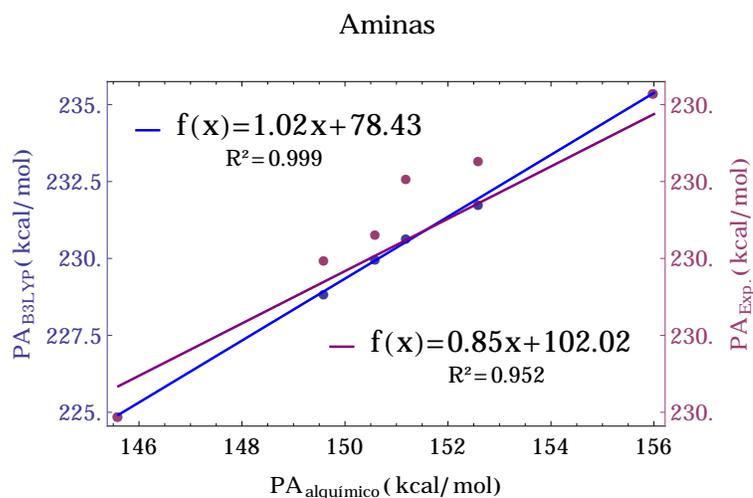


Figura 5.5: Comparación entre los valores de PA (en kcal/mol) *ab initio* versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para el grupo de las aminas.

Ahora bien, si evaluamos la correlación de todo el conjunto de moléculas monopróticas estudiadas (ver figura 5.8), obtenemos un coeficiente de $R^2 = 0.994$ respecto al caso experimental y $R^2 = 0.997$ respecto al caso *ab initio*, lo cual evidencia el poder predictivo de las derivadas alquímicas para este tipo de mutaciones.

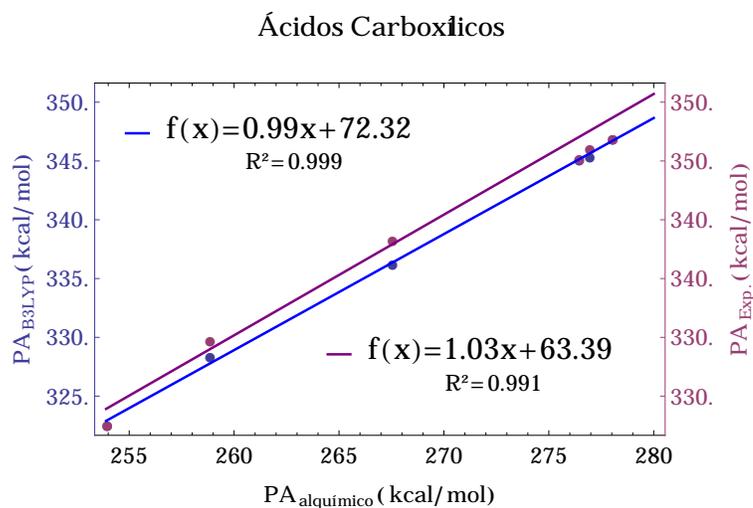


Figura 5.6: Comparación entre los valores de PA (en kcal/mol) *ab initio* versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para el grupo de los ácidos carboxílicos.

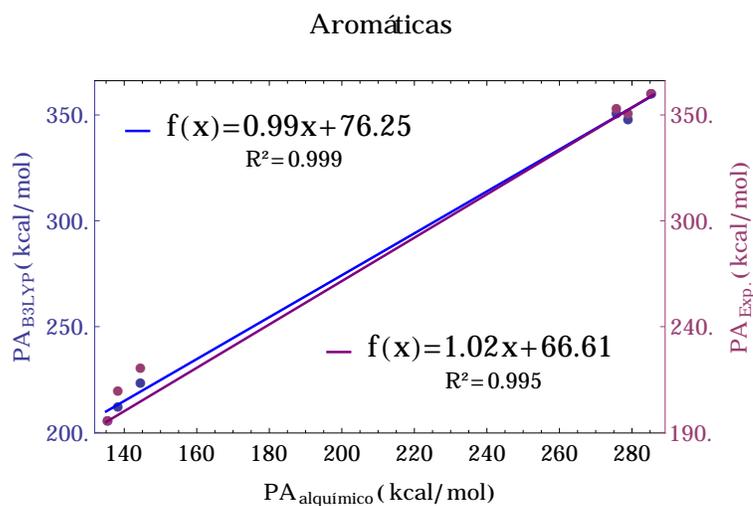


Figura 5.7: Comparación entre los valores de PA (en kcal/mol) *ab initio* versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para el grupo de las moléculas aromáticas.

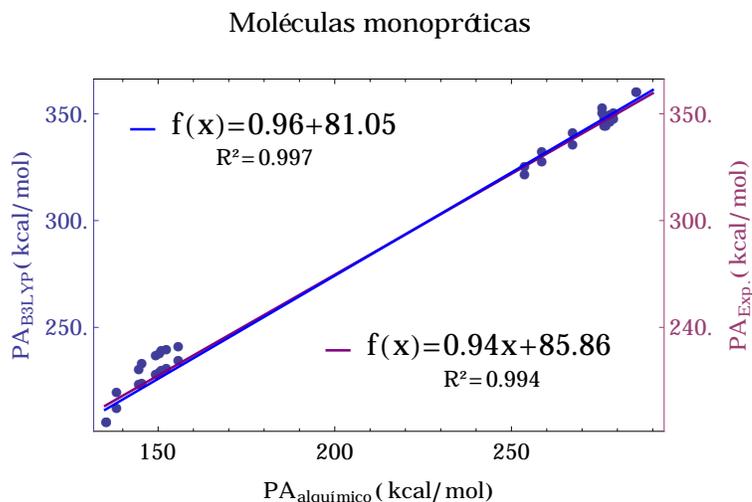


Figura 5.8: Comparación entre los valores de PA (en kcal/mol) *ab initio* versus alquímicos (puntos azules) y experimentales versus alquímicos (puntos púrpuras) para todo el conjunto de moléculas monopróticas: aminas, ácidos carboxílicos y aromáticas.

5.3. Moléculas polipróticas

Otro caso interesante es utilizar las derivadas alquímicas para predecir la acidez relativa en moléculas que son capaces de ceder más de un protón desde diferentes sitios. Un grupo de moléculas polipróticas con gran importancia biológica son las pirimidinas y purinas [136]. Las pirimidinas son compuestos heterocíclicos similares a la molécula de benceno pero con dos átomos de nitrógeno que sustituyen al átomo de carbono en las posiciones 1 y 3 (ver figura 5.9). Las purinas están compuestas por dos anillos, uno de seis átomos y otro de cinco, con cuatro átomos de nitrógeno en total, siendo tres de ellos de carácter básico. La importancia de este tipo de compuestos radica en que constituyen la mayor parte de las moléculas portadoras de la información genética en el ARN y el ADN [126, 137, 138]. Específicamente,

uno de los sistemas en que evaluaremos las derivadas alquímicas son una serie de uracilos (ver figura 5.9) los cuales tienen dos sitios de deprotonación: N_1 y N_3 , y en la posición cinco hay distintos sustituyentes ($X=CH_3, H, F, CHO$ y NO_2) en donde el valor de PA de cada sitio tendrá exclusiva dependencia del tipo de sustituyente dada la estructura electrónica de cada uno de ellos. A este grupo de uracilos se agregan tres moléculas orgánicas que contienen tres o más sitios de deprotonación: la adenina y guanina, que pertenecen al grupo de las purinas, y la citosina que pertenece al grupo de las pirimidinas (ver figura 5.10).

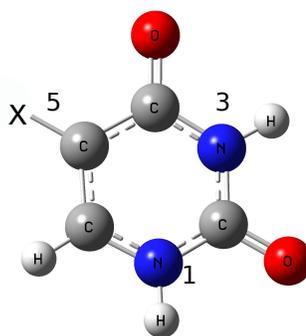


Figura 5.9: Uracilos con dos sitios de deprotonación: N_1 y N_3 . En la posición 5 los sustituyentes son $X=CH_3, H, F, CHO$ y NO_2

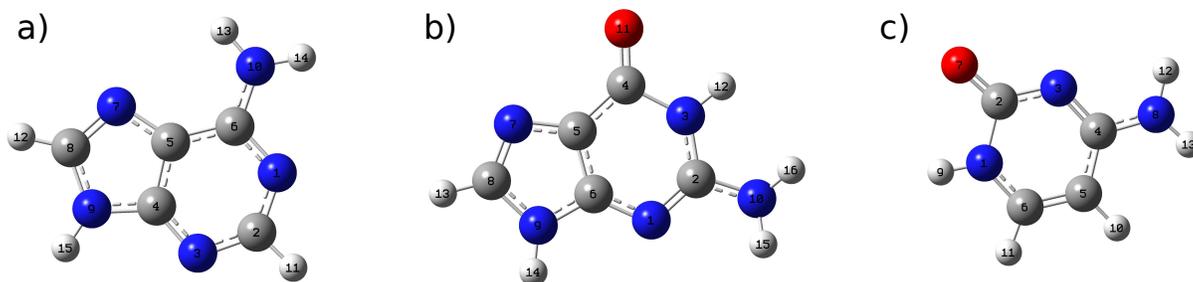


Figura 5.10: Moléculas polipróticas con tres o más sitios de deprotonación: a) adenina, b) guanina y c) citosina.

En la tabla 5.2 se muestran las energías de disociación del protón encontradas para el grupo de moléculas polipróticas mostradas en las figuras 5.9 y 5.10 mediante derivadas alquímicas, DFT(B3LYP) y valores experimentales. Se puede observar que al igual que para el caso de moléculas monopróticas, existe una diferencia de energía entre los valores alquímicos y *ab initio*, que para este caso es aproximadamente 70-76 (kcal/mol). Dentro del grupo de uracilos estudiados, no se encontraron valores experimentales para *trans*-5-formiluracilo, *cis*-5-formiluracilo, 5-nitrouracilo y para el sitio N₃ del 5-fluorouracilo.

Tabla 5.2: Comparación entre las energías de disociación del protón (en kcal/mol) obtenidas con derivadas alquímicas, B3LYP/def2-TZVPPD y experimentales para un conjunto de pirimidinas y purinas.

Molécula	Sitio	Energías de disociación del protón (PA)		
		Alquímico	B3LYP/def2-TZVPPD	Experimental ^a
Timina	N ₁ -H	263.1	336.2	333(2)
	N ₃ -H	274.0	347.5	347(4)
Uracilo	N ₁ -H	261.7	334.6	333(2)
	N ₃ -H	273.7	347.0	347(2)
5-fluorouracilo	N ₁ -H	256.8	329.5	329(3)
	N ₃ -H	267.1	340.4	-
<i>trans</i> -5-formiluracilo	N ₁ -H	247.3	320.9	-
	N ₃ -H	262.5	335.3	-
<i>cis</i> -5-formiluracilo	N ₁ -H	246.6	319.9	-
	N ₃ -H	264.1	336.8	-
5-nitrouracilo	N ₁ -H	240.2	314.2	-
	N ₃ -H	257.4	329.9	-
Adenina	N ₉ -H	264.5	339.7	333(1)
	NH-H	288.2	357.3	352(2)
	NH-H	287.9	357.6	352(2)
	C ₂ -H	329.1	402.3	-
	C ₈ -H	300.1	376.4	-
Guanina	N ₉ -H	264.7	338.5	331(4)
	NH-H	278.3	344.4	336(4)
	NH-H	272.2	339.7	337(4)
	N ₃ -H	267.4	340.5	334(4)
	C ₈ -H	304.8	380.2	-
Citosina	N ₁ -H	274.3	348.9	340(2)
	NH-H	286.4	355.6	346(4)
	NH-H	282.6	350.6	346(4)
	C ₅ -H	304.2	379.6	-
	C ₆ -H	295.7	372.1	-

^aReferencia [138]

Dada la buena correlación encontrada para el grupo de moléculas monopróticas entre los valores de PA alquímicos y los valores experimentales, y considerando que la correlación entre los los valores de PA alquímicos y los encontrados mediante cálculos *ab initio* es aún mejor, se decidió comparar sólo los valores de PA alquímicos versus con los encontrados con DFT(B3LYP), ya que además hay moléculas con un número de sitios sin los datos experimentales correspondientes. Se puede apreciar de la figura 5.11 la buena correlación lineal entre éstas dos cantidades ($R^2 = 0.994$), lo cual constata la capacidad de las derivadas alquímicas para predecir tendencias en los valores de PA en este conjunto de compuestos.

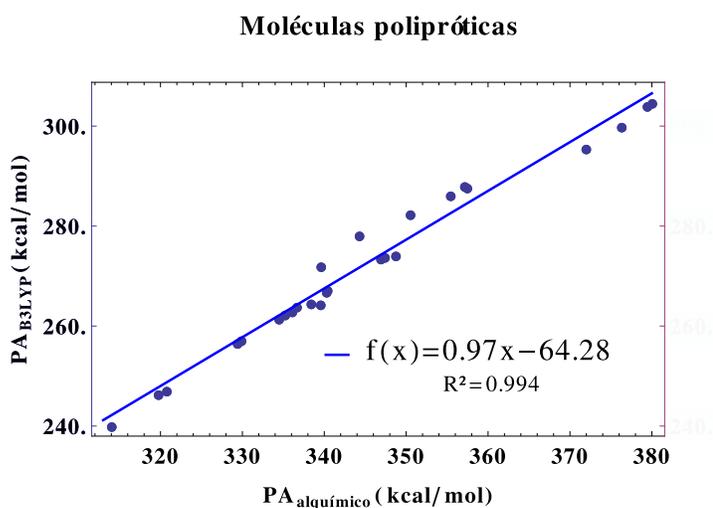


Figura 5.11: Comparación de los valores de afinidad protónica *ab initio* (PA_{B3LYP}) versus alquímicos ($PA_{alquímico}$) (en kcal/mol) para un conjunto de moléculas polipróticas: uracilos, adenina, guanina y citosina.

Un problema más complejo es predecir valores relativos de PA entre diferentes sitios de una molécula. En particular, en síntesis química es importante saber en que orden se deprotonan los diferentes sitios. La complejidad en predecir PA relativas

radica en que la diferencia entre dos sitios es solo una pequeña fracción de la PA de cada sitio. En la figura 5.12 se han graficado los valores relativos de PA calculados con alquimia versus los valores *ab initio* entre los sitios N_1 y N_3 para el caso de los uracilos, N_9 y NH_2 para la adenina, N_9 , N_3 y NH_2 para la guanina, y N_1 y NH_2 para la citosina. Al realizar una regresión lineal entre estas dos cantidades, vemos que la correlación es menor que la encontrada en el gráfico de la figura 5.11, sin dejar de ser estadísticamente importante, con un coeficiente $R^2 = 0.930$. Además, la diferencia persistente de entre 70 y 80 kcal/mol de la PA alquímica contra la *ab initio* desaparece cuando se comparan valores relativos (esto es evidente del pequeño intercepto de la recta en la figura 5.12). Esta cancelación de errores sugieren que la contribución de las derivadas alquímicas de órdenes superiores y el efecto de la relajación geométrica son términos casi contantes.

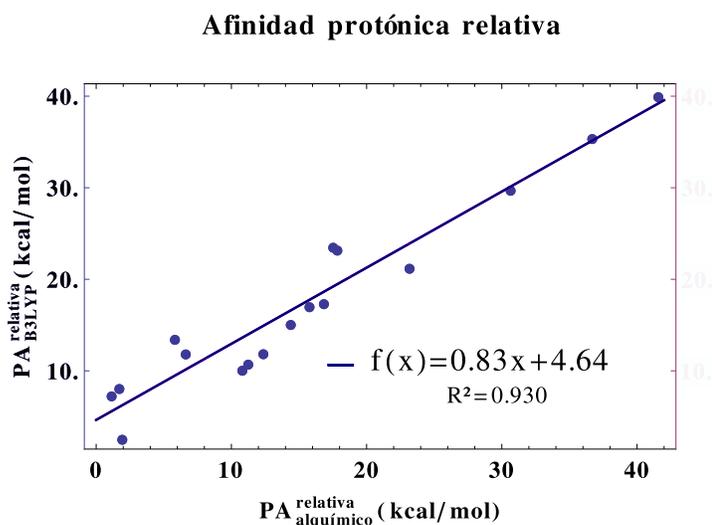


Figura 5.12: Comparación de los valores de afinidad protónica relativa *ab initio* ($PA_{B3LYP}^{relativa}$) versus alquímicos ($PA_{alquímico}^{relativa}$) (en kcal/mol) para un conjunto de moléculas polipróticas: uracilos, adenina, guanina y citosina.

Para el grupo de uracilos en donde se tienen los valores de PA experimentales, específicamente para la timina y el uracilo, vemos que las derivadas alquímicas son capaces de predecir cual es el sitio de más fácil deprotonación, que para ambas moléculas es el sitio N_1 . Para el resto de los uracilos, en que sólo podemos comparar con los valores obtenidos mediante cálculos DFT(B3LYP), también el sitio más probable de deprotonación es el N_1 . Para el resto de las moléculas polipróticas (adenina, guanina y citosina), vemos que las derivadas alquímicas igualmente pueden predecir el sitio más probable de deprotonación. Es importante notar que no esperamos que las derivadas alquímicas sean un método para obtener el valor “real” de la energía *ab initio*, sino que sean capaces de predecir y ordenar de manera correcta cuales son los sitios en donde es más probable que el protón ácido se disocie. Otro aspecto considerable a destacar en el cálculo de acidez en moléculas con más de un sitio de deprotonación, es que el uso de las derivadas alquímicas permite reducir el costo computacional significativamente, ya que para evaluar las derivadas alquímicas necesitamos solo un cálculo de optimización de geometría y frecuencia vibracional, independiente de cuántos sean los sitios ácidos que tenga la molécula. Comúnmente, si un compuesto tiene n sitios de deprotonación, se necesitan $2n + 2$ cálculos, en donde la mitad de ellos corresponde a optimizaciones de geometría y la otra mitad a cálculos de frecuencia. En este capítulo hemos visto la utilidad y eficiencia de las derivadas alquímicas para predecir los sitios de deprotonación en moléculas polipróticas, con una demanda computacional reducida.

En la figura 5.13 se muestra la transformación alquímica adiabática del metanol al anión metóxido. La energía de cada punto de la transformación es relativa a la energía del metanol, por lo que la diferencia de energía entre el punto con $Z_\alpha = 0$ y el punto con $Z_\alpha = 1$ resultará en la afinidad protónica del anión metóxido o acidez en fase gas del metanol. Al realizar el cálculo mediante DFT(B3LYP/def2-TZVPPD) se obtiene que la PA del anión metóxido es de 379.95 (kcal/mol). La transformación adiabática del metanol al anión metóxido se desarrolla variando la carga nuclear del átomo de hidrógeno unido al oxígeno, es decir, el protón va a ir “desapareciendo” a medida que avanza por este camino, tal como se representó en la figura 5.1 al comienzo de este capítulo. El cálculo de la energía al variar la carga nuclear Z_α se realiza con la geometría del metanol, lo que en principio excluye efectos de relajación estructural. Sin embargo, al calcular la energía del anión metóxido relajado, se obtiene una diferencia de 4 (kcal/mol) aproximadamente (ver línea roja punteada en figura 5.13). Para elucidar la diferencia que se obtiene entre los valores *ab initio* y los valores alquímicos de la energía se ha realizado el ajuste de los puntos graficados en la figura 5.13 con un polinomio de segundo orden (curva negra) y con un polinomio de tercer orden (curva verde). Ambos polinomios tienen prácticamente el mismo nivel de correlación (con un $R^2 \approx 1.0$). Sin embargo, al comparar los coeficientes lineales y cuadráticos de ambos ajustes con los valores correspondientes a las derivadas alquímicas evaluadas en $Z = 1$ (curva azul), se ve que es más plausible que el camino adiabático tenga al menos términos cúbicos no despreciables. De hecho, el término cúbico explica 55 (kcal/mol) de la diferencia de 61.1 (kcal/mol) entre el valor *ab initio* y el alquímico con derivadas de segundo orden (ecuación 5.5). Otra fuente que podría explicar las 6 (kcal/mol) restantes es el hecho que el cálculo de cada punto de la transformación se realiza utilizando la misma base que la del metanol,

es decir, no incluye el cambio en el conjunto base debido a la variación de la carga nuclear (error de Pulay [108, 139]). Para poder calcular la energía de los sistemas, la función de onda se expande en una base finita gaussiana que depende de la carga nuclear Z en forma exponencial. Por lo que al variar la carga nuclear en el cálculo, manteniendo la base fija, se tiende a mantener a los electrones alrededor del mismo potencial del estado inicial ($Z = 1$), haciendo que la energía sea mayor a la esperada. Por lo tanto, el camino adiabático de la figura 5.13 es una cota superior al camino exacto. El error de Pulay suele ser menor que 10 (kcal/mol) para energías relativas, lo cual es del orden de las 6 (kcal/mol) que no explica el término cúbico.

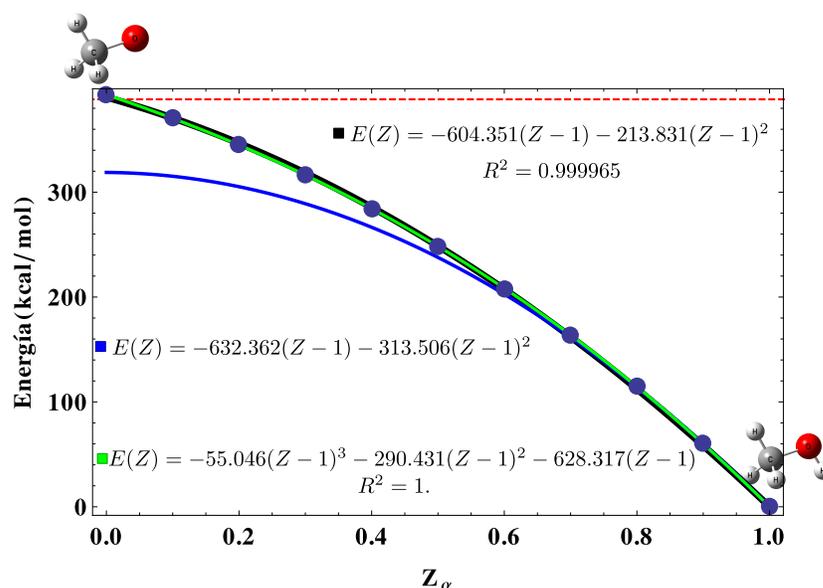


Figura 5.13: Cambio en la energía de la transformación adiabática del metanol al anión metóxido. La curva negra (verde) corresponde al ajuste cuadrático (cúbico) de los puntos de la transformación alquímica adiabática. La curva azul corresponde a los valores de las derivadas alquímicas evaluadas en $Z = 1$. La línea roja punteada corresponde a la energía del anión metóxido relajado.

El artículo de este capítulo “*Predicting deprotonation sites in polyprotic molecules with computational alchemy.*” está en proceso de redacción.

Capítulo 6

Conclusiones

El espacio químico contiene todos los compuestos posibles que se pueden imaginar. Su tamaño colosal descarta la posibilidad de calcular las propiedades de sus elementos mediante cálculos *ab initio*. Sin embargo, solo una pequeña fracción de estos compuestos tiene un interés práctico. El diseño inverso intenta responder a la pregunta de cómo optimizar un compuesto que se ajusta a un conjunto de propiedades deseadas. Esto se puede lograr navegando los “camino” del espacio químico que albergan compuestos óptimos. En este trabajo hemos estudiado cómo las derivadas químicas *ab initio* pueden ser una herramienta eficiente para navegar el espacio químico, siendo capaces de predecir, por ejemplo, los cambios energéticos asociados a los cambios en la composición de cada compuesto.

En el capítulo 3, hemos mostrado cómo las derivadas químicas dependen del modelo de la densidad electrónica y de la aproximación de la función de respuesta lineal utilizada en su evaluación. Dada la complejidad en la evaluación de la función de respuesta lineal exacta, existen distintos tipos de aproximaciones para su cálculo, en donde cada una de ellas incluyen diferentes efectos del cambio en el potencial externo. En este trabajo hemos estudiado distintos tipos de funcionales (BLYP, B3LYP, CAM-B3LYP), primero en un conjunto de moléculas diatómicas en las aproximacio-

nes IPA y RPA para la función de respuesta lineal, y hemos comparado estos valores con los obtenidos mediante la aproximación de diferencias finitas utilizando uno de los métodos más precisos que incluyen la correlación electrónica, *Coupled Cluster*. Para todos los casos estudiados, en este grupo de moléculas diatómicas, la aproximación RPA siempre brinda una mejor descripción de la función de respuesta lineal, independiente del funcional utilizado en su evaluación. No existe una diferencia significativa en los resultados al utilizar un funcional tipo GGA (BLYP) o uno híbrido (B3LYP), los cuales incluyen una parte de intercambio exacto. Por lo tanto, es importante tener en cuenta que para obtener una buena descripción de la función de respuesta lineal, se necesita al menos incluir el efecto de la interacción Coulombiana en el par partícula-hueco. En este capítulo también se ha estudiado las derivadas de segundo orden respecto a la carga nuclear del átomo central para un conjunto de moléculas tetraédricas, con todas las aproximaciones descritas para $\chi(\mathbf{r}, \mathbf{r}')$ en esta tesis. Los resultados obtenidos mediante diferencias finitas tanto para HF(CFD), BLYP(CFD) y B3LYP(CFD) se han validado mediante los obtenidos con los métodos analíticos HF(CPHF), BLYP(CPKS) y B3LYP(CPKS), respectivamente. El siguiente paso fue comparar tanto las aproximaciones IPA y RPA con un método que incluyera de manera confiable la correlación entre los electrones, y al igual que para el caso de las moléculas diatómicas, el método comparativo *gold standard* elegido para este análisis fue CCSD. Como se evidenció en el caso de moléculas diatómicas, la aproximación de partícula independiente siempre subestima el valor de las segundas derivadas respecto a la carga nuclear, independiente del método utilizado para su evaluación. Es importante mencionar, que la subestimación es mucho menor con HF, dada la correcta descripción del intercambio en la región de largo alcance. Para el caso de BLYP y B3LYP, el funcional híbrido es capaz de disminuir esa subestimación, pero

sin dejar de ser suficientemente grande. Por lo tanto podemos decir que para todas las moléculas estudiadas en este trabajo, la aproximación IPA no es la adecuada para describir correctamente las derivadas de segundo orden respecto a la carga nuclear. En el siguiente nivel de aproximación de la función de respuesta lineal estudiado, RPA, pudimos apreciar que siempre existe una sobrestimación de los valores obtenidos para $\chi_{\alpha\alpha}$. Sin embargo, esta sobreestimación es siempre mucho menor, que la subestimación de IPA en los tres métodos estudiados (HF, BLYP y B3LYP). No existe una diferencia sustancial de los valores obtenidos para $\chi_{\alpha\alpha}$ entre los funcionales BLYP y B3LYP (alrededor del 3%). Por lo tanto, para dar una buena descripción de las derivadas de segundo orden respecto a la carga nuclear, es necesario utilizar al menos la aproximación RPA, utilizando un funcional tipo GGA, como BLYP, sin ser necesario utilizar otros funcionales más sofisticados.

En el capítulo 4, vimos como las derivadas alquímicas *ab initio* son una herramienta prometedora para navegar rigurosamente el espacio químico. Aquí llevamos las derivadas alquímicas al límite y evaluamos qué tan confiables son en la predicción del aspecto energético de las transmutaciones composicionales de un clúster semi-metálico. El *cluster* elegido, Al_{13} , y sus mutaciones desafían las derivadas alquímicas porque la fuerza relativa de la perturbación asociada a la mutación es grande, y porque la estructura electrónica de este sistema requiere que se incluyan efectos adicionales a los electrostáticos. Estudiamos dos tipos de transformaciones alquímicas, transformaciones isoelectrónicas en las que los átomos de Al se reemplazan por átomos de Si sin cambiar el número de electrones, y las transformaciones no isoelectrónicas en las que la cantidad de electrones también cambia para mantener el *cluster* neutro. En el caso de las transformaciones no isoelectrónicas, se observa que los efectos de polarización deben incluirse para recuperar la energía relativa correcta

entre los isómeros. De hecho, si se incluye la polarización, se logra un buen *ranking* de energía y se observa una buena correlación con las energías *ab initio* para las sustituciones de un máximo de dos átomos. Cambiar cuatro o más átomos conduce a una falta de convergencia de la expansión de segundo orden y las derivadas químicas pierden su valor predictivo. Se observa una característica atractiva, que es que a medida que aumenta el orden de sustitución, las estructuras más estables son decoraciones de estructuras de baja energía de menor orden de sustitución. Esta característica es fundamental para aplicaciones robustas de derivadas químicas dentro de algoritmos de búsqueda de estructuras estables de estequiometría diversa. Sugerimos que, siempre que sea posible, no se transforme más de un átomo a la vez. Cuando la cantidad de electrones también cambia, es necesario incluir derivadas de transferencia de electrones. Sin embargo, se observa que la serie de perturbaciones converge más lentamente con el número de electrones que con el potencial externo. Incluso agregar dos electrones al sistema debe considerarse como una gran perturbación. En esta situación, se vuelve más importante evitar las transformaciones químicas que cambian más de un átomo a la vez, a menos que se tengan en cuenta los términos de orden superior en ΔN . Sin embargo, es nuestra opinión que la inclusión de estas respuestas de orden superior hace que el cálculo sea más complicado y que es preferible mantener la perturbación pequeña. Es cierto que en aplicaciones para química orgánica $|\Delta N|$ rara vez es mayor que uno. Aunque se espera que los cambios en el número de electrones sean difíciles de capturar con un enfoque perturbativo, es deseable que el poder predictivo de las derivadas químicas para transmutaciones no isoelectrónicas sea igual al de los casos isoelectrónicos.

Finalmente, todavía queda por entender por qué la inclusión de la repulsión nuclear-nuclear de segundo orden “fuera de la diagonal” tiene una gran influencia

en los resultados. Creemos que los términos electrónicos de orden superior se encargarían de este comportamiento inesperado. Aquí adoptamos el enfoque pragmático de dejarlos fuera. Esto se ha hecho consistentemente en la literatura, pero nunca se ha discutido antes.

En el capítulo 5 analizamos la capacidad de las derivadas alquímicas, en una primera instancia, para predecir la energía de enlace del protón en un grupo de moléculas orgánicas monopróticas. Se calculó la afinidad protónica de este grupo de moléculas mediante los métodos alquímicos descritos en el capítulo 2, mediante cálculos DFT y se compararon los valores de las afinidades protónicas con los datos experimentales encontrados en la literatura. Para los tres grupos de moléculas estudiadas, se observó una excelente correlación entre la predicción alquímica y el valor encontrado con DFT de las afinidades protónicas. También la correlación encontrada entre los valores alquímicos y experimentales es sorprendentemente buena. Todo esto se evidenció mediante las regresiones lineales que se realizaron en donde se obtuvieron coeficientes de correlación cercanos a uno. Por lo tanto, dentro del conjunto de moléculas monopróticas estudiadas, la capacidad predictiva de las derivadas alquímicas en el cálculo de las afinidades protónicas es lo suficientemente buena como para pensar que podemos usar este tipo de herramientas para un conjunto mayor de ácidos y bases.

Como un segundo paso, se estudió la capacidad predictiva de las derivadas alquímicas en el proceso de deprotonación de un grupo de moléculas orgánicas polipróticas, específicamente un subconjunto de pirimidinas y purinas. El interés en poner a prueba las derivadas alquímicas en estos sistemas radicó principalmente en dos aspectos, uno de ellos es que tienen una gran importancia biológica ya que estos sistemas constituyen la mayor parte de las moléculas portadoras de la información

genética en el ADN y ARN. Por otro lado, son moléculas en donde existe más de un protón ácido. Como vimos, el uso de las derivadas alquímicas permite reducir el costo computacional significativamente, ya que para evaluar las derivadas alquímicas necesitamos solo un cálculo de optimización de geometría y frecuencia vibracional, independiente de cuántos sean los sitios ácidos que tenga la molécula. Efectivamente, para el grupo de pirimidinas y purinas estudiado, las derivadas alquímicas son capaces de predecir el sitio más probable en donde el protón ácido se disocie, aunque la correlación entre los valores *ab initio* y alquímicos encontrado fue menor en este caso. Como hemos mencionado anteriormente, no esperamos que las derivadas alquímicas sean un método para obtener el valor real de la energía *ab initio*, sino que sean capaces de predecir y ordenar de manera correcta cuales son los sitios en donde es más fácil la deprotonación.

Este trabajo contribuye fundamentalmente en el desarrollo de una metodología basada en un aproximación perturbativa para explorar el espacio químico mediante el uso de derivadas alquímicas, las cuales no solo incluyen transformaciones del potencial externo, sino que también transformaciones en el número de electrones. Este desarrollo nos permite obtener cierto entendimiento sobre la estructura y propiedades de *clusters* y moléculas. Por ejemplo, comprender el balance entre los efectos electrostáticos, de transferencia de carga y de polarización en la estabilidad estructural de *clusters* a través de dopaje, como también nos permite entender mecanismos que favorecen la deprotonación en moléculas orgánicas. Creemos que esta tesis abre las puertas a aplicaciones fascinantes de diseño racional usando métodos enteramente *ab initio*.

Referencias

- [1] E. O. Pyzer-Knapp, C. Suh, R. Gómez-Bombarelli y Aguilera-Iparrag, *Ann. Rev. Mater. Res.* **45**, 195 (2015).
- [2] R. Ramakrishnan, P. O. Dral, M. Rupp y O. A. Von Lilienfeld, *Scientific data* **1**, 140022 (2014).
- [3] J. L. Reymond, *Acc. Chem. Res.* **48**, 722 (2015).
- [4] T. F. Cerqueira, R. Sarmiento-Pérez, M. Amsler, F. Nogueira, S. Botti y M. A. Marques, *J. Chem. Theory Comput.* **11**, 3955 (2015).
- [5] J. R. Morris, D. M. Deaven y K. M. Ho, *Phys. Rev. B* **53**, R1740 (1996).
- [6] A. Franceschetti y A. Zunger, *Nature* **402**, 60 (1999).
- [7] O. A. von Lilienfeld, R. D. Lins y U. Rothlisberger, *Phys. Rev. Lett.* **95**, 153002 (2005).
- [8] O. A. von Lilienfeld y M. E. Tuckerman, *J. Chem. Phys.* **125**, 154104 (2006).
- [9] O. A. von Lilienfeld y M. E. Tuckerman, *J. Chem. Theory Comput.* **3**, 1083 (2007).
- [10] G. H. D. Sheppard y O. A. von Lilienfeld, *J. Chem. Phys.* **133**, 084104 (2010).

- [11] D. J. Wales y H. A. Scheraga, *Science* **285**, 1368 (1999).
- [12] M. Wang, X. Hu, D. Beratan y W. Yang, *J. Am. Chem. Soc.* **128**, 3228 (2006).
- [13] D. Balamurugan, W. Yang y D. N. Beratan, *J. Chem. Phys.* **129**, 174105 (2008).
- [14] S. Keinan, M. J. Therien, D. N. Beratan y W. Yang, *J. Phys. Chem. A* **112**, 12203 (2008).
- [15] F. De Vleeschouwer, W. Yang, D. N. Beratan, P. Geerlings y F. De Prof, *Phys. Chem. Chem. Phys.* **14**, 16002 (2012).
- [16] P. Kirkpatrick y C. Ellis, *Nature* **432**, 823 (2004).
- [17] C. M. Dobson, *Nature* **432**, 824 (2004).
- [18] O. A. von Lilienfeld, *J. Chem. Phys.* **131**, 164102 (2009).
- [19] M. To Baben, J. O. Achenbach y O. A. von Lilienfeld, *J. Chem. Phys.* **144**, 104103 (2016).
- [20] R. S. Bohacek, C. McMartin y W. C. Guida, *Med. Res. Rev.* **16**, 3 (1996).
- [21] T. Allison, O. Coskuner y C. Gonzalez, *Metallic Systems: A Quantum Chemist's Perspective* (CRC Press, 2011).
- [22] S. Goedecker, W. Hellmann y T. Lenosky, *Phys. Rev. Lett.* **95**, 055501 (2005).
- [23] C. Cárdenas, W. Tiznado, P. Ayers y P. Fuentealba, *J. Phys. Chem. A* pág. 3801 (2011).

- [24] R. Balawender, M. A. Welearegay, M. Lesiuk, F. De Proft y P. Geerlings, *J. Chem. Theory Comput.* **9**, 5327 (2013).
- [25] O. A. von Lilienfeld, *Int. J. Quantum Chem.* **113**, 1676 (2013).
- [26] K. S. Chang, S. Fias, R. Ramakrishnan y O. A. von Lilienfeld, *J. Chem. Phys.* **144**, 174110 (2016).
- [27] A. Solovyeva y O. A. von Lilienfeld, *Phys. Chem. Chem. Phys.* **18**, 31078 (2016).
- [28] P. W. Ayers y R. G. Parr, *J. Am. Chem. Soc.* **123**, 2007 (2001).
- [29] J. S. M. Anderson, J. Melin y P. W. Ayers, *J. Chem. Theory Comput.* **3**, 358 (2007).
- [30] J. S. M. Anderson, J. Melin y P. W. Ayers, *J. Chem. Theory Comput.* **3**, 375 (2007).
- [31] C. W. J. Hafner y G. Ceder, *MRS Bull.* **31**, 659 (2006).
- [32] C. Lipinski y A. Hopkins, *Nature* **432**, 855 (2004).
- [33] J. C. Crocker, *Nature* **451**, 528 (2008).
- [34] C. C. Berry y A. S. G. Curtis, *J. Phys. D Appl. Phys.* **36**, R198 (2003).
- [35] J. K. Vasir, M. K. Reddy y V. D. Labhasetwar, *Curr. Nanosci.* **1**, 47 (2005).
- [36] M. A. Kurinovich y J. K. Lee, *Journal of the American Society for Mass Spectrometry* **13**, 985 (2002).
- [37] D. Li y H. Ai, *Journal of Physical Chemistry B* **113** (2009).

- [38] C. R. Moylan y J. I. Brauman, *Annual Review of Physical Chemistry* **34** (1983).
- [39] R. Parr y W. Yang, *Density-Functional Theory of Atoms and Molecules*, International Series of Monographs on Chemistry (Oxford University Press, USA, 1989).
- [40] P. Hohenberg y W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- [41] A. J. Coleman, *Review of Modern Physics* **35** (1963).
- [42] L. H. Thomas, *Mathematical Proceedings of the Cambridge Philosophical Society* **23**, 542548 (1927).
- [43] E. Fermi, *Rend. Accad. Lincei* **6**, 602 (1927).
- [44] W. Kohn y L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [45] C. Dykstra, G. Frenking, K. Kim y G. Scuseria, *Theory and Applications of Computational Chemistry: The First Forty Years* (Elsevier Science, 2011).
- [46] P. W. Ayers, J. S. M. Anderson y L. J. Bartolotti, *International Journal of Quantum Chemistry* **101**, 520 (2005).
- [47] R. A. Donnelly y R. G. Parr, *The Journal of Chemical Physics* **69**, 4431 (1978).
- [48] R. G. Parr y W. Yang, *Journal of the American Chemical Society* **106**, 4049 (1984).
- [49] R. G. Parr y R. G. Pearson, *Journal of the American Chemical Society* **105**, 7512 (1983).
- [50] D. Sheppard, G. Henkelman y O. A. von Lilienfeld, *The Journal of Chemical Physics* **133**, 084104 (2010).

- [51] V. Tevekeliyska, Y. Dong, M. Springborg y V. G. Grigoryan, *The European Physical Journal D* **43**, 19 (2007).
- [52] R. Balawender, M. A. Welearegay, M. Lesiuk, F. De Proft y P. Geerlings, *Journal of Chemical Theory and Computation* **9**, 5327 (2013).
- [53] R. G. Parr, R. A. Donnelly, M. Levy y W. E. Palke, *The Journal of Chemical Physics* **68**, 3801 (1978).
- [54] R. P. Feynman, *Phys. Rev.* **56**, 340 (1939).
- [55] C. Cárdenas, *Chem. Phys. Lett* **513**, 127 (2011).
- [56] P. W. Ayers y M. Levy, *Theoretical Chemistry Accounts* **103**, 353 (2000).
- [57] M. Lesiuk, R. Balawender y J. Zachara, *J. Chem. Phys.* **136**, 034104 (2012).
- [58] D. Peng y W. Yang, *J. Chem. Phys.* (2013).
- [59] M. Berkowitz y R. G. Parr, *The Journal of Chemical Physics* **88** (1988).
- [60] A. L. Fetter y J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Courier Corporation, 2003).
- [61] W. Yang, A. J. Cohen, F. De Proft y P. Geerlings, *The Journal of Chemical Physics* **136** (2012).
- [62] P. W. Ayers, *Faraday Discuss.* **135**, 161 (2007).
- [63] C. Jamorski, M. E. Casida y D. R. Salahub, *The Journal of Chemical Physics* **104** (1996).

- [64] A. Szabo y N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Books on Chemistry Series (Dover Publications, 1996).
- [65] S. Obara y A. Saika, *The Journal of Chemical Physics* **84** (1986).
- [66] S. F. Boys, *Proceedings Mathematical Physical & Engineering Sciences* **200** (1950).
- [67] A. Savin, F. Colonna y M. Allavena, *The Journal of Chemical Physics* **115**, 6827 (2001).
- [68] C. Møller y M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- [69] J. Čížek, *The Journal of Chemical Physics* **45** (1966).
- [70] H. G. Kümmel, *International Journal of Modern Physics B* **17** (2003).
- [71] A. D. Becke, *Physical Review A* **38** (1988).
- [72] C. Lee, W. Yang y R. G. Parr, *Physical Review B* **37** (1988).
- [73] A. D. Becke, *The Journal of Chemical Physics* **98** (1993).
- [74] K. Raghavachari, *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **103** (2000).
- [75] T. Yanai, D. P. Tew y N. C. Handy, *Chemical Physics Letters* **393** (2004).
- [76] D. Rappoport y F. Furche, *The Journal of Chemical Physics* **133**, 134105 (2010).

- [77] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, 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, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, . Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski y D. J. Fox, Gaussian 09 Revision A.1. Gaussian Inc. Wallingford CT 2009.
- [78] F. Heidar-Zadeh, M. Richer, S. Fias, R. A. Miranda-Quintana, M. Chan, M. Franco-Pérez, C. E. González-Espinoza, T. D. Kim, C. Lanssens, A. H. Patel, X. D. Yang, E. Vöhringer-Martinez, C. Cárdenas, T. Verstraelen y P. W. Ayers, *Chem. Phys. Lett.* **660**, 307 (2016).
- [79] T. Verstraelen, P. Tecmer, F. Heidar-Zadeh, K. Boguslawski, M. Chan, Y. Zhao, T. D. Kim, S. Vandenbrande, D. Yang, C. E. González-Espinoza, S. Fias, P. A. Limacher, D. Berrocal, A. Malek y P. W. Ayers, Horton 2.1.0b1, hORTON 2.1.0b1, <http://theochem.github.com/horton/>, 2015.
- [80] E. Prodan y W. Kohn, *Proceedings of the National Academy of Sciences* **102** (2005).

- [81] R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960).
- [82] R. McWeeny, *Phys. Rev.* **126**, 1028 (1962).
- [83] R. McWeeny, *Methods of molecular quantum mechanics*, Theoretical chemistry (Academic Press, 1992), segunda edición.
- [84] J. Hafner, C. Wolverton y G. Ceder, *MRS Bulletin* **31**, 659 (2006).
- [85] C. Lipinski y A. Hopkins, *Nature* **432**, 855 (2004).
- [86] J. C. Crocker, *Nature* **451**, 528 (2008).
- [87] C. C. Berry y A. S. G. Curtis, *Journal of Physics D: Applied Physics* **36**, R198 (2003).
- [88] J. K. Vasir, M. K. Reddy y V. D. Labhasetwar, *Current Nanoscience* **1**, 47 (2005).
- [89] S. Goedecker, W. Hellmann y T. Lenosky, *Phys. Rev. Lett.* **95**, 055501 (2005).
- [90] T. Allison, O. Coskuner y C. Gonzalez, *Metallic Systems: A Quantum Chemist's Perspective* (CRC Press, 2011).
- [91] A. Wranceschetti y A. Zunger, *Nature* **402**, 60 (1999).
- [92] O. A. von Lilienfeld, R. Lins y U. Rothlisberger, *Phys. Rev. Lett.* **95**, 153002 (2005).
- [93] D. Sheppard, G. Henkelman y O. A. von Lilienfeld, *The Journal of Chemical Physics* **133**, (2010).
- [94] D. J. Wales y H. A. Scheraga, *Science* **285**, 1368 (1999).

- [95] A. Laio y M. Parrinello, Proceedings of the National Academy of Sciences **99** (2002).
- [96] N. Metropolis y S. Ulam, Journal of the American Statistical Association **44**, 335 (1949). PMID: 18139350.
- [97] D. G. Truhlar, Computer Physics Communications **74** (1993).
- [98] V. Bach y L. Site, *Many-electron Approaches in Physics, Chemistry and Mathematics: A Multidisciplinary View*, Mathematical Physics Studies (Springer, 2014).
- [99] P. Kirkpatrick y C. Ellis, Nature **432**, 823 (2004).
- [100] C. M. Dobson, Nature **432**, 824 (2004).
- [101] R. S. Bohacek, C. McMartin y W. C. Guida, Medicinal Research Reviews **16**, 3 (1996).
- [102] V. Kumar, S. Bhattacharjee y Y. Kawazoe, Phys. Rev. B **61**, 8541 (2000).
- [103] V. Kumar y V. Sundararajan, Phys. Rev. B **57**, 4939 (1998).
- [104] L. D. Lloyd y R. L. Johnston, Chem. Phys. **236**, 107 (1998).
- [105] J. P. Perdew, K. Burke y M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
- [106] A. D. McLean y G. S. Chandler, J. Chem. Phys. **72**, 5639 (1980).
- [107] R. Krishnan, J. S. Binkley, R. Seeger y J. A. Pople, J. Chem. Phys. **72**, 650 (1980).
- [108] P. Pulay, Molecular Physics **17** (1969).

- [109] F. Weigend, C. Schrodtr y R. Ahlrichs, *J. Chem. Phys.* **121**, 10380 (2004).
- [110] F. Weigend y C. Schrodtr, *Chem-Eur. J.* **11**, 3559 (2005).
- [111] F. Weigend, *J. Chem. Phys.* **141**, 134103 (2014).
- [112] C. Seifried, L. Longo, P. Pollak y F. Weigend, *J. Chem. Phys.* **146**, 034304 (2017).
- [113] C. Morell, A. Grand y A. Toro-Labbe, *Chem. Phys. Lett.* **425**, 342 (2006).
- [114] C. Cárdenas, N. Rabi, P. W. Ayers, C. Morell, P. Jaramillo y P. Fuentealba, *J. Phys. Chem. A* **113**, 8660 (2009).
- [115] P. Fuentealba y C. Cárdenas, *J. Mol. Model.* **19**, 2849 (2013).
- [116] J. Melin, F. Aparicio, M. Galvan, P. Fuentealba y R. Contreras, *J. Phys. Chem. A* **107**, 3831 (2003).
- [117] C. Cárdenas, A. M. Lamsabhi y P. Fuentealba, *Chem. Phys.* **322**, 303 (2006).
- [118] P. Perez, E. Chamorro y P. W. Ayers, *J. Chem. Phys.* **128**, 204108 (2008).
- [119] R. A. Miranda-Quintana y P. W. Ayers, *Theor. Chem. Acc.* **135**, 239 (2016).
- [120] W. R. Knight, *J. Am. Stat. Assoc.* **61**, 436 (1966).
- [121] M. G. Kendall, *Rank correlation methods* (Griffin, London, 1970), cuarta edición.
- [122] M. G. Kendall, *Biometrika* **33** (1945).
- [123] M. Muñoz y C. Cárdenas, *Phys. Chem. Chem. Phys.* **19**, 16003 (2017).

- [124] K. W. Kramarz y J. R. Norton, *Slow Proton-Transfer Reactions in Organometallic and Bioinorganic Chemistry* (John Wiley & Sons, Inc., 2007), págs. 1–65.
- [125] R. Álvarez-Diduk, M. T. Ramírez-Silva, A. Galano y A. Merkoçi, *Journal of Physical Chemistry B* **117**, 12347 (2013).
- [126] Y. H. Jang, L. C. Sowers, T. Çağın y W. A. Goddard, *Journal of Physical Chemistry A* págs. 274–280 (2001).
- [127] A. K. Chandra, T. Uchimaru y T. Zeegers-Huyskens, *Journal of Molecular Structure* **605**, 213 (2002).
- [128] J. N. Brønsted, *Recueil des Travaux Chimiques des Pays-Bas* **42**, 718 (1923).
- [129] T. M. Lowry, *Journal of the Society of Chemical Industry* **42**, 43 (1923).
- [130] J. D. Chai y M. Head-Gordon, *Physical Chemistry Chemical Physics* **10** (2008).
- [131] T. H. Dunning, *The Journal of Chemical Physics* **90**, 1007 (1989).
- [132] E. P. L. Hunter y S. G. Lias, *Journal of Physical and Chemical Reference Data* **27** (1998).
- [133] S. T. Graul, M. E. Schnute y R. R. Squires, *International Journal of Mass Spectrometry and Ion Processes* **96** (1990).
- [134] M. Diaz-Tinoco, J. Romero, J. Ortiz, A. Reyes y R. Flores-Moreno, *J. Chem. Phys.* **138**, 194108 (2013).
- [135] Y. Valadbeigi, *Int. J. Mass Spectrom.* **395**, 49 (2016).

- [136] T. L. Gilchrist, *Heterocyclic chemistry* (Longman Scientific & Technical; Wiley, 1992).
- [137] M. A. Kurinovich y J. K. Lee, *Journal of the American Chemical Society* **122** (2000).
- [138] E. Chen, C. Herder y E. S. Chen, *J. Mol. Struct.* **798**, 126 (2006).
- [139] H. B. Schlegel, *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **103** (2000).