



HAL
open science

Propuesta de un Método Acelerado utilizando Descriptores de Reactividad para Explorar Superficies de Energía Potencial en Clúster Atómicos

Diego Inostroza

► **To cite this version:**

Diego Inostroza. Propuesta de un Método Acelerado utilizando Descriptores de Reactividad para Explorar Superficies de Energía Potencial en Clúster Atómicos. Coordination chemistry. Sorbonne Université; Universidad Andrés Bello (Santiago), 2023. Español. NNT : 2023SORUS514 . tel-04457279

HAL Id: tel-04457279

<https://theses.hal.science/tel-04457279>

Submitted on 14 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Propuesta de un Método Acelerado utilizando Descriptores de Reactividad para Explorar Superficies de Energía Potencial en Clúster Atómicos.

TESIS EN FISICOQUÍMICA MOLECULAR

Diego Estevan Inostroza Mendoza

Directores de tesis:

Dr. William Tiznado.

Dra. Julia Contreras.

Comisión:

Dr. Carlos Cárdenas. (Presidente del Jurado)

Dr. Eduardo Chamorro

Dr. Alexis Markovits

Dra. Tatiana Gómez

AGRADECIMIENTOS

Quiero agradecer a la Universidad Andrés Bello y a Sorbonne Université por brindarme las herramientas necesarias para el desarrollo del presente trabajo doctoral. A las fuentes de financiamiento, la Agencia Nacional de Investigación y Desarrollo (ANID) por su beca Doctorado Nacional 2019-21190427, al proyecto Fondecyt Regular N° 1211128 y al Laboratorio Nacional de Computación de Alto Rendimiento (NLHPC).

A mis tutores William Tiznado y Julia Contreras por guiarme en el desarrollo de mi tesis durante tanto tiempo, sus enseñanzas académicas, por impulsarme a ser más de lo que soy, por sus historias y vivencias en la investigación y por sus consejos de vida. Gracias por la paciencia y confianza depositadas en mi persona.

A los amigos del grupo CTCG, los doctores Osvaldo, Ricardo, Alejandro, Rodrigo, Víctor y Juan que me impulsaron a tomar el doctorado; y los futuros doctores Willy, Dumer, Andrea y Viviana que han hecho este último año uno muy divertido. A los amigos *Laboratoire de Chimie Théorique*, a Trinidad, Andrea, Hugo y Hugo por su apoyo durante mi estancia en París.

Al doctor Luis, que me ayudo a comprender la química cuando estaba perdido, porque siempre estaba disponible para ayudar debido a que nunca aprendió a dormir 8 horas diarias y por brindarme su amistad.

Al doctor José, que también me guió en los conocimientos físicos y químicos que no tenía, por ayudarme a desconcentrarme conversando de fútbol, series y películas, y por brindarme su amistad.

A mis compañeros de generación 2019, la mejor generación. A Jerónimo, Camilo, Leslie, Liz, Franco y Brandon con quienes compartí amablemente las primeras etapas del doctorado.

A los amigos de la vida, los de *Ohana Significa Familia* sin los cuales esta tesis hubiese finalizado mucho antes.

Agradezco a mi familia por su apoyo constante en este proceso, los anteriores y los futuros. A mis padres con sus buenos consejos su preocupación y su paciencia; a Gabriel, mi hermano, que me ayuda a sentirme como niño y Rocío, mi hermana, que tan chiquita me hace reír de sobremanera. A mis abuelos, tíos y primos de mi numerosa familia que mediante historias y vivencias me han mantenido motivado a continuar. Finalmente, a los empachosos, el Marley con sus ladridos, la Cosha con sus maullidos de escape y al Rengarcito que decidió partir antes a dormir en el sol.

INDICE

RESUMEN.....	6
RÉSUMÉ.....	7
SUMMARY.....	8
CAPÍTULO I - INTRODUCCIÓN.....	9
CAPÍTULO II - MARCO TEÓRICO.....	13
2.1 SUPERFICIE DE ENERGÍA POTENCIAL.....	13
2.2 TEORÍA DEL FUNCIONAL DE LA DENSIDAD.....	15
2.2.1 <i>Descriptores de Reactividad</i>	18
2.3 INTERACCIONES NO COVALENTES.....	22
2.3.1 NCI PLOT.....	24
2.4 ALGORITMO DE OPTIMIZACIÓN Y EL PROBLEMA DEL MÍNIMO GLOBAL.....	27
2.4.1 <i>Particle Swarm Optimization (PSO)</i>	30
2.4.2 <i>Ant Colony Optimization</i>	31
2.4.3 <i>Algoritmo Genético (AG)</i>	31
2.4.3.1 Autómata Celular.....	33
2.4.3.2 AUTOMATON.....	34
2.4.4 <i>Algoritmo KICK</i>	36
CAPÍTULO III - HIPÓTESIS Y OBJETIVOS.....	38
3.1 HIPÓTESIS.....	38
3.2 OBJETIVO GENERAL.....	38
3.2.1 <i>Objetivos específicos</i>	38
CAPÍTULO IV - ESTRATEGIAS DE EXPLORACIÓN DE LA SEP.....	39
4.1 GENERALIZACIÓN.....	39
4.2 DETALLES COMPUTACIONALES.....	40
4.3 RESULTADOS.....	41
4.3.1 <i>Mejoras al proceso AUTOMATON</i>	41
4.3.2 <i>Usos MatrixAutomaton -Carbonos Hipercoordinados Planos</i>	43
4.3.3 <i>Usos MatrixAutomaton - NaBH₃</i>	44
4.3.4 <i>Usos MatrixAutomaton - Sistemas Si-Li</i>	44

4.3.5	<i>Usos MatrixAutomaton - Sistemas B</i>	45
4.3.6	<i>Uso MatrixAutomaton y TemplatePermuta - Sistemas Si-C</i>	46
4.4	PERSPECTIVAS	51
CAPÍTULO V - KICK-FUKUI: UN MÉTODO DE PREDICCIÓN DE ESTRUCTURAS MOLECULARES GUIADO POR FUNCIONES DE FUKUI		53
5.1	KICK-FUKUI	53
5.2	SISTEMAS EVALUADOS	56
5.3	DETALLES COMPUTACIONALES	56
5.4	RESULTADOS	57
5.5	PERSPECTIVAS	63
5.5.1	<i>Diseño - Autómata Celular y Función de Fukui</i>	65
CAPÍTULO VI - PREDICCIÓN DE ESTRUCTURAS MOLECULARES MEDIANTE LAS INTERACCIONES NO COVALENTES		70
6.1	GENERALIZACIÓN	70
6.2	SISTEMAS EVALUADOS	71
6.3	DETALLES COMPUTACIONALES	71
6.4	RESULTADOS	72
6.5	PERSPECTIVAS	77
CAPÍTULO VII COMPARACIÓN DE MÉTODOS DESARROLLADOS		82
7.1	RESULTADOS - PTC C-BE-H	82
7.2	PERSPECTIVAS	88
CAPÍTULO VIII DISEÑO DE HERRAMIENTAS		90
8.1	GENERALIZACIÓN	90
8.2	DISEÑO HERRAMIENTAS - MATRIXAUTOMATON	91
8.2.1	<i>MatrixAutomaton</i>	92
8.3	DISEÑO HERRAMIENTAS - TEMPLATEPERMUTA	98
8.4	DISEÑO HERRAMIENTAS - FUNCIÓN DE FUKUI	100
8.5	DISEÑO HERRAMIENTAS - INTEGRAL NCI	104
8.6	DISEÑO HERRAMIENTAS - COMPILATORIO LOOKER	109
CAPÍTULO IX - CONCLUSIONES GENERALES		112
PUBLICACIONES		114
COLABORACIONES		115

CÓDIGOS	115
REFERENCIAS	116
ANEXOS	123
ANEXO 1	123
ANEXO 2	¡ERROR! MARCADOR NO DEFINIDO.
Artículo 1	¡Error! Marcador no definido.
Artículo 2	¡Error! Marcador no definido.
Artículo 3	¡Error! Marcador no definido.
Artículo 4	¡Error! Marcador no definido.
Artículo 5	¡Error! Marcador no definido.
Artículo 6	¡Error! Marcador no definido.
Artículo 7	¡Error! Marcador no definido.
Artículo 8	¡Error! Marcador no definido.
Artículo 9	¡Error! Marcador no definido.
Artículo 10	¡Error! Marcador no definido.
Artículo 11	¡Error! Marcador no definido.
Artículo 12	¡Error! Marcador no definido.
Artículo 13	¡Error! Marcador no definido.
Artículo 14	¡Error! Marcador no definido.
Artículo 15	¡Error! Marcador no definido.

RESUMEN

La creciente complejidad isomórfica de los clústeres atómicos al aumentar su tamaño ha impulsado a la comunidad científica a crear una variedad de técnicas y estrategias innovadoras para analizar eficientemente la superficie de energía potencial (SEP). En este contexto, el presente estudio introduce un conjunto de herramientas computacionales altamente optimizadas, diseñadas para agilizar la identificación de mínimos energéticos en la SEP de sistemas moleculares y clústeres atómicos. Estas herramientas se destacan por su eficiencia computacional, lo que las convierte en una opción atractiva para investigaciones futuras.

Se realizaron evaluaciones exhaustivas de técnicas estocásticas y dirigidas previamente descritas en la literatura científica. El algoritmo AUTOMATON, en particular, fue actualizado con implementaciones significativas que posibilitaron la identificación de sistemas atómicos novedosos, como el primer carbono hexacoordinado en una configuración planar.¹ Las técnicas estocásticas se dirigieron mediante la incorporación de descriptores de reactividad, como la función de Fukui (i), que permitió guiar con precisión la búsqueda de mínimos energéticos y estructuras estables. Además, a través de la Integral del NCI (ii), pudimos demostrar una fuerte correlación entre los valores de integración y la energía del sistema de interés. Este hallazgo establece un punto de partida prometedor para futuras investigaciones en este campo.

La combinación estratégica de técnicas de enjambre de partículas con permutaciones atómicas ha mejorado notablemente nuestra capacidad para explorar la SEP, facilitando la identificación de una amplia gama de soluciones. Estas herramientas han ido más allá de la simple replicación de resultados anteriores, descubriendo nuevos mínimos globales y estructuras estables en el proceso. Al fusionar conceptos químicos fundamentales con métodos de optimización avanzada y algoritmos estocásticos, este trabajo de tesis ha establecido un enfoque altamente predictivo para la exploración en química y el diseño de nuevos materiales.

RÉSUMÉ

La complexité isomorphique croissante des amas atomiques avec l'augmentation de leur taille a incité la communauté scientifique à créer une variété de techniques et de stratégies innovantes pour analyser efficacement la surface d'énergie potentielle (SEP). Dans ce contexte, la présente étude introduit un ensemble d'outils de calcul hautement optimisés conçus pour accélérer l'identification des minima énergétiques dans la SEP des systèmes moléculaires et des amas atomiques. Ces outils se distinguent par leur efficacité de calcul, ce qui en fait une option attrayante pour la recherche future.

Des évaluations complètes des techniques stochastiques et dirigées précédemment décrites dans la littérature scientifique ont été réalisées. L'algorithme AUTOMATON, en particulier, a été mis à jour avec des implémentations significatives qui ont permis l'identification de nouveaux systèmes atomiques, tels que le premier carbone hexacoordonné dans une configuration planaire.¹ Les techniques stochastiques ont été abordées par l'incorporation de descripteurs de réactivité, tels que la fonction Fukui(i), qui ont permis de guider avec précision la recherche de minima énergétiques et de structures stables. En outre, grâce à l'intégrale NCI (ii), nous avons pu démontrer une forte corrélation entre les valeurs d'intégration et l'énergie du système en question. Cette découverte constitue un point de départ prometteur pour de futures recherches dans ce domaine.

La combinaison stratégique des techniques d'essaimage de particules et des permutations atomiques a considérablement amélioré notre capacité à explorer la SEP, facilitant l'identification d'un large éventail de solutions. Ces outils sont allés au-delà de la simple reproduction de résultats antérieurs, découvrant au passage de nouveaux minima globaux et des structures stables. En fusionnant des concepts chimiques fondamentaux avec des méthodes d'optimisation avancées et des algorithmes stochastiques, ce travail de thèse a établi une approche hautement prédictive de l'exploration chimique et de la conception de nouveaux matériaux.

SUMMARY

The increasing isomorphous complexity of atomic clusters with increasing size has prompted the scientific community to create a variety of innovative techniques and strategies to efficiently analyze the potential energy surface (PES). In this context, the present study introduces a set of highly optimized computational tools designed to streamline the identification of energetic minima in the PES of molecular systems and atomic clusters. These tools are noted for their computational efficiency, making them an attractive option for future research.

Comprehensive evaluations of stochastic and directed techniques previously described in the scientific literature were performed. The AUTOMATON algorithm, in particular, was updated with significant implementations that enabled the identification of novel atomic systems, such as the first hexacoordinate carbon in a planar configuration.¹ Stochastic techniques were addressed through the incorporation of reactivity descriptors, such as the Fukui function (i), which allowed to accurately guide the search for energetic minima and stable structures. Furthermore, through the NCI Integral (ii), we were able to demonstrate a strong correlation between the integration values and the energy of the system of interest. This finding establishes a promising starting point for future research in this field.

The strategic combination of particle swarm techniques with atomic permutations has greatly enhanced our ability to explore PES, facilitating the identification of a wide range of solutions. These tools have gone beyond simple replication of previous results, uncovering new global minima and stable structures in the process. By merging fundamental chemical concepts with advanced optimization methods and stochastic algorithms, this thesis work has established a highly predictive approach to chemistry exploration and new materials design.

CAPÍTULO I – INTRODUCCIÓN

El progreso humano a lo largo de los siglos ha estado estrechamente ligado al avance científico, social y cultural de las civilizaciones. Los científicos han enfrentado numerosos desafíos, desde el desarrollo de la agricultura y la ganadería, hasta la comprensión de los fenómenos meteorológicos, el estudio del movimiento de las estrellas, el descubrimiento de las leyes fundamentales que rigen las fuerzas del universo y la investigación de los elementos que constituyen nuestro mundo. En la actualidad, el uso de herramientas computacionales ha revolucionado la investigación científica, destacando en particular los avances en el campo de la Nanotecnología². Esta disciplina se define como el diseño, caracterización y producción de sistemas, materiales y dispositivos a nanoescala, los cuales han permitido avances significativos en diversas áreas, como la medicina, los materiales y la computación.

La Nanociencia se define como el estudio y manipulación de los materiales a escala atómica, molecular y macromolecular, en la cual las propiedades difieren significativamente de las observadas a mayor escala^{2,4}. Este campo de investigación es altamente multidisciplinario, involucrando disciplinas como la Física ⁵, Química ^{6, 7} y Biología ^{8, 9}. La Nanociencia es considerado una rama fundamental de la Nanotecnología y proporciona un marco teórico para el desarrollo de la microelectrónica ¹⁰ y otros avances tecnológicos en diversas áreas de investigación.

Uno de los pilares fundamentales de la nanociencia es el estudio de los conglomerados atómicos (clústeres atómicos) debido a su composición, tamaño y propiedades únicas. Estos clústeres pueden estar constituidos por uno o varios tipos de elementos, y son estudiados bajo diferentes tipos de ambientes como vacío, suspensiones coloidales o sobre superficies¹¹. Los clústeres atómicos exhiben

diferentes comportamientos de interés, como su estabilidad relativa¹² o sus novedosas estructuras^{1, 13-18}.

Los clústeres atómicos son interesantes debido a que presentan propiedades electrónicas, magnéticas y ópticas diferentes a aquellas propias de átomos aislados o del material *bulk*^{19, 20}. Además, sus propiedades dependen del tamaño, la distribución y la composición del clúster. Por lo tanto, pueden considerarse un tipo de material por sí mismos si se encuentran bien definidos. Algunos clústeres pueden incluso ser diseñados para mantener su geometría tras interactuar con otros clústeres, llamados *superátomos*¹², y utilizarse como bloques de construcción para nanomateriales.

Es importante conocer la geometría de los clústeres atómicos, ya que propiedades fundamentales como la energía del sistema o su estructura electrónica dependen estrechamente de estas geometrías.

La búsqueda estructural de los clústeres atómicos plantea un desafío debido a la necesidad de emular su comportamiento cuántico. Resolver la ecuación de Schrödinger de manera exacta para estos sistemas es un problema no trivial requiriendo diferentes aproximaciones y teorías para ser abordada. Algunas de estas aproximaciones son la aproximación de Born-Oppenheimer²¹, que separa las coordenadas electrónicas y nucleares; la teoría de Hartree-Fock, que considera una descripción de los electrones en términos de orbitales individuales; y la teoría del funcional de la densidad, que se basa en la densidad electrónica para describir el sistema.

Aun con la capacidad de calcular con exactitud la energía de cada estructura, persiste la incertidumbre sobre qué clústeres podrían ser experimentalmente viables. Si bien es factible producir clústeres atómicos en el laboratorio y medir ciertas propiedades relevantes²²⁻²⁵, determinar con precisión la estructura real de estos clústeres sigue siendo un desafío considerable.

Con los ajustes adecuados es posible favorecer la estructura termodinámicamente más estable en los experimentos. Esta posibilidad es fundamental para la caracterización estructural mediante estudios teóricos. Es decir, la estructura real de los clústeres atómicos será aquella que posea la menor energía, conocida como mínimo global. Por lo tanto, los estudios teóricos se centran en buscar y predecir las configuraciones con la menor energía, para luego sobre estas simular la propiedad medida experimentalmente y por comparación con su contraparte experimental verificar la estructura real²⁶.

Dentro de la química computacional es posible estudiar, analizar y diseñar clústeres atómicos. En estos estudios, un campo de investigación abierto es la optimización de los procesos de búsqueda de estructuras de mínima energía sobre la SEP. Esto se puede abordar desde distintas direcciones: minimizar la cantidad de cálculos de energía a nivel cuántico, utilización de enfoques heurísticos para caer a buenos resultados de manera dirigida o estocástica, automatización dirigida y conocimiento propio del investigador, entre otros. En este contexto, el uso de diversos descriptores químicos con el fin de guiar la búsqueda de estructuras de mínima energía representa una aproximación que permite combinar el conocimiento químico del investigador con la eficacia computacional de los algoritmos de búsqueda y optimización.

Los métodos más populares son la dinámica molecular²⁷⁻²⁹, algoritmos estocásticos³⁰⁻³³, métodos evolutivos³⁴⁻⁴¹ y estrategias híbridas heurísticas que guían la exploración de mínimos energéticos hacia zonas con mayor probabilidad de presentar un GM⁴²⁻⁴⁴.

Este trabajo de investigación está estructurado en capítulos que abordan temas específicos. El Capítulo II ofrece un resumen conciso del marco teórico relacionado con los métodos clásicos de la mecánica cuántica, plantea los desafíos asociados con la exploración de la SEP y proporciona un breve panorama de los principales

métodos de optimización en uso hoy en día. Por su parte, el Capítulo III presenta la hipótesis y los objetivos que guían este estudio.

Los Capítulos IV, V, VI, VII y VIII inician con resúmenes introductorios, diseñados para ofrecer al lector una rápida comprensión del contenido que se tratará. Específicamente, el Capítulo IV se enfoca en los métodos de búsqueda implementados, detallando las adaptaciones realizadas a cada uno y su aplicación en sistemas químicos de particular interés. Además, se profundiza en cómo algunos de los algoritmos de exploración han sido diseñados con base en la 'intuición química', la cual utiliza conocimiento previo para predecir el comportamiento estructural de los sistemas de interés.

Capítulo V formaliza la 'intuición química' al introducir un descriptor de reactividad química basado en DFT Conceptual, que sirve como guía para predecir clústeres atómicos en algoritmos estocásticos. El Capítulo VI profundiza en otro descriptor químico, la integral de las interacciones no covalentes, y explora su posible correlación con la energía de los clústeres moleculares. En el Capítulo VII se lleva a cabo una evaluación comparativa de las metodologías desarrolladas y estudiadas a lo largo de la tesis, destacando las ventajas y desventajas de cada enfoque. Finalmente, el Capítulo VIII detalla las herramientas diseñadas, abordando las características específicas consideradas durante su implementación y desarrollo.

CAPÍTULO II – MARCO TEÓRICO

2.1 SUPERFICIE DE ENERGÍA POTENCIAL

La superficie de energía potencial (SEP) es una representación gráfica de la energía de un sistema en relación con su geometría. La hiper-superficie de energía potencial es una noción que intenta recalcar la multidimensionalidad de la SEP debido a que en un sistema molecular de N átomos el número de coordenadas independientes que determinan la SEP es de $3N - 6$ ($3N - 5$ para un sistema lineal)⁴⁵.

En términos generales, los puntos críticos en una representación gráfica, como los máximos, mínimos y puntos de transición pueden ser determinados mediante la aplicación de métodos de optimización. Dichos puntos críticos poseen una importancia significativa, ya que ofrecen información valiosa acerca de las propiedades del sistema en estudio, incluyendo su estructura en el estado fundamental, estructuras metaestables y los estados de transición.

La energía que da forma a la SEP puede ser descrita por potenciales atómicos empíricos, como Lennard-Jones^{46, 47}, y/o ecuaciones que incluyen la contribución electrónica como la de Schrödinger independiente del tiempo⁴⁸. Esta última es considerada como el modelo más viable, dentro de la mecánica cuántica, para el estudio de las propiedades físicas de estos sistemas. Sin embargo, su resolución no es trivial por lo cual se recurre a diversas aproximaciones para facilitar su solución.

La ecuación de Schrödinger es el fundamento principal de gran parte de la química cuántica y física de materiales. Permite obtener la energía de un sistema aplicando un operador Hamiltoniano (\hat{H}) a la función de onda multieléctronica (Ψ) del sistema.

$$\hat{H}\Psi = E\Psi \quad (2.1)$$

Bajo la aproximación de Oppenheimer^{21, 49}, el Hamiltoniano no relativístico multieléctronico toma la siguiente forma:

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{i,A} \frac{ZA}{R_{iA}} + \sum_{i>j} \frac{1}{r_{ij}} \quad (2.2)$$

En esta notación, los subíndices i y A se refieren a los electrones y átomos respectivamente; Z representa el número atómico, R_{iA} la distancia entre el núcleo y el electrón y r_{ij} la distancia entre electrones. Cabe destacar que la función de onda multieléctronica contiene toda la información relevante del sistema en estudio. Sin embargo, el último término, que representa la interacción electrón-electrón, plantea un problema de muchos cuerpos que resulta intratable desde un enfoque analítico.

Para abordar la complejidad inherente a este problema, es esencial recurrir a aproximaciones. En este contexto, los métodos convencionales suelen emplear aproximaciones basadas en orbitales. El método de Hartree-Fock (HF)⁵⁰⁻⁵³ utiliza un determinante de Slater como una aproximación a las funciones propias del Hamiltoniano, empleando funciones base para este fin. Mediante el método variacional, que postula que la función de onda del estado fundamental minimiza el valor esperado del Hamiltoniano, se derivan ecuaciones monoeléctronicas. Éstas se resuelven iterativamente a través del método de campo autoconsistente (*Self-Consistent Field*, SCF).

Este procedimiento proporciona resultados precisos en el cálculo de la energía de un sistema al tener en cuenta la componente de intercambio electrónico. No obstante, no tiene en cuenta la correlación electrón-electrón, que es una interacción compleja y crucial en sistemas con múltiples electrones. La correlación electrónica surge debido a las repulsiones y atracciones mutuas entre los electrones, lo que influye significativamente en las propiedades y comportamientos del sistema. Por lo tanto, HF se considera una aproximación de un solo determinante y no capta completamente la correlación electrónica.

No obstante HF sigue siendo considerado un método bien definido y robusto a la hora de calcular la energía de los sistemas atómicos. En el proceso HF, los orbitales moleculares se aproximan mediante un conjunto de funciones bases gracias al método de Combinación Lineal de Orbitales Atómicos (*Lineal Combination of Atomic Orbitals*, LCAO)⁵⁴. Esta aproximación ha impulsado el desarrollo de un campo de estudio dedicado a definir estas funciones base de manera eficiente. Este enfoque involucra el uso de métodos teóricos y/o empíricos para lograr una descripción adecuada y homogénea del comportamiento de los sistemas químicos.

Existen otras aproximaciones basadas en la función de onda, llamadas *post-Hartree-Fock*. Entre ellas, la teoría de perturbación Møller-Plesset^{55, 56}, en especial MP2⁵⁷ y los métodos de cúmulos acumulados (*Coupled Cluster*)^{58, 59}, siendo el formato "Singles Doubles and perturbative Triples" (CCSD(T))⁵⁸ el más utilizado. Estos métodos avanzados permiten tener en cuenta efectos de correlación electrónica más allá de la aproximación de HF, mejorando así la precisión de los cálculos. El método MP2 se basa en una expansión perturbativa de la energía, mientras que los métodos de cúmulos acumulados, como el CCSD(T), incluyen correlaciones electrón-electrón de manera más rigurosa. Estas aproximaciones se utilizan para obtener resultados más precisos a costa del aumento considerable de recursos computacionales.

2.2 TEORÍA DEL FUNCIONAL DE LA DENSIDAD

En 1964 Hohenberg y Kohn⁶⁰ publicaron las bases de lo que hoy se conoce como la teoría del funcional de la densidad (DFT)⁶¹ consistiendo en dos teoremas: (i) que la densidad electrónica, $\rho(r)$, en su estado fundamental determina todas las propiedades del sistema, y (ii) que la densidad electrónica del estado basal es aquella que minimiza el funcional de energía.

$$E[\rho(r)] = F[\rho(r)] + \int \rho(r) v(r) dr \quad (2.3)$$

Donde $E[\rho(r)]$ es el funcional de la energía, $v(r)$ es el potencial externo y $F[\rho(r)]$ es el funcional universal definido como:

$$F[\rho(r)] \equiv T[\rho(r)] + V_{ee}[\rho(r)] \quad (2.4)$$

Con $T[\rho(r)]$ como el funcional de la energía cinética y $V_{ee}[\rho(r)]$ correspondiente a la interacción electrón-electrón. Según DFT, la densidad electrónica proporciona una descripción completa del sistema, y al minimizar el funcional de energía, se pueden obtener las propiedades en su estado fundamental.

Por lo tanto, si se conoce la densidad del sistema es posible conocer la energía.

$$E = E[\rho(r)] \quad (2.5)$$

Al año siguiente Kohn y Sham⁶² propusieron un sistema ficticio donde los electrones son no interactuantes y que tienen como densidad en su estado basal la misma densidad del sistema real de interés, donde los electrones sí interactúan. Bajo este esquema, el funcional de la densidad ($E[\rho(r)]$) queda de la siguiente manera:

$$E[\rho(r)] = T_{ni}[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + \Delta T[\rho(r)] + \Delta V_{ee}[\rho(r)] \quad (2.6)$$

Donde los términos referencian, respectivamente, a la energía cinética de los electrones no interactuantes, interacción núcleo-electrón, repulsión electrón-electrón clásica, la corrección a la energía cinética producto de la naturaleza interactuante de los electrones, y todas las correcciones a la energía de repulsión electrón-electrón. La ecuación 2.6 puede ser reescrita como:

$$E[\rho(r)] = \sum_i^N \left(\left\langle \chi_i \left| -\frac{1}{2} \nabla_i^2 \right| \chi_i \right\rangle - \left\langle \chi_i \left| \sum_k^{nucleo} \frac{Z_k}{|r_i - r_k|} \right| \chi_i \right\rangle \right) + \sum_i^N \left\langle \chi_i \left| \frac{1}{2} \int \frac{\rho(r')}{|r_i - r'|} dr' \right| \chi_i \right\rangle + E_{xc}[\rho(r)] \quad (2.7)$$

Con N como el número de electrones y unificando los términos correctivos en E_{xc} , denominada la energía de intercambio correlación. Luego, continuando con el proceso de identificar los orbitales que minimizan la energía, se llega a que satisfacen la ecuación:

$$h_i^{KS} \chi_i = \varepsilon_i \chi_i \quad (2.8)$$

Con

$$h_i^{KS} = -\frac{1}{2} \nabla_i^2 - \sum_k \frac{Z_k}{|r_i - r_k|} + \int \frac{\rho(r')}{|r_i - r'|} dr' + V_{xc} \quad (2.9)$$

Donde

$$V_{xc} = \frac{\delta E_{xc}}{\delta \rho} \quad (2.10)$$

Como el operador monoeléctrico de Kohn y Sham. Entregando un marco formal para el cálculo energético, donde se utiliza un proceso autoconsistente (SCF) mejorando la precisión de las energías obtenidas y de la densidad electrónica trabajada.

En la teoría, DFT es formalmente exacta, pero en la práctica, su precisión depende en gran medida del funcional de intercambio-correlación utilizado. La principal ventaja de utilizar la densidad electrónica en lugar de las coordenadas electrónicas es que se reduce la cantidad de variables de $3N$ (donde N es el número de átomos) a solo 3, independientemente del número de partículas en el sistema. Esta simplificación en el número de variables ayuda significativamente en el cálculo de la energía electrónica, lo que a su vez mejora la eficiencia del proceso de cálculo.

La desventaja de DFT es el funcional de intercambio y correlación, el cual no tiene una forma exacta conocida y debe aproximarse mediante parámetros teóricos y empíricos. Esta aproximación introduce cierto grado de incertidumbre en los resultados y puede afectar la precisión de las predicciones. Los funcionales de intercambio y correlación se diseñan con el objetivo de obtener resultados coherentes con los experimentos, especialmente en lo que respecta a las energías de los sistemas. Sin embargo, la elección del funcional adecuado para un sistema en particular puede ser un desafío y requiere un conocimiento profundo de las propiedades del sistema y de los funcionales disponibles.

Uno de los primeros funcionales utilizados en DFT es el de la aproximación de la densidad local (LDA)⁶³, que asume una homogeneización de la densidad electrónica en todo el sistema. A pesar de la simpleza de su modelo, ha logrado obtener resultados razonables en términos de energías de sistemas. A medida que se avanzaba en el desarrollo de la DFT, se encontró que incluir información sobre el gradiente de la densidad electrónica mejoraba la precisión de los resultados. Desarrollándose entonces los funcionales de gradiente generalizados (GGA)⁶¹, que consideran tanto la densidad electrónica como su gradiente en su definición.

Posteriormente diferentes aproximaciones fueron desarrolladas, el funcional B3LYP⁶⁴⁻⁶⁷, por ejemplo, combina información empírica con DFT; los funcionales meta-GGA consideran la información de derivadas superiores aumentando al descripción de la densidad electrónica; y los funcionales híbridos integran de manera lineal un funcional DFT con la aproximación HF como por ejemplo PBE0⁶⁸.

2.2.1 DESCRIPTORES DE REACTIVIDAD

La modelización de la reactividad química fue el objetivo principal de lo que hoy se conoce como teoría conceptual de la densidad DFT⁶⁹⁻⁷⁴, desarrollada en la década de los 80 por Parr y Yang. Bajo este nuevo marco formal, varios químicos teóricos fueron capaces de entregar definiciones formales para diversos descriptores de reactividad⁵¹ que durante décadas habían descrito las interacciones químicas como la electronegatividad y el potencial químico por nombrar algunos.

Dependiendo el nivel de respuesta del sistema ante perturbaciones externas, los descriptores de reactividad pueden ser clasificados como descriptores globales o descriptores locales.

Los descriptores globales miden la respuesta global de un sistema frente a perturbaciones, manteniendo el potencial externo constante. Algunos de los descriptores globales más importantes son el potencial químico (μ), la dureza global (η) y la electrofilicidad (ω).

El potencial químico mide la tendencia de los electrones a escapar del equilibrio y por lo tanto se relaciona con la disponibilidad de los electrones para participar en reacciones químicas.

Por el contrario, la dureza global indica la resistencia a la transferencia de electrones²¹, es decir, su rigidez a la hora de interactuar con otros sistemas.

El índice de electrofilicidad representa la energía de estabilización del sistema cuando es saturado por electrones del ambiente.

La DFT Conceptual entregó los fundamentos matemáticos necesarios para definir los descriptores de reactividad⁷⁴. Suponiendo el enfoque matemático es posible rescatar diferentes aproximaciones. Por ejemplo, la aproximación de diferencias finitas es un método matemático para resolver ecuaciones diferenciales. En el contexto de los descriptores de reactividad, se utiliza la información de los sistemas en su estado aniónico (ϵ_{N-1}) y catiónico (ϵ_{N+1}).

La teoría de los orbitales frontera⁷⁵ propone que las reacciones químicas se pueden describir mediante la interacción del orbital molecular más alto ocupado (*Highest Occupied Molecular Orbital*, HOMO) de un reactante y el más bajo desocupado (*Lowest Unoccupied Molecular Orbital*, LUMO) del otro. Dentro de esta teoría, la aproximación de Koopmans⁷⁴ es una simplificación que asume que la ionización de un electrón no afecta significativamente a los demás electrones del sistema, por lo tanto los únicos orbitales necesarios corresponden al HOMO y LUMO.

La tabla 2.1 muestra las ecuaciones según la aproximación utilizada.

Descriptores de Reactividad

	Aproximación Koopmans	Aproximación Diferencias Finitas	Ref
Potencial Químico (μ)	$\mu \approx \frac{1}{2}(\epsilon_L + \epsilon_H)$	$\mu \approx -\frac{1}{2}(I + A)$	21
Dureza (η)	$\eta \approx \frac{1}{2}(\epsilon_L - \epsilon_H)$	$\eta \approx \frac{1}{2}(I - A)$	21, 76
Potencial de Ionización (I)	$I = -\epsilon_H$	$I = \epsilon_{N-1} - \epsilon_N$	77

Electroafinidad (A)	$A = -\epsilon_L$	$A = \epsilon_N - \epsilon_{N+1}$	77
Electronegatividad (X)	$\chi \approx -\frac{1}{2}(\epsilon_L + \epsilon_H)$	$\mu \approx \frac{1}{2}(I + A)$	74, 76
Electrofilicidad (ω)	$\omega = \frac{\mu^2}{2\eta}$	$\omega = \frac{\mu^2}{2\eta}$	21, 74

Tabla 2.1: Ecuaciones de los descriptores de reactividad según aproximación utilizada.

Por otro lado, los descriptores de reactividad local nos indican la respuesta de cada región del sistema ante perturbaciones externas, efectivamente dándonos información local del sistema. La función de Fukui es uno de estos descriptores e indica que sectores de una molécula están asociados con la máxima respuesta del potencial químico. Es decir, representan la respuesta del sistema ante perturbaciones simultáneas en el número total de electrones y el potencial externo debido a cargas positivas compensativas del sistema⁷⁸.

Propuestas por Parr y Yang⁶¹, las funciones de Fukui están basadas en DFT y se describen según la ecuación⁷⁸:

$$f(r) = \left(\frac{\delta\mu}{\delta v(r)} \right)_N = \left(\frac{\partial p(r)}{\partial N} \right)_{v(r)} \quad (2.11)$$

Donde $v(r)$ representa el potencial externo; N , el número total de electrones y μ , el potencial químico.

Debido a la discontinuidad de energía, de la densidad y otras propiedades moleculares por las diferencias mínimas en el número de electrones (N)⁷⁹, la ecuación 2.11 se divide para introducir las derivadas de izquierda y derecha, ambas consideradas para un número dado de electrones N .

$$f^+(r) = \left(\frac{\partial p(r)}{\partial N} \right)_{v(r)}^+ \quad (2.12)$$

$$f^-(r) = \left(\frac{\partial p(r)}{\partial N} \right)_{v(r)}^- \quad (2.13)$$

La ecuación 2.12 es la función de Fukui para un ataque nucleofílico, donde el sistema busca aceptar electrones. Por su parte, la ecuación 2.13 se usa para un

ataque electrofílico, donde el sistema busca donar electrones⁷². Se han propuesto diversas aproximaciones para calcular la función de Fukui como la aproximación de Koopmans^{80, 81}, Diferencias Finitas y *Orbital Weighed*⁸².

La función de Fukui, al estar definida en cada punto de la región del sistema, puede ser complicada de analizar en tres dimensiones debido a su naturaleza espacialmente extensa. Para simplificar su análisis, se utiliza la forma condensada de la función de Fukui, término acuñado por Mortier⁸³. Esta aproximación permite asignar un valor de la función de Fukui a cada átomo del sistema mediante el uso de las cargas de Mulliken. Permitiendo a los investigadores identificar rápidamente qué átomos del sistema son los más reactivos.

Otra aproximación para condensar la función de Fukui está inspirada en el análisis topológico de la densidad electrónica, propuesta por Bader⁸⁴, para definir un átomo dentro de una molécula (*Quantum Theory of Atoms in Molecules*, QTAIM). La función de Fukui es una función escalar en un espacio tridimensional, por lo tanto, un análisis del campo gradiente permite obtener puntos críticos representativos (similar al análisis QTAIM). Este nuevo enfoque permite delimitar atractores que no estén necesariamente ubicados en posiciones atómicas y que son susceptibles a tener un significado químico.

Esta última metodología se utilizó para definir el criterio de *Maximum Matching*⁸⁵ con el fin de maximizar la interacción entre los atractores obtenidos del análisis topológico de la función de Fukui. Se realizaron entonces ensambles de clústeres grandes a partir de fragmentos electrofílicos y nucleofílicos.

Los descriptores de reactividad global y local, aunque se encuentran bien definidos en diversas aproximaciones, suelen tener un uso más descriptivo que predictivo en la química teórica y computacional. Abriendo, implícitamente un desafío a la comunidad científica en esta área.

Este avance debe ir de la mano con un desarrollo computacional acorde con los requerimientos de la investigación, como fue el avance de un algoritmo de

predicción desarrollado por el Yañez⁸⁶ que utiliza ensambles manuales entre especies según la función de Fukui, que entrega información de sectores con alta probabilidad de donar y aceptar electrones, junto con las ventajas de un algoritmo genético.

¿Será posible que la combinación de este tipo de descriptores de reactividad junto con las ventajas de los algoritmos de exploración de la SEP permita enriquecer el nivel de búsqueda, converger en buenos mínimos energéticos y utilizar eficientemente los recursos entregados?

2.3 INTERACCIONES NO COVALENTES

Las interacciones no covalentes (*Non-Covalent Interactions*, NCI)^{87, 88} son interacciones químicas entre átomos o moléculas que no implican el compartir o transferir electrones entre ellos. Son más débiles que los enlaces covalentes, que se crean al compartir electrones. En cambio, las NCI resultan de las fuerzas atractivas o repulsivas entre átomos o moléculas. Estas fuerzas pueden ser electrostáticas (relacionadas con cargas) o no electrostáticas (relacionadas con forma y tamaño). Los ejemplos comunes de NCI incluyen el enlace de hidrógeno, las fuerzas de van der Waals (vdW) y las interacciones iónicas.

El enlace de hidrógeno⁸⁹ es un tipo de fuerza atractiva que ocurre entre un átomo de hidrógeno unido covalentemente a un átomo altamente electronegativo, como oxígeno, nitrógeno o flúor, y otro átomo electronegativo. Este tipo de enlace justifica los altos puntos de ebullición y fusión de algunas sustancias, como por ejemplo el agua.

Las fuerzas de vdW son fuerzas atractivas o repulsivas que ocurren entre moléculas debido a los momentos dipolares fluctuantes de los electrones en las moléculas. Estas fuerzas pueden ser atractivas (fuerzas de London) o repulsivas (fuerzas de Keesom). Por otro lado, las interacciones iónicas son fuerzas de atracción o repulsión electrostáticas entre iones por diferencias electrónicas.

Las NCI juegan un rol importante en muchos procesos biológicos, incluyendo el plegamiento y estabilidad de las proteínas ⁹⁰, la función enzimática y el reconocimiento/unión de ligandos a receptores ⁹¹. También juegan un papel en las propiedades físicas y químicas de las sustancias, como la solubilidad y el punto de ebullición de los compuestos ⁹².

Dado que las NCI juegan un papel importante en la determinación de propiedades y estabilidad química, nace la necesidad de mapear y visualizar estas interacciones de manera eficiente. Los métodos usuales están basados en la densidad electrónica, como es el caso de *Quantum Theory of Atoms in Molecules (QTAIM)*⁸⁴, o mediante la función de localización electrónica (*Electron Localization Function, ELF*)⁹³. Sin embargo, estas metodologías requieren de cálculos previos que muchas veces significan un alto gasto de recursos.

Por último, las NCI pueden ser mapeadas utilizando solamente la densidad electrónica ($\rho(r)$) y su gradiente reducido ($s(r)$). Dependiendo los valores adoptados, es posible advertir la presencia y la intensidad de las NCI en un sistema químico sin la necesidad de realizar cálculos demandantes.

Formalmente el gradiente de densidad reducido tiene la siguiente forma⁹⁴:

$$s(r) = \frac{1}{C_s} \frac{|\nabla\rho(r)|}{\rho(r)^{4/3}} \quad C_s = 2(3\pi^2)^{1/3} \quad (2.14)$$

En lugares donde la densidad electrónica es pequeña (regiones lejanas de la molécula) el gradiente reducido usualmente tendrá un valor alto. Sin embargo, valores cercanos a cero indican zonas relacionadas con enlaces covalentes o vdW. Este fenómeno se explica porque ante la presencia de interacciones débiles, existe un cambio en el gradiente reducido entre átomos interactuantes, lo cual produce puntos críticos de densidad entre fragmentos interactuantes.

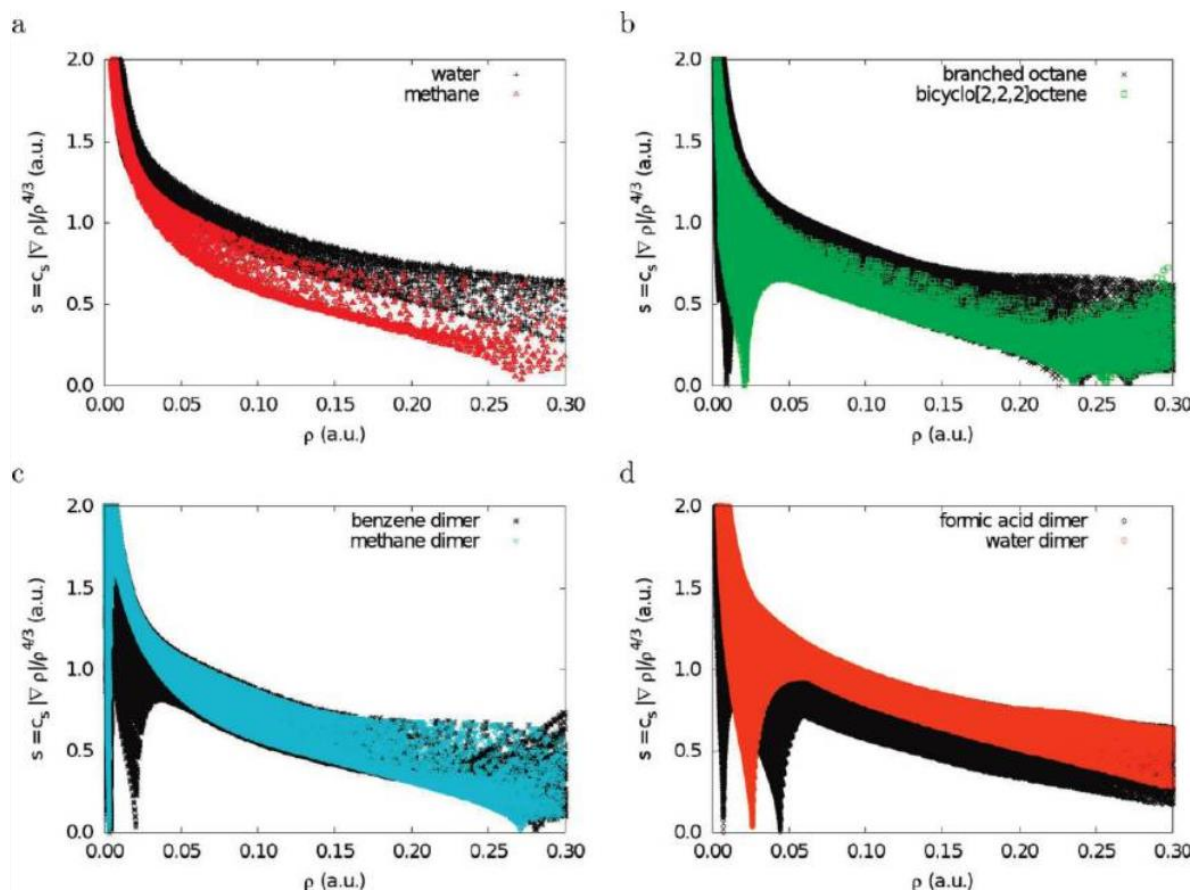


Figura 2.1. Gráficos de densidad electrónica y gradiente reducido para moléculas de agua, metano, octano y octano bicyclo[2,2,2]; y sus dímeros moleculares. Imagen tomada de la referencia ⁹⁵.

El trabajo de Contreras *et al*⁹⁵ concluyó que es posible mapear las NCI utilizando los valores de la densidad electrónica y su gradiente. En específico, cuando ambos descriptores se acercaban a un valor 0 en los sectores de formación de dímeros se infería la existencia de un tipo de NCI como ilustra la Figura 2.1.

Es posible diferenciar los tipos de NCI al analizar el Laplaciano de la densidad, en específico la descomposición en los tres ejes. Uno de los componentes resultantes, λ_2 , revela el tipo de interacción. Donde interacciones enlazantes son aquellas con $\lambda_2 < 0$, repelentes cuando $\lambda_2 > 0$ y valores cercanos a 0 pueden ser considerados como interacciones de vdW.

2.3.1 NCI PLOT

El cálculo del gradiente reducido fue implementado en un software, NCI PLOT⁹⁶, que permite el análisis, clasificación y visualización de las NCI mediante la entrega

de cálculos SCF o por medio de la aproximación promolecular⁹⁷ que solo requiere las coordenadas cartesianas del sistema.

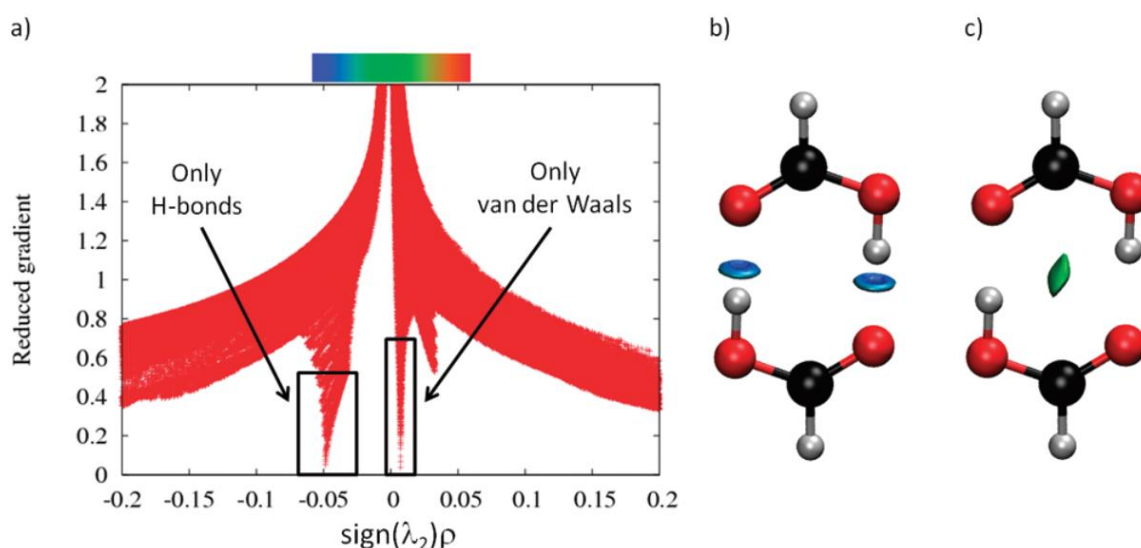


Figura 2.2 Salida del programa NCIPLOT. (a) indica los tipos de NCI según λ_2 y el valor de *cutoff* escogido en el gradiente reducido; (b) representación gráfica de las interacciones no covalentes para el dímero de ácido fórmico. Imagen tomada de la referencia ⁹⁶.

NCIPLOT recibe información del usuario sobre el sistema a analizar, luego engrilla el espacio de trabajo considerando los sectores más críticos de interacción. El valor de la densidad electrónica en cada celda de la grilla permite distinguir los sectores probables a tener NCI. Realiza el análisis con el gradiente reducido de la densidad y hace uso de λ_2 para clasificar cada tipo de NCI. La Figura 2.2 representa un ejemplo de la salida del programa donde se pueden visualizar dos tipos de NCI según el valor *cutoff* del gradiente reducido de la densidad.

Su última iteración, NCIPLOT4⁹⁸, agrega la capacidad de cuantificar las propiedades de las regiones correspondientes a NCI mediante la integración del espacio (I_n)⁹⁵. Esta cuantificación será llamada “Índice de interacciones no covalentes” (NCIi).

Los valores cuantitativos se obtienen al integrar el espacio NCI (Ω_{NCI}) entre densidades promoleculares⁹⁷, en específico sectores donde la densidad proviene de

la interacción entre ambos fragmentos, donde el gradiente de la densidad reducida es pequeño. Así el valor de la integración del NCI queda:

$$I_n = \int_{\Omega_{NCI}} \rho^n(r) dr \quad n = 1.0, 4/3, 1.5, 5/, 2.0, 2.5, 3.0 \quad (2.15)$$

Donde los valores de n varían la correlación entre la energía de interacción y el valor de la integral según explica Contreras *et al*⁹⁸.

No obstante, los análisis con NCI tienen como objetivo explicar y describir problemas moleculares relevantes. Proporciona una comprensión de por qué se producen ciertas interacciones y puede ayudar a determinar, por ejemplo, la validez de un tipo de plegamiento proteico en comparación con otro o la mejor conformación entre dos moléculas de agua.

A pesar de esto, la capacidad predictiva de estas aproximaciones no ha sido demostrada. La naturaleza compleja de las interacciones no covalentes y la diversidad de factores que influyen en ellas hacen que la predicción precisa de estas interacciones sea un desafío.

2.4 ALGORITMO DE OPTIMIZACIÓN Y EL PROBLEMA DEL MÍNIMO GLOBAL

Tal y como se ha discutido a lo largo de esta tesis, existen diversos métodos para el cálculo energético de los sistemas. Los sistemas están descritos por interacciones moleculares, que pueden formarse producto de reacciones fragmentos. En dicho contexto, los descriptores de reactividad son herramientas fundamentales para entender el comportamiento y propiedades de estos sistemas.

No obstante, la búsqueda del mínimo global (GM) necesita de métodos de cálculo energético y una capacidad de obtención de varias estructuras representativa rápidamente. Uno de los enfoques de la química computacional es el barrido de la SEP, esto es distinguir y definir todos los puntos críticos del sistema, incluido el GM, el cual es la representación fidedigna del sistema experimentalmente, considerando que la mayoría de los sistemas químicos tienden a estar en su forma más estable energéticamente.

La imposibilidad, en la práctica, de calcular las energías de un sistema en todas sus configuraciones estructurales posibles convierte la búsqueda del GM, en un problema de optimización global que lograrse identificar buenos resultados sin necesariamente calcular todas las posibilidades.

Los problemas de optimización global no son propios de la química computacional, de hecho, son comunes en un sinnúmero de diversos campos de investigación (logística⁹⁹, diseño de redes¹⁰⁰, ámbitos biológicos¹⁰¹ y químicos), y comparten el mismo objetivo: identificar los puntos críticos relevantes de una función $f(x)$.

Hoy en día se sabe que la cantidad de átomos hace variar el número de mínimos energéticos disponibles en la SEP, incrementándolos de manera exponencial¹⁰², lo que aumenta la cantidad de tiempo y recursos en la búsqueda del GM.

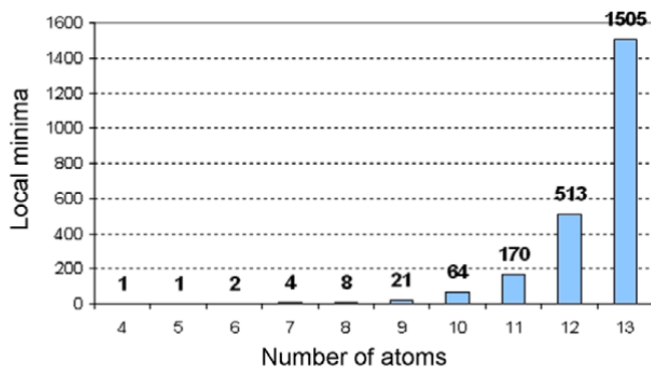
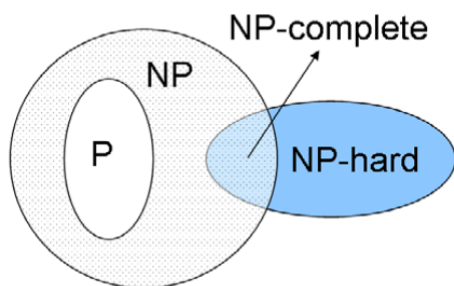


Figura 2.3: A la izquierda se muestra una esquematización de los tipos de problemas P, NP, NP-complete y NP-hard; a la derecha el crecimiento exponencial de mínimos locales a medida que aumenta el tamaño del clúster. Imagen tomada de la referencia ¹⁰³.

Independiente de la importancia en identificar el GM, la tarea en la práctica ha resultado extremadamente difícil de realizar. En la teoría de complejidad computacional, los problemas son clasificados en P, NP, NP-complete y NP-hard según el uso de recursos, como ilustra la Figura 2.3. El trabajo de Wille y Venik¹⁰⁴ clasificó el problema de optimización como NP-hard para clúster homonucleares, siendo generalizado para clúster heteronucleares¹⁰⁵ tiempo después. Esto quiere decir que está comprobado la inexistencia de una solución general que certifique con 100% de precisión el GM. Su identificación dependerá de las características de cada molécula en cuestión.

En las últimas décadas, se han aplicado diferentes métodos heurísticos^{106, 107} que recorren la SEP de manera determinística o estocástica para encontrar el GM de un sistema. En esos casos, las variables más importantes a optimizar son el tiempo de cálculo y los recursos utilizados.

Los métodos determinísticos son aquellos que optimizan el gradiente de la función, el más conocido y utilizado es el método de Newton¹⁰⁸ que utiliza el Hessiano para discernir la dirección a optimizar. No obstante, muchas veces la función optimizada es multivariable, muy complicada o inclusive desconocida, dificultando la tarea de optimización.

Los métodos estocásticos son aquellos que utilizan técnicas con cierto nivel de aleatoriedad para optimizar el problema. Permiten una exploración más amplia de la SEP, pero fallan a la hora de converger en puntos críticos específicos una vez definidos sectores de mínima energía. Para combatir este problema nacen los algoritmos heurísticos que combinan las estrategias estocásticas y junto con información adicional que permita guiar a buenos resultados.

Los algoritmos heurísticos (Figura 2.4) son ideales para la generación de buenos candidatos a solución como son los algoritmos evolutivos¹⁰⁹, *Particle Swarm Optimization*³⁸ y *Ant Colony Optimization*⁴⁰. La medida de calidad que poseen los algoritmos heurísticos se conoce como *fitness*, la cual puede ser resultado de una función matemática, un valor estadístico, una cualidad contable que mejora entre estados, etc. La medida de *fitness* más simple es la comparación de la mejor energía obtenida en un punto del algoritmo contra la energía de un estado más avanzado, si la energía mejora el algoritmo asume buena convergencia y continua, por otro lado si el *fitness* empeora el algoritmo retrocede etapas o incluso da comienzo a su condición de finalización.

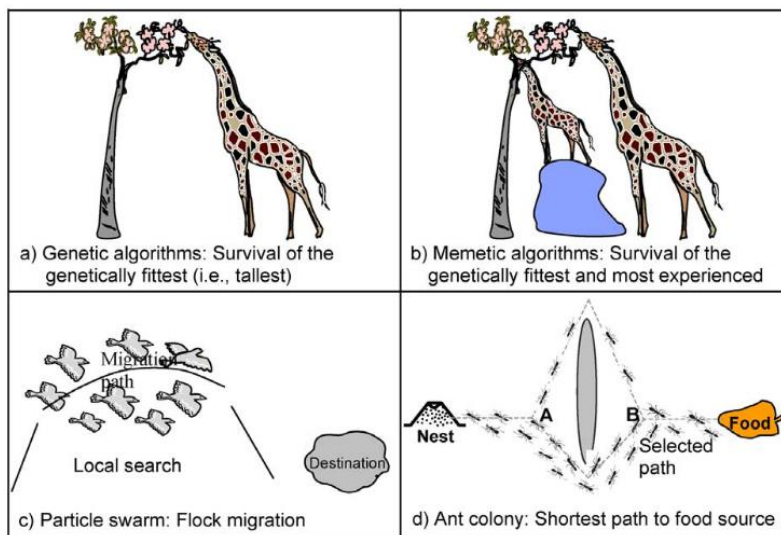


Figura 2.4 Representación simple de algunos algoritmos heurísticos. Imagen tomada de la referencia ⁴¹.

La estrategia más utilizada para explorar la SEP consiste en el uso de un algoritmo heurístico para identificar buenas “zonas” de soluciones y finalizar por medio de un algoritmo determinístico que permita refinar la búsqueda de la solución rápidamente, sin caer en una alta utilización de recursos.

2.4.1 PARTICLE SWARM OPTIMIZATION (PSO)

El algoritmo de enjambre de partículas (PSO, por sus siglas en inglés) es una técnica de optimización heurística que se inspira en el comportamiento social de animales como pájaros y abejas. Es un algoritmo de optimización basado en poblaciones que implica un grupo, o "enjambre", de soluciones potenciales llamadas partículas. Cada partícula representa una solución potencial al problema de optimización y tiene una posición en el espacio de búsqueda que corresponde a los valores de las variables de la solución.

Las partículas se mueven a través del espacio de búsqueda y actualizan sus posiciones basándose en sus propias experiencias y las experiencias de las demás partículas del enjambre, es decir, mediante mutaciones locales el sistema es dirigido hacia un punto crítico de la función con un mejor *fitness*. Las partículas utilizan sus experiencias para orientar su búsqueda hacia áreas del espacio de búsqueda que son más prometedoras. Los movimientos de las partículas están guiados por un conjunto de reglas diseñadas para fomentar la exploración y la explotación del espacio de búsqueda. De esta manera existen individuos que exploran la SEP cercana con el fin de evaluar su calidad, mientras que la mayoría de la población usará la información del mejor de los exploradores como medida de evolución.

Una de las características clave de PSO es que no requiere que el problema de optimización sea diferenciable o incluso continuo, lo que lo convierte en una herramienta útil para optimizar una amplia gama de problemas. También es

relativamente simple de implementar y se puede paralelizar fácilmente, lo que lo convierte en una opción popular para resolver problemas de optimización.

2.4.2 ANT COLONY OPTIMIZATION

Algoritmo de población que tiene como fin optimizar la ruta o camino desde un punto inicial **A** hacia un fin establecido **B**. Nuevamente vemos una técnica que imita la naturaleza, en este caso se inspira en el uso de feromonas por las hormigas trabajadoras para establecer el camino desde el hormiguero hacia la comida.

Una de las diferencias principales con otros algoritmos es la utilización del conjunto de resultados previos para la creación de un camino de solución. Es decir, permite identificar gracias a la mejora constante del “*fitness*” y recordar los descriptores más importantes a la hora de navegar el espacio de solución.

2.4.3 ALGORITMO GENÉTICO (AG).

El algoritmo evolutivo, basado en la teoría de la evolución, bajo la premisa de la sobrevivencia del más apto. A través de múltiples generaciones una población de individuos evoluciona mejorando sus características importantes y disminuyendo sus propiedades menos deseables. Ocupa operadores análogos a los usados en el proceso evolutivo de cruce entre especies, mutaciones y selección natural, para explorar los espacios multidimensionales de la SEP. El más utilizado es el algoritmo genético¹¹⁰ (AG) que le agrega un factor de elitismo al proceso, donde solo los mejores individuos pueden generar descendencia.

Este algoritmo tiene poblaciones de individuos (clúster) representados con un set de cromosomas (coordenadas atómicas). Tales individuos viven y evolucionan sobre la base de la selección natural y las leyes de adaptabilidad. Así, en cada generación la población va mejorando (evolucionando) hacia una estructura cercana al mínimo global, mientras que las conformaciones no aptas se descartan¹⁰³.

En general, el algoritmo genético trabaja en cada generación los siguientes pasos:

- ***Fitness***: es una medida de puntuación a la cual es sometido cada miembro de la población. La medida más simple es el cálculo de la energía potencial de la conformación actual. Sin embargo, existen medidas más elaboradas que toman en consideración otros factores del sistema. La medida *fitness* es, usualmente, normalizada y ocupada como determinante al escoger un miembro de la población para ser utilizado en el proceso de recombinación posterior.
- **Recombinación o Apareamiento**: es el método que permite mezclar la información de los individuos para crear un nuevo miembro en la población. Los individuos para recombinar son los mejores, según *fitness*, de la generación, la recombinación puede también incluir individuos de generaciones anteriores y es usualmente ciega a la predicción de buenos resultados, es decir, no es capaz de recombinar conscientemente buenos cruces que lleven al mejor individuo.
- **Mutación**: para evitar caer en un mínimo local, una cierta cantidad de individuos producidos se les hacen pequeños cambios, creando una generación mutada. De esta forma se puede explorar una región más grande de la SEP. Los cambios realizados en la etapa de mutación comprenden fenómenos de rotación y traslación de átomos, intercambios de posición entre ellos e incluso la generación de un individuo totalmente nuevo.
- **Selección**: es el proceso con el cual se escogen los individuos de la población con mejor *fitness*, los que serán utilizados en el ciclo siguiente. Por ejemplo, pueden ser escogidos los individuos mutados, los individuos hijos o, incluso, algunos individuos padres con buen *fitness*.

El proceso es iterado hasta llegar a una población consenso; es decir, una en la que siempre se escogen las mismas conformaciones. Finalmente, el individuo con el mejor *fitness* correspondería a un mínimo energético global.

El algoritmo ha sufrido modificaciones. Una de las más importantes fue la optimización hecha por Deaven y Ho³⁶, que consiste en agregar una optimización local para cada conformación, transformando la SEP en una superficie “escalera” luego de cada generación.

Son diversos los algoritmos en la química que hacen uso de la heurística para navegar eficientemente la SEP en búsqueda de mejores geometrías, o encontrar alguna propiedad de interés, se resumen algunos en la Tabla 2.2.

Nombre	Descripción	Lenguaje	Ref.
AUTOMATON	Combinación de autómatas celulares y GA.	Perl	35
GEGA	Algoritmo genético guiado por gradiente de energía.	C++	111
TGMin	Algoritmo tipo <i>Basin Hopping</i> , utiliza métodos de simetría.	Python	112, 113
PDECO	Algoritmo Evolutivo en poblaciones paralelas.	C	114
NWPEsSe	Variación de <i>Ant Colony algorithm</i> . Imitación de comportamiento de las abejas sondeando la SEP en busca de miel.	Sin información	115
CALYPSO	Algoritmo PSO de búsqueda estructural bajo diferentes condiciones.	Python	116

TABLA 2.2: ALGORITMOS DE OPTIMIZACIÓN GLOBAL.

2.4.3.1 AUTÓMATA CELULAR

Un autómata celular es un método matemático que imita el comportamiento del universo bajo una serie de reglas preestablecidas. Inventado por Jhon von Neuman¹¹⁷ quien diseñó una grilla bidimensional de celdas. Cada celda tiene asignado un estado y “evoluciona” con el paso del tiempo dependiendo de una función preestablecida. Por ejemplo, una celda **X** siempre está en un estado neutro,

se visualiza de color blanco, pero si uno sus vecinos están en un estado excitado, visualmente negro, su estado cambiará a estado excitado, cambiando de color. Los estados posibles de la celda siempre son neutros o excitado, la función preestablecida son las reglas que permiten el cambio de estado y el paso del tiempo permite evolucionar el proceso como se muestra en la Figura 2.5.

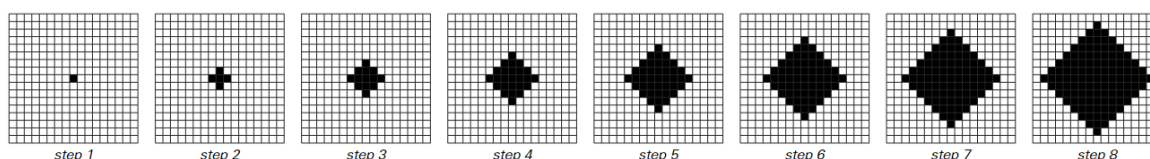


Figura 2.5: Representación evolución de autómatas celulares. Imagen tomada de la referencia ¹¹⁸.

Los autómatas celulares pueden ser diseñados en N dimensiones, en grillas de diferentes formas y con un número infinito de estados. Dependiendo de las reglas establecidas el mapa general de un autómatas celular puede caer en un punto de equilibrio, donde no existen más cambios, o continuar indefinidamente.

2.4.3.2 AUTOMATON

AUTOMATON es un algoritmo de búsqueda global de geometrías desarrollado por Yañez *et al.*,^{35, 119} dentro del grupo de investigación del profesor William Tiznado, como objetivo de su tesis doctoral. El presente tesista también trabajó en el desarrollo de este software que combina un algoritmo genético junto con la teoría de los autómatas celulares.

En AUTOMATON se usaron 3 estados posibles en una grilla de 1, 2 o 3 dimensiones donde las celdas interactúan entre sí. La Figura 2.6 muestra un ejemplo de construcción de un clúster molecular. En resumen, funciona de la siguiente manera.

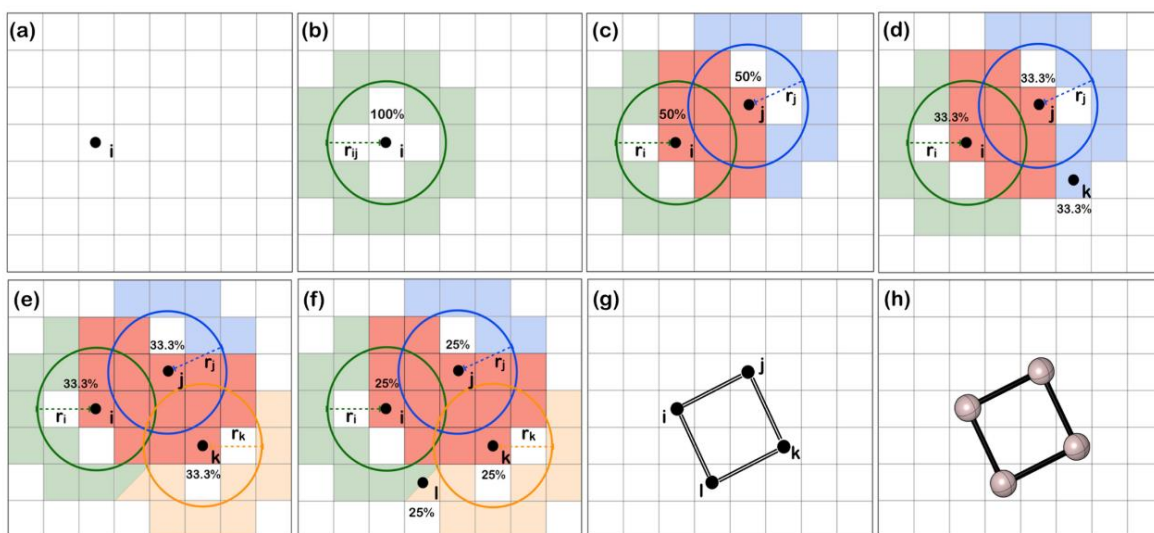


Figura 2.6 Funcionamiento de autómata celular en software AUTOMATON. (a) población de celda aleatoria, (b) generación celdas “posibles vecinos” (en verde), (c) nueva celda es poblada lo que genera nuevas celdas “posibles vecinos” (azul) y celdas “prohibidas” en rojo. (e) y (f) se repite el proceso nuevamente para generar un conjunto de celdas (g) que se traducirán en un clúster de cuatro átomos(h). Imagen tomada de la referencia ³⁵.

1. Una grilla inicial es creada, el tamaño de la grilla corresponderá al tamaño máximo del sistema según la información de los átomos covalentes del mismo dispuestos linealmente.
2. La grilla es dividida en celdas cúbicas de tamaño 0.3 \AA (radio covalente del hidrógeno).
3. Cada celda es asignada como estado “vacío”.
4. Una celda aleatoria es poblada con el primer átomo, su estado cambia a “poblada”. Este paso da inicio al proceso de autómata celular.
5. El radio covalente del átomo recientemente agregado es utilizado para construir una esfera de interacción.
6. Todas las celdas en la superficie de la esfera creada cambian su estado de “neutro” a “posible poblador”.
7. Todas las celdas dentro de la esfera formada cambian su estado a “prohibido” y se mantienen hasta el final del proceso.

8. Se escoge, aleatoriamente, una celda “posible poblador” para cambiar a un estado “poblado”, un átomo es puesto en la celda.
9. Se escoge el siguiente átomo a agregar y se vuelve al paso 5.

Este simple proceso permite la construcción de una población inicial que asegura la conexión de todos los átomos en una estructura única, así como también evita cualquier impedimento estérico. Luego AUTOMATON ocupa esta información en un algoritmo genético completo.

La tecnología del autómata celular puede ser extendida, enriqueciendo los procesos evolutivos de los sistemas iniciales bajo otro tipo de reglas o funciones con el fin de beneficiar la construcción inteligente de clústeres, moléculas o incluso macromoléculas.

2.4.4 ALGORITMO KICK

Un último algoritmo, el algoritmo KICK^{31, 33} es quizás la implementación más sencilla de un algoritmo estocástico, consiste en definir un espacio de interacción donde se ubican aleatoriamente algún fragmento de interés, por ejemplo átomos, moléculas o macromoléculas. Dado un número suficientemente alto de población se puede considerar que toda la SEP fue explorada, y que por lo tanto, se han evaluado todos los mínimos energéticos disponibles.

Sin embargo, este número suficientemente alto es del orden de millones, o miles de millones con un gran porcentaje de estructuras ineficaces, ilógicas o inexistentes. Una implementación mejorada de este algoritmo fue el desarrollado por Addicoat³⁰ donde los fragmentos interactuantes son rotados en sus ejes aleatoriamente previo a ser agregados en la caja de interacción.

Aunque el algoritmo KICK y sus variantes son útiles para explorar la SEP y encontrar mínimos energéticos en sistemas complejos, es importante tener en cuenta que siguen siendo un método estocástico y pueden requerir un tiempo computacional considerable. Además, la calidad de los resultados obtenidos

dependerá en gran medida de la elección adecuada de los parámetros y la interpretación adecuada de los resultados.

CAPÍTULO III – HIPÓTESIS Y OBJETIVOS

3.1 HIPÓTESIS

El uso de descriptores de reactividad química como guía en los métodos de optimización global conduce a la creación de enfoques híbridos altamente eficientes para la exploración exhaustiva de la superficie de energía potencial en clústeres y moléculas.

3.2 OBJETIVO GENERAL

Desarrollar y aplicar estrategias de optimización global utilizando descriptores de reactividad química junto con métodos exploratorios para el estudio de clústeres atómicos.

3.2.1 OBJETIVOS ESPECÍFICOS

1. Evaluar y seleccionar descriptores de reactividad química relevantes para la optimización global de clústeres atómicos.
2. Adaptar y/o desarrollar métodos exploratorios de la superficie de energía potencial en clústeres atómicos.
3. Diseñar algoritmos de optimización global que utilicen descriptores de reactividad y métodos exploratorios seleccionados.
4. Implementar estrategias diseñadas en clústeres atómicos de referencia.
5. Comparar las estrategias desarrolladas con métodos convencionales de exploración de la superficie de energía potencial.
6. Investigar la influencia de las interacciones no covalentes en la optimización global de clústeres atómicos.
7. Discutir la aplicabilidad de las estrategias desarrolladas en otros sistemas moleculares.
8. Identificar posibles mejoras y futuras proyecciones de la optimización global guiada por descriptores de reactividad química.

CAPÍTULO IV – ESTRATEGIAS DE EXPLORACIÓN DE LA SEP

“Este capítulo ofrece una revisión exhaustiva de los métodos desarrollados para explorar la superficie de energía potencial, empleando diversas estrategias de búsqueda. Se analizan múltiples sistemas, incluyendo tanto sistemas comunes como exóticos, átomos hipercoordinados planos, sistemas con interacciones metálicas, y estructuras que varían desde planas hasta tridimensionales, así como sistemas de diferentes tamaños. Además, se evalúa el rendimiento de los programas utilizados, destacando tanto sus mejoras como las desventajas y desafíos inherentes a cada enfoque.”. - Resumen de Capítulo.

4.1 GENERALIZACIÓN

La exploración de la SEP es una tarea imposible de abordar con solo una única solución. Las técnicas de búsqueda se adaptan a isoformas particulares o incluso a familia de sistemas que comparten algún tipo de patrón en común.

Y es este el concepto esencial para todo algoritmo de exploración, “**patrón**”, un patrón puede resultar muy simple para algunos sistemas, por ejemplo, los compuestos aromáticos son, en su mayoría, planos; los clústeres de boro tienden a formar triángulos en sus caras; los átomos de hidrogeno nunca están al centro de una molécula, etc.

Incluso, se puede argumentar que Mendeléyev diseñó la primera tabla periódica siguiendo el patrón de comportamiento de los elementos químicos descubiertos hasta el momento, llegando incluso a predecir las propiedades de aquellos desconocidos¹²⁰. La Teoría de los Enlaces de Valencia¹²¹ de Lewis permite también visualizar un patrón (reglas) de conectividad de los átomos, así como también las isoformas esperables que gobiernan la química orgánica.

Un patrón es entonces capaz de formar la regla, de establecer teorías y, por lo tanto, rigen nuestra forma de ver la realidad. Pero un patrón se puede romper y, así destruir los conceptos establecidos. En química son muchas los ejemplos de

sistemas atómicos o moleculares que no se comportan como se espera^{122, 123}. Las teorías y reglas saltan a la oportunidad de adaptarse para sumar el nuevo extraño comportamiento¹²⁴. Por ejemplo, el concepto de aromaticidad nace con el fin de explicar la estabilización de compuestos cíclicos a través de la deslocalización electrónica. Un área completa de estudio nació solo de este concepto, a través de la definición de deslocalización electrónica.

Entonces, a pesar de que los compuestos químicos pueden romper los patrones establecidos, es posible que en su descubrimiento y definición caigan en un patrón único propio. Es este el corazón de los algoritmos de exploración de la superficie de energía potencial. Encontrar los patrones de compuestos conocidos y explotarlos eficazmente en la identificación de mínimos energéticos.

Pero también existen otros tipos de algoritmos que, a partir de patrones conocidos y seguros, intentan identificar compuestos exóticos y que se alejan de las reglas preconcebidas. Estos algoritmos tienen como objetivo mejorar el conocimiento científico en este campo y establecer nuevas reglas de búsqueda y exploración.

A continuación, se hablará sobre varios métodos de búsqueda en la SEP y cómo se adaptaron para el descubrimiento de sistemas exóticos. Todos los programas mencionados se basan en conceptos químicos bien establecidos, como el radio covalente, las cargas atómicas y el comportamiento clásico de los elementos. Estos conceptos son fundamentales en la química y se utilizan como guías para explorar la diversidad de compuestos y materiales en la SEP.

El detalle de la implementación de los métodos desarrollados se puede leer en el Capítulo VIII.

4.2 DETALLES COMPUTACIONALES

Diferentes trabajos fueron realizados usando MatrixAutomaton y TemplatePermuta, en general todas las estructuras de los sistemas fueron obtenidas de la literatura correspondiente. Se sometieron a un proceso de

optimización utilizando el software químico-computacional Gaussian 16 versión B01¹²⁵ bajo el nivel de teoría PBE0⁶⁸/Def2-TZVP¹²⁶. A menos que sea especificado otro método.

El nuevo software “MatrixAutomaton” fue programado en Python3. Se agregaron las librerías de *numpy*¹²⁷ para cálculos matemáticos y *spicy*¹²⁸, una colección de algoritmos numéricos y estadísticos que permiten el análisis eficiente de características

Todas las exploraciones realizadas por medio de MatrixAutomaton se llevaron a cabo utilizando una población inicial de $5N$, donde N es el número de átomos en el sistema. Además, se utilizó un nivel de teoría PBE0/SDDAll^{129,130} y se necesitaron 9 ciclos para lograr la convergencia adecuada.

Los *template* escogidos en TemplatePermuta son específicos de los clústeres atómicos de interés en cada trabajo. Luego, fueron obtenidos de la literatura y refinados al nivel PBE0/Def2-TZVP.

4.3 RESULTADOS

4.3.1 MEJORAS AL PROCESO AUTOMATON

Paralelo al proceso de actualización del programa AUTOMATON (Capítulo VIII), ideas complementarias fueron desarrolladas con el objetivo de mejorar la búsqueda en clústeres con cualidades especiales.

Una de las ideas es la capacidad del programa de agregar átomos en un orden particular, similar a una muñeca *Matrioshka* (Figura 4.1), es decir, construir clústeres atómicos concéntricos, donde un tipo de átomo se encuentra “atrapado” por un segundo tipo, luego por un tercero, etc.

El sistema escoge un tipo de átomo y lo ubica en una grilla más compacta. Acabados los átomos de un tipo, la grilla es expandida, por lo tanto, los nuevos átomos rodearan a los antiguos. El proceso continuará según la estequiometría solicitada.

En el trabajo “*Evaluation of restricted probabilistic cellular automata on the exploration of the potential energy surface of Be₆B₁₁*”¹¹⁹ (Artículo 2, Anexo 2) se puso a prueba la mejora del programa en el clúster Be₆B₁₁ donde no solo se encuentran los isómeros reportados en literatura, sino que también aumentan la cantidad de mínimos locales y se expone un nuevo mínimo global.

Este trabajo se originó como un ejercicio realizado en las etapas iniciales de prueba de MatrixAutomaton. En trabajos anteriores, Guo y sus colaboradores¹³¹ identificaron los mínimos energéticos del sistema Be₆B₁₁ con una estructura en forma de hélice simétrica como el mínimo global. Ésto se realizó con software reconocidos en la comunidad científica como es Bilatu y con un numero alto de población en un proceso Kick.

Con el objetivo de replicar estos resultados sin utilizar el método KICK (población de 8000), diseñamos la estrategia de búsqueda concéntrica o *Matrioshka*. Aunque esta idea no era original, resultó ser efectiva para encontrar estructuras donde los átomos de boro rodearan a los átomos de berilio. Esta idea solo fue posible por el conocimiento previo del comportamiento de este sistema.

Mediante esta simple implementación, logramos replicar los resultados reportados por Guo y también descubrimos otro mínimo global que resultó ser aproximadamente 1.8 Kcal más estable que el mínimo previamente reportado. Aunque este nuevo mínimo global no exhibe una estructura llamativa o simétrica, mantiene el diseño original en el que los átomos de boro encierran a los átomos de berilio. Estos resultados fueron luego recopilados en un artículo científico¹¹⁹.

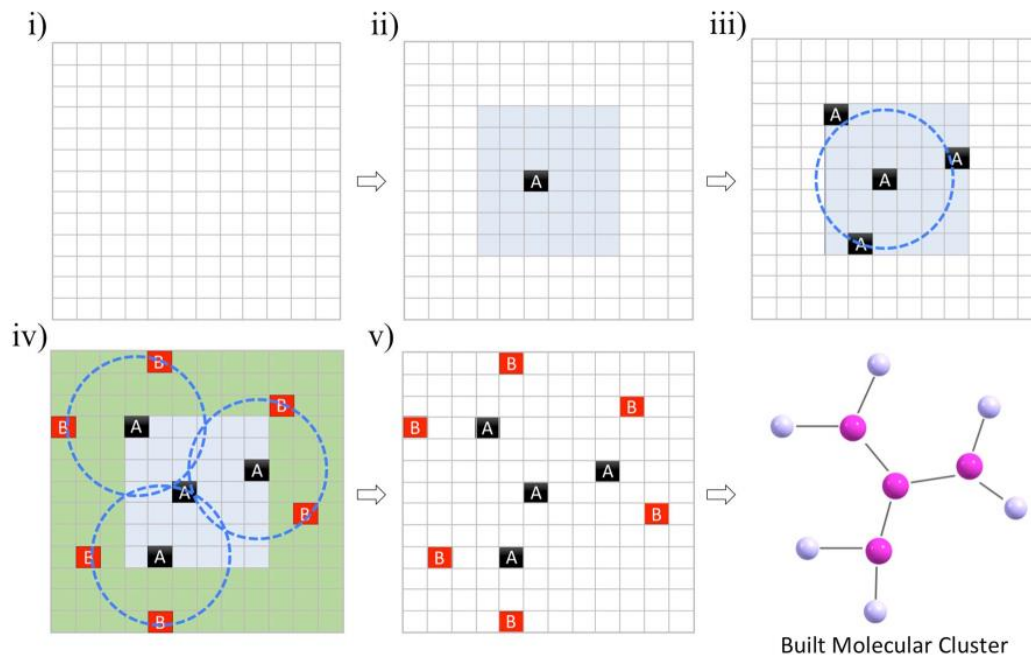


Figura 4.1 Representación 2D del procedimiento de construcción de autómatas celulares probabilísticos. i) definición del espacio de trabajo, ii) y iii) localización de agentes de interés en celdas (A), iv) y v) localización de agentes nuevos (B). Imagen tomada de la referencia ¹¹⁹.

4.3.2 USOS MATRIXAUTOMATON -CARBONOS HIPERCOORDINADOS PLANOS

A finales del año 2020 se trabajó con el compañero tesista Luis Leyva en la primera identificación del esquivo “carbono hexa-coordinado plano” (pHC) en el trabajo “*Planar Hexacoordinate Carbons: Half Covalent, Half Ionic*”¹ (Artículo 3, Anexo 2).

Para esta tarea, se exploró la SEP de los sistemas $CE_3M_3^+$ (E = S-Te, M=Li-Cs), un total de 15 sistemas. El objetivo era encontrar sistemas con simetría D_{3h} que cumplieren con varios análisis de enlace y carga (Wiberg¹³², IQA¹³³, AdNDP^{134, 135}, NPA¹³⁶) que concluyeran la existencia del carbono pHC.

Luego de la identificación del sistema pHC, con Leyva-Parra se publica “*Planar Hypercoordinate Carbons in Alkali Metal Decorated CE_3^{2-} and CE_2^{2-} Dianions*”¹³⁷ (Artículo 4, Anexo 2) donde se utiliza MatrixAutomaton para explorar la SEP de los sistemas $M_mCE_2^p$ (E=S-Te, M=Li-Cs, m=2, 3 and p=m-2) y $M_nCE_3^q$ (E=S-Te, M=Li-Cs, n=1, 2, q=n-2), en total 60 sistemas estudiados en 22100 optimizaciones.

Hay que destacar que el objetivo de este artículo era identificar GMs con átomos de carbono hipercoordinados planos (phCs), los cuales son considerados exóticos ya que se alejan de la coordinación tetraédrica asociada con el átomo de carbono. En este sentido, el software MatrixAutomaton demostró una excelente capacidad de convergencia en los 14 sistemas pentacoordinados planos (ppCs) y 24 tetracoordinados planos (ptCs), lo que indica que su metodología permite una exploración exhaustiva de la superficie de energía potencial (SEP) a pesar de tratarse de conformaciones exóticas, sin discriminar hacia un tipo específico de geometría.

Finalmente, a inicios del 2022 un último trabajo con énfasis en la identificación de phCs "*Persistent Planar Tetracoordinate Carbon in Global mínima Structures of Silicon-Carbon Clusters*"¹³⁸ (Artículo 5, Anexo 2) utilizó MatrixAutomaton para los sistemas tipo S_nC_n ($n=5,8,9$).

En dicho trabajo, se concluye que los sistemas de estudio son efectivamente carbonos tetracoordinados planos (ptC) gracias a análisis de enlace (Wiberg, NPA y AdNDP) y de corriente electrónica (GIMIC^{139, 140}). Además, un anillo de C5 se forma consistentemente en los 3 sistemas analizados que permite la formación del ptC y una corriente de anillo Si-ptC-Si dándole una característica de aromaticidad local σ .

4.3.3 USOS MATRIXAUTOMATON - $NaBH_3$

MatrixAutomaton también fue evaluado en clústeres pequeños, en el trabajo de Rios "*Neither too Classic nor too Exotic: One-Electron Na·B Bond in $NaBH_3^-$ Cluster*"¹⁴¹ (Artículo 6, Anexo 2) se buscaron los sistemas $LiBH_3^-$, KBH_3^- , $MgBH_3$, $CaBH_3$ y $NaBH_3^-$, donde este último es de especial interés por el tipo de enlace que exhibe.

4.3.4 USOS MATRIXAUTOMATON - SISTEMAS SI-LI

Un uso interesante que se le dio a las búsquedas con autómatas celulares fue aquellas que permiten identificar clústeres atómicos que pueden servir de bloques de ensamble para nanomateriales.

Se complementó el trabajo “*Li₈Si₈, Li₁₀Si₉, and Li₁₂Si₁₀: Assemblies of Lithium-Silicon Aromatic Units*”¹⁴² (Artículo 7, Anexo 2) con la identificación de posibles bloques de ensamble: T_d -Li₄Si₄ y C_{2v} -Li₆Si₅ que tienen carácter aromático bajo análisis de respuesta magnética (NICs) y de enlace (Wiberg). Se buscaron los sistemas Li₄Si₄, Li₆Si₅, Li₈Si₈, Li₁₀Si₉ y Li₁₂Si₁₀.

MatrixAutomaton está diseñado para funcionar como constructor de estas nanoestructuras ya que utiliza bloques de ensamble dentro del proceso autómatas, como si fuesen fragmentos moleculares.

Se continuó el trabajo de bloques de construcción basado en Silicio en el artículo “*Si₅-pentagonal rings and Y-shaped Si₄ building blocks in Li₃₂Si₁₈ system: similarities with the crystalline Zintl phase Li₁₂Si₇*”¹⁴³ (Artículo 8, Anexo 2). Los bloques de ensamble corresponden a Si₄ en forma de Y junto a un pentágono de Si₅. Estos se mantienen estables en un alto número de GMs para el conjunto de sistemas tipo Li_nSi₄ (n=4, 6, 8, 10, 12) y Li_mSi₅ (6, 8, 10, 12, 14).

Los bloques de ensamble escogidos fueron Li₁₀Si₄ y Li₆Si₅ por mantener las formas deseadas. Luego utilizando ambos fragmentos se construye la nanoestructura Li₃₂Si₁₈ manteniendo los motivos de ensamble originales. Este proceso se realizó mediante la utilización de un algoritmo tipo KICK.

4.3.5 USOS MATRIXAUTOMATON - SISTEMAS B

El boro resulta un tipo de clúster atómico interesante en la química orgánica. A pesar de no ser tan estudiado como su vecino el carbono si ha establecido un nicho de investigación por sus usos en nanomateriales, semiconductores y como componente biológico con propiedades antivirales, antitumorales y antisépticas¹⁴⁴⁻¹⁵⁰.

El trabajo de Solar *et al* “*Bowl-shaped Cluster CuB₁₂ : A Viable Global Minimum with Twofold Aromaticity*”¹⁵¹ (Artículo 9, Anexo 2) estudia la estructura de B₁₂ ante la adición de un átomo de transición Cu⁻ formando el clúster CuBe₁₂⁻ que toma una forma cuasi-planar con un anillo exterior de Be₁₀ y un anillo interior CuBe₂.

Para explorar la SEP, MatrixAutomaton logra identificar efectivamente el mínimo global para un sistema con un metal de transición, esta prueba no se había realizado anteriormente y sus buenos resultados nos da confianza en la posible viabilidad del programa para con este tipo de sistemas, aunque no representa una prueba definitiva para confirmar la viabilidad de la búsqueda de la SEP con sistemas metálicos.

Solar luego amplió esta metodología en el artículo "*Planar Elongated B₁₂ Structure in M₃B₁₂ Clusters (M = Cu-Au)*"¹⁵² (Artículo 10, Anexo 2) con tres sistemas con metales de transición M = Cu-Au. Se concluye que la adición de M cambia el esqueleto B₁₂ dominante en los clústeres de MBe₁₂ y M₂Be₁₂ por una versión más alargada que naturalmente se encuentra a más de 20 kcal mol⁻¹ de diferencia. Los clústeres de Boro pueden tener efectos inesperados al doparse por este tipo de metales, el cual entrega efectos inesperados en su estructura. También da confianza en la exploración de la SEP realizadas por MatrixAutomaton que fue capaz de encontrar este tipo de estructuras exóticas que no eran esperadas.

4.3.6 USO MATRIXAUTOMATON Y TEMPLATEPERMUTA - SISTEMAS SI-C

A la fecha la herramienta desarrollada ha demostrado un desempeño óptimo en la identificación de clústeres de tamaño pequeño. Desde su implementación han logrado identificar clústeres exóticos como los carbonos hipercoordinados planos; realizar búsquedas con metales de transición (Cu-Au) que poseen una química no clásica y por supuesto en la búsqueda de clústeres conocidos en su análisis de tipo de enlace. Pero ¿Cómo se defiende el programa frente a clústeres de mayor tamaño?

Lamentablemente la herramienta no logra buenos resultados a medida que las estructuras problemas crecen, en el trabajo "*Why an integrated approach between search algorithms and chemical intuition is necessary?*"¹⁵³ (Artículo 11, Anexo 2) ilustra esta problemática con el sistema C₁₂O₂Mg₁₂.

MatrixAutomaton logra converger en un mínimo energético junto con varios mínimos locales en un proceso de más de 1400 optimizaciones, sin embargo, los autores hacen notar que mediante intuición química es posible formar nuevos isómeros. Efectivamente este proceso entregó 15 isómeros en un rango de energía menor a 40 kcal mol⁻¹.

Aunque los resultados de la herramienta desarrollada pueden ser mejorados, no se puede atribuir la responsabilidad completa a la máquina. Es importante tener en cuenta que los isómeros formados manualmente son simplemente mutaciones puntuales de los resultados obtenidos de la búsqueda, lo que demuestra la importancia de combinar la exploración automática con la intuición química. La búsqueda automática puede proporcionar una exploración integral de la SEP, pero el conocimiento humano y la creatividad también son fundamentales para obtener resultados óptimos en la búsqueda de estructuras moleculares complejas.

Gracias a este comentario sobre el comportamiento de nuestras herramientas, y previos trabajos realizados, nos embarcamos en la identificación de estructuras con presencia de dos ptCs. Para esta tarea se realizó primero un proceso de permutación basado en la intuición química junto con una exploración de la SEP con la herramienta.

Un requisito crucial en el diseño de clústeres de anillos aromáticos con ptCs es, valga la redundancia, su carácter aromático. Un rasgo recurrente entre los mínimos globales (GMs) previamente reportados de estos sistemas incluye π -aromaticidad global, semilocal y local, presentándose los dos últimos en anillos fusionados, complementados por σ -aromaticidad local proporcionada por la deslocalización de dos electrones en el enlace deslocalizado E-ptC-E, donde E representa Si en este estudio. Esta característica fue utilizada por vez primera en el trabajo de Yañez¹⁵⁴ donde se estabilizaron hidrocarburos aromáticos planos como se muestra en la Figura 4.2

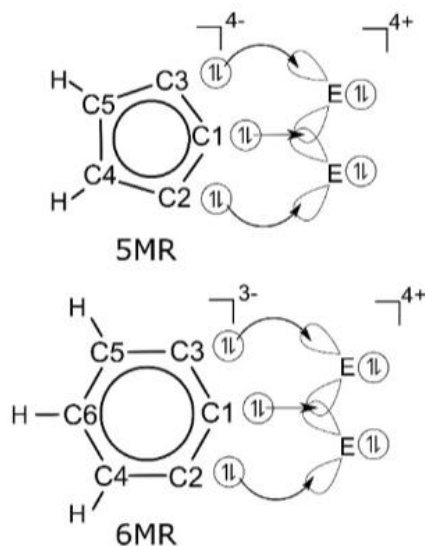


Figura 4.2 Propuesta de diseño para compuestos orgánicos ptC. Imagen tomada de la referencia ¹⁵⁴.

En nuestro estudio, se aplicaron los mismos principios de diseño para identificar una estructura basada en hidrocarburos policíclicos aromáticos que presentara dos ptCs. Utilizando el software TemplatePermuta es posible realizar un intercambio atómico al sistema $C_{18}H_{10}^{2-}$ para formar Si_6C_{18} . Luego, utilizamos la pequeña población generada como guía en la búsqueda de la SEP mediante MatrixAutomaton. Esta estrategia permitió superar la dificultad que MatrixAutomaton tiene con clústeres de mayor tamaño, aprovechando la información previamente obtenida para acelerar la exploración de la SEP.

En el artículo "*Si₆C₁₈: A bispentalene derivative with two planar tetracoordinate carbons*"¹⁵⁵ (Artículo 12, Anexo 2) se permutaron los 10 protones del dianión de un derivado de bispentaleno por 6 dicaciones de Si_2^+ . Nuestra estrategia logra capturar una estructura con dos ptC, aunque corresponde a un mínimo local a 70 kcal mol^{-1} . Mientras tanto el mínimo global corresponde a otra estructura que también se obtuvo por un proceso de permutación de un derivado del trifenileno.

En resumen, el protocolo consistió en utilizar una búsqueda guiada con MatrixAutomaton que además contaba con permutaciones guiadas de estructuras policíclicas carbonadas.

Es importante señalar que las herramientas de optimización utilizadas permiten una buena exploración de la SEP, no obstante, fue el previo conocimiento del investigador y la experiencia para con estos sistemas que se utilizó para formar una población sugerida cercana a los resultados esperados. En este caso, se utilizó el carácter aromático de estos sistemas como reglas de construcción.

Misma estrategia fue luego utilizada en el trabajo "*E₆C₁₅ (E = Si-Pb): Polycyclic Aromatic Compounds with Three planar Tetracoordinate Carbons*"¹⁵⁶ (Artículo 13, Anexo 2) donde se evaluaron un total de 4 estequiometrías cuyos mínimos globales exhiben la presencia de tres carbonos tetracoordinados planos y cuasi-planares.

En este último trabajo, también se utilizó la aromaticidad de estos sistemas en las reglas de diseño, como se muestra en la Figura 4.3

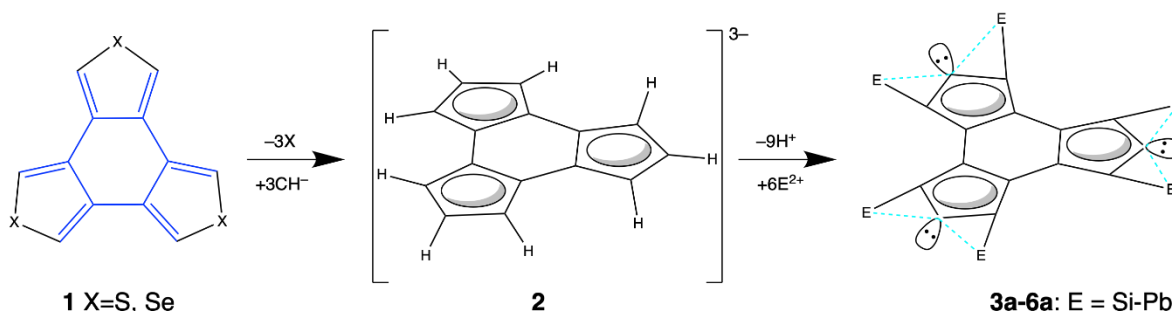


Figura 4.3 Reglas de diseño para sistemas ptCs E_6C_{15} a partir de [6]radialene benzo[1,2-c:3,4-c':5,6-c'']trityfeno (1), pasando por $C_{15}H_9^{3-}$. Imagen tomada de la referencia ¹⁵⁶.

Para cerrar la serie, se investigó un posible sistema con 4 ptCs. En este estudio se examinaron las complejidades de diseñar y explorar las superficies de energía potencial de sistemas que contienen más de veinte átomos con carbonos tetracoordinados planos (ptCs). Para abordar esta cuestión, utilizamos una regla de diseño establecida para diseñar computacionalmente un derivado de nafto [1,2-b:3,4-b':5,6-b'':7,8-b''']tetratiofeno (Figura 4.4). Este proceso comenzó con la sustitución de átomos de S por unidades de CH^- y, a continuación, con la sustitución de tres protones secuenciales por dos unidades de Si_2^+ en el polianión de hidrocarburo aromático policíclico resultante. A pesar de no representar el mínimo global, el sistema Si_8C_{22} con cuatro ptC aportó valiosas ideas para el diseño

estratégico y la exploración de la superficie energética potencial. Nuestros resultados subrayan la importancia de emplear metodologías adecuadas para confirmar la estabilidad de las estructuras moleculares de nuevo diseño que contienen carbonos hipercoordinados planares.

El artículo “*Exploring the Potential Energy Surface of Medium-Sized Aromatic Polycyclic Systems with Embedded Planar Tetracoordinate Carbons: A Guided Approach*”¹⁵⁷ (Artículo 14, Anexo 2) recopila la investigación en detalle, donde sólo se utilizó el software TemplatePermuta para explorar la SEP. MatrixAutomaton no fue utilizado en esta oportunidad.

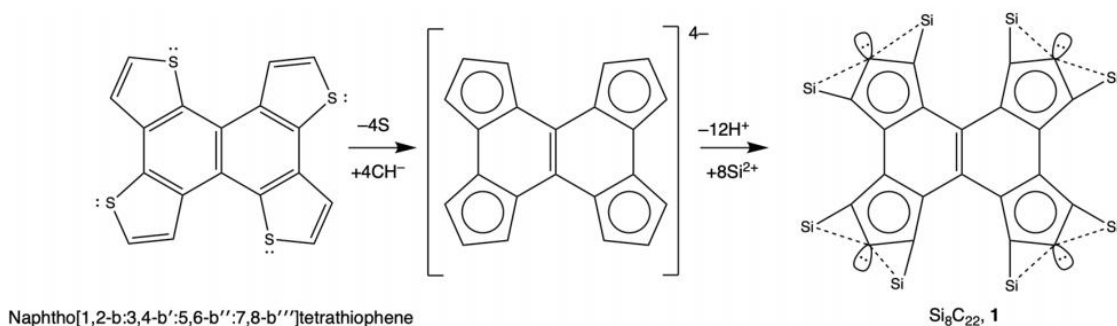


Figura 4.4 Reglas de diseño para el ptC Si₈C₂₂ a partir del sistema nafto [1,2-b:3,4-b':5,6-b'':7,8-b''']tetratiofeno. Imagen tomada de la referencia ¹⁵⁷.

En todos los trabajos realizados se discute a fondo la aromaticidad local, semilocal y global de los sistemas evaluados por medio de análisis de corriente electrónica y de enlace. Sin embargo, estos resultados, si bien interesantes, escapan al objetivo final de esta tesis.

Finalmente, todos los artículos hasta ahora discutidos se pueden leer a totalidad en el Anexo 3.

4.4 PERSPECTIVAS

Los métodos de exploración del SEP están presentes en muchos estudios de clústeres atómicos. Su implementación y desarrollo deben ir ligadas a las propiedades intrínsecas de los sistemas a evaluar, ya sea individualizando un isómero en particular creando un *bias* hacia un método de optimización, como generalizando mediante patrones de búsquedas, que pueden o no estar guiados por conceptos y descriptores químicos como son las interacciones químicas, la aromaticidad, o las similitudes estructurales.

A lo largo de este capítulo se habló de varios trabajos realizados con uno o más métodos desarrollados. MatrixAutomaton fue diseñado para hacer uso de todas las herramientas programadas en AUTOMATON con nuevas funcionalidades como son la implementación de fragmentos de búsqueda, re-escalamiento de variables y eficacia computacional, las cuales se pueden leer en detalle en el Capítulo VIII.

El programa fue puesto a prueba en sistemas exóticos como son los carbonos hipercoordinados planos logrando explorar con eficiencia su SEP e identificando varios mínimos globales, incluido el primero carbono hexacoordinado plano.

Fue utilizado en sistemas clásicos como el NaBH_3 en un contexto de discusión sobre la naturaleza del enlace que domina esta interacción, el cual se aleja del enfoque de esta investigación.

En sistemas de Boro con metales (Cu, Ag, Au) correspondientes a interacciones metálicas en clústeres de gran tamaño, en sistemas de Silicio-Litio identificando patrones comunes que fueron luego utilizados como bloques de ensambles en clústeres de gran tamaño usando KICK de fragmentos, iniciando las estrategias de unión entre metodologías.

Estrategia que fue repetida en la búsqueda de carbonos tetracoordinados planos, los cuales fueron primero ideados bajo las reglas de diseño basadas en aromaticidad (Figura 4.2) y luego explorados por MatrixAutomaton. Rápidamente

se cayó en cuenta de las ventajas de un proceso de permutación isoeléctrica de los hidrocarburos policíclicos permitiría la identificación de los sistemas de interés, cada uno de mayor tamaño (4 ptCs).

Sin embargo, a pesar de los buenos resultados obtenidos gracias a las herramientas diseñadas, en ningún momento se debe afirmar que son perfectas e infalibles en la exploración de la SEP. El objetivo principal de este capítulo es demostrar que las técnicas de optimización diseñadas son efectivas para la exploración dirigida de la SEP, que el conocimiento previo del comportamiento atómico permite adaptar las técnicas de búsqueda resultando en la obtención de buenos mínimos energéticos evitando un alto número de cálculos energéticos, y que la mezcla de metodologías permiten disminuir las desventajas de los métodos individuales y, por lo tanto abarcar una mayor extensión de la SEP.

CAPÍTULO V – KICK-FUKUI: UN MÉTODO DE PREDICCIÓN DE ESTRUCTURAS MOLECULARES GUIADO POR FUNCIONES DE FUKUI

"Este capítulo introduce el método híbrido Kick-Fukui, diseñado para explorar la superficie de energía potencial (SEP) de clústeres y moléculas. El método emplea la integral Coulómbica entre las funciones de Fukui para el primer screening de individuos destacados. Se utilizan pequeños clústeres o moléculas estables como unidades básicas de ensamblaje para formar especies más complejas. Inicialmente, se selecciona un subconjunto de candidatos de una población más grande, generada estocásticamente (Kick), y se ordenan según el valor máximo de la integral Coulómbica entre las funciones de Fukui de los fragmentos. Estos candidatos seleccionados se optimizan posteriormente mediante métodos de gradiente y cálculos de la teoría del funcional de la densidad (DFT). El rendimiento del método ha sido evaluado en diversos sistemas, incluyendo clústeres atómicos y moleculares. En la mayoría de los casos, se ha logrado identificar el mínimo global (GM) con un coste computacional reducido. Aunque la estrategia no ha sido efectiva para identificar el GM en ciertos clústeres de silicio, ha logrado predecir mínimos locales cercanos en energía al GM, que podrían servir como punto de partida para algoritmos evolutivos."-**Abstract de "Kick-Fukui: A Fukui Function-Guided Method for Molecular Structure Prediction" (Anexo 3, Artículo 1)**

5.1 KICK-FUKUI

En este capítulo, se presenta la metodología integrada Kick-Fukui, diseñada para abordar el desafío de la optimización global en clústeres atómicos. Esta metodología combina la función de Fukui como descriptor de reactividad química y el algoritmo KICK basado en ensamblajes aleatorios de fragmentos moleculares. El objetivo principal es lograr una exploración más eficiente y efectiva de la superficie de energía potencial en clústeres atómicos, aprovechando las ventajas de estos dos enfoques complementarios.

Los descriptores de reactividad química, desarrollados dentro del marco teórico del funcional de la densidad conceptual^{72, 74}, se definen como funciones que responden a cambios en la densidad electrónica y energía electrónica. Estos descriptores ofrecen una comprensión detallada de una amplia gama de fenómenos químicos, que abarcan desde reacciones pericíclicas hasta catálisis y toxicología^{158, 159}. Su capacidad para capturar y explicar estos fenómenos los convierte en herramientas poderosas en el estudio de sistemas moleculares y en la predicción de propiedades químicas relevantes.

En el trabajo de Osorio⁸⁵ la función de Fukui fue utilizada como guía en un esquema de teoría perturbacional para determinar la energía de interacción (ΔW_{int}) entre dos fragmentos moleculares. En particular, se enfocaron en la interacción entre un fragmento nucleofílico y otro electrofílico utilizando la ecuación de Berkowitz.

$$\begin{aligned} \Delta W_{int} = & (\mu_{acceptor} - \mu_{donor})\Delta N \\ & + \int \left(\sum_{\alpha \in donor} Z_{\alpha} \delta(r - R_{\alpha}) - \rho_{donor}(r) \right) \Phi_{acceptor}(r) dr \\ & + \Delta N \int (f_{acceptor}^{+}(r) \Phi_{donor}(r) - f_{donor}^{-}(r) \Phi_{acceptor}(r)) dr \\ & - (\Delta N)^2 \iint \frac{f_{donor}^{-}(r) f_{acceptor}^{+}(r')}{|r - r'|} dr dr' \end{aligned} \quad (5.1)$$

Los descriptores de reactividad químicos utilizados corresponden al potencial químico y la función de Fukui. El primero se utiliza para determinar la naturaleza nucleofílica o electrofílica de cada fragmento molecular. En este contexto, el fragmento con un potencial químico más bajo será aquel aceptor de electrones.

La función de Fukui, al estar determinada bajo DFT está definida en todo el espacio molecular. Esto permite realizar un análisis topológico, y como tal, la obtención de puntos críticos, denominados atractores. Dichos atractores representan la función de Fukui condensada, donde su valor refleja la respuesta de

la densidad electrónica en la región espacial definida por el basin correspondiente. Mediante este análisis topológico, es posible obtener información cuantitativa sobre la reactividad de las diferentes regiones de la molécula y su relación con las propiedades químicas observadas.

La ecuación de Berkowitz permite entonces definir la energía de interacción entre dos fragmentos nucleofílicos y electrofílicos. Sin embargo, es importante destacar que, en esta ecuación, solo el componente Coulombiano experimenta cambios significativos según la disposición de los fragmentos. Esta integral, denominada J a partir de ahora, será máxima cuando exista una fuerte interacción o superposición entre los atractores nucleofílicos y electrofílicos correspondientes. En otras palabras, cuanto más cercanos estén los atractores nucleofílicos y electrofílicos más reactivos, mayor será la interacción y, por ende, mayor la energía de interacción.

El trabajo realiza estas superposiciones de basines Fukui entre fragmentos de clústeres conocidos de manera manual, con el fin de maximizar J y formar clústeres de mayor tamaño. A continuación, los clústeres resultantes se someten a un proceso de optimización utilizando un método por gradiente, con el propósito de obtener mínimos energéticos competitivos o, idealmente, el mínimo global de la superficie de energía potencial.

El concepto de usar la Función de Fukui como guía predictiva fue desarrollado y aplicado en el trabajo de Yañez *et al*⁸⁶. En dicho estudio, donde el tesista estuvo involucrado, se agrega un algoritmo genético que utiliza como población los ensamblados obtenidos que maximicen J . El uso de un algoritmo genético permitió explorar de manera eficiente el espacio de búsqueda y seleccionar los ensamblados más prometedores para su posterior optimización.

En el contexto de esta tesis, se pretende evaluar el potencial del descriptor de la función de Fukui como guía exclusiva en la obtención de clústeres atómicos, siguiendo una metodología similar a la descrita previamente. Para ello, se emplearán fragmentos moleculares previamente conocidos y se realizará un

análisis topológico de su función de Fukui para identificar los atractores relevantes. El objetivo principal será maximizar la integral J , que representa la interacción entre los basines de Fukui nucleofílicos y electrofílicos, utilizando un enfoque estocástico en lugar de uno evolutivo. Este enfoque estocástico permitirá explorar de manera aleatoria el espacio de búsqueda y seleccionar ensamblajes moleculares prometedores para su posterior optimización. Al utilizar únicamente la función de Fukui como guía en este enfoque estocástico, se busca simplificar el proceso y evaluar la efectividad de este descriptor como herramienta predictiva en la obtención de clústeres atómicos estables y reactivos.

5.2 SISTEMAS EVALUADOS

Se tomaron sistemas cuyos mínimos energéticos se encuentran bien reportados en literatura. En primer lugar, se consideraron los clústeres de agua desde el dímero hasta el pentámero¹⁶⁰. Estos sistemas representan un ejemplo típico de clústeres moleculares.

Si_nAu^+ ($n=2-15$), un sistema con presencia de átomo metálico, que pueden exhibir propiedades interesantes. Asimismo, nos permite explorar el alcance de la estrategia propuesta para con este tipo de clústeres atómicos.

También se consideraron los clústeres tipo Zintl, los cuales son complejos reportados experimentalmente¹⁶¹. Estos sistemas presentan una organización particular de átomos y pueden participar en oligomerizaciones¹⁶². ¿Podrá la función de Fukui predecir la mejor interacción en este tipo de complejos?

Finalmente, el último tipo de sistema evaluado fueron los clústeres atómicos de $\text{Si}^{37, 85, 163}$, en específico Si_n ($n=12-20$) debido al previo trabajo con estos sistemas en el trabajo de Yañez⁸⁶. Además, estos clústeres tienen interacciones covalentes, donde los sistemas utilizados como fragmentos pueden cambiar drásticamente por el proceso de optimización por gradiente.

5.3 DETALLES COMPUTACIONALES

Todas las estructuras de los sistemas fueron obtenidas de la literatura correspondiente. Se refinaron utilizando el software químico-computacional Gaussian 16 versión B.01 bajo el nivel de teoría PBE0⁶⁸/Def2-TZVP¹²⁶.

Los descriptores de reactividad fueron calculados usando Multiwfn 3.3.6¹⁶⁴ que es una software especializado en el análisis de la función de onda. Con anterioridad, el tesista había desarrollado un software tipo *pipeline* "TAFF"⁸² el cual trivializa el uso de Multiwfn para el cálculo de diferentes descriptores de reactividad global y local. El análisis topológico fue realizado bajo la aproximación de Koopmans. TAFF entrega archivos resumen en formato de coordenadas con información de los atractores, en específico su ubicación y valor.

El programa "KICK-Fukui" fue desarrollado en el lenguaje de programación Perl, debido a su flexibilidad y capacidad para el procesamiento eficiente de datos. Cada búsqueda (ensamble y jerarquización de J) se realizó a nivel PBE0/SDDAll. Posteriormente, las estructuras fueron refinadas a PBE0/Def2-TZVP junto con análisis de frecuencias para verificar la calidad de mínimo energético.

5.4 RESULTADOS

Multiwfn¹⁶⁴ permite realizar el análisis topológico de la función de Fukui a un sistema proporcionado. Sin embargo, este procedimiento resulta tedioso y largo en ciertas aproximaciones. Por ejemplo, si se desea obtener la función de Fukui topológica bajo una aproximación de diferencias finitas se deben ingresar manualmente el número de cada orbital y su valor correspondiente.

TAFF (*Topological Analysis of Fukui Fuction*) fue diseñado (Esquema 3, Anexo 1) con el propósito de automatizar el proceso de obtener el análisis topológico de la función de Fukui para un sistema. Además, haciendo uso de DFT, también calcula varios descriptores de reactividad química, tales como el potencial químico, la dureza y la electronegatividad.

Las Figuras 5.1, 5.2, 5.3 y 5.4 muestran los mínimos energéticos obtenidos por el programa KICK-Fukui. Los clústeres moleculares $(\text{H}_2\text{O})_n$ $n=2-5$ fueron diseñados tomando como fragmentos individuales de una molécula de agua H_2O y un clúster molecular de agua $(\text{H}_2\text{O})_{n-1}$ para obtener la estequiometría deseada. En todos los casos se obtuvieron los mínimos globales reportados en literatura, donde cada molécula de agua no pierde su estructura individual. Esto indica que la función de Fukui desempeñó un papel crucial en la ubicación precisa de estos clústeres moleculares antes de ser sometidos a las optimizaciones por gradiente de energía a pesar de que este tipo de moléculas se encuentran dominadas por interacciones no covalentes.

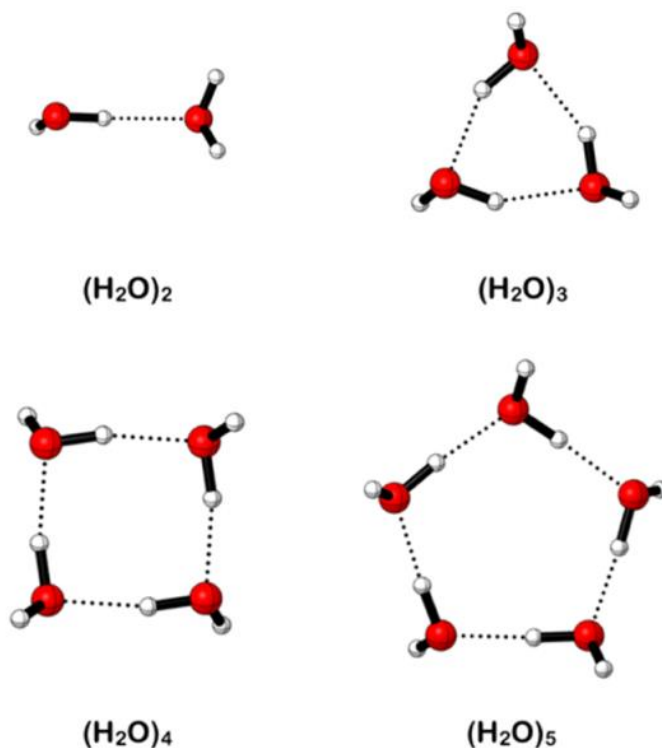


Figura 5.1. Mínimos energéticos encontrados para los clústeres $(\text{H}_2\text{O})_n$ $n=2-5$.

Los sistemas Si_nAu^+ fueron ensamblados utilizando como fragmentos un catión Au^+ y el clúster Si_n correspondiente, teniendo en cuenta el comportamiento sugerido por la función de Fukui para determinar la ubicación óptima del átomo

metálico en un clúster de mayor tamaño. Los resultados obtenidos mostraron que todos los mínimos energéticos encontrados corresponden a GMs reportados en la literatura. (Figura 5.2.) Esto indica que la función de Fukui fue efectiva en guiar la formación de clústeres estables y en identificar las configuraciones de mayor energía mínima. Sin embargo, cabe destacar que existen otras interacciones a considerar en los sistemas metálicos que pueden ofuscar el poder predictivo de la función de Fukui. Por dicha razón solo se realizó una prueba con un átomo único metálico.

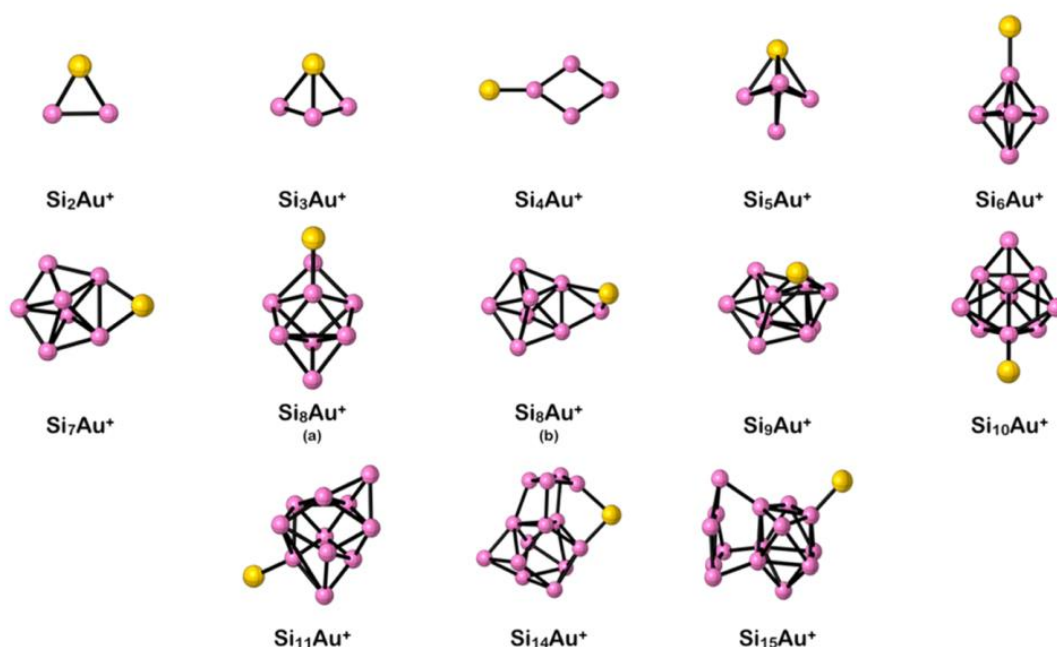


Figura 5.2. Mínimos energéticos encontrados para clústeres Si_nAu^+ . Imagen tomada de la referencia 165.

La Figura 5.3 ilustra la mejor interacción obtenida en el complejo Zintl $[\text{Ge}_9\{\text{Si}(\text{SiMe}_3)_3\}_3\{\text{SnPh}_3\}]$. En este estudio, se exploró la SEP del clúster neutral tetrasustituido mediante el ensamblaje de los fragmentos Ph_3Sn^+ (electrófilo) y $[\text{Ge}_9\{\text{Si}(\text{SiMe}_3)_3\}_3]^-$ (nucleófilo).

Se logró identificar correctamente el mínimo energético global (GM) del complejo, a pesar de su tamaño y complejidad. La ubicación sugerida por la función de Fukui se muestra en tonos más transparentes en la figura, y se observa que está muy

cercana a la ubicación del sistema una vez optimizado. Esto demuestra que la función de Fukui fue efectiva en proporcionar una guía precisa para la formación del complejo y en identificar su configuración más estable.

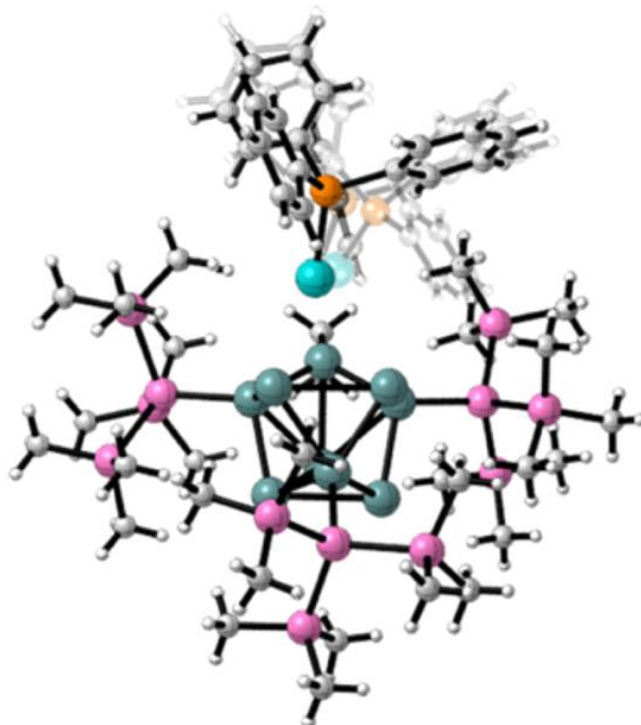


Figura 5.3. Candidato a mínimo energético obtenido por Kick-Fukui. En transparente se muestra la mejor disposición guiada por la función de Fukui previo a su optimización. Imagen tomada de la referencia ¹⁶⁵.

Los sistemas evaluados hasta este punto tienen como característica que sus fragmentos no pierden su identidad estructural luego de ser sometidos a los procesos de optimización. Para verificar el rol predictivo de la función de Fukui se evaluaron clústeres de Silicio, Si_n ($n=12-20$) donde los fragmentos utilizados corresponden a mínimos globales de clústeres de Silicio de menor tamaño. Por ejemplo, el sistema Si_{12} fue ensamblado gracias a la combinación Si_2+Si_{10} , Si_3+Si_9 , Si_4+Si_8 , Si_5+Si_7 y Si_6+Si_6 .

Se observó que los sistemas Si_{15} , Si_{18} y Si_{19} no presentan el mínimo energético reportado en literatura, por otro lado, el Si_{17} corresponde a un mínimo energético más estable que el reportado en la misma. Todos los demás sistemas encuentran el GM reportado (Figura 5.4).

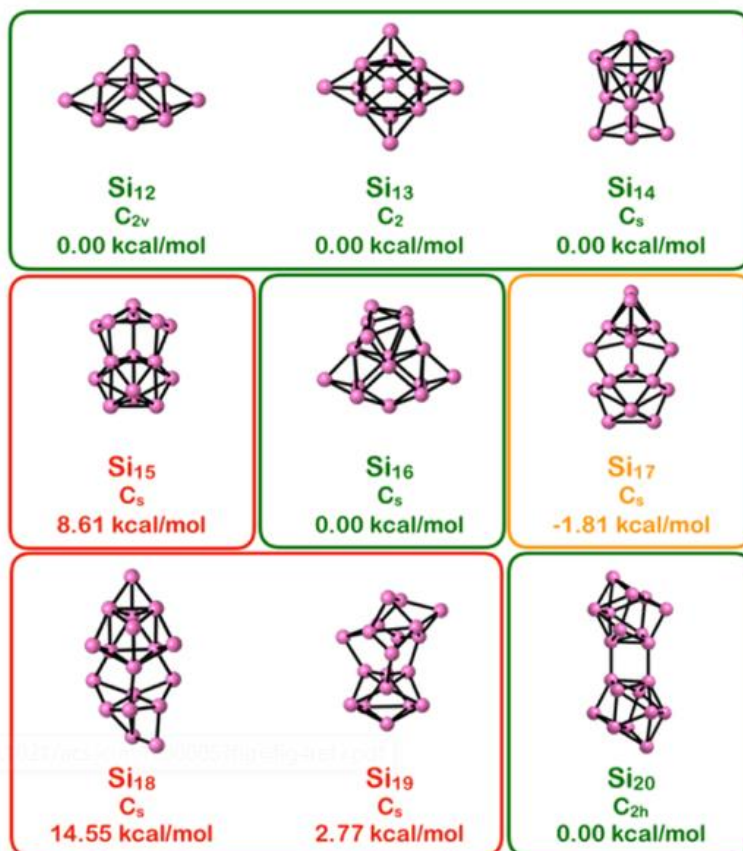


Figura 5.4. Estructuras de mínima energía obtenidas por la metodología. Verde, mínimos globales reportados en literatura; rojo, estructuras menos estables que las reportadas y; amarillo, estructuras más estables que las reportadas en literatura. Imagen tomada de la referencia ¹⁶⁵.

Este tipo de sistemas si pierden la identidad de los fragmentos individuales luego de los procesos de optimización, y es el primer caso donde no se rescatan todos los mínimos globales luego del proceso de KICK-Fukui. Sin embargo, usamos solamente los mínimos globales reportados para clústeres pequeños de silicio como un método de estandarización de la metodología, nada nos prohíbe utilizar mínimos energéticos competitivos como fragmentos, los cuales podrían arrojar mejores resultados. Esta prueba no fue abordada, pero da pie a un posible estudio de los efectos de diferentes tipos de fragmentos en la calidad de los ensambles obtenidos, así como también de los resultados.

La integración de la función de Fukui en un algoritmo estocástico pareciese dar buenos resultados a la hora de predecir buenos mínimos candidatos, sin embargo,

el proceso estocástico es un método de optimización por sí solo. Es conocido estadísticamente que un evento aleatorio se repite suficientes veces eventualmente ocurrirán todas las posibles situaciones, ¿Es posible que los resultados obtenidos hasta ahora sean producto del método estocástico antes que la guía de la función de Fukui?

Para dilucidar esta disyuntiva se hizo competir una búsqueda aleatoria pura (método KICK) y el método KICK-Fukui para el clúster atómico Si₁₂. La Figura 5.5 indica un *boxplot* comparativo de 40 y 20 isómeros con las energías del ensamble “Single-Point” y luego de ser optimizadas “Optimized”. En acuerdo con nuestra hipótesis, los resultados obtenidos aleatoriamente eran inferiores en comparación con aquellos obtenidos con la guía de la función de Fukui, incluso llegando a encontrar el GM.

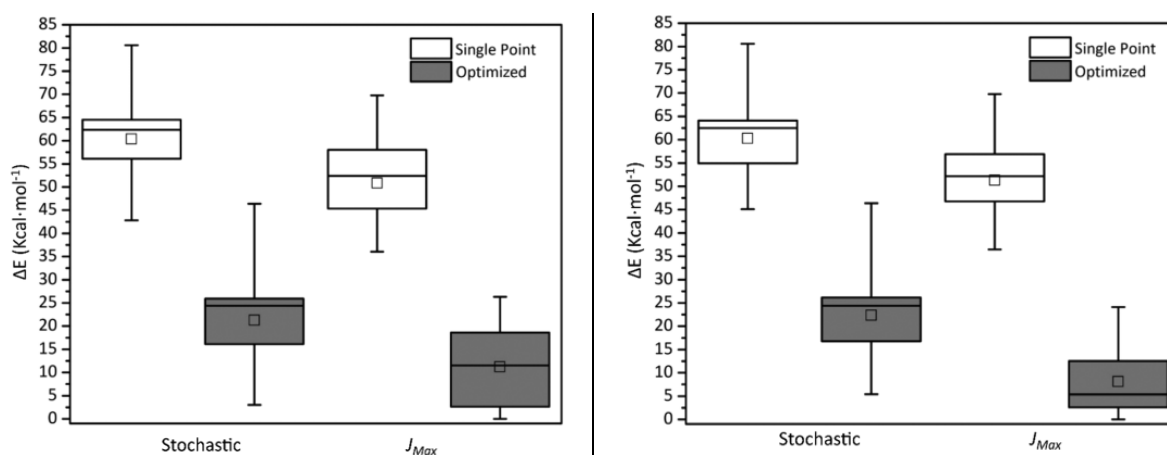


Figura 5.5. *Boxplot* con la distribución energética de Si₁₂. Izquierda, comparación de los 40 mejores isómeros; derecha, comparación de los mejores 20 isómeros. *Boxplot* blancos y grises corresponden a energías *single-point* y optimizadas respectivamente. Imagen tomada de la referencia ¹⁶⁵.

Cabe recalcar que Kick-Fukui produce resultados satisfactorios sin necesidad de utilizar un algoritmo de optimización adicional, lo que significa que este método es sencillo, no requiere de altos costos computacionales y es eficiente en la obtención de resultados.

El artículo *Kick-Fukui: A Fukui Function-Guided Method for Molecular Structure Prediction*¹⁶⁵ (Artículo 1, Anexo 2) abarca los resultado hasta ahora discutidos.

Pero ¿Qué sucede con aquellos clústeres moleculares donde la función de Fukui no entrega información relevante? ¿Qué sucede con aquellos sistemas dominados por interacciones electroestáticas, por ejemplo? Para abordar estas cuestiones, se evaluó un segundo descriptor: la integración del NCI que será discutida en el siguiente capítulo.

5.5 PERSPECTIVAS

Kick-Fukui resulta una herramienta rápida y eficaz para la obtención de buenos mínimos candidatos a estructuras de mínima energía. El punto más difícil de la utilización de esta herramienta consiste en escoger los fragmentos adecuados que conformaran el sistema molecular deseado. Sin embargo, no es para nada infalible y puede cometer errores o simplemente fallar al obtener estructuras de mínima energía como bien dan cuenta los resultados para el caso del Si₁₅, Si₁₈ y Si₁₉.

Una buena solución para realizar una exploración exhaustiva de la SEP es la que tomaron Yañez en su trabajo⁸⁶ al combinar los resultados guiados por Fukui (de manera manual) con un algoritmo genético que permita evolucionar las estructuras conformadas bajo un estándar de energía. Kick-Fukui comprende una mejora al automatizar el proceso completamente, permitiendo la libre creación de miles de ensamblajes rápidamente, y por lo tanto la libre disposición de una fuerte población inicial en cualquier algoritmo de búsqueda y exploración de la SEP.

Existe una preocupación válida respecto a la extrapolación de este protocolo a otras moléculas o macromoléculas, debido a las posibles interferencias de variables químicas adicionales que podrían afectar la capacidad predictiva de la función de Fukui. Factores como el potencial electrostático, diferencias significativas de carga, el tamaño real de los sistemas y otras interacciones presentes en sistemas más complejos que podrían influir en la efectividad de la función de Fukui como guía predictiva.

Si bien, los descriptores de reactividad han sido utilizados en diferentes áreas como la detección de sitios activos¹⁶⁶, estudios de solvatación¹⁶⁷⁻¹⁶⁹ o dentro de la

dinámica molecular^{170, 171}; la utilización de la función de Fukui, dentro de sistemas biológicos, se ha remitido a su carácter descriptivo antes que predictivo¹⁷².

Tomemos como ejemplo las proteínas, que son macromoléculas de gran tamaño, compuestas por cientos de miles de átomos y desempeñan funciones biológicas vitales. La conformación estructural de las proteínas es crucial para su actividad biológica y se mantiene gracias a las interacciones no covalentes, que son relativamente débiles pero fundamentales para estabilizar la estructura tridimensional. En un sistema tan complejo y a gran escala, la función de Fukui podría no ser adecuada como guía debido a su alcance local y la influencia predominante de las interacciones no covalentes.

Es importante reconocer que cada tipo de sistema molecular puede presentar desafíos únicos y requerir enfoques específicos para su estudio. Si bien la función de Fukui ha mostrado utilidad en sistemas más simples, como clústeres atómicos, su aplicabilidad en sistemas más complejos como las proteínas requiere un análisis cuidadoso y consideración de otras variables químicas relevantes.

Finalmente, existe la posibilidad de combinar la capacidad predictiva de la función de Fukui con otros algoritmos o métodos de optimización global que sean menos estocásticos y más dirigidos. Como ejercicio, se ha explorado una posible integración con un autómata celular, donde la evolución de las celdas se ve influenciada por los valores de los atractores de la función de Fukui en clústeres atómicos.

Esta integración presenta varias ventajas. En primer lugar, al no depender de un enfoque estocástico, se puede lograr un mayor determinismo en el proceso de optimización. Además, se pueden guardar "*fingerprints*" de las mejores interacciones entre fragmentos, lo que permite aprovechar la información obtenida en iteraciones anteriores y guiar de manera más efectiva la exploración de la superficie de energía potencial.

Otra ventaja de esta integración es la capacidad de aumentar gradualmente la complejidad del sistema al incluir más fragmentos en el proceso iterativo. Esto permite abordar sistemas más grandes y complejos de manera progresiva, aprovechando la información proporcionada por la función de Fukui en cada etapa.

Es importante destacar que esta propuesta de integración con un autómata celular es un ejemplo ilustrativo y que existen diferentes enfoques y algoritmos que podrían ser adecuados para combinar la función de Fukui con métodos de optimización global más dirigidos. La elección del método dependerá de las características del sistema en estudio y de los objetivos específicos de la investigación.

5.5.1 DISEÑO - AUTÓMATA CELULAR Y FUNCIÓN DE FUKUI

En la implementación propuesta, la función de Fukui se integrará dentro de las reglas del autómata celular que opera en conjunto con un algoritmo genético. Para garantizar un adecuado funcionamiento y aprovechar la capacidad predictiva de la función de Fukui, se deben tener en cuenta los siguientes detalles:

1. **Reglas de crecimiento:** Las reglas que determinan el crecimiento de las celdas en el autómata celular se basarán en los descriptores de la función de Fukui. Estos descriptores proporcionan información sobre las zonas más reactivas y propensas a la interacción en el sistema.
2. **Reglas de evolución:** A diferencia de un enfoque estocástico, en esta implementación las reglas de evolución del autómata celular serán probabilísticas y estarán determinadas por los valores de los descriptores de la función de Fukui. En lugar de tomar decisiones de manera aleatoria, se asignará una probabilidad de evolución basada en la información de los descriptores.

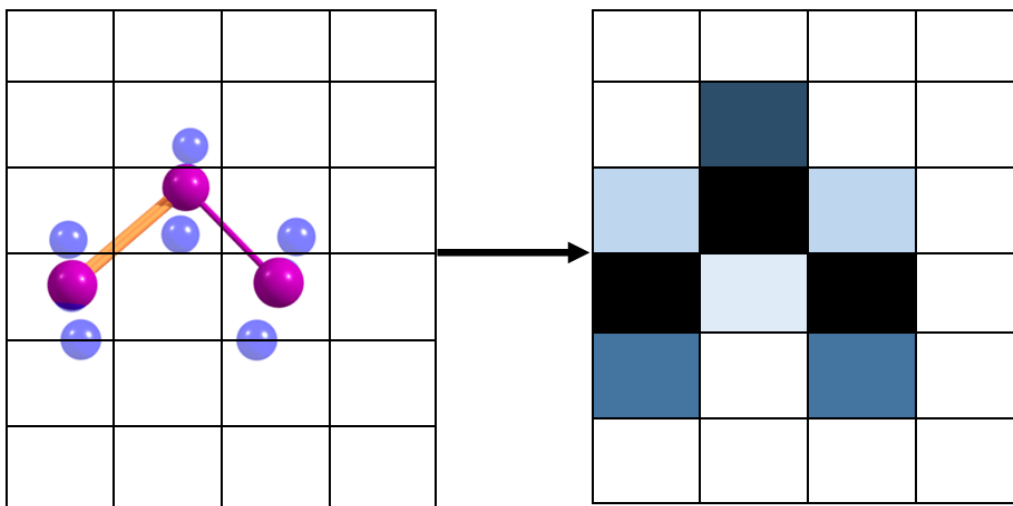


Figura 5.6 Diagrama de transformación sistema cartesiano a autómata celular. En este enfoque cada átomo del sistema cartesiano se le asignará una celda para ser poblada.

La Figura 5.6 ilustra la representación inicial de un clúster en formato legible por el autómata celular, utilizando la información de la función de Fukui. En esta representación, las celdas negras representan la presencia de átomos, mientras que las celdas azules corresponden a la ubicación de los atractores de la función de Fukui condensada. Los diferentes tonos de azul en las celdas indican la intensidad de reactividad asociada con cada atractor, reflejando la probabilidad de generar un ensamble en dicha celda. Esta representación permite visualizar de manera clara la distribución espacial de los átomos y los atractores de la función de Fukui, brindando información valiosa sobre las zonas más reactivas del sistema

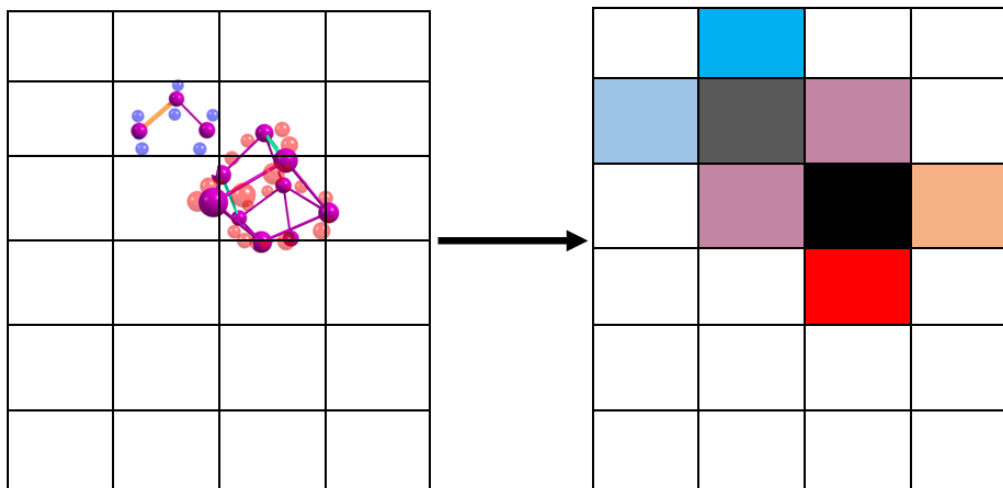


Figura 5.7 Diagrama de transformación sistema cartesiano a autómata celular. En este enfoque cada fragmento del sistema cartesiano se le asignará una celda para ser poblada.

La Figura 5.7 presenta un segundo enfoque para transformar un sistema cartesiano con información de la función de Fukui a un formato legible por el autómata celular. En esta representación, cada celda contiene un fragmento molecular, cuyo tamaño puede ser escalado según las necesidades del sistema. Las celdas colindantes albergan los valores condensados de la función de Fukui correspondientes a ese fragmento molecular.

Es importante destacar que esta estrategia de simplificar el sistema molecular albergando fragmentos en cada celda puede llevar a una pérdida de información detallada sobre la estructura molecular original. Sin embargo, este enfoque tiene la ventaja de proporcionar puntos de inicio adecuados para el proceso evolutivo del autómata celular.

La interacción entre fragmentos en el autómata celular se basará en los valores asignados a las celdas que contienen la información de la función de Fukui. La Figura 5.8 ilustra un ejemplo de cómo se lleva a cabo la interacción al construir el ensamble.

En la figura, cada celda representa una posición en el ensamble, y se asigna un valor de función de Fukui a cada celda, indicando la reactividad asociada. Los colores y tonos de las celdas reflejan los diferentes valores de la función de Fukui.

Durante la construcción del ensamble, las celdas interactúan entre sí de acuerdo con los valores de la función de Fukui. Las celdas con valores más altos indican (color más fuerte) una mayor reactividad y probabilidad de interacción con otras celdas. Por lo tanto, en la figura, las celdas con valores más oscuros o intensos están más involucradas en la formación de interacciones en el ensamble.

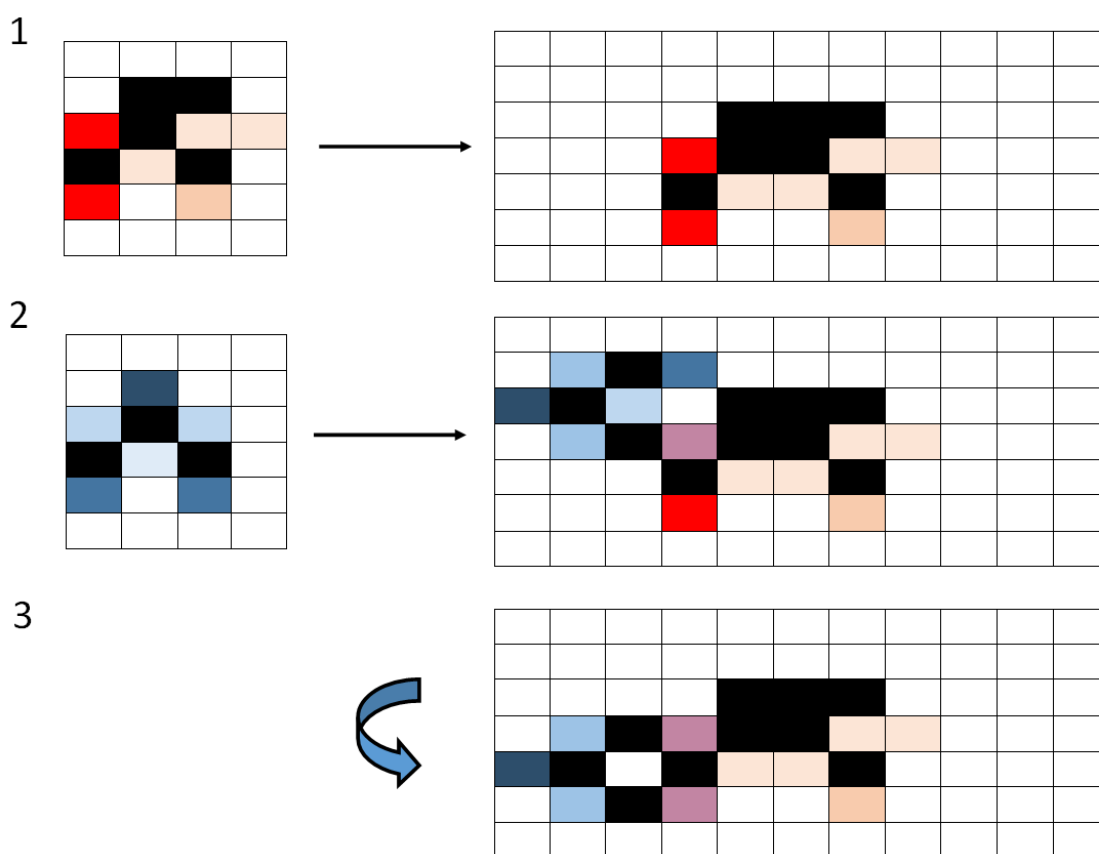


Figura 5.8. Interacción entre dos fragmentos en lenguaje autómatas celulares. Se vislumbra la evolución de las celdas según diferentes tiempos. En el tiempo 1 se agrega el fragmento nucleofílico a la grilla general de interacción. En el tiempo 2 se agrega el fragmento electrofílico a la grilla general, en morado coinciden celdas reactivas correspondientes a la función de Fukui. Tiempo 3 autómatas celulares intenta maximizar la interacción entre ambos fragmentos.

La interacción entre fragmentos se lleva a cabo considerando la reactividad de las celdas adyacentes. Las celdas con valores altos de la función de Fukui tenderán a

interactuar más con las celdas vecinas, formando vínculos y contribuyendo a la estructura global del ensamble.

En resumen, la interacción entre fragmentos en el autómata celular se determina en función de los valores de la función de Fukui asignados a las celdas. Estos valores guían la formación de interacciones entre las celdas vecinas, permitiendo la construcción del ensamble en base a la reactividad de cada fragmento.

CAPÍTULO VI – PREDICCIÓN DE ESTRUCTURAS MOLECULARES MEDIANTE LAS INTERACCIONES NO COVALENTES

"Este capítulo introduce una innovadora herramienta predictiva para la exploración de la superficie de energía potencial en clústeres moleculares. Se demuestra que la integral de las interacciones está directamente relacionada con la energía del Sistema ensamblado. Esto permite el desarrollo de un algoritmo de optimización basado en enjambre de partículas que guía eficientemente nuevas estructuras hacia mínimos energéticos. Sin embargo, se observa que la identificación de estructuras óptimas está vinculada a valores muy bajos de la integral NCI, lo que sugiere la necesidad de emplear algoritmos de optimización más sofisticados." - **Resumen de Capítulo.**

6.1 GENERALIZACIÓN

Como fue abordado con anterioridad las interacciones no covalentes han sido utilizadas descriptivamente para explicar diversas cualidades y propiedades de moléculas y macromoléculas. Son las culpables de la estabilización de proteínas, interacciones con ligandos, y hasta de las propiedades intrínsecas del agua.

Sin embargo, no han sido utilizados como descriptores predictivos en la estabilización de moléculas, en específico en clústeres moleculares. Y este fenómeno se debe principalmente a la falta de cuantificación de este tipo de interacciones. Ciertamente es que mediante técnicas de localización se puede vislumbrar la ubicación del "enlace" o zona de interacción donde algún tipo de NCI se encuentra presente, luego el investigador infiere con bases de la experiencia la tarea de la interacción en la ubicación que le compete, se puede concluir "estabilidad", se puede concluir "intercambio de cargas" u otro tipo de propiedad.

El trabajo de Contreras⁹⁸ realizó una integración a las NCI delimitando un espacio NCI, efectivamente cuantificando el valor de NCI, "NCI-i". Este avance se realizó bajo un esquema promolecular para la densidad electrónica que reduce el costo computacional asociado considerablemente al no necesitar realizar un cálculo SCF;

además de permitir discutir resultados en sistemas donde los sistemas son muy grandes y la función de onda no se encuentra disponible, es decir, entrega escalamiento a macromoléculas manteniendo una precisión relevante¹⁷³.

El nuevo descriptor, NCI-i, presenta un campo de investigación prometedor debido a su posible utilidad en la predicción energética. Se plantean interrogantes sobre cómo se correlaciona con la energía del sistema, si es posible utilizarlo en algoritmos de optimización y cuáles son sus ventajas y desventajas.

En respuesta a estos interrogantes, se ha propuesto un método de optimización global utilizando este nuevo descriptor como guía en la predicción de clústeres moleculares. Este enfoque es similar al abordado anteriormente con la función de Fukui, con la diferencia de que el descriptor NCI-i aún no ha sido sometido a pruebas exhaustivas para evaluar su posible correlación con la energía.

El objetivo de este enfoque es explorar el potencial de NCI-i como una herramienta para la optimización global, es decir, buscar estructuras moleculares que presenten mínimos energéticos utilizando el descriptor NCI-i como guía. Sin embargo, antes de utilizarlo en algoritmos de optimización, es necesario realizar investigaciones adicionales para comprender mejor la posible correlación entre NCI-i y la energía del sistema, así como evaluar las ventajas y desventajas de su implementación.

6.2 SISTEMAS EVALUADOS

Se tomaron sistemas cuyos mínimos energéticos se encuentran bien reportados en literatura y se ven afectados mayoritariamente por NCI. Se consideraron los clústeres de agua desde el dímero hasta el pentámero^{160, 174}. Debido a que estos sistemas representan un ejemplo típico de clústeres moleculares.

6.3 DETALLES COMPUTACIONALES

Todos los fragmentos moleculares fueron tomadas de la literatura y refinadas a ω b97XD¹⁷⁵/Def2-TZVP utilizando el software Gaussian16 versión B.01¹²⁵.

La integración del índice de las interacciones no covalentes se llevó a cabo con el programa NCILOT4⁹⁸, el cual empleó la aproximación promolecular y evaluó solamente interacciones intermoleculares con las variables estándar de *cutoff*.

Todas las optimizaciones energéticas fueron realizadas a PBE0/SDDAll.

6.4 RESULTADOS

El programa de Enjambre de partículas con NCI-i fue desarrollado según los requisitos establecidos en el Capítulo VIII. Donde se ajustaron los parámetros necesarios del programa NCILOT4 gracias a diversas pruebas de concepto, en específico las variables son: "*n*", el tamaño de la población y el sector Ω_{NCI} a integrar según la ecuación 2.15.

Los parámetros del proceso exploratorio de la SEP consideran una población inicial aleatoria de $1000N$ para un clúster molecular de N moléculas; se calculó el valor de NCI-i con $n = 2.5$. Luego, se tomó una muestra de la población ($10N$) distribuidos equitativamente entre los valores máximos y mínimos obtenidos de la integral. La muestra escogida se optimiza mediante métodos DFT y se ordenan según el valor energético.

Con el fin de evaluar nuestra propuesta, se han considerado varios clústeres de agua como sistemas iniciales. Los clústeres de agua son sistemas gobernados por las interacciones no covalentes, y sus estructuras de mínima energía han sido ampliamente documentadas en la literatura.

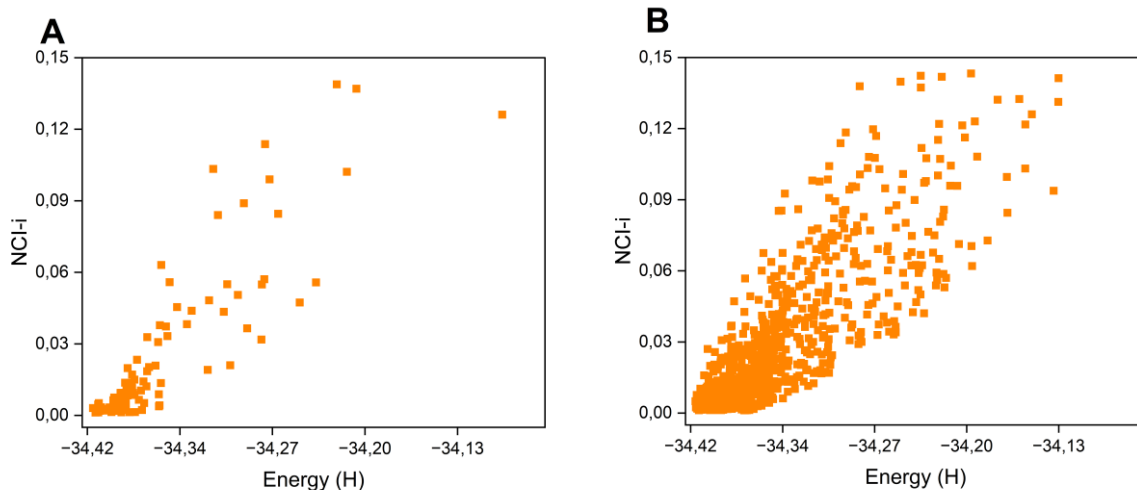


Figura 6.1. Gráfico de dispersión entre los valores de la integral de NCI (Eje Y) vs los valores energéticos (Hartree) *single-point* (Eje X) de cada estructura aleatoria para dímero de agua. **A** población 100, **B** población de 1000.

Primero, se realizó un ensayo con poblaciones de 100N y 1000N para el dímero de agua con el objetivo de estudiar si existe una correlación entre NCI-i y la energía de los clústeres moleculares. Los resultados demuestran una correlación fuertemente proporcional entre la NCI-i y la energía de los clústeres moleculares (Figura 6.1). Específicamente, se observó que a medida que los valores de NCI-i cae a valores cercanos a cero, la energía disminuye proporcionalmente. Esto sugiere que la integración del NCI es un descriptor útil para evaluar la estabilidad de los clústeres moleculares.

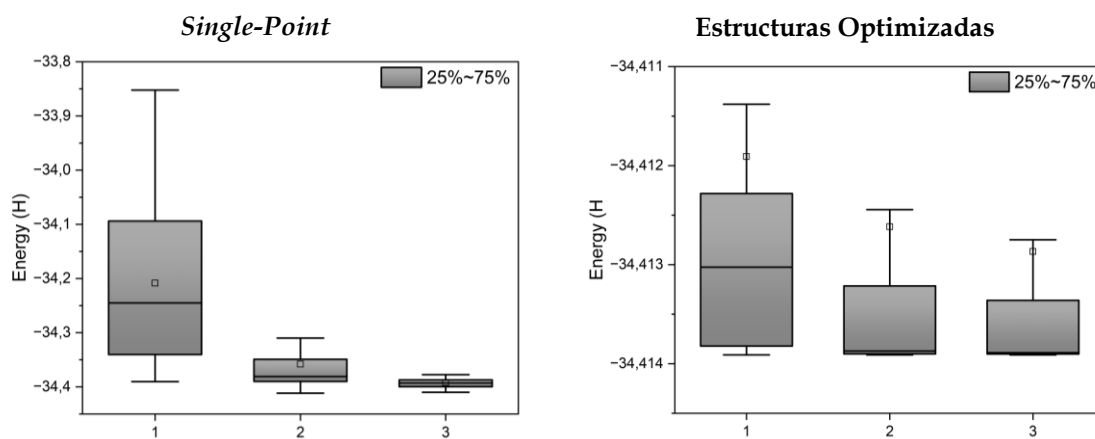
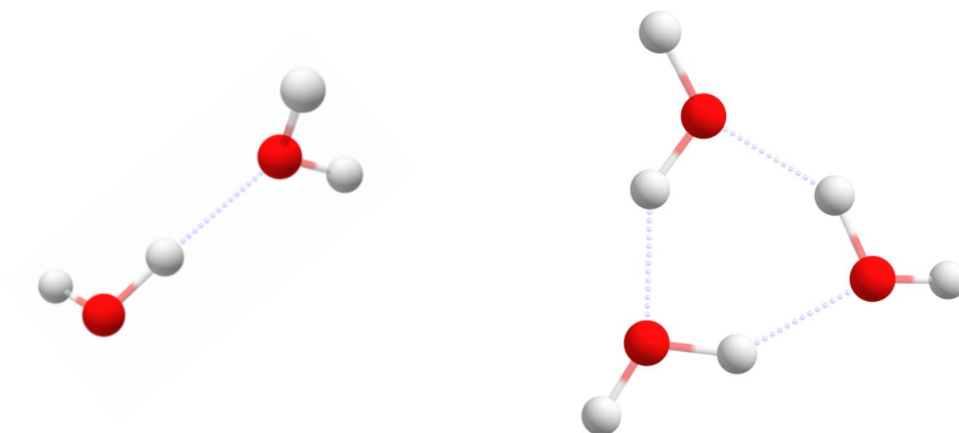


Figura 6.2 Boxplots de la evolución energética (Eje Y) en cada ciclo de optimización (Eje X) guiada por NCI-i. Izquierda, evolución de energía *single-point*; derecha evolución energía optimizada.

En el proceso de búsqueda de clústeres de agua los mínimos globales reportados en literatura fueron correctamente identificados (Figura 6.3). Se escogió un total de 3 generaciones sin cambio energético como requisito para finalizar el proceso exploratorio.

La Figura 6.2 muestra cómo varían las energías *single-point* y las energías optimizadas de la población en cada ciclo de búsqueda para el dímero de agua. Se observa una mejora constante en las energías de *single-point* en cada ciclo de búsqueda, lo cual respalda la hipótesis de que el uso de la NCI-i como descriptor en la formación de clústeres moleculares es efectivo. Sin embargo, al analizar el comportamiento de las energías optimizadas, se observa que esta tendencia no se mantiene de manera consistente. Por ejemplo, el ciclo 2 reporta una mayor cantidad de población a menor energía que aquellas del ciclo 3.

Por otro lado, la exploración de la SEP para el dímero de agua reporta el GM en los 3 ciclos de búsqueda, lo cual puede atribuirse a la simplicidad del sistema estudiado.



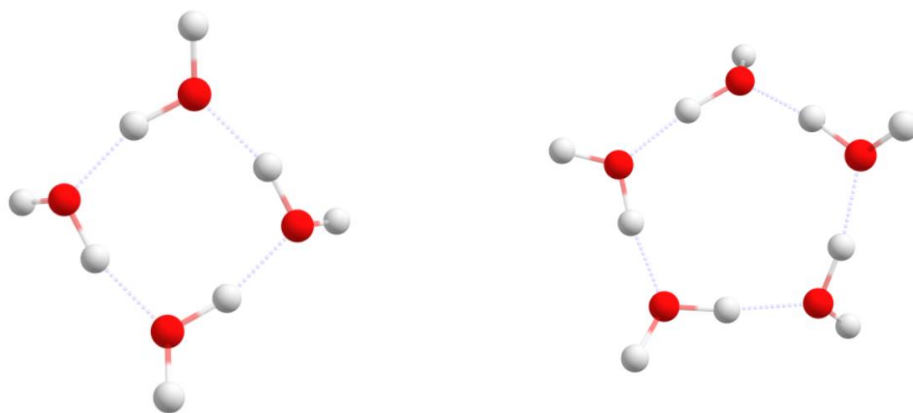


Figura 6.3 Mínimos globales obtenidos para $(\text{H}_2\text{O})_n$ ($n=2-5$) mediante el uso de la integral NCI como descriptor. Tres ciclos de búsqueda iterativa fueron necesarios.

Con el fin de evaluar la contribución real del descriptor NCI-i en la identificación de los mínimos energéticos en clústeres moleculares, se realizó una comparación entre una muestra guiada por el descriptor NCI-i y otra netamente estocástica para los sistemas de agua. Se eligió una población pequeña de 5 isómeros para evitar un posible sesgo en favor de la búsqueda estocástica, la cual tiende a mejorar en poblaciones grandes. Para la muestra guiada por el descriptor de NCI-i, se seleccionaron 5 isómeros del segundo ciclo de optimización.

Los *boxplots* 1 y 3 (Figura 6.4) muestran las energías *single-point* de los 5 isómeros formados en cada método. En general, el método aleatorio muestra buenos resultados en todos los sistemas, excepto en el trímero de agua, donde se observa una gran diferencia de energía entre los isómeros formados, lo cual podría ser fácilmente un artefacto estocástico. Por otro lado, las estructuras formadas por el enfoque guiado NCI-i son menos estables que su contraparte aleatoria. Este fenómeno puede ocurrir y no supone una desventaja para el método guiado por NCI-i, ya que puede ocurrir que una estructura candidata a mínima energética sea alta en energía, pero, a través de un método de optimización por gradiente, lleve a un mínimo energético más estable.

Al considerar la optimización de los isómeros (*Boxplot* 2 y 4), se observa que los isómeros optimizados utilizando el descriptor de NCI-i como guía son más

estables en comparación con los isómeros aleatorios. Esta tendencia se mantiene en la mayoría de los sistemas, con la excepción del dímero de agua donde ambos métodos muestran una convergencia similar, fenómeno que atribuimos a la facilidad de convergencia del sistema.

Además, se observa que a medida que aumenta el número de moléculas de agua, la diferencia en energía entre ambas estrategias es más pronunciada, lo que resalta las ventajas de utilizar el NCIi como descriptor guía en la exploración de la superficie de energía potencial en sistemas moleculares.

Estos resultados dan fe que la elección del descriptor NCI funciona como predictor y que los resultados no se deben a las cualidades de una búsqueda estocástica.

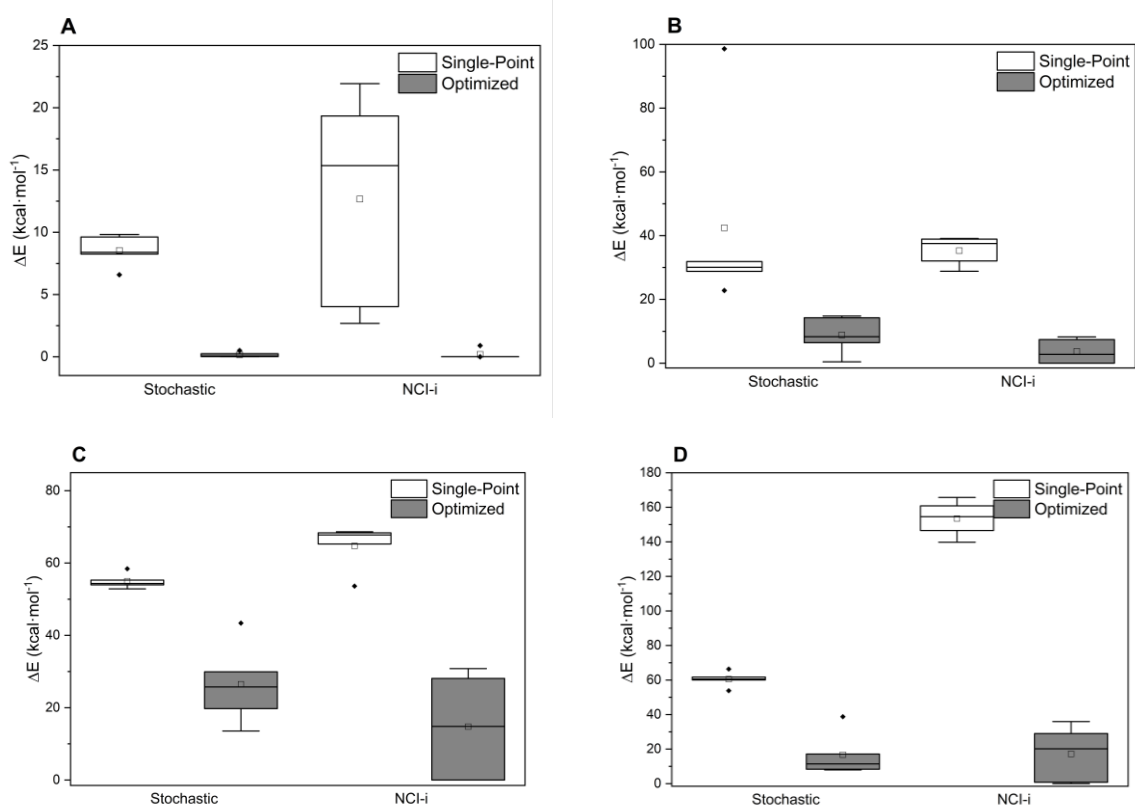


Figura 6.4 Boxplot comparativos entre la energía (Kcal mol^{-1}) de una población dirigida por NCI-i y otra estocástica. A, B, C y D referencian a dímero, trímero, tetrámero y pentámero de agua respectivamente.

Finalmente, en el presente trabajo demostramos que la integración del NCI tiene correlación con la energía de los sistemas moleculares. Mediante un procedimiento simple es posible utilizar el nuevo descriptor para predecir y buscar nuevos clústeres moleculares de menor energía de manera simple y sencilla. En este trabajo se utilizó el método KICK por su sencilla implementación, rapidez de resultados y para verificar que los resultados obtenidos sean por el uso del descriptor y no a la casualidad. Pensamos que utilizar otros métodos de búsqueda con el descriptor propuesto puede resultar en avances en el área de búsqueda de estructuras geométricas en química y biología.

6.5 PERSPECTIVAS

Se presento el primer trabajo realizado en clústeres moleculares predichos por las interacciones no covalentes. El descriptor NCI-i fue utilizado como guía en la exploración de la SEP y se ha demostrado que existe una correlación entre la integración del NCI y la energía de los clústeres moleculares. Se ha desarrollado un programa computacional que automatiza la construcción de inputs para el cálculo del NCI-i, así como la extracción de información relevante de los resultados. Además, se ha implementado un algoritmo de búsqueda basado en el descriptor NCI-i, utilizando el enfoque de enjambre de partículas.

Los resultados obtenidos muestran que el método basado en el NCI-i puede mejorar la estabilidad de los isómeros optimizados en comparación con la selección aleatoria de estructuras. Se ha observado una mejora constante en la energía *single-point* a lo largo de los ciclos de búsqueda, lo que respalda la hipótesis de que el uso del descriptor NCI en la formación de clústeres moleculares es efectivo. Sin embargo, también se observó que la tendencia no se mantiene al considerar las energías optimizadas, lo que indica la necesidad de ajustes adicionales en el algoritmo.

En general, el estudio muestra el potencial del descriptor NCI-i como una herramienta para explorar y predecir la estabilidad de clústeres moleculares. Sin

embargo, es importante tener en cuenta que este trabajo se ha enfocado en sistemas de agua de tamaño limitado y se requiere más investigación para evaluar su aplicabilidad en sistemas más grandes y complejos.

Otra disyuntiva que nace en relación con los resultados es la discrepancia de las energías *single-point* en relación a las energías optimizadas, como se ve en la Figura 6.2. Los resultados hasta ahora muestran que valores cercanos a 0 de NCI-i corresponden a buenas estructuras candidatas, pero que éstas no necesariamente disminuyen los resultados de las estructuras optimizadas. Para abordar este problema se realizó un ejercicio que midió el valor de NCI-i en las estructuras candidatas, tal y como fue diseñada la herramienta en esta tesis, y con los resultados optimizados.

La finalidad de este ejercicio es: (I) verificar que la correlación energía/NCI-i se mantiene con los resultados optimizados y (II), averiguar si el método estocástico es el correcto de implementar.

Para este ejercicio se utilizó una población de $100N$, con N el número de moléculas en clústeres de agua (dímero - pentámero). La Figura 6.5 muestra los resultados obtenidos para los clústeres más grandes, a la izquierda la correlación NCI-i y *single-point* para el conjunto poblacional, mientras que el lado derecho muestra la correlación NCI-i/energías optimizadas.

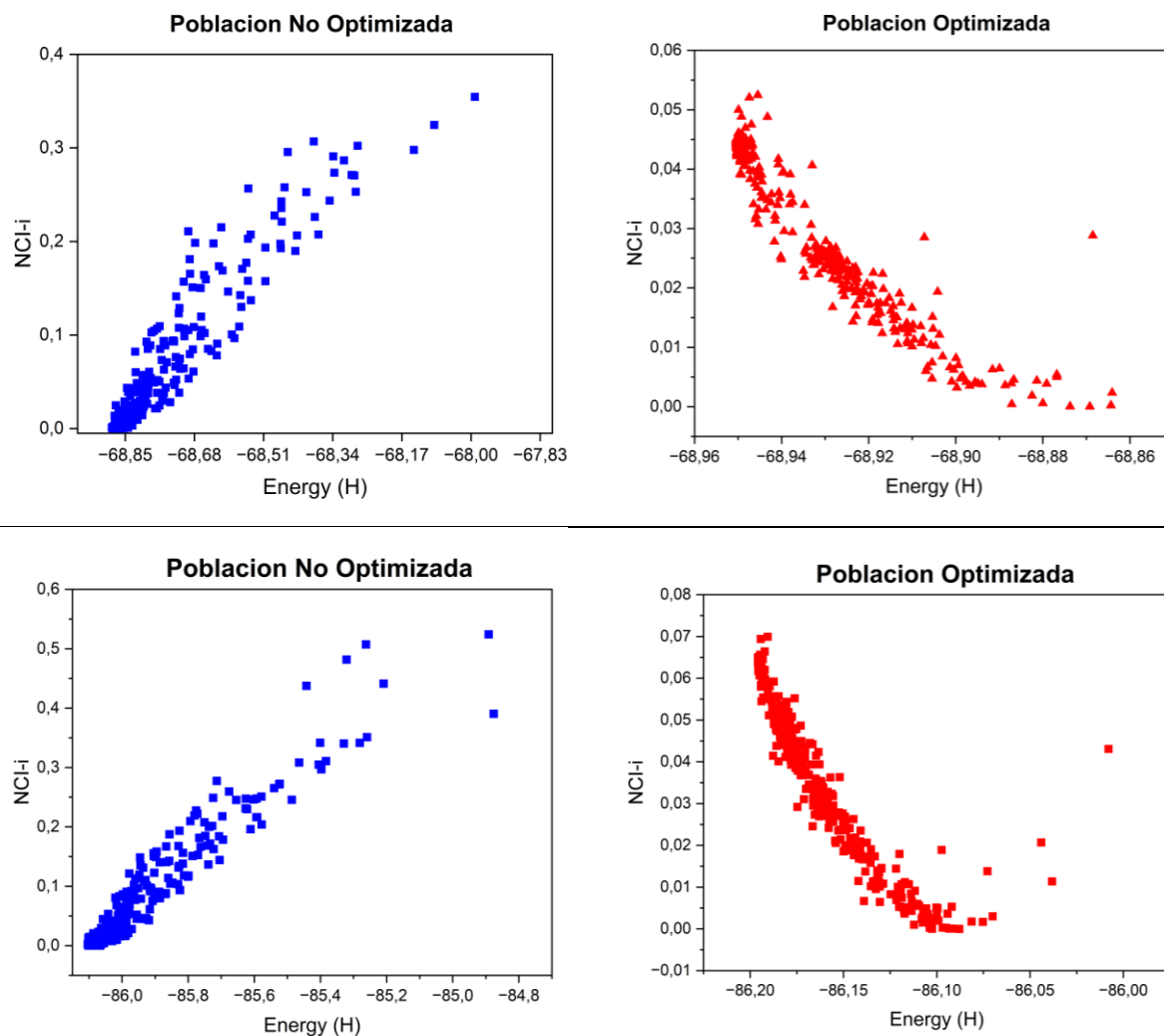


Figura 6.5 Correlación energía/NCIi para clústeres moleculares $(\text{H}_2\text{O})_4$ (Arriba) y $(\text{H}_2\text{O})_5$ (Abajo). A la izquierda correlación de energía *single-point* de la población aleatoria; a la derecha correlación con la energía optimizada. Eje abscisas indica energía en Hartree, eje ordenada valor NCI-i.

Para los clústeres de agua la mejor interacción energía/NCI-i pareciese ser con valores muy cercanos a 0, sin embargo, al estudiar las estructuras optimizadas vemos que los valores con menor energía se alejan de 0 por una diferencia menor. Esta nimia diferencia resulta muy difícil de abordar bajo el método presentado, pues alcanzar este nivel de valores NCI mediante un algoritmo KICK resulta altamente improbable.

Se realizó un análisis comparativo de los promedios de los valores NCI-i para la generación estocástica y los resultados optimizados. En el caso del dímero de agua,

el clúster más simple estudiado, se obtuvo un promedio de NCI-i de 0.041 para la generación estocástica, mientras que para los resultados optimizados se obtuvo un valor de NCI-i de 0.004. Resulta evidente que un enfoque estocástico en la generación de poblaciones de los ciclos de convergencia es insuficiente para alcanzar mínimos energéticos estables sin pasar por un método determinístico por gradiente.

Este ejercicio rápido resalta la dificultad de obtener valores de NCI-i muy cercanos a cero utilizando únicamente el enfoque estocástico, y sugieren la necesidad de combinar diferentes enfoques, como el uso de métodos determinísticos basados en el gradiente, así como también otros métodos de exploración de la SEP que aseguren ensambles entre fragmentos con lógica química.

Finalmente, como ya se habló con anterioridad, las NCI se encuentran bien definidas en sistemas biológicos. Las macromoléculas como proteínas, ADN y ARN, así como las interacciones con la bicapa lipídica de las células poseen varias interacciones y propiedades producto de las diferentes NCI.

La estrategia diseñada en esta tesis permite la predicción de clústeres moleculares utilizando NCI-i como guía, dando buenos resultados en clústeres de agua. Las macromoléculas suponen un desafío más grande en todo sentido, pero no descartamos que esta metodología signifique un avance para el desarrollo científico en estas áreas.

Sin ir más lejos, en el estudio unión entre ligandos y proteínas se basan en las interacciones no covalentes¹⁷⁶, donde los enlaces iónicos suponen las interacciones más fuertes. En el modelo de llave-cerradura, un ligando se une a un sitio activo de una proteína siguiendo varias reglas específicas: (I) estructural, el sitio activo debe poseer una cuenca y/o estar preparada para recibir un ligando, (II) especificidad, cada sitio activo de una proteína es capaz de recibir un tipo de ligando en particular, esta especificidad la entregan los componentes químicos del ligando, así como también del sitio activo y (III) poder de retención que significa que la

proteína asociada con el sitio activo realizara una acción (o dejara de hacer) mientras el ligando esté conectado a ella, por lo cual el poder de retención es esencial para su funcionalidad.

Las interacciones no covalentes están ligadas a cada etapa previamente señalada y, bajo un esquema de mecánica molecular, diferentes drogas e interacciones proteína-proteína se han observado y diseñado. La mecánica molecular, generalizando, simplifica las interacciones químicas a potenciales y campos de fuerza sin tomar en consideración componentes cuánticos. Bajo la estrategia propuesta, que también trabaja con una aproximación promolecular, podrían significar una adición interesante a los protocolos de *docking* ligando proteína y/o diseño de complejos proteicos.

CAPÍTULO VII COMPARACIÓN DE MÉTODOS DESARROLLADOS

"En este capítulo, aplicamos diversas estrategias de exploración para investigar las superficies de energía potencial (SEPs) de los clústeres $CBe_5H_5^+$ y $C_nBe_{3n+2}H_{2n+2}^{2+}$ ($n = 2-4$). Aunque investigaciones anteriores sugirieron que los carbonos pentacoordinados planares (ppCs) representaban los mínimos globales de estos sistemas, nuestros resultados identificaron nuevos mínimos globales putativos e isómeros competitivos, cuestionando algunas de las asignaciones previas. Utilizamos una variedad de métodos, incluidas aproximaciones estocásticas inspiradas en la evolución y guiadas por 'criterios químicos', así como simulaciones de dinámica molecular ab initio a altas temperaturas. Descubrimos que el tamaño de la población explorada tiene un impacto significativo en la eficacia del método evolutivo. Además, observamos que los enfoques restringidos o guiados ofrecen ventajas para identificar mínimos energéticos óptimos en sistemas de mayor tamaño. Este estudio resalta la utilidad de emplear múltiples estrategias complementarias para hallar una gama más amplia de mínimos energéticos. Nuestros resultados aportan perspectivas valiosas para la exploración de SEPs en clústeres, especialmente en aquellos de tamaño medio que podrían actuar como nexos entre clústeres pequeños y nanomateriales."

Abstract de "Revisiting the potential-energy surface of $C_nBe_{3n+2}H_{2n+2}^{2+}$ ($n = 2-4$) clusters: are planar pentacoordinate carbon structures the global minima?"
(Anexo 3, Artículo 15)

7.1 RESULTADOS - PTC C-BE-H

Se realizó un estudio comparativo sobre la eficacia de diferentes métodos para identificar las estructuras GM de los clústeres de $C_nBe_{3n+2}H_{2n+2}^{2+}$ con $n=2-4$, reportados previamente como especies de carbonos pentacoordinados planares (ppCs), apuntando al potencial de crecimiento hacia un nanocintas de C-Be de doble cadena en zigzag¹⁸. Comparamos algoritmos genéticos y estocásticos con restricciones espaciales y búsquedas guiadas basadas en la "intuición química".

Además, realizamos simulaciones de dinámica molecular *ab-initio* a altas temperaturas partiendo de isómeros ppCs y capturando mínimos alternativos en la superficie de energía potencial.

Identificamos nuevos GM putativos para $C_2Be_8H_6^{2+}$ y $C_3Be_{11}H_8^{2+}$ y un mínimo competitivo para $C_4Be_{14}H_{10}^{2+}$. Sin embargo, en este último caso, nuestra búsqueda se limitó únicamente a estrategias guiadas; por tanto, sigue siendo posible que existan estructuras de mínima energía sin detectar.

El clúster atómico $CBe_5H_4^+$ ha sido previamente reportado por Merino, Zhai y colaboradores¹⁷⁷ como mínimo global. Esta estructura tiene simetría C_{2v} , con un ppC con 17 electrones de valencia. Debido a su estabilidad este sistema puede ser utilizado como bloque de ensamble en nanocintas de mayor tamaño. El trabajo de Jia Jia¹⁸ reportó estas construcciones (dímero, trímero y tetramero) como mínimos globales, aunque solo el dímero ($C_2Be_8H_6^{2+}$) se obtuvo por una exploración rápida de la SEP.

Se compararon 5 metodologías diferentes.

- **MatrixAutomaton (i):** Utilizando una población general de $5N$ (N = número de átomos) con un ciclo de convergencia de 9 generaciones.
- **MatrixAutomaton Especial (ii):** Donde se guió al software a construir una población sólo en 2 dimensiones. Además, se construyó el sistema de manera concéntrica, ubicando los átomos de carbono inicialmente, luego los átomos de berilio y finalmente los átomos de hidrogeno, tal y como se propuso en el trabajo de Yañez¹¹⁹ en la búsqueda del sistema Be_6Be_{11} .
- **Kick de 2 dimensiones (iii):** se utilizó el algoritmo KICK en su forma simplificada para explorar la reactividad de los sistemas utilizando fragmentos. En este caso, se emplearon como fragmentos el monómero junto con una cantidad específica de átomos faltantes en la estructura.
- **Método guiado por permutación (iv):** Usando *templates* y patrones estructurales para construir isómeros con a la estequiometría deseada

siguiendo dos criterios: (i) usando como *templates* moleculares los acenos $C_{4n+2}H_{2n+4}$ ($n=2-4$), reemplazando, agregando, eliminando y permutando átomos según la Figura 7.1. Y (ii) usando patrones estructurales identificados como GM putativos en clústeres de menor tamaño, es decir, desde $C_2Be_8H_6^{2+}$ para buscar $C_3Be_{11}H_8^{2+}$ y luego $C_3Be_{11}H_8^{2+}$ para buscar $C_4Be_{14}H_{10}^{2+}$.

- **Dinámica molecular Born-Oppenheimer (v)**¹⁷⁸: a 900K, manteniendo la temperatura constante utilizando el termostato Nose-Hoover^{179, 180}. Se realizó un tiempo de 30ps con un *timestep* de 1ps inicializando de las estructuras ppCs.

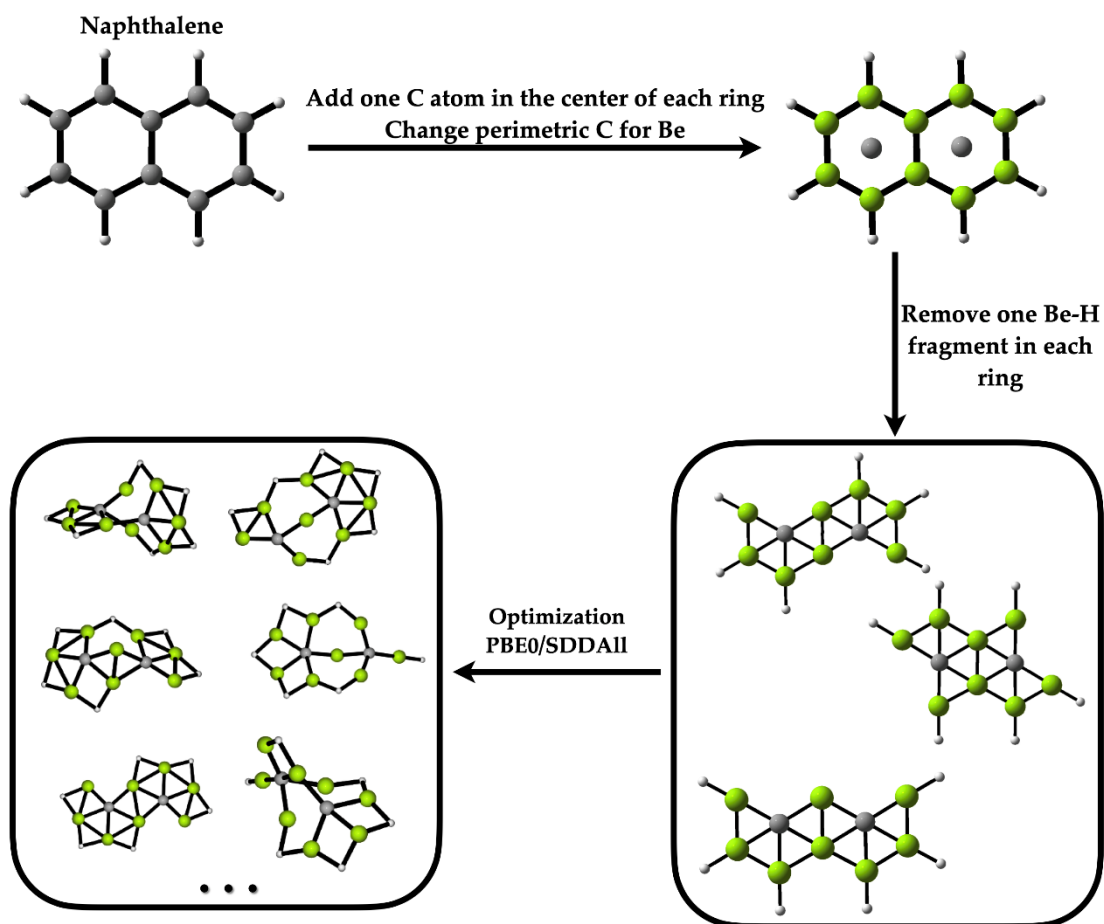


Figura 7.1. Metodología de permutación aplicada a los sistemas $C_{4n+2}H_{2n+4}$ ($n=2-4$); acenos fueron utilizados como sistemas *template*. Átomos de carbono son ubicados al centro de cada anillo, luego el perímetro de carbono se cambia por berilio. Adicionalmente fragmentos de Be-H son eliminados de cada anillo hasta que la estequiometría deseada es alcanzada.

Las Figuras 7.2 y 7.3 presentan los resultados obtenidos en la exploración de la SEP de las nanocintas estudiadas. Cada estrategia utilizada en el estudio tiene sus ventajas y desventajas, y es necesario tener en cuenta estos aspectos para obtener una visión completa de la SEP. La unificación de los resultados de las diferentes estrategias permite abordar de manera más completa y exhaustiva la complejidad de la SEP estudiada. A continuación, se destacarán brevemente las diferencias más importantes entre las metodologías presentadas.

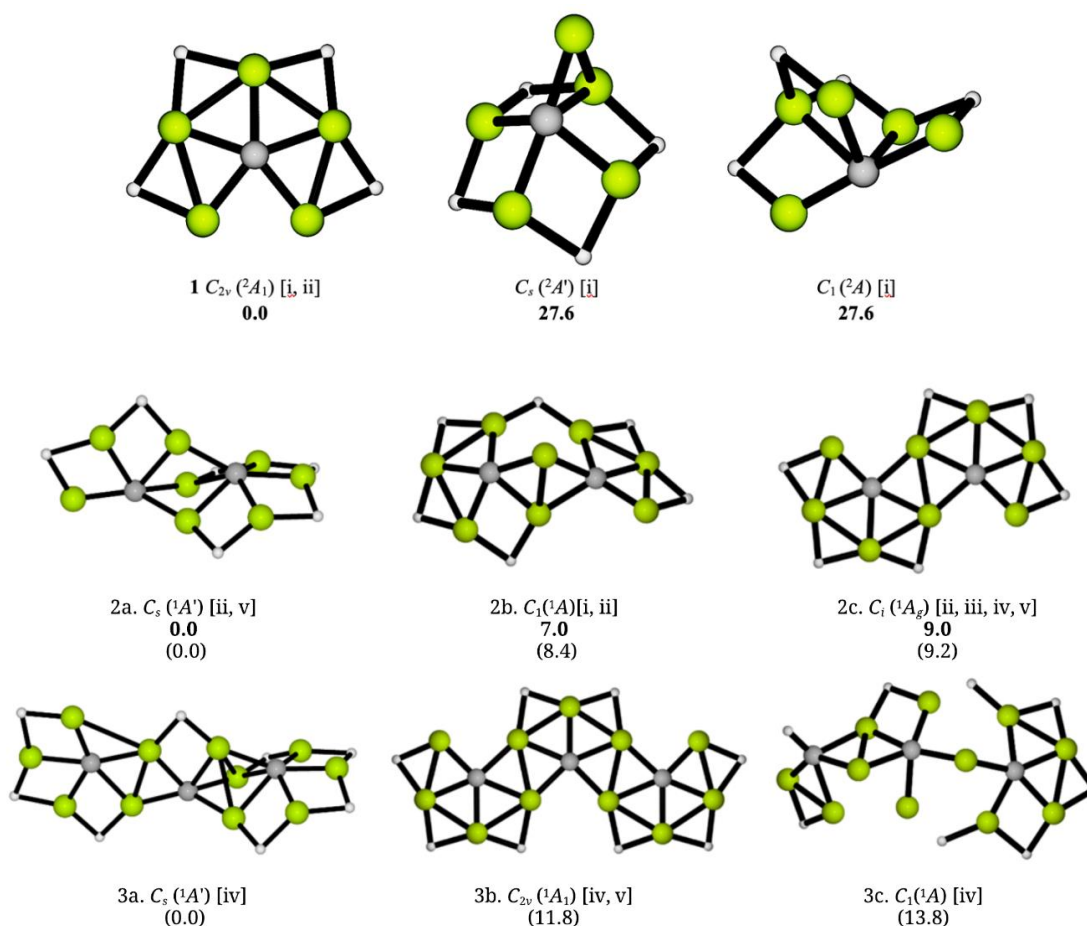


Figura 7.2 Estructuras de mínima energía del monomero, dimero y trimero estudiado. Entre corchetes [] el método de búsqueda responsable de la identificación del isómero. Energías relativas presentadas al nivel DLPNO-CCSD(T)¹⁸¹/CBS^{182, 183}//PBE0-D3¹⁸⁴/Def2-TZVP (negrita) y PBE0-D3/Def2-TZVP incluyendo corrección de punto cero.

Tanto **i** como **ii** fueron capaces de converger sin problemas en el mínimo global reportado en la literatura para el monómero. Brindándonos confianza en que estos métodos son capaces de explorar de manera efectiva sistemas con esta complejidad, lo cual no había sido investigado previamente. Cabe agregar que los métodos **iii** y **iv** no fueron aplicados a este sistema, pero si se utilizó el monómero como fragmento en el proceso KICK (**iii**).

El mínimo global reportado para $C_2Be_8H_6^{+2}$ es un C_i cuasi planar que es la extensión del monómero $CBe_5H_4^+$. Nuestros resultados identifican dos mínimos energéticos más estables por hasta 9 Kcal mol⁻¹. Fueron identificados por los métodos **ii** y **v**. El mínimo reportado con anterioridad fue también identificado por los métodos **ii**, **iii**, **iv** y **v**.

Cabe destacar que **i** no identificó esta estructura, no así **ii** que es el método guiado con intuición del investigador. Este resultado resalta el valor del conocimiento previo sobre el comportamiento de estos sistemas y las mejoras implementadas en MatrixAutomaton para la exploración de la SEP.

El sistema $C_3Be_{11}H_8^{+2}$ C_{2v} se identifica gracias a **iv** y **v**; el método TemplatePermuta utilizó el antraceno como sistema base formando buenos sistemas candidatos, no obstante, se identificó un mejor mínimo energético con simetría C_s por **iv**. El modo fue extender el mínimo global $C_2Be_8H_6^{+2}$ en un proceso guiado que lograra comprender las cualidades de los mejores sistemas en la familia y como afectarían a los resultados generales.

El sistema de mayor envergadura de este estudio $C_4Be_{14}H_{10}^{+2}$ corresponde al mínimo energético reportado en la literatura, aunque por su tamaño solo el método guiado (**iv**) y la BOMD (**v**) fueron realizadas, no se descarta mejores isómeros. La Figura 7.3 recoge los resultados obtenidos, donde el segundo mínimo energético es altamente competitivo a tan solo 0.2 Kcal mol⁻¹ de energía.

El artículo “Revisiting the Potential Energy Surface of $C_nBe_{3n+2}H_{2n+22+}$ ($n=2-4$) Clusters: Are Planar Pentacoordinate Structures the Global Minima?”¹⁸⁵ (Artículo 15, Anexo 2) abarca los resultado hasta ahora discutidos.

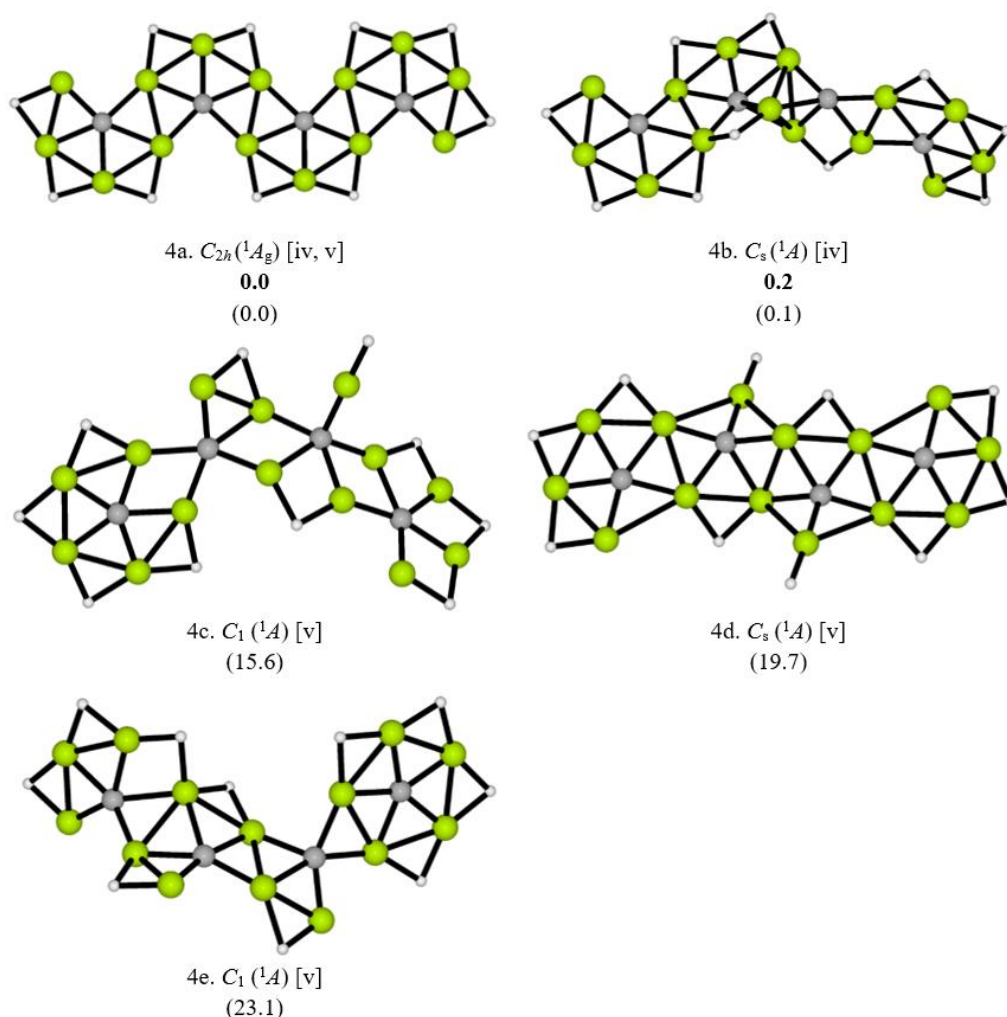


Figura 7.3 Optimized structures, point groups, method of search, and relative energies (kcal mol^{-1}) of low-lying $C_2Be_8H_6^{2+}$ and $C_3Be_{11}H_8^{2+}$ isomers at DLPNO-CCSD(T)/CBS//PBE0-D3/Def2-TZVP (bold) and PBE0-D3/def2-TZVP level, including zero-point energy (ZPE) corrections.

La Tabla 7.1 muestra una comparación de los tiempos que cada estrategia tomó para explorar la SEP del trímero. El método **iv**, que utiliza la estrategia de permutación selectiva y guiada por el conocimiento previo, logró identificar 10 mínimos energéticos, incluyendo el mínimo global, en tan solo una hora. Sin

embargo, es importante destacar que esta estrategia obtuvo buenos resultados debido a la elección adecuada del modelo a permutar.

Método	Tiempo de Cálculo	Población Total Evaluada
MatrixAutomaton (i)	28 Hrs	1550
MatrixAutomaton Especial (ii)	7 Días 19 Hrs	2640
KICK (iii)	7 Hrs 30'	255
Guiado Permutación/Intuición (iv)	1 Hr 20'	10

Tabla 7.1 Comparativa de tiempos para la exploración de la SEP del sistema $C_3Be_{11}H_8^{2+}$ y la cantidad de isómeros arrojados.

Por otro lado, la búsqueda estocástica (iii), aunque no produce mínimos competitivos en términos de energía, puede escalar en tiempo a medida que se identifican más mínimos. Es decir, a medida que se exploran más estructuras, el tiempo requerido para el análisis aumenta.

En resumen, el método iv fue más eficiente en términos de tiempo al identificar mínimos energéticos en comparación con la búsqueda estocástica. Se debe tener en cuenta que la estrategia utilizada en el método iv fue clave para obtener resultados favorables. Finalmente, la exploración realizada por MatrixAutomaton entrega una alta cantidad de estructuras, pero su costo en tiempo escala de sobremanera y no logra identificar el GM.

Este resultado pareciese sugerir que la metodología i y ii no entrega buenos resultados en sistemas grandes haciendo evidente la necesidad de combinar diferentes estrategias de búsqueda junto con la intuición química del científico.

7.2 PERSPECTIVAS

En este estudio se compararon diversas estrategias para explorar las SEP de los clústeres $CBe_5H_5^+$ y $C_nBe_{3n+2}H_{2n+2}^{2+}$ ($n=2-4$). Trabajos anteriores propusieron que los mínimos globales de estos cúmulos contienen carbonos pentacoordinados

planares (ppCs), con la aclaración de que para $n = 2$, existe una ligera desviación de la planaridad.

MatrixAutomaton, MatrixAutomaton guiado, método estocástico y guiados por “criterio químico” fueron comparados bajo los resultados obtenidos, tiempo de cálculo y población total evaluada.

Nuestro estudio identificó nuevos mínimos globales putativos para $n = 2$ y 3; y un isómero competitivo con la estructura ppC para $n = 4$, con una diferencia energética menor a 1 kcal.mol^{-1} . Sin embargo, nuestros hallazgos contradicen las asignaciones de los trabajos anteriores de los sistemas ppC a los mínimos globales putativos.

MatrixAutomaton funcionó bien para sistemas pequeños de hasta $n = 2$, pero no identificó el mínimo global putativo para $n = 3$. Estos resultados sugieren que el tamaño de la población explorada debe crecer exponencialmente con el número de átomos del sistema estudiado para lograr una búsqueda de alta calidad. Los métodos restringidos o guiados mostraron una ventaja a la hora de identificar mejores mínimos para sistemas más grandes.

Nuestro estudio demuestra que usar múltiples estrategias complementarias resulta en una mayor variedad de mínimos identificados. Búsquedas guiadas por “intuición química” son una alternativa excelente para explorar sistemas de mayor tamaño. Aunque siempre dependerá de la calidad de la información utilizada.

No existe un método superior, así como tampoco es posible identificar todos los mínimos energéticos mediante el uso de una estrategia única.

CAPÍTULO VIII DISEÑO DE HERRAMIENTAS

"En este capítulo, se describen detalladamente las herramientas desarrolladas durante la realización de esta tesis. Se pone especial énfasis en las características que se buscaban, los algoritmos de optimización que se han implementado y la importancia de la información química utilizada, especialmente en relación con su capacidad predictiva." - **Resumen de Capítulo.**

8.1 GENERALIZACIÓN

Los mecanismos utilizados a lo largo de esta tesis se obtuvieron gracias a herramientas y metodologías diseñadas para la exploración exhaustiva de la SEP. A continuación, se recorrerá de manera rápida las características y detalles de cada implementación evaluada en esta investigación. El uso de cada herramienta se puede leer en detalle en los capítulos anteriores.

Primero se presenta al lector el diseño para la mejora del programa AUTOMATON, con un cambio de enfoque programático que le permitirá la construcción más rápida de sistemas, mejorar el escalamiento con átomos de mayor enlace covalente y la adición de construcción mediante fragmentos atómicos.

Luego, se hablará brevemente de un nuevo software computacional de intercambio atómico dirigido, donde gracias a una molécula modelo será posible la construcción de nuevos sistemas que compartan características isomórficas con el modelo original.

Los siguientes diseños presentados combinan simples implementaciones estocásticas con información química relevante, en específico, la función de Fukui que entrega una guía para sistemas covalentes; y la integral de las interacciones no covalentes para sistemas dominados con este tipo de fuerzas.

Finalmente, a modo de ejercicio se presenta un diseño general de una suite compilatoria con los métodos hasta aquí presentados, que permitiría al usuario escoger el método ideal para la exploración de la SEP de su sistema problema particular.

8.2 DISEÑO HERRAMIENTAS - MATRIXAUTOMATON

La implementación de los algoritmos genéticos se realizará mediante una actualización del software computacional AUTOMATON, desarrollado por el grupo de investigación, a una nueva versión que permitirá una mejor conexión con diferentes tipos de información. La actualización permitirá establecer una sinergia entre los fragmentos moleculares, los tipos de mutaciones, las permutaciones atómicas guiadas, entre otros aspectos relevantes.

Los cambios propuestos a AUTOMATON son:

- Cambio de lenguaje de programación de Perl a Python3: Se migrará el código del software de Perl a Python3, lo cual permitirá aprovechar las ventajas y funcionalidades de este lenguaje de programación más moderno y ampliamente utilizado en la comunidad científica.
- Adaptación de la grilla del autómatas celular: Se modificará la estructura de la grilla del autómatas celular, pasando de un enfoque cartesiano a uno basado en matrices. Permitiendo un manejo más eficiente y flexible de los elementos en la grilla, facilitando su manipulación y análisis.
- Escalamiento de celdas de interacción según los tipos de átomos: Se implementará un mecanismo de escalamiento de las celdas de interacción de acuerdo con los tipos de átomos/moléculas presentes en el sistema. Garantizando un tratamiento adecuado de las interacciones entre los diferentes tipos de átomos y mejorará la precisión de los resultados.
- Población de celdas con moléculas: Además de los átomos, se permitirá la población de las celdas de la grilla con moléculas completas. Ampliando la

capacidad del software para simular sistemas moleculares más complejos y realistas.

- Mayor control del usuario en variables del programa: Se brindará a los usuarios un mayor control sobre las variables y parámetros propios del programa.

La implementación del protocolo de matrices permite un enfoque de posiciones en el espacio independiente del objeto utilizado para poblar la matriz, es decir, AUTOMATON ya no estaría limitado a trabajar solamente con posiciones atómicas en el plano cartesiano, si no que puede ser poblado con fragmentos moleculares, clústeres atómicos, formas moleculares, e incluso artefactos biológicos como proteínas y enzimas.

Con el fin de diferenciar la mejora de AUTOMATON con el original, un cambio de nombre fue propuesto: **MatrixAutomaton**.

8.2.1 MATRIXAUTOMATON

AUTOMATON es un software de exploración de la SEP que combina un autómeta celular con un algoritmo genético. En resumen, AUTOMATON está compuesto por las siguientes etapas:

- **Algoritmo de Autómata Celular.**
 - Creación de la grilla de trabajo.
 - Posicionamiento de los pobladores en grilla.
 - Control de celdas prohibidas.
 - Evolución del autómeta celular.
- **Algoritmo Genético**
 - Asignación del *Fitness*.
 - Recombinación de población.
 - Mutación de población.
- **Envío de cálculos químicos**
 - Creación de archivos de entrada para programa Gaussian16.

- Envió en clúster SGE.
- Control de errores.
- **Recopilación de información.**
 - Lectura de archivos de salida de Gaussian16.
 - Jerarquización de resultados.
- **Control de repetidos.**

Que se encuentran programados bajo un paradigma lineal en un gran código escrito en Perl.

La primera etapa consistió en la distribución de las tareas del algoritmo bajo un paradigma orientado a objetos. Además, se hizo un énfasis en dividir cada tarea en *scripts* particulares, los cuales interaccionan bajo las directrices de un código padre, gestor de tareas (Tabla 8.1). El nuevo programa, denominado MatrixAutomaton, fue programado completamente en Python3.

Nombre <i>Script</i>	Descripción	Interacción
Gestionador	Programa principal, gestiona tareas de cada <i>script</i> , control de evolución del programa, control de orden de término del programa, lectura de archivos de inicio.	<ul style="list-style-type: none"> ● MatrixAutomaton ● Impresora ● Lector ● Lanzador ● Var ● Genetic
MatrixAutomaton	Algoritmo de autómeta celular, reglas de evolución, control de población de celdas.	<ul style="list-style-type: none"> ● Gestionador
Lector	<i>Script</i> de lectura de cualquier archivo necesario (XYZ, salidas de Gaussian16), recopila	<ul style="list-style-type: none"> ● Gestionador ● Lanzador

	información relevante y la entrega en formato legible al programa.	
Impresora	<i>Script</i> de escritura de cualquier archivo necesario (XYZ, entrada de Gaussian16, archivos tipo log).	<ul style="list-style-type: none"> • Gestionador
Genetic	Algoritmo genético, tareas de recombinación de individuos, mutación de individuos y control de convergencia.	<ul style="list-style-type: none"> • Gestionador • Impresora • Lector
Var	Archivo de variables globales y locales necesarias para el funcionamiento del programa, por ejemplo: número mínimo de ciclos sin cambios en el mínimo energético encontrado, nombre de archivos de salida, tamaño población, nombre del sistema.	<ul style="list-style-type: none"> • Gestionador
Lanzador	Simple script que llama a programas externos de optimización.	<ul style="list-style-type: none"> • Gestionador

Tabla 8.1: Desglose de los *scripts* usados en el software MatrixAutomaton.

La implementación más importante en MatrixAutomaton fue el diseño del algoritmo autómeta celular. AUTOMATON utiliza varias funciones de construcción de la grilla de trabajo basado en el plano cartesiano:

- Cálculo del valor máximo posible para un sistema lineal tomando los átomos del sistema a buscar. Ejemplo: para el agua el valor máximo posible

será la sumatoria de los radios covalentes de 2 átomos de hidrogeno (0.31 Å) y un átomo de oxígeno (0.66 Å), en total 1.28 Å.

- Construcción del espacio cartesiano de búsqueda de tamaño tomado el doble del valor máximo posible. Proceso se hace para conjuntos en 1, 2 y 3 dimensiones.
- El espacio de interacción cartesiano es dividido en celdas de tamaño 0.31 Å (átomo hidrogeno), cada celda se le asocia un espacio en memoria con su posición y estado.

Este proceso completo se repite cada vez que el proceso de autómatas celulares es requerido durante la ejecución del programa. `MatrixAutomaton` cambia el paradigma de este proceso al utilizar la librería matemática *Numpy* de Python3, donde la construcción de la grilla, la interacción de las celdas y el estado de estas se hacen utilizando los objetos de matrices.

El nuevo proceso se resume en:

- Se toman los radios covalentes de cada átomo en el sistema entregado.
- Los radios son escalados según el menor radio covalente presente. Luego son redondeados al número entero más cercano. Ejemplo: Para agua el radio covalente menor corresponde al átomo de hidrogeno (0.31 Å). Escalando el hidrogeno a un valor 1.0 provoca que el átomo de oxígeno (0.66 Å) se encuentre a una razón de 2.13. Finalmente, al redondear quedamos con los valores covalentes ajustados para el átomo de hidrogeno como 1.0 y el oxígeno como 2.0.
- El valor máximo será la suma de los átomos covalentes ajustados del sistema.
- Se guarda en memoria el objeto matriz de tamaño acorde al valor máximo del sistema.

Donde no existe procesamiento de creación de la grilla de interacción; no se guardan en memoria cada celda con su posición y estado; tampoco se trabaja con

números decimales. Los objetos matrices en programación están diseñados con el fin de guardar información de manera eficiente, rápida y simple.

En el autómata celular de AUTOMATON, uno de los procesos clave es la definición del radio de interacción de una celda poblada para determinar las celdas probables en la grilla. Este radio de interacción se establece utilizando el concepto de radio covalente.

AUTOMATON, al estar diseñado bajo un enfoque cartesiano, calcula la distancia cartesiana entre la celda poblada y todas las celdas de la grilla. Luego, se determina qué celdas se encuentran dentro del radio covalente definido. Estas celdas que se encuentran dentro del radio covalente son asignadas como celdas probables, lo que indica que pueden interactuar o formar parte del sistema en estudio.

MatrixAutomaton, a diferencia de AUTOMATON, no requiere calcular la distancia cartesiana con todas las celdas de la grilla de trabajo. En su lugar, se utiliza el algoritmo del círculo de punto medio¹⁸⁶ (*Midpoint Circle Algorithm*) para determinar de manera eficiente las celdas que se encuentran a una distancia específica de la celda poblada.



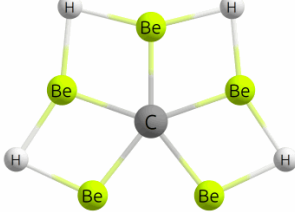
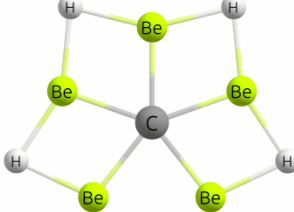
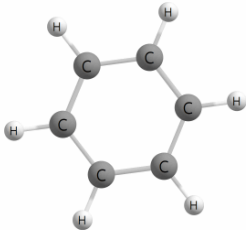
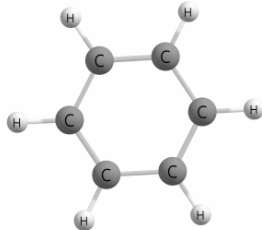
El algoritmo del círculo de punto medio permite trazar rápidamente los puntos en una circunferencia dada su posición central y su radio. En el contexto de MatrixAutomaton, este algoritmo se utiliza para identificar las celdas que están a una distancia X de la celda poblada. En lugar de calcular la distancia cartesiana con todas las celdas, se traza una circunferencia con centro en la celda poblada y radio X , y se marcan las celdas que intersecan esta circunferencia.

Finalmente, AUTOMATON define celdas con un estado prohibido que son aquellas que no pueden ser pobladas por impedimento estérico entre átomos. Para definir las nuevamente se calcula la distancia cartesiana entre todas las celdas pobladas en cada iteración del autómata celular. MatrixAutomaton simplemente define las celdas prohibidas guardando su ubicación en listas de prohibición desde

su descubrimiento en el algoritmo de círculo medio, siendo innecesaria su cálculo en cada iteración.

Se evaluó el nivel de complejidad de ambos programas, AUTOMATON tiene una complejidad $O(n^7)$ mientras que MatrixAutomaton de $O(n^5 \log_2 n)$.

Se hicieron búsquedas comparativas de los sistemas de agua, benceno, borazina, CBe_5H_4^+ y $\text{C}_2\text{Be}_{10}\text{H}_8^{2+}$ evaluando el mínimo energético encontrado y el tiempo de ejecución (Tabla 8.2).

Sistema	Tiempo de ejecución		Mínimo Global Putativo	
	AUTOMATO N	MatrixAutomaton	AUTOMATON	MatrixAutomaton
H_2O N = 3	21 minutos	8 minutos	-17.2 H 	-17.2H 
CBe_5H_4^+ N = 10	1 día, 6 horas, 35 minutos	14 horas, 38 minutos	-13.2 H 	-13.2 H 
Benceno C_6H_6 N = 12	2 días, 9 horas, 1 minuto	8 horas, 2 minutos	-37.6 H 	-37.6 H 

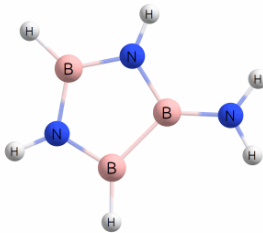
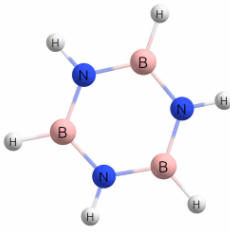
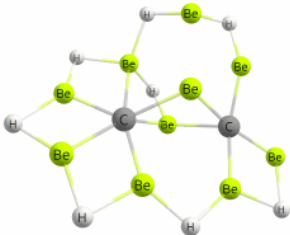
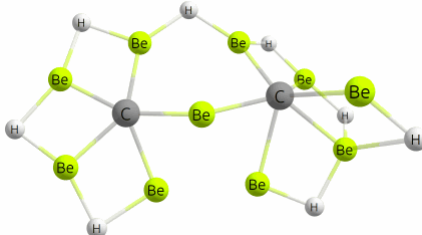
Borazina $B_3N_3H_6$ N = 12	19 horas, 20 minutos	9 horas, 1 minuto	-41.96 H		-42.03 H	
C₂B₁₀H₈²⁺ N = 22	14 días, 3 horas, 40 minutos	2 días, 12 horas, 47 minutos	-26.43 H		-26.46 H	

Tabla 8.2: resultados de la comparación entre los softwares AUTOMATON y MatrixAutomaton.

Se demuestra entonces que MatrixAutomaton supone un ahorro de recursos computacionales importante en relación con su predecesor. Además, la calidad de los resultados obtenidos no se vio afectada.

8.3 DISEÑO HERRAMIENTAS - TEMPLATEPERMUTA

Dentro de los procesos genéticos de un algoritmo genético, la mutación tiene un papel fundamental porque permite una exploración dirigida de la SEP evitando caer en agujeros incorrectos. Existen diversos tipos de mutación como movimientos de átomos, intercambios de posiciones atómicas y nuevas estructuras con formas específicas (anillos, coronas, helicoidales).

Sin embargo, muchas veces el conocimiento del investigador permite la predicción del tipo de estructura que debiese formar un sistema (planos, coronas, esferas). Muchas de ellas por compartir las propiedades isoelectrónicas con estructuras ya conocidas, lo que entrega una idea del comportamiento general de los sistemas.

Siguiendo esta idea, los procesos de permutación desarrollados consisten en el intercambio isoelectrónico de un sistema modelo (*template*) hacia un sistema objetivo. Aunque, debido a las posibles diferencias en la ecuación estequiométrica

entre el modelo y el objetivo, puede ser necesario eliminar algunas posiciones atómicas durante el proceso de permutación. Para abordar este desafío, se ha establecido un conjunto de reglas que determina qué átomos deben ser eliminados y cuáles deben mantenerse en el proceso de permutación. Estas reglas guían la generación de sistemas objetivo a partir de un sistema modelo único, ampliando la diversidad de soluciones posibles en la búsqueda de clústeres moleculares. La Figura 8.1 ilustra este concepto y muestra cómo se generan diferentes sistemas objetivo a partir de un sistema modelo mediante la aplicación de las reglas de permutación eliminatoria. Este enfoque permite explorar un espacio más amplio de soluciones y mejorar la capacidad para descubrir clústeres moleculares con propiedades novedosas.

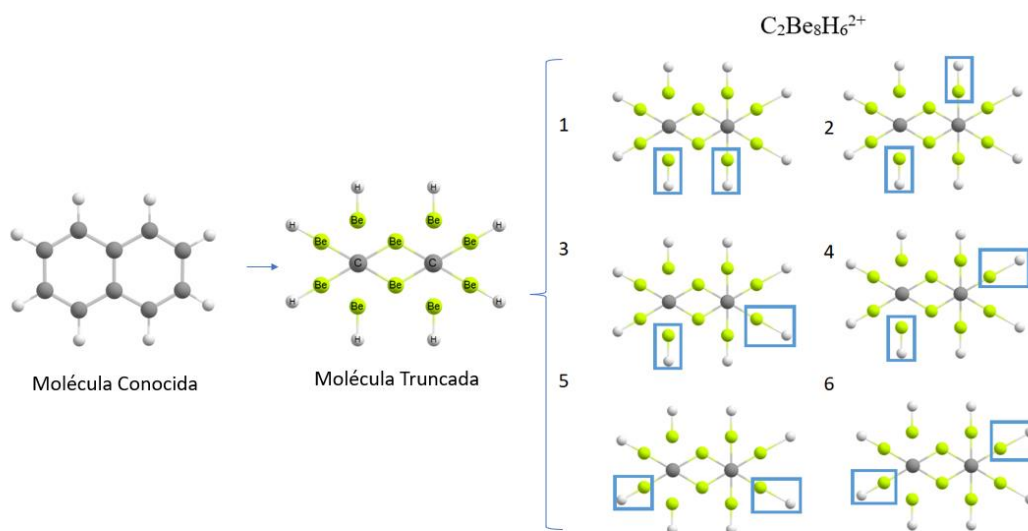


Figura 8.1. Esquema del procedimiento de permutación. Para encontrar el sistema $C_2Be_8H_6^{+2}$ se utiliza como *modelo* la molécula de Naftaleno. Se intercambian los átomos de C por Be; se agrega un átomo de C en el centro de los anillos. La molécula troncada sufre procesos de permutación dirigida para alcanzar la estequiometría objetivo. Como regla se elimina un dímero $-BeH$ por cada anillo. Un total de 6 permutaciones son posibles bajo estas reglas.

El proceso de permutación se lleva a cabo en los siguientes pasos:

- Selección del *template*: Se elige una estructura modelo o *template* que será sometida a permutación. La elección de este *template* es guiada por el objetivo de la investigación.
- Selección de átomos para permutación: Se determinan los átomos del *template* que serán permutados por otros tipos de átomos. Esta selección puede basarse en criterios específicos, como la posición y las propiedades de los átomos.
- Adición de posiciones específicas: Opcionalmente, se pueden agregar posiciones específicas de interés en el *template*, como los puntos centrales de anillos de carbono u otros elementos clave.
- Establecimiento de reglas de eliminación atómica: Se definen reglas que determinan qué átomos deben ser eliminados durante el proceso de permutación para lograr la estequiometría buscada. Estas reglas son importantes para garantizar la correcta formación de los sistemas objetivo.
- Permutación de átomos: Se realiza la permutación de átomos según las reglas establecidas. Este proceso implica intercambiar los átomos seleccionados en el paso 2 por otros tipos de átomos, siguiendo las reglas de eliminación y manteniendo la estructura y conectividad del sistema.

Estos pasos permiten generar diferentes sistemas objetivo a partir del *template* inicial, ampliando así la diversidad de estructuras que se pueden explorar en la búsqueda de clústeres moleculares o la modificación de sistemas existentes.

8.4 DISEÑO HERRAMIENTAS - FUNCIÓN DE FUKUI

Se diseñó una herramienta computacional que utiliza la función de Fukui como descriptor predictivo en un algoritmo de optimización estocástico. Contiene los siguientes requerimientos:

- **Cálculo de la Función de Fukui:** La herramienta calcula la Función de Fukui y realiza el análisis topológico de la misma. Será capaz de realizar el cálculo

utilizando tanto la aproximación de diferencias finitas como la aproximación de Koopmans.

- **Claridad de resultados:** La herramienta presenta los resultados de manera clara y comprensible para el usuario. Los valores de la Función de Fukui y otros descriptores relevantes serán entregados en un formato fácil de leer. Además, se ofrecerá la opción de guardar los resultados en formato de lenguaje máquina si es necesario.
- **Valores del potencial químico:** La herramienta proporciona al usuario los valores del potencial químico para cada fragmento molecular. Permitiendo identificar los fragmentos nucleofílicos y electrofílicos de manera precisa.
- **Implementación del algoritmo KICK:** Desarrollo de un software similar a KICK que genere ensambles aleatorios entre fragmentos moleculares. Los fragmentos moleculares se seleccionarán considerando la información de la Función de Fukui. La caja de interacción KICK se ajustará automáticamente según los fragmentos utilizados. Además, el programa será capaz de detectar impedimentos estéricos en los sistemas formados y garantizar una interacción efectiva entre los atractores de la Función de Fukui.
- **Calcular la integral Coulombica (J):** Entre las funciones de Fukui electrofílica y nucleofílica del ensamble formado. Esta información nos permite evaluar la calidad del ensamble.
- **Generación de múltiples ensambles:** Realizar múltiples ensambles con el fin de realizar una buena exploración de la SEP.
- **Jerarquización de resultados:** Los resultados obtenidos se jerarquizan según los valores de J . Permitiendo identificar los ensambles que presenten una mayor interacción entre los fragmentos y, supuestamente, tengan un mayor potencial para ser mínimos energéticos competitivos.
- **Impresión de resultados:** Los mejores resultados obtenidos se imprimen en formato de coordenadas y se proporcionarán como archivos de entrada para programas de optimización adicionales.

El Diagrama 1 (Anexo 1) muestra el diseño conceptual de la herramienta con claras entradas y salidas de información según lo explicado anteriormente. El Esquema 8.1 (izquierda) ilustra en detalle los diferentes pasos que integran el método descrito. En el ejemplo representado se tienen como fragmentos los clústeres BH_3 y NH_3 que, al reaccionar forman H_3NBH_3 .

En una primera etapa (I) se utiliza el potencial químico para definir el agente nucleofílico y electrofílico. En el siguiente paso (II) la función de Fukui es calculada para cada fragmento siguiendo las Ecuaciones 2.12 y 2.13. Además, se realiza el análisis topológico correspondiente para asignar los valores condensados de la función de Fukui a los atractores, los cuales tienen asignadas posiciones espaciales específicas.

En el siguiente paso (III), se muestra el comportamiento estocástico del algoritmo KICK, el cual coloca aleatoriamente cada fragmento en una caja de interacción. Tanto las coordenadas atómicas como las posiciones de los atractores se mantienen para preservar su conectividad integral. Un espacio cúbico de interacción se crea automáticamente cuyo tamaño escalará según la suma de los fragmentos dispuestos. Nuestra implementación de KICK genera una rotación aleatoria de cada fragmento alrededor de su centro de masa y su posterior ubicación en el origen de coordenadas, donde se aplica una "patada" (KICK) siguiendo un vector de desplazamiento aleatorio que asegura que el fragmento permanezca dentro de la caja de interacción. Un rápido escaneo de las ubicaciones resultantes permite verificar posibles impedimentos estéricos. Luego se realiza el cálculo de J para el ensamble aleatorio formado y se agrega el valor obtenido a la base de datos final.

El proceso se repite hasta obtener el número deseado de ensamblajes (IV), y se guarda el valor de J calculado para cada uno de ellos. Finalmente, los ensamblajes con mejores valores de J son seleccionados como posibles candidatos a mínimos energéticos y son sometidos a un proceso de optimización utilizando algún método por gradiente.

Este enfoque permite evaluar múltiples ensam3bles aleatorios y jerarquizarlos según los valores de la integral Coulombica, lo que proporciona una guía efectiva para la identificación de posibles mínimos energéticos en los clústeres estudiados.

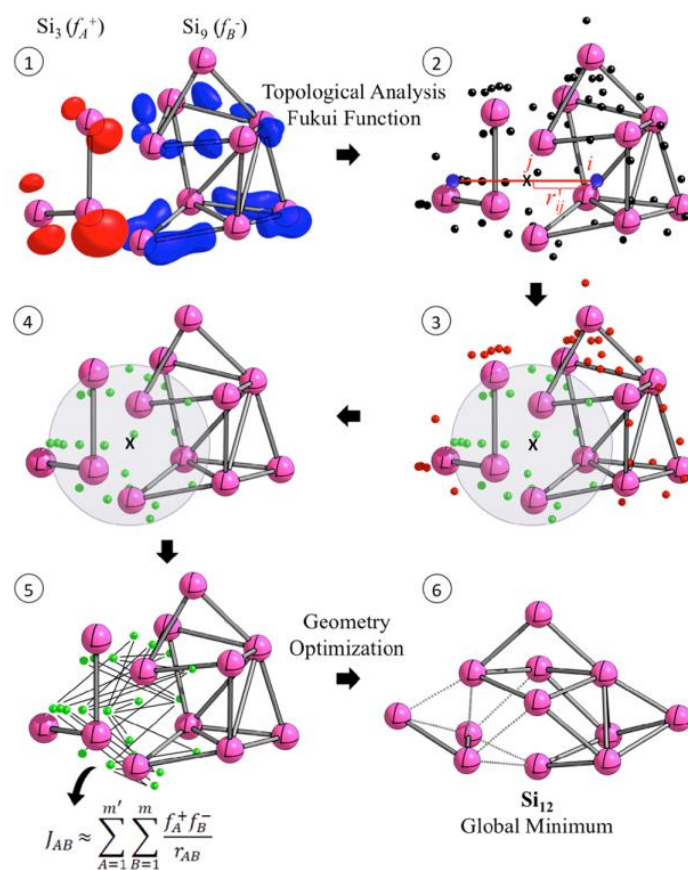
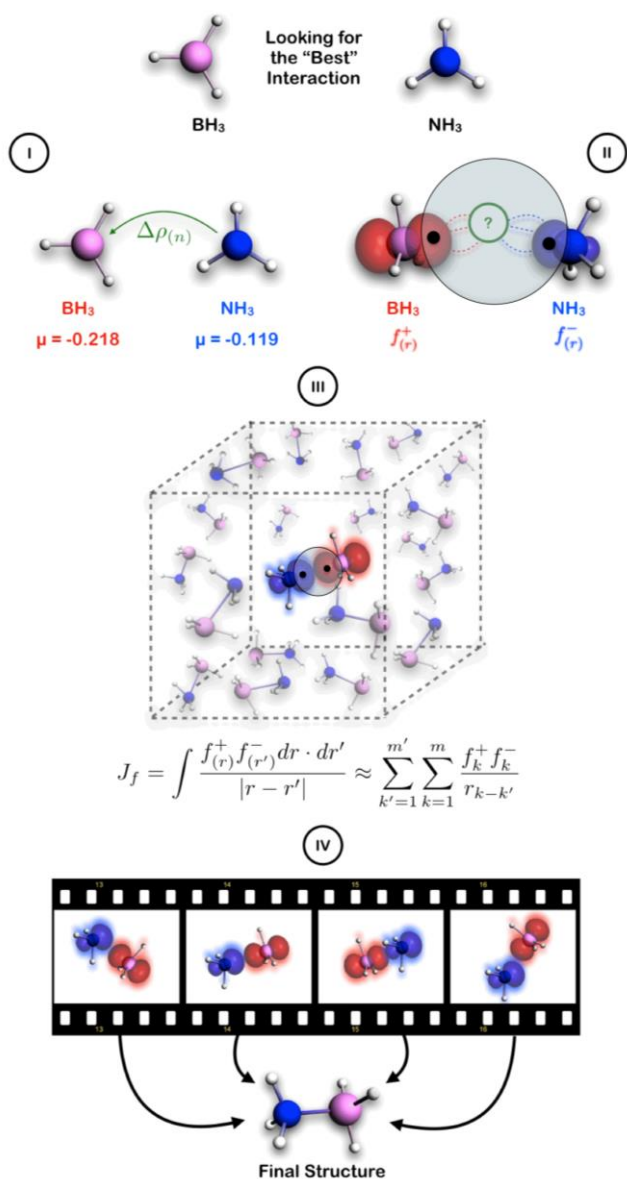
Con el objetivo de garantizar la mejor interacción entre fragmentos y optimizar el solapamiento de los atractores nucleofílicos y electrofílicos, se debe asegurar que sólo aquellos atractores que realmente interactúan en el ensamble sean considerados. Para abordar esta cuestión, hemos desarrollado una estrategia específica, que se ilustra en el Esquema 8.1 (Derecha).

En esta aproximación, utilizamos una "esfera de interacción" que se encuentra centrada entre los centros de masa de cada fragmento en el ensamble. Esta esfera define un espacio limitado donde se consideran los atractores que interactúan en el cálculo de J , dividiendo así los atractores entre aquellos que participan en la interacción (interactuantes en J) y aquellos que no lo hacen (no interactuantes en J).

Esta estrategia nos permite enfocarnos en los atractores que realmente contribuyen a la interacción entre fragmentos, evitando considerar aquellos que están obstruidos por la misma molécula y que no participarían en una reacción química normal. Al restringir el análisis a los atractores dentro de la esfera de interacción, podemos obtener una evaluación más precisa y significativa de la integral Coulombica (J) y, por lo tanto, de la calidad de la interacción entre los fragmentos en el ensamble.

Esta aproximación nos brinda una mayor selectividad en la evaluación de la interacción entre fragmentos, permitiendo una mejor identificación de los ensambles con los mejores valores de J y, en última instancia, de los posibles mínimos energéticos en el sistema estudiado.

El software contiene todas las especificaciones hasta ahora discutidas, permite una fácil recolección de información y rápida obtención de resultados de candidatos a mínimos energéticos competitivos. Por su implementación estocástica y guiada mediante la función de Fukui el programa es llamado "Kick-Fukui".



Esquema 8.1. Izquierda; Representación esquemática del procedimiento de exploración de la SEP. Derecha; Selección de basines reactivos en el cálculo de J .

8.5 DISEÑO HERRAMIENTAS - INTEGRAL NCI

Se diseñó una herramienta computacional que utiliza la integral del NCI como descriptor predictivo en un algoritmo de optimización estocástico. Contiene los siguientes requerimientos:

- Automatización de la construcción de los *inputs* necesarios para ejecutar el programa NCILOT4. Permite realizar cálculos de la integral del NCI de manera eficiente.
- Extracción de información relevante de la salida del programa NCILOT4. Se implementa una funcionalidad para extraer los valores de la integral del NCI y cualquier otra información importante para su posterior análisis.
- Implementación de un software tipo KICK para la generación de sistemas moleculares aleatorios. Los sistemas moleculares están compuestos por sus coordenadas cartesianas y la caja de interacción se escalará según el tamaño del sistema final. El programa verifica la presencia de posibles impedimentos estéricos y permite descartar sistemas que no cumplan con los parámetros establecidos.
- Construcción de archivos de entrada para programas de optimización. Se desarrolló una funcionalidad que genera los archivos de entrada necesarios para programas de optimización, utilizando la información de las estructuras moleculares generadas anteriormente.
- Ejecución de programas de optimización. La herramienta envía los programas de optimización seleccionados para buscar las estructuras moleculares optimizadas y recopilará la información correspondiente.
- Recopilación de las estructuras optimizadas y su correspondiente información en una base de datos. Se implementa una base de datos donde se almacenan las estructuras moleculares optimizadas junto con la información relevante obtenida de los programas de optimización.
- Implementación del algoritmo de enjambre de partículas. El algoritmo utiliza la integral del NCI como descriptor para guiar la búsqueda de estructuras moleculares óptimas. Los mejores valores de la integral del NCI (*Alphas*) se asocian con las mejores energías de los sistemas. Estos valores *Alphas* se guardan para su posterior análisis. Además, el programa busca nuevas estructuras aleatorias utilizando los valores *Alphas* como guía.

- Identificación de la convergencia del programa cuando no se mejoren las energías optimizadas. Se establece un criterio de convergencia para determinar cuándo el programa ha alcanzado una solución óptima y dejará de buscar nuevas estructuras.

El Diagrama 2 (Anexo 1) muestra el diseño general de la herramienta confeccionada. El Esquema 8.2 detalla el algoritmo PSO con NCI-i como guía en la optimización de clústeres moleculares.

En el primer paso de la herramienta, se lleva a cabo la generación de la población inicial, también conocida como población de origen. Con este fin, se utiliza el algoritmo KICK. Como se ha mencionado anteriormente, el algoritmo KICK genera sistemas moleculares aleatorios dentro de un espacio de interacción confinado. El tamaño de este espacio de interacción se determina mediante la suma de cada fragmento que compone el clúster molecular.

Una vez que se ha alcanzado el tamaño deseado para la población de origen, esta es enviada al software NCILOT4 para calcular el índice NCI de cada individuo en la población. Se mantiene una base de datos actualizada con la información relevante, donde los individuos se jerarquizan en función de sus valores de NCI-i.

En el Esquema 8.2, se utiliza el ejemplo de un pentámero de agua. Supongamos que se tiene una población de $1000N$ ($N = N^\circ$ moléculas de H_2O), lo que corresponde a 5000 estructuras aleatorias de clústeres de agua. Sin embargo, solo se seleccionarán 50 estructuras para ser optimizadas, lo que representa el 1% de la población. Para seleccionar estas estructuras, se elige una muestra representativa de los valores de NCI-i, tomando sectores equitativos entre el valor máximo y mínimo de NCI-i que se encuentran en la base de datos. Esta selección garantiza una diversidad en las estructuras a optimizar y permite evaluar un rango amplio de reactividad y energías asociadas.

En la parte inferior del esquema se muestra el proceso iterativo de búsqueda. Los 50 isómeros optimizados son ordenados según su energía, lo que permite obtener

las primeras estructuras *Alfas* que corresponden a aquellas con menor energía. Los valores de NCI-i asociados a estas estructuras, en teoría, representan el valor "correcto" de menor energía. Con base en estos valores, se forma una segunda población que consiste en estructuras con valores de NCI-i "cercaños" a los valores *Alfas* obtenidos.

Se define un *delta* ($\Delta = \mp 0.2\text{GAP}$, $\text{GAP} = \text{NCI}_{\text{max}} - \text{NCI}_{\text{min}}$) para limitar la cercanía a los valores *alfas*. Iterativamente, en cada ciclo, nuevamente se escogen 10N estructuras, se optimizan y se reevalúa los nuevo valores *Alfas* asociados a nuevas estructuras de menor energía según NCI. El proceso iterativo se repite con esta nueva población, generando y optimizando estructuras adicionales hasta que se alcance una convergencia, es decir, hasta que no se observe una mejora significativa en las energías optimizadas. Este enfoque de búsqueda basado en el descriptor NCI-i permite encontrar estructuras con energías más bajas y, por lo tanto, más estables en el sistema molecular en estudio.

Se evaluaron diferentes valores de ρ^n , ($n = 1.0, 1.5, 2.0, 2.5, 3.0, 4/3, 5/3$) según la ecuación 2.15 buscando como afectaría la correlación con la energía. Para este ejercicio se utilizaron los clústeres de agua descritos en el capítulo V. No se observaron cambios significativos en la correlación, por lo que se decidió utilizar como valor constante de $n = 2.5$ para todos los sistemas evaluados (Figura 8.2). Además, en el trabajo de Contreras se realizó un análisis similar con el set de datos S66^{187, 188} llegándose a la misma conclusión y utilizando el mismo valor de n .

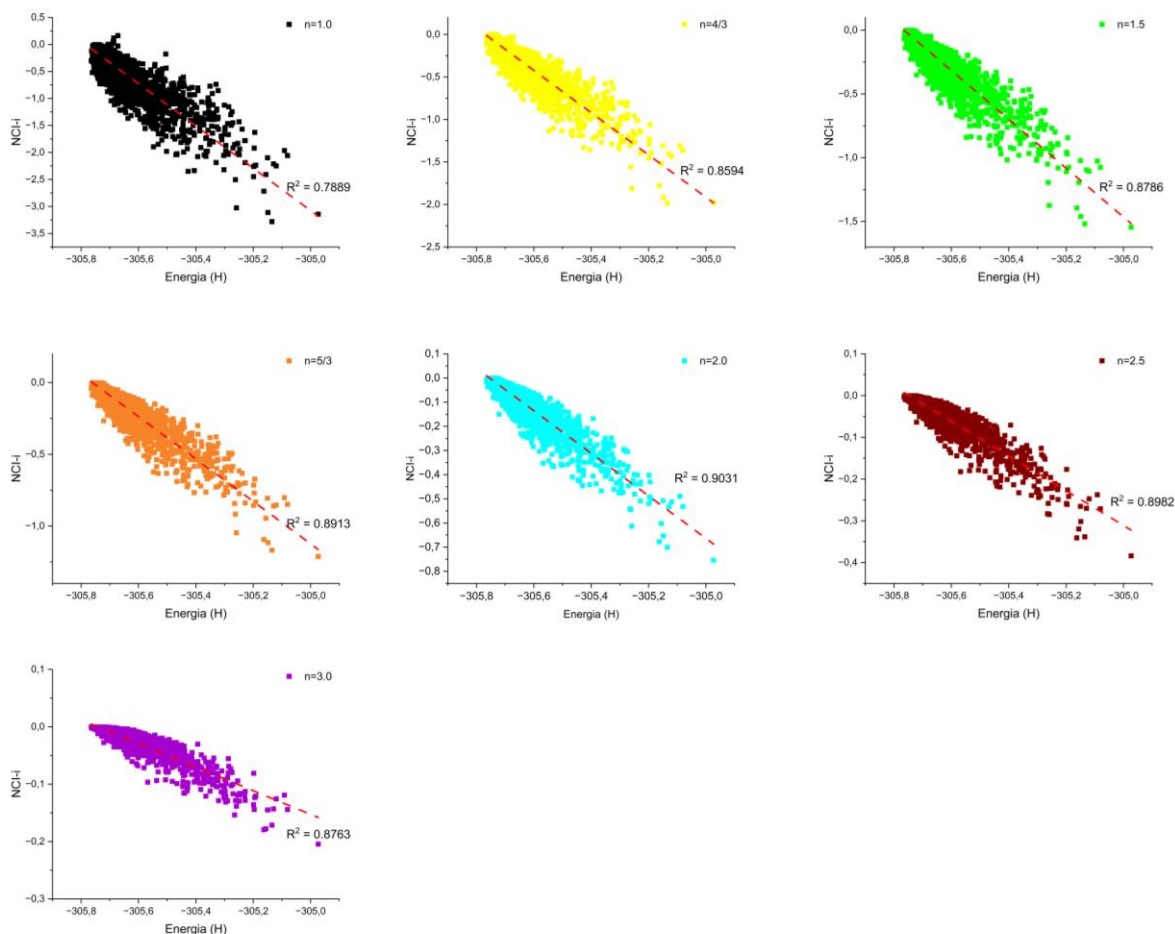
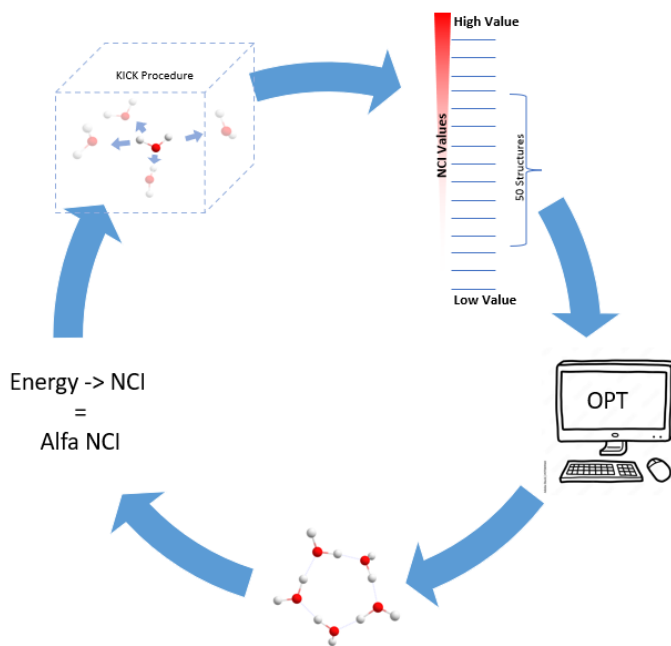
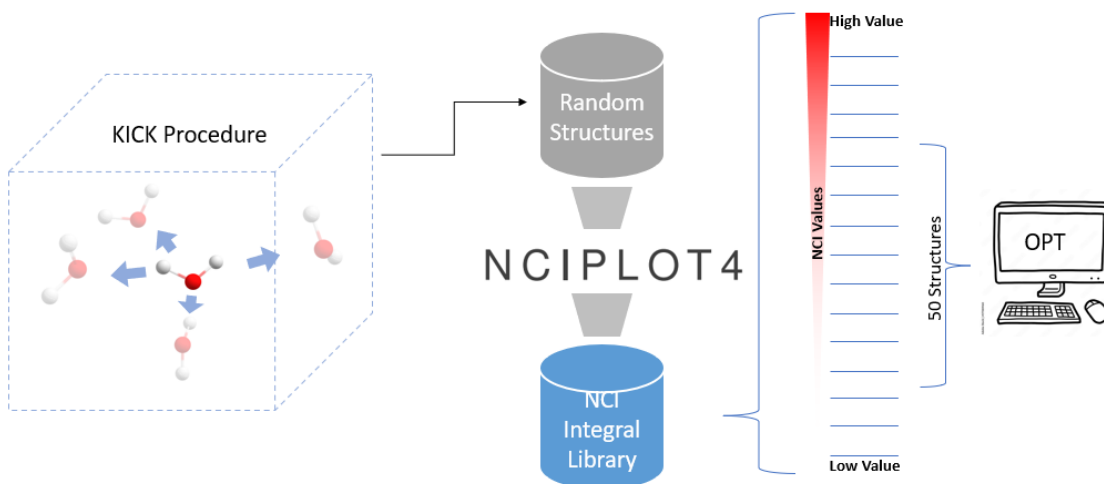


Figura 8.2. Correlación entre valores de NCI (Eje Y) y energía *single-point* (Eje X) del tetrámero de agua, en azul son estructuras aleatorias obtenidas por el procedimiento KICK, en naranja el mínimo global reportado en literatura.

Finalmente, se tomó en consideración si el signo de λ_2 afectaría la correlación NCI-i y la energía de los clústeres moleculares. Se compararon dos casos: uno en el que se utilizó el espacio ρ^n independientemente del valor de λ_2 (representado en la Figura 5.1), y otro en el que se tomó en consideración el signo de λ_2 (representado en la Figura 5.2). Los resultados muestran que la correlación no se ve afectada por la elección del espacio de integración.

Por lo tanto, se decidió continuar utilizando la integración del NCI en todo el volumen Ω_{NCI} , sin tener en cuenta el signo de λ_2 . Esta elección se realizó debido a la simplicidad y para trabajar siempre con valores positivos en la integración del NCI.



Esquema 8.2 Protocolo de trabajo propuesto donde los clústeres atómicos son formados mediante un procedimiento KICK, son evaluados con NCIPILOT 4 para luego escoger y enviar a optimizar según los valores de la integral NCI. Abajo se muestra el diagrama iterativo donde en cada ciclo los isómeros de menor energía son evaluados para establecer los valores *alfas* de NCI.

8.6 DISEÑO HERRAMIENTAS – COMPILATORIO *LOOKER*

El compilatorio computacional con las herramientas desarrolladas a lo largo de esta tesis, se quiere dejar diseñado el compilatorio en miras de un posible desarrollo a futuro que puede estar atada a la presente tesis.

- Soporte para varias herramientas de optimización de la química computacional: El compilatorio debe ser compatible con diferentes programas de optimización utilizados en la química computacional, permitiendo su integración y utilización en la exploración del SEP.
- Integración con diferentes arquitecturas de sistemas operativos: El software compilatorio debe ser compatible con distintas arquitecturas de sistemas operativos, como Windows, macOS y Linux, para asegurar su disponibilidad en diferentes entornos de trabajo.
- Integración con MatrixAutomaton: El compilatorio debe permitir la integración y el uso conjunto con la herramienta MatrixAutomaton, aprovechando su capacidad de exploración del SEP y sus mejoras implementadas.
- Integración con el algoritmo de enjambre de partículas NCI: El software compilatorio debe integrar el algoritmo de enjambre de partículas basado en el descriptor NCI, permitiendo su utilización como una estrategia de búsqueda adicional en la exploración del SEP.
- Capacidad de elección del usuario para escoger el mejor algoritmo: El compilatorio debe proporcionar al usuario la capacidad de seleccionar y configurar los algoritmos de exploración del SEP a utilizar, permitiéndoles adaptar el proceso según sus necesidades y preferencias.
- Integración con bases de datos para la obtención de estructuras modelo: El software compilatorio debe tener la capacidad de conectarse a bases de datos que contengan estructuras modelo relevantes, permitiendo su utilización como punto de partida en la exploración del SEP.

- Entrega clara de resultados finales: El compilatorio debe presentar de manera clara y comprensible los resultados obtenidos en la exploración del SEP, facilitando la interpretación y el análisis de estos.

En resumen, el compilatorio propuesto combina varias herramientas desarrolladas en esta tesis, como MatrixAutomaton, algoritmos aleatorios tipo KICK y guiados tipo permutación, y el algoritmo de enjambre de partículas NCI, para proporcionar una búsqueda eficiente y precisa en la exploración de la superficie de energía potencial de sistemas moleculares. Además, cumple con los requerimientos básicos mencionados para garantizar su funcionalidad y utilidad en la comunidad científica.

CAPÍTULO IX - CONCLUSIONES GENERALES

Durante la etapa de esta tesis, se publicaron un total de 16 artículos relacionados con el desarrollo de los objetivos de esta, que evidencian las ventajas y aplicaciones de los algoritmos de búsqueda y optimización de la superficie de energía potencial (SEP). Además, se publicaron 5 artículos en colaboración con otros proyectos, donde se aplicaron estas herramientas de búsqueda. Estos trabajos abordaron:

- El estudio de descriptores de reactividad química basados en DFT conceptual y su capacidad para guiar una exploración de la SEP.
- La propuesta y desarrollo de múltiples estrategias para clústeres atómicos de interés tanto para el autor como para el grupo de investigación.
- Actualizaciones al software AUTOMATON, con la implementación de MatrixAutomaton, que redujo la complejidad y aumentó la velocidad para crear poblaciones iniciales.
- La creación del software TemplatePermuta, que utiliza un sistema modelo para la construcción dirigida de clústeres atómicos mediante cambios atómicos.
- La implementación de dos softwares estocásticos dirigidos por descriptores químicos: Kick-Fukui y el programa NCI-i.

Además, se evaluó el nuevo descriptor de integración del NCI en un posible uso predictivo de clústeres moleculares mediante un algoritmo de enjambre (PSO) rápido y eficaz.

A partir de estos estudios, se pueden destacar las siguientes conclusiones:

1. La función de Fukui demostró ser un descriptor predictivo eficaz en clústeres atómicos, proporcionando información crucial sobre sus propiedades y reactividad.

2. La función de Fukui supera a una búsqueda aleatoria, resaltando la importancia de herramientas de optimización guiadas por descriptores químicos.
3. Se identificó una correlación entre los valores de integración del NCI y la energía de los clústeres moleculares, sugiriendo su potencial como descriptor predictivo.
4. Aunque el algoritmo PSO con NCI-i es prometedor, se evidenció la necesidad de metodologías de exploración más avanzadas.
5. MatrixAutomaton ha mejorado significativamente la exploración de la SEP en comparación con AUTOMATON.
6. MatrixAutomaton mostró flexibilidad en la exploración de sistemas moleculares complejos, pero enfrenta desafíos con clústeres de gran tamaño.
7. La integración de intuición química y técnicas avanzadas, como la permutación atómica guiada, enriquece la búsqueda en la SEP.

Es esencial señalar que el objetivo principal de esta tesis no fue diseñar una metodología infalible para la exploración de la SEP, lo que es imposible por definición. Las técnicas propuestas buscan acercar al usuario a mínimos energéticos óptimos con costos computacionales reducidos, utilizando una población pequeña pero efectiva, y destacando la importancia de la integración con otros descriptores o intuición química.

PUBLICACIONES

1. Yañez, O., **Inostroza, D.**, Usuga-Acevedo, B., Vásquez-Espinal, A., Pino-Rios, R., Tabilo-Sepulveda, M., ... & Tiznado, W. (2020). Evaluation of restricted probabilistic cellular automata on the exploration of the potential energy surface of Be 6 B 11-. *Theoretical Chemistry Accounts*, 139(3), 1-8.
2. Manrique-de-la-Cuba, M. F., Leyva-Parra, L., **Inostroza, D.**, Gomez, B., Vásquez-Espinal, A., Garza, J., ... & Tiznado, W. (2021). Li8Si8, Li10Si9, and Li12Si10: Assemblies of Lithium-Silicon Aromatic Units. *ChemPhysChem*.
3. Leyva-Parra, L., Diego, L., Yañez, O., **Inostroza, D.**, Barroso, J., Vásquez-Espinal, A., ... & Tiznado, W. (2021). Planar hexacoordinate carbons: half covalent, half ionic. *Angewandte Chemie International Edition*, 60(16), 8700-8704.
4. Yañez, O., Báez-Grez, R., **Inostroza, D.**, Pino-Rios, R., Rabanal-León, W. A., Contreras-García, J., ... & Tiznado, W. (2021). Kick-Fukui: A Fukui Function-Guided Method for Molecular Structure Prediction. *Journal of Chemical Information and Modeling*, 61(8), 3955-3963.
5. Pino-Rios, R., **Inostroza, D.**, & Tiznado, W. (2021). Neither too Classic nor too Exotic: One-Electron Na· B Bond in NaBH3- Cluster. *Angewandte Chemie International Edition*, 60(23), 12747-12753.
6. Leyva-Parra, L., Diego, L., **Inostroza, D.**, Yañez, O., Pumachagua-Huertas, R., Barroso, J., ... & Tiznado, W. (2021). Planar Hypercoordinate Carbons in Alkali Metal Decorated CE32- and CE22- Dianions. *Chemistry-A European Journal*.
7. Thimmakondur, V. S., Sinjari, A., **Inostroza, D.**, Vairaprakash, P., Thirumoorthy, K., Roy, S., Anoop, A., & Tiznado, W. (2022). Why an integrated approach between search algorithms and chemical intuition is necessary? *Physical Chemistry Chemical Physics*, 24(19), 11680-11686.
8. Leyva-Parra, L., **Inostroza, D.**, Yañez, O., Cruz, J. C., Garza, J., García, V., & Tiznado, W. (2022). Persistent Planar Tetracoordinate Carbon in Global Minima Structures of Silicon-Carbon Clusters. *Atoms*, 10(1), 27.
9. Solar-Encinas, J., Leyva-Parra, L., Yañez, O., **Inostroza, D.**, Barrios-Llacuachaqui, J. R., Vásquez-Espinal, A., Orellana, W., & Tiznado, W. (2022). Bowl-shaped CuB12-Cluster. A viable Global Minimum with Twofold Aromaticity. *Chemphyschem: a European Journal of Chemical Physics and Physical Chemistry*.
10. **Inostroza, D.**, Leyva-Parra, L., Yañez, O., Cruz, J. C., Garza, J., García, V., Thimmakondur, V. S., Ceron, M. L., & Tiznado, W. (2022). Si6C18: A bispentalene derivative with two planar tetracoordinate carbons. *International Journal of Quantum Chemistry*, 123(1), e27008.
11. Yañez, O., **Inostroza, D.**, Leyva-Parra, L., Solar-Encinas, J., Cruz, J. C., Garza, J., Vásquez-Espinal, A., Pino-Rios, R., Orellana, W., & Tiznado, W. (2022). Si 5-pentagonal rings and Y-shaped Si 4 building blocks in Li 32 Si 18 system: similarities with the crystalline Zintl phase Li 12 Si 7. *Molecular Systems Design & Engineering*.
12. **Inostroza, D.**, Leyva-Parra, L., Vásquez-Espinal, A., Contreras-García, J., Cui, Z.-H., Pan, S., Thimmakondur, V. S., & Tiznado, W. (2022). E 6 C 15 (E= Si-Pb): polycyclic aromatic compounds with three planar tetracoordinate carbons. *Chemical Communications*, 58(94), 13075-13078.
13. Solar-Encinas, J., Vásquez-Espinal, A., Leyva-Parra, L., Yañez, O., **Inostroza, D.**, Valenzuela, M. L., Orellana, W., & Tiznado, W. (2023). Planar Elongated B12 Structure in M3B12 Clusters (M = Cu-Au). *Molecules*, 28(1), 236.
14. **Inostroza, D.**, Leyva-Parra, L., Yañez, O., Solar-Encinas, J., Vásquez-Espinal, A., Valenzuela, M. L., & Tiznado, W. (2023). Searching for Systems with Planar Hexacoordinate Carbons. *Atoms*, 11(3), 56.
15. **Inostroza, D.**, Vásquez-Espinal, A., Leyva-Parra, L., García-Argote, W., Cerón, M.

- L., Yañez, O., & Tiznado, W. (2023). Revisiting the potential-energy surface of $C_nBe_3n+2H_2n+22+$ ($n = 2-4$) clusters: are planar pentacoordinate carbon structures the global minima? *Physical Chemistry Chemical Physics*, 25(30), 20235-20240.
16. **Inostroza, D.**, Leyva-Parra, L., Yañez, O., Cooksy, A. L., Thimmakonda, V. S., & Tiznado, W. (2023). Exploring the Potential Energy Surface of Medium-Sized Aromatic Polycyclic Systems with Embedded Planar Tetracoordinate Carbons: A Guided Approach. *Chemistry*, 5(3), 1535-1545.

COLABORACIONES

1. Báez-Grez, R., **Inostroza, D.**, García, V., Vásquez-Espinal, A., Donald, K. J., & Tiznado, W. (2020). Aromatic ouroboroi: heterocycles involving a σ -donor-acceptor bond and $4n+2$ π -electrons. *Physical Chemistry Chemical Physics*, 22(4), 1826-1832.
2. Islas, R., **Inostroza, D.**, Arias-Olivares, D., Zúñiga-Gutiérrez, B., Poater, J., & Solà, M. (2020). Analysis of the electronic delocalization in some isoelectronic analogues of B 12 doped with beryllium and/or carbon. *Physical Chemistry Chemical Physics*, 22(21), 12245-12259.
3. **Inostroza, D.**, García, V., Yañez, O., Torres-Vega, J. J., Vásquez-Espinal, A., Pino-Rios, R., Báez-Grez, R., & Tiznado, W. (2021). On the NICS limitations to predict local and global current pathways in polycyclic systems. *New Journal of Chemistry*, 45(18), 8345-8351.
4. Pino-Rios, R., Yañez, O., **Inostroza, D.**, Báez-Grez, R., Cárdenas, C., & Tiznado, W. (2023). Structure prediction using reactivity descriptors. *Chemical Reactivity*, 449-462.
5. R. Báez-Grez, **D. Inostroza**, A. Vásquez-Espinal, R. Islas and R. Pino-Rios, *RSC Advances*, 2023, **13**, 24499-24504.

CÓDIGOS

El código principal del software MatrixAutomaton se encuentra en:

- <https://github.com/Rufox/MatrixAutomaton/tree/py3>

El código de KICK-Fukui se encuentra en:

- <https://github.com/HumanOsv/Kick-Fukui>

REFERENCIAS

1. L. Leyva-Parra, L. Diego, O. Yañez, D. Inostroza, J. Barroso, A. Vásquez-Espinal, G. Merino and W. Tiznado, *Angewandte Chemie International Edition*, 2021, **60**, 8700-8704.
2. A. Dowling, *Uncertainties, A Report by The Royal Society & The Royal Academy of Engineering, London*, 2004.
3. A. Castleman and K. Bowen, *The Journal of Physical Chemistry*, 1996, **100**, 12911-12944.
4. C. P. Poole Jr and F. J. Owens, *Introduction to nanotechnology*, John Wiley & Sons 2003.
5. K. L. Wang, *Journal of Nanoscience and Nanotechnology*, 2002, **2**, 235-266.
6. G. M. Whitesides, *Small*, 2005, **1**, 172-179.
7. F. X. Redl, K. S. Cho, C. B. Murray and S. O'Brien, *Nature*, 2003, **423**, 968-971.
8. R. F. Service, *Science*, 2002, **298**, 2322-2323.
9. M. Carella, S. Volinia and P. Gasparini, *Journal of Nephrology*, 2003, **16**, 597-602.
10. R. J. Martín-Palma, J. Martínez-Duart and F. Agulló-Rueda, *Nanotechnology for microelectronics and optoelectronics*, Elsevier 2006.
11. F. Baletto and R. Ferrando, *Reviews of modern physics*, 2005, **77**, 371.
12. S. N. Khanna and P. Jena, *Phys Rev B Condens Matter*, 1995, **51**, 13705-13716.
13. X. Li, H. F. Zhang, L. S. Wang, G. D. Geske and A. I. Boldyrev, *Angewandte Chemie*, 2000, **112**, 3776-3778.
14. Y. Pei, W. An, K. Ito, P. V. R. Schleyer and X. C. Zeng, *J Am Chem Soc*, 2008, **130**, 10394-10400.
15. Y.-B. Wu, H.-G. Lu, S.-D. Li and Z.-X. Wang, *The Journal of Physical Chemistry A*, 2009, **113**, 3395-3402.
16. R. Grande-Aztatzi, J. L. Cabellos, R. Islas, I. Infante, J. M. Mercero, A. Restrepo and G. Merino, *Physical Chemistry Chemical Physics*, 2015, **17**, 4620-4624.
17. Z.-h. Cui, V. Vassilev-Galindo, J. L. Cabellos, E. Osorio, M. Orozco, S. Pan, Y.-h. Ding and G. Merino, *Chemical Communications*, 2017, **53**, 138-141.
18. J.-J. Li, Y. Mu, X. Tian, C. Yuan, Y.-B. Wu, Q. Wang, D. Li, Z.-X. Wang and S.-D. Li, *Journal of Materials Chemistry C*, 2017, **5**, 408-414.
19. P. Jena and Q. Sun, *Chemical reviews*, 2018, **118**, 5755-5870.
20. P. Jena, S. N. Khanna and B. Rao, *Physics and chemistry of finite systems: from clusters to crystals*, Springer Science & Business Media 2013.
21. C. J. Cramer, *Essentials of computational chemistry: theories and models*, John Wiley & Sons 2013.
22. V. Bondybey and J. English, *The Journal of Chemical Physics*, 1982, **76**, 2165-2170.
23. W. A. De Heer, *Reviews of Modern Physics*, 1993, **65**, 611.
24. Y. Zhang, L. Shuiqing, S. Deng, Q. Yao and D. T. Stephen, *Journal of Aerosol Science*, 2012, **44**, 71-82.
25. H. Scheibel and J. Porstendo, *Journal of Aerosol Science*, 1983, **14**, 113-126.
26. A. I. Boldyrev and L.-S. Wang, *Chemical reviews*, 2005, **105**, 3716-3757.
27. B. J. Alder and T. E. Wainwright, *The Journal of Chemical Physics*, 1959, **31**, 459-466.
28. B. J. Alder and T. E. Wainwright, *The Journal of Chemical Physics*, 1960, **33**, 1439-1451.
29. B. Alder, *The Journal of chemical physics*, 1964, **40**, 2724-2730.
30. M. A. Addicoat and G. F. Metha, *Journal of computational chemistry*, 2009, **30**, 57-64.
31. P. P. Bera, K. W. Sattelmeyer, M. Saunders, H. F. Schaefer and P. v. R. Schleyer, *The Journal of Physical Chemistry A*, 2006, **110**, 4287-4290.
32. M. Saunders, *J Am Chem Soc*, 1987, **109**, 3150-3152.
33. M. Saunders, *Journal of computational chemistry*, 2004, **25**, 621-626.

34. M. D. Wolf and U. Landman, *The Journal of Physical Chemistry A*, 1998, **102**, 6129-6137.
35. O. Yanez, R. Baez-Grez, D. Inostroza, W. A. Rabanal-Leon, R. Pino-Rios, J. Garza and W. Tiznado, *J Chem Theory Comput*, 2019, **15**, 1463-1475.
36. D. M. Deaven and K.-M. Ho, *Physical review letters*, 1995, **75**, 288.
37. L. Marim, M. Lemes and A. Dal Pino Jr, *Physical Review A*, 2003, **67**, 033203.
38. R. Hodgson, Particle swarm optimization applied to the atomic cluster optimization problem, 2002.
39. Q. Tong, J. Lv, P. Gao and Y. Wang, *Chinese Physics B*, 2019, **28**, 106105.
40. M. Dorigo and G. Di Caro, Ant colony optimization: a new meta-heuristic, Washington, DC, 1999.
41. E. Elbeltagi, T. Hegazy and D. Grierson, *Advanced engineering informatics*, 2005, **19**, 43-53.
42. C. Zacharias, M. Lemes and A. Dal Pino Jr, *Journal of Molecular Structure: THEOCHEM*, 1998, **430**, 29-39.
43. D. Karaboga, *scholarpedia*, 2010, **5**, 6915.
44. W. Gao, S. Liu and L. Huang, *Journal of Computational and Applied Mathematics*, 2012, **236**, 2741-2753.
45. P. Muller, *Pure and Applied Chemistry*, 1994, **66**, 1077-1184.
46. J. E. Jones, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 1924, **106**, 441-462.
47. J. E. Lennard-Jones, *Proceedings of the Physical Society*, 1931, **43**, 461.
48. E. Schrödinger, *Phys Rev*, 1926, **28**, 1049-1070.
49. M. Born and J. Oppenheimer, *Annalen der Physik*, 1927, **389**, 457-484.
50. V. Fock, *Zeitschrift für Physik*, 1930, **61**, 126-148.
51. C. Roothaan, *Reviews of modern physics*, 1960, **32**, 179.
52. J. Binkley, J. Pople and P. Dobosh, *Molecular Physics*, 1974, **28**, 1423-1429.
53. W. J. Hehre, *Accounts of Chemical Research*, 1976, **9**, 399-406.
54. J. Huheey, E. Keiter and R. Keiter, Harper Collins, New York 1993.
55. C. Møller and M. S. Plesset, *Physical review*, 1934, **46**, 618.
56. R. Krishnan and J. A. Pople, *International Journal of Quantum Chemistry*, 1978, **14**, 91-100.
57. M. Head-Gordon, J. A. Pople and M. J. Frisch, *Chemical physics letters*, 1988, **153**, 503-506.
58. G. D. Purvis III and R. J. Bartlett, *The Journal of Chemical Physics*, 1982, **76**, 1910-1918.
59. K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chemical Physics Letters*, 1989, **157**, 479-483.
60. P. Hohenberg and W. Kohn, *Physical review*, 1964, **136**, B864.
61. R. G. Parr and W. Yang, *Journal of the American Chemical Society*, 1984, **106**, 4049-4050.
62. W. Kohn and L. J. Sham, *Physical review*, 1965, **140**, A1133.
63. V. Sahni, K.-P. Bohnen and M. K. Harbola, *Physical Review A*, 1988, **37**, 1895.
64. S. Vosko, L. Wilk and M. Nusair, *Can J Phys*, 1980, **58**, 1200-1211.
65. C. Lee, W. Yang and R. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
66. A. D. Becke, *Journal of Chemical Physics*, 1993, **98**, 1372-1377.
67. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *The Journal of physical chemistry*, 1994, **98**, 11623-11627.
68. C. Adamo and V. Barone, *The Journal of chemical physics*, 1999, **110**, 6158-6170.
69. R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules*, Oxford University Press ; Clarendon Press, New York, Oxford England, 1989.
70. R. G. Parr and W. Yang, *Annual review of physical chemistry*, 1995, **46**, 701-728.
71. H. Chermette, *Journal of computational chemistry*, 1999, **20**, 129-154.
72. P. Geerlings, F. De Proft and W. Langenaeker, *Chemical reviews*, 2003, **103**, 1793-1874.

73. P. W. Ayers, J. S. Anderson and L. J. Bartolotti, *International Journal of Quantum Chemistry*, 2005, **101**, 520-534.
74. P. Fuentealba and C. Cárdenas, in *Chemical modelling2014*, pp. 151-174.
75. K. Fukui, T. Yonezawa and H. Shingu, *The Journal of Chemical Physics*, 1952, **20**, 722-725.
76. R. G. Parr and R. G. Pearson, *Journal of the American chemical society*, 1983, **105**, 7512-7516.
77. R. G. Pearson, *Proceedings of the National Academy of Sciences*, 1986, **83**, 8440-8441.
78. P. Kolandaivel, G. Praveena and P. Selvarengan, *Journal of Chemical Sciences*, 2005, **117**, 591-598.
79. J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz Jr, *Physical Review Letters*, 1982, **49**, 1691.
80. K. Fukui, *science*, 1982, **218**, 747-754.
81. P. W. Ayers and M. Levy, *Theoretical Chemistry Accounts*, 2000, **103**, 353-360.
82. R. Pino-Rios, O. Yañez, D. Inostroza, L. Ruiz, C. Cardenas, P. Fuentealba and W. Tiznado, *Journal of Computational Chemistry*, 2017, **38**, 481-488.
83. W. Yang and W. J. Mortier, *Journal of the American Chemical Society*, 1986, **108**, 5708-5711.
84. R. Bader, Oxford1994.
85. E. Osorio, M. B. Ferraro, O. B. Oña, C. Cardenas, P. Fuentealba and W. Tiznado, *Journal of Chemical Theory and Computation*, 2011, **7**, 3995-4001.
86. O. Yañez, A. Vásquez-Espinal, D. Inostroza, L. Ruiz, R. Pino-Rios and W. Tiznado, *Journal of Computational Chemistry*, 2017, **38**, 1668-1677.
87. P. Hobza and K. Müller-Dethlefs, *Non-covalent interactions: theory and experiment*, Royal Society of Chemistry2010.
88. A. O. de la Roza and G. A. DiLabio, *Non-covalent Interactions in Quantum Chemistry and Physics: Theory and Applications*, Elsevier2017.
89. E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg and P. Hobza, *Pure and applied chemistry*, 2011, **83**, 1637-1641.
90. K. A. Dill, *Biochemistry*, 1990, **29**, 7133-7155.
91. S. K. Panigrahi and G. R. Desiraju, *Proteins: Structure, Function, and Bioinformatics*, 2007, **67**, 128-141.
92. S. V. Prigent, H. Gruppen, A. J. Visser, G. A. Van Koningsveld, G. A. De Jong and A. G. Voragen, *Journal of Agricultural and Food Chemistry*, 2003, **51**, 5088-5095.
93. A. D. Becke and K. E. Edgecombe, *The Journal of chemical physics*, 1990, **92**, 5397-5403.
94. J. P. Perdew, K. Burke and M. Ernzerhof, *Physical review letters*, 1996, **77**, 3865.
95. E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen and W. Yang, *Journal of the American Chemical Society*, 2010, **132**, 6498-6506.
96. J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan and W. Yang, *Journal of chemical theory and computation*, 2011, **7**, 625-632.
97. M. Spackman and E. Maslen, *The Journal of Physical Chemistry*, 1986, **90**, 2020-2027.
98. R. A. Boto, F. Peccati, R. Laplaza, C. Quan, A. Carbone, J.-P. Piquemal, Y. Maday and J. Contreras-García, *Journal of Chemical Theory and Computation*, 2020, **16**, 4150-4158.
99. M. Barbaty, G. Bruno and A. Genovese, *Expert Systems with Applications*, 2012, **39**, 6020-6028.
100. R. Asorey-Cacheda, A.-J. Garcia-Sanchez, F. García-Sánchez and J. García-Haro, *Journal of Network and Computer Applications*, 2017, **82**, 1-20.
101. D. R. Penas, D. Henriques, P. González, R. Doallo, J. Saez-Rodriguez and J. R. Banga, *PLoS One*, 2017, **12**, e0182186.
102. F. H. Stillinger, *Physical Review E*, 1999, **59**, 48.

103. G. Rossi and R. Ferrando, *Journal of Physics: Condensed Matter*, 2009, **21**, 084208.
104. L. T. Wille and J. Vennik, *Journal of Physics A: Mathematical and General*, 1985, **18**, L419.
105. A. B. Adib, *Journal of Physics A: Mathematical and General*, 2005, **38**, 8487.
106. S. Luke, *Essentials of metaheuristics*, Lulu Raleigh2013.
107. J. Zhang and V. A. Glezakou, *International Journal of Quantum Chemistry*, 2021, **121**, e26553.
108. N. M. Temme, *Acta Numerica*, 2007, **16**, 379-478.
109. I. Rechenberg, in *Simulationmethoden in der Medizin und Biologie*, Springer1978, pp. 83-114.
110. H. Holland John, *Ann Arbor: University of Michigan Press*, 1975.
111. A. N. Alexandrova and A. I. Boldyrev, *Journal of chemical theory and computation*, 2005, **1**, 566-580.
112. Y. Zhao, X. Chen and J. Li, *Nano Research*, 2017, **10**, 3407-3420.
113. X. Chen, Y. F. Zhao, Y. Y. Zhang and J. Li, *Journal of Computational Chemistry*, 2019, **40**, 1105-1112.
114. Z. Chen, X. Jiang, J. Li, S. Li and L. Wang, *Journal of computational chemistry*, 2013, **34**, 1046-1059.
115. J. Zhang, V.-A. Glezakou, R. Rousseau and M.-T. Nguyen, *Journal of Chemical Theory and Computation*, 2020, **16**, 3947-3958.
116. Y. Wang, J. Lv, L. Zhu and Y. Ma, *Physical Review B*, 2010, **82**, 094116.
117. J. Von Neumann and A. W. Burks, *IEEE Transactions on Neural Networks*, 1966, **5**, 3-14.
118. S. Wolfram, *A New Kind of Science*, Wolfram Media, Inc., Champaign, 2002.
119. O. Yañez, D. Inostroza, B. Usuga-Acevedo, A. Vásquez-Espinal, R. Pino-Rios, M. Tabilo-Sepulveda, J. Garza, J. Barroso, G. Merino and W. Tiznado, *Theoretical Chemistry Accounts*, 2020, **139**, 1-8.
120. D. Mendeleev, *Journal of the Russian Chemical Society*, 1871, **3**, 25–56.
121. L. Pauling, *J Am Chem Soc*, 1932, **54**, 3570-3582.
122. P. v. R. Schleyer, E. U. Wuerthwein and J. A. Pople, *J Am Chem Soc*, 1982, **104**, 5839-5841.
123. W. T. Borden, H. Iwamura and J. A. Berson, *Accounts of chemical research*, 1994, **27**, 109-116.
124. A. D. Zdetsis, *The Journal of Physical Chemistry A*, 2021, **125**, 6064-6074.
125. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, Wallingford, CT2016.
126. F. Weigend and R. Ahlrichs, *Physical Chemistry Chemical Physics*, 2005, **7**, 3297-3305.
127. C. R. Harris, K. J. Millman, S. J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg and N. J. Smith, *Nature*, 2020, **585**, 357-362.
128. C. Hill and C. Hill, *Learning Scientific Programming with Python*. Cambridge, United Kingdom: Cambridge University Press, 2015.

129. P. Fuentealba, L. Von Szentpaly, H. Preuss and H. Stoll, *Journal of Physics B: Atomic and Molecular Physics (1968-1987)*, 1985, **18**, 1287.
130. A. Bergner, M. Dolg, W. Küchle, H. Stoll and H. Preuß, *Molecular Physics*, 1993, **80**, 1431-1441.
131. J. C. Guo, L. Y. Feng, Y. J. Wang, S. Jalife, A. Vásquez-Espinal, J. L. Cabellos, S. Pan, G. Merino and H. J. Zhai, *Angewandte Chemie International Edition*, 2017, **56**, 10174-10177.
132. K. B. Wiberg, *Tetrahedron*, 1968, **24**, 1083-1096.
133. A. M. Pendás, M. Blanco and E. Francisco, *The Journal of chemical physics*, 2004, **120**, 4581-4592.
134. D. Y. Zubarev and A. I. Boldyrev, *Physical chemistry chemical physics*, 2008, **10**, 5207-5217.
135. D. Y. Zubarev and A. I. Boldyrev, *The Journal of Organic Chemistry*, 2008, **73**, 9251-9258.
136. A. E. Reed, R. B. Weinstock and F. Weinhold, *The Journal of Chemical Physics*, 1985, **83**, 735-746.
137. L. Leyva-Parra, L. Diego, D. Inostroza, O. Yañez, R. Pumachagua-Huertas, J. Barroso, A. Vásquez-Espinal, G. Merino and W. Tiznado, *Chemistry—A European Journal*, 2021, **27**, 16701-16706.
138. L. Leyva-Parra, D. Inostroza, O. Yañez, J. C. Cruz, J. Garza, V. García and W. Tiznado, *Atoms*, 2022, **10**, 27.
139. J. Jusélius, D. Sundholm and J. Gauss, *J. Chem. Phys.*, 2004, **121**, 3952-3963.
140. H. Fliegl, S. Taubert, O. Lehtonen and D. Sundholm, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20500-20518.
141. R. Pino-Rios, D. Inostroza and W. Tiznado, *Angewandte Chemie*, 2021, **133**, 12857-12863.
142. M. F. Manrique-de-la-Cuba, L. Leyva-Parra, D. Inostroza, B. Gomez, A. Vásquez-Espinal, J. Garza, O. Yañez and W. Tiznado, *ChemPhysChem*, 2021, **22**, 906-910.
143. O. Yañez, D. Inostroza, L. Leyva-Parra, J. Solar-Encinas, J. C. Cruz, J. Garza, A. Vásquez-Espinal, R. Pino-Rios, W. Orellana and W. Tiznado, *Molecular Systems Design & Engineering*, 2022.
144. S. L. Bud'ko, G. Lapertot, C. Petrovic, C. Cunningham, N. Anderson and P. Canfield, *Physical Review Letters*, 2001, **86**, 1877.
145. W. Chen, Y. Li, G. Yu, C.-Z. Li, S. B. Zhang, Z. Zhou and Z. Chen, *J Am Chem Soc*, 2010, **132**, 1699-1705.
146. H. Nakamura, *Yakugaku Zasshi: Journal of the Pharmaceutical Society of Japan*, 2013, **133**, 1297-1306.
147. Q. An and W. A. Goddard III, *The Journal of Physical Chemistry C*, 2017, **121**, 11831-11838.
148. G. Akopov, M. T. Yeung and R. B. Kaner, *Advanced Materials*, 2017, **29**, 1604506.
149. J. P. Scheifers, Y. Zhang and B. P. Fokwa, *Accounts of chemical research*, 2017, **50**, 2317-2325.
150. P. J. Robinson, G. Liu, S. Ciborowski, C. Martinez-Martinez, J. R. Chamorro, X. Zhang, T. M. McQueen, K. H. Bowen and A. N. Alexandrova, *Chemistry of Materials*, 2017, **29**, 9892-9896.
151. J. Solar-Encinas, L. Leyva-Parra, O. Yanez, D. Inostroza, J. R. Barrios-Llacuachaqui, A. Vasquez-Espinal, W. Orellana and W. Tiznado, *Chemphyschem: a European Journal of Chemical Physics and Physical Chemistry*, 2022.
152. J. Solar-Encinas, A. Vásquez-Espinal, L. Leyva-Parra, O. Yañez, D. Inostroza, M. L. Valenzuela, W. Orellana and W. Tiznado, *Molecules*, 2023, **28**, 236.
153. V. S. Thimmakonda, A. Sinjari, D. Inostroza, P. Vairaprakash, K. Thirumoorthy, S. Roy, A. Anoop and W. Tiznado, *Physical Chemistry Chemical Physics*, 2022, **24**, 11680-11686.

154. O. Yañez, A. Vásquez-Espinal, R. Pino-Rios, F. Ferraro, S. Pan, E. Osorio, G. Merino and W. Tiznado, *Chemical Communications*, 2017, **53**, 12112-12115.
155. D. Inostroza, L. Leyva-Parra, O. Yañez, J. C. Cruz, J. Garza, V. García, V. S. Thimmakondur, M. L. Ceron and W. Tiznado, *International Journal of Quantum Chemistry*, 2022, **123**, e27008.
156. D. Inostroza, L. Leyva-Parra, A. Vásquez-Espinal, J. Contreras-García, Z.-H. Cui, S. Pan, V. S. Thimmakondur and W. Tiznado, *Chemical Communications*, 2022, **58**, 13075-13078.
157. D. Inostroza, L. Leyva-Parra, O. Yañez, A. L. Cooksy, V. S. Thimmakondur and W. Tiznado, *Chemistry*, 2023, **5**, 1535-1545.
158. U. Sarkar, D. Roy, P. Chattaraj, R. Parthasarathi, J. Padmanabhan and V. Subramanian, *Journal of Chemical Sciences*, 2005, **117**, 599-612.
159. W. Yang and R. G. Parr, *Proceedings of the National Academy of Sciences*, 1985, **82**, 6723-6726.
160. T. James, D. J. Wales and J. Hernández-Rojas, *Chemical Physics Letters*, 2005, **415**, 302-307.
161. F. Li, A. Muñoz-Castro and S. C. Sevov, *Angewandte Chemie*, 2012, **124**, 8709-8712.
162. S. C. Sevov and J. M. Goicoechea, *Organometallics*, 2006, **25**, 5678-5692.
163. L. Marim, M. Lemes and A. Dal Pino Jr, *Journal of Molecular Structure: THEOCHEM*, 2003, **663**, 159-165.
164. T. Lu and F. Chen, *Journal of computational chemistry*, 2012, **33**, 580-592.
165. O. Yañez, R. Báez-Grez, D. Inostroza, R. Pino-Rios, W. A. Rabanal-León, J. Contreras-García, C. Cardenas and W. Tiznado, *Journal of Chemical Information and Modeling*, 2021, **61**, 3955-3963.
166. J. Faver and K. M. Merz Jr, *Journal of chemical theory and computation*, 2010, **6**, 548-559.
167. P. Fuentealba, P. Pérez and R. Contreras, *The Journal of Chemical Physics*, 2000, **113**, 2544-2551.
168. P. Pérez, A. Toro-Labbé and R. Contreras, *J Am Chem Soc*, 2001, **123**, 5527-5531.
169. J. Padmanabhan, R. Parthasarathi, U. Sarkar, V. Subramanian and P. Chattaraj, *Chemical physics letters*, 2004, **383**, 122-128.
170. S. Liu, *Journal of Chemical Sciences*, 2005, **117**, 477-483.
171. A. Zhong, C. Rong and S. Liu, *The Journal of Physical Chemistry A*, 2007, **111**, 3132-3136.
172. J. Oller, D. A. Saez and E. Vöhringer-Martinez, *The Journal of Physical Chemistry A*, 2020, **124**, 849-857.
173. J. Contreras-García, R. A. Boto, F. Izquierdo-Ruiz, I. Reva, T. Woller and M. Alonso, *Theoretical Chemistry Accounts*, 2016, **135**, 1-14.
174. A. Malloum, J. J. Fifen, Z. Dhaouadi, S. G. N. Engo and J. Conradie, *New Journal of Chemistry*, 2019, **43**, 13020-13037.
175. J.-D. Chai and M. Head-Gordon, *Physical Chemistry Chemical Physics*, 2008, **10**, 6615-6620.
176. G. L. Patrick, *An introduction to medicinal chemistry*, Oxford university press 2023.
177. J.-C. Guo, L.-Y. Feng, J. Barroso, G. Merino and H.-J. Zhai, *Chemical Communications*, 2020, **56**, 8305-8308.
178. J. M. Millam, V. r. Bakken, W. Chen, W. L. Hase and H. B. Schlegel, *The Journal of chemical physics*, 1999, **111**, 3800-3805.
179. S. Nosé, *The Journal of chemical physics*, 1984, **81**, 511-519.
180. W. G. Hoover, *Physical review A*, 1985, **31**, 1695.
181. C. Riplinger, P. Pinski, U. Becker, E. F. Valeev and F. Neese, *The Journal of chemical physics*, 2016, **144**, 024109.
182. D. G. Truhlar, *Chemical Physics Letters*, 1998, **294**, 45-48.
183. F. Neese, A. Hansen and D. G. Liakos, *The Journal of chemical physics*, 2009, **131**, 064103.

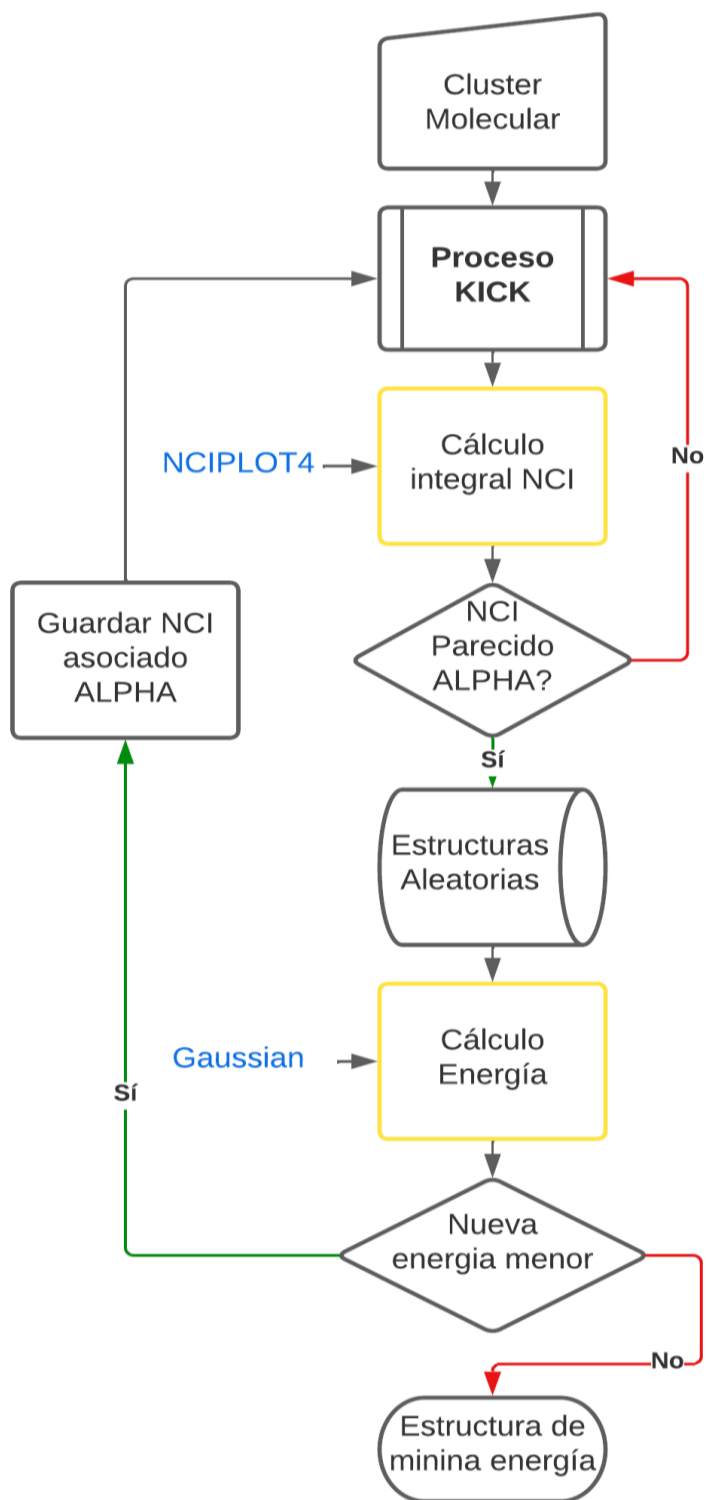
184. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *The Journal of chemical physics*, 2010, **132**, 154104.
185. D. Inostroza, A. Vásquez-Espinal, L. Leyva-Parra, W. García-Argote, M. L. Ceron, O. Yañez and W. Tiznado, *Physical Chemistry Chemical Physics*, 2023.
186. M. L. Pitteway, *The Computer Journal*, 1967, **10**, 282-289.
187. J. Rezáč, K. E. Riley and P. Hobza, *Journal of chemical theory and computation*, 2011, **7**, 2427-2438.
188. J. Řezáč, K. E. Riley and P. Hobza, *Journal of Chemical Theory and Computation*, 2014, **10**, 1359-1360.

ANEXOS

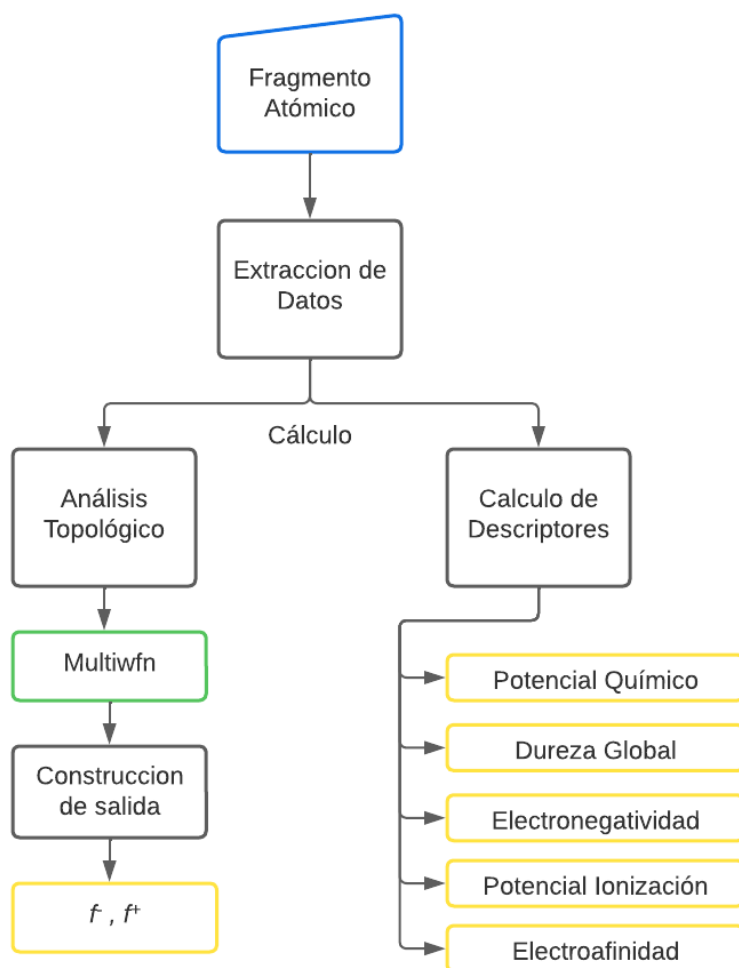
ANEXO 1



Esquema 1. Diseño de herramienta para la construcción automática de ensamblés guiados por la función de Fukui entre fragmentos atómicos. Dos subprocesos son evaluados, el primero consiste en el análisis topológico de la función de Fukui para la obtención de puntos críticos; el segundo es el proceso KICK en un volumen de interacción acorde. El resultado esperado son los ensamblés finales jerarquizados según el valor de la integral Coulombica.



Esquema 2. Diseño de herramienta para la integral del NCI propuesta donde los clústeres atómicos son formados mediante un procedimiento KICK, son evaluados con NCIPLOT 4 para luego ser optimizados. El proceso itera según valores de NCI "Alphas" asociados a estructuras de baja energía que sirven de guía para la formación de nuevos sistemas. El proceso finaliza cuando no se formen más estructuras de menor energía.



Esquema 3. Diseño de la herramienta *Topological Analysis of Fukui Function* TAFF donde a partir de la información de un fragmento atómico es posible calcular los descriptores de reactividad descritos en DFT. TAFF funciona como un pipeline entre el usuario y el software Multiwfn acelerando y automatizando la obtención del análisis topológico de la función de Fukui.

