



## BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS QUÍMICAS

CENTRO DE QUÍMICA - INSTITUTO DE CIENCIAS

# TESIS

LA DIMENSIÓN FRACTAL COMO MEDIDA  
CARACTERIZADORA DE LA ESTRUCTURA  
DE PROTEÍNAS

PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS QUÍMICAS EN EL ÁREA DE  
FISICOQUÍMICA

PRESENTA:

QUÍM. ÉDGAR GARCÍA JUÁREZ

DIRECTOR DE TESIS:

DR. JUAN MANUEL SOLANO ALTAMIRANO

CODIRECTORA DE TESIS:

DRA. VIRIDIANA VARGAS CASTRO

PUEBLA, PUEBLA A 30 DE OCTUBRE DE 2025



*Amo todo lo que fue,  
todo lo que ya no es,  
el dolor que ya no duele,  
la antigua y errónea fe,  
el ayer que me dejó un dolor,  
lo que dejó la alegría  
sólo porque pasó, voló  
y hoy ya es otro día.*

***Poesías Inéditas (1930-1935)***

***Fernando Pessoa***

*De los dioses: el tener buenos abuelos, buenos progenitores, buena hermana, buenos maestros, buenos amigos íntimos [...]*

***Meditaciones***

***Marco Aurelio***

# Agradecimientos

Agradezco a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el financiamiento otorgado a través de una beca de posgrado con número de CVU 1260790, destinada a mis estudios de maestría.

Asimismo a la Vicerrectoría de Docencia y Estudios de Posgrado (VIEP-BUAP) por el apoyo económico proporcionado que hizo posible la difusión de este trabajo en los siguientes congresos:

- XXII Reunión Mexicana de Fisicoquímica Teórica (RMFQT).
- XXIII Congreso de la Sociedad Latinoamericana de Biología Matemática (SoLaBiMa).
- XXIII Reunión Mexicana de Fisicoquímica Teórica (RMFQT).

Mi gratitud también se dirige a mis asesores: Dr. Juan Manuel Solano Altamirano y Dra. Viridiana Vargas Castro, cuya orientación y respaldo constante, han sido pilares fundamentales para el desarrollo de esta tesis.

De igual forma, extiendo mi sincero agradecimiento a los miembros del comité revisor: Dr. Julio Manuel Hernández Pérez, Dr. Humberto Laguna Galindo, Dr. José Luis Carrillo Estrada y Dr. Minhuy Hô, por su dedicación al evaluar este trabajo y por sus aportaciones, las cuales enriquecieron notablemente este manuscrito.

# Índice general

|  |          |
|--|----------|
| <b>1. Introducción</b>   | <b>1</b> |
| <b>2. Objetivos</b>  | <b>4</b> |
| 2.1. Objetivo general . . . . .  | 4        |
| 2.1.1. Objetivos particulares . . . . .  | 4        |
| <b>3. Marco teórico</b>  | <b>6</b> |
| 3.1. Fractales . . . . .   | 6        |
| 3.2. Ley de potencia . . . . .   | 7        |
| 3.3. Dimensión fractal . . . . .   | 8        |
| 3.3.1. Dimensión fractal de masa . . . . .   | 8        |
| 3.4. Multifractalidad . . . . .  | 10       |
| 3.5. Objetos multifractales y sistemas magnetoreológicos . . . . .                     | 11       |
| 3.6. Niveles de organización en la estructura de las proteínas . . . . .               | 12       |
| 3.7. Medidas para comparar estructuras proteicas . . . . .                             | 14       |
| 3.7.1. Geometría fractal en proteínas . . . . .  | 16       |
| 3.8. Metodologías para la determinación de la dimensión fractal en proteínas . . . . . | 17       |
| 3.8.1. Primeras aproximaciones y enfoques pioneros (1980–1990) . . . . .               | 17       |

---

## ÍNDICE GENERAL

---

|  |           |
|--|-----------|
| 3.8.2. Consolidación de métodos geométricos y autoafines (1990–2000) . . . . . | 19        |
| 3.8.3. Enfoques dinámicos y de simulación (2000–2025) . . . . .                | 21        |
| <b>4. Resultados y discusión</b>   | <b>27</b> |
| 4.1. Algoritmo (diseño del programa) . . . . .                                 | 27        |
| 4.1.1. Pasos 1, 2, 3 y 4 del diagrama de flujo general . . . . .               | 28        |
| 4.1.2. Paso 5 del diagrama de flujo general . . . . .                          | 30        |
| 4.1.3. Paso 6 y 7 del diagrama de flujo general . . . . .                      | 31        |
| 4.1.4. Paso 8 del diagrama de flujo general . . . . .                          | 32        |
| 4.1.5. Paso 9 del diagrama de flujo general . . . . .                          | 33        |
| 4.1.6. Paso 10 del diagrama de flujo general . . . . .                         | 34        |
| 4.2. Metodología . . . . .   | 35        |
| 4.2.1. Densidad de partículas promedio $\rho_P$ . . . . .                      | 35        |
| 4.2.2. Pruebas de fractalidad en proteínas . . . . .                           | 38        |
| 4.2.3. Elección de sistemas . . . . .  | 39        |
| 4.2.4. Elección de medida de análisis . . . . .                                | 40        |
| 4.2.5. Determinación de $M(r)$ . . . . .                                       | 41        |
| 4.2.6. Análisis multifractal . . . . .   | 42        |
| 4.3. Discusión de resultados . . . . .   | 43        |
| 4.3.1. Dimensiones fractales en distintas etapas . . . . .                     | 44        |
| 4.3.2. Efecto de iones en la dimensión fractal . . . . .                       | 45        |
| 4.3.3. Caso particular de estudio:tubulinas . . . . .                          | 47        |
| <b>5. Conclusión</b>   | <b>48</b> |
| <b>A. Funcionamiento del programa</b>  | <b>50</b> |
| <b>B. Conjunto de 9 proteínas</b>  | <b>52</b> |

## ÍNDICE GENERAL

---

**Referencias**

**62**

# Índice de figuras

---

## ÍNDICE DE FIGURAS

---

|   |    |
|---|----|
| 4.2. Diagrama 2D del cálculo de la dimensión fractal de masa en proteínas mediante <i>molfractaldim</i> . (1) Se define un radio máximo en la proteína ( $R_{max}$ ) alrededor del centro de masa del sistema, (2) se determina si un centro local es adecuado ( $\rho_p \geq 0.34$ ). Estas cantidades se utilizan para calcular la masa promedio de las partículas $\bar{M}(r_k)$ .   | 31 |
| 4.3. Gráficos obtenido mediante el programa <i>molfractaldim</i> de $\log_{10} r$ vs $\log_{10} Z(r)$ donde $\log_{10} Z(r)$ corresponde a $\langle M(r_k) \rangle$ , $\langle N(r_k) \rangle$ o $\langle Rg(r_k) \rangle$ .  | 34 |
| 4.4. Distribución de la densidad de partículas promedio ( $\rho_P$ ) en función del índice de proteínas (ldx) para los archivos PDB de las Tablas 4.1 y 4.2. La línea azul horizontal discontinua representa la media ( $\nu$ ) de $\rho_P$ , mientras que las líneas rojas horizontales discontinuas indican los límites de dos desviaciones estándar ( $\nu \pm 2\sigma$ ) de $(\rho_P)$ .  | 37 |
| 4.5. Gráficos de $\log_{10} r$ vs $\log_{10} N(r)$ correspondientes la proteína <i>Tubulina mutada</i> después de una dinámica molecular de 10 ns, estableciendo el número de medidas en $N_{means} = 1$ y 200 (izquierda a derecha). Donde $N_{means}$ establece el número de medidas para obtener $\langle N(r_k) \rangle$ , $\langle M(r_k) \rangle$ y $\langle Rg(r_k) \rangle$ .   | 41 |
| 4.6. Valores de dimensión fractal de masa $D$ vs identificadores de proteínas (IdPDB) para las proteínas ilustradas en la Tabla 4.3, correspondientes a cuatro etapas de procesamiento: (1) Adición de átomos de hidrógeno al sistema proteico, (2) al minimizar la energía de la estructura molecular, (3) equilibrando el sistema bajo condiciones termodinámicas controladas y (4) después de una dinámica molecular de 1 ns. Las líneas azules, verdes y rojas representan $D$ en los intervalos de 1 a 2 Å, 2 a 6 Å y 6 a 20 Å, respectivamente. | 43 |
| 4.7. Gráficos de $\log_{10} r$ vs $\log_{10} M(r)$ correspondientes a dos etapas de procesamiento de la proteína <i>Translocon EsA</i> de la Tabla 4.3 (con iones $\text{Na}^+$ en su estructura) después de (1) equilibrar el sistema bajo condiciones termodinámicas controladas y (2) de una dinámica molecular de 1 ns. Los valores de $b$ en cada gráfica representan la dimensión fractal de masa en los intervalos 1 a 2 Å, 2 a 6 Å y 6 a 20 Å.  | 45 |

---

## ÍNDICE DE FIGURAS

---

|  |    |
|--|----|
| 4.8. Gráficos de $\log_{10} r$ vs $\log_{10} M(r)$ correspondientes a dos etapas de procesamiento de la proteína <i>Translocon EsA</i> de la Tabla 4.3 (sin iones $\text{Na}^+$ en su estructura) de la Tabla 4.3 después de (1) equilibrar el sistema bajo condiciones termodinámicas controladas y (2) de una dinámica molecular de 1 ns. Los valores de $b$ en cada gráfica representan la dimensión fractal de masa en los intervalos 1 a 2 Å, 2 a 6 Å y 6 a 20 Å. . . . . | 45 |
| 4.9. Gráficos de $\log_{10} r$ vs $\log_{10} M(r)$ correspondientes a dos proteínas diferentes: <i>Tubulina nativa</i> (alanina) y <i>Tubulina mutada</i> (treonina) después de una dinámica molecular de 10 ns. Los valores de $b$ en cada gráfica representan la dimensión fractal de masa en los intervalos 1 a 2 Å, 2 a 6 Å y 6 a 20 Å. . . . .  | 47 |
| <br>B.1. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la primera proteína con <i>IdPDB:1a2b</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . .  | 53 |
| B.2. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la séptima proteína con <i>IdPDB:1b3e</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . .  | 54 |
| B.3. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la octava proteína con <i>IdPDB:11gs</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . .   | 55 |

---

## ÍNDICE DE FIGURAS

---

|   |    |
|---|----|
| B.4. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la sexta proteína con <i>IdPDB:1auk</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . .   | 56 |
| B.5. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la cuarta proteína con <i>IdPDB:1a8m</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . .  | 57 |
| B.6. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la tercera proteína con <i>IdPDB:1a52</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . . | 58 |
| B.7. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la segunda proteína con <i>IdPDB:1a3n</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . . | 59 |
| B.8. Regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ correspondiente a cuatro etapas de procesamiento de la quinta proteína con <i>IdPDB:1a9w</i> de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . .  | 60 |

## ÍNDICE DE FIGURAS

---

- B.9. Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la novena proteína con *IdPDB:7khw* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns. . . . . 61

# Índice de tablas

|  |    |
|--|----|
| 4.1. Índice de proteínas (ldx), identificadores del <i>Protein Data Bank</i> (IdPDB) y nombre de las proteínas (Parte 1). . . . .  | 35 |
| 4.2. Índice de proteínas (ldx), identificadores del <i>Protein Data Bank</i> (IdPDB) y nombre de las proteínas (Parte 2). . . . .  | 36 |
| 4.3. Nuevo índice de proteínas (ldx), Identificadores del <i>Protein Data Bank</i> (PDB), nombre de la proteína y número de átomos presentes en cada estructura. . . . . | 39 |
| 4.4. Resumen de datos de las regresiones lineales de $\log_{10} r$ vs $\log_{10} M(r)$ de las Figuras 4.7 y 4.8. . . . .   | 46 |
| A.1. Banderas de entrada para el programa <i>molmassfractaldim</i> . . . . .   | 51 |
| B.1. Nuevo índice de proteínas (ldx), Identificadores del <i>Protein Data Bank</i> (PDB), nombre de la proteína y número de átomos presentes en cada estructura. . . . . | 52 |

Capítulo 1

## Introducción

Las proteínas desempeñan un papel central en los sistemas biológicos, siendo indispensables para una vasta gama de funciones que sustentan la vida. Su estudio es crucial en múltiples disciplinas, desde la comprensión de procesos fisiológicos hasta el diseño de enzimas con aplicaciones biotecnológicas. La síntesis proteica, aunque no completamente elucidada, sigue un mecanismo ampliamente aceptado que implica varias etapas de plegamiento y agregación molecular. Inicialmente, los aminoácidos (unidades estructurales básicas) se organizan en estructuras secundarias localizadas ( $\alpha$  hélices y láminas  $\beta$ ). Estas, a su vez, se ensamblan en una conformación tridimensional específica (estructura terciaria), que determina la función biológica de las proteínas [3]. Finalmente, las proteínas emergen como polímeros funcionales, cuya estabilidad y actividad dependen de su correcto plegamiento. Si bien, la estructura proteica que se captura en un cristal es representativa, en ella sólo se observa una de las muchas formas que una proteína puede adoptar durante el transcurso de su función biológica. Por lo tanto, la noción tradicional que concibe a las proteínas como objetos tridimensionales compactos resulta insuficiente para describir la complejidad estructural de estos sistemas macromoleculares. Es por lo anterior que, en los últimos años se ha señalado la posibilidad de que las proteínas puedan estar mejor caracterizadas mediante conceptos provenientes de la geometría fractal en lugar de recurrir únicamente a modelos tridimensionales convencionales [4–7].

---

## INTRODUCCIÓN

---

En este contexto, la dimensión fractal emerge como un parámetro relevante para cuantificar la complejidad geométrica inherente en las estructuras proteicas. Investigaciones como las de Enright *et al.* [1] han calculado la dimensión fractal de varias proteínas de forma global, sin el análisis de varias etapas de agregación. Es decir, sin analizar cómo cambian los patrones de agregación cuando se cambia la escala de medición de la dimensión fractal de masa. Otros autores como Banerji y Ghosh [27] han proporcionado una revisión en las metodologías de la dimensión fractal en las proteínas, pero no reportan explícitamente el análisis multifractal en dichos sistemas.

Sin embargo, al igual que otros sistemas complejos y como ya se dijo, las proteínas presentan múltiples niveles de agregación y conformación que no se distribuyen homogéneamente en el espacio. Por lo que esta heterogeneidad podría sugerir la existencia de comportamientos multifractales, como se ha observado en sistemas magnetorreológicos (piedras, polvos magnetizados o geles envejecidos), en los que la dimensión fractal masa varía a lo largo de diferentes etapas del proceso de formación. Lo anterior sucede porque cuando se mide la dimensión fractal de masa a diferentes escalas, se puede obtener información acerca de los patrones de agregación. El análisis multifractal encontrado por Carrillo *et al.* [2], aplicado a diferentes tipos de estructuras (no simplemente a geles envejecidos), podría servir para extraer información de los patrones de agregación de otros objetos en diferentes etapas.

Por otra parte, el problema de comparar diferentes estructuras de una misma proteína continúa siendo un desafío relevante en la química computacional [9]. Una de las métricas más utilizadas para este propósito es la desviación cuadrática media (*rmsd*, por sus siglas en inglés). Esta medida resulta fundamental en áreas como la bioquímica computacional y dinámica molecular, ya que permite evaluar el grado de similitud entre estructuras obtenidas por experimentos o simulaciones moleculares. Sin embargo, el *rmsd* presenta limitaciones importantes, especialmente cuando las diferencias estructurales son locales.

---

## INTRODUCCIÓN

Por ejemplo, cuando dos estructuras con diferencias en segmentos específicos presentan valores globales de *rmsd*, ocultando variaciones funcionalmente relevantes. Por esta razón, es necesario desarrollar métodos alternativos o complementarios que permitan comparar estructuras a un nivel local, enfocándose en regiones críticas que eviten pequeñas fluctuaciones estructurales que distorsionan la comparación global [9].

Se sabe que, en las proteínas existe más de una etapa de agregación o de conformación, esto plantea la siguiente pregunta; ¿Se podrían detectar estos patrones de agregación en las proteínas midiendo la dimensión fractal de masa en este esquema de objetos multifractales? De ser cierto lo anterior, ¿sería esta medida lo suficientemente precisa para distinguir entre dos proteínas con secuencias genéticas similares pero estructuralmente diferentes? Para responder a estas interrogantes, es fundamental desarrollar una herramienta computacional que nos permita calcular la dimensión fractal de masa y llevar a cabo pruebas exhaustivas. Es importante resaltar que, de ser cierto lo anterior, se podría abrir la puerta a futuras aplicaciones en enfermedades. No obstante, aún no se sabe con certeza si es posible observar la multifractalidad en proteínas, por lo que este trabajo se enfocará en determinar si se observa.

Capítulo **2**

# Objetivos

## 2.1. Objetivo general

Determinar si las proteínas presentan características multifractales. De ser así, analizar dicha multifractalidad con el fin de utilizar una o más variantes de la dimensión fractal como medida(s) caracterizadora(s) de la estructura y patrones de agregación de las proteínas.

### 2.1.1. Objetivos particulares

- Desarrollar una herramienta de cómputo que permita optimizar el cálculo de la dimensión fractal de masa en sistemas proteicos.
- Seleccionar un conjunto de proteínas de trabajo para calibrar el método y validar la medida de multifractalidad.
- Adaptar la determinación de la dimensión fractal de masa a las proteínas, a partir de otros casos como geles o materiales ferromagnéticos.
- Explorar el uso potencial de la dimensión fractal de masa como una herramienta para identificar o resolver entre proteínas genéticamente similares pero con diferencias estructurales notables.

## OBJETIVOS

---

- Seleccionar un conjunto de proteínas de estudio que tengan potencial para aplicaciones posteriores. Por ejemplo proteínas nativas y mutadas, proteínas que han mantenido su función biológica pero que se han diferenciado estructuralmente, entre otros casos. Se priorizarán los casos en que se cuente con información experimental (*PDB*).

Capítulo **3**

## Marco teórico

### 3.1. Fractales

Diversos físicos que trabajan en múltiples campos han reconocido que muchas de las estructuras observadas en sus experimentos exhiben un tipo particular de complejidad geométrica, cuya comprensión ha sido fuertemente influenciada por los trabajos de Benoit Mandelbrot quien llamó la atención sobre las propiedades geométricas de diversos objetos [6].

Los fractales son objetos geométricos que presentan una estructura fragmentada y aparentemente irregular que se manifiestan en muchos contextos, tanto naturales como los copos de nieve, como artísticos por ejemplo, en las obras de M. C. Escher, así como en fenómenos físicos, *verbi gratia* cúmulos de galaxias, entre otras áreas. El término fractal proviene del latín *fractus* que significa “roto o quebrado” y fue acuñada por el matemático polaco B. Mandelbrot, padre de los fractales. La definición de fractal depende del área en la que se esté trabajando, sin embargo existen varias propiedades que tienen estos objetos y si alguna de ellas se cumple entonces se dice que es un objeto fractal:

- Es autosimilar.
- Su dimensión de *Hausdorff-Besicovich* es mayor que su dimensión topológica.
- No es diferenciable en ningún punto y tiene una longitud o complejidad infinita.

La propiedad mas importante en fractales es la autosimilitud. La autosimilitud implica observar como las mismas propiedades o características de un objeto se replican a distintas escalas. En fractales matemáticos o puros como el conjunto de Mandelbrot, la autosimilitud es exactamente igual sin importar el número de veces que se amplie, siempre se verá lo mismo. A la autosimilitud anterior se le conoce como autosimilitud exacta. Otro tipo de autosimilitud, es la autosimilitud aproximada que normalmente aparece en objetos donde sus partes o conjuntos de esas partes son muy similares (aunque no idénticas) y en su mayoría, aparecen en la naturaleza, como en las arterias o venas de la retina de un ojo. A diferencia de la autosimilitud exacta, la autosimilitud aproximada solo se puede replicar una cierta cantidad de veces y después, se pierde esa autosimilitud.

Dicho todo lo anterior, podríamos decir que un fractal es un objeto geométrico en el que una misma estructura irregular o aparentemente fragmentada se repite a diferentes escalas y tamaños.

### 3.2. Ley de potencia

Existe una estrecha relación entre la ley de potencia y el estudio de fractales. La ley de potencia es una relación matemática de la siguiente forma:

$$y(x) = cx^a \quad (3.1)$$

Donde  $c$  es una constante,  $a$  es el exponente de la ley de potencias,  $x$  y  $y$  son las variables dependientes e independientes. La ecuación 3.1 es una función matemática que relaciona dos cantidades, donde el cambio en una cantidad da como resultado un cambio en la otra cantidad que es proporcional al cambio elevado a un exponente constante. Es decir, una cantidad varía como potencia de otra. El cambio es independiente del tamaño inicial de dichas cantidades [10].

### 3.3. Dimensión fractal

En geometría fractal existe un concepto fundamental conocido como dimensión fractal, que es una generalización del concepto de dimensión en geometría Euclíadiana. La dimensión fractal ( $D$ ) es una medida de cuánto aparenta llenar el espacio un objeto conforme aumentamos o disminuimos la escala de análisis. De la amplia variedad de dimensiones fractales que existen, la definición de la dimensión fractal de *Hausdorff-Besicovich* es probablemente la más usada [6, 10]. Sin embargo, es necesario medir o calcular cantidades que puedan demostrar que están relacionadas con la dimensión fractal de los objetos que se analizan. En este contexto, existen tres tipos de enfoques principales para la determinación de la dimensión fractal: (1) el experimental, (2) el teórico y (3) el informático [6].

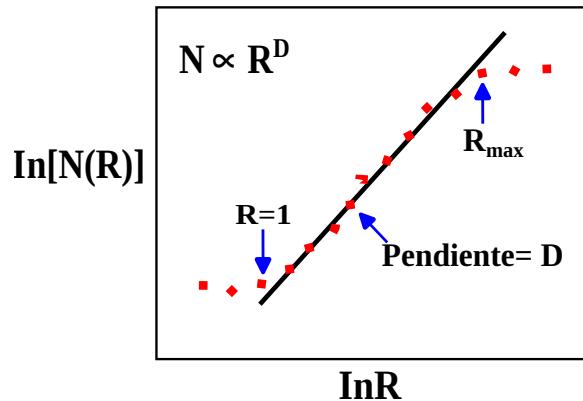
El enfoque informático se obtienen digitalizando imágenes o por procedimientos numéricos. Respecto a este último, los datos generados numéricamente suelen producirse mediante variaciones del método Monte Carlo y haciendo uso de datos obtenidos experimentalmente [6]. Para hacer estimaciones precisas generalmente se calcula la dimensión fractal para muchos grupos y se promedia sobre los resultados. A continuación hablamos de un método particular para calcular  $D$ .

#### 3.3.1. Dimensión fractal de masa

La dimensión fractal de masa es útil para estimar la dimensión de objetos similares como redes, vasos sanguíneos o sistemas de agregación. Este método consiste en seleccionar un punto que pertenece al objeto de estudio (normalmente el centro de masa) y contar el número de partículas  $N$  o sitios que pertenecen al objeto dentro de una secuencia de esferas de radio  $R$ . La dimensión fractal de masa  $D$  esta definida por la ley de potencia:

$$M \sim R^D \quad (3.2)$$

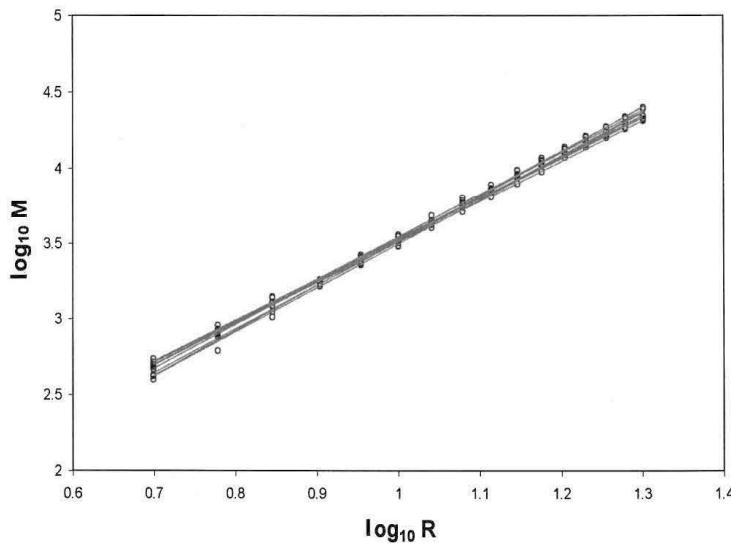
La dimensión fractal de masa  $D$  puede calcularse ajustando una línea recta de los datos  $\ln[N(R)]$  contra  $\ln R$  dando como resultado una curva cuya pendiente asintótica es igual a  $D$ , ver Figura 3.1 . Por conveniencia, en lo que sigue usaremos con frecuencia el término partícula para referirnos a un sitio del sistema que pertenece al fractal y *clúster* para los objetos compuestos por partículas conectadas [6].



**Figura 3.1:** Gráfico de  $\ln[N(R)]$  vs  $\ln(R)$  del número de partículas  $N(R)$  que pertenecen a un fractal que se encuentran dentro de una esfera de radio  $R$ . La dimensión fractal se obtiene ajustando una línea recta a los datos en la región de escala.

## ¿Qué es un objeto monofractal?

Un objeto monofractal es una estructura que puede describirse mediante una única dimensión fractal. Esto implica que su complejidad o irregularidad es **uniforme** en todas las escalas y regiones del objeto. Por lo tanto, su pendiente es constante y el mismo valor de dimensión fractal describe la estructura en todas las partes del objeto. Un caso concreto de monofractalidad es el estudio presentado por Enright *et al.* [1], donde no se observa un cambio de pendiente (veáse la Figura 3.2), probablemente por la pequeña cantidad de puntos que se presentan en su análisis y la distancia que hay entre los puntos. Además, Enright *et al.* calculó la dimensión fractal de masa como un promedio espacial (usando esferas concéntricas y  $R_g$ ), lo que posiblemente captura esta monofractalidad.



**Figura 3.2:** Gráfico de  $\log_{10} M$  vs  $\log_{10} R$  para la proteína Sialidase, donde los valores de  $M$  son las masas encerradas por esferas concéntricas de radio  $R$  centradas en un átomo. Imagen tomada de Enright *et al* [1].

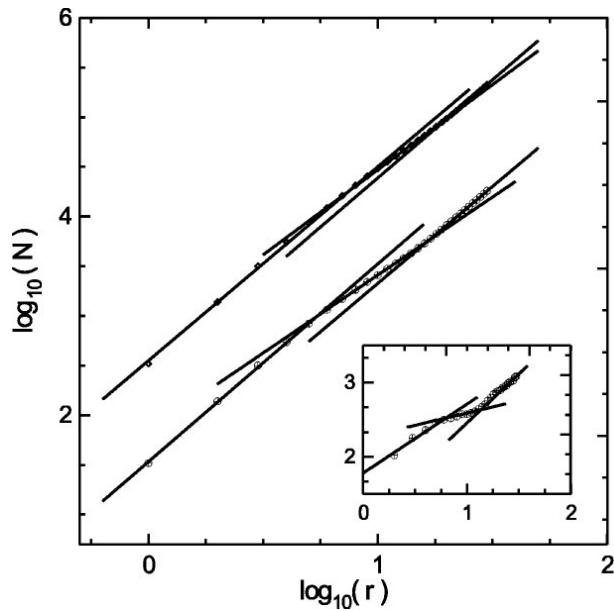
### 3.4. Multifractalidad

Otro concepto desarrollado por B. Mandelbrot en su libro *Fractal Geometry of Nature* fue la multifractalidad. La multifractalidad es una propiedad de ciertos sistemas o estructuras complejas que presentan un crecimiento fractal pero de manera heterogénea en distintas regiones o escalas. A diferencia de los fractales, que son descritos con una única dimensión fractal (un valor constante que representa la relación entre el detalle del patrón y la escala), en un sistema multifractal existen múltiples dimensiones fractales que reflejan la variabilidad de la distribución y concentración de sus elementos.

En términos simples, un sistema multifractal está compuesto de distintas subestructuras que tienen diferentes grados de “irregularidad”, lo que significa que cada región del sistema podría necesitar una dimensión fractal específica para describir su complejidad.

### 3.5. Objetos multifractales y sistemas magnetoreológicos

Un objeto multifractal es una estructura fractal que requiere un conjunto de dimensiones fractales para ser descrita completamente. La complejidad del objeto varía localmente y está distribuida de manera heterogénea, es decir, existen múltiples pendientes que dependen de la región del objeto o de la medida que se esté analizando y es común en fenómenos naturales o en sistemas caóticos. Un ejemplo claro es el trabajo realizado por Carrillo *et al.* [2], donde los sistemas magnetorreológicos y otros conglomerados formados por procesos de agregación, mostraron multifractalidad, que se manifiesta en la variación de la dimensión fractal a lo largo de diferentes etapas del proceso de formación de estructuras (véase la Figura 3.3). En términos simples, en la Figura 3.3, se observa que en los puntos de  $\log_{10} N(r)$  vs  $\log_{10} r$  se pueden clasificar en tres conjuntos, cada conjunto sigue una relación lineal con una pendiente definida, esto es, hay tres dimensiones fractales.



**Figura 3.3:** Gráfico de  $\log_{10} N$  vs  $\log_{10} r$  para estructuras formadas en un sistema magnetorreológico de tres etapas. La multifractalidad es una propiedad de ciertos sistemas complejos donde la estructura exhibe autosimilitud a múltiples escalas. Modificado de Carrillo *et al* [2].

Respecto a los sistemas magnetoreológicos, son fluidos complejos que contienen partículas magnéticas (óxidos de metales) suspendidos en un líquido (como aceite de silicona). La característica principal de estos sistemas es que sus propiedades mecánicas cambian drásticamente cuando se aplica un campo magnético externo. En ausencia de este campo, el fluido se comporta como un líquido ordinario, pero al aplicar un campo magnético, las partículas magnéticas se alinean y forman estructuras organizadas (como cadenas o fibras) que transforman el fluido en una especie de gel envejecido o sólido semirrígido en cuestión de milisegundos.

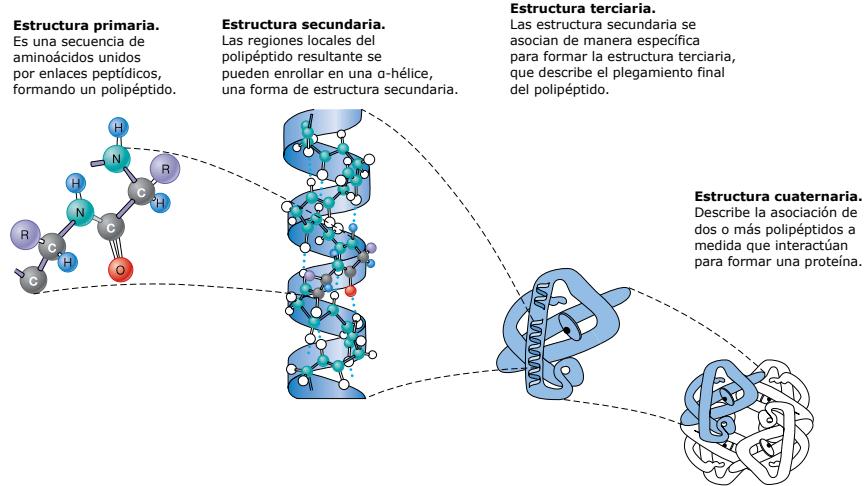
### **3.6. Niveles de organización en la estructura de las proteínas**

Las proteínas están presentes en todos los sistemas vivos, desde estructuras como la hemoglobina o el tejido cerebral y una cantidad considerable de esas proteínas se han cristalizado para posteriormente, caracterizarse por métodos como RMN, R-X, ME. Una vez hecho lo anterior, los datos son enviados y son revisados por expertos biocuradores, después de ser aprobados se ponen a disposición de forma gratuita bajo algún dominio como el *Protein Data Bank* [11].

Como es bien sabido, las proteínas son polímeros lineales formados por aminoácidos y aunque en la célula se han identificado más de 60 aminoácidos diferentes, solo 20 de ellos son incorporados de manera habitual en la síntesis de proteica.

Cada aminoácido presente en las proteínas tiene una estructura básica compuesta por un grupo amino ( $-NH_2$ ), un grupo carboxilo ( $-COOH$ ), un átomo de hidrógeno ( $-H$ ) y una cadena lateral o grupo R, todos ellos unidos a un átomo de carbono central quiral (conocido como carbono  $\alpha$ ). La unión covalente entre el grupo carboxilo de un aminoácido y el grupo amino de otro, con la liberación de una molécula de agua ( $H_2O$ ), da lugar al denominado enlace peptídico, el cual constituye la base estructural de las cadenas polipeptídicas. Podemos describir a las proteínas en cuatro niveles jerárquicos (véase la Figura 3.4):

- **Estructura primaria:** Corresponde a la secuencia lineal de aminoácidos que constituyen la cadena polipeptídica. Su descripción especifica el orden de los aminoácidos desde el extremo amino-terminal (N-terminal) hasta el extremo carboxilo-terminal (C-terminal) de la molécula.
- **Estructura secundaria:** Consiste en patrones regulares y repetitivos en la disposición espacial local de la cadena, estabilizados principalmente por puentes de hidrógeno. Estas interacciones dan lugar a dos conformaciones secundarias principales: las  $\alpha$ -hélices y láminas  $\beta$  plegadas.
- **Estructura terciaria:** Resulta de interacciones entre los grupos laterales (radicales) de los aminoácidos, los cuales presentan diversas propiedades químicas (hidrofobicidad, polaridad, carga eléctrica). Estos enlaces provocan el plegamiento, enrollamiento y torsión de la cadena en una conformación tridimensional específica, conocida como estructura nativa, que generalmente representa la configuración más estable para una secuencia dada.
- **Estructura cuaternaria:** Es el nivel de organización que implica la asociación y ensamblaje de múltiples cadenas polipeptídicas (subunidades) para formar una proteína funcional. Esta estructura es característica de proteínas oligoméricas, comúnmente aquellas con un peso molecular superior a 50,000 uma.



**Figura 3.4:** Niveles de organización en la estructura de las proteínas. Imagen extraída de Hardin *et al.* [3].

### 3.7. Medidas para comparar estructuras proteicas

Para evaluar y comparar estructuras proteicas a los niveles de organización proteica antes mencionados, se utilizan diferentes medidas y aunque el problema parece sencillo su cuantificación es compleja y sigue evolucionando [9]. Una de esas medidas es la desviación cuadrática media (por su acrónimo en inglés *root mean square deviation, rmsd*) es un parámetro fundamental para caracterizar transformaciones conformacionales en sistemas moleculares [12]. Esta medida cuantifica la discrepancia estadística promedio entre una configuración molecular en un tiempo  $t$  y una estructura de referencia definida en el tiempo inicial  $t_0$ .

$$\text{rmsd}(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\mathbf{r}_i(t) - \mathbf{r}_i(t_0)\|^2} \quad (3.3)$$

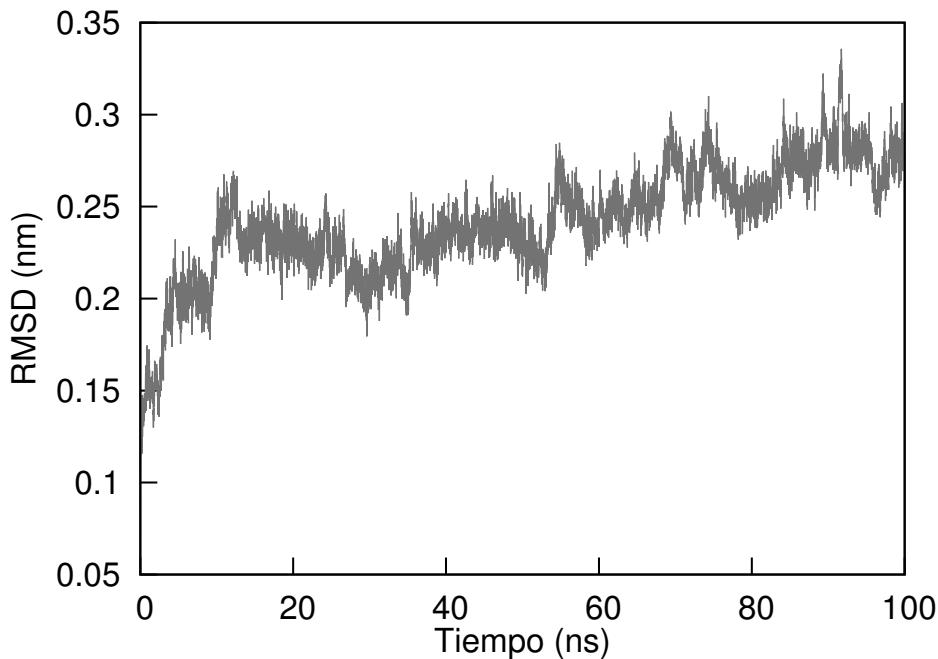
En la ecuación 3.3, el vector  $\mathbf{r}_i(t)$  representa las coordenadas espaciales del i-ésimo átomo en el instante  $t$ , mientras que  $\mathbf{r}_i(t_0)$  corresponde a su posición de referencia. La suma se realiza sobre el conjunto de  $N$  partículas seleccionadas para el análisis. El uso del la  $rmsd$  se extiende a diversas propiedades dinámicas como velocidades atómicas y fuerzas interatómicas. Para comparaciones estructurales entre distintas moléculas, es necesario realizar operaciones de superposición geométrica (traslación y rotación) que optimicen el alineamiento conformacional. En estudios proteicos, es común restringir el cálculo a los carbonos  $\alpha$ , ya que estos forman parte de la estructura polipeptídica y proporcionan una representación robusta del esqueleto proteico. Además, la evolución temporal del  $rmsd$  exhibe comportamientos diferenciados según el estado de agregación: en sistemas sólidos mantiene valores reducidos pero en fase líquida presenta un crecimiento lineal inicial que eventualmente se estabiliza tras periodos prolongados. La gráfica que se observa en la Figura 3.5 es producida mediante dinámica molecular y en ella, se estudia cómo evolucionan las posiciones de los átomos de una proteína. Es importante señalar que la  $rmsd$  es una medida global. En la Figura 3.5 no analiza si la estructura proteica se está deshaciendo, se está expandiendo o se está torciendo. Solo es, cuánto en promedio se desvían los átomos de su posición original.

Otra medida altamente utilizada es el radio de giro ( $R_g$ ). Esta métrica describe el equilibrio conformacional (estabilidad y flexibilidad estructural) de una proteína en un entorno biológico, también evalúa la compactación de la estructura durante una simulación de dinámica molecular [13]. Un valor pequeño de  $R_g$  indica una estructura rígida, mientras que un valor alto de  $R_g$ , implica un incremento en la flexibilidad de la estructura. Matemáticamente, el  $R_g$  se representa como:

$$R_g = \sqrt{\frac{1}{M} \sum_{i=1}^n m_i (r_i - R)^2} \quad (3.4)$$

Donde  $M$  es la masa total de la proteína,  $m_i$  es la masa de cada átomo,  $r_i$  es la posición de cada átomo,  $R$  es el centro de masa de la proteína.

En este contexto, encontrar métodos alternativos al  $rmsd$  y al  $R_g$  o hallar enfoques complementarios para comparar estructuras proteicas sigue siendo un reto para la comunidad científica.



**Figura 3.5:** Evolución temporal del *rmsd* a 90 ns de la proteína *Tubulina mutada (treonina)*.

### 3.7.1. Geometría fractal en proteínas

Tras establecer la definición de dimensión fractal, multifractalidad (ver sección 3.5) y presentar un método para su determinación. Surge la siguiente pregunta ¿La dimensión fractal tiene alguna utilidad en sistemas proteicos?, la respuesta recae en varios experimentos en lo que se ha empleado la dimensión fractal para cuantificar algunos aspectos de la morfología proteica. Dado que la estructura de las proteínas tiene una forma tan compleja, solo se puede analizar adecuadamente mediante el enfoque de la geometría fractal, por ejemplo se puede utilizar para caracterizar la estructura terciaria de las proteínas y las enzimas [5]. Aunque se pueden construir fractales iterados que son perfectamente autosimilares, la autosimilitud de una cadena de proteínas, es similar solo en un sentido estadístico, es decir, no siempre se verá exactamente como el todo. Otra diferencia entre las macromoléculas biológicas y los objetos ideales es que una macromolécula no es autosimilar desde cualquier escala. Hay límites de tamaño superiores e inferiores más allá de los cuales una macromolécula ya no es un fractal. Por lo que la investigación de la geometría fractal en proteínas es un campo activo [5].

## 3.8. Metodologías para la determinación de la dimensión fractal en proteínas

El estudio de la dimensión fractal en proteínas ha evolucionado a través de una combinación de diferentes enfoques geométricos, estadísticos y dinámicos con diferentes objetivos. La diversidad de la metodología en la literatura responde tanto a la disponibilidad de datos estructurales tridimensionales como a la interpretación física de la geometría fractal en sistemas moleculares complejos. A continuación se muestra una breve revisión al estado del arte de la dimensión fractal en sistemas proteicos.

### 3.8.1. Primeras aproximaciones y enfoques pioneros (1980–1990)

Las primeras tentativas de cuantificar la naturaleza fractal de las proteínas se remontan a los trabajos de Stapleton *et al.* [14] en 1980, quienes investigaron la estructura fractal de las proteínas mediante estudios de relajación de espín electrónico con hierro ( $\text{Fe}^{3+}$ ). Los autores midieron la dependencia térmica de la tasa de relajación Raman y encontraron que la densidad de estados vibracionales sigue una ley de potencia. La dimensión fractal encontrada fue  $d \approx 1.65$  que posteriormente, fue confirmada con datos de rayos X de la mioglobina. Este trabajo fue pionero en demostrar experimentalmente que las proteínas tienen una estructura fractal. Aunque este estudio establece una base sólida para la determinación de la dimensión fractal en proteínas, no explora la multifractalidad ni variaciones locales en esas estructuras.

Cuatro años más tarde, Helman *et al.* [15] en 1984, propusieron una metodología basado en la dimensión fractal de fractones y modos vibracionales  $d_{fr}$  apoyada en la teoría de Alexander-Orbach [16], que depende de la topología del sistema (incluyendo los puentes de hidrógeno) y la dimensión fractal geométrica  $d_f$ . Su modelo teórico enfatiza la importancia de la topología proteica en la determinación de las propiedades fractales, pero se mantiene en el régimen monofractal sin abordar la multifractalidad.

Poco después, Isogai *et al.* [17] desarrollaron un análisis geométrico de la estructura terciaria de 43 proteínas utilizando la teoría fractal. Su enfoque estadístico-geométrico no considera mecanismos físicos como la relajación de espín. En su investigación lograron identificar dos regímenes fractales distintos: (1) La dimensión fractal  $S$  regida por repulsiones estéricas (presentes en estructuras locales) y (2) la dimensión fractal  $L$  regida por fuerzas atractivas (presentes en estructuras globales). Además, sugieren que la teoría fractal será particularmente útil para estudiar propiedades emergentes como la estabilidad conformacional, que dependen de la organización global más que de detalles locales. Si bien el estudio no es multifractal, la dependencia escalar de  $S$  y  $L$  indican heterogeneidad estructural que podría ser explorada mediante un análisis multifractal.

Un año más tarde, Wagner *et al.* en 1985 [18] parte del grupo de Stapleton *et al.*, desarrollaron un método para calcular dimensiones fractales  $\bar{d}$  a partir de coordenadas de carbonos  $\alpha$  en 70 proteínas, estableciendo un rango de  $1.27 \leq \bar{d} \leq 1.87$  y demostrando su correlación con el contenido de estructuras secundarias:  $\alpha$ -hélices y giros- $\beta$  aumentan  $\bar{d}$ , mientras que las láminas- $\beta$  lo disminuyen. Por último, descubrieron que el solvente influye en la dinámica fractal. Este trabajo estableció una conexión entre geometría fractal y dinámica proteica, pero se centró en la dimensión fractal del *backbone* y no analizó la dimensión fractal de masa ni la multifractalidad.

De manera experimental, Sow-Hsin y Teixeira [19] en 1986 introdujeron un enfoque innovador para caracterizar la estructura fractal de proteínas mediante dispersión de neutrones a bajo ángulo (SANS) en complejos proteína-detergente. Observaron la formación de complejos tipo “collar de perlas” donde las micelas de detergente se distribuyen de forma fractal a lo largo de la cadena polipeptídica desnaturizada. Por último, desarrollaron un modelo teórico que considera explícitamente la longitud de correlación finita, permitiendo extraer simultáneamente la dimensión fractal  $D$ , el tamaño micelar y el grado de desnaturización. Encontraron que  $D$  disminuye progresivamente desde 2.30 hasta 1.76 conforme aumenta la desnaturización de la proteína. En resumen, Sow-Hsin y Teixeira cuantificaron la dimensión fractal global de la distribución de micelas, análoga a la dimensión fractal de masa pero no aborda la multifractalidad a proteínas.

### 3.8.2. Consolidación de métodos geométricos y autoafines (1990–2000)

Xin Wang *et al.* [20] en 1990 presentaron un método refinado para determinar las dimensiones fractales de 90 proteínas provenientes del *Protein Data Bank* que abarcan cuatro clases estructurales: (1) dominios  $\alpha$  antiparalelos, (2) dominios  $\beta$  antiparalelos, (3) dominios  $\alpha - \beta$  paralelos y (4) dominios ricos en disulfuro o ricos en metales (SD). Analizaron la relación entre la dimensión fractal y la estructura terciaria de las proteínas. El valor medio de la dimensión fractal  $D_2$  para la estructura global de las proteínas es 1.65, muy cercano al valor teórico  $\frac{5}{3}$ , asociado con un *random walk* en el espacio euclíadiano tridimensional. Por último, encontraron que la dimensión fractal es un indicador en la evolución para proteínas homólogas y que varía de forma característica entre las clases de proteínas que analizaron, proporcionando una herramienta cuantitativa para la clasificación estructural. Aunque Xin Wang *et al.* proporcionan un análisis riguroso de la dimensión fractal global de la cadena peptídica, no abordan la dimensión fractal de masa ni la multifractalidad y en su lugar, se centran únicamente en la topología de la cadena principal.

Cserzö y Vicsek en 1991 [7], introdujeron un enfoque revolucionario al analizar a las proteínas como superficies autoafines en lugar de fractales autosimilares. Transformaron las estructuras proteicas en mapas de distancias  $C\alpha$ - $C\alpha$  que analizaron como superficies fractales mediante el cálculo del exponente de rugosidad  $H$ . Su método reveló que las proteínas nativas exhiben dos regímenes de escalado: A escalas pequeñas ( $H \approx 0.32$ ) reflejan flexibilidad local y estructuras secundarias, mientras que a escalas grandes ( $H \approx 0.12$ ) manifiestan compactación global. Esta dualidad de escala captura la naturaleza de las proteínas: flexibles de forma local pero compactas de forma global. En adición, desarrollaron modelos determinísticos en una red cúbica que reproducen este comportamiento y demostraron que el método es suficientemente sensible para detectar diferencias entre proteínas mesófilas y termófilas, así como para monitorizar procesos de despliegamiento simulado. El trabajo de Cserzö y Vicsek representó un avance conceptual significativo al mostrar que la autoafinidad (no la autosimilitud) es un marco teórico apropiado para describir la organización estructural de las proteínas, estableciendo un puente entre la flexibilidad local y la compactación global. Por lo tanto, su trabajo no habla de la distribución de masa y tampoco explora la multifractalidad.

En el mismo año, Li *et al.* [21] marca un punto de inflexión al superar la limitación para describir la superficie mediante una única dimensión fractal. Los autores introducen los conceptos *fat fractal* y multifractal, argumentando que la superficie de una proteína, siendo rugosa pero de volumen finito, constituye un fractal "gordo". Asimismo, criticaron los métodos de cálculo tradicionales por su inestabilidad en sistemas autoafines y proponen el método de variación para determinar la dimensión fractal superficial ( $D_G$ ) de forma más fiable. Su contribución fundamental reside en demostrar la naturaleza multifractal de estas superficies a través del espectro  $f(\alpha)$ , revelando una heterogeneidad intrínseca donde coexisten múltiples subconjuntos fractales con distintos grados de singularidad. Esto permite explicar que regiones específicas, como el sitio activo de una proteína, presentan una corrugación mayor que el promedio global, optimizando procesos como la catálisis. Es importante mencionar que este es uno de los pocos trabajo que explora explícitamente la multifractalidad en proteínas, aunque se centra en superficies y no en la masa total de la proteína.

A principios de 1993, Dewey [22] realizó un análisis pionero de las múltiples dimensiones fractales en proteínas globulares, estableciendo las bases para su caracterización como objetos fractales. Determinó la dimensión fractal del esqueleto ( $D = 1/\nu = 2.86 \pm 0.03$ ) mediante regresión de  $R_g \sim N^\nu$  sobre 45 proteínas, confirmando el comportamiento como polímeros compactos ( $\nu = 0.35 \pm 0.03$ ). Simultáneamente, calculó la dimensión fractal de superficie ( $D_s = 2.16$ ) a partir de la relación área-volumen ( $A \sim V^{D_s/3}$ ), aplicando la ley de aditividad de codimensiones para explicar la rugosidad superficial. En el ámbito dinámico, analizó la dimensión espectral ( $\tilde{d}$ ) reportando valores de 1.34 para proteínas con Fe-S y 1.67 para hemoproteínas, que interpretó mediante modelos de relajación con puentes locales. Utilizó un análisis de Hurst, para revelar correlaciones de largo alcance en factores de *Debye-Waller* ( $H = 0.77$  para *backbone* vs  $H = 0.67$  para cadenas laterales), simulando este fenómeno mediante un modelo de Ising unidimensional desordenado que genera correlaciones tipo ley de potencia. Aunque este trabajo sienta las bases conceptuales para un estudio multifractal al demostrar la multiplicidad de dimensiones fractales en proteínas, no desarrolla formalmente un análisis multifractal ni calcula espectros de singularidad.

La investigación de Dewey constituye una contribución fundacional para la física fractal de proteínas, estableciendo el paradigma de que estos sistemas exhiben propiedades de escala en múltiples niveles estructurales y dinámicos.

### 3.8.3. Enfoques dinámicos y de simulación (2000–2025)

A principios del siglo XXI Moret *et al.* [23] investigaron las propiedades multifractales de la hipersuperficie de energía potencial en polipéptidos y proteínas. El comportamiento multifractal se obtuvo a partir del espectro  $f(\alpha)$ , cuya función describe algunas propiedades estructurales de una proteína (ofreciendo una explicación alternativa a la paradoja de Le-vinthal). Además, encontraron que es necesario tomar en cuenta la formación de enlaces de hidrógeno ya que influye directamente en la dimensión fractal del sistema. Por otro lado, explican como el disolvente puede perturbar la estructura terciaria de una proteína y por lo tanto, modificar el comportamiento multifractal.

Burioni *et al.* en 2004 [24], investigaron la inestabilidad térmica topológica en proteínas de un solo dominio mediante el Modelo de Red Gaussiana (GNM), focalizándose en la dimensión espectral ( $\tilde{d}$ ) como parámetro descriptor de la conectividad global. Establecieron una correlación significativa entre  $\tilde{d}$  y la longitud de la cadena proteica ( $N$ ), formalizada mediante la relación  $\frac{2}{\tilde{d}} = a + \frac{b}{\ln(N)}$ , donde los parámetros de ajuste  $a = 0.63$  y  $b = 2.61$  (para  $R_0 = 7$  Å) reflejan cómo la topología nativa se ajusta para mantener estabilidad termodinámica. Reportaron que valores de  $\tilde{d} < 2$  previenen la divergencia de fluctuaciones atómicas según la relación  $\langle r^2 \rangle \propto N^{2/\tilde{d}-1}$ , con valores típicos en el rango  $1.56 < \tilde{d} < 2.10$  para proteínas entre 100-3600 residuos. Este trabajo vinculó la dimensión espectral con límites de estabilidad estructural, demostrando que proteínas más largas adoptan topologías menos compactas. Aunque el estudio profundiza en las consecuencias termodinámicas de la dimensión fractal dinámica, no explora la dimensión fractal de masa ni aborda la multifractalidad, centrándose exclusivamente en propiedades vibracionales globales.

Un año más tarde, nuevamente el grupo de Moret *et al.* [25] investigó las propiedades fractales de 5526 cadenas proteicas diferentes a partir de un análisis de la dimensión fractal con un valor de  $\delta = 2.47$ , dicho grupo sostiene que este valor de dimensión proporciona una medida de la compacidad de la proteína. Indicando que el análisis multifractal puede describir algunas propiedades estructurales de las proteínas y corrobora la explicación sobre la multifractalidad en la hipersuperficie energética. Por último, sugieren que los experimentos realizados, son cruciales para determinar la mejor descripción del mecanismo de plegamiento de proteínas específicas.

Enright *et al.*, publicaron dos estudios que guardan una estrecha relación con el presente tema de investigación. En el primer artículo publicado en 2005 [1], se estableció una metodología para calcular la dimensión fractal de masa a partir de estructuras tridimensionales obtenidas del *Protein Data Bank* usando un conjunto de 200 proteínas que abarcan desde 100 hasta 10,000 aminoácidos y examinaron la variación de  $D$  con el tamaño de la proteína. El valor promedio  $\bar{D} = 2.5$ , que es significativamente menor que un polímero tridimensional completamente compacto. También observaron que, en promedio, una proteína en su configuración nativa del *Protein Data Bank* ocupa menos del 80% de volumen en la proteína, lo que concuerda con una dimensión fractal de masa inferior a 3. La masa de la proteína también se ajustó al radio de giro ( $R_g$ ), con un exponente de 2.5 para el conjunto de proteínas antes mencionado que corresponde estrechamente con las dimensiones fractales de masa que obtuvieron al estudiar las vibraciones de varias proteínas mediante las relaciones de *Alexander-Orbach* [16].

Nuevamente, Enright *et al.* [8] esta vez en 2006 ampliaron su estudio sobre la dimensión fractal de masa en proteínas al incorporar el efecto de la hidratación, utilizando simulaciones de dinámica molecular para analizar cómo influye el agua en la difusión vibracional. Este trabajo permitió evidenciar que la dimensión fractal varía en función del contenido de agua, y que la propagación de energía en proteínas hidratadas presenta un comportamiento anómalo, característico de sistemas con estructura fractal. Así, se estableció una conexión entre la geometría fractal de las proteínas y su dinámica energética. Como parte del estudio, se calcularon las dimensiones fractales de masa de un conjunto de 200 proteínas del *Protein Data*

*Bank*, con tamaños que oscilan entre 100 y 11,000 aminoácidos, considerando estructuras con y sin las moléculas de agua asociadas. Los resultados mostraron que la inclusión de las moléculas de agua internas y de la primera capa de hidratación produce un ligero aumento en el valor promedio de la dimensión fractal, pasando de 2.49 (sin agua) a 2.52 (con agua). Adicionalmente, se realizaron simulaciones de dinámica molecular en solución acuosa para 20 proteínas de entre 100 y 4000 aminoácidos. En este caso, al considerar explícitamente la hidratación de los sistemas, la dimensión fractal de masa promedio se incrementó hasta 2.87, valor que se mantuvo constante sin depender del tamaño de la proteína. Se observó, además, que variaciones estructurales moderadas (como aquellas derivadas de un despliegamiento parcial que modifica el radio de giro en un 10%) no afectan significativamente la dimensión fractal.

Chang-Yong Lee en el año de 2006 [26], investigó las propiedades de dimensión fractal de masa en el ribosoma bacteriano, implementando la relación fundamental  $N \propto R^{D_M}$  donde  $N$  representa el número de átomos contenidos dentro de esferas concéntricas de radio  $R$ . Calculó  $D_M$  para ambas subunidades ribosómicas y sus rRNAs constituyentes, revelando una dicotomía estructural fundamental: la subunidad 30S y el rRNA 16S exhiben dimensiones fractales de  $2.58 \pm 0.06$  y  $2.82 \pm 0.07$  respectivamente, mientras que la subunidad 50S y el rRNA 23S presentan valores cercanos a 3 ( $3.07 \pm 0.08$  y  $3.11 \pm 0.07$ ). Esta divergencia en compactación fractal se correlaciona directamente con la función biológica: un menor valor de  $D_M$  de la subunidad 30S refleja su naturaleza dinámica y activa en la decodificación del mRNA, mientras que la mayor  $D_M$  de la subunidad 50S indica una estructura rígida especializada. El estudio implementó un análisis considerando 27 orígenes diferentes para minimizar sesgos en el cálculo de  $D_M$ . Aunque este trabajo proporciona un análisis cuantitativo riguroso de dimensión fractal de masa en un complejo macromolecular crucial, no explora la multifractalidad ni analiza las posibles variaciones locales en las propiedades de escalado. La investigación de Lee establece un precedente importante para relacionar dimensiones fractales globales con dinámica funcional en complejos proteína-RNA.

En el año de 2011, Banerji y Ghosh [27] realizaron una revisión de las metodologías de dimensión fractal aplicadas al interior de proteínas, estableciendo aproximaciones que incluyen dimensiones de masa, correlación, espectral y teorías de renormalización. En el contexto de la dimensión fractal de masa, los autores formalizaron el concepto mediante la relación  $M \sim R^{D_m}$ , donde  $D_m$  cuantifica cómo la masa atómica se distribuye espacialmente dentro de la proteína, relacionándose directamente con su grado de compactibilidad. Reportaron que los valores típicos de  $D_m$  en proteínas globulares ( $2.00 < D_m < 3.00$ ) reflejan un estado intermedio entre polímeros extendidos y compactos, en estudios específicos reportaron valores alrededor de  $2.489 \pm 0.172$ . Este trabajo vinculó la dimensión fractal de masa con propiedades biofísicas clave como hidrofobicidad y polarizabilidad, revelando que diferentes clases estructurales de proteínas ( $\alpha, \beta, \alpha/\beta, \alpha + \beta$ ) exhiben simetrías de escala distintivas en la distribución de estas propiedades. Aunque la revisión explora múltiples facetas de la fractalidad proteica, no aborda explícitamente el análisis multifractal, centrándose principalmente en caracterizar monofractales. El trabajo de Banerji y Ghosh representa una metodología para la cuantificación de la autosimilitud estructural en proteínas, estableciendo las bases para estudios más especializados en análisis fractal de sistemas proteicos.

En mayo de 2013 Peng *et al.* [28] publicaron un estudio de las propiedades fractales en estructuras proteicas, introduciendo y comparando la dimensión fractal local ( $D_L$ ) y la dimensión fractal del esqueleto ( $D_B$ ) para 750 proteínas de diferentes clases estructurales. Determinaron que las proteínas exhiben comportamiento fractal en el rango  $1 \leq N \leq 15$ , con  $D_B$  consistentemente mayor que  $D_L$  para una misma proteína, indicando mayor complejidad estructural global versus local. Cuantificaron un orden característico de dimensión fractal entre clases estructurales:  $\alpha$  ( $D_L = 1.59, D_B = 1.61$ )  $> \alpha/\beta$  ( $D_L = 1.55, D_B = 1.59$ )  $> \alpha + \beta$  ( $D_L = 1.51, D_B = 1.56$ )  $> \beta$  ( $D_L = 1.47, D_B = 1.53$ ), estableciendo así, una relación directa entre tipo de estructura secundaria y complejidad fractal. Los autores también desarrollaron un modelo de orbital híbrido fractal que relaciona  $D_B$  con los estados de hibridación atómica en la cadena principal, con valores promedio que sugieren orbitales del tipo  $sp^{4.889}$  para proteínas globulares.

Aunque este trabajo proporciona un análisis comparativo de múltiples dimensiones fractales en proteínas, no aborda la multifractalidad ni explora la heterogeneidad en las propiedades de escalado a través de la estructura proteica.

En 2024 Petreus *et al.* [29] publicaron un estudio en que se analizaban los aspectos fractales de las estructuras de las proteínas S100 para comprender mejor su complejidad. Para el cálculo de las dimensiones fractales de masa y superficie, se consideraron 33 estructuras en solución y 18 estructuras cristalinas correspondientes a las proteínas S100 humanas. El valor de la dimensión fractal de masa obtenido fue de  $D_m = 1.54$ , lo que confirmó la conformación extendida de los dímeros de estas proteínas según su investigación. El valor medio de la dimensión fractal de superficie es  $D_s = 2,35 \cdot 0,09$  cuando se calcula utilizando estructuras en solución y  $D_s = 2,23 \cdot 0,05$  cuando se calcula utilizando estructuras cristalinas, lo que revela las irregularidades superficiales en las proteínas S100. También se registró cambios en las dimensiones fractales de superficie de dichas proteínas debido a cambios en el pH del entorno, mutaciones en sus secuencias que alteran el plegamiento de la proteína y sus interacciones con iones o ligandos que reflejan los reordenamientos estructurales producidos tras la unión. Estos cambios pueden influir significativamente en la actividad biológica de la proteína, lo que convierte a la dimensión fractal de superficie en un parámetro valioso para el estudio de las funciones proteicas, sus interacciones y su posible orientación terapéutica.

Czub *et al.* en enero de 2025 [30] investigaron la separación de fases de la proteína *Bik1*, un componente clave de los complejos +TIP en levaduras. Mediante una combinación de dispersión de rayos X a bajo ángulo (SAXS), microscopía y espectrometría de masas, caracterizando la organización supramolecular de los condensados proteicos. Uno de los hallazgos de esta investigación fue que los condensados de *Bik1* exhiben una estructura fractal con dimensión fractal de masa  $d \approx 2$ , determinada a partir del decaimiento de la señal de SAXS ( $I(q) \sim q^{-2}$ ). Esta fractalidad se extiende en un rango de escalas de  $\sim 30$  a  $300$  nm, revelando una red heterogénea con regiones densas en las proteínas y vacíos ricos en solvente. Los autores proponen que esta organización surge de un proceso de percolación, donde los dímeros de *Bik1* forman oligómeros extendidos que dan lugar a la red fractal observada. El estudio se centra en una dimensión fractal global y no explora explícitamente la multifractalidad.

En el transcurso de 20025, Hughes *et al.* [31] caracterizaron la evolución estructural de redes en hidrogeles formadas por albúmina sérica bovina (BSA) force-lábil, empleando dispersión de rayos X a bajo ángulo (SAXS). Ellos cuantificaron la dimensión fractal de masa ( $D_f$ ), observando que el desplegamiento *in situ* induce transiciones de  $D_f = 2.48 \pm 0.01$  (BSA mecánicamente robusta) a  $D_f = 2.75 \pm 0.01$  (BSA force-lábil), indicativo de clusters más densos y compactos. Demostraron que este incremento en  $D_f$  correlaciona con un aumento de 3-4 veces el módulo de almacenamiento ( $G'$ ), estableciendo una relación directa entre compactación fractal y la rigidez mecánica. Metodológicamente, implementaron un modelo de ensamblaje triple que distingue la formación primaria ( $k^\alpha = 0.78 \pm 0.01\text{min}^{-1}$ ), secundaria ( $k^\beta = 0.32 \pm 0.01\text{min}^{-1}$ ) y relajación por desplegamiento ( $\tau = 3200 \pm 100$  s). Aunque este trabajo proporciona un análisis cuantitativo riguroso de la dimensión fractal de masa en dichos sistemas, no explora la multifractalidad en la distribución espacial de masa a nivel de proteína individual. La investigación de Hughes *et al.* representa un avance metodológico para el estudio de transiciones estructurales en sistemas proteicos, demostrando la sensibilidad de  $D_f$  para capturar cambios conformacionales inducidos por desplegamiento.

Existen otros estudios que no guardan una relación directa con el enfoque específico de esta investigación, pero abordan temáticas vinculadas al análisis proteico mediante medidas fractales. Para el lector interesado, se recomienda la consulta de las referencias [32–34]. Dichos trabajos no se discuten en detalle en el presente estudio, dado que sus objetivos y, en particular, sus metodologías, se encuentran considerablemente distantes del enfoque adoptado en esta tesis.

# Resultados y discusión

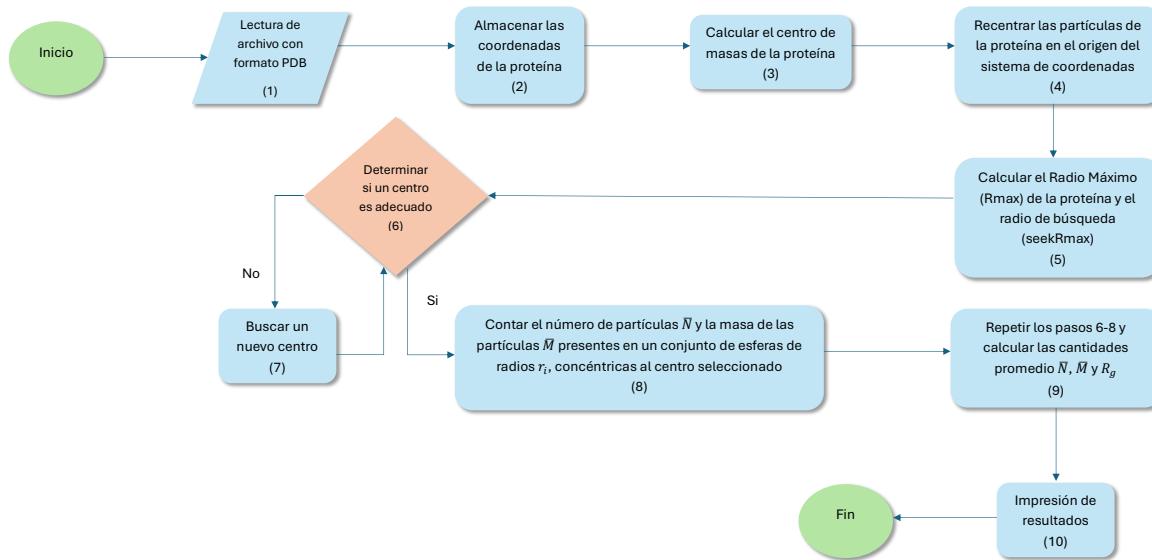
Uno de los principales objetivos de esta tesis fue desarrollar un programa de cómputo científico para explorar la aplicación del análisis multifractal en estructuras proteicas. Por lo tanto, el programa podría ser una herramienta útil para caracterizar las distintas etapas del plegamiento proteico, describiéndolas en términos de la dinámica de formación de agregados estructurales, lo que permitiría una comprensión más profunda de estos procesos. En la siguiente sección se describe el algoritmo del programa *molmassfractaldim*.

## 4.1. Algoritmo (diseño del programa)

El propósito principal del programa *molmassfractaldim* es determinar la dimesión fractal de masa mediante la caracterización de la geometría interna de un sistema proteico. Esto se logra construyendo una tabla de  $r_k$ ,  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle Rg(r_k) \rangle$ : donde;

- $r_k$  es el radio de medida centrado en posiciones aleatorias de una proteína con valores mínimos y máximos definidos como  $r_{min}$  y  $r_{max}$ .
- $\langle N(r_k) \rangle$  es el número promedio de partículas contenidas dentro de un radio  $r_k$ .
- $\langle M(r_k) \rangle$  es el número de masa promedio de las partículas contenidas en el radio  $r_k$ .
- $\langle Rg(r_k) \rangle$  es el radio de giro promedio de las partículas contenidas en el radio  $r_k$ .

Posteriormente, los valores de  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle Rg(r_k) \rangle$  pueden analizarse mediante regresiones lineales para determinar la dimensión fractal de masa. El algoritmo general del programa se describe en el diagrama de flujo de la figura 4.1, cuyos pasos se explican en las siguientes subsecciones. Para una consulta general sobre las opciones disponibles de *molmassfractaldim*, véase el apéndice A.



**Figura 4.1:** Diagrama de flujo general del programa *molmassfractaldim*.

#### 4.1.1. Pasos 1, 2, 3 y 4 del diagrama de flujo general

- **Paso 1:** Sea  $\mathcal{M}$  un sistema molecular compuesto por un conjunto de partículas  $N_{\text{at}}$ . Cada partícula  $i$  está descrita por la terna de la ecuación 4.1.

$$\mathcal{M}(\mathbf{r}_i, m_i, Z_i) \quad (4.1)$$

Donde  $\mathbf{r}_i$  es el vector de posición,  $m_i$  la masa y  $Z_i$  el número atómico. El sistema completo se representa por la ecuación 4.2.

$$\mathcal{M} = \{(\mathbf{r}_i, m_i, Z_i) \mid i = 1, \dots, N_{\text{at}}\} \quad (4.2)$$

- **Paso 2:** Para usar las posiciones de las partículas  $i$  de  $\mathcal{M}$  se define una matriz  $\mathbf{X}$ , que servirá como base para todos los cálculos posteriores, véase la ecuación 4.3.

$$\mathbf{X} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_{N_{\text{at}}} & y_{N_{\text{at}}} & z_{N_{\text{at}}} \end{bmatrix} \in \mathbb{R}^{N_{\text{at}} \times 3} \quad (4.3)$$

Donde la  $i$ -ésima fila es la traspuesta  $\mathbf{r}_i^T$ .

- **Paso 3:** El centro de masas  $\mathbf{r}_{\text{cm}}$  del sistema  $\mathcal{M}$  se calcula con la ecuación 4.4.

$$\mathbf{r}_{\text{cm}} = \frac{\sum_{i=1}^{N_{\text{at}}} m_i \mathbf{r}_i}{\sum_{i=1}^{N_{\text{at}}} m_i} \in \mathbb{R}^3 \quad (4.4)$$

$\mathbf{r}_{\text{cm}}$  representa el punto de equilibrio mecánico del sistema  $\mathcal{M}$ . Donde  $\sum_{i=1}^{N_{\text{at}}} m_i \mathbf{r}_i$  es la suma de las posiciones ponderadas por las masas y  $\sum_{i=1}^{N_{\text{at}}} m_i$  es la masa total del sistema.

- **Paso 4:** Se trasladan las coordenadas de  $\mathcal{M}$  a su centro de masa usando las ecuaciones 4.5 y 4.6.

$$\mathbf{r}'_i = \mathbf{r}_i - \mathbf{r}_{\text{cm}}, \quad i = 1, \dots, N_{\text{at}} \quad (4.5)$$

Esto garantiza que

$$\sum_{i=1}^{N_{\text{at}}} m_i \mathbf{r}'_i = \mathbf{0} \quad (4.6)$$

Donde  $\mathbf{r}'_i$  es la posición de la partícula  $i$  medida con respecto al centro de masa del sistema  $\mathcal{M}$  y  $r_{cm}$  es el vector del centro de masa del sistema  $\mathcal{M}$ .

#### 4.1.2. Paso 5 del diagrama de flujo general

- **Radio máximo.** Se define el radio máximo del sistema  $\mathcal{M}$ , calculado como la distancia máxima desde el origen hasta la partícula más lejana del sistema, véase la ecuación 4.7.

$$R_{\max} = \max_i \|\mathbf{r}'_i\| \quad (4.7)$$

- **Radio de búsqueda.** Se establece un radio que asegure que los puntos de medición no se encuentren cerca de los límites del sistema (ecuación 4.8), con un valor del 75 % tamaño de  $\mathcal{M}$ , donde  $\alpha = 0.75$ .

$$R_{\text{seek}} = \alpha \cdot R_{\max} \quad (4.8)$$

- **Radio de medición.** Se describe un conjunto de radios de medición como  $n_r$  con un incremento radial  $dr$  (ecuaciones 4.9 y 4.10).

$$n_r = 50 \quad (4.9)$$

$$dr = \frac{r_{\max} - r_{\min}}{n_r - 1} \quad (4.10)$$

Donde

$$r_{\min} = 0.15 R_{\max}, \quad r_{\max} = 20 \quad (4.11)$$

- **Radio de medición local.** Para cada centro de muestreo se especifica un centro local, véase la ecuación 4.12.

$$r_{\text{loc}} = \beta R_{\max}, \quad \beta = 0.15 \quad (4.12)$$

$\beta$  limita el tamaño de las esferas a 15 % de  $R_{\max}$ .

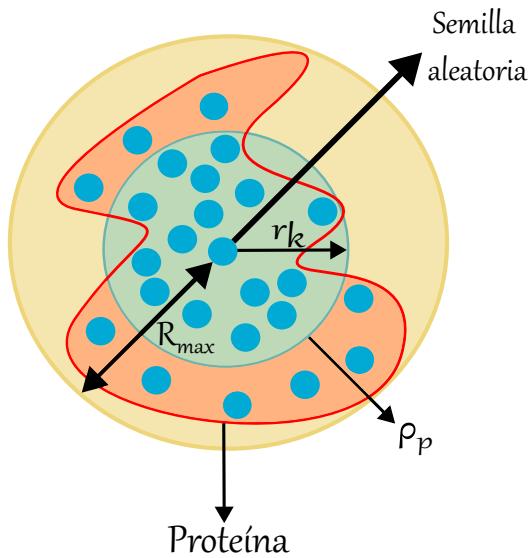
- **Centros de medición válidos.** Un centro  $\mathbf{r}_c$  puede usarse como centro sólo si se cumple la condición 4.13.

$$\|\mathbf{r}_c\| + r_{\text{loc}} \leq R_{\text{seek}} \quad (4.13)$$

La condición anterior asegura que la esfera de radio  $r_{loc}$  quede completamente dentro de la región de muestreo.

#### 4.1.3. Paso 6 y 7 del diagrama de flujo general

En este paso se utiliza un criterio llamado densidad de partículas promedio  $\rho_P$  que tiene como objetivo determinar cuando un centro local ( $r_c$ ) (ver paso 6 del diagrama de flujo presentado en la Fig. 4.1) no es adecuado para realizar la medición de la dimensión fractal masa. Por ejemplo, si una esfera de cierto radio centrada en algún punto de la proteína tiene una  $\rho_P < 0.034$  (en la sección 4.2.1 se discute a detalle la razón del valor anterior), sería altamente probable que la esfera intersecte en una zona considerada como un hueco de la proteína. Por lo tanto, un centro con tales características podría ser descartado hasta que se encuentre otro centro adecuado (ver bucle 6  $\longleftrightarrow$  7 en la Figura 4.1). Visualmente este criterio se observa en la Figura 4.2.



**Figura 4.2:** Diagrama 2D del cálculo de la dimensión fractal de masa en proteínas mediante *molfractaldim*. (1) Se define un radio máximo en la proteína ( $R_{max}$ ) alrededor del centro de masa del sistema, (2) se determina si un centro local es adecuado ( $\rho_p \geq 0.34$ ). Estas cantidades se utilizan para calcular la masa promedio de las partículas  $\bar{M}(r_k)$ .

Además de lo anterior,  $\rho_P$  permite refinar la medición de  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle Rg(r_k) \rangle$ , que a su vez son las relaciones que se utilizan para medir las dimensiones fractales y así, determinar si existe o no multifractalidad en estructuras proteicas. El criterio  $\rho_P$  calcula dos valores clave: (1) El volumen de cada esfera de radio  $r_k$  y (2) la densidad de partículas promedio presente en cada esfera de radio  $r_k$ . El volumen de cada esfera se calcula usando el radio  $r_k$  de las esferas definidas anteriormente y a partir de la ecuación 4.14.

$$V(r_k) = \frac{4}{3}\pi r_k^3 \quad (4.14)$$

La densidad de partículas promedio se calcula a partir de la ecuación 4.15. Donde  $\rho_P$  es la densidad de partículas promedio presentes en la esfera de radio  $r_k$ .  $\bar{N}(r_k)$  es la cantidad promedio de partículas presentes en una esfera de radio  $r_k$ .  $V_k$  es el volumen de una esfera de radio  $r_k$ .

$$\rho_P = \frac{\bar{N}(r_k)}{V_k} \quad (4.15)$$

#### 4.1.4. Paso 8 del diagrama de flujo general

Para cada radio  $r_k$  en la malla  $\{r_{\min}, r_{\min} + \Delta r, \dots, r_{\max}\}$ , se definen las ecuaciones 4.16, 4.17 y 4.18.

$$N(\mathbf{r}_c, r_k) = \sum_{i=1}^{N_{\text{at}}} \Theta(r_k - \|\mathbf{r}'_i - \mathbf{r}_c\|) \quad (4.16)$$

$$M(\mathbf{r}_c, r_k) = \sum_{i=1}^{N_{\text{at}}} m_i \Theta(r_k - \|\mathbf{r}'_i - \mathbf{r}_c\|) \quad (4.17)$$

$$R_g(\mathbf{r}_c, r_k) = \sqrt{\frac{\sum_{i=1}^{N_{\text{at}}} \|\mathbf{r}'_i - \mathbf{r}_c\|^2 \Theta(r_k - \|\mathbf{r}'_i - \mathbf{r}_c\|)}{N(\mathbf{r}_c, r_k)}} \quad (4.18)$$

Donde  $\mathbf{r}_c$  es el centro de muestreo dentro del sistema  $\mathcal{M}$ ,  $\mathbf{r}'_i$  es la posición del partícula  $i$  respecto al centro de masa,  $m_i$  es la masa del partícula  $i$ .  $\Theta(x)$  es la función de Heaviside que permite contar sólo los átomos dentro de la esfera, igual a 1 si  $x \geq 0$  y 0 si  $x < 0$ ,  $N(\mathbf{r}_c, r_k)$  es número de partículas dentro de la esfera de radio  $r_k$ ,  $M(\mathbf{r}_c, r_k)$  es la masa total dentro de la esfera y  $R_g(\mathbf{r}_c, r_k)$  es radio de giro de las partículas de la esfera.

#### 4.1.5. Paso 9 del diagrama de flujo general

**Cálculo de valores promedio sobre múltiples centros.** Se seleccionan aleatoriamente centros de muestreo en  $\{\mathbf{r}_{c_j}\}$  (veáse la ecuación 4.19) que puede modificarse a conteos menores (donde los promedios fluctuarán más) o conteos mayores (los promedios serán muy parecido entre sí).

$$N_{\text{meas}} = 30 \quad (4.19)$$

Por último, para calcular los valores de  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle R_g(r_k) \rangle$  se utilizan las ecuaciones 4.20, 4.21 y 4.22.

$$\langle N(r_k) \rangle = \frac{1}{N_{\text{meas}}} \sum_{j=1}^{N_{\text{meas}}} N(\mathbf{r}_{c_j}, r_k) \quad (4.20)$$

$$\langle M(r_k) \rangle = \frac{1}{N_{\text{meas}}} \sum_{j=1}^{N_{\text{meas}}} M(\mathbf{r}_{c_j}, r_k) \quad (4.21)$$

$$\langle R_g(r_k) \rangle = \frac{1}{N_{\text{meas}}} \sum_{j=1}^{N_{\text{meas}}} R_g(\mathbf{r}_{c_j}, r_k) \quad (4.22)$$

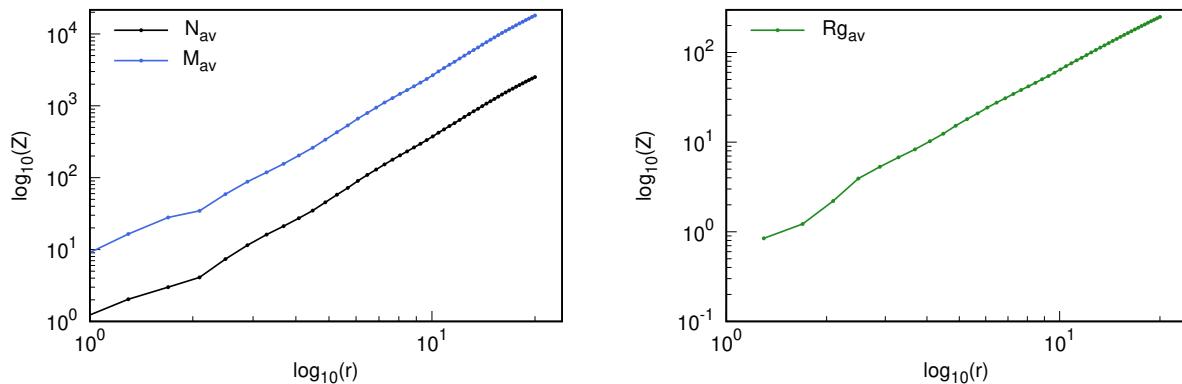
Donde  $N_{\text{meas}}$  es el número de centros aleatorios sobre los que se promedia.  $\mathbf{r}_{c_j}$  es el vector de posición del  $j$ -ésimo centro de medición y  $\langle \cdot \rangle$  es el promedio sobre todos los centros.

#### 4.1.6. Paso 10 del diagrama de flujo general

**Determinación de la dimensión fractal.** Para calcular la dimensión fractal de masa  $D$ , se asume un comportamiento de ley de potencia, ver ecuación 4.23.

$$\langle Z(r_k) \rangle \propto r_k^D \quad (4.23)$$

El exponente de escala  $D$  relaciona  $\langle Z(r) \rangle$  con  $r_k$  donde  $\langle Z(r) \rangle$  representa  $\langle M(r_k) \rangle$ ,  $\langle N(r_k) \rangle$  o  $\langle Rg(r_k) \rangle$ . Por último, se crea una gráfica de  $\log_{10} Z(r_k)$  vs  $\log_{10} r_k$  (Figura 4.3) y se aplica una regresión lineal en escala logarítmica.



**Figura 4.3:** Gráficos obtenido mediante el programa *molfractaldim* de  $\log_{10} r$  vs  $\log_{10} Z(r)$  donde  $\log_{10} Z(r)$  corresponde a  $\langle M(r_k) \rangle$ ,  $\langle N(r_k) \rangle$  o  $\langle Rg(r_k) \rangle$ .

## 4.2. Metodología

### 4.2.1. Densidad de partículas promedio $\rho_P$

| Idx | IdPDB | Nombre de proteína            | Idx | IdPDB | Nombre de proteína            |
|-----|-------|-------------------------------|-----|-------|-------------------------------|
| 1   | 11gs  | Glutatión S-transferasa       | 21  | 1a27  | 17 $\beta$ -hidroxiesteroides |
| 2   | 12ca  | Anhidrasa carbónica II        | 22  | 1a28  | Receptor de progesterona      |
| 3   | 13gs  | Glutatión S-transferasa       | 23  | 1a2b  | Proteína transformante RhoA   |
| 4   | 14gs  | Glutatión S-transferasa       | 24  | 1a2c  | Trombina                      |
| 5   | 16gs  | Glutatión S-transferasa P     | 25  | 1a31  | Topoisomerasa I               |
| 6   | 1a00  | Hemoglobina                   | 26  | 1a35  | Topoisomerasa I               |
| 7   | 1a01  | Hemoglobina                   | 27  | 1a36  | ADN topoisomerasa I           |
| 8   | 1a02  | Factor nuclear de células     | 28  | 1a3b  | Trombina                      |
| 9   | 1a07  | Tirosina quinasa src          | 29  | 1a3e  | Trombina                      |
| 10  | 1a08  | Tirosina quinasa src          | 30  | 1a3n  | Hemoglobina                   |
| 11  | 1a09  | Tirosina quinasa src          | 31  | 1a3o  | Hemoglobina                   |
| 12  | 1a0l  | Triptasa beta-1               | 32  | 1a3q  | Factor nuclear $\kappa$ -B    |
| 13  | 1a0u  | Hemoglobina                   | 33  | 1a42  | Anhidrasa carbónica II        |
| 14  | 1a0z  | Hemoglobina                   | 34  | 1a46  | Trombina                      |
| 15  | 1a12  | Regulador de cromosoma        | 35  | 1a4i  | Dihidrofolato reductasa       |
| 16  | 1a15  | Factor 1- $\alpha$ de estroma | 36  | 1a4r  | Proteína de unión G25K        |
| 17  | 1a17  | Fosfatasa ser/thr             | 37  | 1a4y  | Inhibidor de ribonucleasa     |
| 18  | 1alm  | Antígeno HLA clase I          | 38  | 1a52  | Receptor de estrógeno         |
| 19  | 1a1o  | Antígeno HLA clase I          | 39  | 1a66  | Núcleo NFATC1                 |
| 20  | 1a22  | Hormona de crecimiento        | 40  | 1a7a  | Adenosilhomocisteinasa        |

**Tabla 4.1:** Índice de proteínas (Idx), identificadores del *Protein Data Bank* (IdPDB) y nombre de las proteínas (Parte 1).

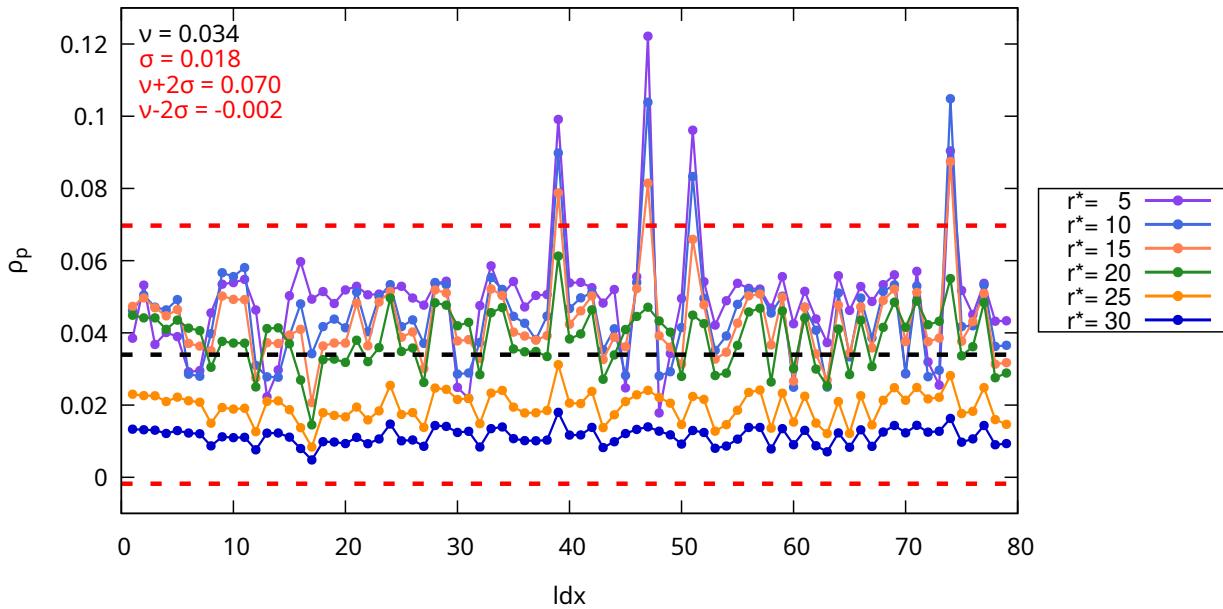
| Idx | IdPDB | Nombre de proteína                   | Idx | IdPDB | Nombre de proteína             |
|-----|-------|--------------------------------------|-----|-------|--------------------------------|
| 41  | 1a7c  | Inhibidor del activador plasminógeno | 61  | 1awb  | Mioinositol monofosfatasa      |
| 42  | 1a8m  | Factor de necrosis tumoral           | 62  | 1azx  | Antitrombina III               |
| 43  | 1a9b  | Antígeno HLA-B                       | 63  | 1b09  | Regulador de transcripción CRP |
| 44  | 1a9e  | Antígeno HLA-B                       | 64  | 1b3e  | Transferrina sérica            |
| 45  | 1a9w  | Hemoglobina                          | 65  | 1b3o  | Inosina monofosfato            |
| 46  | 1aax  | Tirosina fosfotasa 1B                | 66  | 1b41  | Acetilcolinesterasa            |
| 47  | 1ab2  | Tirosina quinasa C-ABL               | 67  | 1b47  | CBL                            |
| 48  | 1abw  | Hemoglobina                          | 68  | 1b59  | Metionina aminopeptidasa       |
| 49  | 1aby  | Hemoglobina                          | 69  | 1b5g  | Trombina $\alpha$              |
| 50  | 1agn  | Alcohol deshidrogenasa               | 70  | 1b86  | Hemoglobina                    |
| 51  | 1ah1  | CTLA-4                               | 71  | 1ba8  | Trombina                       |
| 52  | 1ald  | Aldolasa fructosa-bifosfato          | 72  | 1bab  | Hemoglobina                    |
| 53  | 1ant  | Antitrombina-III                     | 73  | 1bbb  | Hemoglobina                    |
| 54  | 1aos  | Argininosuccinato liasa              | 74  | 1bcn  | Interleucina-4                 |
| 55  | 1ap5  | Superóxido dismutasa                 | 75  | 1bhg  | Beta-glucuronidasasa           |
| 56  | 1apy  | Aspartilglucosaminidasa              | 76  | 1bhs  | 17- $\beta$ -deshidrogenasa    |
| 57  | 1apz  | Aspartilglucosaminidasa              | 77  | 1bxh  | Trombina                       |
| 58  | 1aud  | U1A 102                              | 78  | 6oj0  | Proteína estructural VP4       |
| 59  | 1auk  | Arilsulfatasa A                      | 79  | 7khw  | Translocon EspA                |
| 60  | 1avo  | Regulador 11S                        |     |       |                                |

**Tabla 4.2:** Índice de proteínas (Idx), identificadores del *Protein Data Bank* (IdPDB) y nombre de las proteínas (Parte 2).

La implementación del criterio, densidad de partículas promedio  $\rho_P$  (ver subsección 4.1.3) fue realizada seleccionando 79 proteínas, véase las Tablas 4.1 y 4.2 (provenientes del *Protein Data Bank* con fecha de acceso: agosto de 2024) y con el uso del programa *molmassfractaldim* (sin la aplicación del bucle 6  $\longleftrightarrow$  7 del diagrama de flujo de la Figura 4.1) se procesó cada proteína mencionada.

Una vez completada esta fase, se extrajeron los valores de los radios  $r^* = 5, 10, 15, 20, 25, 30 \text{ \AA}$  junto con el número de partículas promedio  $\langle N(r^*) \rangle$  de las datos generados a partir del programa *molmassfractaldim*. Con esta información, se procedió al cálculo de volúmenes y densidades de partículas promedio presentes en las 79 proteínas.

A continuación, se creó una gráfica para observar la distribución de los valores de  $\rho_P$  en función del índice de la proteína *ldx* de las Tablas 4.1 y 4.2, así como la densidad doblemente promediada, es decir,  $v \equiv \bar{\rho}_P$  y la desviación estándar. Tanto  $v$  como  $\sigma$  se calcularon sobre los distintos  $\rho_P(r^*)$ .



**Figura 4.4:** Distribución de la densidad de partículas promedio ( $\rho_P$ ) en función del índice de proteínas (*ldx*) para los archivos *PDB* de las Tablas 4.1 y 4.2. La línea azul horizontal discontinua representa la media ( $v$ ) de  $\rho_P$ , mientras que las líneas rojas horizontales discontinuas indican los límites de dos desviaciones estándar ( $v \pm 2\sigma$ ) de ( $\rho_P$ ).

De la Figura 4.4 se observó que la mayoría de las proteínas de las Tablas 4.1 y 4.2 presentan una  $\rho_P$  entre 0.002 y 0.070. Se distingue, además, una distribución moderadamente sesgada de  $\rho_P$  hacia valores de radios menores ( $r^* \leq 20$ ). Es decir, los valores de  $\rho_P$  se encuentran cerca de la media (0.034) especialmente con valores de  $r^* = 5, 10, 15, 20$ . Pero existe un rango considerable que se extiende hacia valores menores a partir de  $r^* = 25, 30$ .

Se puede inferir que  $\rho_p$ , disminuye conforme aumenta el radio. cuyo comportamiento resulta consistente con el esperado en sistemas físicos, donde las partículas tienden a estar más densamente distribuidas en las regiones próximas al centro de un volumen esférico.

Es importante mencionar que pocos valores caen fuera del intervalo  $v \pm 2\sigma$  (11 valores con  $r^* = 5, 10, 15$ ) por lo tanto, las variaciones significativas son limitadas y pueden representar anomalías o efectos físicos menos frecuentes en tales proteínas.

De la información anterior, rescatamos que al analizar una proteína arbitraria, se puede utilizar el valor de la densidad de partículas  $\rho_p \geq 0.34$  como criterio para determinar cuándo un centro no es adecuado para realizar la medición de la dimensión fractal de masa.

#### 4.2.2. Pruebas de fractalidad en proteínas

El análisis fractal fue realizado tomando como punto de partida 9 de las 79 estructuras que se utilizaron anteriormente aplicando la restricción de la densidad de partículas promedio ( $\rho_p \geq 0.034$ ) como criterio para determinar cuando un centro no es adecuado para realizar el conteo de partículas y posteriormente, la cuantificación de la masa como función del radio de las esferas. Además, se hizo uso del programa de dinámica molecular *GROMACS* [35, 36] que permite realizar simulaciones moleculares, con el fin de comparar diferentes estados de los sistemas proteicos. Los pasos esenciales de este proceso fueron los siguientes: (1) Preparación de la topología del sistema. (2) Creación de la caja y solvatación, (3) adición de iones para la neutralización de la carga, (4) minimización de la energía del sistema, (5) equilibrio del sistema, (6) simulación de producción y (7) resultados.

Hecho lo anterior, se extrajeron las estructuras correspondientes a la última geometría de los siguientes procesos, para una comparación posterior.

1. Después de adicionar átomos de hidrógeno al sistema proteico.
2. Tras finalizar el proceso de minimización de la energía.
3. Luego de completarse la etapa de equilibrio del sistema.
4. Completada la dinámica molecular durante un 1 ns.

#### 4.2.3. Elección de sistemas

| Idx | IdPDB | Nombre de proteína                             | Número de átomos |
|-----|-------|--|------------------|
| 1   | 1a2b  | Transformadora RHOA                            | 2831             |
| 2   | 1b3e  | Transferrina sérica                            | 5037             |
| 3   | 11gs  | Glutatión s-transferasa                        | 6536             |
| 4   | 1auk  | Arilsulfatasa A                                | 7086             |
| 5   | 1a8m  | Factor de necrosis tumoral $\alpha$            | 7104             |
| 6   | 1a52  | Receptor de estrógeno                          | 7765             |
| 7   | 1a3n  | Hemoglobina humana desoxidante cadena $\alpha$ | 8734             |
| 8   | 1a9w  | Hemoglobina embrionaria cadena $\alpha$        | 8820             |
| 9   | 7khw  | Translocon EspA                                | 131200           |

**Tabla 4.3:** Nuevo índice de proteínas (Idx), Identificadores del *Protein Data Bank* (PDB), nombre de la proteína y número de átomos presentes en cada estructura.

La selección de los sistemas se realizó aleatoriamente (utilizando como punto de partida las 79 proteínas de las Tablas 4.1, 4.2 y creando un nuevo índice de proteínas para cada una de ellas), bajo la condición de que todos sistemas contaran con la secuencia completa de aminoácidos y evitando componentes externos como ligandos o solventes que introducirían ruido en el cálculo de la dimensión fractal masa intrínseca de las proteínas. Los sistemas analizados se listan en la Tabla 4.3.

#### 4.2.4. Elección de medida de análisis

Nuestro estudio se ha enfocado en analizar la dimensión fractal de masa  $M(r)$  debido a las siguientes razones:

1. La dimensión fractal de masa  $M(r)$  captura la distribución real de masa en un sistema.

En una proteína, se presentan átomos que tienen masas atómicas muy diferentes, además de localizadas y centradas en los núcleos (ejemplo de lo anterior son: H  $\sim 1\text{ g mol}^{-1}$ , C  $\sim 12\text{ g mol}^{-1}$ , N  $\sim 14\text{ g mol}^{-1}$ , O  $\sim 16\text{ g mol}^{-1}$ , S  $\sim 32\text{ g mol}^{-1}$ ). Ignorar esta heterogeneidad, como hace la dimensión fractal del número de partículas  $N(r)$ , arrojaría un análisis incompleto o diferente. Dos posibles escenarios al realizar una medida de  $M(r)$  y  $N(r)$  son los siguientes:

- Una zona con gran cantidad de átomos de hidrógeno (baja masa) contribuirá mayormente a la medida de  $N(r)$  pero menor a la medida  $M(r)$ .
- Una región densa con átomos de azufre o con la presencia de un grupo hemo contribuirá masivamente a  $M(r)$  pero no a  $N(r)$ .

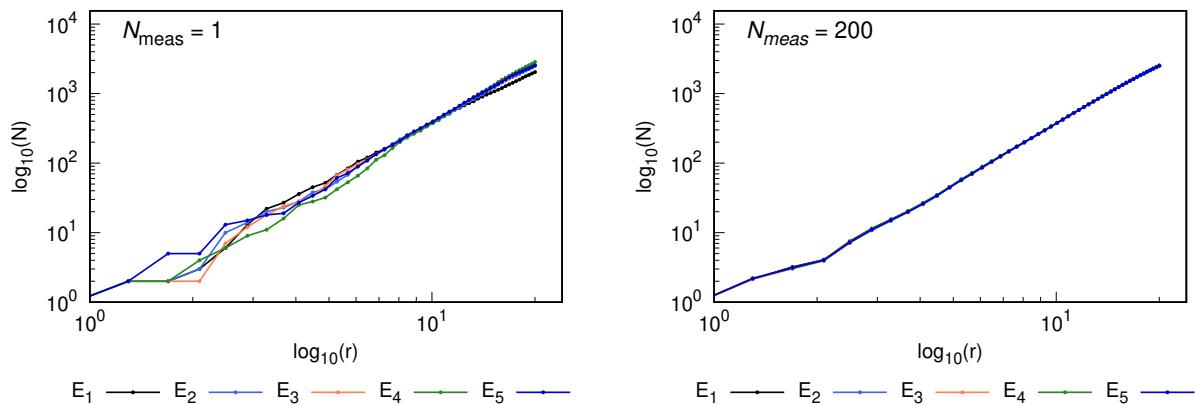
2. Relación con multifractalidad.

La multifractalidad surge de la heterogeneidad en un objeto de análisis. Al usar la masa como medida  $M(r)$ , se asegura una heterogeneidad adicional y físicamente significativa respecto a solo contar partículas ( $N(r)$ , es decir, tratar a todos los átomos por igual). Por lo tanto, la medida de  $M(r)$  incluye información de las masas atómicas.

#### 4.2.5. Determinación de $M(r)$

Se calculó la masa promedio de las partículas contenidas dentro de las esferas de radio  $r_k$ , definida como  $\langle M(r_k) \rangle$ . Para llevar a cabo este análisis, se hizo uso de la herramienta *molfractaldim*. Ajustando dos parámetros clave: (1)  $N_{meas}$  (ecuación 4.19) que define cuántos valores se promedian para obtener el valor de  $\langle M(r_k) \rangle$ , cuando  $N_{meas} = 1$ , los valores promediados presentan una mayor variabilidad, como puede observarse en la primera gráfica de la Figura 4.5. En cambio, al aumentar el número de mediciones a  $N_{meas} = 200$ , las fluctuaciones disminuyen considerablemente, y los valores promediados tienden a ser muy similares entre sí, como se muestra en la segunda gráfica de la Figura 4.5. Y (2)  $r_{max}$  (ecuación 4.11) que determina cuántos radios  $r_k$  se toman para la medición.

En este análisis  $N_{meas} = 200$  y  $r_{max} = 200$  fue para asegurar resultados representativos y detallados. Este procedimiento se aplicó en cada uno de los sistemas de la Tabla 4.3, correspondientes a los pasos descritos en la sección 4.2.2.



**Figura 4.5:** Gráficos de  $\log_{10} r$  vs  $\log_{10} N(r)$  correspondientes la proteína *Tubulina mutada* después de una dinámica molecular de 10 ns, estableciendo el número de medidas en  $N_{meas} = 1$  y 200 (izquierda a derecha). Donde  $N_{meas}$  establece el número de medidas para obtener  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle Rg(r_k) \rangle$ .

#### 4.2.6. Análisis multifractal

Se realizó un análisis multifractal sobre los datos extraídos del paso anterior, para calcular la dimensión fractal de masa  $D$ , donde el exponente de escala  $b$  relaciona la masa promedio  $\langle M(r_k) \rangle$  con el radio  $r_k$  mediante la ecuación 4.24.

$$\langle M(r_k) \rangle \sim r_k^b \quad (4.24)$$

Para un sistema multifractal,  $b$  varía según la escala espacial, reflejando heterogeneidad en la distribución de masa.

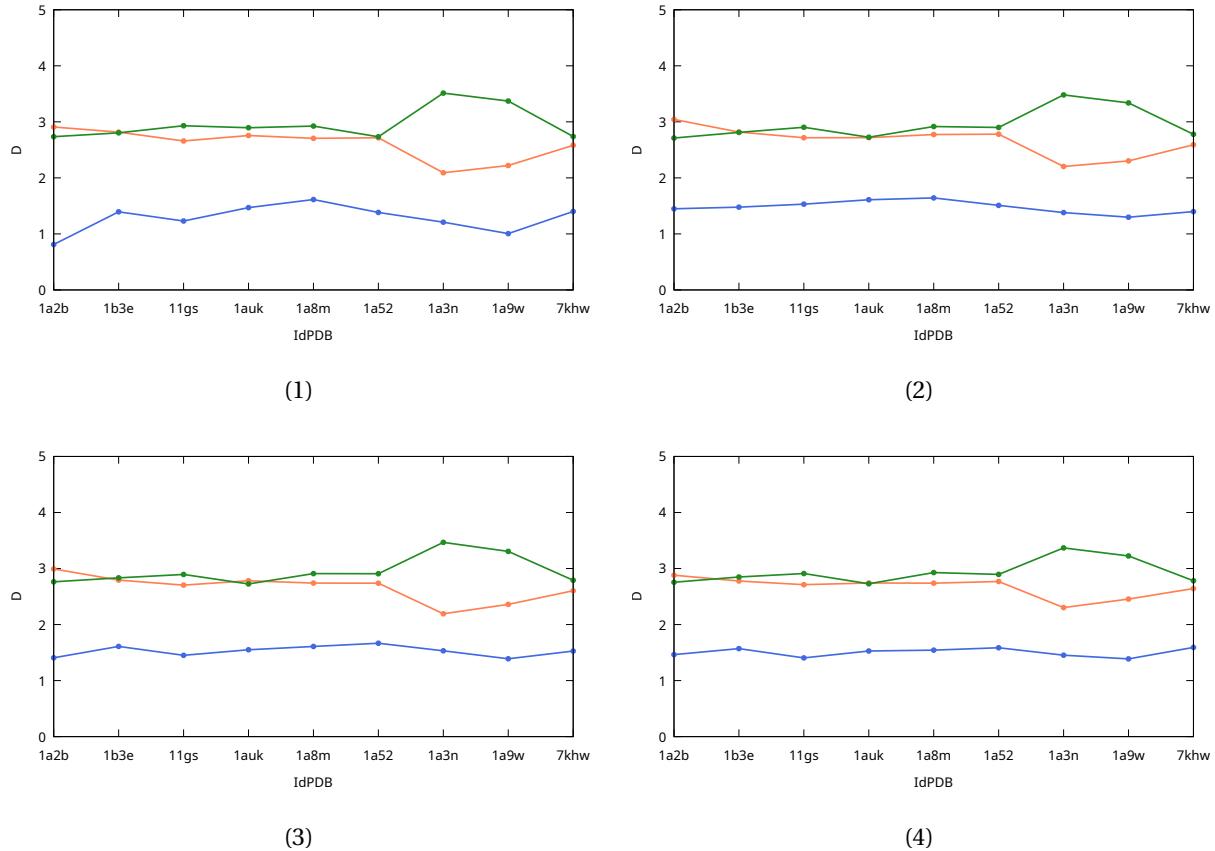
■ **Pendiente  $b$ :**

- Si  $b$  muestra un único valor en el intervalo general 1–20 Å: Existe un comportamiento monofractal (estructura homogénea).
- Si  $b$  tiene distintos valores en el intervalo general de 1–20 Å: Existe multifractalidad (estructura heterogénea).

A continuación, se llevaron a cabo múltiples regresiones lineales sobre cada archivo de datos numéricos, en intervalos donde era altamente probable encontrar distintos valores de dimensión fractal. Estos intervalos fueron: 1 a 2 Å, 2 a 6 Å y 6 a 20 Å [1, 37].

La selección del intervalo 2–6 Å utilizado para la medición o implementación de ajustes lineales en el análisis multifractal tiene su origen en el diámetro promedio de los aminoácidos ( $d_p$ ) reportado por Siyuan Liu *et al.* [37] siendo  $d_p \approx 4,188$  Å. Posteriormente, la elección de los intervalos parciales 1–2 Å y 6–20 Å fueron elegidos heurísticamente, observando las gráficas de  $\log_{10} r$  vs  $\log_{10} M(r)$  de cada proteína de la Tabla 4.3.

### 4.3. Discusión de resultados



**Figura 4.6:** Valores de dimensión fractal de masa  $D$  vs identificadores de proteínas (IdPDB) para las proteínas ilustradas en la Tabla 4.3, correspondientes a cuatro etapas de procesamiento: (1) Adición de átomos de hidrógeno al sistema proteico, (2) al minimizar la energía de la estructura molecular, (3) equilibrando el sistema bajo condiciones termodinámicas controladas y (4) después de una dinámica molecular de 1 ns. Las líneas azules, verdes y rojas representan  $D$  en los intervalos de 1 a 2 Å, 2 a 6 Å y 6 a 20 Å, respectivamente.

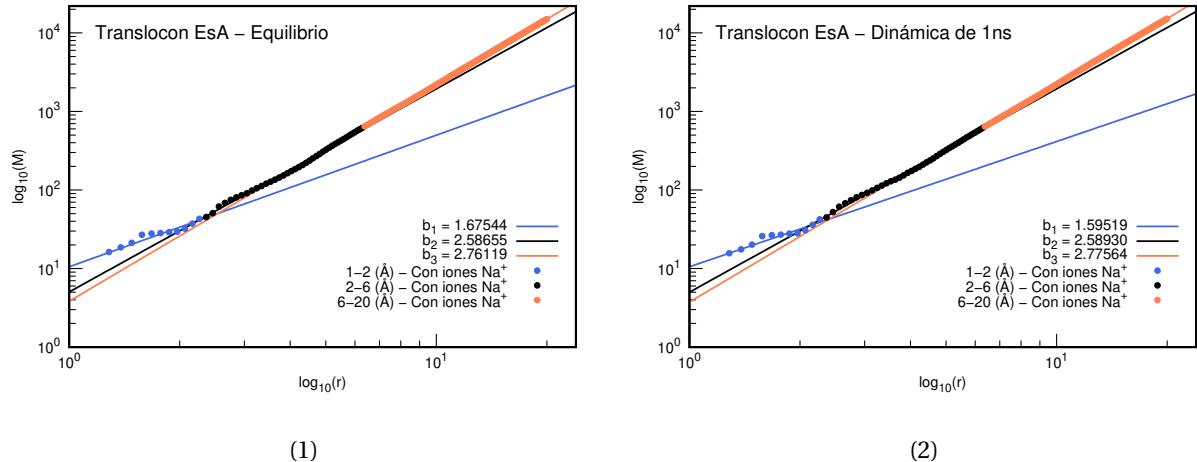
#### 4.3.1. Dimensiones fractales en distintas etapas

En la Figura 4.6, las proteínas Hemoglobina humana desoxidante cadena  $\alpha$  y Hemoglobina embrionaria cadena  $\alpha$  (1a3n y 1a9w) exhiben valores de dimensión fractal  $D \approx 3.5$  en todas las etapas, con una reducción ( $\Delta D < 0.2$ ) posterior al equilibrio termodinámico y una dinámica molecular de 1ns. El valor  $D \approx 3.5$  es contraintuitivo, pero podría deberse a que nuestro estudio se ha enfocado en analizar la dimensión fractal de masa (se toman en cuenta las masas promedio de los átomos en la medida) y no la dimensión fractal del número de partículas (es decir, que no depende de la masa promedio de las partículas presentes en la estructura proteica). Además, se identificó una correlación inversa (*relación tipo espejo*) entre los intervalos 2–6 Å y 6–20 Å, particularmente marcada durante la etapa de equilibrio del sistema.

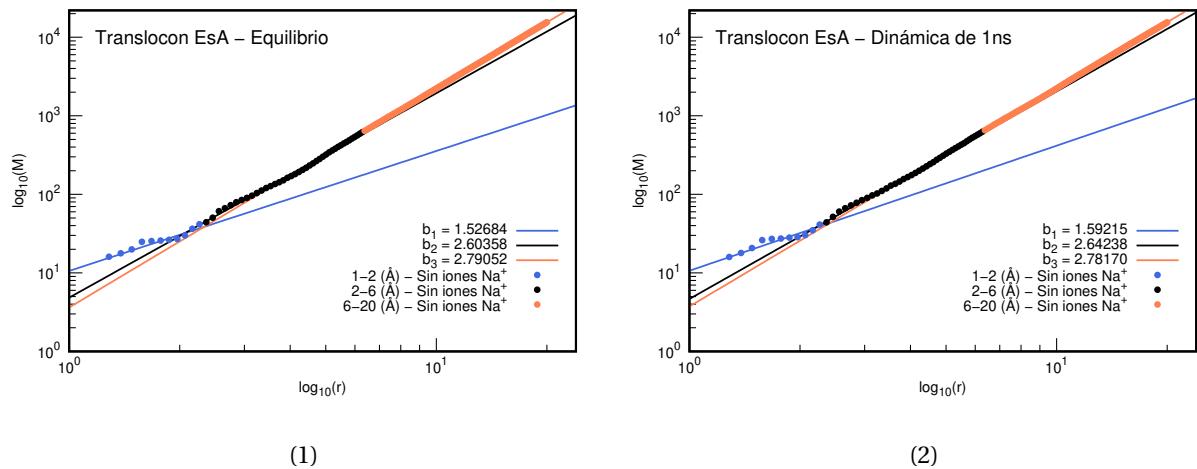
Este comportamiento sugiere la coexistencia de al menos, dos dimensiones fractales diferenciadas en las proteínas: (1a2b) Transformadora RHOA, (1b3e) Transferrina sérica, (11gs) Glutatión s-transferasa, (1auk) Arilsulfatasa A, (1a8m) Factor en necrosis tumoral  $\alpha$ , (1a52) Receptor de estrógeno y (7khw) Translocon EspA en los intervalos de 1–2 Å y 2–6 Å que se muestran en todas las fases de simulación. La variación de  $D$  evidenció un comportamiento multifractal, donde la dimensión fractal de masa evoluciona progresivamente desde organizaciones locales hasta globales. Estos hallazgos confirman que cada etapa computacional (adicción, minimización, equilibrio y dinámica molecular) inciden en la reconfiguración de la dimensión fractal proteica debido al tamaño del sistema.

Es importante mencionar que en los sistemas proteicos: (1a3n) Hemoglobina humana desoxidante cadena  $\alpha$  y (1a9w) Hemoglobina embrionaria cadena  $\alpha$ , tienen una estructura cuaternaria, es decir, para llevar a cabo su función requieren de cuatro monómeros perfectamente plegados que se unen por interacciones no covalentes para adquirir un carácter funcional y a diferencia de los siete sistemas anteriores, (1a3n) y (1a9w) muestran tres dimensiones fractales en los intervalos de 1–2 Å, 2–6 Å y 6–20 Å persistentes en todas las fases de simulación. Lo anterior sugiere que la medición podría ser lo suficientemente sensible como para detectar niveles superiores de organización proteica.

### 4.3.2. Efecto de iones en la dimensión fractal



**Figura 4.7:** Gráficos de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondientes a dos etapas de procesamiento de la proteína *Translocon EsA* de la Tabla 4.3 (con iones  $\text{Na}^+$  en su estructura) después de (1) equilibrar el sistema bajo condiciones termodinámicas controladas y (2) de una dinámica molecular de 1 ns. Los valores de  $b$  en cada gráfica representan la dimensión fractal de masa en los intervalos 1 a 2  $\text{\AA}$ , 2 a 6  $\text{\AA}$  y 6 a 20  $\text{\AA}$ .



**Figura 4.8:** Gráficos de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondientes a dos etapas de procesamiento de la proteína *Translocon EsA* de la Tabla 4.3 (sin iones  $\text{Na}^+$  en su estructura) de la Tabla 4.3 después de (1) equilibrar el sistema bajo condiciones termodinámicas controladas y (2) de una dinámica molecular de 1 ns. Los valores de  $b$  en cada gráfica representan la dimensión fractal de masa en los intervalos 1 a 2  $\text{\AA}$ , 2 a 6  $\text{\AA}$  y 6 a 20  $\text{\AA}$ .

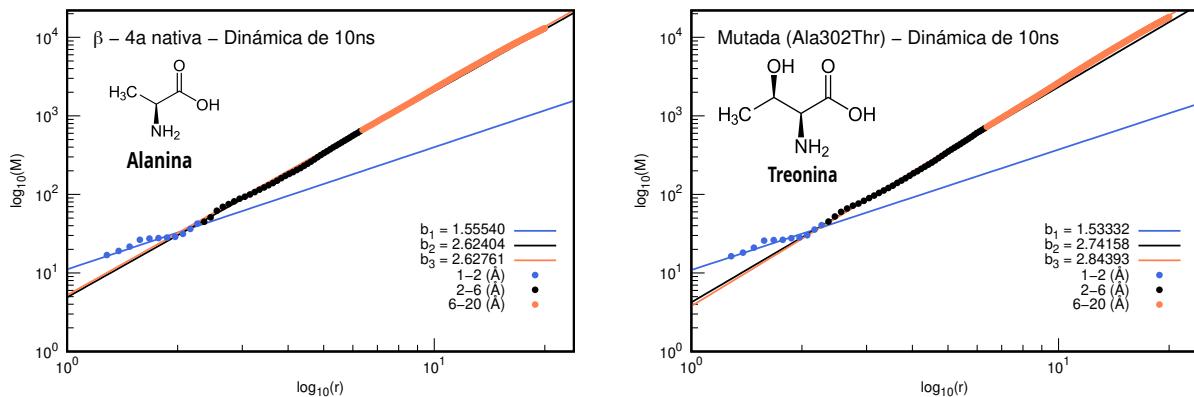
**Proteína Translocon EsA (idPDB: 7khw)**

| <b>Estado de la proteína</b> | <b>Etapa</b>   | <b>Intervalo (Å)</b> | $R^2$    | <b>Pendiente (b)</b> |
|------------------------------|----------------|----------------------|----------|----------------------|
| Con $\text{Na}^+$            | Equilibrio     | 1-2                  | 0.933264 | 1.67544              |
|                              |                | 2-6                  | 0.997049 | 2.58655              |
|                              |                | 6-20                 | 0.999965 | 2.76119              |
|                              | Dinámica (1ns) | 1-2                  | 0.938494 | 1.59519              |
|                              |                | 2-6                  | 0.997471 | 2.5893               |
|                              |                | 6-20                 | 0.999977 | 2.77564              |
|                              | Equilibrio     | 1-2                  | 0.930791 | 1.52684              |
|                              |                | 2-6                  | 0.996961 | 2.60358              |
|                              |                | 6-20                 | 0.999988 | 2.79052              |
| Sin $\text{Na}^+$            | Dinámica (1ns) | 1-2                  | 0.929409 | 1.59215              |
|                              |                | 2-6                  | 0.99807  | 2.64238              |
|                              |                | 6-20                 | 0.999938 | 2.78170              |

**Tabla 4.4:** Resumen de datos de las regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  de las Figuras 4.7 y 4.8.

En las Figuras 4.7 y 4.8, se compara la distribución de masa de la proteína *Translocon EsA* (7khw) en dos etapas: Tras el equilibrio estructural y luego de una dinámica molecular de 1 ns. La Figura 4.7 considera la estructura proteica con iones presentes, revelando tres pendientes en la regresión de  $\log_{10} r$  vs  $\log_{10} M(r)$ , lo que sugiere un comportamiento multi-fractal en 3 intervalos. Sin embargo, en la Figura 4.8, 250 iones de sodio que no pertenecían a la estructura intrínseca de la proteína fueron eliminados, con el fin de validar la reproducibilidad de los resultados. Sorprendentemente, en el último caso solo se observaron dos pendientes, aunque la estructura original contenía un excedente de 250 átomos de sodio (aproximadamente el 0.19 % comparado con los 131,200 átomos presentes en la estructura proteica). Por lo tanto, esta comparación muestra que la estimación de la dimensión fractal de masa es sensible a modificaciones estructurales mínimas.

### 4.3.3. Caso particular de estudio:tubulinas



**Figura 4.9:** Gráficos de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondientes a dos proteínas diferentes: *Tubulina nativa* (alanina) y *Tubulina mutada* (treonina) después de una dinámica molecular de 10 ns. Los valores de  $b$  en cada gráfica representan la dimensión fractal de masa en los intervalos 1 a 2 Å, 2 a 6 Å y 6 a 20 Å.

La comparación presentada en la Figura 4.9 examina dos variantes estructurales: (1) Una proteína nativa con 6,812 átomos (estructura de referencia) y (2) una variante mutada con 6,816 átomos (reemplazo de un residuo aminoacídico).

Los resultados obtenidos demostraron diferencias en las pendientes de la regresión lineal analizadas específicamente en la etapa final del procesamiento (después de una dinámica molecular de 10 ns). Esta diferencia estructural se observa a pesar de la alta similitud global entre ambas variantes. La detección de estas diferencias sutiles, valida la sensibilidad del método propuesto para discriminar modificaciones estructurales mínimas en sistemas proteicos.

Con base en la evidencia discutida en esta subsección y las anteriores (subsección 4.3.1 a 4.3.3), se puede establecer que las proteínas, en general, presentan un comportamiento multifractal. Este fenómeno se manifiesta en intervalos específicos, los cuales han sido caracterizados en el análisis. Las proteínas restantes del estudio exhiben un comportamiento análogo. Para una revisión detallada, se remite al lector al Apéndice B para examinar los detalles particulares de cada sistema.

Capítulo **5**

## Conclusión

Este trabajo establece que las proteínas presentan un comportamiento multifractal. Esta multifractalidad puede correlacionarse con la información dinámica y funcional de las proteínas. Se especula que en el proceso evolutivo de las proteínas la multifractalidad es una propiedad que proporciona un nuevo marco cuantitativo para entender su complejidad jerárquica. En conclusión:

- La medida del cálculo de la dimensión fractal de masa demuestra sensibilidad en cambios estructurales. Por consiguiente, podría utilizarse como medida de caracterización estructural de las proteínas.
- Las nueve proteínas analizadas exhiben multifractalidad, evidenciada por la variación significativa en  $D$  entre los intervalos 1–2 Å, 2–6 Å y 6–20 Å.
- Estructuras cuaternarias como las proteínas *Hemoglobin humana* y *Hemoglobina embrionaria* podrían arrojar una dimensión fractal de masa mayor a tres ( $D > 3$ ).
- La dimensión fractal de masa puede cuantificar los diferentes ambientes fisicoquímicos simulados durante una dinámica molecular, lo cual se manifiesta en las modificaciones estructurales de las conformaciones proteicas evaluadas en distintos estados del sistema.

# Perspectivas

Establecida la base de los resultados que confirman la existencia de multifractalidad en sistemas proteicos y la posibilidad de utilizar la dimensión fractal como medida caracterizadora de la estructura en las proteínas, se plantean las siguientes líneas de investigación:

1. **Análisis de la dimensión fractal de masa en función del radio de giro ( $R_g$ ).**
2. **Caracterización de la multifractalidad en sistemas proteicos exclusivamente en átomos pesados.** La eliminación de átomos de hidrógeno permitirá aislar la contribución del esqueleto proteico y las cadenas laterales masivas a la multifractalidad. Esto podría ayudar a discriminar entre la heterogeneidad estructural intrínseca y la aportada por la distribución de los átomos ligeros, simplificando el sistema para un modelado más preciso.
3. **Evaluación de la distribución fractal de átomos ligeros y su relación con la formación de puentes de hidrógeno.** Un estudio centrado en la dimensión fractal de átomos de hidrógeno, podría revelar patrones de agregación espacial asociados a la formación de puentes de hidrógeno. La hipótesis nace de las regiones con una alta densidad de dichos puentes y podrían exhibir un valor de dimensión fractal de masa significativamente diferente, reflejando su papel en el plegamiento y la estabilidad de estructuras secundarias y terciarias.

## Funcionamiento del programa

El propósito principal del programa *molmassfractaldim* es determinar la dimensión fractal de masa mediante la caracterización de la geometría interna de un sistema proteico. Esto se logra construyendo una tabla de  $r_k$ ,  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle Rg(r_k) \rangle$ : donde;

- $r_k$  es el radio de medida centrado en posiciones aleatorias de una proteína con valores mínimos y máximos definidos como  $r_{min}$  y  $r_{max}$ .
- $\langle N(r_k) \rangle$  es el número promedio de partículas contenidas dentro de un radio  $r_k$ .
- $\langle M(r_k) \rangle$  es el número de masa promedio de las partículas contenidas en el radio  $r_k$ .
- $\langle Rg(r_k) \rangle$  es el radio de giro promedio de las partículas contenidas en el radio  $r_k$ .

Posteriormente, los valores de  $\langle N(r_k) \rangle$ ,  $\langle M(r_k) \rangle$  y  $\langle Rg(r_k) \rangle$  pueden analizarse mediante regresiones lineales para determinar la dimensión fractal de masa.

## Banderas de entrada

El programa *molmassfractaldim*, incorpora banderas de entrada al código para tener distintas opciones en el cálculo de la dimensión fractal de masa. Las banderas incorporadas se muestran en la Tabla A.1.

| Opción | Descripción  |
|--------|--|
| -a A   | Establece el número de medidas a realizar. Esto se refiere a la cantidad de veces que el programa calculará el promedio de $\langle N(r_k) \rangle$ , $\langle M(r_k) \rangle$ y $\langle R_g(r_k) \rangle$ . Es útil si se desea promediar los resultados sobre diferentes configuraciones o pruebas. Para más detalles sobre esta opción, véase la sección ??. |
| -n k   | Define el número de radios $r_k$ que se utilizarán. Es decir, el número de filas en la tabla de $r_k$ , $\langle N(r_k) \rangle$ , $\langle M(r_k) \rangle$ y $\langle Rg(r_k) \rangle$ .  |
| -l     | Permite espaciar los radios multiplicativamente. Es decir, los radios de medición están igualmente separados en una escala logarítmica, en lugar de una escala lineal.   |
| -q     | Modo silencioso. Al activar esta opción, la mayoría de los mensajes informativos que normalmente se mostrarían en la salida estándar se suprimirán.  |
| -r rmn | Establece el radio mínimo $r_{min}$ para las mediciones. Esto es útil para limitar el rango de los radios de las esferas que se tomarán en cuenta en los cálculos; $r$ se refiere al radio de la esfera.   |
| -R rmx | Establece el radio máximo $r_{max}$ para las mediciones. Similar a la opción anterior, permite definir el rango máximo de los radios de las esferas.   |
| -V     | Muestra la versión del programa.   |
| -h     | Muestra el menú de ayuda con la descripción de las opciones y el uso del programa.   |

---

**Tabla A.1:** Banderas de entrada para el programa *molmassfractaldim*.

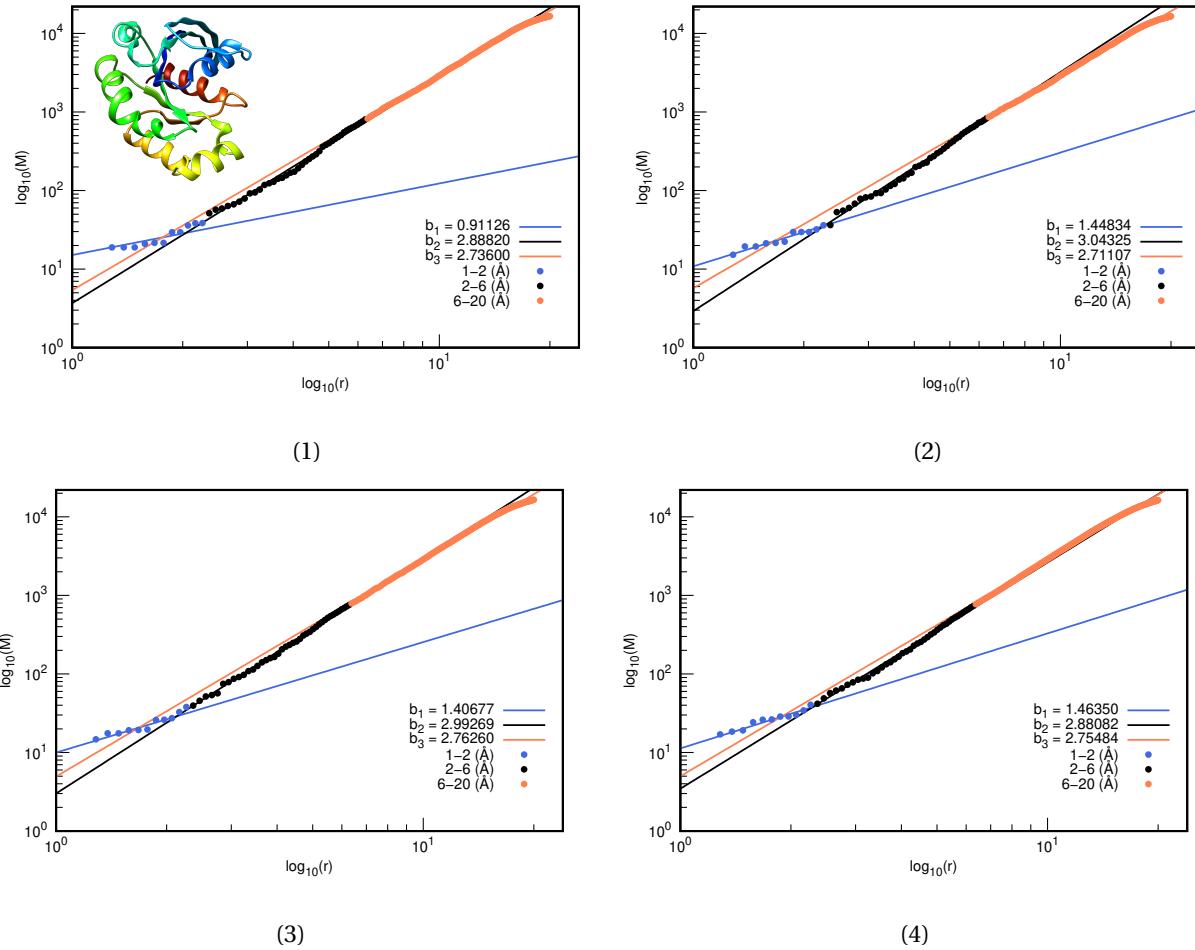
## Conjunto de 9 proteínas

El análisis global del conjunto proteico (9 proteínas) presentado en la sección 4.2.2, demuestran un comportamiento multifractal en toda la muestra. Dado que este patrón se manifiesta de manera consistente (aunque con una intensidad variable) en todos los sistemas, el análisis individual no aportan información adicional para verificar la existencia del fenómeno. Por consiguiente, se ha decidido ubicar los resultados en este apéndice (junto con una imagen de cada sistema), donde el lector podrá examinar los detalles particulares de cada sistema. Para comodidad del lector, nuevamente se expone la Tabla 4.3.

| Idx | IdPDB | Nombre de proteína                             | Número de átomos |
|-----|-------|--|------------------|
| 1   | 1a2b  | Transformadora RHOA                            | 2831             |
| 2   | 1b3e  | Transferrina sérica                            | 5037             |
| 3   | 1lgs  | Glutatión s-transferasa                        | 6536             |
| 4   | 1auk  | Arilsulfatasa A                                | 7086             |
| 5   | 1a8m  | Factor de necrosis tumoral $\alpha$            | 7104             |
| 6   | 1a52  | Receptor de estrógeno                          | 7765             |
| 7   | 1a3n  | Hemoglobina humana desoxidante cadena $\alpha$ | 8734             |
| 8   | 1a9w  | Hemoglobina embrionaria cadena $\alpha$        | 8820             |
| 9   | 7khw  | Translocon EspA                                | 131200           |

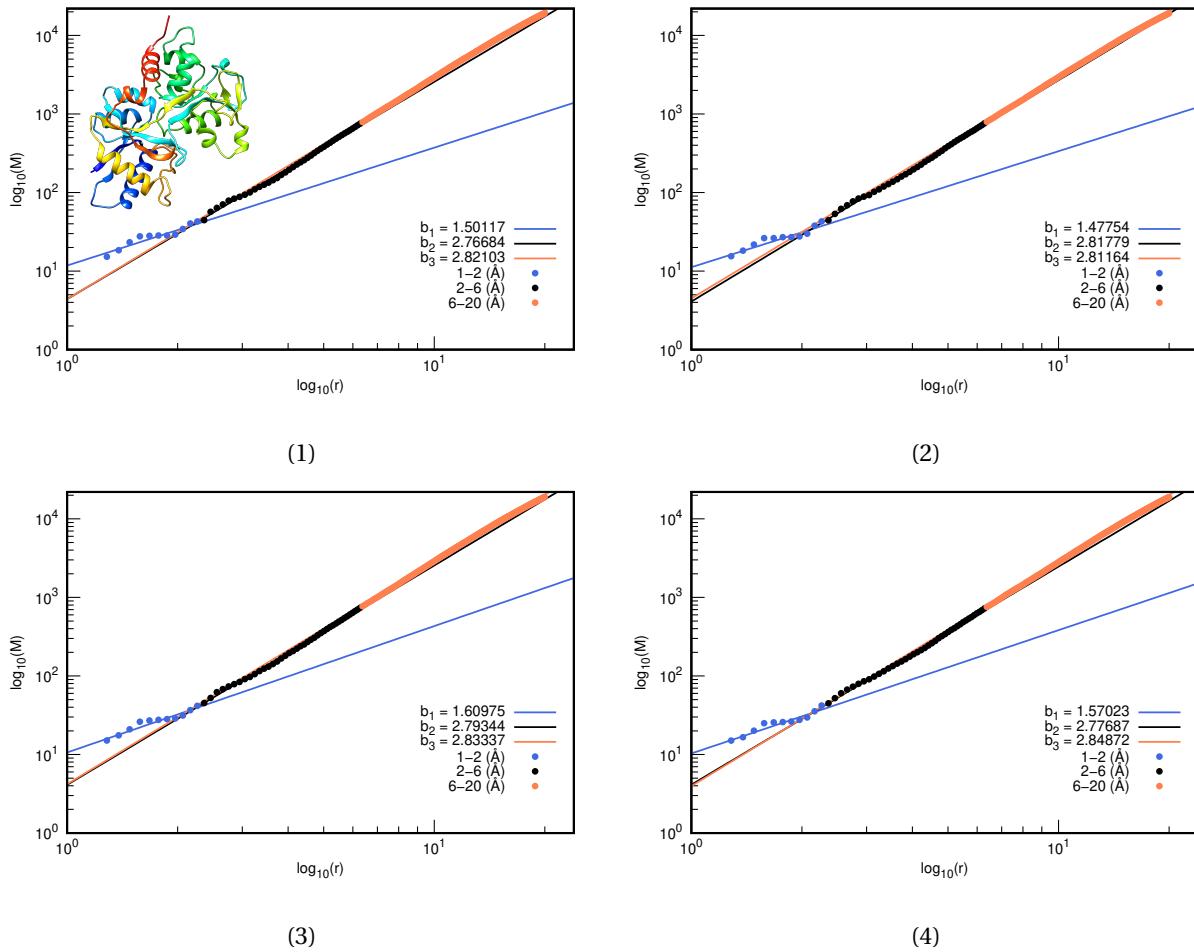
**Tabla B.1:** Nuevo índice de proteínas (Idx), Identificadores del *Protein Data Bank* (PDB), nombre de la proteína y número de átomos presentes en cada estructura.

### Proteína Transformadora RHOA (IdPDB:1a2b)



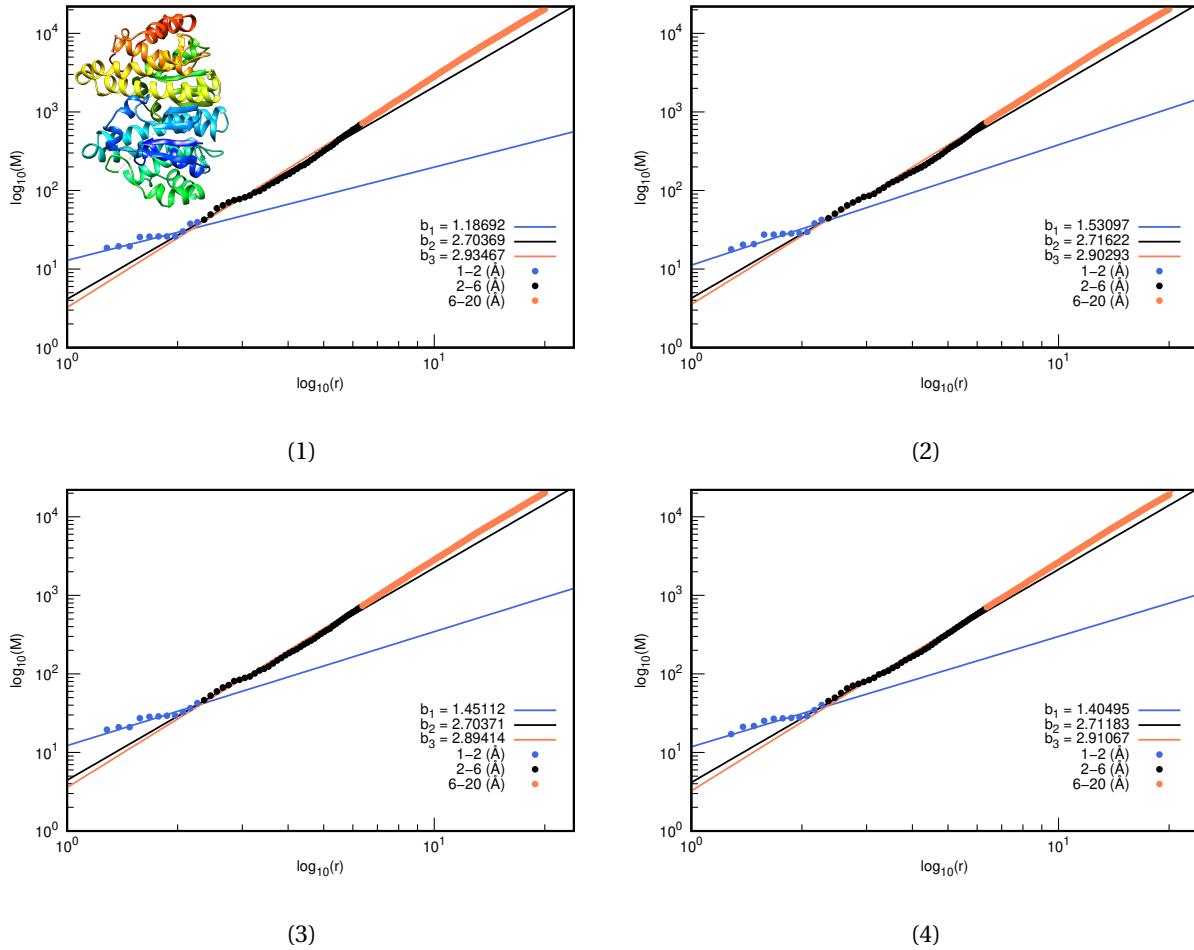
**Figura B.1:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la primera proteína con *IdPDB:1a2b* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

## Proteína Transferrina sérica (IdPDB:1b3e)



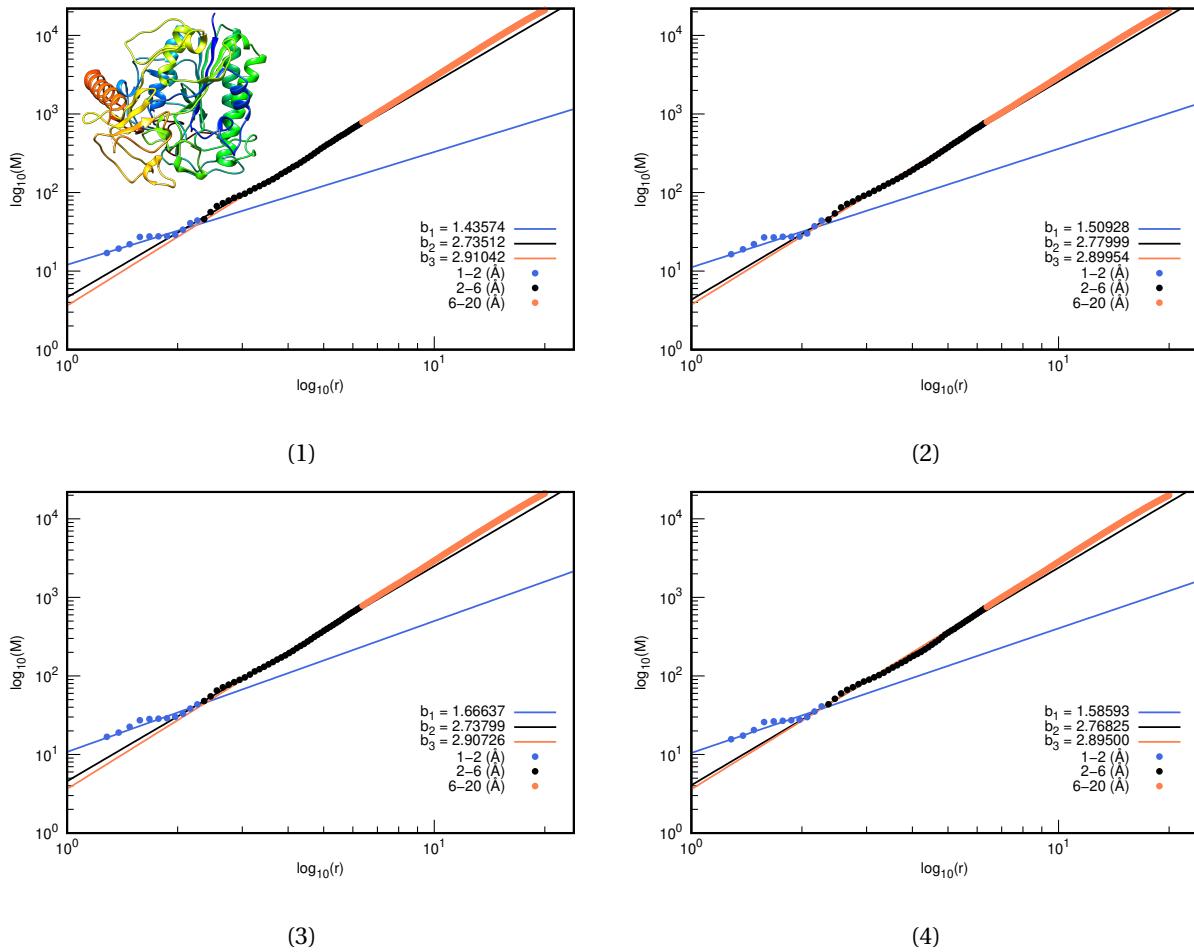
**Figura B.2:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la séptima proteína con *IdPDB:1b3e* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

### Proteína Glutatióñ s-transferasa (IdPDB:11gs)



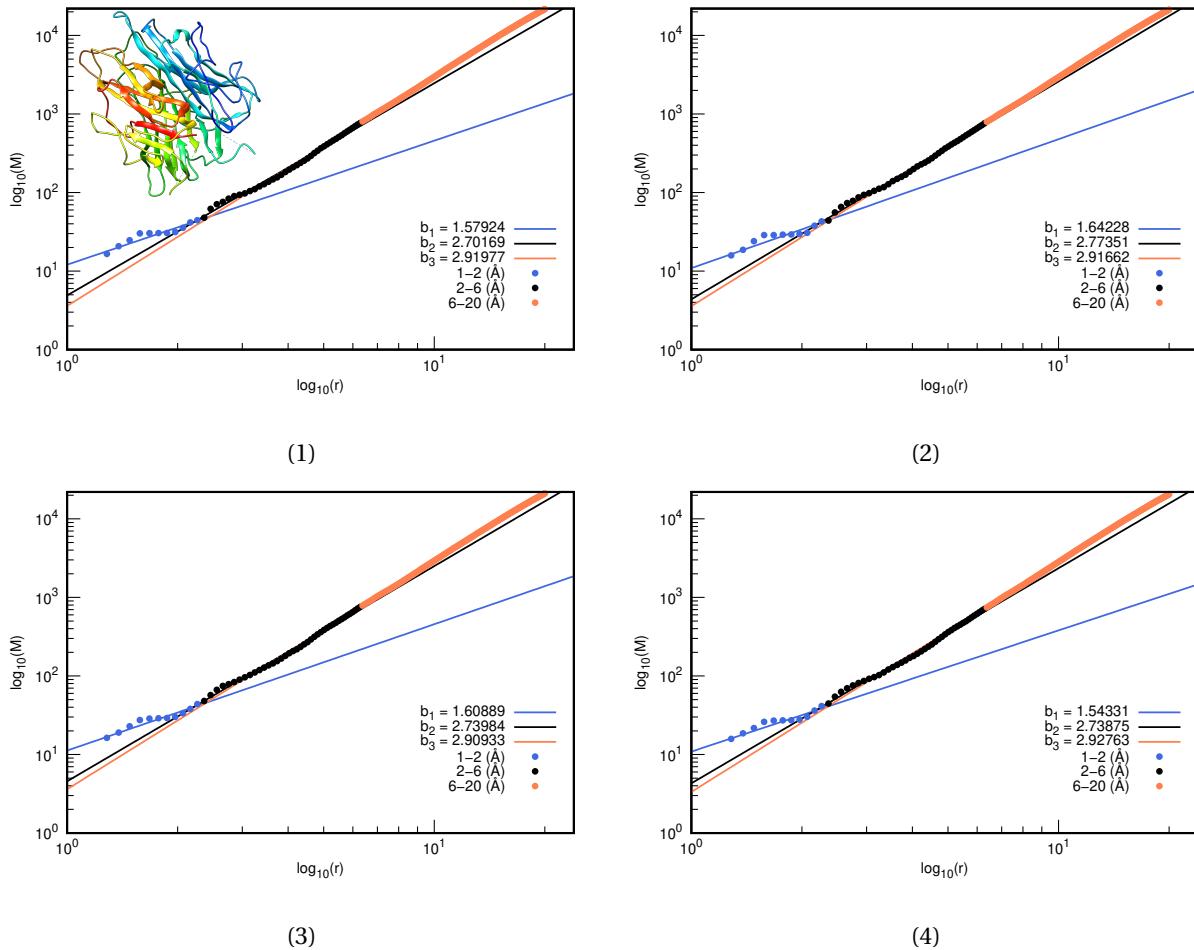
**Figura B.3:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la octava proteína con *IdPDB:11gs* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

## Proteína Arilsulfatasa A (IdPDB:1auk)



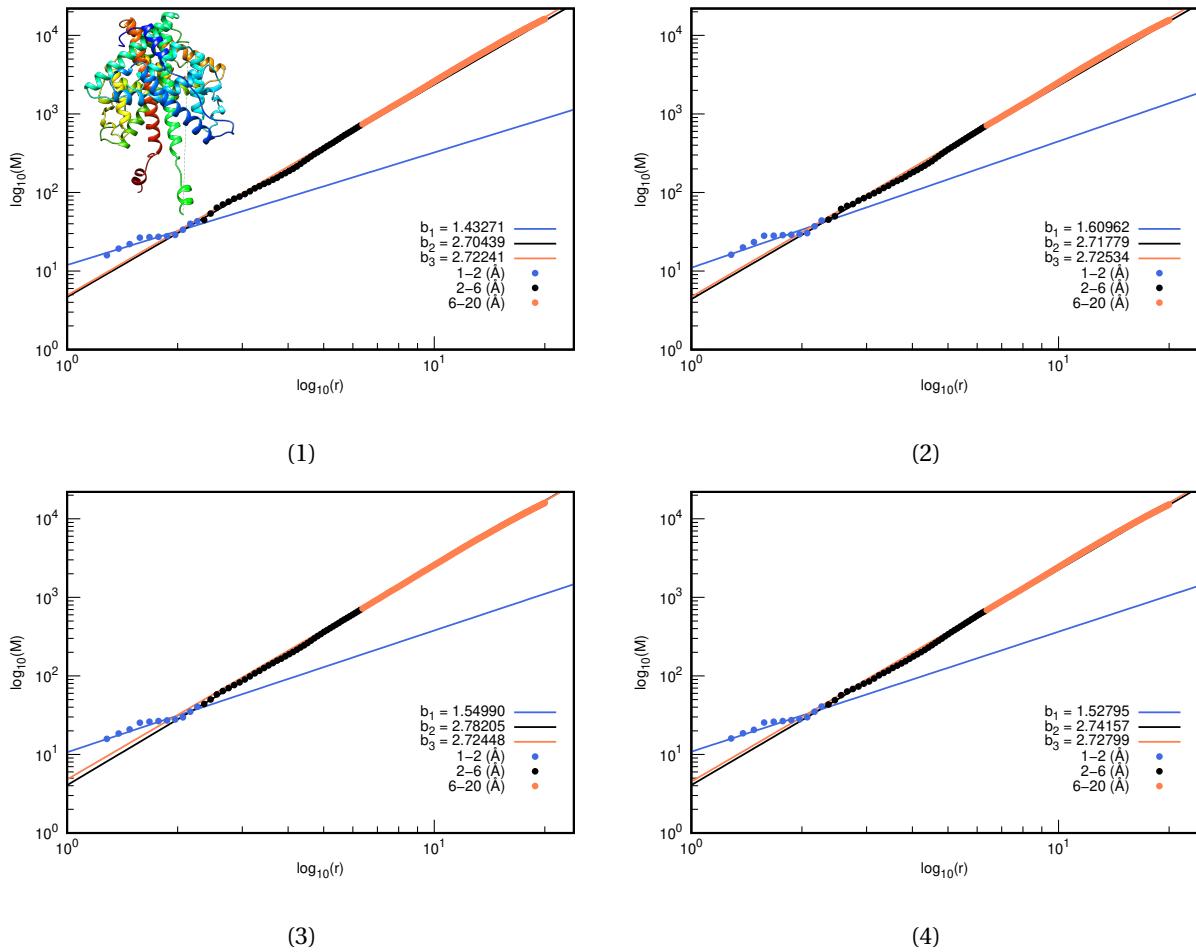
**Figura B.4:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la sexta proteína con *IdPDB:1auk* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

### Proteína Factor de necrosis tumoral $\alpha$ (IdPDB:1a8m)



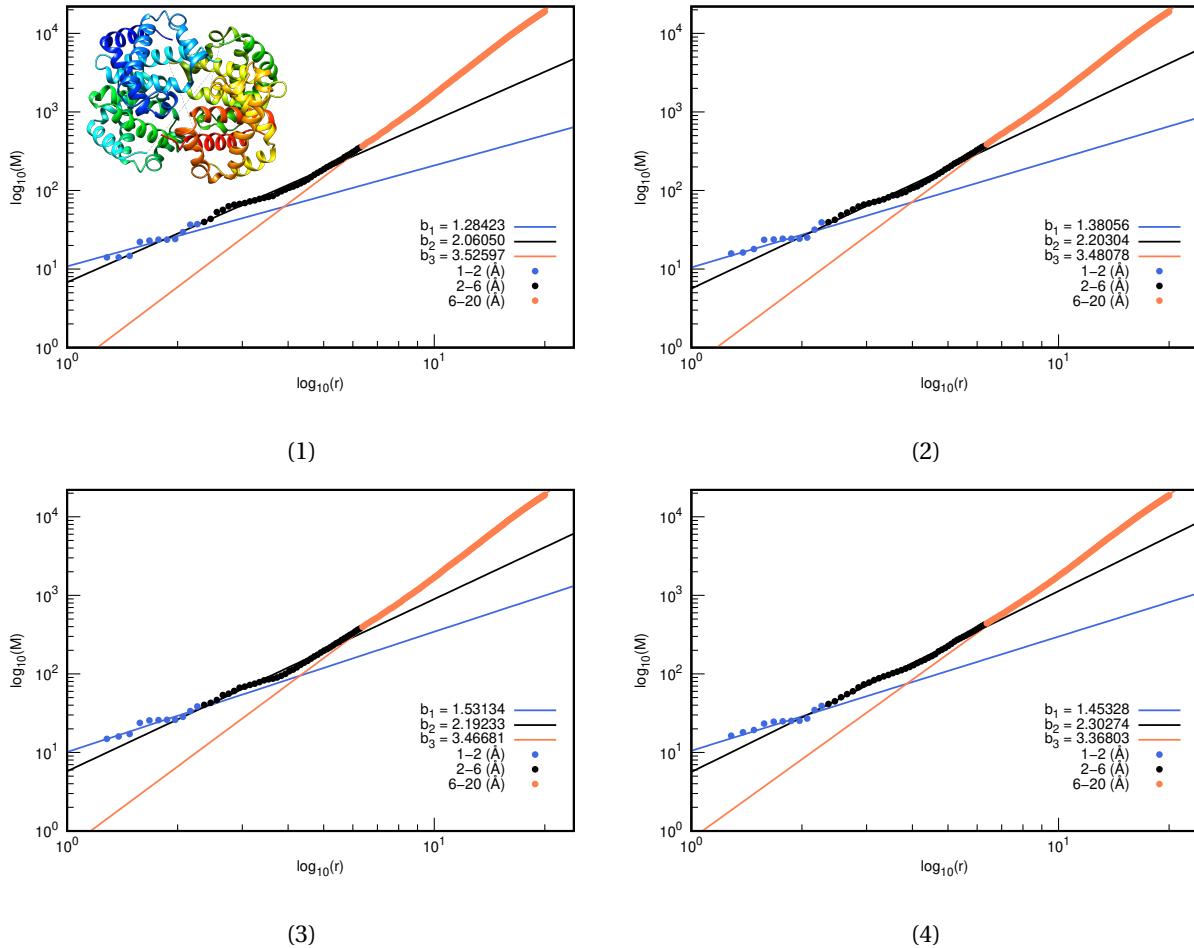
**Figura B.5:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la cuarta proteína con *IdPDB:1a8m* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

### Proteína Receptor de estrógeno (*IdPDB:1a52*)



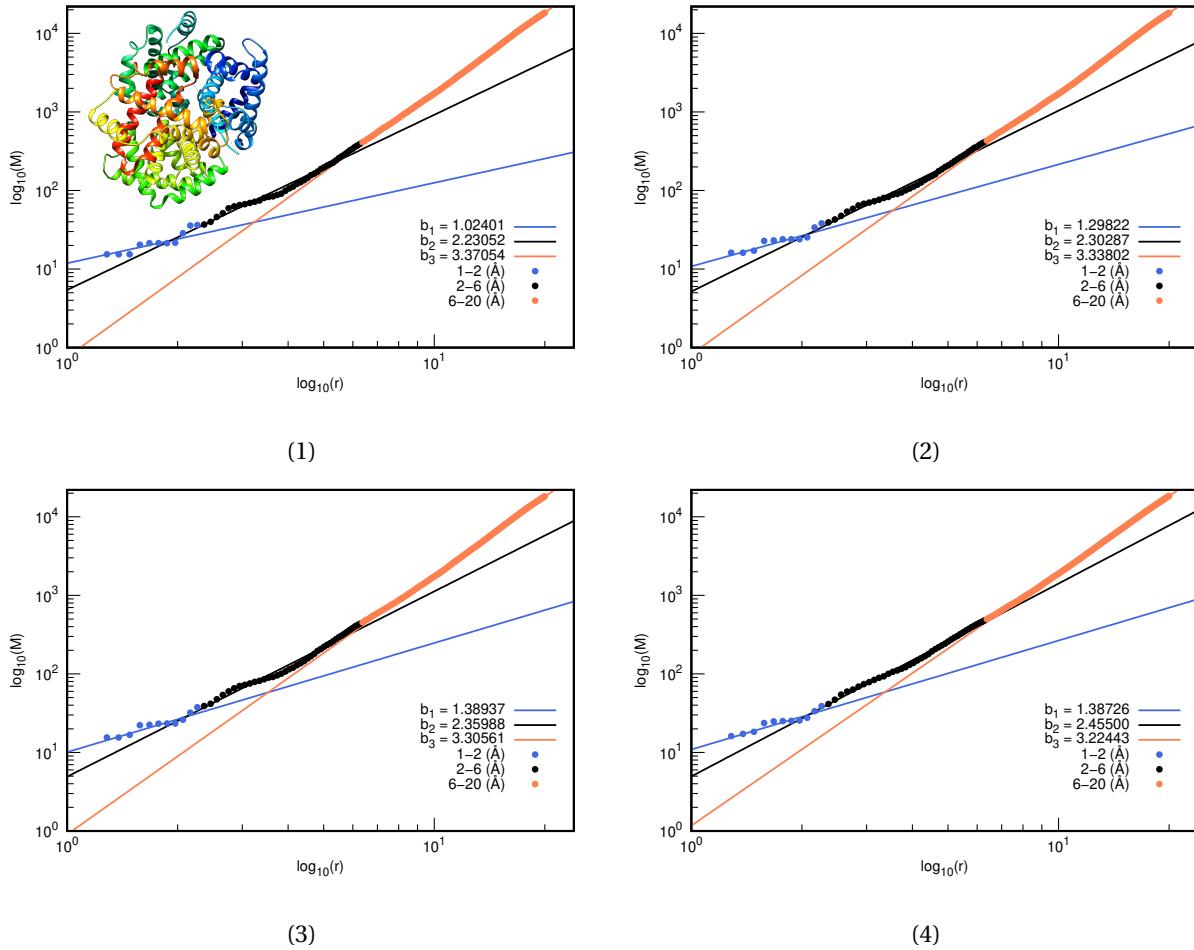
**Figura B.6:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la tercera proteína con *IdPDB:1a52* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

### Proteína Hemoglobina humana desoxidante cadena $\alpha$ (IdPDB:1a3n)



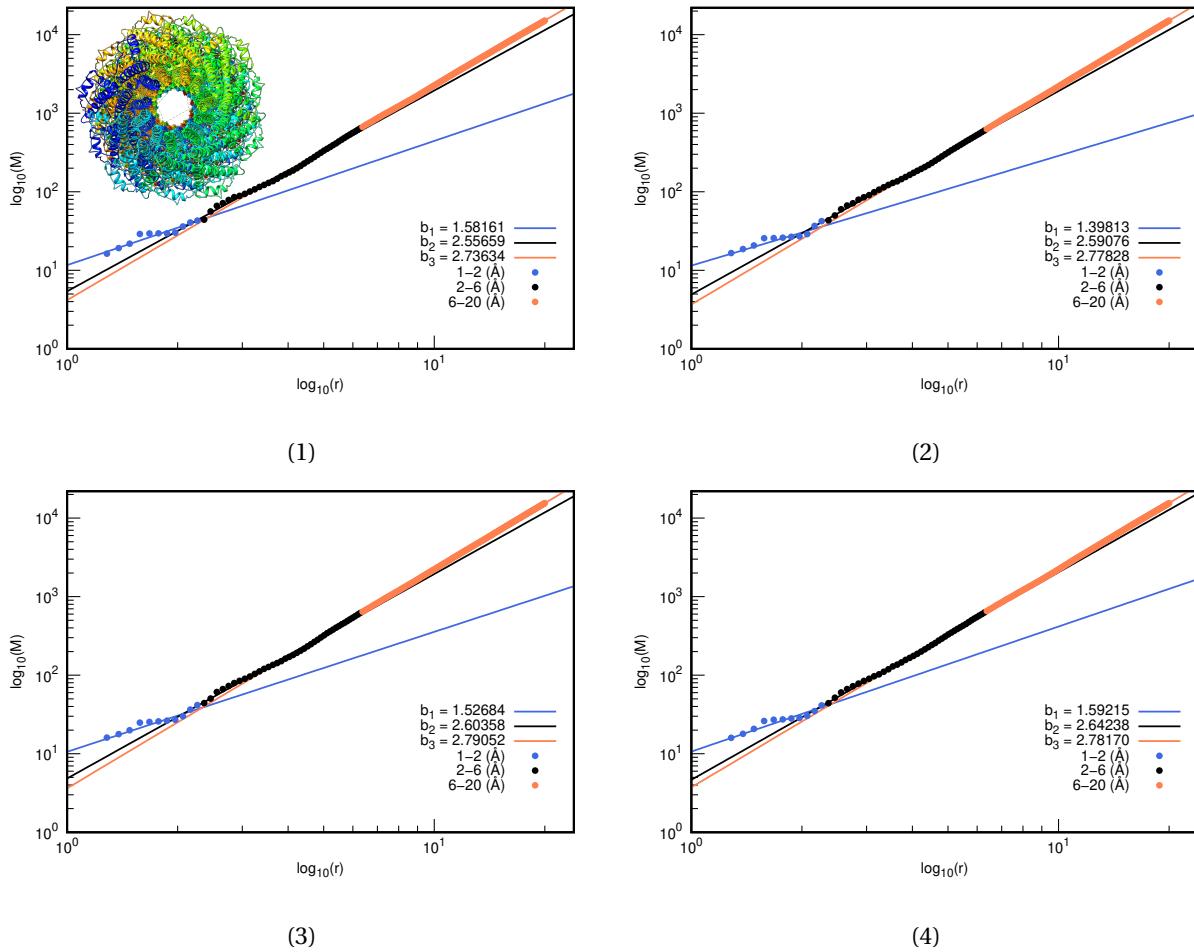
**Figura B.7:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la segunda proteína con *IdPDB:1a3n* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

### Proteína Hemoglobina embrionaria cadena $\alpha$ (IdPDB:1a9w)



**Figura B.8:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la quinta proteína con *IdPDB:1a9w* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

### Proteína Translocon EspA (IdPDB:7khw)



**Figura B.9:** Regresiones lineales de  $\log_{10} r$  vs  $\log_{10} M(r)$  correspondiente a cuatro etapas de procesamiento de la novena proteína con *IdPDB:7khw* de la Tabla 4.3: (1) Adición de átomos de hidrógeno al sistema proteico; (2) al minimizar la energía de la estructura molecular; (3) equilibrando el sistema bajo condiciones termodinámicas controladas; y (4) después de una dinámica molecular de 1 ns.

# Referencias

- [1] Enright M. B. y Leitner D. M., *Phys. Rev. E* **71**, 011912 (2005).
- [2] Carrillo J. L., Donado F, y Mendoza M. E., *Phys. Rev. E* **68**, 061509 (2003).
- [3] Hardin J. y Lodolce J. P., *Becker's World of the Cell*, Pearson Education, 10 edition, 2022.
- [4] Dewey T. G., *Fractals in Molecular Biophysics*, Oxford University Press, 1997.
- [5] Iannaccone P. M. y Khokha M., *Fractal Geometry in Biological Systems: An Analytical Approach*, CRC Press, Inc., 1996.
- [6] Vicsek T., *Fractal Growth Phenomena*, World Scientific Publishing Company, 1992.
- [7] Cserzö M. y Vicsek T., *Chaos, Solitons and Fractals* **1**, 431 (1991).
- [8] Enright M. B., Yu X., y Leitner D. M., *Phys. Rev. E* **73**, 051905 (2006).
- [9] Irina K. y Ruben A., *Homology Modeling: Methods and Protocols*, Humana Press, Totowa, NJ, 2012.
- [10] Meakin P., *Fractals, Scaling and Growth Far from Equilibrium*, Cambridge University Press, 1998.
- [11] RCSB Protein Data Bank (RCSB PDB), <https://www.rcsb.org/>.
- [12] Santamaria R., *Molecular Dynamics*, Springer Cham, 1 edition, 2023.

---

## REFERENCIAS

---

- [13] Saudagar P. y Tripathi T., *Protein Folding Dynamics and Stability*, Springer Singapore, 1 edition, 2023.
- [14] Stapleton H. J., Allen J. P., Flynn C. P., Stinson D. G., y Kurtz S. R., *Phys. Rev. Lett.* **45**, 1456 (1980).
- [15] Helman J. S., Coniglio A., y Tsallis C., *Phys. Rev. Lett.* **53**, 1195 (1984).
- [16] Alexander S. y Orbach R., *Journal de Physique Lettres* **43**, 625 (1982).
- [17] Isogai Y. y Itoh T., *Journal of the Physical Society of Japan* **53**, 2162 (1984).
- [18] Wagner G. C., Colvin J. T., Allen J. P., y Stapleton H. J., *Journal of the American Chemical Society* **107**, 5589 (1985).
- [19] Chen S.-H. y Teixeira J., *Phys. Rev. Lett.* **57**, 2583 (1986).
- [20] Wang C. X., Shi Y. Y., y Huang F. H., *Phys. Rev. A* **41**, 7043 (1990).
- [21] Li H., Chen S., y Zhao H., *International Journal of Biological Macromolecules* **13**, 210 (1991).
- [22] Dewey T. G., *Fractals* **01**, 179 (1993).
- [23] Moret M. A., Pascutti P. G., Mundim K. C., Bisch P. M., y Nogueira E., *Phys. Rev. E* **63**, 020901 (2001).
- [24] Burioni R., Cassi D., Cecconi F., y Vulpiani A., *Proteins* **55**, 529 (2004).
- [25] Moret M. A., Miranda J. G. V., Nogueira E., Santana M. C., y Zebende G. F., *Phys. Rev. E* **71**, 012901 (2005).
- [26] Lee C.-Y., *Phys. Rev. E* **73**, 042901 (2006).
- [27] Banerji A. y Ghosh I., *Cellular and Molecular Life Sciences* **68**, 2711 (2011).
- [28] Peng X., Qi W., Wang M., Su R., y He Z., *Communications in Nonlinear Science and Numerical Simulation* **18**, 3373 (2013).

---

## REFERENCIAS

---

- [29] Petreus D. E. y Isvoran A., *Applied Sciences* **14** (2024).
- [30] Czub M. P., Uliana F., Grubić T., Padeste C., Rosowski K. A., Lorenz C., Dufresne E. R., Menzel A., Vakonakis I., Gasser U., y Steinmetz M. O., *Nature Communications* **16**, 1165 (2025).
- [31] Hughes M. D. G., Cussons S., Borumand A., Tyler A. I. I., Brockwell D. J., y Dougan L., *Soft Matter* **21**, 1748 (2025), Open Access, first published 05 Feb 2025.
- [32] Shen T. Y., Tai K., y McCammon J. A., *Phys. Rev. E* **63**, 041902 (2001).
- [33] Banerji A. y Navare C., *Journal of Molecular Recognition* **26**, 201 (2013).
- [34] Sendker F. L., Lo Y. K., Heimerl T., Bohn S., Persson L. J., Mais C.-N., Sadowska W., Paczia N., Nußbaum E., del Carmen Sánchez Olmos M., Forchhammer K., Schindler D., Erb T. J., Benesch J. L. P., Marklund E. G., Bange G., Schuller J. M., y Hochberg G. K. A., *Nature* **628**, 894 (2024).
- [35] Lemkul J. A., *The Journal of Physical Chemistry B* **128**, 9418 (2024), PMID: 39305267.
- [36] Abraham M. J., Murtola T., Schulz R., Páll S., Smith J. C., Hess B., y Lindahl E., *SoftwareX* **1-2**, 19 (2015).
- [37] Liu S., Xiang X., Gao X., y Liu H., *Scientific Reports* **10**, 4371 (2020).