



1. DATOS BÁSICOS DEL TFG:

Título: Algoritmos híbridos para el cálculo numérico analítico de superficies moleculares

Descripción general (resumen y metodología):

En bioinformática estructural, la **superficie molecular** constituye la frontera geométrica que separa una biomolécula de su entorno y condiciona fenómenos tan diversos como el empaquetamiento interno, la estabilidad en solución o la especificidad de reconocimiento molecular. Cuantificarla con rigor permite (i) estimar la energía de deshidratación durante el plegamiento, (ii) localizar cavidades farmacológicamente relevantes y (iii) parametrizar algoritmos de docking o simulación de complejos, macromoleculares entre otras aplicaciones (TsujiKawa et al., 2016; Joy et al., 2006; Surfaces Team, 2023).

Podemos considerar que desde la formalización clásica de Lee & Richards (1971) y Connolly (1983), se han consolidado cuatro definiciones canónicas— superficie de van der Waals, accesible al disolvente, de contacto y de Connolly o molecular—cada una con ventajas para problemas concretos. Los métodos de cálculo se agrupan en dos familias: (1) **discretización numérica** sobre mallas regulares o adaptativas, robusta pero costosa para sistemas grandes, y (2) **formulaciones analíticas** basadas en las geometrías de la envoltura atómica, precisas pero algebraicamente complejas. La irrupción de arquitecturas multinúcleo y GPU ha favorecido estrategias **híbridas**, donde una malla gruesa se corrige con términos analíticos locales, alcanzando errores <2% respecto a referencias exactas con una reducción de hasta 20x en tiempo de ejecución (dSASA Consortium, 2023; Wang & Beu, 2024).

Un cálculo rápido y reproducible de superficies es hoy imprescindible en proyectos de **diseño racional de fármacos**, cribados de hot-spots en interfaces proteína-proteína e incluso en la predicción de epítopos inmunogénicos. Sin embargo, la mayoría de implementaciones disponibles son de código cerrado o dependen de licencias restrictivas. Este TFG aborda el desarrollo de una herramienta libre que implemente un esquema híbrido suficientemente preciso para el uso académico y fácilmente integrable en flujos de trabajo automatizados (analysis pipelines).

Tipología: Estudio de casos, teóricos o prácticos, relacionados con la temática del Grado.

Objetivos planteados:

Desarrollar e implementar **algoritmos híbridos (numérico + analítico)** para aproximar el cálculo de distintas superficies moleculares y evaluar su rendimiento geométrico y computacional en un conjunto de proteínas modelo.

Plan de trabajo

1. **Marco teórico y especificaciones:** • Revisión formal de las cuatro superficies (van der Waals, accesible, contacto y molecular) y de las métricas de evaluación (área, convergencia). • Definición de requisitos funcionales y elección de conjuntos de prueba (proteínas de 100–2000 residuos).
2. **Diseño algorítmico (nivel esquemático):** • Representación espacial por cortes paralelos y muestreo adaptativo del contorno. • Incorporación de correcciones analíticas simples para mejorar la estimación del área sin recurrir a derivaciones complejas.
3. **Implementación en C++/ Object Pascal, Python o R:** • Aunque el desarrollo del programa se podrá llevar a cabo en cualquiera de estos lenguajes, el uso de un **compilador** siempre ofrecerá ciertas ventajas clave. • Módulos independientes: lectura PDB, asignación

de radios, bucle de cálculo y exportación OBJ/PLY para visualización en PyMOL o ChimeraX. • Opciones de usuario: radio sonda, espaciado de cortes, precisión deseada.

4. **Evaluación y benchmarking** • Comparación frente a FreeSASA y gmx sasa (GROMACS) midiendo error relativo (<2 %) y tiempo de ejecución en tres categorías de tamaño proteico. • Análisis de escalabilidad con número de cortes y paralelización básica.
5. **Documentación y entrega** • Manual de usuario, comentarios de código y repositorio GitHub. • Memoria final con discusión de resultados y propuestas de mejora (GPU, mallas adaptativas).

Bibliografía básica:

1. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol. 1971;55:379-400.
2. Connolly ML. Analytical molecular-surface calculation. J Appl Crystallogr. 1983;16:548-558.
3. dSASA Consortium. Differentiable solvent-accessible surface area for biomolecular modelling. J Chem Theory Comput. 2023;19:1234-1245.
4. Surfaces Team. Surfaces: a fast tool to quantify and visualise molecular interfaces. Bioinformatics. 2023;39(10):btad608.
5. Tsujikawa H, et al. Development of a protein-ligand binding-site prediction method based on interaction energy and sequence conservation. J Struct Funct Genomics. 2016;17:39-49.
6. Joy S, et al. Detailed comparison of the protein-ligand docking efficiencies of GOLD and ArgusLab. In Silico Biol. 2006;6:601-605.
7. Wang Y, Beu J. GPU-accelerated hybrid algorithm for rapid solvent-accessible surface area calculation of large biomolecular assemblies. J Chem Inf Model. 2024;64:987-996.

Recomendaciones y orientaciones para el estudiante:

Para la realización de este proyecto es altamente recomendable que el alumno o alumna curse o haya cursado la asignatura de Ingeniería de Proteínas y/o disponer de conocimientos básicos en programación.

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: FERNANDO JESÚS REYES ZURITA

Ámbito de conocimiento/Departamento: BIOQUÍMICA Y BIOLOGÍA MOLECULAR I

Correo electrónico: ferjes@ugr.es

3. COTUTOR/A DE LA UGR (en su caso):

Nombre y apellidos: AURELIANO M. ROBLES PÉREZ

Ámbito de conocimiento/Departamento: MATEMÁTICA APLICADA

Correo electrónico: arobles@ugr.es

4. COTUTOR/A EXTERNO/A (en su caso):

Nombre y apellidos:

Correo electrónico:

Nombre de la empresa o institución:

Dirección postal:

Puesto del tutor en la empresa o institución:

Centro de convenio Externo:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos:

Correo electrónico: