



UNIVERSIDAD DE MURCIA

FACULTAD DE QUÍMICA

Simulación de Propiedades en Disolución de Polímeros

D. Ricardo Rodríguez Schmidt

2012



UNIVERSIDAD DE
MURCIA

D. José García de la Torre, Catedrático de Universidad del Área de Química Física en el Departamento de Química Física, AUTORIZA:

La presentación de la Tesis Doctoral titulada “*Simulación de Propiedades en Disolución de Polímeros*”, realizada por D. Ricardo Rodríguez Schmidt, bajo mi inmediata dirección y supervisión, y que presenta para la obtención del grado de Doctor por la Universidad de Murcia.

En Murcia, a 25 de Julio de 2012



Facultad de Química
Departamento de Química Física

Campus Universitario de Espinardo. 30100 Murcia
T. 868 887 421 – F. 868 884 148 – www.um.es/dp-quimica-fisica/



UNIVERSIDAD DE
MURCIA

D. José Ginés Hernández Cifre, Profesor Titular de Universidad del Área de Química Física en el Departamento de Química Física, AUTORIZA:

La presentación de la Tesis Doctoral titulada “*Simulación de Propiedades en Disolución de Polímeros*”, realizada por D. Ricardo Rodríguez Schmidt, bajo mi inmediata dirección y supervisión, y que presenta para la obtención del grado de Doctor por la Universidad de Murcia.

En Murcia, a 25 de Julio de 2012

José Ginés Hernández C.



Facultad de Química
Departamento de Química Física

Campus Universitario de Espinardo. 30100 Murcia
T. 868 887 421 – F. 868 884 148 – www.um.es/dp-quimica-fisica/



D. Manuel Hernández de Córdoba, Catedrático de Universidad del Área de Química Analítica y **Presidente Comisión Académica programa doctorado** * en Química, INFORMA:

Que una vez evaluado, de conformidad con el procedimiento establecido en el artículo 21 del Reglamento de doctorado de la Universidad de Murcia, el expediente completo de la tesis doctoral titulada "Simulación de Propiedades en Disolución de Polímeros", realizada por D. Ricardo Rodríguez Schmidt, bajo la inmediata dirección y supervisión de D. José García de la Torre y D. José Ginés Hernández Cifre, esta Comisión Académica, en sesión celebrada en fecha 10 de septiembre de 2012, ha dado su autorización para su presentación ante la Comisión General de Doctorado.

Murcia, a 14 de septiembre de 2012

Doctorando: D. Ricardo Rodríguez Schmidt

**Informe del Departamento para alumnos del RD 778/1998.*

**Informe de La Comisión Académica del Programa para alumnos del RD 56/2005 y RD 1393/2007.*

“To err is human, but to really foul things up you need a computer.”

— Paul R. Ehrlich.

ÍNDICE

ÍNDICE	I
AGRADECIMIENTOS	v
1. INTRODUCCIÓN	1
1.1. Organización de la presente memoria.	2
2. FUNDAMENTOS Y METODOLOGÍA	5
2.1. Propiedades en disolución.	6
2.1.1. Radio de giro.	6
2.1.2. Coeficiente de difusión traslacional.	7
2.1.3. Viscosidad intrínseca.	8
2.1.4. Radios equivalentes y relaciones de radios.	9
2.2. Modelos de bolas y muelles.	12
2.3. Procedimiento de parametrización multiescala.	16
2.3.1. Parametrización del radio de las esferas	18
2.3.2. Parametrización de conectores y ángulos	19
2.4. Técnicas de simulación: dinámica browniana.	24
2.4.1. Ecuaciones para la dinámica browniana.	25
2.4.2. La interacción hidrodinámica.	27
2.5. Técnicas de simulación: Monte Carlo.	32
2.6. Magnitudes reducidas.	34

3. EQUIPO Y SOFTWARE	37
3.1. Equipos informáticos de alto rendimiento.	38
3.1.1. <i>Cluster</i> del grupo.	39
3.1.2. Supercomputador Ben Arabi.	40
3.2. Esquemas de trabajo en entornos de cálculo en paralelo.	42
3.3. Adaptación del software de simulación a entornos de trabajo en paralelo.	44
3.4. El paquete SIMUFLEX.	46
3.5. El programa MONTEHYDRO.	47
3.6. Programas auxiliares: MULTISIMUFLEX.	48
4. METODOLOGÍAS PARA CÁLCULO DE D_T	51
4.1. Introducción.	51
4.2. Base teórica.	52
4.2.1. Coeficiente de difusión de Einstein.	52
4.2.2. Coeficiente de difusión de Kirkwood.	53
4.2.3. Tratamiento del Cuerpo Rígido.	54
4.2.4. Coeficiente de difusión a tiempo largo.	55
4.3. Modelo y metodología.	56
4.3.1. Modelo de bolas y muelles.	56
4.3.2. Detalles de la simulación y el análisis.	58
4.4. Resultados.	61
4.5. Conclusiones del Capítulo.	66
5. ALGORITMOS PARA TRATAMIENTO DE DB-IH	69
5.1. Introducción.	69
5.2. Base Teórica.	72
5.2.1. Generalidades.	72
5.2.2. Algoritmo de Ermak y McCammon. La descomposición de Cholesky.	74
5.2.3. Desarrollo en serie de Chebyshev y algoritmos derivados.	75

5.2.4. Desarrollo truncado de Geyer y Winter.	83
5.3. Detalles sobre las simulaciones.	88
5.3.1. Modelos empleados en la simulación.	88
5.3.2. Simulaciones y cálculo de propiedades.	91
5.4. Resultados.	93
5.4.1. Comparación de propiedades.	94
5.4.2. Comparación de velocidad.	105
5.5. Conclusiones del Capítulo.	109
6. POLÍMEROS HIPERRAMIFICADOS	111
6.1. Introducción.	111
6.2. Base Teórica.	113
6.2.1. Índices topológicos.	113
6.3. Correspondencia entre estructura y topología.	120
6.4. Algoritmo generador de topología.	121
6.5. Índices topológicos de las estructuras generadas.	123
6.5.1. Influencia de N en DB y W	126
6.6. Modelado, parametrización y simulación.	127
6.6.1. Generación de topologías.	127
6.6.2. Modelado y parametrización.	128
6.7. Resultados.	130
6.7.1. Polibenziléter (PBzE).	130
6.7.2. Poliaminoamida (PAMAM).	132
6.7.3. Polimetilalilcarbósilano (PCS3).	135
6.7.4. Polimetilundecilcarbósilano (PCS11).	145
6.7.5. Polidimetil 5-(4-hidroxibutoxi)isofталato (PDHBI).	149
6.8. Discusión de resultados.	152
6.8.1. Variación de las propiedades con el peso molecular.	152

6.8.2. Función de distribución de radios de giro.	155
6.8.3. Comparación con los datos experimentales.	156
6.9. Conclusiones del capítulo.	162
7. COPOLÍMEROS ANFIFÍLICOS TERMOSENSIBLES	165
7.1. Introducción.	165
7.2. Modelo y parametrización.	168
7.2.1. Parametrización de las bolas.	169
7.2.2. Parametrización de los muelles.	170
7.2.3. Parametrización del volumen excluido.	178
7.3. Resultados.	182
7.3.1. Copolímeros dibloque.	182
7.3.2. Copolímeros tribloque.	197
7.4. Conclusiones del Capítulo	206
8. CONCLUSIONES / CONCLUSIONS	209
GLOSARIO	213
BIBLIOGRAFÍA	219
SUMMARY	S-1
1. Comparison of approaches to D_t	S-3
2. Algorithms for BD-IH	S-7
3. Multiscale simulation of hyperbranched polymers.	S-13
4. Simulation of thermoresponsive amphiphilic copolymers.	S-21

AGRADECIMIENTOS

Desde que me las encontré en un libro de literatura castellana, hace ya muchos años, siempre he creído en las palabras de Ortega y Gasset: “yo soy yo y mis circunstancias”, y también en las que Machado dedicaba a su caminante, y que ya conoce todo el mundo: que no hay camino, sino que se hace el camino al andar. Uniendo ambas reflexiones, me quedo con que una parte muy importante de cada uno nosotros responde a las circunstancias y los momentos que hemos pasado a lo largo de nuestro camino.

Durante los últimos años, mis circunstancias han estado irremediabilmente marcadas por el desarrollo los trabajos cuyos resultados y conclusiones se recogen en las muchas páginas y varios Capítulos que componen esta Tesis. Sin embargo, esta etapa de mi vida ha dado mucho más de sí que las meras cuestiones científicas que aquí se presentan, pues he estado rodeado de mucha gente estupenda que me ha acompañado, guiado, enseñado, y por qué no decirlo, también a veces aguantado, a lo largo de todo este tiempo. Y es por toda esta gente por lo que ahora, al echar la vista atrás, sólo puedo recordar momentos buenos y mejores, pues incluso los que *a priori* parecían malos, han terminado por ser buenos recuerdos, y no los cambiaría por otros.

Así, creo firmemente que toda esta gente no sólo ha formado parte de mis circunstancias, sino que también lo son de mí mismo, de lo que soy a día de hoy. Y como tal, no sería justo por mi parte olvidarme de ellos ahora que ha llegado el momento de recopilar en este volumen el fruto de estos cuatro años de trabajo. Por este motivo, quiero dejar constancia de mi más profundo agradecimiento a toda esta gente, aunque sea tan sólo con unas breves palabras (creo que nunca sería capaz de encontrar las suficientes como para agradecerles todo lo que me han aportado), ni mucho menos tener espacio suficiente para plasmarlas.

En primer lugar, les debo el mayor de los agradecimientos a Carmina y a Pepe, que, en primer lugar me abrieron las puertas del Departamento de Química Física, tu-

telándome primero en la obtención del Máster y luego durante la elaboración de esta Tesis, brindándome su apoyo y sus consejos.

En segundo lugar, y no menos importante, a Ginés, que desde el primer día estuvo ahí, no sólo encauzando mi trabajo en esta Tesis en la dirección adecuada, sino que también me acompañó en tantas ocasiones fuera de lo académico: en el deporte, las celebraciones y alguna que otra pequeña conspiración.

Por supuesto, también a mis padres, que, cada uno a su manera, me han apoyado y animado de forma irremplazable para seguir adelante, no sólo durante estos cuatro años, sino durante toda mi vida.

A Ramón, por mostrarme que había vida más allá de la “mazmorra”, y estar siempre ahí para darme un empujón, un consejo, echar una mano o cualquier otra cosa que hiciera falta.

A toda la “panda” de becarios y compañeros de la facultad con los que he coincidido, tanto a los del departamento de Química Física: Gustavo, Diego, Ana, Fran Menchón, Álvaro, Encarni, Edu, Carmen, Antonio Vázquez, Miguel, Cerón, Javi Cerezo, Marwa y el Bola; como a los de los demás departamentos: Luisa, Marina, M^a José y Virginia (Química Inorgánica); Jaime, Magalí, Jesús, Magda y Alejandro (Bioquímica A); Jose María (Bioquímica B); Loli, Kike, Jose Antonio y Rosalía (Química Orgánica) y Yésica (Química Analítica). Gracias a todos por el día a día, por liarme y dejaros liar en tantas y tantas ocasiones. También a Horacio y a Javi Pina, con quienes no he coincidido en la Universidad, pero que se han convertido en grandes amigos, compañeros de juergas, carreras a pie, rutas en bici y muchas risas y buenos momentos.

A todos mis compañeros de los cursos de japonés del Servicio de Idiomas, en especial Alfredo, Jose Padilla, Cristina Serrano, Cristina Sánchez, M^a José Bernal, Jose Mora, Juanfran, Miriam, Alan, Eduardo, y muy especialmente, a Pablo, o “Maez”, como se le conoce ahora que se está haciendo conocido en el mundo de la música; a nuestra excelente profesora, ^{Akaboshi Michiko}赤星みちこ, y a Cristina Ortiz, nuestra becaria favorita del Servicio de Idiomas. A todos vosotros, どうも ありがとう ございます!

A los integrantes del grupo del profesor Bo Nyström, del Departamento de Química de la Universidad de Oslo, allá en Noruega: Anna-Lena Kjøniksen, Kaizheng Zhu, Atoosa Maleki, Loan Thi Trinh, Hang Thuy Thi Phan, Solmaz Bayati, Shirin Fanaian y muy especialmente a Nodar Al-Manasir. También a los miembros del Departamento de Farmacia, Wei Wang, Therese Klemetsrud, Helene Jonassen, Sara Bekhradnia, Teresa Calejo y

Lilia Ulanova. A todos, gracias, por cantarme el “Happy Birthday” al terminar mi charla durante nuestro primer *meeting* en el 2010, por volver a visitarnos el año siguiente, y por lo bien que nos recibieron y nos trataron a Ginés y a mí cuando fuimos a visitarles para el *meeting* del 2011. Thank you very much, guys, you are awesome! Tusen takk!

Al profesor Helmut Cölfen de la Universidad de Constanza, y a todo su grupo, por acogerme durante mi periodo de estancia en el extranjero; muy especialmente a Rose Rosenberg y Dirk Haffke, mis “maestros” a lo largo de la estancia. También a Johannes Dingler y Daniela Hrzán, del Welcome Center de la Universidad de Konstanza, por la bienvenida, los consejos, y los eventos que organizan para que los que venimos de fuera no nos sintamos solos. También al grupo “DokNet” por permitirme participar en los estupendos eventos que organizan, ponerme en contacto con otros doctorandos en Konstanz, y en general, por todos los buenos ratos. A Enara Pérez de Heredia, María López Quijorna, Izaro Arroita, Esteban Freidin, Kaia Magnusen, John Berg e Iva Radisavljević, a los que considero grandes amigos, y gracias a quienes me sentí como en casa. Y sobre todo a Nuria Mancebo, que es casi tan cabezota como yo, y que a fuerza de todo el tiempo que pasamos juntos por allí me hizo pasar tanto los mejores ratos como los mayores “mosqueos” que he tenido en Constanza. Vielen Dank für alles, Leute!

A todo el profesorado del departamento de Química Física, por el estupendo ambiente (y no sólo de trabajo) que mantienen, en especial a Guillermo, que, al día de escribir estas líneas, todavía nos debe el cumplimiento de su palabra de participar en la carrera popular de la universidad que él mismo instauró durante su etapa Vicerrector de Extensión. También a Pedro y Úrsula, sin cuyo trabajo el departamento se vendría abajo.

A Asun Hidalgo, que a pesar de estar siempre hiper-ocupada siempre ha tenido un momento para darme su apoyo.

A Jordi S. Vicente Alonso, Mónica Gil, Sergio Fernández y Marichu Gómez, peregrinos, jubilares, y sobre todo, grandes amigos a los que me “encontré” por el Camino, después volví a ver en el Moncayo y espero volver a ver muchas veces.

A todos aquellos con los que nos encontramos y / o enfrentamos en una pista de fútbol sala durante todo este tiempo.

A Eduardo Zaragoza y José Muñoz, enormes amigos y compañeros de madrugones los domingos para salir a la montaña y ganar cumbres.

A los que acudieron a las sesiones diurnas y también a las nocturnas del JIP 2010, y que comparten la “culpa” de que a aquel congreso lo recordemos como el “Tuneti 2010” (y por supuesto, a la dueña de aquel bar, que aún debe estar esperando a que volvamos).

A todos aquellos que, en los momentos de desesperación, me prestaron un oído al que contar mis penas, un hombro sobre el que apoyarme, y me dieron una palmada en el hombro o una patada en el culo para que siguiera adelante.

Y para todos aquellos que en este momento se me escapan, además de mi agradecimiento por haber estado ahí, una enorme disculpa por no haber caído a tiempo en mentarlos.

CAPÍTULO 1

INTRODUCCIÓN

Hoy en día la importancia de las macromoléculas, tanto naturales como artificiales, está fuera de toda duda. Desde hace bastantes años, elementos poliméricos como fibras y plásticos forman parte de nuestra vida cotidiana, mientras que otras macromoléculas como proteínas y ácidos nucleicos, que siempre nos han acompañado, pues sin ellas no sería posible la vida, son cada vez más utilizadas en procesos artificiales.

Conociendo esta presencia casi universal de las macromoléculas en la vida moderna es fácil comprender la gran importancia de la investigación de las mismas que se lleva a cabo desde ámbitos y puntos de vista muy variados.

En particular, el estudio de disoluciones diluidas de macromoléculas desde un punto de vista físicoquímico es interesante puesto que permite estudiar cambios conformacionales en la estructura macromolecular [1, 2]. A partir de informaciones como la variedad conformacional y el tamaño de la macromolécula en disolución puede obtenerse información acerca de las propiedades y el comportamiento de moléculas aisladas (es decir, en ausencia de interacción con otras macromoléculas). De esta manera, se obtiene información directamente relacionada con la estructura de la macromolécula: viscosidad intrínseca, tamaño promedio, flexibilidad de la cadena, etc.

Así, el estudio de las disoluciones diluidas de macromoléculas ha recibido bastante atención durante décadas tanto a nivel experimental como teórico [3]. Si bien el trabajo experimental no está exento de dificultades, el tratamiento teórico de esta cuestión es muy complicado, pues han de tenerse en cuenta los aspectos estadísticos debidos a la variabilidad conformacional de las macromoléculas en disolución, así como la interacción hidrodinámica que se establece por la fricción del disolvente sobre la macromolécula, y la

transmisión de interacciones de unos segmentos de la macromolécula a otros a través de las moléculas disolvente que los rodean.

Ambos aspectos contribuyen notablemente al aumento de la complejidad del tratamiento matemático necesario para realizar estudios teóricos de este tipo, por lo que la simulación de la dinámica de macromoléculas en disolución debe evaluarse necesariamente mediante métodos informatizados. Esta Tesis Doctoral se centra en el estudio computacional de las propiedades de tales disoluciones, poniendo especial énfasis en los aspectos metodológicos de dicha simulación computacional. Para ello es necesaria la creación de modelos que representen de forma adecuada las moléculas que sean objeto del estudio, así como el desarrollo de aplicaciones informáticas adecuadas a la situación que se desee estudiar [4, 5].

1.1. Organización de la presente memoria.

Tanto las técnicas de simulación como los modelos empleados en esta Tesis se describen en el Capítulo 2. Allí se desarrollan los fundamentos teóricos y metodológicos de las simulaciones de tipo Monte Carlo y de tipo dinámica browniana, así como también se describen las características de los modelos de cadenas flexibles que se emplearán en la realización de las simulaciones realizadas en los Capítulos siguientes. También se incluyen aquí las definiciones de las propiedades en disolución que se emplean en el resto de la Tesis, así como el procedimiento de parametrización multiescala que se emplea en los Capítulos en los que se estudian sistemas que representan a macromoléculas reales.

En el siguiente Capítulo, el 3, se realiza una breve presentación de los equipos informáticos de alto rendimiento y el software, tanto a nivel de sistema operativo como de gestión de colas, que han sido utilizados para llevar a cabo las simulaciones correspondientes a los trabajos realizados durante la elaboración de esta Tesis. En la segunda parte de este Capítulo se presentan también los paquetes de software de simulación de macromoléculas en disolución que han sido realizados por nuestro grupo y que han sido empleados para realizar algunas de las simulaciones de esta Tesis.

En los dos Capítulos que se han mencionado se introducen los dos elementos necesarios para la realización de las simulaciones: los algoritmos que se implementarán en los

programas de simulación y los equipos informáticos que se encargarán de la realización de los cálculos necesarios. Parece evidente que una mejora en la potencia de los equipos informáticos empleados redundará en una mejora de la velocidad con la que se obtienen los resultados de las simulaciones, al mismo tiempo que se podrán abordar estudios de sistemas que por su complejidad serían inviables con equipos de menor potencia, pues como se describe en el Capítulo 2, la potencia de cálculo es un factor limitante en cuanto al tamaño de las moléculas que pueden emplearse en una simulación.

Dentro de las simulaciones convencionales, un caso muy crítico en cuanto a necesidad de potencia de cálculo corresponde a las simulaciones de dinámica browniana con interacción hidrodinámica. Por este motivo, desde que Ermak y McCammon [6] presentaran su algoritmo de dinámica browniana con inclusión de la interacción hidrodinámica se han desarrollado nuevos algoritmos, que, merced a ciertas aproximaciones y limitaciones, aspiran a conseguir una mayor eficiencia en los cálculos. Sin embargo, hasta este momento no se ha realizado ninguna comparación exhaustiva entre ellos, por lo que en esta Tesis se ha realizado una comparación de los mismos en cuanto a la simulación de cadenas flexibles de diferentes tamaños. Esta comparativa, además de los detalles de cada uno de los algoritmos evaluados, se presenta en el Capítulo 5.

De la misma manera que existen diferentes opciones en cuanto a algoritmos a la hora de realizar una simulación, para posteriormente calcular determinadas propiedades, como por ejemplo, el coeficiente de difusión traslacional, existen diferentes métodos para su cálculo teórico, los cuales no siempre producen resultados exactamente coincidentes entre ellos. Por este motivo, en el Capítulo 4 se han comparado diferentes procedimientos para obtener el coeficiente de difusión traslacional, tanto a partir de simulaciones de tipo Monte Carlo como de dinámica browniana.

Tras estos dos Capítulos, en los que nos planteamos objetivos a un nivel teórico, a saber, la comparación de metodologías de simulación y el estudio de la relación entre los valores de una propiedad obtenida mediante diferentes métodos, se presentan otros dos Capítulos cuyo objetivo es la aplicación de las técnicas de simulación a estudios de sistemas reales de macromoléculas.

En el Capítulo 6 se busca estudiar la aplicación de una metodología desarrollada en un trabajo anterior de este grupo [7] a una nueva familia de macromoléculas. En dicho trabajo se describió una metodología para la parametrización *a priori* (es decir, sin emplear parámetros ajustables) y simulación de macromoléculas, y se aplicó con éxito a la simulación de Monte Carlo de los polímeros conocidos como “dendrímeros”, moléculas con

un alto grado de ramificación y regularidad. En la presente Tesis, empleando los nuevos métodos y equipos de computación distribuida descritos en el Capítulo 3, se ha intentado extender esta metodología de trabajo a la simulación de polímeros hiperramificados, moléculas que, al igual que los dendrímeros, poseen un alto grado de ramificación, pero que carecen de la regularidad de los anteriores.

El último Capítulo, dedicado a la simulación de moléculas reales, el Capítulo 7, versa sobre un trabajo realizado en colaboración con el grupo del profesor Nyström, del Departamento de Química de la Universidad de Oslo. En el laboratorio del profesor Nyström se sintetizaron y caracterizaron ciertos copolímeros termosensibles, en concreto, dos familias de compuestos dibloque y tribloque de poli(N-isopropilacrilamida) (PNIPAAM) y metoxipolietilenglicol (MPEG), cuya hidrofobicidad depende de la temperatura en un rango limitado de ésta.

El trabajo presentado en el Capítulo 7, complementario al realizado por el grupo del profesor Nystrom, contiene los resultados obtenidos mediante el estudio de simulación computacional realizado para esclarecer el comportamiento de moléculas aisladas de los mencionados copolímeros en el rango de temperaturas en el cual estas familias de compuestos muestran su respuesta anfifílica.

Finalmente, en el Capítulo 8, redactado en inglés con el objeto de conseguir la “Mención Internacional” en el título de doctor, se recopilan las conclusiones que se han obtenido en cada uno de los Capítulos anteriores.

CAPÍTULO 2

FUNDAMENTOS Y METODOLOGÍA

El presente Capítulo tiene por objetivo realizar una ligera introducción a los fundamentos teóricos en los que se basan las técnicas de simulación que se han empleado en el desarrollo de esta Tesis, así como presentar otros elementos y cuestiones relacionados con los temas que se abordan en el resto de Capítulos de esta Tesis.

En la primera de las secciones que comprende este Capítulo se hace una breve descripción de las propiedades que se obtienen en los trabajos de simulación presentados más adelante. También se presenta el concepto de “radio equivalente” y las relaciones entre estos radios, conceptos ambos útiles a la hora de construir modelos moleculares y comparar las propiedades de distintas moléculas.

A continuación se presentan los modelos de “bolas y muelles” que se emplean en las simulaciones que se han llevado a cabo en esta Tesis, así como el procedimiento de parametrización multiescala que se desarrolló en [7], y que en esta Tesis se empleará también en la parametrización de los modelos empleados en los Capítulos 6 y 7.

También se desarrollan aquí los fundamentos de dos de las técnicas más ampliamente empleadas en simulación computacional, la dinámica browniana y el método de Monte Carlo. Ambos métodos no son totalmente equivalentes. Así, el método de Monte Carlo se emplea en estudios de sistemas en equilibrio mientras que la dinámica browniana puede utilizarse tanto para realizar simulaciones de tipo dinámico, en las que el estado del sistema en estudio variará con el tiempo, como simulaciones en equilibrio, en cuyo caso debe producir resultados equivalentes a los obtenidos mediante la técnica de Monte Carlo.

2.1. Propiedades en disolución.

2.1.1. Radio de giro.

Un indicador adecuado del tamaño promedio de una cadena macromolecular en disolución, cualquiera que sea su topología, es la propiedad conformacional conocida como radio de giro cuadrático medio, $\langle s^2 \rangle$, definida mediante la expresión [1, 8]:

$$\langle s^2 \rangle = \frac{1}{N} \sum_{i=1}^N \langle s_i^2 \rangle, \quad (2.1)$$

siendo s_i^2 el módulo cuadrático del vector definido entre el centro de masas de la cadena y la posición ocupada por el eslabón i en una conformación instantánea. Los corchetes angulares $\langle \dots \rangle$ indican el promedio sobre todas las conformaciones posibles de la macromolécula, aunque en la práctica, ante la imposibilidad de tener en cuenta toda la variabilidad conformacional de una molécula, suele emplearse un espacio conformacional lo suficientemente amplio. Dependiendo del modelo con el que se trabaje, la definición de un “eslabón” puede comprender desde un átomo en un modelo atómico, hasta un grupo de miles de átomos si se trata una bola en un modelo de bolas y muelles (que se describirá en la siguiente sección).

Existe una forma alternativa, pero equivalente, de expresar el radio de giro cuadrático medio en términos de vectores r_{ij} definidos entre los elementos i, j de la cadena [1, 8]:

$$\langle s^2 \rangle = \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N \langle r_{ij}^2 \rangle. \quad (2.2)$$

De ambas expresiones, la 2.2 es probablemente la más conocida, pero computacionalmente es la más ventajosa 2.1 por el menor número de operaciones matemáticas que implica.

Por simplicidad, es común emplear el término “radio de giro”, definido como $R_g = \sqrt{\langle s^2 \rangle}$ en el caso de una molécula flexible, o simplemente $R_g = s$ para una macromolécula rígida.

2.1.2. Coeficiente de difusión traslacional.

Es posible observar la difusión de partículas de soluto en una disolución mediante diversos montajes de laboratorio, basados, en general, en la observación de los cambios de concentración de soluto en una porción concreta de la disolución a lo largo del tiempo. La variación espacio-temporal de la concentración es la solución al sistema de ecuaciones diferenciales constituido por la primera y la segunda ley de Fick para unas condiciones iniciales determinadas. De los perfiles de concentración observados se puede obtener el coeficiente de difusión traslacional D_t [2, 9, 10].

La difusión macroscópica es una manifestación del movimiento microscópico – el movimiento browniano, del que se hablará más adelante – de las partículas. La conexión entre ambos aspectos se establece mediante la teoría de Einstein [11], y se traduce en la expresión formulada por el propio Einstein del recorrido cuadrático medio $\langle d^2 \rangle$ de una partícula browniana en un tiempo t [2, 12]:

$$\langle d^2 \rangle = 6D_t t. \quad (2.3)$$

Además de esta expresión, existen otras técnicas para determinar el coeficiente de difusión traslacional que también se basan en la teoría del movimiento browniano, como por ejemplo, la dispersión dinámica de luz [2, 13],

Otro de los aspectos de la teoría de Einstein relaciona el coeficiente de difusión con el coeficiente de fricción de una partícula que se encuentre inmersa en un disolvente. Debido a la viscosidad del disolvente, al moverse la partícula en el seno del mismo, estará sujeta a una fuerza de fricción $\mathbf{F}^{(\text{fri})}$, que, en el caso de una partícula esférica que se mueva con una velocidad \mathbf{u} , puede expresarse como

$$\mathbf{F}^{(\text{fri})} = f_t \mathbf{u}, \quad (2.4)$$

donde f_t es el coeficiente de fricción traslacional. Este coeficiente f_t es un escalar, y para el caso concreto de una partícula esférica, su valor viene dado por la ley de Stokes:

$$f_t = 6\pi\eta_0\sigma, \quad (2.5)$$

donde η_0 es la viscosidad del disolvente puro, y σ es el radio de la esfera.

La relación entre los coeficientes de fricción y difusión traslacional fue demostrada por Einstein utilizando una versión macroscópica (termodinámica) de la teoría del movimiento browniano, y puede expresarse como

$$D_t = \frac{k_B T}{f_t}, \quad (2.6)$$

donde $k_B T$ es el factor de Boltzmann. Cuando f_t es sustituida por su valor para la esfera según la ley de Stokes (ecuación 2.5), la ecuación 2.6 se conoce como la ecuación de Stokes-Einstein.

Si en vez de partículas con simetría esférica se habla de partículas de forma arbitraria, f_t y D_t deben sustituirse por los correspondientes tensores de fricción y difusión. En este caso, el tratamiento matemático se hace bastante más complicado, pero no presenta novedades conceptuales [14–16].

Además de la ecuación fundamental 2.3, existen otras aproximaciones al cálculo del coeficiente de difusión que se usan con frecuencia en el ámbito de la simulación. En el Capítulo 4 se realizará una comparación entre algunas de estas expresiones alternativas y los valores del coeficiente de difusión que se obtiene para modelos sencillos de diferentes características.

2.1.3. Viscosidad intrínseca.

En general, al añadir cualquier soluto a una disolución se observa un aumento, aunque sea pequeño, de su viscosidad. Sin embargo, en el caso concreto de las disoluciones de macromoléculas, éstas se caracterizan por poseer una viscosidad, η , considerablemente superior a la del disolvente puro, η_0 . Este aumento de viscosidad es apreciable incluso a concentraciones notablemente pequeñas. Esto es debido a que la presencia de macro-

moléculas hace que aumenten de manera importante las fuerzas de rozamiento internas, y por tanto, la viscosidad.

Existen diversas formas de expresar el aumento de viscosidad de una disolución de concentración c :

a) Viscosidad relativa:

$$\eta_{rel} = \frac{\eta}{\eta_0}. \quad (2.7)$$

b) Viscosidad específica:

$$\eta_{sp} = \frac{\eta - \eta_0}{\eta_0} = \eta_{rel} - 1. \quad (2.8)$$

c) Viscosidad reducida:

$$\eta_{red} = \frac{\eta_{sp}}{c}. \quad (2.9)$$

Además de las definiciones anteriores se suele trabajar con la propiedad conocida como “viscosidad intrínseca” o “viscosidad inherente”, $[\eta]$, que es una forma de expresar la contribución infinitesimal del soluto al aumento de la viscosidad de la disolución. Experimentalmente, se suele determinar mediante la medición de la viscosidad de disoluciones de un mismo soluto a diferente concentración, y la extrapolación a concentración cero; de acuerdo a este procedimiento, la viscosidad intrínseca puede definirse como

$$[\eta] = \lim_{c \rightarrow 0} \frac{\eta - \eta_0}{\eta_0 c} = \lim_{c \rightarrow 0} \frac{\eta_{sp}}{c}. \quad (2.10)$$

Para el caso de una partícula esférica de radio σ , Einstein [17, 18] demostró que la expresión de la viscosidad intrínseca resulta en la expresión:

$$[\eta] = \frac{10\pi\sigma^3}{3m} = \frac{5}{2} \frac{V}{m}, \quad (2.11)$$

siendo m la masa de la esfera y $V = (4/3)\pi\sigma^3$ su volumen.

2.1.4. Radios equivalentes y relaciones de radios.

Un concepto que nos será de utilidad tanto en la elaboración de los modelos como en la comparación de los resultados obtenidos en las simulaciones es el de “radio de la esfera

equivalente”, propuesto y desarrollado por nuestro grupo [19, 20]. El valor de un radio de esfera equivalente puede calcularse a partir de múltiples propiedades de una partícula, y representa el tamaño de una hipotética partícula esférica que tuviera el mismo valor de la propiedad que se ha empleado para su cálculo que la partícula a partir de la cual se ha obtenido tal valor.

La determinación del valor de dicho radio se realiza mediante el empleo de las ecuaciones para el cálculo de las distintas propiedades en disolución para una geometría esférica, aunque la partícula en cuestión no sea rígida ni esférica, y determinando el radio de dicha esfera ficticia a partir del valor ya conocido de la propiedad elegida.

Dada la variedad de propiedades a las que puede aplicarse el cálculo de radios equivalentes, para facilitar el empleo de este tipo de magnitudes, se emplea una notación unificada: los radios equivalentes se denotan como a_X , donde X representa, mediante una letra mayúscula, a la propiedad respecto a la cual se ha calculado el radio equivalente. En la tabla 2.1 se presentan algunas propiedades para las que pueden calcularse radios equivalentes, las letras por las que se denotan y la expresión que se emplea para su cálculo.

Propiedad	Símbolo	Letra	Radio Equivalente
Coefficiente de fricción traslacional	f_t	T	$a_T = f_t/6\pi\eta_0$
Tiempos de relajación rotacionales	τ	R	$a_R = (3k_B T/4\pi)^{1/3}(\tau/\eta_0)^{1/3}$
Viscosidad intrínseca	$[\eta]$	I	$a_I = (3/10\pi N_A)^{1/3}([\eta]M)^{1/3}$
Covolumen	u	C	$a_C = (3/32\pi)^{1/3}u^{1/3}$
Volumen	V	V	$a_V = (3/4\pi)^{1/3}V^{1/3}$
Radio de Giro	R_g	G	$a_G = (5/3)^{1/2}R_g$
Dimensión máxima	L	L	$a_L = L/2$

TABLA 2.1: Radios equivalentes para las distintas propiedades.

Así, los radios equivalentes cumplen con tres propiedades importantes:

1. Para una partícula esférica compacta homogénea, de densidad uniforme, los radios equivalentes para las distintas propiedades coincidirán en un único valor: el radio de la propia partícula esférica.
2. Para una partícula no esférica, los radios equivalentes para las distintas propiedades no serán iguales, pero no serán muy diferentes entre sí.

3. El valor de los radios equivalentes depende del tamaño de la partícula expresado como una dimensión lineal, pero no volumétrica.

El uso de los radios equivalentes a la hora de comparar propiedades de partículas supone una ventaja importante. Al comparar directamente los valores de las propiedades de las partículas en disolución, es necesario tener una idea de la variabilidad de cada una de las propiedades para poder decidir si la diferencia es más o menos grande.

En cambio, al utilizar los radios equivalentes, todas las propiedades se ven reducidas a una magnitud común: el tamaño de una partícula esférica equivalente. De esta manera, comparar el grado de semejanza entre las propiedades de diferentes partículas, o la influencia de un determinado efecto sobre una propiedad, es tan simple como comparar el tamaño de las esferas equivalentes.

Por este motivo, el empleo de radios equivalentes, en lugar de las propiedades a partir de los que se calculan, está especialmente indicado cuando se analizan conjuntamente varias propiedades para varios tipos de solutos, muestras, etc, como nuestro grupo demostró en [20].

Además, existe una ventaja adicional en cuanto al uso de este tipo de magnitudes. En el estudio de las disoluciones macromoleculares se emplean habitualmente “funciones universales”, entendiéndose por tales combinaciones adimensionales de dos propiedades en disolución que adoptan valores numéricos característicos que dependen del tipo de conformación, pero no de la constitución química, tamaño, o peso molecular de la macromolécula, como por ejemplo, la función β de Scheraga-Manfalkern, las funciones P o Φ de Flory, etc [2]. Basándonos en los radios equivalentes también pueden definirse combinaciones adimensionales de este tipo, definiendo cocientes de dos radios equivalentes cualesquiera, de forma que podemos definir una relación del tipo:

$$XY \equiv a_X/a_Y. \quad (2.12)$$

Como consecuencia de las propiedades de los radios de giro equivalentes, es fácil entender que para una partícula esférica homogénea, todas las posibles relaciones de radios equivalentes sean iguales a la unidad, y cuando más difiere la conformación de la forma esférica, más se desvían de la unidad estas relaciones.

2.2. Modelos de bolas y muelles.

Para poder realizar simulaciones del comportamiento de macromoléculas es necesario, en primer lugar, elaborar un modelo gobernado por expresiones matemáticas que represente fielmente a la molécula física. Estos modelos pueden elaborarse con distintos niveles de detalle, que resultarán en un grado de complejidad creciente cuanto mayor sea el número de elementos implicados en el modelo.

Dada la línea de trabajo habitual del Grupo en el que se ha realizado, en esta Tesis se emplearán modelos de “grano grueso” (o “*coarse-grained*”, en inglés), es decir, de un grado de detalle bajo, en los que un elemento del modelo representa un número variable de átomos, según resulte apropiado en cada caso. En concreto, a los efectos de emplear las técnicas de simulación que se detallarán en las secciones siguientes se empleará el tipo de modelo conocido como “de bolas y muelles” [21].

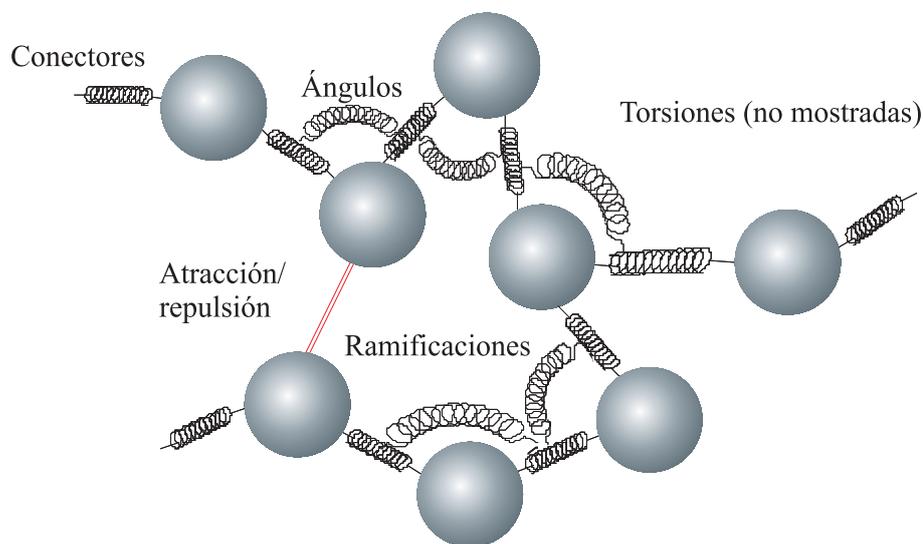


FIGURA 2.1: Ejemplo de un modelo de bolas y muelles.

Esta clase de modelos, como su nombre indica, están compuestos por dos tipos de elementos: las **bolas** y los **muelles**.

Las primeras representan la entidad física de la cadena simulada, siendo los lugares donde ocurren las interacciones, y poseerán como características unas coordenadas de posición, r_i , un radio α , y una masa m_i . La extensión de molécula representada por una

bola es variable, y dependerá de la parametrización que se haga de la misma, pudiendo corresponder a un grupo de átomos, residuos de aminoácidos, monómeros, subcadenas, etc.

En cuanto a los muelles, éstos definen la topología de la molécula, conectando las bolas en una secuencia determinada, así como al mismo tiempo representan las interacciones de origen principalmente entrópico existentes entre las partes de la molécula que representan las bolas y que se oponen a que éstas se alejen de la situación de equilibrio. Para ello, se definen diferentes tipos de muelles, que representan diferentes tipos de interacciones y que se caracterizan por una determinada expresión de la energía potencial V . Una conformación concreta representada por un modelo de bolas y muelles será tanto más estable cuando menor sea el potencial total de la misma.

En esta Tesis se emplearán muelles que conectan las bolas consecutivas (muelles conectores), los cuales forman el esqueleto de la cadena, y, en ocasiones, muelles que conectan tres bolas vecinas (muelles angulares). Además, las bolas no conectadas por muelles de los tipos anteriores pueden interactuar a través de diversos potenciales que tienen en cuenta el efecto de volumen excluido.

En el esquema que se muestra a continuación se reúnen los diferentes tipos de potenciales de muelles y potenciales de volumen excluido que se han empleado en esta Tesis:

1. Los **muelles conectores**: estos muelles son los que definen la topología de la molécula, conectando cada bola con la siguiente (o siguientes, en el caso de que exista una ramificación). Los conectores empleados en esta Tesis son de los siguientes tipos:

- **Muelle gaussiano**: definido por un potencial de tipo hookeano (es decir, cuadrático respecto a su elongación) que representa subcadenas con una distribución gaussiana de la distancia extremo a extremo. Su expresión corresponde a:

$$V^{con}(l) = \frac{1}{2}Hl^2, \quad (2.13)$$

donde H es la constante de fuerza del muelle y l la elongación instantánea del conector.

- **Muelle Fraenkel [22] o hookeano duro**: como su nombre indica, también viene definido por un potencial de tipo hookeano, apropiado para modelos en los que la longitud de los conectores se desvía poco de la longitud de equilibrio.

$$V^{con}(l) = \frac{1}{2}H(l - l_e)^2, \quad (2.14)$$

donde, de nuevo, H representa a la constante de fuerza del muelle, l a la elongación instantánea del conector, y l_e es la longitud en equilibrio del conector.

- **Muelle FENE** (de “Muelle de Extensibilidad Finita y Elongación No-Lineal”, por sus siglas en inglés) [23]; se trata, como puede observarse a partir de la expresión de su potencial, de un muelle cuya extensibilidad es limitada, y además, el potencial acumulado aumenta de forma no lineal.

$$V^{con}(l) = -\frac{1}{2}Hl_{max}^2 \ln \left(1 - \frac{l^2}{l_{max}^2} \right); \quad (2.15)$$

como en los casos anteriores, H es la constante del muelle, l_{max} es la máxima elongación del muelle y l es su longitud instantánea.

- **Muelles “Hard-FENE”**. Se trata de un desarrollo original por parte de nuestro Grupo que representa un comportamiento híbrido entre el de un hookeano duro, y un muelle de tipo FENE. Ha mostrado buenos resultados en trabajos anteriores, como el de que se ha mencionado sobre simulación de dendrímeros [24]. Su expresión para el potencial en función de su longitud instantánea l corresponde a:

$$V^{con}(l) = -\frac{1}{2}Hl_{max}^2 \ln \left(\frac{l_{max}^2 - l^2}{l_{max}^2 - l_e^2} \right) - \frac{1}{2}Hl_{max}l_e \ln \left[\frac{(l_{max} + l)(l_{max} - l_e)}{(l_{max} - l)(l_{max} + l_e)} \right]. \quad (2.16)$$

Como puede deducirse de la expresión, el comportamiento se aproxima al de un muelle FENE cuando la longitud de equilibrio, l_e , es próxima a cero, y a un muelle hookeano duro cuando la longitud máxima, l_{max} , tiende a infinito.

2. Los **Muelles angulares**: se emplean para definir restricciones de tipo angular entre grupos de tres bolas consecutivas. Generalmente, suele emplearse un potencial de tipo cuadrático en función del ángulo determinado por los dos conectores que unen las tres bolas:

$$V^{ang}(\alpha) = \frac{1}{2}Q(\alpha - \alpha_e)^2. \quad (2.17)$$

La expresión se evalúa respecto al valor instantáneo del ángulo externo o suplementario definido por ambos conectores, α ; α_e representa el valor de equilibrio de este ángulo, y Q es una constante de flexión que rige la desviación de α respecto a α_e .

3. **Interacciones de largo alcance: efecto de Volumen Excluido.** Este tipo de interacciones son las que se dan entre pares de bolas i, j que no están directamente relacionadas mediante conectores o ángulos, y tienden a impedir el solapamiento entre partes alejadas a lo largo de la secuencia de la cadena pero que, ocasionalmente, pueden estar próximas en el espacio. En la mayoría de los casos, independientemente del tipo de potencial empleado, suele definirse un “radio de corte” o *cut-off* a partir del cual se considera que el potencial no tiene efectos apreciables, y se omite su cálculo para ahorrar tiempo de procesador.

Para este tipo de interacción también hay múltiples elecciones posibles en cuanto al tipo de potencial. Las que se usan a lo largo de esta Tesis, son las siguientes:

- **Potencial de esferas duras** : se trata de un potencial simple, puramente repulsivo, que se define como infinito, $V^{VE} = \infty$, en el caso de que las dos bolas que interactúan se solapen, es decir, si $|r_j - r_i| < 2\sigma_{HS}$, siendo σ_{HS} el radio que se asigne a las esferas, mientras que en cualquier otro caso ($|r_j - r_i| \geq 2\sigma_{HS}$), el potencial no actúa ($V^{VE} = 0$).

A pesar de la sencillez de este potencial, a la hora de implementarlo existen dificultades, como por ejemplo, la complejidad de implementar el concepto de infinito en un programa, por lo que en la práctica, se recurre a sustituir el valor de infinito por un valor constante muy elevado. Otro inconveniente consiste en que, al no ser un potencial derivable no puede usarse directamente en simulaciones de dinámica browniana, ya que en este caso, en lugar de con expresiones de potencial, se trabaja con fuerzas (que se obtienen derivando los potenciales), ya que puede producir problemas numéricos durante la simulación.

- **Potencial de Lennard-Jones**, ampliamente conocido y descrito; se trata de un potencial de uso muy generalizado debido a su simplicidad y la característica de contemplar tanto atracciones como repulsiones. Su expresión,

$$V^{VE}(r_{ij}) = 4\epsilon_{LJ} \left[\left(\frac{\sigma_{LJ}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{LJ}}{r_{ij}} \right)^6 \right], \quad (2.18)$$

se define en base a los parámetros ϵ_{LJ} , la energía potencial mínima atractiva alcanzable, σ_{LJ} , la distancia a la que el potencial se anula, y r_{ij} , la distancia entre los centros de las dos bolas implicadas.

- **Potencial “suave”**: se trata de un potencial puramente repulsivo, interesante por su sencillez y porque presenta una dependencia suave con la distancia, lo que supone una ventaja en simulaciones de dinámica browniana. La expresión que define el potencial es la siguiente [25]:

$$V^{VE}(r_{ij}) = A \exp(-\beta r_{ij}). \quad (2.19)$$

En ella, r_{ij} representa la distancia entre los centros de las bolas que interactúan, y A y β son dos parámetros arbitrarios.

- **Potencial “armónico”:** este potencial se ha utilizado puntualmente para comparar los algoritmos de dinámica browniana con interacción hidrodinámica. Este potencial implica definir muelles hookeanos duros entre todas las bolas:

$$V^{VE}(r_{ij}) = \frac{1}{2} H (r_{ij} - 2\sigma)^2, \quad (2.20)$$

donde r_{ij} es la distancia entre los centros de las bolas relacionadas, H una constante de fuerza, y σ el radio de las bolas del modelo. Al igual que el potencial de esferas duras, es siempre repulsivo, y actúa sólo cuando $r_{ij} < 2\sigma$, es decir, cuando las bolas se solapan. Dado que este potencial sí es derivable, su uso en simulaciones de dinámica browniana sí es posible.

2.3. Procedimiento de parametrización multiescala.

Las simulaciones de propiedades estructurales y dinámicas de un sistema químico pueden, en principio, ser realizadas mediante modelos atómicos. Los actuales avances teóricos y computacionales, en diversos campos tales como la química cuántica, la termodinámica estadística, y la creciente potencia de los equipos informáticos, propician el trabajo a ese nivel, y hoy en día, la simulación de dinámica molecular con detalle atómico es un método ampliamente utilizado.

Una macromolécula en disolución puede ser modelada con detalle atómico sin excesiva dificultad. El problema reside en elevadísimo número de átomos que compondrían el sistema, pues con facilidad pueden encontrarse macromoléculas compuestas por miles de átomos, a los que habría que añadir varios miles más de moléculas de disolvente para representar un sistema diluido. Utilizando métodos de dinámica molecular atómica, la simulación de la evolución de un sistema de unos 10^4 átomos durante un microsegundo puede requerir varios meses de cálculo en equipos de cálculo de moderada potencia.

Algunos fenómenos de interés en la dinámica de moléculas suceden en un espacio de tiempo de entre un picosegundo y un nanosegundo, por lo que no requieren simulaciones tan largas, y para los cuales la escala atómica es factible. Sin embargo, la dinámica global de grandes macromoléculas tiene lugar en tiempos mucho más largos, llegando a estudiarse periodos de tiempo del orden de los milisegundos. Evidentemente, esos tiempos son inaccesibles a la dinámica molecular con detalle atómico y disolvente explícito.

Desde hace tiempo, la dinámica de macromoléculas en disolución ha sido estudiada teóricamente, y luego computacionalmente, mediante modelos simplificados en los que cada elemento representa no un átomo, sino una porción bastante grande de la molécula, a la vez que se considera al disolvente como un fluido viscoso continuo. Estos modelos con bajo grado de detalle corresponden a los ya mencionados modelos de “grano grueso”.

Por medio del uso de este tipo de modelos se consigue reducir de forma muy importante el número de elementos en el modelo frente al que habría en uno equivalente con detalle atomístico, de forma que en la práctica, el número de elementos en una simulación suele oscilar entre veinte y doscientos.

Los modelos construidos están gobernados por una serie de expresiones matemáticas como las que ya se han descrito en la sección 2.2 para el modelo de “bolas y muelles”, y contienen diversos parámetros ajustables, cuyo valor se suele fijar en base a propiedades medidas en el laboratorio. Esta forma de proceder contrasta con la de la dinámica molecular atómica en la que (una vez adoptado un campo de fuerzas), no hay parámetros ajustables.

Para aprovechar las respectivas ventajas de los modelos de grano grueso y atomísticos, y superar algunos de los inconvenientes de ambos, se está extendiendo la tendencia de efectuar simulaciones multiescala [26, 27]. En nuestro Grupo de Investigación, se ha desarrollado y empleado con éxito uno de estas estrategias multiescala en un trabajo anterior de simulación de dendrímeros [24], y también será empleado en esta Tesis en los Capítulos 6 y 7.

En concreto, en este tipo de procedimientos se emplea un modelo de grano grueso para representar la dinámica global, debido a la ventaja de poder estudiar sistemas que a escala atómica serían inabordables y los menores requisitos de potencia de cálculo. En cambio, los parámetros de este modelo, en lugar de ser ajustados frente a propiedades experimentales, serán calculados mediante simulaciones con detalle atómico de las entidades químicas que representan las bolas y muelles que constituyen los modelos de grano grueso, pues como estas entidades son mucho menores que la macromolécula global en cuanto a número de

átomos, su simulación mediante técnicas de dinámica molecular con detalle atómico sí es factible.

A continuación, se describe el procedimiento de parametrización multiescala empleado en esta Tesis, que puede dividirse en dos fases: la parametrización de las bolas, a través de su radio, y la parametrización de los muelles desde el punto de vista de las interacciones a las que representan. En cuanto a las interacciones de largo alcance, debido a las particularidades de cada una de las expresiones empleadas a lo largo de esta Tesis, se ha optado por incluir los detalles sobre la parametrización de cada caso en su correspondiente sección en lugar de hacerlo aquí.

2.3.1. Parametrización del radio de las esferas

Para el cálculo de las propiedades hidrodinámicas es necesario asignar un radio σ a las esferas que componen el modelo. Pueden darse casos en los que sea conveniente construir el modelo con más de un tipo de bolas. Por ejemplo, en el Capítulo 7 se presenta el caso de copolímeros compuestos por dos tipos de monómeros, de manera que los modelos correspondientes se construyen también con dos tipos de bolas diferentes. En estos casos es necesario determinar por separado el radio de cada tipo de bola, siguiendo el proceso que se describe a continuación para cada uno de los tipos.

Puesto que en el procedimiento multiescala no van a emplearse parámetros ajustables, este radio ha de estimarse a partir de detalles atómicos de la estructura a la que representa. De esta manera, al estar compuesto el modelo por bolas repetidas, se puede considerar que la molécula simulada está compuesta por repeticiones de esta estructura, siendo ésta una especie de “pseudo-monómero” (que no tiene por qué corresponder con la estructura del monómero real, en caso de que la macromolécula en cuestión sea un polímero).

Dicho pseudo-monómero presentará una cierta variabilidad conformacional, dentro de la cual existirán ciertas variaciones en el tamaño y otras propiedades de las conformaciones adoptadas, y por tanto, es necesario llevar a cabo una simulación (atomística, en este caso) que permita barrer un espacio conformacional representativo y poder calcular así un valor promedio de σ . En nuestro procedimiento, se emplea para ello una dinámica molecular de Langevin mediante el programa comercial HYPERCHEM [28], utilizando el campo de fuerzas “AMBER”.

De las conformaciones que componen la trayectoria obtenida (listados temporales compuestos por las coordenadas de todas las bolas del modelo en cada instante de la simulación) se extrae una muestra representativa, por ejemplo, tomando una de cada X conformaciones. En general los pseudo-monómeros serán estructuras poco flexibles, y su variabilidad conformacional será limitada, por lo que no es necesario tomar un número demasiado elevado de conformaciones para tener una muestra representativa.

Sobre estas conformaciones se calculan las propiedades hidrodinámicas de cada una de ellas aplicando el tratamiento de cuerpo rígido [29] mediante el programa HYDROPRO [30], desarrollado por nuestro Grupo de Investigación, y empleando valores estándar para la temperatura, los radios atómicos, y otros parámetros de simulación.

Los resultados de los cálculos con HYDROPRO dependen del valor asignado al radio atómico de los elementos del modelo, “AER” (“*atomic equivalent radius*”) [30]. Se sabe [31] que para las moléculas pequeñas, una opción apropiada para este radio es el radio de Van der Waals de los átomos extendidos (con los hidrógenos implícitos). Para los átomos típicos de C, de O y de N, este radio es cercano a 1.8 Å [32], y éste será el valor que se utilice en los trabajos de esta Tesis.

HYDROPRO proporciona, entre otras propiedades, el radio de Stokes (que corresponde al radio equivalente respecto al coeficiente de fricción traslacional, a_T) y el radio de giro R_g de la estructura. A partir de este último se obtiene su radio equivalente, a_G . De acuerdo con lo enunciado al presentar los radios equivalentes en la sección 2.1.4, los radios a_T y a_G deben ser bastante similares (recordemos que las relaciones de radios suelen ser próximas a uno), por lo que puede calcularse un valor medio y tomarlo como el radio equivalente del pseudo-monómero.

2.3.2. Parametrización de conectores y ángulos

Como se indicaba en la sección 2.2, en los modelos de bolas y muelles las bolas vecinas están unidas por conectores parcialmente flexibles, representados por muelles. Éstos tienen asociado un potencial $V^{con}(l)$ en función de l , la longitud del conector en una conformación instantánea. La probabilidad de que un conector tenga una longitud instantánea determinada, $p(l)$, viene determinada por la exponencial de Boltzmann de este potencial:

$$p(l) = A^{con} l^2 \exp\left(\frac{-V^{con}(l)}{k_B T}\right), \quad (2.21)$$

donde A^{con} es una constante de normalización.

En el caso de emplearse restricciones angulares en el modelo, éstas también deberán parametrizarse de una manera análoga. Los ángulos determinados por dos conectores vecinos están restringidos por un potencial de flexión $V^{ang}(\alpha)$ (de nuevo, ver sección 2.2). Si se tiene una terna de bolas i, j, k , donde j es la bola central, definimos α como el ángulo externo entre los vectores \mathbf{ij} y \mathbf{jk} . La probabilidad asociada a este potencial de flexión, $p(\alpha)$, es, otra vez, el producto del factor de Boltzmann por una constante de normalización, A^{ang} , pero en este caso es necesario añadir también el producto por un factor geométrico $\sin(\alpha)$:

$$p(\alpha) = A_{ang} \sin(\alpha) \exp\left(\frac{-V^{ang}(\alpha)}{k_B T}\right). \quad (2.22)$$

Tanto en el caso de los muelles conectores como en los angulares, es necesario determinar los valores apropiados de los parámetros de las expresiones de sus potenciales. Al igual que se hace con el radio de las bolas, el valor de estos parámetros se determinará a partir de simulaciones atomísticas realizadas con HYPERCHEM.

Como ya se ha mencionado en la parametrización de las esferas, en un mismo modelo puede haber más de un tipo de bolas. En tal caso, las interacciones representadas por muelles conectores y angulares se darán no sólo entre las bolas de cada uno de los tipos, sino que también habrá casos de interacciones “mixtas” entre bolas de diferentes tipos. Así, será necesario determinar los parámetros de las expresiones de potencial para los conectores y los ángulos determinados por las bolas de cada uno de los tipos, y también los parámetros para los conectores y ángulos que se establezcan entre bolas de tipos diferentes.

Para la parametrización de los conectores y ángulos, la estructura simulada en HYPERCHEM estará compuesta por varios pseudo-monómeros, formando una estructura mínima en la que se dé cada una de las interacciones que sea necesario parametrizar.

En la figura 2.2 se presenta el esquema de una de estas estructuras y el correspondiente modelo de HYPERCHEM extraídos del trabajo realizado para el Capítulo 6. En concreto, esta estructura mínima está compuesta por cuatro pseudo-monómeros dispuestos en forma de “Y” (en la figura, la “Y” está dispuesta en horizontal). De esta forma, en esta estructura

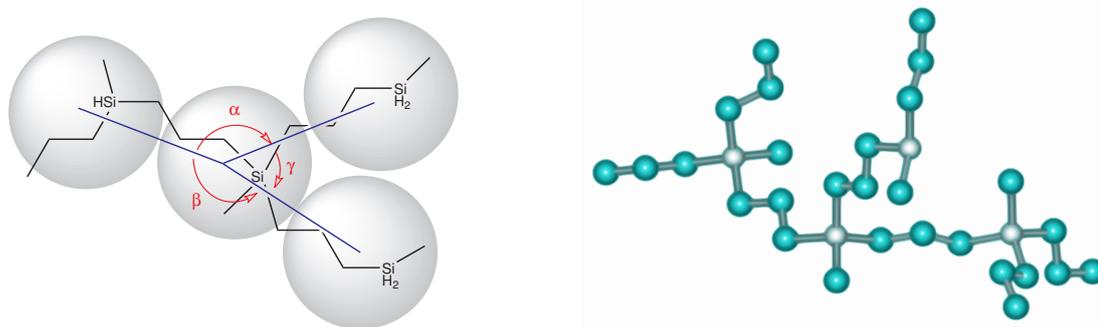


FIGURA 2.2: Esquema del modelo y modelo con hidrógenos omitidos en HYPERCHEM de la estructura mínima de PCS3 (ver Capítulo 7).

se definen tres muelles conectores (líneas en azul), y tres ángulos (representados por los arcos de circunferencia en rojo), de los que dos son ángulos “externos” (α y β , definidos entre el “pie” y un “brazo” de la “Y”), y un ángulo “interno” (γ , definido entre los dos “brazos” de la “Y”). Estos cuatro tipos de muelles son los mismos que presenta todo el modelo global que representará a la molécula completa en la posterior simulación de “grano grueso”, y la simulación y el análisis de esta estructura mínima permitirá obtener datos suficientes para parametrizarlos.

De la trayectoria generada a partir de la simulación de la estructura mínima se extraen las longitudes instantáneas de los conectores y los ángulos instantáneos entre dos conectores consecutivos. Por analogía con los modelos de grano grueso, en los que los muelles se definen entre los centros de las bolas, los conectores definidos deberían extenderse de un centro de bola a otro, aunque en nuestro Grupo se ha encontrado que, dado que estas estructuras son poco flexibles, el punto exacto sobre el que se sitúa el origen de los conectores tiene poca importancia (ya sea el centro de cada bola, o un átomo en concreto), siempre y cuando este punto sea el mismo en todas las bolas.

Con los datos extraídos de la trayectoria se construyen las distribuciones de probabilidad de longitud de los conectores, $p(l)$, y de los ángulos, $p(\alpha)$. Los datos de dichas distribuciones se ajustan a la expresión de la función de probabilidad del tipo de muelle en cuestión (obtenida utilizando las ecuaciones 2.21 o 2.22, combinadas con la del potencial del muelle apropiado).

Para realizar este ajuste, empleamos el programa FUNFIT, un programa auxiliar desarrollado por nuestro Grupo que permite calcular los valores de los parámetros de una función suministrada por el usuario que hacen que dicha función se ajuste lo mejor posible a unos datos de referencia. El programa funciona barriendo los rangos de valores

que define el usuario para cada parámetro, en busca de la combinación que produzca la menor desviación cuadrática respecto a los datos introducidos como referencia. Como salida, el programa proporciona los valores óptimos encontrados, además de un listado de las desviaciones estándar calculadas para cada combinación de parámetros, de forma que pueden realizarse gráficos en forma de “mapas de desviación”, sobre los que se puede estudiar la influencia de cada parámetro en el ajuste.

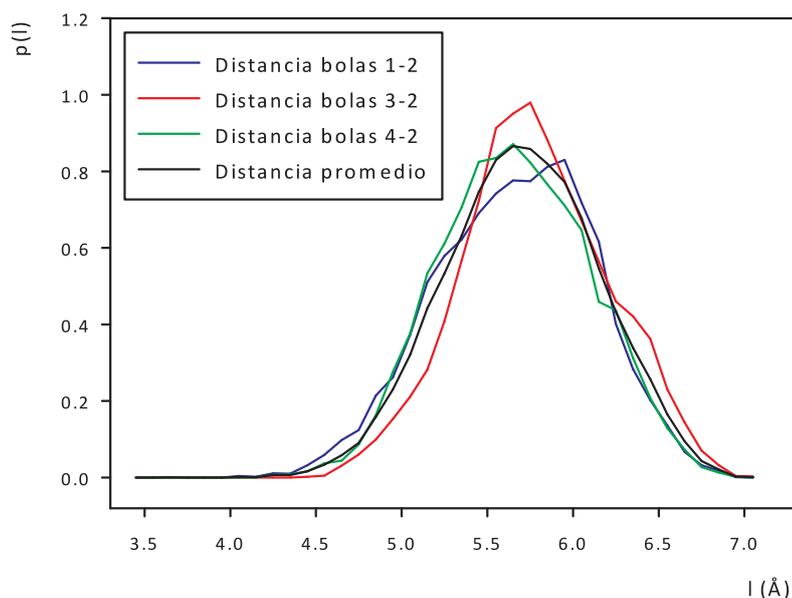


FIGURA 2.3: Funciones de distribución de la longitud de los conectores observados en PCS3.

Obviamente, las ecuaciones de los potenciales de los muelles (tanto los conectores como los angulares) no pueden reproducir los detalles más finos de las distribuciones de probabilidad obtenidas en las simulaciones con detalle atómico, no obstante, proporcionan una buena representación de la estadística de los conectores y los ángulos, como se mostrará en la aplicación de este procedimiento de parametrización en los Capítulos 6 y 7. En las figuras 2.3, 2.4 y 2.5 se presentan, respectivamente, los ejemplos de la distribución de probabilidad de longitudes de los conectores, el gráfico obtenido del ajuste obtenido mediante FUNFIT respecto al promedio de la distribución de probabilidades, y el “mapa de desviación” obtenidos para la estructura mínima presentada en la figura 2.2.

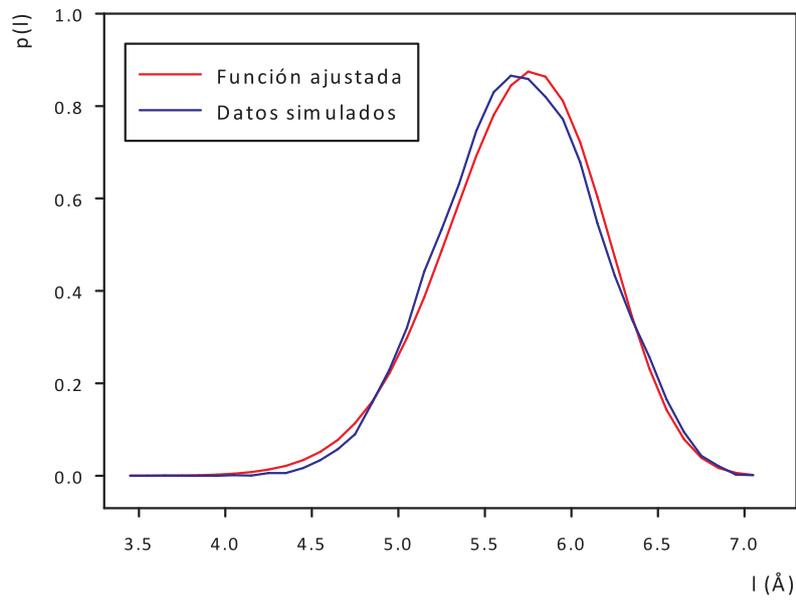


FIGURA 2.4: Ajuste de los parámetros de los conectores en PCS3.

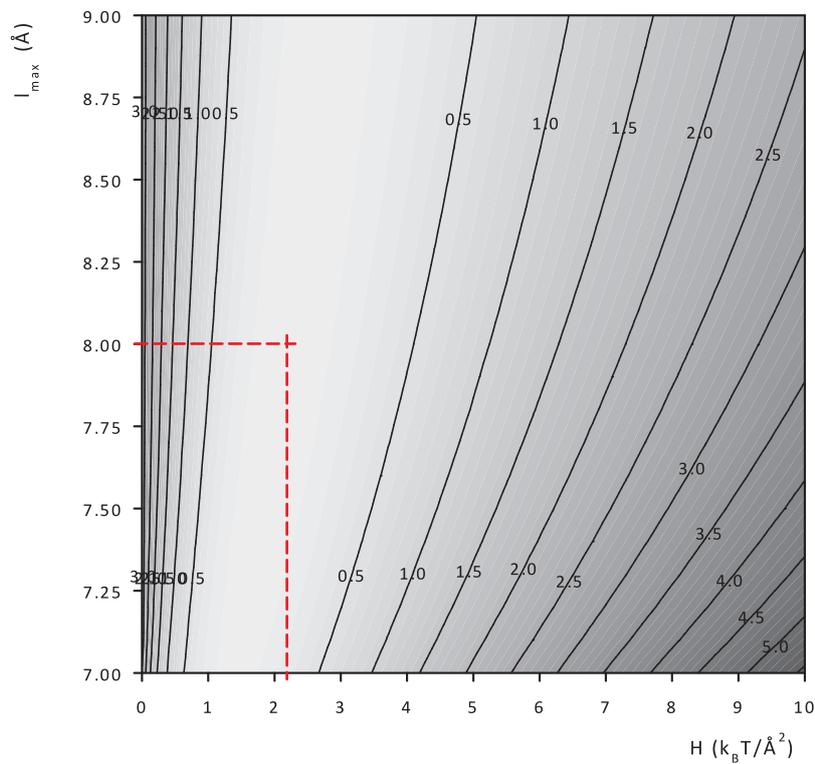


FIGURA 2.5: Desviaciones cuadráticas respecto a los parámetros de los conectores en PCS3. Las líneas discontinuas marcan los valores de H y l_{\max} que hacen que la desviación cuadrática sea mínima, por lo que serán utilizados en la simulación.

2.4. Técnicas de simulación: dinámica browniana.

El movimiento browniano se conoce como tal en honor al biólogo escocés Robert Brown, quien fue el primero en describir este fenómeno y estudiarlo sistemáticamente en 1827, aunque ya había sido observado previamente, incluso en la antigüedad clásica. Brown estudiaba al microscopio suspensiones en agua de granos de polen y esporas de musgo cuando observó la existencia de pequeñas partículas que se separaban de los granos de polen y ejecutaban un continuo movimiento aleatorio. Para explicar este movimiento, Brown formuló inicialmente la hipótesis de que los granos de polen tenían vida. Más adelante, retiraría esta hipótesis al observar el mismo comportamiento en partículas de polvo en suspensión acuosa, y finalmente sería incapaz de postular el origen de estos movimientos.

A pesar de que a lo largo de los años posteriores se hicieron descripciones de otros fenómenos relacionados con el movimiento browniano, e incluso se llegó a formular alguna teoría matemática al respecto de éste, no fue hasta 1905 cuando Einstein publicó un artículo [11] en el que presentaba una ecuación que caracterizaba, por medio de un coeficiente de difusión, el movimiento browniano de una partícula inmersa en un fluido. Fue, precisamente, con este artículo y su validación experimental años después, con lo que Einstein conseguía demostrar definitivamente la existencia de átomos y moléculas, de cuya existencia física aún se dudaba en aquella época.

Desde entonces, la teoría sobre el movimiento browniano se ha desarrollado, y actualmente, para cualquier partícula inmersa en un fluido, se conoce la existencia de las llamadas “fuerzas brownianas”, causantes del movimiento aleatorio observado por Brown, y cuyo origen se encuentra en los choques aleatorios de las moléculas del fluido sobre la superficie de la partícula, como producto de la agitación térmica.

La consideración de dicha fuerza aleatoria, $\mathbf{F}_i^{(Bro)}$, en la ecuación de la Ley de Newton, junto a las fuerzas internas de la partícula, $\mathbf{F}_i^{(int)}$, y la fuerza de fricción con el disolvente al producirse el movimiento, $\mathbf{F}_i^{(fri)}$, conducen a la formulación de la ecuación de Langevin,

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i^{(int)} + \mathbf{F}_i^{(fri)} + \mathbf{F}_i^{(Bro)}. \quad (2.23)$$

El término inercial del primer miembro de esta expresión está relacionado con fenómenos muy rápidos y de muy corto alcance, como es, por ejemplo, la vibración de los enlaces químicos. Este tipo de fenómenos ocurren en intervalos de tiempo ínfimos, y al promediarse sus efectos sobre un intervalo de tiempo de orden de magnitud superior, como los que suelen ser empleados en la aplicación de esta ecuación a las simulaciones del tipo de esta Tesis, se anulan entre sí. Por este motivo, la particularización de esta expresión para el caso de moléculas de cierto tamaño, como las macromoléculas que se tratan en esta Tesis, conlleva que el término inercial sea despreciable frente al resto y se tome como nulo el primer miembro de la ecuación.

Hecha esta primera consideración, la ecuación resultante representa un sistema de ecuaciones diferenciales de tipo estocástico, debido a la naturaleza aleatoria de las fuerzas brownianas y a la necesidad de darles tratamiento de tipo estadístico.

Mediante la discretización de dicho sistema de ecuaciones pueden construirse algoritmos que permiten representar fielmente el movimiento de una partícula en disolución, y calcular y registrar los datos que describen la posición de sus elementos en cada instante. La generación de tal registro mediante uno de estos algoritmos se conoce como simulación de dinámica browniana.

2.4.1. Ecuaciones para la dinámica browniana.

El algoritmo de Ermak-McCammon.

Una solución particular de la ecuación de Langevin es la presentada por Ermak y McCammon en 1978 [6]. Estos autores presentan dos ecuaciones, una de ellas incluyendo el término inercial y la otra sin él. Como ya se ha comentado, en los casos que se recogen en esta Tesis el término inercial tiene muy poca importancia, por lo que en adelante se entenderá que nos referimos a la ecuación que lo ignora, que es la que en rigor se conoce como dinámica browniana.

La integración numérica de la ecuación de Langevin siguiendo un esquema de primer orden da como resultado la siguiente expresión:

$$\mathbf{r}_i = \mathbf{r}_i^0 + \frac{\Delta t}{k_B T} \sum_{j=1}^N \mathbf{D}_{ij}^0 \cdot \mathbf{F}_j^{(int)(0)} + \Delta t \sum_{j=i}^N \left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}_j} \right)^0 + \mathbf{R}_i + (\Delta t) \mathbf{v}_i^0. \quad (2.24)$$

En esta expresión, particularizada para un modelo de bolas y muelles como los desarrollados en la sección 2.2, \mathbf{r}_i representa las coordenadas de la bola i del modelo (habrá una ecuación similar para cada bola del modelo); \mathbf{D}_{ij} es cada uno de los bloques 3×3 que forman el llamado “tensor de difusión” (sobre el que se tratará más adelante, en la sección 2.4.2 de este mismo Capítulo); $\mathbf{F}_j^{(int)}$ se refiere al vector resultante del cómputo de las fuerzas internas que actúan sobre la bola j (incluyendo los conectores y las interacciones de volumen excluido); Δt es la amplitud del paso temporal o paso browniano empleado en la discretización de la ecuación de Langevin; \mathbf{R}_i es un vector aleatorio con distribución gaussiana tal que su promedio y su covarianza sea, respectivamente, $\langle \mathbf{R}_i \rangle = \mathbf{0}$ y $\langle \mathbf{R}_i \mathbf{R}_j \rangle = 2D_{ij}\Delta t$, con $i, j = 1 \dots N$ (donde N representa al número de bolas del modelo); y el término $(\Delta t)\mathbf{v}_i$ es una modificación que refleja la posibilidad de tener en cuenta un flujo macroscópico en la disolución, representado por el vector velocidad \mathbf{v}_i de la porción de fluido que ocuparía la posición de la bola i . Finalmente, la incorporación del superíndice 0 a una magnitud indica que ésta toma el valor correspondiente a su estado inicial anterior a producirse el paso de simulación.

Así pues, esta expresión de la ecuación de Ermak y McCammon constituye un algoritmo numérico de simulación de primer orden, en el mismo sentido en el que lo es la solución de Euler en el ámbito habitual de las ecuaciones diferenciales.

El algoritmo de Iniesta-García de la Torre.

Al igual que en el campo de la ecuaciones diferenciales el método de segundo orden de Runge-Kutta supone una mejora de la solución de Euler, Iniesta y García de la Torre [33] proponen un método de simulación de “pseudo-segundo orden”, basado en el procedimiento de Runge-Kutta y la ecuación de Langevin.

En este algoritmo, de tipo “predictor-corrector”, cada paso de simulación se realiza en dos etapas: en primer lugar, se hace un cálculo “predictivo” empleando en la ecuación de primer orden de Ermak y McCammon las coordenadas iniciales \mathbf{r}_i^0 junto con los correspondientes valores de los vectores de fuerzas $\mathbf{F}_j^{(int)(0)}$, los bloques \mathbf{D}_{ij}^0 del tensor de difusión y el vector velocidad de flujo \mathbf{v}_i^0 , para obtener una estimación de las coordenadas finales, \mathbf{r}'_i :

$$\mathbf{r}'_i = \mathbf{r}_i^0 + \frac{\Delta t}{k_B T} \sum_{j=1}^N \mathbf{D}_{ij}^0 \cdot \mathbf{F}_j^{(int)(0)} + \Delta t \sum_{j=i}^N \left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}'_j} \right)^0 + \mathbf{R}_i + (\Delta t)\mathbf{v}_i^0. \quad (2.25)$$

Esta estimación de las coordenadas finales sirve para recalcular el valor del vector de fuerzas $\mathbf{F}_j^{(int)}$, los bloques \mathbf{D}'_{ij} del tensor de difusión y el vector de flujo \mathbf{v}'_j . Con estos valores estimados y los anteriores calculados a partir de la posición inicial, se obtienen unos valores promedio que se emplean para volver a calcular el paso de simulación, de nuevo a partir de las coordenadas iniciales, y obteniendo esta vez las coordenadas definitivas en un paso “corrector”, \mathbf{r}_j :

$$\begin{aligned} \mathbf{r}_i = & \mathbf{r}_i^0 + \frac{\Delta t}{k_B T} \frac{1}{2} \sum_{j=1}^N (\mathbf{D}_{ij}^0 \cdot \mathbf{F}_j^{(int)(0)} + \mathbf{D}'_{ij} \cdot \mathbf{F}_j^{(int)}) \\ & + \Delta t \frac{1}{2} \sum_{j=i}^N \left[\left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}_j} \right)^0 + \left(\frac{\partial \mathbf{D}_{ij}}{\partial \mathbf{r}_j} \right)' \right] + \mathbf{R}'_i + \Delta t \frac{1}{2} (\mathbf{v}_i^0 + \mathbf{v}'_i). \end{aligned} \quad (2.26)$$

Obsérvese que en este caso se emplea un nuevo vector aleatorio \mathbf{R}'_i , también con media cero, y en este caso, covarianza $\langle \mathbf{R}_i \mathbf{R}_j \rangle = (\mathbf{D}_{ij}^0 + \mathbf{D}'_{ij}) \Delta t$.

Con este algoritmo, *a priori*, el tiempo de CPU necesario para calcular un paso de simulación es aproximadamente el doble que empleando el algoritmo de Ermak y McCammon, y por lo tanto, de emplearse las mismas condiciones de simulación, la duración del proceso sería mucho más lento. Sin embargo, la naturaleza del algoritmo “predicor-corrector” permite usar un paso browniano de amplitud superior al doble de la que puede usarse con el algoritmo de primer orden, por lo que para una simulación de la misma duración será necesario realizar un número de pasos inferior a la mitad, lo que resultaría en un menor tiempo de simulación [33–35].

2.4.2. La interacción hidrodinámica.

En la definición de las fuerzas brownianas se indicaba que el movimiento browniano es el resultado de los choques de las moléculas del disolvente con las partículas inmersas en éste, como consecuencia del movimiento aleatorio de estas moléculas debido a la temperatura.

Sin embargo, ésta no es la única forma de interacción existente entre una partícula y el disolvente en el que se halla inmersa. Al moverse la partícula de forma global, o al producirse un movimiento relativo entre diferentes partes de la partícula, en caso de que ésta sea flexible, la partícula desplaza a su vez a las moléculas del disolvente, transmitiendo

a éstas cierto movimiento que puede traducirse en nuevos choques con otras partes de la molécula, esta vez de naturaleza no aleatoria. En resumen, la interacción de la partícula con el fluido puede traducirse en una transferencia de movimiento de unas partes de la partícula a otras por medio del disolvente. Este efecto es lo que se conoce como interacción hidrodinámica (IH).

En simulaciones de dinámica browniana de modelos de bolas y muelles se tiene en cuenta este tipo de interacción a través del “tensor de difusión”, \mathbf{D} , que consiste en una supermatriz de dimensión $3N \times 3N$, que puede considerarse como compuesta por $N \times N$ submatrices \mathbf{D}_{ij} , de tamaño 3×3 , que definen la interacción entre cada par de bolas i,j en las tres coordenadas espaciales.

Este tensor de difusión \mathbf{D} se construye a partir de lo que se conoce como “tensor de interacción hidrodinámica”, \mathbf{T} , que es un tensor de movilidad traslacional (inverso de la fricción traslacional) que a su vez tiene la misma estructura que \mathbf{D} , consistiendo por tanto en una supermatriz de dimensión $3N \times 3N$, que puede subdividirse en $N \times N$ submatrices \mathbf{T}_{ij} de tamaño 3×3 , de las que cada una especifica la interacción hidrodinámica existente entre cada par de bolas i,j del modelo simulado. La relación entre los bloques que componen ambos tensores es por tanto

$$\mathbf{D}_{ij} = k_B T \left[\delta_{ij} \left(\frac{1}{\zeta_i} \right) \mathbf{I} + (1 - \delta_{ij}) \mathbf{T}_{ij} \right], \quad (2.27)$$

con $i, j = 1 \dots N$, donde δ_{ij} es la función delta de Kronecker. \mathbf{I} es la matriz identidad 3×3 , y ζ_i es el coeficiente de fricción de la bola i .

Como resultado, cada una de las submatrices \mathbf{D}_{ij} indica la relación existente entre los desplazamientos en la dirección de cada una de las tres coordenadas espaciales de cada par de bolas i,j . Además, dado que existe simetría en las interacciones (la interacción entre el par i,j , representada por \mathbf{D}_{ij} , es simétrica respecto a la del par j,i , representada por $\mathbf{D}_{ji} = \mathbf{D}_{ij}^T$), los tensores de difusión serán siempre matrices simétricas.

En la simulación de dinámica browniana de macromoléculas, el uso de un tensor de difusión que incluya el efecto de la interacción hidrodinámica entre los elementos se ha demostrado necesario para que las trayectorias calculadas reflejen de forma realista la dinámica de las macromoléculas [36]. Sin embargo, la incorporación de este tipo de interacción a los cálculos supone un alto coste computacional debido al tamaño de los tensores de difusión y lo costoso de su manipulación.

Hace algunas décadas, antes de la generalización del uso de ordenadores, debido a la complejidad del tratamiento de la interacción hidrodinámica, se empleaban tensores “pre-promediados” que se construyen de forma previa a la simulación, se mantienen constantes durante la misma y que representan el efecto promediado sobre una determinada variabilidad conformacional dentro de la que se espera que se encuentre la trayectoria calculada. De esta manera, el tratamiento del tensor de difusión se hace bastante más ligero, aunque se corre el riesgo de que las conformaciones a lo largo de la simulación difieran demasiado de las empleadas para generar el tensor pre-promediado, y los resultados de la simulación no sean veraces.

Actualmente, y gracias a la proliferación de los medios informáticos, se prefiere el uso de la interacción hidrodinámica “fluctuante”, en la que el tensor de difusión se recalcula con frecuencia para prevenir la situación de que el tensor no sea adecuado a la conformación instantánea adoptada por la macromolécula. Sin embargo, como ya se ha indicado, la inclusión de procedimientos para el tratamiento de la interacción hidrodinámica fluctuante en las simulaciones implica un alto coste computacional, por lo que se han desarrollado diferentes estrategias para el tratamiento de la misma. Durante la realización de esta Tesis se han ensayado y comparado algunas de estas estrategias, y los resultados de este estudio se presentan en el Capítulo 5.

El tensor de Oseen.

Una primera aproximación para el tratamiento de la interacción hidrodinámica consiste en utilizar el tensor de Oseen en la ecuación 2.27 para el cálculo de \mathbf{D} . Éste está formulado sobre la hipótesis de que los elementos constituyentes del modelo son puntuales y carecen de radio, por lo que el uso de este tensor comienza a presentar problemas cuando las distancias entre los elementos del modelo son del orden del tamaño de las bolas.

Habitualmente, el tensor toma la forma de una matriz definida positiva, es decir, una matriz simétrica cuyos valores propios son estrictamente positivos, y como tal, puede descomponerse en factores, mediante, por ejemplo, la factorización de Cholesky, en la cual se basa el algoritmo de simulación de Ermak-McCammon. Cuando el orden de magnitud de las distancias entre los elementos del modelo y el tamaño de las bolas se aproximan demasiado, los tensores de difusión calculados ya no cumplen con la propiedad de ser siempre definidos positivos, y por tanto operaciones como la factorización o la inversión no son aplicables.

Cuando en la ecuación 2.27 \mathbf{T} es el tensor de Oseen, las submatrices o bloques 3×3 que lo componen pueden clasificarse en dos tipos: el primero de ellos corresponde a los bloques identificados por los índices i,i , que estarían situados sobre la diagonal del tensor. Para éstos, la expresión que define las submatrices corresponde a

$$\mathbf{D}_{ii} = (k_B T / \zeta_i) \mathbf{I}, \quad (2.28)$$

donde k_B es la constante de Boltzmann, $\zeta_i = 6\pi\eta_0\sigma_i$ es el coeficiente de fricción de Stokes de una bola de radio σ_i inmersa en un fluido de viscosidad η_0 , e \mathbf{I} es la matriz identidad de dimensión 3×3 .

El segundo tipo de submatrices corresponde a los bloques situados fuera de la diagonal, asociados a los pares de bolas i,j (con $i \neq j$). Para éstos, la expresión para el cálculo de las submatrices 3×3 es

$$\mathbf{D}_{ij} = \frac{k_B T}{8\pi\eta_0 r_{ij}} \left(\mathbf{I} + \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij}} \right), \quad (2.29)$$

donde r_{ij} representa la distancia entre los centros de las bolas i y j , y $\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}$ representa el producto diádico o tensorial del vector \mathbf{r}_{ij} por sí mismo.

El tensor de Rotne-Prager-Yamakawa.

Debido a las limitaciones del tensor de Oseen, estos autores presentan una modificación del mismo [37, 38] que lo mejora. Esta modificación tiene en cuenta el volumen de los elementos, y la posibilidad de que éstos solapen.

En el caso de utilizar el tensor de Rotne-Prager-Yamakawa en la ecuación 2.27, la expresión para los bloques situados sobre la diagonal del tensor de difusión es, evidentemente, la misma empleada cuando se utiliza el tensor de Oseen (ecuación 2.28), mientras que para el caso de los bloques fuera de la diagonal existen dos expresiones. La primera corresponde a un par de bolas i, j tal que $i \neq j$ que no se solapan, es decir, se aplica en el caso en el que la distancia $r_{ij} > 2\sigma$,

$$\mathbf{D}_{ij} = \frac{k_B T}{8\pi\eta_0 r_{ij}} \left(\mathbf{I} + \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij}^2} + \frac{2\sigma^2}{r_{ij}^2} \left(\frac{1}{3} \mathbf{I} - \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij}^2} \right) \right), \quad (2.30)$$

mientras que para el caso en el que existe solapamiento y $r_{ij} < 2\sigma$, se utiliza la expresión

$$\mathbf{D}_{ij} = \frac{k_B T}{6\pi\eta_0\sigma} \left(1 - \frac{9}{32} \frac{r_{ij}}{\sigma} \right) \mathbf{I} + \frac{3}{32} \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij}\sigma}. \quad (2.31)$$

La principal ventaja de este tensor respecto al de Oseen es que es siempre definido positivo, requisito indispensable para el correcto funcionamiento de los algoritmos de dinámica browniana.

El uso de este tensor presenta además una ventaja adicional, que consiste en que el término $\partial \mathbf{D}_{ij} / \partial \mathbf{r}_j$ de la ecuación de Ermak-McCammon se anula, con el consiguiente ahorro en cálculos.

El tensor de García de la Torre y Bloomfield.

Esta última definición del tensor de difusión, fue propuesta por sus autores [39] a partir de una modificación adicional al tensor de Rotne-Prager-Yamakawa, que consiste en la incorporación de la posibilidad de que las bolas del modelo no sean de un tamaño uniforme. De esta forma, se añade esa posibilidad a las ventajas inherentes a esa definición.

En este tercer caso, de nuevo la expresión para los bloques situados sobre la diagonal del tensor de difusión coincide con las anteriores, sin embargo, conviene recordar que es posible que las bolas no sean del mismo tamaño, lo que debe tenerse en cuenta en el cálculo de los respectivos coeficientes de fricción ζ_j .

Finalmente, para el caso de los bloques \mathbf{D}_{ij} (con, $i \neq j$) y $r_{ij} > (\sigma_i + \sigma_j)$, donde σ_i y σ_j son los radios de las bolas,

$$\mathbf{D}_{ij} = \frac{k_B T}{8\pi\eta_0 r_{ij}} \left(\mathbf{I} + \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij}^2} + \frac{\sigma_i^2 + \sigma_j^2}{r_{ij}^2} \left(\frac{1}{3} \mathbf{I} - \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij}^2} \right) \right), \quad (2.32)$$

y, para el caso de \mathbf{D}_{ij} (con $i \neq j$) y $r_{ij} < (\sigma_i + \sigma_j)$, es decir, con bolas solapantes,

$$\mathbf{D}_{ij} = \frac{k_B T}{6\pi\eta_0 \bar{\sigma}} \left(1 - \frac{9}{32} \frac{r_{ij}}{\bar{\sigma}} \right) \mathbf{I} + \frac{3}{32} \frac{\mathbf{r}_{ij} \otimes \mathbf{r}_{ij}}{r_{ij} \bar{\sigma}}; \quad (2.33)$$

tomando $\bar{\sigma} = (\sigma_i + \sigma_j)/2$ [39].

2.5. Técnicas de simulación: Monte Carlo.

Como se adelantaba en la introducción a este Capítulo, el método de Monte Carlo, a diferencia de la dinámica browniana, sólo es aplicable al estudio de sistemas en equilibrio, es decir, sistemas en los que no existe una variación temporal de sus propiedades. Esto se debe a que, particularizando para el caso de las macromoléculas en disolución, en lugar de estudiar las causas de los cambios conformacionales, las simulaciones se hacen generando aleatoriamente conformaciones de la cadena mediante un procedimiento estadísticamente correcto y comprobando posteriormente la validez de las mismas.

Sin embargo, esta técnica presenta una ventaja importante en aquellos casos en los que es aplicable: dado que la posición de los elementos del modelo se determina de forma aleatoria, no es necesario tener en cuenta la interacción hidrodinámica (que aparecerá a posteriori en los cálculos de las propiedades hidrodinámicas), y por lo tanto, al no tener que manejar el tensor de difusión, la potencia computacional necesaria es mucho menor, y las simulaciones son mucho más rápidas que sus equivalentes realizadas mediante dinámica browniana.

Existen diferentes métodos válidos para generar conformaciones para una simulación de tipo Monte Carlo. Uno de los más usados, y en concreto, el que se usa en los trabajos de esta Tesis es el método de Metropolis [40]. El método consiste en la generación sucesiva de conformaciones a partir de una conformación anterior. Para ello, en cada paso de simulación, las coordenadas $r_i, i = 1 \dots N$ de los N elementos de la cadena se obtienen realizando un desplazamiento aleatorio de cada elemento del modelo respecto a su posición previa, es decir:

$$r_i = r_i^0 + \delta_i, \quad (2.34)$$

donde δ_i es un número aleatorio comprendido entre $-\Delta/2$ y $+\Delta/2$, siendo Δ la máxima amplitud del desplazamiento aleatorio. Los valores de δ_i se generan como

$$\delta_i = (U_i - \frac{1}{2})\Delta, \quad (2.35)$$

donde U_i son números aleatorios con una distribución uniforme en el intervalo (0,1).

Una vez generada la conformación, para comprobar su validez, se evalúa la energía potencial total, que, para una cadena de bolas y muelles, viene dada por la expresión

$$V = \sum_{i=1}^{N-1} V_i^m + \sum_i^N \sum_{j,ne}^N V_{ij}^{ne}. \quad (2.36)$$

En la expresión, V_i^m representa a la suma de los potenciales de todos los muelles a los que la bola i está unida, dados por las ecuaciones correspondientes al tipo de muelle en concreto, y V_{ij}^{ne} es el potencial de interacción de la bola i con cualquier otra bola j con la que no esté directamente enlazada por un muelle.

Una vez calculado el potencial total de la nueva conformación, V , y dado que ya se conoce el valor del potencial de la conformación de partida (calculada en el paso anterior), V^0 , tiene lugar la decisión acerca de la validez de la conformación, que es el aspecto más importante del método de Monte Carlo.

Si el potencial no cambia o disminuye, es decir, $V \leq V^0$, la nueva conformación será tan estable como la anterior o más, y por tanto, se acepta y se continúa la simulación empleándola como punto de partida para el siguiente paso.

En el caso contrario, con $V > V^0$, la nueva conformación es menos estable que la anterior. Si se rechazase directamente, al repetir el procedimiento durante un número elevado de veces, el algoritmo convergería hacia la conformación de menor energía posible y una vez alcanzada, la simulación se estancaría en ella. Para evitar esto, es necesario introducir cierta entropía en la simulación, imitando el efecto de las fluctuaciones térmicas, y aceptando algunas conformaciones de potencial mayor al anterior.

Para decidir si la nueva conformación se acepta a pesar del aumento de potencial, se genera un nuevo número aleatorio U , también con distribución uniforme en el intervalo (0,1), y se compara con la exponencial de Boltzmann del incremento de potencial:

- Si $U \leq \exp\left[\frac{(V^0-V)}{k_B T}\right]$, entonces, la nueva conformación se acepta.
- Si $U > \exp\left[\frac{(V^0-V)}{k_B T}\right]$, entonces, la conformación generada se rechaza, y se emplea un duplicado de la conformación anterior.

Es fácil concluir que, cuanto menor sea el aumento de potencial respecto a la conformación anterior, mayores son las probabilidades de que la conformación sea aceptada.

Siguiendo este procedimiento, se genera una colección de conformaciones o trayectoria, que en el caso de ser lo suficientemente extensa como para ser representativa de la variedad conformacional de la molécula, puede emplearse para calcular las propiedades de la misma.

2.6. Magnitudes reducidas.

Por sencillez en su manejo, y también debido a algunas ventajas computacionales, es conveniente que en la simulación se trabaje con lo que se denominan “magnitudes reducidas” o adimensionales, las cuales se identifican añadiendo un asterisco como superíndice a la propiedad correspondiente; por ejemplo, al emplear magnitudes adimensionales, en lugar del vector de posición físico \mathbf{r} se emplearía su equivalente adimensional, \mathbf{r}^* .

Para transformar las unidades físicas en unidades reducidas es necesario definir un sistema de unidades de referencia adecuadamente elegidas que sirvan para hacer la transformación. En el caso de nuestro Grupo, las magnitudes de referencia que empleamos habitualmente son las siguientes:

- **La unidad de longitud** será la distancia de equilibrio característica entre dos esferas consecutivas del modelo empleado en cada caso, que denotaremos como b . Para un modelo basado en muelles gaussianos, la longitud cuadrática promedio de los muelles en equilibrio corresponde al valor b^2 . Así pues, por ejemplo, si \mathbf{r} representa el vector de posición, su equivalente reducido \mathbf{r}^* se calcularía como

$$\mathbf{r}^* = \frac{\mathbf{r}}{b}. \quad (2.37)$$

- **La unidad de energía** se establece como $k_B T$, donde k_B es la constante de Boltzmann y T es la temperatura absoluta del sistema. Todos las energías y potenciales reducidos se calcularán de forma análoga a

$$V^* = \frac{V}{k_B T}. \quad (2.38)$$

- **La unidad de fricción** equivale al coeficiente de fricción traslacional de la primera esfera del modelo en el fluido disolvente, ζ_1 , dado por la Ley de Stokes, $\zeta_1 = 6\pi\eta_0\sigma_1$, de manera que:

$$\zeta^* = \frac{\zeta}{\zeta_1} = \frac{\zeta}{6\pi\eta_0\sigma_1}, \quad (2.39)$$

donde η_0 es la viscosidad del disolvente en cuestión, y σ_1 el radio de la esfera de referencia.

Por estudios previos [41] se conoce que una correcta representación del efecto de la interacción hidrodinámica entre cadenas gaussianas tiene lugar cuando el tamaño hidrodinámico de la bola se asigna como $\sigma = 0.256b$, por lo que $\sigma^* = \sigma/b = 0.256$.

Estas tres magnitudes reducidas fundamentales nos permiten deducir las relaciones de normalización necesarias para convertir a la forma reducida cualquier otra magnitud relevante en nuestras simulaciones, por ejemplo:

- Cualquier **coeficiente de difusión** (o cualquier componente de un tensor de difusión) se expresará en unidades de *energía partido por fricción*:

$$D^* = \frac{D}{k_B T / \zeta_1} = D \frac{6\pi\eta_0\sigma_1}{k_B T}. \quad (2.40)$$

- El **tiempo** se convierte en adimensional a través del factor *longitud al cuadrado por fricción partido por energía*:

$$t^* = \frac{t}{b^2\zeta_1/k_B T} = t \frac{k_B T}{6\pi\eta_0\sigma_1 b^2}. \quad (2.41)$$

- Las **fuerzas** se expresarán en unidades de *energía partido por longitud*:

$$\mathbf{F}^* = \frac{\mathbf{F}}{k_B T / b} = \mathbf{F} \frac{b}{k_B T}. \quad (2.42)$$

Finalmente, para el caso de una simulación en unidades reducidas sería necesario adaptar también las ecuaciones empleadas. Esto puede hacerse sin mayores complicaciones siguiendo las indicaciones dadas para las diferentes magnitudes. Como ejemplo, la ecuación de Ermak-McCammon (ecuación 2.24) en unidades reducidas tomará la forma:

$$\mathbf{r}_i^* = \mathbf{r}_i^{*(0)} + \Delta t^* \sum_{j=1}^N \mathbf{D}_{ij}^{*(0)} \cdot \mathbf{F}_j^{*(int)(0)} + \Delta t^* \sum_{j=i}^N \left(\frac{\partial \mathbf{D}_{ij}^*}{\partial \mathbf{r}_j^*} \right)^0 + \mathbf{R}_i^* + (\Delta t^*) \mathbf{v}_i^{*(0)}. \quad (2.43)$$

CAPÍTULO 3

EQUIPO Y SOFTWARE

Durante mucho tiempo, el aumento de la potencia de los equipos informáticos ha estado ligado al desarrollo de procesadores con una arquitectura más eficiente y funcionasen a frecuencias de reloj cada vez mayores. Sin embargo, esta estrategia presenta una importante desventaja: el funcionamiento de los procesadores a unas velocidades tan altas implica el uso de voltajes cada vez más altos y la generación de una importante cantidad de calor. De esta forma, de no dotar al equipo de una capacidad de refrigeración adecuada, el núcleo del procesador podía llegar a fundirse, provocando una avería irreparable.

Alrededor del año 2004 se produce un cambio de estrategia, y se enfoca el desarrollo de nuevos procesadores hacia la integración de más núcleos de proceso en un solo procesador, desarrollándose los primeros procesadores de “doble núcleo”, para, a continuación y con el paso del tiempo, ir aumentando progresivamente el número de núcleos, así como su respectiva potencia de cálculo.

Con los equipos de un solo núcleo ya era posible realizar varios procesos simultáneos en un mismo equipo informático; sin embargo, esta “multitarea” se basaba en compartir el tiempo de procesador, ejecutando cada vez una sola tarea, pero cambiando con frecuencia de una tarea a otra. Por contra, en los equipos con más de un núcleo de proceso, la “multitarea” es real: pueden ejecutarse simultáneamente varias tareas, una por cada núcleo del que disponga el procesador.

De esta manera, si la idea clásica de una simulación computacional correspondía al esquema “un equipo = una simulación”, la proliferación en los últimos tiempos de equipos con múltiples núcleos de proceso y los *clusters* de cálculo de altas prestaciones, o equipos de computación distribuida, hacen que este esquema quede obsoleto. Actualmente es factible

tanto llevar a cabo varias simulaciones al mismo tiempo en un solo equipo, como el llevar a cabo una sola simulación en la que trabajen, al mismo tiempo, varios núcleos de cálculo.

Para disponer de esta capacidad, por supuesto que en primer lugar es necesario disponer de equipos con varios núcleos de proceso, además de que es deseable el poder interconectarlos para gestionarlos de forma centralizada y automatizar algunos procesos. También es necesario disponer del software adecuado: tanto para la preparación y el procesado en paralelo de los trabajos, como para gestionar la cooperación de varios núcleos en un mismo proceso. Y por supuesto, también es necesario tener en cuenta esta forma de trabajo a la hora de diseñar y desarrollar el software de simulación.

Por este motivo, y puesto que una parte importante del trabajo de este Grupo de Investigación se centra en la elaboración de software de simulación, se han implementado en nuestro software las modificaciones necesarias para hacer funcionar los programas de simulación siguiendo algunos de estos esquemas, mientras que también se han elaborado algunos programas auxiliares destinados a facilitar las tareas de preparación de simulaciones de este tipo.

En las secciones siguientes se describen tanto los equipos de computación distribuida que se han empleado durante esta Tesis, como el software (tanto de gestión de los equipos como de simulación) y las adaptaciones realizadas encaminadas al uso del mismo en entornos de computación en paralelo.

3.1. Equipos informáticos de alto rendimiento.

Durante la realización de esta Tesis se han empleado principalmente dos *clusters* de cálculo: el que es propiedad del Grupo de Investigación, y el supercomputador “Ben Arabi”, del Centro de Supercomputación del Parque Científico de Murcia.

3.1.1. Cluster del grupo.

Se trata de un pequeño *cluster* de cálculo, compuesto por nueve servidores HP Proliant DL160 G5 instalados en un armario *rack*. Cada uno de ellos está dotado de dos procesadores de cuatro núcleos que funcionan a una frecuencia de 2.66 gigahercios, modelo Intel Xeon E5430. Cada uno de los equipos está a su vez dotado con 9 gigabytes de memoria RAM y un disco duro de 160 gigabytes para el sistema operativo y almacenamiento local. Además, uno de los equipos (el *maestro*) está dotado de dos discos duros adicionales de 750 gigabytes de capacidad, que se destinarán al almacenamiento de datos ya procesados. Todos los equipos están interconectados entre sí mediante una red Ethernet exclusiva de 1 gigabaudios.



FIGURA 3.1: *Cluster* de cálculo del Grupo de Investigación.

Debido a que el equipo está exclusivamente orientado a la producción (esto es, a la realización de simulaciones), se busca que los equipos ofrezcan la máxima potencia de cálculo, a la vez que el entorno sea lo más estable posible para evitar problemas a la hora de lanzar cálculos de duración prolongada. Por estos motivos se optó instalar en el *cluster* un sistema operativo GNU/Linux, concretamente, la distribución conocida como **Scientific Linux** versión 5. Este sistema operativo está basada en el código fuente (abierto) de la distribución comercial **Red Hat Enterprise Linux** versión 5, siendo compilado y mantenida en conjunto por personal del CERN, el Fermilab y otras entidades académicas. La distribución **Red Hat Enterprise Linux**, tiene una larga trayectoria y un uso ampliamente extendido en entornos de producción, en los que la fiabilidad y la estabilidad de los sistemas cobra una gran importancia, lo que se traduce en un código fuente con unas funcionalidades y seguridad muy cuidadas, que hereda la distribución **Scientific Linux**, además de paquetes de software exclusivos útiles en entornos científicos.

En cuanto a la gestión de los trabajos en *cluster*, la gestión manual de los mismos puede resultar bastante tediosa, por lo que se emplea un software de gestión de colas que distribuya los trabajos entre los nodos de trabajo, enviando trabajos nuevos a medida que los anteriores finalicen. También para esta tarea se opta por un proyecto de código abierto, en este caso el **Sun Grid Engine**, versión 6.1u5, desarrollado y distribuido por **Sun Microsystems**. Se trata de un gestor de colas, que, aunque sigue en desarrollo, ofrece

ya una funcionalidad bastante completa y muy configurable, por lo que sus posibilidades sobrepasan de largo las necesidades que planteaban los trabajos recogidos en esta Tesis; sin embargo, la flexibilidad que ofrece este software puede volverse interesante para futuros proyectos que pudieran presentar unas necesidades más complejas.

Por las características de este software de gestión se estructura el *cluster* de forma que el equipo dotado de los discos duros adicionales sea el equipo “maestro” al que accedan los usuarios, y desde el que se realice la preparación y el envío de los “lotes” de trabajo así como la gestión de las colas, mientras que los ocho equipos restantes, con un total de 64 procesadores, serán los nodos “esclavos” en los que se ejecutarán los trabajos.

3.1.2. Supercomputador Ben Arabi.

Durante la realización de esta Tesis comenzó entró en funcionamiento el Centro de Supercomputación de la Fundación Parque Científico de Murcia (<http://sc.fpcmur.es>), y gracias a los acuerdos entre éste y la Universidad de Murcia, nuestro grupo de investigación tuvo acceso al supercomputador tanto en su fase de pruebas, como una vez que éste entró en pleno funcionamiento.

El centro de Supercomputación de Murcia pone a la disposición de sus usuarios el supercomputador Ben Arabi que funciona bajo sistema operativo SUSE Linux Enterprise, aunque soporta cualquier sistema Microsoft Windows y variantes de Unix mediante virtualización. El supercomputador está compuesto por dos sistemas con arquitecturas diferentes e independientes, llamadas “Ben” y “Arabi”:

En primer lugar, el nodo central “Ben”, consiste en un sistema de memoria compartida HP Integrity Superdome SX2000 con 128 núcleos del procesador Intel Itanium-2 Dual-Core Montvale a 1.6 gigahercios. Está dotado con 1.5 terabytes de memoria DDR2, común a todos los núcleos, y 5.84 terabytes en discos SAS para *scratch*. El rendimiento real conseguido, según el *benchmark* “Linpack” de 737 gigaflops.

En la arquitectura de memoria compartida todos los procesadores del sistema pueden acceder directamente a todas las posiciones de memoria. El espacio de memoria física es único



FIGURA 3.2: Sistema HP Integrity Superdome SX2000 (nodo “Ben”).

y global. La utilización de este espacio de memoria común evita la duplicación de datos y el lento trasvase de información entre los procesos.

El modelo de programación de memoria compartida es una extensión natural del modelo convencional (secuencial). La comunicación entre tareas/procesos es sencilla, por medio de variables compartidas. Establecer una distribución de datos adecuada afecta a la eficiencia de ejecución, pero no a su corrección.

La memoria principal es compartida por todos los nodos siguiendo una arquitectura ccNUMA. El sistema está organizado en células o nodos, las cuales están formadas por un procesador y una memoria. Los procesadores pueden acceder a la memoria de las demás células a través de la red de interconexión. La coherencia de caché es mantenida por el hardware, por lo que puede ser programado como una arquitectura de memoria compartida. El acceso a la memoria local (en la misma célula) es más rápido que el acceso remoto (en una célula diferente).

En cuanto al sistema “Arabi”, se trata de un *cluster* computacional de nodos finos en formato *Blade* de alta eficiencia energética. Formado por 102 nodos de cálculo, ofrece un total de 816 núcleos del procesador Intel Xeon Quad-Core L5450, a 3.0 gigahercios. Distribuido entre todos los nodos, ofrece un total de 1072 gigabytes de memoria FBD DDR2 y 11.52 terabytes en discos duros SATA para *scratch*. Este sistema obtiene un rendimiento pico máximo de 9.792 teraflops en el *benchmark* “Linpack”.

El “BladeSystem” de HP es una infraestructura adaptable en un chasis. Consolida los elementos esenciales de un centro de datos moderno (alimentación, refrigeración, conectividad, redundancia, seguridad, computación y almacenamiento) dentro de una plataforma modular que puede acelerar la integración y optimización.

En este sistema, el espacio de memoria es distribuido, cada procesador tiene su espacio de direcciones exclusivo. Es preciso el reparto de datos de entrada a cada procesador a través de la red que interconecta los *Blades*. Existe un nodo que funciona como “maestro”, gestionando la comunicación y el envío los datos a los nodos “esclavos” y no toma parte directamente en el cálculo. La necesidad de sincronización y la gestión de las comu-



FIGURA 3.3: Cluster de Blades HP (nodo “Arabi”).

nicaciones dificulta la programación porque la paralelización se realiza de forma manual mediante la inclusión de funciones y código específico en los programas.

3.2. Esquemas de trabajo en entornos de cálculo en paralelo.

La modificación del software preparado para correr en equipos de un solo núcleo para su adaptación a entornos de múltiples núcleos no es una cuestión trivial. Para ello, primero es necesario tener una idea de los diferentes niveles de trabajo en paralelo que pueden encontrarse. Brevemente, podrían resumirse en lo siguiente:

- **Funcionamiento independiente:** básicamente, se trata de lanzar varios trabajos al mismo tiempo en una máquina con varios núcleos de cálculo, de forma que aunque estos procesos se ejecutan en un mismo equipo compartiendo los recursos del mismo, no interaccionan entre sí de ninguna manera. Cada proceso lanzado está limitado a los recursos de memoria RAM, disco y el núcleo de proceso que le han sido asignados en exclusiva por el sistema operativo.
- **Integración a nivel de máquina:** en este caso, un proceso corre en un solo equipo con varios núcleos, y puede disponer de varios de ellos, o todos a la vez, de forma que éstos “colaboren” y se repartan el trabajo, pero no existe interacción con otras máquinas.
- **Integración total:** en este nivel de integración, el más alto de los tres, todos los núcleos de cálculo del *cluster*, que se hallan distribuidos en diferentes equipos físicos, están de una u otra manera implicados en el desarrollo de la misma tarea.

Por supuesto, al aumentar el grado de integración también aumenta la complejidad de los programas empleados.

En el nivel de operación independiente, los programas requieren pocas o ninguna modificación respecto a los que funcionarían en una máquina de un único procesador; al fin y al cabo, se trata de lanzar múltiples instancias de cada programa (tantas como núcleos de cálculo haya, si se desea un aprovechamiento óptimo del equipo), y que éstas

funcionen de forma independiente unas de otras. La única precaución necesaria es la de asegurarse de que la ejecución de uno de estos procesos no interfiera con otros, por ejemplo, sobrescribiendo archivos temporales de trabajo o archivos de salida de datos.

Para el caso que se ha denominado como “integración a nivel de máquina” el funcionamiento de los programas ya es más complejo: en principio, existe un solo programa en ejecución; sin embargo, en puntos específicos del funcionamiento de éste, se crean múltiples hilos de trabajo, que se dividen entre los núcleos de cálculo disponibles. De esta manera, el trabajo a realizar (deshablamiento de bucles, cálculos en paralelo, etc), se reparte entre varios hilos, de forma que el trabajo se reparte y se lleva a cabo al mismo tiempo en los diferentes núcleos, con la correspondiente ventaja en cuanto a tiempo de ejecución.

Para conseguir este reparto del trabajo es necesario realizar modificaciones en el código fuente y emplear directivas específicas para estos propósitos. Esto puede conseguirse de dos maneras diferentes, según el grado de especificidad del código que se desea paralelizar. Para el caso del código más específico no queda más opción que la de diseñar e implementar el código de paralelización mediante, por ejemplo, la interfaz OpenMP a la misma vez que se desarrolla el código del programa. Por el contrario, para el caso de muchas funciones de carácter general, como por ejemplo, aquellas que realizan operaciones matemáticas complejas, existen librerías de funciones cuyo código ya ha sido preparado y optimizado por sus desarrolladores para correr en paralelo. De esta manera, se simplifica bastante implementar el trabajo en paralelo, pues sólo es necesario llamar a las funciones de estas librerías y establecer algunos parámetros para obtener un código paralelizado.

Durante el desarrollo de los programas específicos de los que hace uso esta Tesis se han combinado ambas opciones: la paralelización “manual”, por lo delicado de la tarea, se ha reservado para unos pocos casos específicos (por ejemplo, funciones de correlación), mientras que el uso de funciones de librerías matemáticas, en concreto, la librería **Intel Math Kernel Library (MKL)** versión 10.3, se ha usado de forma más extensa, en concreto, en la generación de números aleatorios y en el trabajo con matrices y vectores de gran tamaño.

Por último, como se puede esperar, el nivel más alto de integración es el que requiere un mayor esfuerzo para adaptar, no sólo los programas, sino también el entorno de funcionamiento de los mismos, cosa que no se requiere en ninguno de los otros dos niveles. Los programas que se emplean en este nivel deben estar especialmente preparados, no sólo para repartir el trabajo entre varios hilos en un mismo equipo, sino también entre diferentes equipos.

En cada equipo del *cluster* se ejecutará una instancia propia del programa, y éstas no sólo deben ser capaces de comunicarse con otras instancias que se ejecutan en el resto de los equipos implicados en la tarea, sino que, además, las comunicaciones de datos entre los equipos deben estar correctamente diseñadas y sincronizadas de manera que la comunicación no suponga un cuello de botella que retrase la ejecución de toda la tarea en conjunto.

Para realizar esta gestión, típicamente, una de las instancias en ejecución suele actuar coordinando el trabajo de todas las demás. Para ello, el código del programa, y especialmente debe tener implementadas las funciones específicas de coordinación y comunicación entre equipos. Aunque existen librerías con tales funciones (librerías MPI), es necesario optimizar al detalle el funcionamiento de cada programa y cada *cluster* específico para identificar los posibles cuellos de botella e implementar las soluciones adecuadas. En definitiva, este nivel de integración supone un esfuerzo muy importante para adaptar el software de simulación al entorno.

Debido a la complejidad de este tipo de adaptaciones, aunque el entorno del Ben Arabi está preparado para ello, no se ha llegado a emplear este tipo de paralelización durante el desarrollo del software empleado en esta Tesis.

3.3. Adaptación del software de simulación a entornos de trabajo en paralelo.

Una de las actividades más destacadas de nuestro grupo de investigación es la producción de software de simulación y estimación de diversas propiedades de las macromoléculas que luego es puesto a disposición del público de forma gratuita a través de la página web del grupo, <http://leonardo.inf.um.es/macromol>.

En lo que se refiere al contenido de esta Tesis, de entre todo el conjunto de software producido por el grupo en los últimos años, tienen especial importancia los paquetes de simulación SIMUFLEX [42] y MONTEHYDRO [43]. Ambos programas se diseñaron inicialmente para funcionar sobre equipos de escritorio tradicionales con un sólo procesador, y aunque actualmente está prevista su adaptación a entornos de múltiples procesadores (en lo que se

ha llamado “integración a nivel de máquina” en la sección anterior), durante la realización del trabajo correspondiente a esta Tesis doctoral dicha versión no estaba aún disponible.

Por este motivo, y dada la disponibilidad de equipos de computación de alto rendimiento, como el *cluster* del Grupo o el supercomputador Ben Arabi, para mejorar el rendimiento de estos programas se realizaron las modificaciones necesarias para poder lanzar simultáneamente varias instancias de estos programas en funcionamiento independiente, de forma que se pudieran procesar varias simulaciones al mismo tiempo.

En el caso de las simulaciones que se contemplan en esta Tesis, esta estrategia supone una importante ventaja, ya que, desde el punto de vista del cálculo de propiedades en promedio a partir de simulaciones en equilibrio, son equivalentes el realizar una única simulación muy larga, o varias más cortas que sean estadísticamente independientes y en conjunto, sumen el mismo número de pasos que la primera, más larga.

Así pues, un trabajo de simulación relativamente largo, que, con las versiones actuales de los programas sólo podría ejecutarse un sólo núcleo de cálculo, puede dividirse en una serie de subsimulaciones independientes que pueden distribuirse entre la totalidad de núcleos de proceso disponibles. De esta forma, dado que podemos realizar varias simulaciones más cortas simultáneamente, la duración del trabajo total se acorta de manera prácticamente proporcional al número de subsimulaciones en las que se divide la original.

Es necesario aclarar que, a efectos prácticos, esto no es completamente cierto, puesto que al realizar varias simulaciones simultáneas, tanto el uso de memoria volátil como el volumen de datos que han de ser volcados a disco aumenta, y esto puede ralentizar el proceso de simulación, aunque en cualquier caso, el proceso seguirá siendo bastante más rápido que el realizar una única simulación larga.

Además, también hay que tener en cuenta que, debido al tipo de trabajo que suponen las simulaciones de esta Tesis, este esquema de funcionamiento también es más ventajoso que la integración de varios núcleos de cálculo en el mismo proceso de simulación.

Esta ventaja se debe principalmente a que los procesos de simulación son secuenciales, y cada nuevo paso generado depende del resultado del anterior, por lo que no pueden calcularse simultáneamente varios pasos de simulación en distintos núcleos.

De esta manera, el principal factor de aceleración mediante el uso de múltiples procesadores queda limitado a la cooperación de varios núcleos en la solución de operaciones complejas, como puede ser, por ejemplo, la factorización de grandes matrices, mientras

que en las operaciones sencillas no se obtendrá ningún beneficio, pudiendo incluso incurrir en penalizaciones, pues la creación de nuevos hilos de cálculo y la recolección de sus resultados y el cierre de los hilos también conlleva un breve tiempo de proceso (este es uno de los motivos por los que la ventaja obtenida por emplear un mayor número de núcleos de cálculo para una simulación no es directamente proporcional al número de núcleos de cálculo adicionales empleados).

Por este motivo, el uso de más de un procesador en este tipo de simulaciones, especialmente en las de tipo Monte Carlo, en las que no se manejan tensores de difusión, como ocurre en la dinámica browniana, no supone una gran ventaja y puede incluso suponer un inconveniente.

Sin embargo, si se trabaja con varias instancias independientes del programa de simulación, se estarán ejecutando simultáneamente varios pasos de simulación (cada uno perteneciente a una simulación independiente de las demás), obteniendo así una ganancia de rendimiento mayor que con la cooperación de varios procesadores (esta afirmación también tiene sus matices, recordemos que el mayor uso de RAM y el acceso a disco también pueden degradar el rendimiento).

En cambio, desde el punto de vista de la preparación de este tipo de trabajos, el trabajar con varias simulaciones independientes tiene un inconveniente, que consiste en la necesidad de preparar los archivos necesarios para todo el lote de simulaciones, así como el posterior tratamiento de los datos obtenidos de cada una de las simulaciones. Por este motivo, se ha preparado una serie de programas auxiliares destinados a facilitar estas tareas.

3.4. El paquete SIMUFLEX.

El paquete SIMUFLEX [42] está ideado para la simulación y cálculo de propiedades en disolución de macromoléculas flexibles, representadas por modelos de bolas y muelles de topología arbitraria. El paquete está compuesto por los programas BROWFLEX y ANAFLEX. El funcionamiento de ambos programas es altamente adaptable mediante archivos de configuración para adaptar la simulación y el cálculo de propiedades a las necesidades específicas del usuario.

BROWFLEX es capaz de trabajar con modelos de conectividad arbitraria enlazados mediante muelles de tipos variados, entre los que destacan los reseñados en la sección 2.2, y realizar simulaciones mediante la técnica de la dinámica browniana, en las que puede o no tenerse en cuenta la interacción hidrodinámica fluctuante, a elección del usuario, generando los correspondientes archivos de trayectorias. El programa puede emplearse para realizar simulaciones tanto en equilibrio como aquellas en las que existe una variación temporal de las propiedades.

ANAFLEX, en cambio, tiene por finalidad análisis las trayectorias generadas, y el cálculo a partir de ellas de múltiples propiedades en disolución, incluyendo propiedades promediadas sobre la trayectoria y funciones de correlación con el tiempo.

3.5. El programa MONTEHYDRO.

Si el paquete SIMUFLEX permite realizar simulaciones de macromoléculas flexibles mediante dinámica browniana, MONTEHYDRO [43] permite llevar a cabo una tarea análoga empleando la técnica de Monte Carlo con muestreo por importancia de tipo Metropolis, descrito también en el Capítulo anterior, en la sección 2.5.

En este caso, y debido a la propia limitación de la técnica de Monte Carlo, no es posible obtener funciones dependientes del tiempo, y sólo pueden calcularse propiedades en equilibrio. Para este fin se aplica el conocido como tratamiento de cuerpo rígido, propuesto por Zimm [44], que consiste en calcular las propiedades de cada una de las conformaciones generadas como si éstas fueran moléculas rígidas, para a continuación promediar las propiedades calculadas sobre todo el espacio conformacional analizado.

Debido a que en estas simulaciones no se tiene en cuenta la interacción hidrodinámica para la generación de conformaciones, sino que ésta se incluye a posteriori, las propiedades de tipo hidrodinámico calculadas mediante esta técnica sólo pueden considerarse buenas aproximaciones a los valores reales.

3.6. Programas auxiliares: MULTISIMUFLEX.

Para que la división de una simulación en múltiples subsimulaciones independientes sea eficiente, se requiere que estas simulaciones sean completamente independientes entre sí. Para que esta condición se cumpla en las simulaciones realizadas con **SIMUFLEX** y **MONTEHYDRO** existen dos alternativas: el uso de diferentes conformaciones de partida para cada una de las subsimulaciones, o el uso de una semilla de números pseudo-aleatorios diferente (el uso de números aleatorios en programas informáticos nunca es completamente aleatorio, pues está regido por una variedad de algoritmos, de forma que los números aleatorios siempre están, de alguna manera, relacionados con un número “semilla” que se usa como base para la generación).

Ambos casos suponen la creación de un número de archivos de entrada apropiados al número de subsimulaciones que se desee realizar, además de que tras las simulaciones es necesario reunir el mismo número de archivos de trayectorias, analizarlos y calcular los promedios de las propiedades con sus correspondientes incertidumbres. Con este objetivo, se preparó el paquete de programas auxiliares **MULTISIMUFLEX**, compuesto por los siguientes elementos:

- **MULTIMONTESEED** y **MULTIBROWSEED**: Estos programas generan, a partir de un archivo de entrada genérico (similar en cada caso a los de **MONTEHYDRO** o **SIMUFLEX**), una serie de archivos de entrada identificados por un número de serie de tres cifras (el máximo de subsimulaciones está limitado, por lo tanto, a 1000) que se diferencian en la semilla del generador de aleatorios incluida en cada uno. Además de los archivos de entrada, estos programas generan un listado con los nombres de los archivos de entrada, de forma que puedan ser fácilmente manejados e incluidos en el sistema de colas de los *clusters* de cálculo.
- **MULTIMONTEINIT** y **MULTIBROWINIT**: De forma análoga a los dos programas anteriores, la finalidad de estos dos programas es la generación de los múltiples archivos de entrada de las subsimulaciones y un archivo con el listado de los mismos, aunque en este caso la diferencia entre los archivos estriba en que si bien las semillas de los números aleatorios son las mismas, las conformaciones iniciales proporcionadas son diferentes. Dado que la generación de conformaciones iniciales es un proceso com-

plejo, debe ser el usuario el que genere estas conformaciones iniciales, que después se incluirán en los archivos de entrada.

- **MULTIBATCH**: este pequeño programa consiste en un *script* para el *shell bash*, que facilita el lanzamiento en bloque de las simulaciones a la cola del gestor de trabajos del *cluster*. Debido a que se trata de una tarea muy específica, el script debe ser también específico para el gestor de tareas que se encuentre instalado en el *cluster* en el que se esté trabajando.
- **MULTITRACOLLECT**. En el caso de trabajar con **SIMUFLEX**, como ya se ha comentado, es necesario el uso de **ANAFLEX** para el análisis de las trayectorias calculadas. Este proceso es bastante más rápido que el de generación de las trayectorias, y además, **ANAFLEX** viene dotado de la posibilidad de trabajar en serie con lotes de archivos de trayectorias. Por estos motivos, aunque también es posible lanzar en paralelo varios análisis (teniendo que cambiar en los archivos de entrada únicamente el nombre del archivo de trayectoria a procesar), no se ha considerado necesario producir una herramienta para este propósito. Sin embargo, para el modo de trabajo en lote de **ANAFLEX** sí que resulta conveniente disponer de una herramienta como **MULTITRACOLLECT**, capaz de generar un archivo de trabajo en lote para **ANAFLEX** que especifique los archivos de trayectorias a analizar y los parámetros con los que deben ser analizados.
- **MULTIMONTERES** y **MULTIANARES**: Por último, estos dos programas son los encargados de realizar el cálculo de los promedios y las incertidumbres de las propiedades que se han obtenido, respectivamente, bien a través de la ejecución de **MONTEHYDRO** o bien ejecutando sucesivamente **SIMUFLEX**, y a continuación **ANAFLEX**. En principio, es una tarea sencilla, sin embargo, resulta útil disponer de programas que sean capaces de procesar rápidamente y con la mínima intervención del usuario la serie de archivos de resultados generados, especialmente, debido al formato específico de cada uno de los dos tipos de archivos de resultados.

Está previsto que al igual que los paquetes de simulación **MONTEHYDRO** y **SIMUFLEX**, este paquete de herramientas auxiliares y su documentación sean puestos a disposición del público en la página web del Grupo, <http://leonardo.inf.um.es/macromol>.

CAPÍTULO 4

COMPARACIÓN DE METODOLOGÍAS PARA EL CÁLCULO DEL COEFICIENTE DE DIFUSIÓN TRASLACIONAL

4.1. Introducción.

Como se ha explicado con anterioridad, la difusión de cadenas de polímero en disolución es esencial para comprender muchos procesos relacionados con la dinámica de los sistemas macromoleculares. Debido a ello, desde hace tiempo esta cuestión ha sido objeto de un importante foco de atención [45] en lo que se refiere al desarrollo teórico de expresiones y algoritmos computacionales que permitan el cálculo teórico del coeficiente de difusión de macromoléculas individuales.

A lo largo del tiempo, se han propuesto diferentes expresiones para el cálculo de esta propiedad hidrodinámica. Por ejemplo, en el mencionado Capítulo 2 se definía de forma rigurosa el coeficiente de difusión y se mostraba una expresión matemática para su cálculo basada en el desplazamiento cuadrático medio del centro de masas a lo largo del tiempo. Sin embargo, existen otras expresiones que permiten calcular el coeficiente de difusión a partir de la estructura de modelos computacionales, de forma que se evita tener que recurrir al cálculo de la función de correlación temporal del desplazamiento del centro de masas.

Como puede suponerse, la existencia de estas diferentes expresiones se debe al uso de diferentes consideraciones e hipótesis, que suponen que las expresiones no sean completamente equivalentes. A pesar de que las diferencias entre ellas normalmente son pequeñas, y de poca relevancia en comparación con los errores experimentales, es importante tener en cuenta estas diferencias y conocer las circunstancias bajo las que pueden volverse relevantes.

En las siguientes secciones se presentan algunas de estas expresiones alternativas a la presentada en el Capítulo 2, y se intenta realizar un estudio sistemático de la influencia de la flexibilidad intrínseca de la cadena sobre los valores que se obtienen empleando las diferentes expresiones.

4.2. Base teórica.

4.2.1. Coeficiente de difusión de Einstein.

Como se recordará de la Sección 2.1.2, el coeficiente de difusión traslacional global de un grupo de partículas, como los que representa un modelo de bolas y muelles, puede calcularse de forma sencilla a partir de una simulación de dinámica browniana con interacción dinámica mediante la relación de Einstein-Schmoluchowsky, enunciada mediante la ecuación 2.3, que por conveniencia en este capítulo expresaremos como

$$\langle(\Delta\mathbf{r}_{CM})^2\rangle = 6D_E t, \quad (4.1)$$

donde $\Delta\mathbf{r}_{CM}$ es el desplazamiento del centro de masas respecto de su posición anterior a lo largo de un intervalo de tiempo t . D_E es el coeficiente de difusión traslacional, denotado con el subíndice E por haber sido calculado mediante la expresión de Einstein.

Para el cálculo de esta propiedad a lo largo de un intervalo de tiempo prolongado, como la duración de una trayectoria de simulación como las que se emplean a lo largo de esta Tesis, es necesario discretizar la ecuación anterior y el desplazamiento cuadrático promedio se calcula sobre todos los posibles intervalos de tiempo de duración τ , lo que

equivale a la función de autocorrelación de los desplazamientos del centro de masas [25] de manera que el cálculo que se implementa en los programas de simulación realmente corresponde a

$$\langle [\mathbf{r}_{CM}(t) - \mathbf{r}_{CM}(t + \tau)]^2 \rangle = 6D_E\tau. \quad (4.2)$$

4.2.2. Coeficiente de difusión de Kirkwood.

El coeficiente de difusión de una cadena formada por N elementos esféricos o bolas también se puede calcular a partir de la teoría de fenómenos de transporte irreversibles en disoluciones macromoleculares desarrollada por Kirkwood y Riseman [46, 47]. Para aplicar dicha teoría, se considera que las bolas son elementos de fricción a los que se aplica individualmente la ley de Stokes, mientras que la interacción hidrodinámica entre ellas se representa mediante el tensor de Oseen (ecuaciones 2.28 y 2.29). En versiones posteriores de la teoría, dicho tensor se reemplaza por el tensor de Rotne-Prager-Yamakawa (ecuaciones 2.28, 2.30 y 2.31).

Un aspecto importante de la teoría de Kirkwood y Riseman es que permite calcular las propiedades hidrodinámicas de forma aproximada, no rigurosa, pero sí sencilla, mediante expresiones simples que relacionan términos que dependen de las distancias entre los N elementos del modelo. Así, utilizando la denominada fórmula de Kirkwood aproximada [46], el coeficiente de difusión traslacional de Kirkwood, que denotaremos de forma distintiva como D_K , se puede calcular a partir de la expresión:

$$D_K = \frac{1}{3N^2} \sum_i \sum_j Tr \langle \mathbf{D}_{ij} \rangle, \quad (4.3)$$

donde $Tr \langle \mathbf{D}_{ij} \rangle$ representa la traza de cada bloque de tamaño 3×3 que relaciona las bolas i y j en el tensor de difusión \mathbf{D} , que se vuelve dependiente de la conformación al tener en cuenta la interacción hidrodinámica en la simulación. Los corchetes angulares, $\langle \dots \rangle$, indican que, si la macromolécula es flexible, dicho tensor proviene de realizar un promedio conformacional (en el caso del trabajo original de Kirkwood y Riseman, se refiere a un pre-promediado). Independientemente de que la interacción hidrodinámica sea descrita por el tensor de Oseen o el de Rotne-Prager-Yamakawa (que es el que se utiliza en esta

Tesis), la expresión para el cálculo de D_K se reduce a la siguiente [15], mucho más práctica desde el punto de vista computacional:

$$D_K = \frac{k_B T}{6\pi\eta_0} \left(\frac{1}{N\sigma} + \frac{2}{N^2} \sum_{i>j} \sum \left\langle \frac{1}{r_{ij}} \right\rangle \right), \quad (4.4)$$

donde $k_B T$ es el factor de Boltzmann, σ es el radio de Stokes de cada elemento esférico (entendiendo que todos son del mismo tamaño), y r_{ij} es la distancia entre los elementos i y j de la cadena. Si la cadena es flexible, $1/r_{ij}$ será también un promedio conformacional.

Es conocido [44, 48] que el uso de la aproximación de Kirkwood mediante el doble sumatorio de la ecuación 4.4) produce resultados de D_K que suponen una cota superior del coeficiente de difusión. Esto es debido a que el coeficiente de difusión de Kirkwood describe el comportamiento traslacional en una escala de tiempo menor que el tiempo de relajación global de la cadena, τ_Z (tiempo de relajación de Zimm), es decir, cuando $t \ll \tau_Z$ [49].

4.2.3. Tratamiento del Cuerpo Rígido.

Actualmente, mediante el uso de ordenadores, es posible realizar un tratamiento riguroso de las ecuaciones propuestas por Kirkwood y Riseman, de forma que se evite el uso de pre-promediados. Así, cuando se trabaja con macromoléculas flexibles, la aplicación rigurosa de la teoría de Kirkwood y Riseman sirve como base para el procedimiento conocido como tratamiento del “Cuerpo Rígido” (o “*Rigid Body*” en inglés). A lo largo de este Capítulo, a los coeficientes de difusión obtenidos mediante este procedimiento los denominaremos D_{RB} .

Dicho tratamiento consiste en la generación de una muestra representativa de conformaciones en equilibrio de la macromolécula flexible (por ejemplo, mediante una simulación de Monte Carlo [44]) y en la obtención de las propiedades hidrodinámicas de cada conformación individual, considerando éstas como instantáneamente rígidas, para, finalmente, calcular el promedio de las propiedades obtenidas a lo largo de toda la trayectoria. Como se ha podido comprobar, los valores de D_{RB} obtenidos de esta manera constituyen una cota inferior del coeficiente de difusión traslacional [48] (ver también la tabla I de [25]).

Por lo general, la solución rigurosa de las ecuaciones de la teoría de Kirkwood y Riseman requiere la inversión del tensor de difusión que se construye para cada conformación

a partir del tensor de interacción hidrodinámica correspondiente (ver ecuación 2.27 del Capítulo 2). Por este motivo, es necesario emplear un tensor de interacción hidrodinámica que sea siempre definido positivo, como es el caso del tensor de Rotne-Prager-Yamakawa [15], también presentado en el Capítulo 2.

En los trabajos realizados en este Capítulo, el algoritmo empleado para la realización del tratamiento del Cuerpo Rígido es el mismo que se encuentra implementado en el conjunto de programas HYDRO, desarrollado por nuestro grupo de investigación y puesto a disposición pública [50] en la página web de nuestro grupo.

4.2.4. Coeficiente de difusión a tiempo largo.

Como se mencionaba con anterioridad, D_K supone una cota superior para del coeficiente de difusión, debido a que la escala de tiempo contemplada era tal que $t \ll \tau_Z$.

Para escalas de tiempo mayores ($t \gg \tau_Z$), el movimiento difusivo está caracterizado por un coeficiente de difusión que denominaremos D_L (por “difusión a tiempo largo”) que, en general, es diferente a D_K debido a que ya ha tenido lugar la relajación de la dinámica interna de la macromolécula. La diferencia entre D_K y D_L suele ser pequeña, aunque se manifiesta de forma algo más evidente cuando las cadenas son largas y flexibles [49]. Fixman [51] encontró que $D_L < D_K$, pudiendo escribirse la siguiente relación

$$D_L = D_K - D_1, \quad (4.5)$$

donde D_1 es una cantidad positiva que se calcula a partir de la función de autocorrelación del vector velocidad $\mathbf{A}(t)$,

$$D_1 = \frac{1}{3N^2} \int_0^\infty dt \langle \mathbf{A}(0) \cdot \mathbf{A}(t) \rangle, \quad (4.6)$$

siendo $\mathbf{A}(t)$ un vector tridimensional dependiente del tiempo, tal que para cada instante t ,

$$\mathbf{A} = \frac{1}{k_B T} \sum_i \sum_j \mathbf{D}_{ij} \cdot \mathbf{F}_i, \quad (4.7)$$

donde \mathbf{F}_i es la fuerza que actúa sobre la bola i y \mathbf{D}_{ij} son los bloques 3×3 del tensor de difusión.

Liu y Dünweg [49] han estudiado en profundidad el significado de D_K y D_L , tanto a nivel teórico como mediante simulaciones de dinámica browniana, alcanzando en estas últimas un grado de precisión suficiente como para reflejar las diferencias entre D_K y D_L , por lo que sus resultados se han empleado en este trabajo como referencia para el afinado de nuestras metodologías de cálculo.

4.3. Modelo y metodología.

4.3.1. Modelo de bolas y muelles.

Hemos utilizado esencialmente dos tipos de modelos de cadena para dilucidar la influencia de la flexibilidad y en las diferentes formas de calcular el coeficiente de difusión. Se han construido varias cadenas de cada uno de estos modelos, de forma que las simulaciones pongan de manifiesto la influencia del tamaño de la macromolécula en los coeficientes de difusión calculados.

Además, y debido a que, por una parte, el objetivo de este Capítulo no es la comparación de los resultados de las simulaciones con los obtenidos a partir de ningún polímero real, y que, por otra parte, los únicos datos bibliográficos de referencia de los que disponemos se refieren a magnitudes reducidas (ver la sección 2.6), también se emplearán éstas en las simulaciones.

De esta manera, los modelos empleados en este trabajo quedan definidos de la siguiente manera:

1. Modelos de tipo 1: modelos flexibles con muelles gaussianos y volumen excluido “suave” entre bolas no enlazadas, igual que el que se emplea en el trabajo de Liu y Dünweg [49]. Las expresiones matemáticas de estos elementos son las que corresponden, respectivamente, a las ecuaciones 2.13 y 2.19 del Capítulo 2.

Respecto al uso de estas expresiones es necesario mencionar dos salvedades: la primera se refiere al uso de magnitudes reducidas, que se traduce únicamente en la sustitución en las ecuación de cada una de las magnitudes por su contrapartida reducida. En segundo lugar, en los trabajos de este Capítulo se emplean sobre todo simulaciones de dinámica browniana, en la que se manejan fuerzas internas que actúan sobre cada uno de los elementos del modelo, por lo que será necesario derivar las dos expresiones anteriores para obtener la ecuación de la fuerza correspondiente.

Teniendo en cuenta lo anterior, los valores de los parámetros de este tipo de modelos consisten en una constante para los muelles conectores gaussianos $H^* = 3$, y unos parámetros del potencial de volumen excluido “suave” $A^* = 75$ y $\beta^* = 4$; además, para este potencial, se considera un radio de corte o “*cutoff*”, tal que para ahorrar tiempo de cálculo, se omitirá (por considerarla despreciable) la fuerza entre elementos separados una distancia mayor que $r_{cut}^* = 0.512$. El valor de todos estos parámetros coincide con los empleados en [49].

2. Modelos de tipo 2: modelos semirrígidos, con muelles de tipo hookeano duro (Fraenkel), definidos por la ecuación 2.14, entre bolas adyacentes y muelles angulares (ecuación 2.17) definidos por cada terna de bolas unida por dos muelles consecutivos.

La mayor o menor flexibilidad de las cadenas o de segmentos de las mismas se puede regular mediante el valor de la constante del muelle angular Q^* , la cual está relacionada con la longitud de persistencia de la cadena (y por tanto, con la flexibilidad intrínseca de la macromolécula) mediante la expresión $P^* = 2Q^*/l_{eq}^*$.

La flexibilidad real de la cadena dependerá del cociente entre su longitud de contorno, L_{cont}^* , y la longitud de persistencia, P^* (a mayor L_{cont}^*/P^* , mayor flexibilidad). Un ángulo de equilibrio $\alpha_{eq} = \pi$ y una constante Q^* elevada sirven para representar una varilla rígida.

Para este tipo de modelos se emplearán constantes elevadas, de forma que los modelos resultantes presenten una flexibilidad muy limitada: para los muelles conectores, una constante de $H^* = 100$, junto con una longitud de equilibrio $l_{eq}^* = 1$, mientras que para los muelles angulares se define una constante de $Q^* = 100$ y un ángulo de equilibrio de $\alpha_0 = \pi$. El uso de estos parámetros implica que el modelo se comporte como una varilla rígida, aunque en los modelos de mayor tamaño se llega a apreciar cierta curvatura.

Además de los parámetros anteriores, en ambos casos se emplea un radio hidrodinámico de $\sigma^* = 0.256$ para las bolas del modelo, que, como se comentó en la sección 2.6 del

Capítulo 2, proporciona una representación adecuada del efecto de la interacción hidrodinámica en cadenas gaussianas que resulta independiente del tamaño de cadena [4].

4.3.2. Detalles de la simulación y el análisis.

Una vez que se han definido los elementos y parámetros que compondrán los modelos empleados, es necesario construir los mismos. El método más sencillo para construir estos modelos consiste en colocar en una posición arbitraria las sucesivas bolas que se van añadiendo al modelo, a una distancia igual a la longitud de equilibrio de la bola anterior. De esta manera, si bien aunque es probable que la cadena generada no se encuentre en equilibrio, al menos se consigue que no exista un solapamiento importante entre las bolas del mismo.

Para que la simulación comience a partir de un modelo en equilibrio, previamente se realiza un equilibrado rápido en el que se obvia el efecto de la interacción hidrodinámica. Mediante este equilibrado previo se aleatoriza la conformación del modelo, a la vez que se consigue que éste alcance el equilibrio. Este equilibrado tiene lugar durante 10^8 pasos brownianos, de una duración $\Delta t^* = 0.0001$, independientemente del tamaño del modelo. Mientras que para los modelos de mayor tamaño (100 y 200 bolas) esta duración puede incluso ser algo corta, debido a los grandes tiempos de relajación, para modelos pequeños se trata de un número muy alto de pasos, ya que el equilibrio se alcanza muchísimo antes, pero dado que el equilibrado transcurre en apenas unos segundos para el modelo más pequeño, y en algunos minutos para modelos algo mayores, no merece la pena detenerse demasiado en el estudio de la longitud óptima de este proceso.

Tras el equilibrado se lleva cabo el cálculo de la trayectoria definitiva, en la que sí se tiene en cuenta la interacción hidrodinámica. La duración de estas trayectorias es también de 10^8 pasos brownianos para todas las simulaciones de modelos semirrígidos del tipo 2, y casi todos los flexibles del tipo 1. Sólo en los modelos flexibles de mayor tamaño, $N = 100$ y $N = 200$, se emplean simulaciones de mayor duración: 10^9 y 1.8×10^9 pasos respectivamente. Independientemente del tamaño de la molécula simulada, el tamaño del paso browniano empleado es de $\Delta t^* = 0.0001$.

En el cálculo de las simulaciones, a diferencia del equilibrado previo, para tener en cuenta la interacción hidrodinámica entre los elementos del modelo, se usa el tensor de Rotne-Prager-Yamakawa [37, 38], descrito en el Capítulo 2. De esta manera, al tener

que realizar el tratamiento del tensor de difusión, el proceso de simulación se vuelve notablemente más lento que el proceso de equilibrado.

El método empleado para realizar la factorización del tensor de difusión es la utilización del algoritmo de Ermak-McCammon [6], que implementa la descomposición de Cholesky (sobre éste y otros métodos de tratamiento del tensor de difusión se hablará en el Capítulo 5). Debido al elevado coste computacional de este procedimiento, especialmente en los modelos de mayor tamaño, se hace deseable emplear algún tipo de estrategia para acelerar la obtención de resultados de las simulaciones. En concreto se emplean dos estrategias diferenciadas: en primer lugar, se implementan en los programas de simulación las funciones de la librería `Intel Math Kernel Library` (MKL), que están preparadas para el trabajo en paralelo sobre múltiples núcleos de cálculo, además de paralelizar en el propio código directivas `OpenMP` los bucles de cálculo relacionados con las funciones de autocorrelación necesarias para el cálculo de los parámetros D_E^* y D_1^* . En segundo lugar, se mantiene el mismo tensor de difusión durante un espacio de diez pasos brownianos consecutivos. Al mantener el tensor y no actualizarlo, se evita el tener que realizar nuevamente esta operación, pues ésta es la que supone el mayor consumo de tiempo de CPU en este algoritmo. Esta estrategia es factible y no afecta en gran medida a los resultados finales, puesto que, como se ha demostrado con anterioridad [52], las conformaciones macromoleculares, y por tanto el tensor de difusión, cambian muy poco en un espacio de tiempo tan breve como el que representan diez pasos de simulación con un paso browniano tan pequeño.

Durante la simulación, cada cien pasos de simulación se calculan y se almacenan en memoria volátil los siguientes datos: el valor del vector velocidad \mathbf{A} (necesario para el posterior cálculo de D_1^*), el coeficiente de difusión $D_{RB,i}^*$ de la conformación instantánea actual i , obtenido mediante la solución exacta de las ecuaciones de Kirkwood y Riseman (es decir, la aplicación del tratamiento del Cuerpo Rígido), y que más adelante servirá para obtener un valor promedio de D_{RB}^* ; finalmente, se registran también otros parámetros auxiliares relevantes para el posterior cálculo de los otros coeficientes de difusión que se han mencionado en la sección anterior.

Debe observarse que el proceder descrito en el párrafo anterior supone una excepción al procedimiento habitual de escribir a disco las coordenadas de posición de los elementos del modelo. El motivo de que se almacene directamente el valor de las propiedades en lugar de las trayectorias se debe a que para el cómputo de D_1^* se requiere un elevado número de registros y que éstos estén separados por un Δt^* pequeño (ya que la función de autocorrelación del vector \mathbf{A} decae rápidamente). Esto se traduce en un acceso al disco

muy frecuente y en un alto volumen de datos que almacenar, por lo que, para acelerar la simulación y el análisis de los datos, se prefiere conservar los datos en memoria volátil y permitir que sea el sistema operativo el que gestione el volcado de datos al disco (mediante técnicas de memoria virtual o *swapping*) en caso de que fuera necesario.

Una vez que se ha terminado de calcular la trayectoria al completo, se dispone de los datos necesarios para realizar el cálculo de D_1^* . Para un tiempo finito, t^* , la ecuación 4.6, adopta la forma

$$D_1^*(t^*) = \frac{1}{3N^2} \int_0^{t^*} d\tau \langle \mathbf{A}(0) \cdot \mathbf{A}(\tau^*) \rangle, \quad (4.8)$$

la cual se discretiza sustituyendo la integral entre 0 y t^* por un sumatorio sobre los $n = t^*/\Delta t^*$ intervalos de tiempo de duración Δt^* :

$$D_1^*(t^*) = \frac{\Delta t^*}{6N^2} \langle \mathbf{A}^2(0) \rangle + \frac{\Delta t^*}{3N^2} \sum_{k=1}^n \langle \mathbf{A}(0) \cdot \mathbf{A}(k\Delta t^*) \rangle. \quad (4.9)$$

La ecuación 4.9 requiere el cálculo de la función de autocorrelación del vector \mathbf{A} a lo largo de la trayectoria global. Para obtener además una estimación de la incertidumbre de los datos se realiza la autocorrelación de dicho vector a lo largo de cinco subtrayectorias correspondientes a la división de la trayectoria original en cinco partes iguales.

En segundo lugar, se procede a calcular el coeficiente D_E^* , esta vez a través de la autocorrelación temporal del desplazamiento promedio cuadrático del centro de masas a lo largo de la trayectoria calculada mediante la expresión 4.2. Para ello se calcula la función de autocorrelación hasta alcanzar un valor del tiempo de correlación de cien veces el tiempo entre registros, o lo que es lo mismo, $\tau^* = 100\Delta t_{reg}^* = 10^4\Delta t^*$. Y de nuevo, para obtener una estimación de la incertidumbre, se calcula el valor de la correlación para las cinco subtrayectorias ya mencionadas.

En tercer lugar, se obtiene el dato de D_K^* , calculado mediante la expresión 4.4, empleando los datos de promedio de distancia entre las bolas del modelo que se han almacenado a lo largo de la trayectoria.

Y finalmente, se recuperan los datos de $D_{RB,i}^*$ calculados durante la trayectoria, y se promedian para obtener un valor final promediado de D_{RB}^* para la trayectoria calculada.

4.4. Resultados.

Hemos llevado a cabo simulaciones para cadenas con diferente longitud de los modelos 1 y 2 descritos anteriormente. Para cada trayectoria, hemos evaluado los coeficientes de difusión definidos en la sección de teoría: D_E^* (ecuación 4.2), D_K^* (ecuación 4.4), D_L^* (ecuación 4.5 con D_1^* dado por la ecuación 4.9), y D_{RB}^* . Es importante destacar que en el caso de D_K^* se han empleado datos obtenidos de simulación numérica en lugar de ecuaciones que manejen datos pre-promediados.

N	D_K^*	D_1^*	$D_L^* = D_K^* - D_1^*$	D_E^*	D_{RB}^*
11	0.2504 ± 0.0003	0.0043 ± 0.0001	0.2461 ± 0.0002	0.250 ± 0.002	0.2409 ± 0.0002
21	0.1785 ± 0.0002	0.0038 ± 0.0002	0.1747 ± 0.0003	0.174 ± 0.003	0.1662 ± 0.0003
50	0.1084 ± 0.0002	0.0027 ± 0.0003	0.1057 ± 0.0002	0.1054 ± 0.0003	0.1000 ± 0.0001
100	0.0727 ± 0.0001	0.0018 ± 0.0002	0.0709 ± 0.0003	0.0719 ± 0.0006	0.0668 ± 0.0003
200	0.0505 ± 0.0002	0.00145 ± 0.00005	0.0490 ± 0.0002	0.0495 ± 0.0002	0.0459 ± 0.0002

TABLA 4.1: Coeficientes de difusión traslacional de cadenas lineales flexibles (gaussianas) correspondientes al modelo 1.

N	D_K^*	D_1^*	$D_L^* = D_K^* - D_1^*$
11	0.184007 ± 0.000003	0.00221 ± 0.00002	0.18180 ± 0.00002
21	0.111575 ± 0.000001	0.00122 ± 0.00001	0.11035 ± 0.00001
50	0.055697 ± 0.000001	0.000568 ± 0.000006	0.05531 ± 0.00001
100	0.031492 ± 0.000001	0.00039 ± 0.00001	0.03110 ± 0.00001
200	0.017578 ± 0.000001	0.000137 ± 0.000004	0.01744 ± 0.00005

N	D_E^*	D_{RB}^*
11	0.181 ± 0.001	0.181673 ± 0.000003
21	0.1102 ± 0.0006	0.109966 ± 0.000001
50	0.0547 ± 0.0002	0.054716 ± 0.000001
100	0.0311 ± 0.0001	0.030793 ± 0.000002
200	0.01753 ± 0.00001	0.017176 ± 0.000001

TABLA 4.2: Coeficientes de difusión traslacional de cadenas lineales semirrígidas (vermiformes) correspondientes al modelo 2.

En las tablas 4.1 y 4.2 se recogen los valores de los coeficientes de difusión obtenidos a partir de las simulaciones de dinámica browniana de varias cadenas lineales de diferente tamaño. Los valores de la tabla 4.1 se han obtenido a partir de modelos flexibles (tipo 1), y se corresponden bastante bien con los datos obtenidos por otros autores [25, 49] para modelos del mismo tipo. La tabla 4.2 recoge los valores obtenidos a partir de modelos semirrígidos (vermiformes) del tipo descrito como 2.

Como es lógico, los valores de los coeficientes de difusión disminuyen con el tamaño global de la macromolécula en disolución. Este efecto se observa tanto al aumentar el tamaño de la molécula, N , como al comparar los datos de las tablas 4.1 y 4.2 para un mismo tamaño de cadena: debido a su rigidez, los modelos semirrígidos siempre presentarán una conformación más extendida que los flexibles, lo que se manifiesta en sus propiedades.

Por otra parte, se comprueba que, tal como se adelantaba en la sección teórica de este capítulo, D_K^* supone efectivamente una cota superior para el valor de los coeficientes de difusión traslacional calculados, tanto para los modelos flexibles como para los semirrígidos. Sin embargo, el caso de D_{RB}^* , que teóricamente debería ser una cota inferior, no es tan claro: mientras en los modelos flexibles se aprecia claramente que D_{RB}^* es una cota inferior, en el caso de los modelos vermiformes esto no es estrictamente cierto, siendo en todo caso más difícil de apreciar, debido a que los valores de los diferentes coeficientes de difusión son muy próximos entre sí. Esto se debe a que este modelo se encuentra próximo al caso extremo de una partícula completamente rígida, caso en el que, en teoría, todos estos valores deberían ser coincidentes. Merced a esta proximidad en los valores del coeficiente de difusión, para mayor generalidad, aceptaremos la premisa anterior de que D_{RB}^* supone una cota inferior al valor de este parámetro. En base a lo anterior, D_L^* y D_E^* pueden considerarse equivalentes independientemente del modelo considerado, dado que ambos tienen valores muy próximos y se encuentran acotados entre los valores de D_K^* y D_{RB}^* .

Sin embargo, entre D_L^* y D_E^* existe una leve preferencia a favor de D_E^* , por calcularse directamente a partir del desplazamiento del centro de masas. Por este motivo, se considera que las aproximaciones más exactas para calcular el coeficiente de difusión son, en principio, aquellas que producen un valor adecuado de D_E^* (o el considerado equivalente D_L^*). Por lo tanto, en las comparaciones que van a establecerse, se empleará el valor de D_E^* como referencia.

Así, podemos calcular la desviación respecto a D_E^* que suponen las otras aproximaciones. En la tabla 4.3 se muestran tales desviaciones en forma de porcentaje de desviación.

N	Modelo flexible (tipo 1)			Modelo semirrígido (tipo 2)		
	D_K	D_L	D_{RB}	D_K	D_L	D_{RB}
11	2.9 %	1.04 %	3.1 %	1.66 %	0.44 %	0.37 %
21	2.7 %	0.51 %	4.3 %	1.29 %	0.18 %	0.17 %
50	2.9 %	0.33 %	5.1 %	1.82 %	0.79 %	0.03 %
100	1.08 %	1.41 %	7.2 %	1.30 %	0.04 %	0.95 %
200	2.02 %	1.01 %	7.3 %	0.29 %	0.50 %	2.01 %

TABLA 4.3: Diferencia entre los coeficientes de difusión de las tablas 4.1 y 4.2 y los correspondientes valores de D_E tomados como referencia.

La observación más evidente que se extrae de los datos de esta tabla ya se ha mencionado con anterioridad: los coeficientes de los modelos semirrígidos (tabla 4.2) son más próximos entre sí que los de los modelos flexibles (tabla 4.1). Esto parece indicar que cuanto mayor es la flexibilidad del modelo analizado, mayor es la diferencia entre D_E y las otras aproximaciones al coeficiente de difusión que se han estudiado.

Al aumentar el número de elementos de la cadena, y por tanto, su longitud, también aumenta la flexibilidad de la misma, en base a la relación entre la longitud de contorno L y la longitud de persistencia P que se ha mencionado con anterioridad. Sin embargo, no se aprecia una influencia importante de este efecto sobre el porcentaje de desviación de las demás aproximaciones con respecto a D_E .

Si se observan los resultados correspondientes al modelo 1, ha de tenerse en cuenta que éste es un modelo casi gaussiano (la única diferencia es la presencia del volumen excluido en nuestro modelo), y que incluso en las cadenas de menor tamaño ya existe bastante flexibilidad. Esto se manifiesta en el coeficiente D_{RB} , que parece ser el más sensible a la variación de la flexibilidad, de manera que la diferencia entre éste y D_E es creciente con N , especialmente en los casos de menor tamaño. El caso donde menos se aprecia esto es al pasar de $N = 100$ a $N = 200$, y podría deberse a que, pese a duplicarse el tamaño de la cadena, la flexibilidad del modelo de cien bolas ya es tan alta que el cambio de longitud no produce un efecto importante en la flexibilidad inherente de la molécula.

Por el contrario, las cadenas del modelo 2 son todas bastante rígidas, alcanzando un máximo de $L/P \approx 1$ para $N = 200$, lo que indica aún una importante rigidez, pues para que se aprecien de manera importante los efectos de la flexibilidad es necesario alcanzar

valores de $L/P \gg 1$. Únicamente en el caso de $N = 200$, y quizás también levemente en el de $N = 100$, se aprecia un ligero crecimiento de D_{RB} , que, como ya se ha mencionado, es el valor más sensible al aumento de la flexibilidad.

A partir de la observación de los D_{RB} de ambas tablas se concluye en primer lugar que, efectivamente, es el valor más afectado por el aumento de la flexibilidad. En segundo lugar, se observa también que el tratamiento del Cuerpo Rígido, lógicamente, es muy apropiado para el caso de trabajar con partículas de rigidez importante (de hecho, es la aproximación que produce los valores más cercanos a D_E). Sin embargo, al trabajar con cadenas grandes e intrínsecamente flexibles empeora su fiabilidad, aunque en el caso de los ensayos realizados, sus valores son representativos del coeficiente de difusión esperado.

Por otro lado, a pesar de estar basado en la misma teoría de Kirkwood y Riseman, el coeficiente D_K no se ve apenas afectado por la flexibilidad de las moléculas simuladas, y en ningún caso parece sufrir un empeoramiento de los resultados como consecuencia del aumento de la misma, por lo que esta aproximación puede seguir siendo válida incluso en los casos de moléculas grandes y flexibles en los que D_{RB} empieza a producir desviaciones apreciables respecto a D_E .

La explicación de este fenómeno puede deberse a que el coeficiente D_K se obtiene en relación a tiempos cortos, y está relacionado con movimientos locales en la cadena, en general, fenómenos independientes de la longitud de la cadena, mientras que el coeficiente D_{RB} está relacionado con el movimiento global de la cadena, y al igual que D_L se relaciona con el tiempo de relajación más largo, por lo que requiere de muestras muy amplias para ser determinado con cierto grado de exactitud.

En cualquier caso, todas las aproximaciones al cálculo del coeficiente de difusión traslacional que se han analizado parecen buenas opciones, pues en casi todos los casos, las diferencias con respecto a D_E no superan el 5% de desviación, por lo que pueden considerarse del mismo orden de magnitud que los errores experimentales habituales. Sólo para el caso de cadenas grandes, mayores que las empleadas habitualmente en simulaciones numéricas, el error puede volverse lo suficientemente relevante.

En general es D_L la aproximación más cercana a D_E , seguida de D_{RB} en el caso de moléculas rígidas (para las que esta aproximación es la mejor en casi todos los casos), y de D_K en el caso de las flexibles.

Finalmente, considerando D_K y D_{RB} como los límites que definen el rango entre los que se encuentran los coeficientes de difusión traslacional, podemos intentar extraer algunas

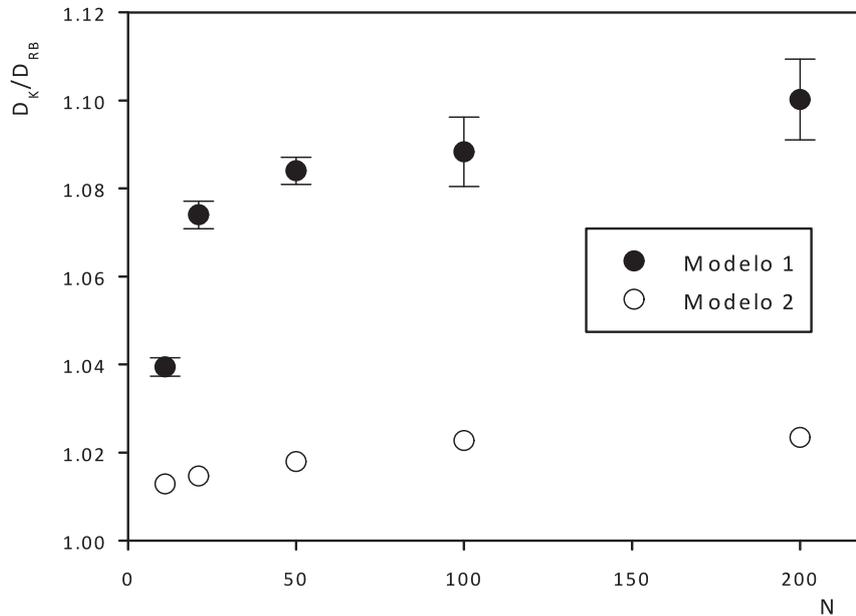


FIGURA 4.1: Variación del rango D_K/D_{RB} respecto al tamaño de la molécula para cada tipo de modelo. Las barras de error para el modelo semirrígido están incluidas en el gráfico, pero son demasiado pequeñas para ser apreciadas.

conclusiones sobre cómo varía este rango en función de la flexibilidad de la cadena. Con esta finalidad, definiremos este rango por medio de la relación D_K/D_{RB} . En la figura 4.1 se muestra la evolución de este parámetro frente al tamaño del modelo, según el tipo de modelo.

Tal como se había observado en los datos anteriores, el rango delimitado por estas dos cotas es mayor para las cadenas flexibles que para las semirrígidas, pero además se puede observar cómo en el caso de las cadenas flexibles este rango crece rápidamente con el tamaño de la cadena: para cadenas de pequeño tamaño (N bajo), el crecimiento es mucho mayor que para cadenas de mayor tamaño, donde llega a estabilizarse en un valor casi constante. Esto se debe a que el aumento de longitud de las cadenas sobre la flexibilidad tiene una mayor incidencia en las cadenas cortas, afectando al valor de D_{RB} , mientras que para las cadenas de mayor longitud apenas se aprecia ninguna influencia de la longitud sobre la flexibilidad.

En las cadenas rígidas se observa algo similar, aunque los valores del cociente son más pequeños que para cadenas flexibles. Así, para las cadenas de pequeño tamaño se

observa un ligero crecimiento de D_K/D_{RB} , mientras que para tamaños de cadena mayores, se alcanza un valor casi constante. En este caso, el rango de variabilidad D_K/D_{RB} para todas las cadenas es bastante similar debido a que todas ellas son bastante rígidas y no hay grandes diferencias entre las distintas aproximaciones al coeficiente de difusión.

4.5. Conclusiones del Capítulo.

A través de las simulaciones realizadas, se ha podido confirmar que existe una relación entre las diferentes aproximaciones estudiadas al cálculo del coeficiente de difusión traslacional de una macromolécula en disolución, que puede resumirse en la expresión

$$D_K > D_L \approx D_E > D_{RB}. \quad (4.10)$$

Esta expresión es válida para cualquier modelo de cadena macromolecular no rígida, e implica la existencia de unas cotas superior (D_K) e inferior (D_{RB}) del rango en el que puede encontrarse el coeficiente de difusión. En el caso límite de una molécula completamente rígida, los valores obtenidos mediante estas aproximaciones deben ser coincidentes, y como prueba de ello, al aumentar la rigidez del modelo, los valores obtenidos son más próximos entre sí.

Al mismo tiempo, se ha confirmado que D_L y D_E son equivalentes pues, aunque no producen resultados numéricos idénticos, la diferencia entre los mismos siempre es muy pequeña, e inferior al error que puede obtenerse en determinaciones experimentales.

En cuanto a la flexibilidad de los modelos estudiados, se ha encontrado que de las diferentes aproximaciones que se han estudiado, es el tratamiento del Cuerpo Rígido el que presenta una mayor sensibilidad al aumento de la flexibilidad de la molécula. Esta sensibilidad se manifiesta en un lento aumento de la divergencia respecto a los valores de D_E (considerado como el valor de referencia en base a lo intuitivo de su cálculo a partir de los desplazamientos del centro de masa), que para tamaños de cadena aún mayores que los estudiados aquí podría llegar a invalidar el uso de esta aproximación para calcular el valor del coeficiente de difusión.

Por otro lado, a pesar de estar basado en la misma teoría, el coeficiente D_K supone una aproximación mejor que D_{RB} para el caso de moléculas flexibles, pues al estar relacionado con fenómenos asociados a tiempos de relajación cortos soporta mejor el aumento de flexibilidad en las moléculas simuladas, siendo una alternativa mejor que el tratamiento del Cuerpo Rígido en estos casos.

CAPÍTULO 5

COMPARATIVA DE ALGORITMOS PARA LA DINÁMICA BROWNIANA CON INTERACCIÓN HIDRODINÁMICA

5.1. Introducción.

En el estudio del comportamiento de disoluciones diluidas de polímeros por medio de la simulación de dinámica browniana (DB) es esencial la consideración de la interacción hidrodinámica (IH) [4] entre los elementos del modelo (bolas) para reflejar correctamente la dinámica real de los polímeros [36].

Por medio de una gran cantidad de simulaciones se ha confirmado que el uso de una IH dependiente de la conformación instantánea, lo que también se conoce como IH fluctuante o IH no prepromediada, es esencial para estudiar de forma realista algunos fenómenos implicados en la dinámica de macromoléculas en periodos de tiempo largos, como el estiramiento y la orientación de polímeros sometidos a flujo [36, 53–55], la dinámica del ADN [34, 56] o el plegamiento de proteínas [57].

Sin embargo, el uso de IH fluctuante supone un alto coste computacional. Más aún, el tiempo de CPU aumenta dramáticamente con la longitud de la cadena, o lo que es lo mismo, el número de bolas del modelo, N . Esto se debe a que el método de DB-IH requiere la factorización del tensor de difusión de dimensión $3N \times 3N$ asociado a

cada conformación de la cadena macromolecular, lo que supone una operación costosa en términos de tiempo de CPU que debe repetirse cada vez que se actualiza el tensor de difusión. Como resultado de ello, esta factorización se convierte en el mayor cuello de botella del proceso de simulación.

La forma más común de construir el tensor de difusión macromolecular es emplear el tensor de Rotne-Prager-Yamakawa [37], que, como ya se ha explicado en la sección 2.4.2, posee la propiedad de ser definido positivo para cualquier conformación de la cadena, y se convierte en el tensor de Oseen cuando el cociente entre el radio de las bolas y la distancia entre las mismas tiende a cero.

El uso de la descomposición de Cholesky para la factorización del tensor de difusión, propuesta por Ermak y McCammon en su trabajo pionero en el uso de un algoritmo de DB con IH [6], supone un método de factorización exacta del tensor de difusión. Dicho método se ha implementado en posteriores algoritmos de DB-IH, como la modificación del algoritmo de Ermak-McCammon basada en predicción-corrección propuesta por Iniesta y García de la Torre [33], y se ha usado ampliamente simulaciones de DB-IH [34, 36, 54–57].

Como consecuencia del uso de esta descomposición, a nivel teórico se predice que, para N grande, el número de operaciones, y por tanto, el tiempo de CPU, t_{CPU} , consumido por algoritmos de este tipo crece en proporción a N^3 ; sin embargo, en la práctica sólo puede abordarse la simulación de moléculas de hasta cierto tamaño, o valor de N , pues los tiempos de cálculo de simulaciones de moléculas mayores se vuelven inasumibles. Así pues, en la práctica el valor de este exponente de escalado suele ser algo menor al que predice la teoría.

En 1986 Fixman propuso un procedimiento alternativo [58], basado en la aproximación de la factorización del tensor por medio del uso de un desarrollo en serie de Chebyshev. Con este procedimiento, el exponente del factor de escala de t_{CPU} con el tamaño de la cadena se reduce sustancialmente, y presenta un valor teórico aproximado de 2.25 (según [59], $9/4$), lo que es muy interesante a la hora de trabajar con sistemas de gran tamaño. El método de Fixman ha sido comprobado en varios trabajos [59–61], y también ha sido bastante empleado en diferentes simulaciones de DB-IH [62, 63].

Además de estos dos métodos principales, se han diseñado otras aproximaciones para el tratamiento de la DB-IH que pueden suponer ciertas ventajas para aplicaciones específicas [64, 65]. Sin embargo, actualmente la mayor parte de las simulaciones de DB que tienen en cuenta la IH siguen implementando, o bien la descomposición exacta de Cholesky, o bien la aproximación que emplea el desarrollo de Chebyshev propuesta por Fixman.

Recientemente, Geyer y Winter [61] han presentado una nueva aproximación que, según afirman sus autores, mejora a todas las anteriores, pues entre otras ventajas, t_{CPU} escalaría aproximadamente en proporción a N^2 sin suponer una pérdida de precisión significativa.

Aunque estos tres procedimientos generales ya han sido estudiados con anterioridad [59–61], hasta donde nosotros conocemos, sólo se han comunicado algunos aspectos de una comparación directa entre los tres procedimientos en la referencia [61]. Dado que creemos que esta cuestión merece algo más de atención, este Capítulo se centrará en realizar una comparación de su eficiencia, desde el punto de vista de la duración de las simulaciones, y precisión, entendida como la desviación respecto a los resultados obtenidos en una simulación empleando la descomposición exacta de Cholesky. Para esta última comparación, sobre las trayectorias calculadas con cada algoritmo se calcularán dos propiedades: una propiedad conformacional, el radio de giro, R_g , y una propiedad hidrodinámica, el coeficiente de difusión traslacional, D_t , y en base a ambas se establecerán las comparaciones.

Debido a que actualmente el paquete de simulación SIMUFLEX, presentado en la sección 3.4, sólo dispone de la implementación del algoritmo de Cholesky y que el objetivo de este Capítulo es la evaluación frente a éste último del método de Geyer y Winter y dos adaptaciones diferentes del método de Fixman, los algoritmos de Jendrejack y col. y el de Kröger y col., fue necesario crear nuevos programas de simulación, específicos para ese trabajo, que implementaran estos tres algoritmos, y también otro que implementara el algoritmo de Cholesky para que pudiese ser empleado como referencia en unas condiciones de máxima similitud con los anteriores.

La parte más general del código de los programas (inicialización de variables, gestión de la entrada y salida de datos, cálculo del tensor de difusión, cálculo del término determinista de la dinámica, etc) es común a estos programas, siendo diferente en cada uno de ellos únicamente la parte del código dedicada a la implementación de algoritmo empleado para el cálculo del término estocástico, que incluye la factorización del tensor de difusión. En las secciones subsiguientes se explican brevemente los detalles de la implementación de cada uno los algoritmos ensayados.

En estos programas, las operaciones matemáticas más generales, como la descomposición de Cholesky, el cálculo de los valores propios del tensor, o las operaciones entre matrices y vectores se realizan por medio de las correspondientes subrutinas del conocido y testado paquete de funciones matemáticas de dominio público LAPACK. La implementación de estas subrutinas se ha llevado a cabo a través del uso de la librería comercial Intel Math Kernel Library (MKL) versión 10.3, que incluye versiones optimizadas de

las funciones de LAPACK que además están preparadas para su uso en equipos de computación de alto rendimiento, permitiendo tanto el cálculo en paralelo en un mismo equipo como el cálculo distribuido en varios equipos.

En general, el uso de estas funciones para trabajo en paralelo supone una importante ventaja cuando se trabaja con matrices de gran tamaño, puesto que de esta manera es posible dividir el gran número de operaciones a realizar entre varios núcleos de cálculo, acelerando por tanto las operaciones.

Por ejemplo, en el cuerpo de código común compartido por los cuatro programas de simulación realizados se han podido implementar varias de estas funciones LAPACK optimizadas. En concreto, en esta parte común del código se emplean las funciones SAXPY, que realiza la operación $\mathbf{y} = a * \mathbf{x} + \mathbf{y}$, donde a es un escalar y \mathbf{x} e \mathbf{y} son vectores, y SSPMV, que calcula el producto de un vector por una matriz simétrica de almacenamiento empaquetado. En esta parte del código también se emplea la función VSRNGGAUSSIAN y sus auxiliares, exclusivas de la librería MKL (y por tanto, no presentes en LAPACK), que realizan la generación optimizada de vectores de números aleatorios con distribución gaussiana. Además, en cada uno de los algoritmos pueden emplearse algunas funciones más específicas que también aceleran los cálculos.

5.2. Base Teórica.

5.2.1. Generalidades.

La ecuación diferencial estocástica que gobierna el movimiento browniano de una estructura no inercial compuesta por N elementos de fricción (bolas esféricas) puede escribirse como:

$$d\mathbf{R} = [\mathbf{K} \cdot \mathbf{R} + (k_B T)^{-1} \mathbf{D} \cdot \mathbf{F}] dt + \sqrt{2\mathbf{B}} \cdot d\mathbf{w}, \quad (5.1)$$

cuya integración da lugar al conjunto de ecuaciones [2.24](#) presentado en el [Capítulo 2](#).

En esta forma de la ecuación, dt es el intervalo de tiempo infinitesimal sobre el que se define el paso de simulación. \mathbf{R} y \mathbf{F} son los vectores que, respectivamente, contienen las $3N$ coordenadas de posición de las bolas y las fuerzas internas deterministas que actúan sobre ellas. $k_B T$ es el factor de Boltzmann. La matriz \mathbf{K} , de dimensiones $3N \times 3N$, es el tensor gradiente de velocidad que se define cuando la disolución está sometida a flujo laminar homogéneo, tal que $\mathbf{K} \cdot \mathbf{R} = \mathbf{v}(\mathbf{R})$ es el vector que define la velocidad que tendría el fluido disolvente en las posiciones que ocupan las bolas si éstas no estuvieran presentes.

El término $(k_B T)^{-1} \mathbf{D} \cdot \mathbf{F}$ representa el movimiento determinista de la estructura debido a las fuerzas internas de la misma: \mathbf{D} es el tensor de difusión $3N \times 3N$, que según se definió en la sección 2.4.2, está compuesto por $N \times N$ bloques \mathbf{D}_{ij} de dimensiones 3×3 que definen la interacción entre cada par de bolas i, j con $i, j = 1 \dots N$. Recordemos que los elementos fuera de la diagonal de los bloques \mathbf{D}_{ij} son ceros, mientras que los de la diagonal son los coeficientes de difusión traslacional de las bolas individuales, $D_{t,i} = k_B T / \zeta_i$. Los bloques \mathbf{D}_{ij} con $i \neq j$ dependen de la forma adoptada para expresar la interacción hidrodinámica, siendo la forma más común el tensor de Rotne-Prager-Yamakawa [37], descrito en la sección 2.4.2, y que también se adoptará en este trabajo, dado que todas las bolas de los modelos empleados serán del mismo tamaño y tendrán los mismos valores de σ y ζ .

Finalmente, el último término de la expresión 5.1, $\sqrt{2} \mathbf{B} \cdot d\mathbf{w}$, representa el movimiento estocástico debido a los choques aleatorios de las moléculas del disolvente contra la estructura simulada. $d\mathbf{w}$ es un vector aleatorio de componentes reales y de distribución gaussiana, tal que su media es cero y su varianza, dt . \mathbf{B} es la matriz $3N \times 3N$ resultante de la factorización de \mathbf{D} , de forma que, como requiere el teorema de fluctuación-disipación:

$$\mathbf{D} = \mathbf{B} \cdot \mathbf{B}^T. \quad (5.2)$$

Para un \mathbf{D} cualquiera, es decir, para cualquier conformación dada del sistema de bolas, \mathbf{B} no es única; cualquier matriz \mathbf{B} que satisfaga la ecuación 5.2 puede ser utilizada en la ecuación 5.1. Éste es uno de los principales puntos de partida que dan origen los diferentes algoritmos de DB-IH.

5.2.2. Algoritmo de Ermak y McCammon. La descomposición de Cholesky.

Como ya se ha descrito en la sección 2.4.1, en este método se emplea el algoritmo de descomposición de Cholesky para realizar la factorización del tensor de difusión \mathbf{D} [6]. A lo largo de este Capítulo nos referiremos a este algoritmo como algoritmo de Cholesky para tener siempre presente la característica que estamos evaluando.

La descomposición de Cholesky se basa en el descubrimiento del matemático André-Louis Cholesky de que una matriz simétrica definida positiva puede ser descompuesta en el producto de una matriz triangular inferior por su traspuesta (o lo que es lo mismo, una matriz triangular superior por su correspondiente traspuesta). El procedimiento es una modificación de la factorización LU, de la que supone una derivación más eficiente.

En el ámbito de este trabajo, se considera que la descomposición del tensor de difusión \mathbf{D} resulta en una matriz triangular superior, \mathbf{C} , que satisface la ecuación 5.2:

$$\mathbf{D} = \mathbf{C} \cdot \mathbf{C}^T, \quad (5.3)$$

de forma que los procedimientos basados en Cholesky emplean $\mathbf{B} = \mathbf{C}$ en la ecuación 5.1 para realizar el cálculo de los desplazamientos estocásticos.

En este caso, la implementación de este procedimiento en los programas realizados se hace mediante el uso de la función de LAPACK “SPPTRF”, que realiza la factorización de Cholesky de matrices simétricas de almacenamiento empaquetado. Además del uso de esta función también resulta ventajoso el uso de la función “STPMV”, que calcula el producto de un vector por una matriz triangular empaquetada, que se usa para multiplicar la matriz \mathbf{C} resultante de la factorización por el vector aleatorio $d\mathbf{w}$ y obtener así el término estocástico.

Aunque no se utilizará en este Capítulo, una estrategia eficiente para la reducción del tiempo de proceso [52, 66] consiste en admitir que, debido a que las fluctuaciones en la interacción hidrodinámica son lentas, las diferencias entre las conformaciones a largo de una serie de pasos de simulación, típicamente, entre 5 y 50, son lo suficientemente pequeñas como para que pueda admitirse el uso del mismo tensor de difusión \mathbf{D} como representativo de todas ellas. Esta consideración implica poder conservar \mathbf{D} y \mathbf{B} durante dicha serie de pasos, con el consiguiente ahorro de tiempo de cálculo (el ahorro es prácticamente

proporcional al número de pasos durante el que se conserva el tensor), y sin afectar de manera importante a los resultados de la simulación.

5.2.3. Desarrollo en serie de Chebyshev y algoritmos derivados.

Como se mencionaba en la introducción de este Capítulo, como alternativa a la descomposición de Cholesky, Fixman [58] propuso un procedimiento extremadamente elegante, basado en una aproximación mediante un desarrollo en serie de Chebyshev para evaluar la matriz raíz cuadrada, \mathbf{S} , dada por

$$\mathbf{D} = \mathbf{S} \cdot \mathbf{S}, \quad (5.4)$$

que también puede ser empleada en la ecuación 5.1 tal que $\mathbf{B} = \mathbf{S}$, puesto que dada la simetría de \mathbf{D} se sigue que $\mathbf{S} = \mathbf{S}^T$. Nótese que en este caso, la matriz \mathbf{S} no es triangular, como en el caso de la descomposición de Cholesky.

A continuación se hace un pequeño resumen del procedimiento seguido para el cálculo de los desplazamientos aleatorios haciendo uso de la aproximación polinómica de Chebyshev propuesta por Fixman.

Sean λ_{min} y λ_{max} los valores propios mínimo y máximo de \mathbf{D} . Para cualquier tensor de difusión puede escribirse una ecuación tal que

$$\mathbf{E} = k_1 \mathbf{D} + k_2 \mathbf{I}, \quad (5.5)$$

de forma que la matriz transformada \mathbf{E} tendrá sus valores propios en el intervalo $[-1, 1]$ cuando los coeficientes k_1 y k_2 se obtengan de la manera:

$$k_1 = \frac{2}{\lambda_{max} - \lambda_{min}}, \quad (5.6)$$

$$k_2 = -\frac{\lambda_{max} + \lambda_{min}}{\lambda_{max} - \lambda_{min}}. \quad (5.7)$$

Al mismo tiempo, la matriz raíz cuadrada \mathbf{S} puede desarrollarse hasta un orden L mediante una serie de $L+1$ términos tensoriales de Chebyshev \mathbf{C}_l sobre la variable tensorial \mathbf{E} y coeficientes a_l ; es decir:

$$\mathbf{S} = \sum_{l=0}^L a_l \mathbf{C}_l \quad (5.8)$$

con $\mathbf{C}_0 = \mathbf{I}$, $\mathbf{C}_1 = \mathbf{E}$, \dots , $\mathbf{C}_{l+1} = 2\mathbf{E} \cdot \mathbf{C}_l - \mathbf{C}_{l-1}$.

En el desarrollo en serie representado por la ecuación 5.8, los escalares a_l están dados por la expresión [67]:

$$a_l = \frac{2}{L+1} \sum_{k=1}^{L+1} \cos\left(\frac{\pi l(k - \frac{1}{2})}{L+1}\right) f\left[\cos\left(\frac{\pi(k - \frac{1}{2})}{L+1}\right)\right], \quad (5.9)$$

donde la función $f(e)$ es la función raíz cuadrada de la variable escalar transformada e , tal que $f(e) = [(e - k_2)/k_1]^{1/2}$, o, lo que es lo mismo, $f(e) = [a_+ + a_- \cdot e]^{1/2}$ si se definen $a_+ = (\lambda_{max} + \lambda_{min})/2$ y $a_- = (\lambda_{max} - \lambda_{min})/2$.

A partir de este desarrollo en serie puede obtenerse la aproximación a la matriz \mathbf{S} , que a continuación debería multiplicarse por el vector $d\mathbf{w}$ para obtener el término estocástico de la ecuación de la dinámica browniana. Sin embargo, Fixman observó que es más interesante introducir también el vector $d\mathbf{w}$ en el desarrollo en serie, y obtener así, de forma directa, una aproximación del producto $\mathbf{y} = \mathbf{S} \cdot d\mathbf{w}$, que desarrollado en serie hasta el término L podría expresarse como

$$\mathbf{y}_L = \left[\sum_{l=1}^L a_l \mathbf{x}_l \right] + \frac{a_0}{2} \mathbf{x}_0, \quad (5.10)$$

donde los vectores \mathbf{x}_l se corresponden con:

$$\mathbf{x}_0 = d\mathbf{w}, \quad (5.11)$$

$$\mathbf{x}_1 = \mathbf{E} \cdot d\mathbf{w}, \quad (5.12)$$

$$\mathbf{x}_{l+1} = 2\mathbf{E} \cdot \mathbf{x}_l - \mathbf{x}_{l-1}. \quad (5.13)$$

Tal como puede imaginarse, este método también requiere un número importante de operaciones de cálculo, pues en principio, sería necesario determinar los valores propios del tensor cada vez que éste se actualiza (lo cual también es una operación costosa en términos de tiempo de CPU), para a continuación realizar el desarrollo en serie. Por este motivo, es necesario desarrollar estrategias que permitan reducir el número de operaciones, pues de lo contrario este método emplearía tiempos de CPU superiores a los empleados por el método de Cholesky.

En este trabajo se analizan dos algoritmos basados en el desarrollo de Chebyshev, el algoritmo de Jendreck y col., y el de Kröger y col. Las estrategias en las que se basa cada uno, junto con una visión general de ellos, se detallan a continuación en sendas subsecciones.

Algoritmo de Jendreck y col.

El funcionamiento de este algoritmo [60], que en adelante, para abreviar denominaremos únicamente como “algoritmo de Jendreck”, es el más complejo de los cuatro analizados. Se basa en el método de Fixman, y utiliza el desarrollo polinómico de Chebyshev para aproximar la raíz cuadrada del tensor de difusión. Este algoritmo emplea un número variable de términos en el desarrollo, ajustado en cada paso de simulación en función de la precisión obtenida en la aproximación a $\mathbf{S} \cdot d\mathbf{w}$. Dicha precisión se estima por medio de un parámetro E_f que representa un error relativo de la aproximación, calculado mediante la expresión:

$$E_f = \sqrt{\frac{|\mathbf{y} \cdot \mathbf{y} - d\mathbf{w} \cdot \mathbf{D} \cdot d\mathbf{w}|}{d\mathbf{w} \cdot \mathbf{D} \cdot d\mathbf{w}}}, \quad (5.14)$$

en la que todas las magnitudes ya han sido definidas anteriormente (ver ecuaciones 5.1 y 5.10). E_f sirve como una medida del grado en el que se viola el teorema de fluctuación-disipación. El cálculo de E_f durante cada paso la simulación es una operación cuyo gasto en tiempo de CPU es pequeño en comparación con la factorización de \mathbf{D} , y gracias a este parámetro se puede estimar si la aproximación conseguida es suficiente, y por tanto se trunca el desarrollo en serie en el término actual, o no lo es, y se calcula el siguiente término del desarrollo.

La información que está disponible al comienzo del cálculo de cada paso de simulación es la siguiente:

En primer lugar, entre los datos de entrada de la simulación se dispone de un valor límite de E_f , al que denominaremos $E_{f,lim}$, por debajo del cual se considera que la precisión alcanzada en el cálculo de \mathbf{y} es suficiente. Además al comienzo de cada paso de simulación se calcula el tensor de difusión \mathbf{D} correspondiente a la nueva conformación, y se genera un nuevo vector aleatorio $d\mathbf{w}$.

También se dispone de la información con la que se ha realizado el paso previo, en concreto, se conocen los valores de λ_{min} , λ_{max} y L empleados. Dado que los pasos de simulación son cortos, se supone que los cambios en el tensor de difusión son pequeños, y los valores de λ_{min} y λ_{max} pueden ser válidos también para el nuevo tensor de difusión, por lo que se emplearán también para el nuevo paso. En caso de que esta hipótesis no fuese válida, y los valores propios, y por lo tanto también el número de términos del desarrollo, no fuesen aceptables para el nuevo tensor, se pondrá de manifiesto a lo largo del cálculo, y serán recalculados. Además, esta hipótesis es válida para todos los pasos de la simulación salvo el primero, para el que, lógicamente, no hay disponible información sobre el paso anterior. En este caso particular se establecen, de forma arbitraria, unos valores de $\lambda_{min} = 0.8$, $\lambda_{max} = 1.2$ y $L = 3$, con los que se puede realizar el cálculo del paso. Lo más probable es que estos valores no correspondan a los de la conformación inicial, por lo que el cálculo del paso de simulación fallará, y λ_{min} y λ_{max} serán recalculados, tal como se indicaba anteriormente.

Disponiendo de estos datos, el procedimiento por el que se calcula el paso de simulación es como sigue:

1. En primer lugar, se calcula y almacena el valor del producto $d\mathbf{w} \cdot \mathbf{D} \cdot d\mathbf{w}$, empleado como referencia en el cálculo de E_f . Este valor es independiente de los valores propios y del número de términos del desarrollo, por lo que puede calcularse *a priori* y conservarse durante todo el paso de simulación.
2. En segundo lugar, a partir de los datos que se conservan del paso anterior, se calculan los coeficientes a_i hasta el término L , y se obtienen los coeficientes k_1 y k_2 , junto con la matriz transformada \mathbf{E} . Aunque finalmente no fuese necesario utilizar todos los L términos del desarrollo para dar el paso, se obtienen todos los coeficientes a_i ($i = 1 \dots L$), debido a que es más sencillo calcularlos todos que calcularlos uno por uno en el caso de ser necesarios.
3. Se obtienen los vectores \mathbf{x}_0 y \mathbf{x}_1 de acuerdo a las relaciones 5.11 y 5.12. Con ellos se obtiene \mathbf{x}_2 mediante la relación recurrente 5.13 que también se usará para todos los términos siguientes.

4. Con esos tres primeros vectores y los correspondientes coeficientes se calcula una primera aproximación $\mathbf{y}_3 \approx \mathbf{S} \cdot d\mathbf{w}$.

5. Se calcula y evalúa el valor de E_f correspondiente a esta primera aproximación.

Si la precisión es suficiente ($E_f \leq E_{f,lim}$), se acepta \mathbf{y}_l ; se conservan los valores de λ_{min} y λ_{max} actuales, y se fija el valor de L para el siguiente paso como el número de términos actual del desarrollo, l , mas un margen de seguridad de tres términos. Hecho esto, se actualiza el valor de las coordenadas de las bolas y se prosigue con el siguiente paso, reiniciando el algoritmo.

En caso contrario ($E_f > E_{f,lim}$), si el número de términos empleados para aproximar \mathbf{y} es inferior a L , se calcula y añade al desarrollo el siguiente término del mismo mediante la expresión 5.13, y aplicando el correspondiente coeficiente a_l . Este procedimiento se repite tantas veces como sea necesario hasta que, o bien se cumpla la condición mencionada en el párrafo anterior y se alcance la precisión suficiente y se proceda de la forma estipulada, o bien hasta que el número de términos utilizados sea igual a L , en cuyo caso se prosigue con el siguiente punto.

6. Si se han calculado L términos del desarrollo y no se ha alcanzado la precisión suficiente, se asume que la conformación ha cambiado lo suficiente como para que los valores de λ_{min} y λ_{max} del paso anterior ya no sean válidos, y será necesario recalcularlos diagonalizando el tensor de difusión \mathbf{D} correspondiente a la conformación actual.

En cuanto al valor de L , se sabe con seguridad que, en algún momento, el desarrollo a partir de los nuevos valores propios producirá una aproximación válida de \mathbf{y} , pero se desconoce cuántos términos serán necesarios, por lo que sería conveniente no limitar el número de términos del desarrollo ($L = \infty$), o establecer un límite muy alto, pues, de no actualizar el valor de L , podría darse la circunstancia de que incluso con los nuevos valores propios se llegase a calcular un número de términos del desarrollo $l = L$ sin alcanzar el límite de precisión dado. En tal caso, el algoritmo volvería a llegar hasta el presente punto, se recalcarían los valores propios λ_{min} y λ_{max} , y dado que la conformación no ha cambiado, se obtendrían los mismos valores que la vez anterior, de forma que el programa entraría en un bucle sin fin.

Sin embargo, por otro lado, el valor de L está implicado en el cálculo de los coeficientes del desarrollo (ver ecuación 5.9), por lo que no puede tomarse como infinito, y además, si se fijara un límite del número de términos del desarrollo L demasiado elevado, ralentizaría el cálculo de estos coeficientes y también la simulación.

Por este motivo, el algoritmo de Jendreck emplea una solución de compromiso: en caso de recalcar λ_{min} y λ_{max} , se establece el valor de L como el doble de su valor

previo. De esta forma, si por segunda vez no se alcanzara la precisión suficiente, volvería a duplicarse el número máximo de términos del desarrollo, cuadruplicando el valor inicial, y así sucesivamente, hasta superar un límite preconvencido de $L = L_{max} = 1000$, caso en el que se consideraría que el algoritmo ha fracasado. Este último caso es altamente inusual, y normalmente sería indicativo de algún problema en el algoritmo o en la simulación, y afortunadamente, nunca se llegó a producir en las simulaciones que se realizaron para este trabajo.

En este algoritmo, las funciones matemáticas susceptibles a ser evaluadas mediante funciones de la librería MKL son las funciones empleadas para el cálculo de los valores propios del tensor de difusión, **SSPTRD** y **SSTERF**, que, respectivamente, sirven para tridiagonalizar una matriz simétrica de almacenamiento empaquetado y hallar todos los valores propios de una matriz de este tipo. También son de utilidad las ya mencionadas funciones **SSPMV** y **SAXPY**, empleadas en este caso para el cálculo del desarrollo de Chebyshev. Finalmente, en el cálculo de E_f se emplea la función **SDOT** que calcula el producto escalar de dos vectores.

Algoritmo de Kröger y col.

Este segundo algoritmo basado en el método de Fixman [59], al que de nuevo por simplicidad denominaremos “algoritmo de Kröger”, también hace uso del desarrollo en serie de Chebyshev. Sin embargo, en este caso se emplea un número constante de términos del desarrollo L , determinado de antemano, y se utilizan los mismos valores propios durante toda la simulación, sin que se realice ninguna clase de chequeo sobre la precisión del método.

Kröger y col. basan su método en la observación que aparece en el propio trabajo de Fixman de que no es necesario que los valores de λ_{min} y λ_{max} empleados en el desarrollo sean realmente los valores propios mínimo y máximo del tensor de difusión, sino que es suficiente con que sean una cota inferior y otra superior de los mismos.

Sobre esta base, Kröger y col. desarrollan unas ecuaciones teóricas que permiten la predicción de los valores propios mínimo y máximo del tensor de difusión promediado en equilibrio para un modelo de Zimm (es decir, una cadena con interacción hidrodinámica y sin interacciones de volumen excluido), Λ_{max}^{Zimm} y Λ_{min}^{Zimm} , en función de los parámetros del modelo N (número de bolas) y σ (radio de las bolas) [59]:

$$\Lambda_{min}^{Zimm} \approx 1 - 1.71\sigma \quad (5.15)$$

$$\Lambda_{max}^{Zimm} \approx 1 + \pi N^{1/2}\sigma \quad (5.16)$$

Así mismo, postulan que los valores propios de los tensores de difusión instantáneos a lo largo de una simulación cercana al equilibrio, caracterizados por sus valores máximo y mínimo, respectivamente, Λ_{max}^{inst} y Λ_{min}^{inst} , están acotados entre los valores propios obtenidos en el equilibrio mediante las relaciones $2\Lambda_{min}^{inst} > \Lambda_{min}^{Zimm}$ y $\Lambda_{max}^{inst} < 2\Lambda_{max}^{Zimm}$, por lo que los valores

$$\Lambda_{min} = \frac{1 - 1.71\sigma}{2} \quad (5.17)$$

$$\Lambda_{max} = 2(1 + \pi N^{1/2}\sigma) \quad (5.18)$$

son una cota inferior y una superior válidas para los valores propios de los tensores de difusión que pueden darse a lo largo de una simulación cercana al equilibrio de un modelo de Zimm de N bolas de radio σ .

Teniendo en cuenta estas estimaciones, y dando crédito a la observación hecha por Fixman de que el número de términos necesarios en un desarrollo de Chebyshev para obtener una precisión determinada es proporcional a la raíz cuadrada del cociente de los valores propios [58], Kröger y col. proponen calcular el número de términos del desarrollo de manera que L tome el valor entero truncado de la expresión

$$L = \sqrt{\frac{\Lambda_{max}}{\Lambda_{min}}} + 1 \quad (5.19)$$

definida para obtener una precisión del orden de $\epsilon_F = 10^{-2}$ [59], entendiendo el error de la manera definida por Fixman [58]:

$$\epsilon_F = \frac{norm(\mathbf{S}_L - \mathbf{S})}{norm(\mathbf{S})} \quad (5.20)$$

donde \mathbf{S} es el valor considerado exacto del desarrollo, \mathbf{S}_L la aproximación conseguida empleando L términos en el desarrollo y la operación norma de una matriz se toma como la suma de los valores absolutos de todos sus elementos.

De esta manera, como ya se ha dicho, el método de Kröger no comprueba la precisión del método a lo largo de la simulación, sino que emplea unas cotas superior e inferior de los valores propios, y un número de términos del desarrollo de Chebyshev constante a lo largo de toda la simulación.

Debido a que los modelos estudiados en este Capítulo, a diferencia de los estudiados por Kröger, sí tienen interacciones de volumen excluido no es posible utilizar las ecuaciones 5.17 y 5.18 para estimar las cotas de los valores propios. Por este motivo, éstas se obtienen a partir de un tensor de difusión prepromediado, adecuado a nuestras cadenas con volumen excluido, calculado a partir de simulaciones previas sin interacción hidrodinámica. Dicho tensor de difusión se construye utilizando las expresiones para el tensor de Oseen prepromediado, que sí es definido positivo. En nuestras unidades reducidas, las expresiones para el cálculo de los bloques de la diagonal, \mathbf{D}_{ii} , y de fuera de la diagonal, \mathbf{D}_{ij} , del tensor prepromediado son:

$$\mathbf{D}_{ii} = \mathbf{I} \quad (5.21)$$

$$\mathbf{D}_{ij} = \sigma \left\langle \frac{1}{r_{ij}} \right\rangle \mathbf{I}. \quad (5.22)$$

Siguiendo a Kröger y col., para asegurarnos que efectivamente estamos tomando cotas reales de los valores propios de los tensores de difusión que se den a lo largo de la simulación, usaremos como cotas superior e inferior el doble del mayor valor propio del tensor prepromediado y la mitad del menor, respectivamente.

Así, una vez que se han calculado al comienzo de la simulación las cotas de los valores propios, a partir de ellas se obtienen de forma inmediata el número de términos del desarrollo, L , y los coeficientes a_l . A partir de ahí, el problema del cálculo del término estocástico de la simulación se reduce a calcular, en cada paso de la simulación, la matriz \mathbf{E} derivada del tensor de difusión, obtener el vector aleatorio $d\mathbf{w}$, y los vectores \mathbf{x}_l .

En el caso estricto del algoritmo de Kröger, las funciones de LAPACK/MKL empleadas son las funciones SSPMV y SAXPY, ya referidas anteriormente, y que en este caso también se emplean en el cálculo del desarrollo en serie de Chebyshev.

Dada la necesidad de calcular los valores propios del tensor de difusión prepromediado, se emplean las funciones `SSPTRD` y `SSTERF` (también mencionadas en el algoritmo de Jendrejack) para obtener los valores propios del mismo; sin embargo, dado que la diagonalización y obtención de los valores propios se realiza una única vez al principio del cálculo, el uso de las librerías optimizadas en este caso no tiene prácticamente ninguna influencia, debido a que el tiempo de CPU requerido por esta diagonalización inicial es muy pequeño frente al total de tiempo de cálculo de la trayectoria.

5.2.4. Desarrollo truncado de Geyer y Winter.

A diferencia de los métodos de Ermak-McCammon y de Fixman, el método desarrollado por Geyer y Winter, al que denominan “aproximación del desarrollo truncado” o TEA (las siglas de “*Truncated Expansion Ansatz*”), no se basa en el uso de un algoritmo matemático que permita una factorización más rápida del tensor de difusión \mathbf{D} , sino que se basa en un desarrollo teórico apoyado en ciertas hipótesis que permiten la simplificación del cálculo del término $\sqrt{2}\mathbf{B} \cdot d\mathbf{w}$ de la ecuación 5.1 [61]. Cabe destacar que en este tratamiento se descompone el tensor \mathbf{D} en sus elementos individuales D_{ij} , entendiendo que en esta sección, a diferencia del resto de los Capítulos de esta Tesis, los subíndices i y j se refieren a coordenadas, y no a partículas.

El punto de partida de este desarrollo consiste en la idea de realizar un tratamiento del término estocástico de la ecuación de la dinámica browniana 5.1 similar al que se da al término determinista. Tomando este último término de la ecuación 2.24, y transformándolo en base a lo propuesto por Geyer y Winter se obtiene:

$$\Delta r_i^{det} = \frac{\Delta t}{k_B T} \sum_{j=1}^{3N} D_{ij}^0 \cdot F_j^0 = \frac{D_{ii}^0 \Delta t}{k_B T} F_i^{(0)(eff)}, \quad (5.23)$$

donde, según los autores, $F_i^{(0)(eff)}$ representa una “fuerza efectiva hidrodinámicamente corregida”,

$$F_i^{(0)(eff)} = \sum_{j=1}^{3N} \frac{D_{ij}^0}{D_{ii}^0} F_j^0. \quad (5.24)$$

Al hacer, como se propone, un tratamiento similar del término estocástico de la ecuación 2.24, se reformula el término de los desplazamientos aleatorios en función una fuerza aleatoria no correlacionada f_j ,

$$\Delta r_i^{aleat} = \frac{D_{ii}^0 \Delta t}{k_B T} C_i \sum_{j=1}^N \beta_{ij} \frac{D_{ij}^0}{D_{ii}^0} f_j = \frac{D_{ii}^0 \Delta t}{k_B T} f_i^{(0)(eff)}. \quad (5.25)$$

Para elaborar esta expresión se propone la introducción de dos nuevos parámetros, C_i y β_{ij} . De esta manera, se da lugar a la definición de una “fuerza aleatoria hidrodinámicamente corregida”, $f_i^{(0)(eff)}$, cuya expresión sería:

$$f_i^{(0)(eff)} = C_i \sum_{j=1}^N \beta_{ij} \frac{D_{ij}^0}{D_{ii}^0} f_j. \quad (5.26)$$

Comparando las definiciones de $F_i^{(0)(eff)}$ y $f_i^{(0)(eff)}$, es importante destacar que, mientras la primera, por su definición a partir del término determinista, es función de \mathbf{D}^0 , la segunda, por estar relacionada con el término estocástico, debe estar relacionada con $\sqrt{\mathbf{D}^0}$. Para tener en cuenta numéricamente esta diferencia se introducen los escalares β_{ij} , mientras que C_i funciona como un factor de escala cuya función es garantizar que preserve el valor de los coeficientes de difusión D_{ii}^0 para cada coordenada i .

Como se expone a lo largo del desarrollo, C_i es un conjunto de factores de escala, diferentes para cada coordenada, mientras que por razones de simetría sólo existen dos valores para los coeficientes β_{ij} : un valor β_{ii} para los elementos correspondientes a la diagonal del tensor de difusión, a la que se denotará simplemente como β , y otro valor $\beta_{i \neq j}$ para los elementos situados fuera de la diagonal, que de aquí en adelante se denotará como β' .

Si se imponen a las fuerzas aleatorias no correlacionadas f_j las restricciones tales que

$$\langle f_j \rangle = 0, \quad (5.27)$$

$$\langle f_i f_j \rangle = \frac{2(k_B T)^2}{D_{ii} \Delta t} \delta_{ij}, \quad (5.28)$$

entendiendo por δ_{ij} como la función delta de Kronecker, es fácil ver que automáticamente se cumplirá también la primera de las condiciones que debe cumplir el término estocástico Δr_i^{aleat} de la ecuación de la dinámica browniana (ver sección 2.4.1), tal que $\langle \Delta r_i^{aleat} \rangle = 0$.

Para la comprobación de la segunda condición es necesario partir del producto $\Delta r_i^{aleat} \Delta r_j^{aleat}$, que según lo presentado en la ecuación 5.25, valdrá

$$\Delta r_i^{aleat} \Delta r_j^{aleat} = \left(\frac{\Delta t}{k_B T} \right)^2 C_i C_j \sum_{k,l} \beta_{ik} \beta_{jl} D_{ik} D_{jl} f_k f_l. \quad (5.29)$$

Si se aplica a la expresión anterior la condición 5.28 mencionada antes y se eliminan los términos en los que $k \neq l$, pues las fuerzas \mathbf{f}_j no están correlacionadas, al combinar el resultado con la restricción impuesta a la covarianza de los desplazamientos aleatorios de la ecuación de la dinámica browniana, se tiene que:

$$\langle \Delta r_i^{aleat} \Delta r_j^{aleat} \rangle = 2\Delta t C_i C_j \sum_k \beta_{ik} \beta_{jk} \frac{D_{ik} D_{jk}}{D_{kk}} = 2D_{ij} \Delta t. \quad (5.30)$$

Si se particulariza esta igualdad para el caso en el que $i = j$, se obtiene una expresión que permite determinar las constantes de normalización C_j :

$$\left(\frac{1}{C_i} \right)^2 = \sum_k \beta_{ik}^2 \frac{D_{ik}^2}{D_{ii} D_{kk}}. \quad (5.31)$$

Llegados a este punto, Geyer y Winter fijan de forma arbitraria el valor de $\beta_{ii} = \beta = 1$, lo que puede hacerse sin pérdida de generalidad. Al hacer esto, es posible extraer el término correspondiente a $k = i$ del sumatorio de la ecuación 5.31, y realizar una última transformación:

$$\left(\frac{1}{C_i} \right)^2 = \beta_{ii}^2 \frac{D_{ii}^2}{D_{ii} D_{ii}} + \sum_{k \neq i} \beta_{ik}^2 \frac{D_{ik}^2}{D_{ii} D_{kk}} = 1 + \sum_{k \neq i} \beta_{ik}^2 \frac{D_{ik}^2}{D_{ii} D_{kk}}. \quad (5.32)$$

A partir de este punto, Geyer y Winter afirman que no existe una forma eficiente de determinar un juego de coeficientes que satisfaga la ecuación anterior y se cumplan las condiciones impuestas. Por este motivo, los autores se ven obligados a recurrir a una aproximación para reducir la complejidad de la expresión hasta un nivel que permita obtener los parámetros adecuados.

Dicha aproximación consiste en considerar que la interacción hidrodinámica es débil, de forma que los elementos situados fuera de la diagonal del tensor de difusión, $D_{ij} (\forall i \neq j)$, son mucho menores que los situados en la diagonal, D_{ii} , de forma que si se define el cociente ϵ como el promedio de la fuerza de acoplamiento relativa, tal que

$$\epsilon = \langle D_{ij}/D_{ii} \rangle, \quad (5.33)$$

se puede tomar la aproximación

$$\sqrt{1 + \epsilon} \approx 1 + \epsilon/2, \quad (5.34)$$

y de esta forma, la ecuación 5.32 puede aproximarse a

$$\frac{1}{C_i} \approx 1 + \frac{1}{2} \sum_{k \neq i} \beta_{ik}^2 \frac{D_{ik}^2}{D_{ii} D_{kk}}, \quad (5.35)$$

lo que trasladado a la expresión 5.30 se traduce en la aproximación

$$\frac{1}{C_i C_j} \approx 1 + \frac{1}{2} \sum_{k \neq i} \frac{\beta_{ik}^2 D_{ik}^2}{D_{ii} D_{kk}} + \frac{1}{2} \sum_{l \neq j} \frac{\beta_{jl}^2 D_{jl}^2}{D_{jj} D_{ll}} + O\left(\frac{D_{ij}}{D_{ii}}\right)^4. \quad (5.36)$$

Si en esta última expresión se tiene en cuenta la aproximación establecida de que $D_{ij} \ll D_{ii}$, entonces el término elevado a la cuarta potencia es despreciable. Reescribiendo la ecuación en función del parámetro ϵ definido con anterioridad, la expresión se simplifica a:

$$\frac{1}{C_i C_j} \approx 1 + (3N - 1) \beta^2 \epsilon^2. \quad (5.37)$$

Por otra parte, partiendo directamente de la expresión 5.30, se deduce la siguiente relación exacta (sin recurrir a ninguna clase de aproximación):

$$\frac{1}{C_i C_j} = \sum_k \beta_{ik} \beta_{jk} \frac{D_{ik} D_{jk}}{D_{kk} D_{ij}}. \quad (5.38)$$

En esta nueva expresión, del término de la derecha pueden extraerse del sumatorio los dos términos en los que $k = i$ y $k = j$, pues son independientes de la interacción hidrodinámica (corresponden a la interacción de una bola consigo misma), y haciendo esto, y reescribiendo la expresión en función del parámetro ϵ (teniendo además en cuenta que $\beta_{ik} = \beta'$ para todo i, k),

$$\frac{1}{C_i C_j} = 2\beta' + (3N - 2)\beta'^2 \epsilon. \quad (5.39)$$

Igualando las expresiones 5.37 y 5.39, pues ambas se refieren al valor de $(C_i C_j)^{-1}$, se obtiene una ecuación cuadrática en función del número de coordenadas, $3N$, y ϵ ,

$$\beta'^2 \left[(3N - 1)\epsilon^2 - (3N - 2)\epsilon \right] - 2\beta' + 1 = 0 \quad (5.40)$$

a partir de la cual es trivial obtener el valor de β' . Sobre esta expresión, Geyer y Winter apuntan que no está definida en ausencia de interacción hidrodinámica, y que en caso de tener una interacción hidrodinámica débil, es decir, con $\epsilon \rightarrow 0$, el valor de β' converge hacia 0.5 (el valor que toma en caso de que $\epsilon = 0$).

La mayor complejidad de este método reside en la deducción de las ecuaciones teóricas 5.40 y 5.32, pues la implementación en los programas de simulación mediante dinámica browniana (a la que a partir de ahora, por comodidad, nos referiremos como “algoritmo de Geyer”) es muy sencilla. Además, su aplicación supone la gran ventaja de que su consumo de tiempo de CPU escala en proporción a $O(N^2)$, aunque por contra está limitado, debido a las consideraciones hechas en su desarrollo, a los casos en los que la interacción hidrodinámica entre las partículas del modelo es débil.

El funcionamiento del algoritmo una vez implementado en los programas de simulación podría resumirse en lo siguiente:

En primer lugar, se calcula el parámetro ϵ analizando los elementos individuales del tensor de difusión, previamente calculado para obtener el valor del desplazamiento determinista, mediante la expresión 5.33.

A continuación, para obtener β' se resuelve la ecuación de segundo grado planteada por 5.40:

$$\beta' = \frac{1 - \sqrt{1 - [(N-1)\epsilon^2 - (N-2)\epsilon]}}{[(N-1)\epsilon^2 - (N-2)\epsilon]} \quad (5.41)$$

y finalmente, para obtener los parámetros de escalado C_i , se emplea la igualdad 5.32. Una vez que se dispone de todos estos datos, el desplazamiento estocástico puede obtenerse de forma sencilla mediante el uso de la expresión 5.25.

Debido a que en este método no se trabaja con matrices, sino que se trabaja con los componentes individuales del tensor de difusión, no es posible emplear funciones de librería específicas para su cálculo. Sin embargo, sí que sería posible emplear directivas de trabajo en paralelo de tipo OpenMP para dividir los bucles de cálculo de los diferentes parámetros en múltiples hilos de cálculo. Sin embargo, debido a la gran rapidez de este algoritmo, este método de optimización resulta ineficiente para los tamaños del modelo analizados, pues en todos los casos resulta en aumentos del tiempo de cálculo, por lo que se evita este tipo de paralelización.

5.3. Detalles sobre las simulaciones.

5.3.1. Modelos empleados en la simulación.

Para la parte principal de la comparación de estos algoritmos se construyeron una serie de modelos de diferentes tamaños, comprendidos entre $N = 10$ y $N = 210$, con las condiciones más desfavorables posibles, de manera que las diferencias entre los resultados producidos por cada uno de los procedimientos fueran tan evidentes como fuera posible. Estas condiciones corresponden a cadenas con conectores de tipo gaussiano, sin restricciones angulares e interacciones de volumen excluido de tipo Lennard-Jones entre las bolas que no están directamente conectadas. Este modelo, al que nos referiremos de aquí en adelante como “modelo 1a”, fue elegido debido a que, gracias a tratarse de un modelo muy flexible, posee una gran variedad conformacional, y admite un cierto grado de solapamiento entre las bolas que lo componen.

Los elementos que componen este modelo ya han sido descritos, y las expresiones de los potenciales que los gobiernan se han presentado en la sección 2.2; sin embargo,

para ser implementados en el algoritmo de DB es necesario derivar las expresiones de los potenciales para obtener las ecuaciones de las fuerzas internas a las que se asocian. Así, para el caso del muelle gaussiano le correspondería la expresión:

$$F_{Gauss} = Hl, \quad (5.42)$$

donde H es la constante del muelle y l es la longitud instantánea del muelle.

En cuanto al volumen excluido de Lennard-Jones, su expresión en términos de la fuerza aplicada entre un par de bolas i y j separadas una distancia r_{ij} viene dado por

$$F_{LJ} = \frac{24\epsilon_{LJ}}{r_{ij}} \left[2 \left(\frac{\sigma_{LJ}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{LJ}}{r_{ij}} \right)^6 \right], \quad (5.43)$$

donde, al igual que en la expresión del potencial asociado (ecuación 2.18), ϵ_{LJ} y σ_{LJ} representan la energía potencial mínima de atracción alcanzable y la distancia a la que el potencial se anula.

La elección de un potencial de Lennard-Jones para representar las interacciones de largo alcance y su identificación con el caso más desfavorable posible se debe a tres razones, fundamentalmente.

En primer lugar, los autores de cada uno de los algoritmos utilizan un potencial de volumen excluido diferente en cada caso: Kröger y col. [59] lo obviaban (aunque evitaban el solapamiento de las bolas utilizando un flujo de cizalla débil); Jendrejack y col. [60] lo tenían en cuenta mediante el uso de un potencial repulsivo suave (ecuación 2.19), similar al propuesto por Rey y col. [25], mientras que Geyer y Winter [61] usan un potencial armónico repulsivo entre todas las bolas, análogo al usado por Ermak y McCammon [6]. Es de esperar que las condiciones en las que se efectúa la demostración de cada uno de éstos algoritmos sea ventajosa para su aplicación, mientras que nosotros nos proponemos evaluar los diferentes algoritmos frente a un modelo común, que no necesariamente represente las condiciones idóneas de aplicación de estos algoritmos.

Por otra parte, el volumen excluido facilita la implementación del método de Fixman dado que dificulta en gran medida que las bolas no consecutivas se solapen [60]. Los tres trabajos citados, así como éste, implementan la IH a través del tensor de Rotne-Prager-Yamakawa (ver sección 2.4.2), el cual está preparado para trabajar con bolas solapantes.

Sin embargo, si el solapamiento es muy frecuente, además de la pérdida del significado físico de la aproximación de Rotne-Prager-Yamakawa, el método de Fixman puede volverse poco práctico, dado el incremento necesario en el orden del desarrollo polinómico para conseguir un determinado grado de precisión [60].

Y finalmente, se decidió utilizar un potencial de Lennard-Jones porque posee la capacidad de emular todo el rango de condiciones de calidad del medio disolvente, y en concreto, se escogieron los valores de los parámetros ϵ_{LJ} y σ_{LJ} que reproducen condiciones theta apropiadas para el trabajo con cadenas lineales gaussianas [68]. Esta elección se debe a que dado que bajo estas circunstancias las leyes de escala que relacionan las propiedades de las macromoléculas en disolución con el peso molecular están bien establecidas teóricamente, lo que será útil para comprobar la calidad de los resultados obtenidos.

Para terminar de definir el modelo 1a, dado que el propósito de este trabajo es evaluar el rendimiento de los algoritmos de DB-IH sin simular ningún sistema real concreto, y para que los resultados sean comparables a los de la bibliografía consultada, todas las magnitudes presentadas en este Capítulo serán empleadas en forma reducida, utilizando las unidades descritas en la sección 2.6.

En este sistema de unidades, la constante de muelle gaussiano se convierte en $H^* = 3$, y los valores de los parámetros de Lennard-Jones que representan las condiciones theta del disolvente son $\epsilon_{LJ}^* = 0.3$ y $\sigma_{LJ}^* = 0.8$ [68]. También se usa un radio hidrodinámico de las bolas adimensional, $\sigma^* = 0.256$, que corresponde a un parámetro hidrodinámico $h^* = 0.25$, el cual minimiza el efecto de la longitud de cadena [4, 59]. El lapso de tiempo utilizado para la integración será de $\Delta t^* = 10^{-4}$, tal como requieren las severas condiciones fijadas por el potencial de Lennard-Jones. Los asteriscos, como se indicó anteriormente, denotan que se trata de magnitudes adimensionales.

Además del modelo 1a, se han empleado algunos otros para responder a cuestiones que fueron surgiendo durante el desarrollo del trabajo.

El primero de estos modelos alternativos es un modelo similar al 1a, con la diferencia de que los muelles empleados en este caso son “pseudo-gaussianos”, en los que se ha alterado el valor de la constante de los muelles de forma que sea $H^* = 6$. Esta modificación hace que los muelles sean algo más rígidos, favoreciendo ligeramente el solapamiento de las bolas vecinas (que, recordemos, no están relacionadas mediante el volumen excluido). A este modelo, por su similitud con el anterior, lo denominaremos **modelo 1b**.

En segundo lugar, se emplea también un modelo de cadena vermiforme o de Kratky-Porod [69], un modelo habitualmente empleado para representar moléculas con cierta rigidez, como por ejemplo, polisacáridos y otras moléculas de origen biológico. El modelo, al que denominaremos **modelo 2**, se construye con muelles hookeanos duros (ecuación 2.14, convenientemente derivada para obtener la expresión para la fuerza), en los que se usa una constante de muelle alta, $H^* = 100$, y una longitud de equilibrio $l_e^* = 1$. También se definen muelles angulares entre cada terna de bolas consecutivas, caracterizados por la derivada de la ecuación 2.17 y los parámetros $Q^* = 10$ y $\alpha_e = 0$. Si recordamos que el ángulo α se define como el ángulo suplementario al definido por los conectores, estos parámetros implican que el ángulo determinado por cada pareja de conectores consecutivos será próximo al ángulo llano.

El último modelo empleado fue definido a imitación del empleado por Geyer y Winter en la comprobación de su algoritmo TEA [61]. En este modelo, al que denominaremos **modelo 3**, se emplean muelles hookeanos duros tanto para conectar las bolas consecutivas, como para representar el volumen excluido, que en este modelo se define entre todos los pares de bolas, estén o no conectadas. Para conseguir la máxima similitud con el trabajo de Geyer y Winter, a diferencia de los modelos anteriores, en este caso se define $\sigma^* = 1$. Para ambas instancias de los muelles hookeanos se establece que $H^* = 100$; además, para el caso de los muelles conectores se define $l_e^* = 3\sigma^*$, mientras que para el volumen excluido $l_e^* = 2\sigma^*$ y $r_{cut-off} = 2\sigma^*$.

5.3.2. Simulaciones y cálculo de propiedades.

Para obtener la máxima similitud posible en las condiciones de las simulaciones realizadas con cada uno de los procedimientos, se recurre a realizar, para cada tamaño de cadena y modelo, la repetición de la misma simulación (los mismos parámetros y la misma conformación inicial) con cada uno de los algoritmos para el tratamiento de la interacción hidrodinámica presentados en la sección siguiente.

Las conformaciones iniciales necesarias para las simulaciones se generaron aleatoriamente, evitando el solapamiento de las bolas. Las conformaciones generadas se llevaron al equilibrio corriendo simulaciones lo suficientemente largas sin tener en cuenta la IH (en caso contrario el tiempo de equilibrado se torna excesivo). A continuación, las conformaciones resultantes en forma de ovillo al azar se toman como conformaciones iniciales para las simulaciones con IH.

Como base para realizar la comparación entre los resultados de las simulaciones realizadas con los diferentes algoritmos se eligieron las propiedades radio de giro y el coeficiente de difusión traslacional, que han sido presentadas en el Capítulo 2, con la salvedad de que al usar magnitudes reducidas en las simulaciones, también las propiedades calculadas se obtienen en forma reducida: R_g^* y D_t^* . Por simplicidad en los cálculos, para el caso del coeficiente de difusión traslacional, se emplea la definición de Kirkwood, la cual ha sido empleada en trabajos anteriores de verificación de algoritmos [60], que en la presente Tesis se estudia en el Capítulo 4, y se define de forma más extensa en la sección 4.2.2. A efectos de implementación, la expresión empleada para calcular dicha propiedad es la ecuación 4.4.

Debido a que las cadenas flexibles de gran tamaño con interacciones de volumen excluido presentan tiempos de relajación altos, las simulaciones deben tener una longitud considerable para que pueda barrerse un espacio conformacional lo suficientemente amplio como para que los resultados obtenidos sean representativos de la macromolécula. Dado que estas simulaciones conllevan unos tiempos muy largos, para las simulaciones realizadas para la comparación de la precisión de los algoritmos, se decide llevar a cabo la estrategia planteada en la sección 3.3 de dividir las simulaciones largas en varias subsimulaciones más cortas y estadísticamente independientes que serán ejecutadas en paralelo en los *clusters* de cálculo a los que tenemos acceso. De este modo, se llegó a tener en funcionamiento hasta 128 simulaciones al mismo tiempo. La independencia de unas subsimulaciones de otras se consigue empleando distintas semillas del generador de aleatorios en los archivos de cada subsimulación. Para la generación de los archivos de entrada, la gestión de los mismos y la recopilación de los resultados se emplearon los programas auxiliares descritos en la sección 3.6.

Para comprobar que las trayectorias generadas eran lo suficientemente largas, se incrementó la duración de las mismas, y se comprobó que pese a la prolongación, el valor de las propiedades calculadas (R_g^* y D_t^*) no variaba apreciablemente, lo que confirmaba que las simulaciones se encontraban en equilibrio y los resultados obtenidos son representativos de la dinámica de la macromolécula.

El número de subsimulaciones, la longitud en pasos de cada una de ellas, y el número total de pasos empleados en las simulaciones para cada tamaño de molécula son los que se reflejan en la tabla 5.1.

En cambio, para las simulaciones realizadas para llevar a cabo la comparación de los algoritmos desde el punto de vista de la eficacia de los mismos, se fijó la duración de las

N	Subsimulaciones	n° pasos / subsim.	n° pasos totales
10	1	10^7	10^7
30	1	10^7	10^7
70	8	10^7	$8 \cdot 10^7$
120	30	10^7	$3 \cdot 10^8$
160	60	10^7	$6 \cdot 10^8$
210	60	10^7	$6 \cdot 10^8$

TABLA 5.1: Número de subsimulaciones, pasos en cada subsimulación y pasos totales para cada tamaño de modelo en las simulaciones realizadas para la comparación de precisión en los resultados de los diferentes procedimientos.

simulaciones a 10^6 pasos brownianos, independientemente del tamaño del modelo. Este número de pasos es insuficiente para producir una trayectoria representativa del espacio conformacional de los modelos simulados, pero dado que estas simulaciones no se utilizarán para el cálculo de propiedades, sino que únicamente se analizará su duración en términos de tiempo de CPU, no es necesario prolongar más su duración.

5.4. Resultados.

Como se ha adelantado, la comparación realizada entre los cuatro métodos de tratamiento del tensor de interacción hidrodinámica se llevará a cabo teniendo en cuenta dos factores: la precisión de los resultados obtenidos, entendiendo por tal la divergencia de los resultados obtenidos mediante los algoritmos de Geyer, Jendrejack y Kröger respecto a los obtenidos mediante el algoritmo de Cholesky, así como la rapidez de cada ellos, evaluada a partir de la duración de simulaciones de 10^6 pasos cada una.

La necesidad de contrastar la precisión de los resultados obtenidos empleando cada uno de los algoritmos es evidente: salvo el algoritmo original de Ermak-McCammon, que emplea la factorización de Cholesky para producir una raíz exacta del tensor de difusión, los demás métodos citados se sirven de diversas aproximaciones para obtener el desplazamiento estocástico, introduciendo por lo tanto, un cierto grado de inexactitud en

las trayectorias calculadas, y por tanto, también en las propiedades. Así pues, para realizar el estudio de la precisión de los resultados producidos mediante el empleo de cada uno de los métodos propuestos, se tomarán como referencia los resultados producidos mediante el método de Cholesky (es decir, Ermak-McCammon).

Una vez comprobada la precisión de los métodos, se evaluará la rapidez de cada uno de ellos, pues el objetivo último de este Capítulo es el de encontrar alternativas al método de Cholesky que permitan reducir la duración de las simulaciones.

5.4.1. Comparación de propiedades.

Modelos con solapamiento.

En las tablas 5.2 y 5.3 se presentan los resultados de las simulaciones realizadas con el modelo 1a (cadenas de muelles gaussianos, $H^* = 3$, y volumen excluido de Lennard-Jones). Tanto para los coeficientes de difusión, como para los radios de giro obtenidos se ha calculado la desviación de cada dato respecto al valor obtenido a partir de la simulación equivalente realizada mediante el método de Cholesky.

En las tablas se puede observar que la incertidumbre de todos los resultados es bastante inferior al 1%, lo que nos reafirma en la conclusión de que las simulaciones han sido lo suficientemente largas. En segundo lugar, puede observarse una doble tendencia muy clara, especialmente en los datos del coeficiente de difusión: por una parte están los datos provenientes de las simulaciones llevadas a cabo mediante los métodos de Cholesky y Geyer, que muestran un elevado grado de coincidencia entre ellos, y por el otro lado están los datos obtenidos mediante los métodos basados en el de Fixman (Jendrejack y Kröger), en este caso, con una coincidencia casi total de los resultados. Sin embargo, estas dos tendencias se diferencian entre sí, y la diferencia crece cuanto mayor es el tamaño de las cadenas simuladas, llegando a diferir hasta casi un 20% en el caso de $N = 210$.

Una primera conclusión que se puede obtener de estos datos consiste en que, comparando los métodos de Kröger y Jendrejack, ambos producen resultados muy similares, por lo que se puede afirmar que la hipótesis utilizada por Kröger acerca de las cotas de los valores propios es válida, y, o bien no se producen violaciones de las cotas, o estas son muy puntuales y no influyen en la estadística conformacional global. En base a esto, se puede concluir que mediante este método es posible obtener buenos resultados de las simulaciones, aun a pesar de que durante el procedimiento de simulación no se realiza ningún tipo

Algoritmo	N = 10		N = 30	
	D_t^*	Dif. Chol.	D_t^*	Dif. Chol.
Cholesky	0.2576 ± 0.0002	-	0.1378 ± 0.0001	-
Geyer	0.2596 ± 0.0002	0.8 %	0.1387 ± 0.0001	0.7 %
Jendrejack	0.2635 ± 0.0003	2.3 %	0.1429 ± 0.0001	3.7 %
Kröger	0.2634 ± 0.0003	2.3 %	0.1433 ± 0.0001	3.9 %

Algoritmo	N = 70		N = 120	
	D_t^*	Dif. Chol.	D_t^*	Dif. Chol.
Cholesky	0.0870 ± 0.0003	-	0.0658 ± 0.0002	-
Geyer	0.0875 ± 0.0001	0.6 %	0.0673 ± 0.0001	2.3 %
Jendrejack	0.0933 ± 0.0001	7.3 %	0.0740 ± 0.0001	12.5 %
Kröger	0.0938 ± 0.0001	7.8 %	0.0740 ± 0.0001	12.4 %

Algoritmo	N = 160		N = 210	
	D_t^*	Dif. Chol.	D_t^*	Dif. Chol.
Cholesky	0.0563 ± 0.0001	-	0.0499 ± 0.0001	-
Geyer	0.0577 ± 0.0001	2.6 %	0.0512 ± 0.0001	2.6 %
Jendrejack	0.0638 ± 0.0001	13.4 %	0.0587 ± 0.0001	17.7 %
Kröger	0.0641 ± 0.0001	14.0 %	0.0594 ± 0.0001	19.0 %

TABLA 5.2: Coeficientes de difusión traslacional y diferencia respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y modelo 1a de muelles gaussianos ($H^* = 3$).

de comprobación de la violación de los límites de los valores propios del tensor de difusión ni de la bondad de la aproximación.

En cuanto a las diferencias entre los resultados de los algoritmos basados en el desarrollo de Chebyshev y los otros, Fixman [58] ya encontraba algunos problemas al trabajar con sistemas en los que se produce solapamiento entre bolas del modelo, afirmando que la presencia de solapamiento implicaba aumentar el número de términos del desarrollo de Chebyshev necesario para alcanzar una cierta precisión en la aproximación. Basándonos en esta observación, nos proponemos comprobar si la divergencia de los resultados obtenidos mediante los algoritmos de Jendrejack y Kröger respecto a los de Cholesky (que tomaremos como referencia por ser obtenidos mediante un método de factorización exacto) pudiera estar relacionada con este factor.

Algoritmo	N = 10		N = 30	
	R_g^*	Dif. Chol.	R_g^*	Dif. Chol.
Cholesky	1.45 ± 0.01	-	2.75 ± 0.01	-
Geyer	1.42 ± 0.01	2.1 %	2.68 ± 0.01	2.4 %
Jendrejack	1.40 ± 0.01	3.3 %	2.61 ± 0.01	5.1 %
Kröger	1.40 ± 0.01	3.2 %	2.60 ± 0.01	5.3 %

Algoritmo	N = 70		N = 120	
	R_g^*	Dif. Chol.	R_g^*	Dif. Chol.
Cholesky	4.33 ± 0.03	-	5.61 ± 0.03	-
Geyer	4.22 ± 0.01	2.5 %	5.23 ± 0.01	6.7 %
Jendrejack	3.89 ± 0.01	10.2 %	4.68 ± 0.01	16.6 %
Kröger	3.85 ± 0.01	11.2 %	4.70 ± 0.01	16.2 %

Algoritmo	N = 160		N = 210	
	R_g^*	Dif. Chol.	R_g^*	Dif. Chol.
Cholesky	6.57 ± 0.01	-	7.16 ± 0.01	-
Geyer	6.10 ± 0.01	7.1 %	6.63 ± 0.01	7.4 %
Jendrejack	5.43 ± 0.01	17.4 %	5.63 ± 0.01	21.3 %
Kröger	5.39 ± 0.01	18.0 %	5.54 ± 0.01	22.6 %

TABLA 5.3: Radios de giro y diferencia respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y modelo 1a de muelles gaussianos ($H^* = 3$).

Para ello se harán varios ensayos con modelos diferentes. En primer lugar, se utilizará un modelo similar al “1a”, el modelo “1b”, que, recordemos, emplea muelles levemente más duros que los gaussianos ($H^* = 6$), mientras que conserva una longitud de equilibrio nula. Este ligero endurecimiento de los muelles se traducirá en que las bolas consecutivas del modelo se atraerán con más fuerza, y por lo tanto, aumentará el número de solapamientos presentes. Debido a la duración de las simulaciones con mayor número de bolas, en este caso sólo se simularán modelos entre $N = 10$ y $N = 120$, lo que debería ser suficiente como para que se manifestase la influencia del solapamiento sobre los resultados.

Los resultados de estas simulaciones se presentan en las tablas 5.4 y 5.5. En ellas puede observarse cómo el endurecimiento de los muelles conduce a un importante aumento en la divergencia entre los resultados obtenidos mediante los métodos que implementan el

desarrollo de Chebyshev y el valor de referencia obtenido mediante el método de Cholesky, mientras que la diferencia en los resultados obtenidos mediante el método de Geyer no varía excesivamente. Mientras el error de éste último aumenta entorno a un 1% en los casos de tamaño intermedio ($N = 70$ y $N = 120$), el error en el coeficiente de difusión obtenido mediante los métodos basados en el desarrollo de Chebyshev para cadenas de $N = 120$ pasa de un 13% en los resultados de las simulaciones para muelles gaussianos, a un 22% de diferencia en los resultados obtenidos de las simulaciones con el modelo 1b.

Una vez que se ha comprobado que un aumento en el grado de solapamiento (a través del endurecimiento de los muelles) presente en las simulaciones aumenta la divergencia de los resultados de los métodos basados en el desarrollo de Chebyshev frente a los obtenidos por el método basado en la descomposición de Cholesky, también pueden extraerse algunas conclusiones respecto al crecimiento de estas diferencias al aumentar el tamaño de la cadena simulada.

Por una parte, al aumentar el tamaño de la cadena crecen en gran medida la variabilidad conformacional y los tiempos de relajación, y también se hace mayor la relación entre los valores propios máximo y mínimo del tensor de difusión. En concreto, el solapamiento hace que el valor propio mínimo decrezca, y por lo tanto según la ecuación 5.19, L aumente (según dicha ecuación, el caso extremo de un solapamiento completo de las bolas haría que $\lambda_{min} = 0$, y por tanto, $L = \infty$). Por otro lado, dado que el modelo de muelles gaussianos implica cierto grado de solapamiento entre bolas adyacentes (en las no adyacentes el volumen excluido lo impide), al aumentar N , también aumenta la posibilidad de que se produzcan solapamientos entre bolas vecinas.

Considerando simultáneamente estos dos efectos, se puede esperar que, como efectivamente ocurre, el valor de L aumente al crecer el tamaño de la molécula simulada. En la tabla 5.6 se presentan los datos de L extraídos de las simulaciones llevadas a cabo con los algoritmos basados en el desarrollo de Chebyshev. Es necesario recordar que, si bien para el procedimiento de Kröger el valor de L es constante a lo largo de la simulación, en el de Jendrejack L es variable en cada paso, por lo que se han tomado como datos representativos el valor promedio de L a lo largo de la simulación, L_{medio} , y el máximo número de términos que se llega a usar a lo largo de la simulación, L_{max} .

Tal como se esperaba, los datos de la tabla confirman la hipótesis del crecimiento del número de términos del desarrollo con el tamaño de las cadenas simuladas. Además, salta a la vista la relación existente entre el valor de L en las simulaciones basadas en el método de Kröger y el L promedio de Jendrejack, siendo este último, en todos los casos, muy

Algoritmo	N = 10		N = 30	
	D_t^*	Dif. Chol.	D_t^*	Dif. Chol.
Cholesky	0.2913 ± 0.0002	-	0.1607 ± 0.0001	-
Geyer	0.2937 ± 0.0002	0.8 %	0.1618 ± 0.0001	0.7 %
Jendrejack	0.2991 ± 0.0003	2.7 %	0.1683 ± 0.0001	4.8 %
Kröger	0.2991 ± 0.0002	2.7 %	0.1689 ± 0.0002	5.1 %

Algoritmo	N = 70		N = 120	
	D_t^*	Dif. Chol.	D_t^*	Dif. Chol.
Cholesky	0.1047 ± 0.0003	-	0.0811 ± 0.0003	-
Geyer	0.1065 ± 0.0001	1.7 %	0.0842 ± 0.0001	3.3 %
Jendrejack	0.1173 ± 0.0001	11.9 %	0.0986 ± 0.0001	21.6 %
Kröger	0.1176 ± 0.0001	12.3 %	0.0989 ± 0.0001	22.0 %

TABLA 5.4: Coeficientes de difusión traslacional y diferencia respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y modelo 1b ($H^* = 6$).

Algoritmo	N = 10		N = 30	
	R_g^*	Dif. Chol.	R_g^*	Dif. Chol.
Cholesky	1.19 ± 0.01	-	2.22 ± 0.01	-
Geyer	1.16 ± 0.01	2.0 %	2.15 ± 0.01	2.9 %
Jendrejack	1.14 ± 0.01	3.7 %	2.05 ± 0.01	7.7 %
Kröger	1.14 ± 0.01	3.8 %	2.04 ± 0.01	8.2 %

Algoritmo	N = 70		N = 120	
	R_g^*	Dif. Chol.	R_g^*	Dif. Chol.
Cholesky	3.37 ± 0.02	-	4.29 ± 0.03	-
Geyer	3.20 ± 0.01	5.3 %	3.89 ± 0.01	8.4 %
Jendrejack	2.77 ± 0.01	17.9 %	3.06 ± 0.01	28.7 %
Kröger	2.74 ± 0.01	18.6 %	3.03 ± 0.01	29.3 %

TABLA 5.5: Radios de giro y diferencia respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y modelo 1b ($H^* = 6$).

N	Jendrejack	Kröger
	$L_{\text{medio}} / L_{\text{max}}$	L
10	14 / 75	5
30	18 / 84	7
70	24 / 114	8
120	27 / 112	9
160	29 / 115	10
210	32 / 126	11

TABLA 5.6: Número de términos calculados en los desarrollos de Chebyshev de los algoritmos de Jendrejack y Kröger para las simulaciones con el modelo 1a.

próximo al triple del primero, lo que tendrá importantes repercusiones sobre el tiempo de cálculo, como se verá más adelante.

Especialmente en el caso de las simulaciones con el algoritmo de Jendrejack se debería tener en cuenta que el alto número de términos en el desarrollo puede llevar asociado una cierta acumulación de errores debidos al redondeo. Teniendo en cuenta que el orden de magnitud de los sucesivos términos del desarrollo es decreciente, es probable que a partir de cierto punto, los errores de cálculo incluidos en un término comiencen a ser superiores a la mejora del resultado aportado por dicho término, o dicho de otra manera, llegará un momento en el que los términos adicionales del desarrollo no mejoren la precisión del resultado o incluso lleguen a empeorarlo.

Sin embargo, como se ha mencionado, en promedio, a pesar de que el número de términos del desarrollo en las simulaciones con el algoritmo de Kröger es próximo a la tercera parte de los de las simulaciones con el algoritmo de Jendrejack, las diferencias respecto a los resultados obtenidos en las simulaciones que emplean el método de Cholesky son muy similares, por lo que no se puede atribuir esta diferencia a la acumulación de errores de redondeo.

Modelos sin solapamiento.

Si anteriormente se había ensayado un modelo con un solapamiento mayor que el del modelo gaussiano, en esta ocasión se recurrió al caso contrario: se simularon cadenas

de tipo vermiforme bastante rígidas, correspondientes al modelo 2, que apenas varían su conformación, e impiden prácticamente cualquier solapamiento. De esta manera, se descarta que la rigidez de los muelles sea la causa de los malos resultados del algoritmo de Chebyshev, pues en este caso, se emplearán muelles mucho más duros que en el modelo 1b, en el que se había observado que el grado de divergencia aumentaba respecto al modelo 1a.

En concreto, se simularon únicamente cadenas con $N = 70$, compuestas por muelles conectores duros con $H^* = 100$ y $l_{eq}^* = 1$, y muelles angulares duros con $Q^* = 10$ y $\alpha_0 = 0$. Ambos elementos y las expresiones relacionadas se describieron en la sección 2.2. En este modelo ya no se emplean interacciones de volumen excluido, puesto que la rigidez de los conectores junto a una distancia de equilibrio entre bolas bastante mayor que su radio, dificulta en gran medida la posibilidad de que se produzcan solapamientos. Los resultados obtenidos se presentan en las tablas 5.7 y 5.8.

Algoritmo	N = 70	
	D_t^*	Dif. Chol.
Cholesky	0.04797 ± 0.00003	-
Geyer	0.04757 ± 0.00003	0.8 %
Jendrejack	0.04776 ± 0.00003	0.4 %
Kröger	0.04760 ± 0.00003	0.8 %

TABLA 5.7: Coeficientes de difusión traslacional y diferencia respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y el modelo 2 (vermiforme).

Algoritmo	N = 70	
	R_g^*	Dif. Chol.
Cholesky	11.35 ± 0.02	-
Geyer	11.89 ± 0.02	4.8 %
Jendrejack	11.67 ± 0.02	2.8 %
Kröger	11.78 ± 0.02	3.8 %

TABLA 5.8: Radios de giro y diferencia respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y el modelo 2 (vermiforme).

En primer lugar se observa que las diferencias entre los valores obtenidos con cada uno de los métodos y el método de referencia de Cholesky son menores que las observadas tanto para los modelos 1a como 1b, siendo los datos obtenidos prácticamente idénticos sea cual sea el tratamiento de la interacción hidrodinámica empleado para obtenerlos. Curiosamente, en este caso la mayor desviación respecto a los datos del método de Cholesky es la que presenta el radio de giro calculado mediante el uso del algoritmo de Geyer, que presenta un valor algo más alto de lo observado con anterioridad en los casos de $N = 70$, aunque no llega a ser una diferencia significativa.

Estos resultados descartan la dureza de los muelles como causa de las desviaciones respecto a los resultados del método de Cholesky, e indica que, como ya se suponía, las diferencias encontradas al trabajar con los algoritmos basados en el desarrollo de Chebyshev, se deben al solapamiento. Por lo tanto, teniendo en cuenta esta limitación, este algoritmo sería una buena opción para simular, por ejemplo, macromoléculas biológicas de carácter semiflexible, que suelen modelarse de manera similar a las cadenas empleadas en esta simulación.

Finalmente, como comprobación adicional, y verificación de los resultados obtenidos en el caso anterior mediante el algoritmo de Geyer, se llevaron a cabo las simulaciones empleando el modelo armónico (modelo 3), utilizado por Geyer y Winter para poner a prueba su algoritmo [61]. Inicialmente, estas simulaciones se emplearon para verificar nuestras implementaciones de los algoritmos de simulación, sin embargo, también nos sirvieron para confirmar definitivamente el comportamiento de los algoritmos al emplear sistemas donde se impide el solapamiento. En este caso, se realizaron simulaciones de modelos con $N = 70$ y $N = 160$.

Este modelo, al ser una reproducción del estudiado por Geyer y Winter, no presenta ninguna novedad desde el punto de vista de los valores de las propiedades estudiadas, más allá de corresponderse perfectamente con los datos publicados por ellos. Como puede verse en la tabla 5.9, tanto los métodos basados en el desarrollo de Chebyshev, como el método de Geyer, producen resultados con un error en torno al 5%. Este nivel de error coincide con el aceptado por Geyer y Winter para su algoritmo [61].

Por otro lado, desde el punto de vista del número de términos del desarrollo de Chebyshev requerido por los algoritmos de Kröger y Jendreck cuando se aplican al modelo 3, se encontró que, para el caso de las simulaciones basadas en el método de Kröger, L es muy similar al obtenido para el modelo 1a (gaussiano), encontrando para $N = 160$ un valor de $L = 9$, mientras que $L = 10$ para el modelo gaussiano, como se aprecia en la tabla 5.6.

D_t^* , % Diferencia respecto a Cholesky			
N	Algoritmo		
	Kröger	Jendrejack	Geyer
70	1.1	1.2	0.7
160	3.2	5.4	2.6

R_g^* , % Diferencia respecto a Cholesky			
N	Algoritmo		
	Kröger	Jendrejack	Geyer
70	1.8	1.3	3.9
160	4.0	5.5	6.2

TABLA 5.9: Coeficientes de difusión y radios de giro respecto al valor de referencia (Cholesky) obtenidos mediante cada uno de los algoritmos y el modelo 3.

Sin embargo, para las simulaciones basadas en el método de Jendrejack sí se aprecia cierta diferencia en el valor de L ; si bien L_{medio} disminuye algo (el caso más destacado es en $N = 160$: para el modelo 3, $L_{medio} = 22$ frente a $L_{medio} = 29$ para el modelo 1a), lo más destacable es la diferencia en L_{max} . Mientras para el modelo 1a (gaussiano) con $N = 160$ se tenía $L_{max} = 115$, para el modelo 3 $L_{max} = 33$. Esta diferencia se atribuye a la menor variabilidad conformacional del modelo utilizado por Geyer.

Conclusiones sobre la influencia del solapamiento.

Anteriormente se había hablado del error acumulado en el cálculo de los desarrollos de Chebyshev. Si anteriormente ya se había dado poco crédito a esta causa como responsable de la desviación de los resultados, debido a que los resultados obtenidos a partir de los métodos de Jendrejack y Kröger son iguales, a pesar de las diferencias en los números de términos empleados para los desarrollos, una vez estudiado el caso de las simulaciones con el modelo 3 esta hipótesis se descarta por completo, ya que el número de términos de los desarrollos no varía mucho al utilizar el modelo 1a (que permite solapamiento) o el modelo 3 (que no lo permite), especialmente en el caso del algoritmo de Kröger, y sin embargo, sólo al emplear el modelo 1a se observa la divergencia de los resultados respecto a los obtenidos a partir de las simulaciones realizadas con el algoritmo de Cholesky.

Sin embargo, en base a las pruebas realizadas sí que se puede atribuir la divergencia en los resultados a la frecuencia de aparición de situaciones de solapamiento a lo largo de la simulación. De hecho, estudiando el grado de solapamiento (entendiendo como tal el número de veces que se producen solapamientos a lo largo de la simulación) en las diferentes trayectorias, se observa que este solapamiento es independiente del método de simulación empleado, y depende únicamente del modelo.

Estudiando las trayectorias generadas para los modelos 1a y 1b, se descubre que el número de solapamientos, tal como se sospechaba, aumenta con el tamaño del modelo simulado, existiendo una relación prácticamente lineal entre ambos factores. De hecho se observó que el porcentaje de solapamientos, definido como el número de solapamientos respecto al número de bolas del modelo, se mantenía, para todos los tamaños simulados, cualquiera que fuera el algoritmo empleado, entorno a un valor de un 10%. Este valor, a pesar de ser significativo, no es excesivamente grande, y por lo tanto, no es posible extrapolar directamente las conclusiones de este estudio a otros ámbitos en los que se trabaje con modelos con un solapamiento mayor.

En lo que se refiere al algoritmo de Geyer, en las tablas 5.2 y 5.3 se puede observar que el porcentaje de diferencia respecto al método de Cholesky fluctúa en torno a unos valores bastante inferiores al que presentan los métodos basados en la factorización de Chebyshev, aunque también parece haber un cierto crecimiento con N . A partir de esto concluimos que, aunque parece que este método también se ve afectado por la presencia de solapamientos, la influencia de éstos es menor que en el caso de los procedimientos basados en Chebyshev.

De hecho, creemos que las desviaciones encontradas son más bien debidas al incumplimiento de una de las hipótesis necesarias para el desarrollo de este algoritmo, concretamente, la suposición de un acoplamiento hidrodinámico débil. Para valores lo suficientemente grandes de N , puede aumentar el grado de interacciones de tipo hidrodinámico entre las bolas de la cadena, e invalidar así la hipótesis del acoplamiento hidrodinámico débil, produciendo resultados que se desvíen en mayor medida de los que se emplean como referencia.

En cualquier caso, vista la magnitud de las desviaciones observadas para el rango de tamaños de cadena estudiados, no se espera que, para el rango de tamaños habitualmente empleado en simulaciones de este tipo, las desviaciones que el uso de este algoritmo pueda producir sean significativas.

Para terminar con el estudio comparativo de las propiedades, si se representan en escala doble logarítmica los datos de las tablas 5.2 y 5.3, se obtienen las figuras 5.1 y 5.2. Sobre estos gráficos se observa claramente la desviación existente entre los datos procedentes de las simulaciones basadas en el método de Chebyshev y los datos de referencia obtenidos mediante el método de Cholesky, mientras que los datos procedentes de las simulaciones con el método de Geyer, a pesar de presentar cierta desviación, son muy próximos a los de referencia, especialmente en el caso de D_t^* .

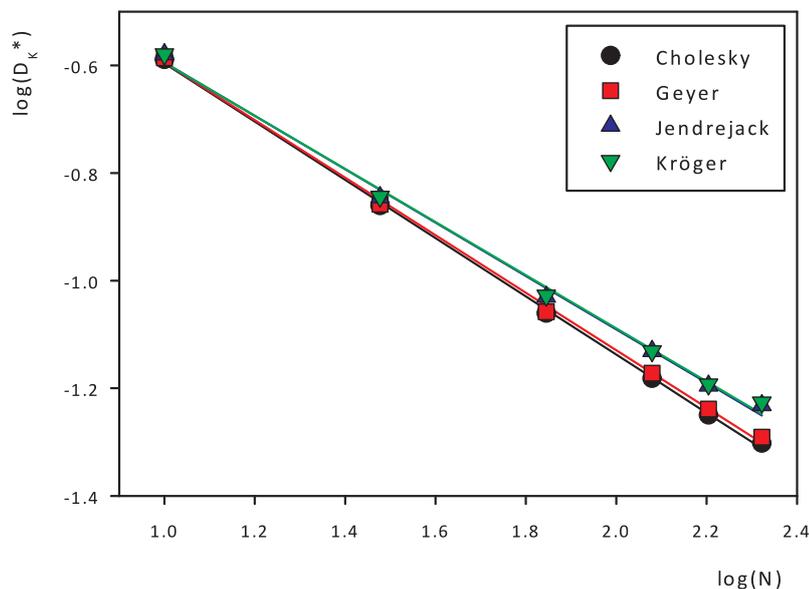


FIGURA 5.1: Representación de la evolución de D_t^* con N en cadenas gaussianas (modelo 1a) de tamaño creciente, empleando los diferentes algoritmos de simulación.

Relacionando los valores de las propiedades estudiadas con el tamaño de cadena pueden calcularse las correspondientes leyes de escala. Los valores de los parámetros de éstas están recogidas en las tablas 5.10 y 5.11. Si las leyes de escala obtenidas se comparan con las conocidas teóricas para un modelo similar a nuestro 1a en condiciones theta [68], en las que $R_g^* \propto N^{0.5}$ y $D_t^* \propto N^{-0.5}$, se observa que existe cierta desviación en las mismas. Esta desviación se debe a que estas relaciones teóricas están enunciadas para un modelo en el que no existe volumen excluido. El uso de un potencial de Lennard-Jones para imitar condiciones theta debería aproximarse a estas condiciones, pues las interacciones de atracción y repulsión deberían anularse en promedio, pero para eso es necesario que las cadenas tengan un tamaño suficiente, condición que quizás no se cumpla en todos

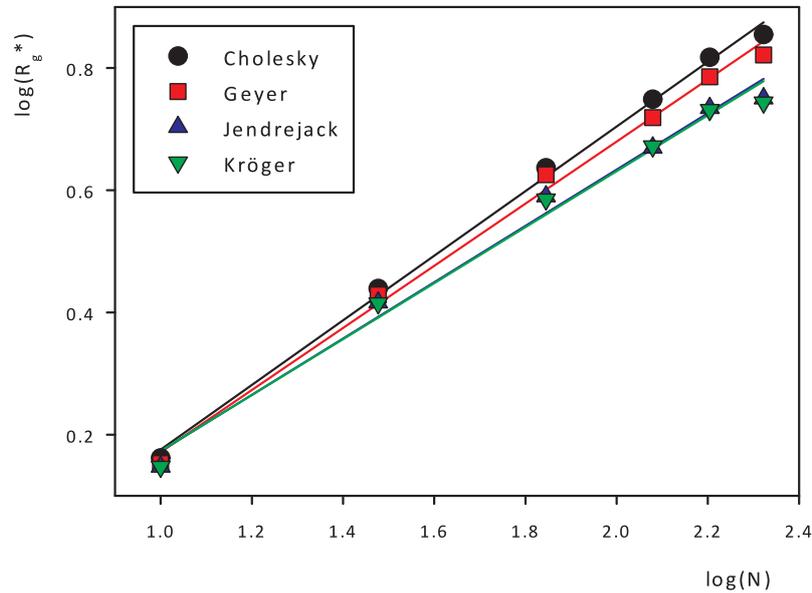


FIGURA 5.2: Representación de la evolución de R_g^* con N en cadenas gaussianas (modelo 1a) de tamaño creciente, empleando los diferentes algoritmos de simulación.

los modelos estudiados en este capítulo, produciéndose por tanto desviaciones como las observadas.

Algoritmo	α_D	$\log(k_D)$
Cholesky	-0.5255	-0.0874
Geyer	-0.5141	-0.1027
Jendrejack	-0.4628	-0.1676
Kröger	-0.4601	-0.1709

TABLA 5.10: Factor de escala de D_t^* para cada uno de los algoritmos ensayados: $D_t = k_D N^{\alpha_D}$

5.4.2. Comparación de velocidad.

Para evaluar la rapidez de los algoritmos se llevaron a cabo simulaciones de una duración de 10^6 pasos mediante cada uno de los algoritmos estudiados, sin emplear ningún tipo de aceleración (no se usó cálculo en paralelo sobre múltiples procesadores, ni tam-

Algoritmo	α_G	$\log(k_G)$
Cholesky	0.4996	-0.2926
Geyer	0.4682	-0.2536
Jendrejack	0.4047	-0.1711
Kröger	0.4016	-0.1678

TABLA 5.11: Factor de escala de R_g^* para cada uno de los algoritmos ensayados: $D_t = k_G N^{\alpha_G}$

poco librerías matemáticas precompiladas). Los tiempos de cálculo de cada una de estas trayectorias se presentan en la tabla 5.12 y se representan en la figura 5.3, tanto a escala doble logarítmica (figura 5.3b) como a escala lineal (figura 5.3a), para dar una mejor idea de la diferencia entre los tiempos empleados en las diferentes simulaciones.

N	Cholesky	Geyer	Jendrejack	Kröger
	t_{CPU} (min)	t_{CPU} (min)	t_{CPU} (min)	t_{CPU} (min)
10	0.55	0.38	1.53	0.40
30	7.39	2.77	11.81	3.14
70	72.77	13.94	93.75	16.28
120	336.59	39.08	387.00	48.40
160	773.78	69.05	975.52	93.06
210	1737.75	123.64	2072.62	169.98

TABLA 5.12: Tiempos de cálculo de una trayectoria de 10^6 pasos brownianos para cadenas gaussianas (modelo 1a) de diferentes tamaños empleando los diferentes algoritmos.

De los datos de la tabla 5.12 se puede observar cómo las simulaciones realizadas mediante el algoritmo de Jendrejack presentan tiempos de simulación incluso superiores a los de las simulaciones que emplean el método de Cholesky. En cambio, las simulaciones que emplean los algoritmos de Geyer y Kröger presentan tiempos de cálculo mucho menores.

Aún más importante que los tiempos de cálculo en sí mismos es establecer cómo escalan éstos con la longitud de cadena: $t_{\text{CPU}} = kN^\alpha$. La relación de parámetros de las leyes de escala obtenidas se recoge en la tabla 5.13. Dichos parámetros se calculan ajustando los datos empleados en la elaboración de la figura 5.3b a la ecuación de una recta. Para el cálculo de estas leyes de escala se han omitido los datos correspondientes a $N = 10$, pues

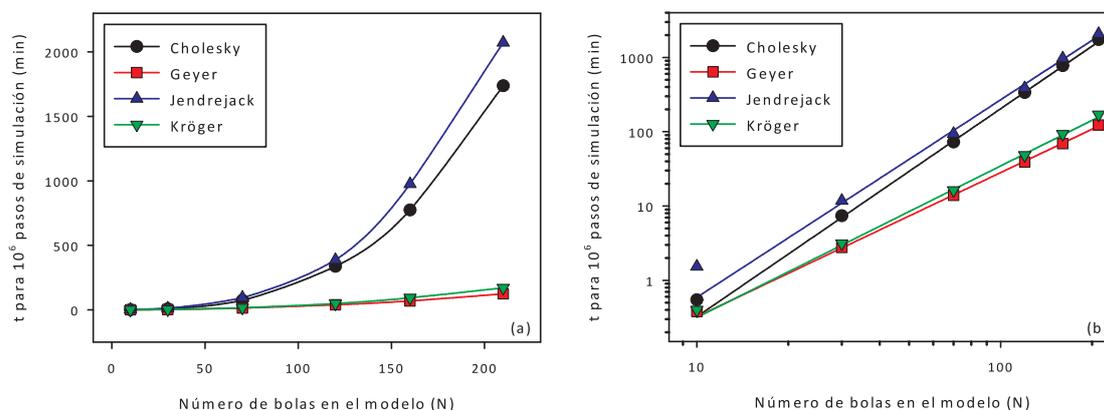


FIGURA 5.3: Representación de los tiempos de cálculo de una trayectoria de 10^6 pasos, usando el modelo 1a, frente al número de bolas N , empleando los diferentes algoritmos: (a) en escala lineal y (b) en escala logarítmica.

en este caso, el número de bolas es pequeño y los datos obtenidos desvirtúan el ajuste de los datos a la ecuación.

Algoritmo	α	$\log(k)$
Cholesky	2.80	-3.29
Geyer	1.94	-2.43
Jendrejack	2.66	-2.89
Kröger	2.04	-2.54

TABLA 5.13: Leyes de escala del tiempo de cálculo para 10^6 pasos de simulación con cada uno de los algoritmos: $t_{\text{CPU}} = kN^\alpha$

Como se mencionaba en la introducción teórica, el algoritmo de Cholesky presenta una ley de escala teórica del tiempo de cálculo proporcional a N^3 . Sin embargo, en nuestros resultados el exponente de la ley de escala es de tan solo 2.8. Esto se debe a que la ley de escala teórica sólo se observa cuando el tamaño de las moléculas simuladas es lo suficientemente grande, y en nuestro caso, para el cálculo de esta ley de escala se han empleado valores relativamente pequeños (desde $N = 30$ hasta $N = 210$). Además, es necesario tener en cuenta que los tiempos de cálculo obtenidos incluyen también el tiempo invertido en el volcado a disco de la trayectoria generada, y el volumen de esta información aumenta de forma importante al hacerlo el tamaño de la cadena.

Para el caso de las simulaciones basadas en el método de Fixman se obtienen resultados dispares. Teóricamente, el exponente de la ley de escala debería ser de $9/4$, pero los valores obtenidos son de 2.66 en el caso del algoritmo de Jendrejack, y 2.04 para el caso de Kröger. Esta diferencia es lógica, puesto que cada uno de los algoritmos se basa en una estrategia diferente para ahorrar cálculos y reducir el tiempo de CPU. Mientras que para el caso de Kröger puede hacerse la misma argumentación que se ha hecho para el método de Cholesky, el mal resultado del método de Jendrejack sólo puede explicarse teniendo en cuenta la conclusión alcanzada en el apartado anterior: la presencia de los solapamientos conduce a aumentar el número de términos necesarios en el desarrollo de Chebyshev para alcanzar la precisión requerida, y esto a su vez redundante en un tiempo de cálculo mucho más alto, más aún si tenemos en cuenta que al elevado valor de L empleado se llega como consecuencia de un elevado número de pasos no válidos realizados con valores de L menores.

Como se observaba en la tabla 5.6, los valores de L_{medio} de las simulaciones de Jendrejack son próximas al triple del L fijo empleado en las simulaciones del método de Kröger. Pero además también hay que tener en cuenta los altos valores alcanzados por L_{max} , que son entre 10 y 15 veces superiores a los valores de L de Kröger, y, aunque son poco frecuentes a lo largo de la trayectoria, implican que en algunos puntos de la simulación se da un alto número de fallos en la evaluación de la precisión del desarrollo, con la consiguiente necesidad de repetir en varias ocasiones el cálculo de los valores propios del tensor de difusión para dar un solo paso, una operación que supone un importante factor de consumo de tiempo de CPU.

Comparando el algoritmo de Jendrejack y el de Cholesky, se observa que, a pesar de que el exponente de la ley de escala del primero es inferior, los tiempos de cálculo obtenidos son superiores. Esto se debe a que el factor preexponencial obtenido para el método de Jendrejack es casi un orden de magnitud mayor que el obtenido para Cholesky, y aunque el exponente sea menor, la “ventaja” en los tiempos de cálculo no se manifiesta en los tamaños de cadena estudiados. A partir de las leyes de escala obtenidas puede estimarse el punto a partir del cual la “ventaja” en tiempo del método de Jendrejack sobre el de Cholesky comenzaría a observarse. Para ello, basta igualar ambas leyes y obtener el valor de N para el que ambas predican el mismo tiempo de CPU. Este valor es próximo a $N \approx 720$, lo que supone un valor muy por encima de los valores habituales en las aplicaciones prácticas de las simulaciones de este tipo.

Finalmente, en cuanto al algoritmo de Geyer, se puede observar que los tiempos de cálculo requeridos por este método son los menores que se han obtenido en todos los

casos. Además, se observa que el factor de escala es de 1.94, prácticamente coincidente con la observación realizada por Geyer y Winter en sistemas sin solapamiento, para los que había obtenido un crecimiento del tiempo de cálculo proporcional a N^2 . Es decir, el grado de solapamiento del modelo simulado parece no tener apenas influencia sobre la rapidez del método, al contrario que en el método de Jendrejack.

5.5. Conclusiones del Capítulo.

Se han comparado diferentes algoritmos comúnmente usados para el tratamiento de la interacción hidrodinámica en simulaciones de dinámica browniana. Estos algoritmos están basados en diferentes aproximaciones matemáticas al cálculo de la raíz cuadrada del tensor de difusión: la descomposición de Cholesky, el desarrollo en serie de Chebyshev (con dos consideraciones diferentes para reducir el número de cálculos) y el algoritmo TEA de Geyer y Winter. Mediante implementaciones de los diferentes algoritmos se han realizado simulaciones en equilibrio de cuatro tipos de modelos de bolas y muelles, con diferentes grados de flexibilidad y diferentes tipos de interacciones de volumen excluido, y hasta un tamaño de 210 bolas. A partir de estas simulaciones se han calculado las propiedades radio de giro y coeficiente de difusión traslacional (ambos en unidades reducidas) y las correspondientes leyes de escala.

La comparación de propiedades y las leyes de escala obtenidas, tomando como referencia los resultados proporcionados por las simulaciones basadas en la factorización de Cholesky, conducen a la conclusión de que el método más ventajoso es el propuesto por Geyer y Winter, pues el que produce los resultados con la mejor aproximación, tanto en modelos que permiten solapamiento como en los que no lo permiten.

En cambio, los algoritmos basados en el desarrollo de Chebyshev solamente producen buenos resultados en los casos en los que se evita el solapamiento entre los elementos del modelo, ya sea mediante la aplicación de fuerzas externas que lo impidan, o utilizando moléculas con un cierto grado de rigidez, suficiente como para que no se produzca el solapamiento, tal como suele suceder en la simulación de macromoléculas biológicas.

En cuanto a la eficiencia de los algoritmos, también en este caso el algoritmo TEA de Geyer y Winter es la mejor opción: es el algoritmo más rápido, y también el que consume

menos recursos computacionales, pues evita el uso de operaciones entre grandes matrices y vectores. Sin embargo, el algoritmo de Kröger también es bastante competitivo en este aspecto, aunque su aplicabilidad está limitada a casos en los que no haya solapamiento, y cuenta con la desventaja adicional de que es necesario encontrar unas cotas superior e inferior de los valores propios de los tensores de difusión correspondientes a las sucesivas conformaciones instantáneas generadas a lo largo de la simulación.

En conclusión, el algoritmo de Geyer y Winter supone una alternativa muy atractiva al algoritmo clásico de Ermak y McCammon, pues tanto en la precisión de los resultados como en lo que se refiere a su eficiencia se comporta de forma adecuada. Por el contrario, el algoritmo de Chebyshev, empleado habitualmente hasta ahora en la simulación de grandes sistemas macromoleculares parece ser útil únicamente en el caso de que apenas se produzcan solapamientos.

Sin embargo, es necesario destacar que estos resultados se han obtenido para cadenas lineales flexibles, en los que el grado de solapamiento es pequeño (en torno al 10 %, como se mencionaba anteriormente). De esta manera, estas conclusiones no tienen porqué ser necesariamente extrapolables al uso de otros modelos más sofisticados, como los usados para representar macromoléculas estructuradas como es el caso de muchas proteínas, en las que existe un grado importante de solapamiento. Para este tipo de modelos con solapamientos “biológicamente realistas” sería necesario realizar un estudio más amplio y concienzudo sobre el rendimiento de estos algoritmos.

Finalmente, destacar que los algoritmos basados en la factorización de Cholesky, que posee la ventaja de proporcionar una solución exacta a la descomposición del tensor de difusión, pueden acelerarse mediante el uso de ciertas estrategias, como la conservación del tensor de difusión durante varios pasos de simulación, y especialmente, mediante el uso de técnicas computacionales avanzadas como el uso del cálculo paralelizado sobre múltiples procesadores o la computación distribuida entre múltiples equipos.

CAPÍTULO 6

SIMULACIÓN MULTIESCALA DE POLÍMEROS HIPERRAMIFICADOS

6.1. Introducción.

La determinación de propiedades conformacionales e hidrodinámicas de polímeros en disolución diluida es de gran interés pues están íntimamente relacionadas con la estructura y, por extensión, con la funcionalidad de estas moléculas. Por tanto, el estudio de las propiedades en disolución diluida de polímeros hiperramificados, un tipo de polímeros relacionados con las moléculas dendríticas pero con geometría no regular, es de gran utilidad para comprender su comportamiento y desarrollar potenciales aplicaciones.

Este tipo de compuestos se ha venido sintetizando y estudiando desde finales de los años 80, y con bastante intensidad en los últimos diez años. Algunos de los esfuerzos experimentales realizados han estado dirigidos a determinar propiedades de estos polímeros en disolución diluida. Una de las familias más sintetizadas y caracterizadas han sido los poliésteres [70–74] y, en menor medida polifenilenos [75], poliamidoaminas [76] y carbosilanos [77].

El estudio y caracterización de los polímeros hiperramificados es más complejo que el de los dendrímeros (no así su síntesis) debido a la variabilidad que presentan tanto en el grado de polimerización como en el grado de ramificación. Dicho de otro modo, durante la síntesis de estos polímeros se da lugar a una gran variedad de pesos moleculares y de

topologías, lo que influye en la variedad de propiedades y comportamientos encontrados. Así, por ejemplo, Hobson y Feast [76] encuentran que, a semejanza de los dendrímeros, los polímeros hiperramificados presentan un máximo en la variación de la viscosidad intrínseca $[\eta]$ con el grado de polimerización. En cambio, en otros trabajos [73, 74] se encuentra un incremento monótono de $[\eta]$ con N como ocurre en polímeros lineales a los que se les puede aplicar la relación de Mark-Houwink.

Por lo expuesto anteriormente, en el caso de estos polímeros, la simulación computacional se presenta como una herramienta muy útil para esclarecer la relación entre estructura y propiedades. Hasta ahora se han realizado simulaciones de polímeros hiperramificados en disolución diluida a partir de las cuales se han obtenido fundamentalmente radios de giro R_g (una propiedad conformacional) y viscosidades intrínsecas $[\eta]$ (una propiedad hidrodinámica) y se han estudiado las leyes de potencia que relacionan estas propiedades, tanto con el grado de polimerización, representado por el número de elementos del modelo, como con el grado de ramificación, representado principalmente por medio de los índices topológicos “índice de Wiener”, W , y “grado de ramificación” o “*degree of branching*”, DB . Para ello se han utilizado modelos de bolas y varillas (conectores rígidos), y técnicas de simulación tanto de tipo Monte Carlo como de dinámica browniana.

En los trabajos en los que se usa la técnica de Monte Carlo no suele incluirse el efecto de volumen excluido, y la viscosidad intrínseca se calcula a partir del radio de giro, empleando el modelo de esfera equivalente [78, 79]. Otros autores han utilizado la técnica de dinámica browniana incluyendo los efectos de volumen excluido e interacción hidrodinámica, y han calculado la viscosidad intrínseca a partir de la extrapolación a flujo cero de los resultados de simulaciones de moléculas hiperramificadas sometidas a flujo de cizalla [80, 81]. Recientemente, los mismos autores han publicado trabajos de dinámica browniana en la línea de los anteriores en los que se incluye la presencia de cargas [82, 83]. Los resultados obtenidos mediante estas simulaciones han servido para evaluar la influencia que tienen por separado la ramificación y el grado de polimerización en las propiedades estudiadas. Sin embargo, en estos trabajos no se utilizan unidades físicas y sólo se ha evaluado la concordancia de los resultados simulados con ciertas predicciones teóricas procedentes de teorías de campo medio y con resultados procedentes de otras simulaciones, pero no se comparan explícitamente con datos experimentales.

En nuestro trabajo empleamos una metodología algo más sofisticada, ya empleada con éxito con anterioridad en la simulación de macromoléculas dendríticas [24], para simular el comportamiento de macromoléculas hiperramificadas en disolución diluida, obtener diversas propiedades conformacionales e hidrodinámicas (radio de giro, R_g , coeficiente de

difusión, D_t , viscosidad intrínseca, $[\eta]$) y comparar los resultados obtenidos (inicialmente en forma de leyes de potencia) con datos experimentales.

Para ello se han empleado modelos de bolas y muelles “hard-FENE” representando polímeros hiperramificados reales. Para asignar valores a los parámetros de esos modelos se utilizó un procedimiento multiescala basado en simular a nivel atómico las pequeñas porciones de la macromolécula que se identifican con los elementos del modelo, tal como se describió en la sección 2.3 del Capítulo 2. Una vez construido el modelo, se utilizó la técnica de simulación de Monte Carlo, considerando el efecto del volumen excluido, para generar una muestra estadísticamente representativa de moléculas con diferente topología pero igual grado de polimerización (peso molecular o número de elementos en el modelo, N) a partir de la cual se obtienen las distintas propiedades estacionarias en disolución diluida como un doble promedio sobre distintas conformaciones y sobre distintas topologías, manteniendo un mismo peso molecular a través del uso de un número de bolas del modelo, N , fijo. Realizando simulaciones análogas y variando N se puede estudiar la dependencia con el peso molecular. Las propiedades hidrodinámicas se calcularon mediante un procedimiento más riguroso que el utilizado en trabajos anteriores, el tratamiento hidrodinámico del cuerpo rígido con la inclusión de determinadas correcciones que ya se han mostrado eficaces a la hora de obtener viscosidades intrínsecas de dendrímeros [84].

6.2. Base Teórica.

6.2.1. Índices topológicos.

Consideramos en este apartado una estructura ramificada compuesta por N nodos (los círculos en la figura 6.1, presentada más adelante), unidos entre sí por conectores (los segmentos en la misma figura). El número de conectores, para una estructura que no contenga ciclos, debe ser $N - 1$, pues, como se puede verificar por construcción, añadir un nodo siempre requiere un nuevo conector.

Al dibujar una de estas estructuras se puede considerar que existe una direccionalidad en el sentido de que un nuevo nodo se puede considerar “hijo” de aquel al que se une. Y así es en el proceso químico real en el que se sintetizan estos polímeros: al crecer una macro-

molécula hiperramificada, cada nueva subunidad se conecta a una previamente integrada en la macromolécula, siendo “hija” de ésta. Con este punto de vista, los conectores serían vectores dirigidos hacia la subunidad hija. Pero desde el punto de vista de la estructura final resultante, el aspecto principal de la topología es una conectividad no direccional que indica que pares de nodos son vecinos, mediante un conector entre ellos.

Consideraremos estructuras compuestas por nodos de conectividad máxima (funcionalidad) F , los cuales pueden estar unidos a un cierto número de nodos, que será:

- 1, tratándose de nodos terminales, que señalaremos como “ T ”
- 2, en los que denominaremos nodos lineales, identificados como “ L ”
- f , entre 3 y F , en los que llamaremos nodos ramificados o dendríticos, D_f

Para polímeros resultantes de monómeros tipo AB_2 , los nodos serán trifuncionales ($F = 3$), y los dendríticos estarán unidos a otros tres nodos, cabiendo solamente la posibilidad $f = F = 3$. Para $F = 4$ y superiores cabe la posibilidad de encontrar nodos ramificados parcial o totalmente, unidos a otros $f = 3 \dots F$ nodos.

Para una estructura ramificada arbitraria se pueden definir diversos índices topológicos que proporcionan información sobre su grado de ramificación. En general, la relación entre índice topológico y estructura no es unívoca, es decir, varias estructuras isómeras pueden tener el mismo valor de un determinado índice topológico. Esta característica se conoce como degeneración, y es variable de un índice a otro, siendo mayor cuanto mayor es el número de estructuras que pueden ser representadas por un mismo valor del índice topológico.

A continuación se describen tres posibles índices topológicos, de los cuales los dos primeros son ampliamente utilizados en el estudio de polímeros hiperramificados.

Grado de ramificación.

Es el índice más usado en la bibliografía, si bien hay ciertas, ligeras, diferencias en cómo lo definen unos y otros autores (ver por ejemplo [72, 80]). En inglés se conoce como *Degree of Branching*, por lo que lo representaremos por DB .

Se trata de un índice acotado entre 0 y 1, valores que corresponden, respectivamente, a los casos extremos de moléculas completamente lineales, y moléculas donde todos los nodos están totalmente ramificados excepto, obviamente, los terminales. Su expresión más general, diseñada de manera que valga 0 para una estructura lineal y 1 para la de máxima ramificación, según la formulan Hölder y col. [85], es el cociente entre el parámetro de ramificación R , definido como el número de ramificaciones presentes en la estructura, y su valor máximo R_{max} para una estructura del mismo número de nodos:

$$DB = \frac{R}{R_{max}}. \quad (6.1)$$

Los nodos lineales (y obviamente los terminales), no aportan ramificaciones, y un nodo de funcionalidad f aporta $f - 2$ ramificaciones (una para $f = 3$, dos para $f = 4$, etc). Esto se puede expresar como

$$R = \sum_{i=1}^N (f_i - 2) \quad \forall f_i \geq 2, \quad (6.2)$$

siendo i el índice de los nodos.

En la figura 6.1 se presentan varios ejemplos para una estructura de $N = 15$ nodos de $F = 5$. En esta figura,

- (A) presenta dos nodos de $f = 3$, luego $R = 2$;
- (B) tiene un nodo de $f = 4$ y dos de $f = 3$, siendo $R = 4$;
- (C) tiene uno de $f = 5$ y tres de $f = 3$, con $R = 6$;
- (D) tiene tres de $f = 5$ con $R = 9$;
- (E) tiene dos de $f = 5$, uno de $f = 4$ y otro de $f = 3$, con $R = 9$.

La ramificación máxima posible para esta funcionalidad y este número de nodos es la de (D) y (E) con $R = R_{max} = 9$, de manera que los valores de DB son $2/9$, $4/9$, $6/9$, 1 y 1, respectivamente.

La descripción topológica se hace complicada cuando los nodos tienen $F > 3$, pues puede haber ramificaciones parciales o totales. Afortunadamente, muchos polímeros ramificados, como veremos, están constituidos por subunidades idénticas tipo AB_2 con $F = 3$, y

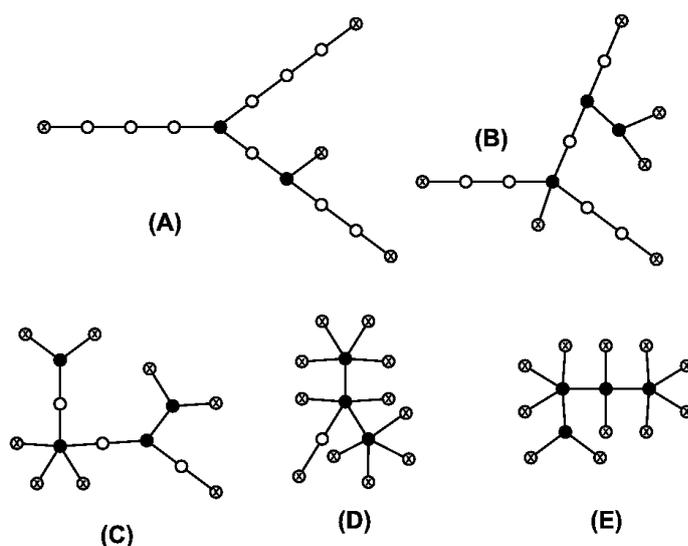


FIGURA 6.1: Ejemplos de topologías de moléculas hiperramificadas con $F = 5$ y $N = 15$. Los nodos lineales, ramificados y terminales se distinguen por puntos blancos, negros y con aspa respectivamente.

por lo tanto solo hay un tipo de nodos (totalmente) ramificados, de $f = F = 3$, pudiéndose hablar de un número de tales nodos D , sin distinción entre ellos.

Continuamos la descripción para este caso concreto de $F = 3$. Si T , L y D indican los nodos terminales, lineales, y ramificados, evidentemente

$$N = D + L + T. \quad (6.3)$$

Cada nodo ramificado aporta a R una unidad, luego $R = D$. R_{max} será entonces igual al valor D_{max} que puede haber en una estructura de N nodos que será aquel que maximice el número de nodos ramificados. Para $N = 1, 2, 3, 4, 5, 6, 7, 8, 9, \dots$ se ve por simple inspección (para cualquier estructura que no tenga ningún nodo L) que $D_{max} = 0, 0, 0, 1, 1, 2, 2, 3, 3, \dots$ y en general

$$D_{max} = \text{int}[(N - 2)/2], \quad (6.4)$$

donde $\text{int}()$ indica la división entera, de manera que desarrollando la ecuación 6.1,

$$DB = \frac{D}{\text{int}[(N - 2)/2]}. \quad (6.5)$$

Para N grande, $D_{max} \approx N/2$, resultando

$$DB \approx \frac{2D}{N} = \frac{2D}{D+L+T}. \quad (6.6)$$

Cabe otro argumento para determinar D_{max} (aportado por Hölder y col. [85]; el previo es original nuestro). Una estructura con $N = D+L+T$, se puede modificar para maximizar su ramificación reconvirtiendo cada pareja de nodos L en un nodo T y un nodo D . El valor de D en la estructura resultante, que no tendría ningún nodo L (si L fuese par) o un solo nodo L (si L fuese impar), resultaría ser $D_{max} = D + \text{int}(L/2) \approx D + L/2$, de manera que

$$DB \approx \frac{D}{D+L/2} = \frac{2D}{2D+L}. \quad (6.7)$$

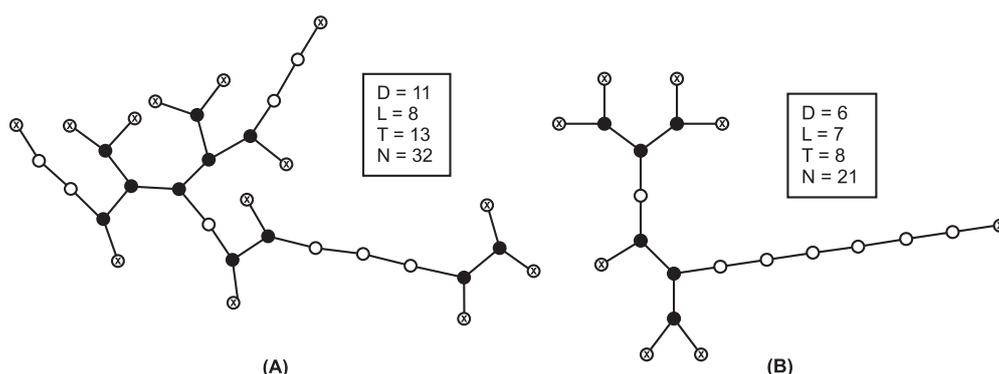


FIGURA 6.2: Topologías de moléculas hiperramificadas con $F = 3$ para el cálculo de DB.

Aplicando las expresiones anteriores a las dos estructuras de la figura 6.2, en ambos casos se encuentran resultados coincidentes: utilizando la ecuación 6.5, $DB_A = 22/30$ y $DB_B = 6/9$, mientras que utilizando la ecuación 6.6, $DB_A = 22/30$ y $DB_B = 12/19$. Como se puede observar, en el caso de la estructura “A”, ambas ecuaciones dan idéntico resultado, mientras que para la estructura “B” son prácticamente coincidentes ($6/9 \approx 12/19$). Las fracciones obtenidas por un procedimiento u otro pueden diferir por una unidad en el denominador (cuando L es impar), lo cual carece de importancia para N suficientemente elevado.

Siendo correctas las expresiones enunciadas, debe cumplirse (aunque no lo hemos demostrado directamente) que $2D + L \approx D + L + T$, lo cual implica $D \approx T$.

Para una macromolécula lineal, $D = 0$ de manera que $DB = 0$. Por otra parte, para una molécula dendrítica totalmente ramificada $L = 0$, y $DB = 1$ en la ecuación 6.6.

La expresión

$$DB = \frac{D + T}{N} = \frac{D + T}{D + L + T}, \quad (6.8)$$

que fue propuesta en los trabajos iniciales (por ejemplo, en [72]), fue desaconsejada por Hölter y col. [85], y ha caído en desuso. Dicha ecuación produce valores algo desviados respecto a las ecuaciones 6.5 y 6.6. Siguiendo el ejemplo anterior, se obtendrían los valores $DB_A = 24/32$ y $DB_B = 14/21$, respectivamente, para las estructuras (A) y (B) de la figura 6.2.

En cualquier caso, a pesar de ser éste un parámetro bastante atractivo por su sencilla definición y su accesibilidad experimental (es muy sencillo de determinar empleando RMN [72]), es un índice bastante degenerado y no discrimina demasiado la geometría de la estructura hiperramificada. Así, por ejemplo, un $DB = 1$ puede estar asociado tanto a una estructura dendrítica perfectamente regular, como a una estructura hiperramificada de crecimiento irregular.

Índice de Wiener.

Este índice, W , aunque fue propuesto por Wiener para estudiar las relaciones entre puntos de ebullición de parafinas [86], es otro índice topológico ampliamente utilizado en el estudio de los polímeros ramificados. En términos sencillos, podría decirse que da una idea del grado de linealidad de la molécula: cuanto mayor es el valor de este parámetro más lineal es la molécula.

Su expresión más general es la siguiente:

$$W = \frac{1}{2} \sum_{j=1}^N \sum_{i=1}^N d_{ij}, \quad (6.9)$$

siendo d_{ij} la distancia en conectores que separa los nodos i y j . Nótese que en ausencia de ciclos, solo hay un camino posible entre i y j , y d_{ij} está unívocamente definido.

Siempre para el caso de moléculas en las que no existan ciclos, la expresión anterior puede convertirse en otra más sencilla y computacionalmente más eficiente [81]:

$$W = \sum_{s=1}^{N-1} (V_{L,s} \cdot V_{R,s}). \quad (6.10)$$

Esta ecuación consiste en la suma, para todos los conectores (que serán $N - 1$), del producto del número de nodos “a la izquierda” (“*left*”) y “a la derecha” (“*right*”) del conector o segmento s ($V_{L,s}$ y $V_{R,s}$, respectivamente).

Para el caso de la molécula lineal, que presenta el máximo índice de Wiener para un tamaño dado, se encuentra fácilmente que [79]

$$W_{max} = N \frac{(N-1)N}{2} - \frac{(N-1)N[2(N-1)+1]}{6} \approx \frac{N^3}{6}, \quad (6.11)$$

siendo el resultado aproximado válido para N grande.

Como se desprende de la ecuación 6.11, el valor de W no está acotado, como ocurría con DB , sino que su magnitud depende de forma muy importante del tamaño de la estructura, N . Sin embargo, para un N fijo, el índice de Wiener está más relacionado con la geometría de la molécula (está menos degenerado) que DB .

Fracción de nodos terminales.

Ciertas propiedades del polímero hiperramificado pueden depender esencialmente de la fracción de nodos terminales, T/N . Sobre la base que $T \approx D$, de la ec. 6.6 resultaría $T/N = DB/2$, presentando un valor máximo de $1/2$.

6.3. Correspondencia entre estructura y topología.

Los polímeros producidos a partir de monómeros tipo AB_2 se corresponderán con topologías de nodos con $F = 3$. En la reacción química de polimerización, la incorporación de una nueva unidad monomérica se realiza por unión de su grupo A a un grupo B libre en una macromolécula en crecimiento, lo cual añade dos nuevos grupos B libres al esqueleto de la macromolécula.

En la figura 6.3 se muestra un esquema de una molécula compuesta por monómeros de tipo AB_2 que responde a la estructura de la figura 6.2-B. En la figura se muestra una numeración que correspondería con el orden de incorporación de los distintos monómeros a la molécula durante su crecimiento, así como la conectividad entre los grupos A (segmentos azules) y B (segmentos rojos) que resultaría. Sin embargo, desde el punto de vista del estudio de la molécula, dicha numeración es irrelevante, y tampoco es relevante que la conectividad se produzca entre A y B. Basta con especificar para lo que sigue que habrá nodos con dos, uno o ningún conector disponible (T , L y D , respectivamente).

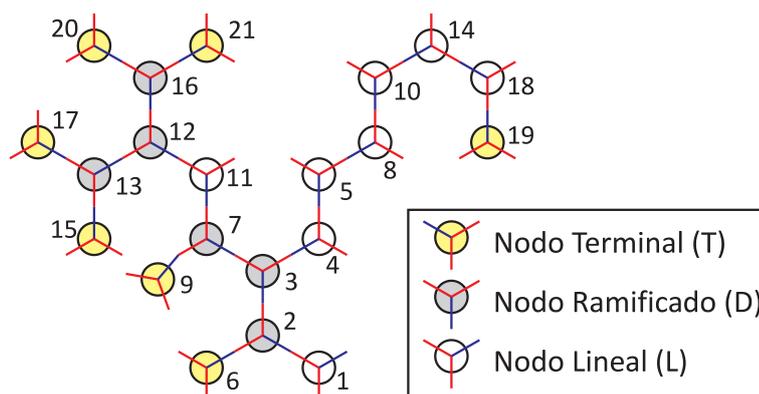


FIGURA 6.3: Modelo de una macromolécula correspondiente a la topología de la figura 6.2-B.

6.4. Algoritmo generador de topología.

Describimos a continuación un algoritmo que genera posibles topologías de una macromolécula ramificada con nodos trifuncionales, $F = 3$, como corresponde a monómeros tipo AB_2 .

La estructura se construye por pasos, en cada uno de los cuales se añade un nodo, conectándolo al azar a uno de los nodos previos que tenga uno o dos conectores vacantes.

Para regular el grado de ramificación de la estructura resultante, vamos a utilizar el parámetro p , que puede tomar valores entre cero e infinito, tal que la probabilidad de que el nuevo nodo se una a un conector de un nodo que tiene los dos conectores vacantes sea 1, mientras que la probabilidad de que se una al único conector vacante de un nodo que tiene el otro ocupado será p . Valores de p muy pequeños, o muy grandes, producirán, respectivamente, estructuras muy lineales o muy ramificadas. Es importante destacar que si bien a p puede asignársele un valor de cero (lo que generaría estructuras exclusivamente lineales), no existe una cota superior que garantice la ramificación en el cien por cien de los casos.

La variable $\text{node}(i, k)$, $k=1, 2$, indicará el estado de ocupación de los dos conectores de cada nodo i de la estructura (los adicionales a aquél por el que se unió a ella al ser generado). Si el conector k está ocupado por el nodo n , $\text{node}(i, k)=n$, y si está aún vacante, $\text{node}(i, k)=0$. Además, por argumentos derivados de la síntesis química se considera la posibilidad de que el nodo inicial tenga un solo conector, y en tal caso, hacemos $\text{node}(1, 2)=-1$. Sin embargo, esta es sólo una posibilidad, y el segundo conector del nodo inicial también podría ser funcional, y podría producirse el crecimiento de la molécula también en esa dirección. El resultado será una lista con registros conteniendo i , $\text{node}(i, 1)$ y $\text{node}(i, 2)$ que especifica completamente la topología de la estructura.

Por ejemplo, la siguiente lista correspondería a la topología de la figura 6.3:

A partir de una lista como ésta, se puede obtener una relación de la conectividad, esto es, pares de nodos entre los que hay un conector. En el ejemplo serían (1,2), (2,3), (2,6), (3,4), (3,7), (4,5), (7,9), etc. También es inmediato clasificar los nodos en terminales ($\text{node}(i, 1)=0$, $\text{node}(i, 2)=0$) y también ($\text{node}(1, 1)=2$, $\text{node}(1, 2)=-1$), si el nodo

i	node(i,1)	node(i,2)	i	node(i,1)	node(i,2)
1	2	-1	12	13	16
2	3	6	13	15	17
3	4	7	14	18	0
4	5	0	15	0	0
5	8	0	16	20	21
6	0	0	17	0	0
7	9	11	18	19	0
8	10	0	19	0	0
9	0	0	20	0	0
10	14	0	21	0	0
11	12	0			

TABLA 6.1: Lista de conexiones correspondiente a la figura 6.3.

inicial sólo admite un conector, como en el caso del ejemplo), lineales ((node(i,1)=x , node(i,2)=0)), y ramificados ((node(i,1)=x, node(i,2)=y), salvo en el caso de que $i = 1$, puesto que el nodo inicial es lineal cuando tiene ocupados sus dos conectores. En los casos anteriores, los valores x e y corresponden a los índices de los nodos a los que está unido el nodo i . En el ejemplo, son terminales 1, 6, 9, 15, 17, 19, 20 y 21 ($T = 8$); lineales 4, 5, 8, 10, 11, 14 y 18 ($L = 7$), y ramificados 2, 3, 7, 12, 13 y 16 ($D = 6$).

Esquemáticamente, el algoritmo de generación de la cadena hiperramificada es como sigue:

- **Iniciación:**

Caso (a): El nodo inicial 1 sólo admite un conector, enganchado necesariamente al nodo 2, y éste, lo está, al menos, al nodo 3:

node(1,1)=2; node (1,2)=-1;

node(2,1)=3; node (2,2)=0;

node(3,1)=0; node (3,2)=0;

inode=3,

Caso (b): El nodo inicial 1 puede estar conectado a uno o dos nodos - lo estará al menos al 2:

node(1,1)=2; node (1,2)=0;

node(2,1)=0; node (2,2)=0;

inode=2

En ambos casos, la variable `inode` es un contador sobre el índice de nodos que indica cuál es el último nodo añadido a la estructura. Nótese que durante la generación los conectores a los que se asigna el valor cero son los posibles puntos de crecimiento y/o ramificación.

- Propagación, hasta completar número de nodos:
 1. Generar la lista de conectores vacantes (aquellos que tengan el valor `node(i,k)=0`), identificando el nodo al que pertenecen, así como su probabilidad frente al crecimiento, 1 o p .
 2. Sortear el conector sobre el que se añadirá el nuevo nodo. El sorteo se hace de la manera convencional para seleccionar aleatoriamente un evento (uno de los `ncon` conectores vacantes en ese momento) de acuerdo con sus probabilidades: se construye una serie de probabilidades acumuladas, `probcum(icon)`, `icon=1, ...ncon`, siendo `probcum(icon)` la suma de las probabilidades 1 o p de los conectores $1, \dots, icon$. En esta serie, el conector k ($k=1,2$) del nodo j tiene el índice `kcon`, y probabilidad `probcum(kcon)`. Entonces, se genera un número aleatorio uniforme, a , entre 0 y `probcum(ncon)`, y se selecciona el conector `kcon` que cumpla `prob(kcon-1) < a < prob(kcon)`. Por supuesto, en el caso de iniciación de tipo (a), el segundo conector, inhábil, del nodo 1 no entra en el sorteo.
 3. Actualizar los datos con el nuevo nodo, de forma que `node(j,k)=inode`, y dado que el nuevo nodo, al menos por el momento es T , `node(inode,1)=0`; `node(inode,2)=0`;
 4. Hacer (`inode=inode+1`), y pasar al siguiente nodo, o terminar la generación.

6.5. Índices topológicos de las estructuras generadas.

Es preciso establecer una correspondencia entre el parámetro de probabilidad de ramificación, p , y los índices topológicos DB y W de las estructuras resultantes de la generación de un hiperramificado en tales condiciones. Para un cierto valor de p son posibles muy diversas estructuras de un N determinado. Por ello, los índices deben considerarse como

p	N = 50		N = 500	
	$\langle DB \rangle \pm \sigma_{DB}$	$(\langle W \rangle \pm \sigma_W) \times 10^{-3}$	$\langle DB \rangle \pm \sigma_{DB}$	$(\langle W \rangle \pm \sigma_W) \times 10^{-6}$
0.001	0.02 ± 0.03	20.00 ± 1.00	0.04 ± 0.01	10.00 ± 1.00
0.003	0.04 ± 0.03	19.00 ± 2.00	0.06 ± 0.01	7.90 ± 0.90
0.01	0.10 ± 0.04	16.00 ± 2.00	0.12 ± 0.01	5.40 ± 0.50
0.03	0.18 ± 0.05	14.00 ± 1.00	0.20 ± 0.02	3.90 ± 0.30
0.1	0.31 ± 0.06	12.00 ± 1.00	0.33 ± 0.02	2.90 ± 0.20
0.3	0.48 ± 0.06	9.90 ± 0.80	0.48 ± 0.02	2.30 ± 0.10
1	0.64 ± 0.06	8.70 ± 0.50	0.66 ± 0.02	1.91 ± 0.07
3	0.78 ± 0.05	8.10 ± 0.30	0.81 ± 0.02	1.79 ± 0.05
10	0.89 ± 0.04	8.00 ± 0.30	0.92 ± 0.01	1.78 ± 0.05
30	0.94 ± 0.03	8.00 ± 0.30	0.97 ± 0.01	1.80 ± 0.06
100	0.96 ± 0.02	8.00 ± 0.40	0.99 ± 0.01	1.81 ± 0.06
300	0.97 ± 0.01	8.00 ± 0.40	0.99 ± 0.01	1.82 ± 0.06
1000	0.98 ± 0.01	8.00 ± 0.30	1.00 ± 0.01	1.82 ± 0.06

TABLA 6.2: Índices topológicos promedio $\langle DB \rangle$, $\langle W \rangle$ para dos tamaños de moléculas, $N = 50$ y $N = 500$.

valores medios, $\langle DB \rangle$, $\langle W \rangle$, etc, calculados sobre un buen número de topologías generadas con nuestro algoritmo. El barrido de p incluye desde valores próximos a cero (que favorecen la linealidad), hasta valores mucho mayores que la unidad (que favorecen la ramificación total). En la tabla 6.2 presentamos resultados para $\langle DB \rangle$, $\langle W \rangle$, etc (incluyendo su desviación estándar) en función de p para dos N , uno pequeño $N = 50$ y uno grande, $N = 500$. Las estadísticas están calculadas sobre una muestra de 1000 moléculas para cada caso.

Los datos, que aparecen representados en la figura 6.4, se ajustan relativamente bien a la ecuación 6.12 en el intervalo $0.007 \leq p \leq 10^4$, de forma que en este intervalo el error relativo en el ajuste, ϵ_R , cumple que $\epsilon_R \leq 10\%$:

$$DB = \frac{p^2 + 0.35p}{p^2 + 1.06p + 0.02} \quad (6.12)$$

También es útil obtener la ecuación 6.13, que nos permite obtener el valor aproximado de p que ha de emplearse para generar moléculas de un determinado DB .

$$p = \frac{0.35 - 1.06DB - \sqrt{0.98DB^2 - 0.66DB + 0.12}}{2(DB - 1)} \quad (6.13)$$

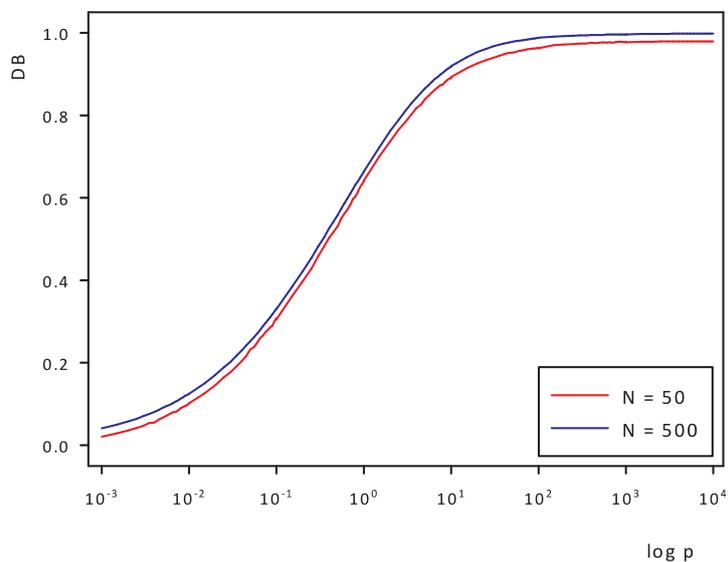


FIGURA 6.4: Representación de DB de hiperramificados con $N = 50$ y $N = 500$ frente a $\log p$.

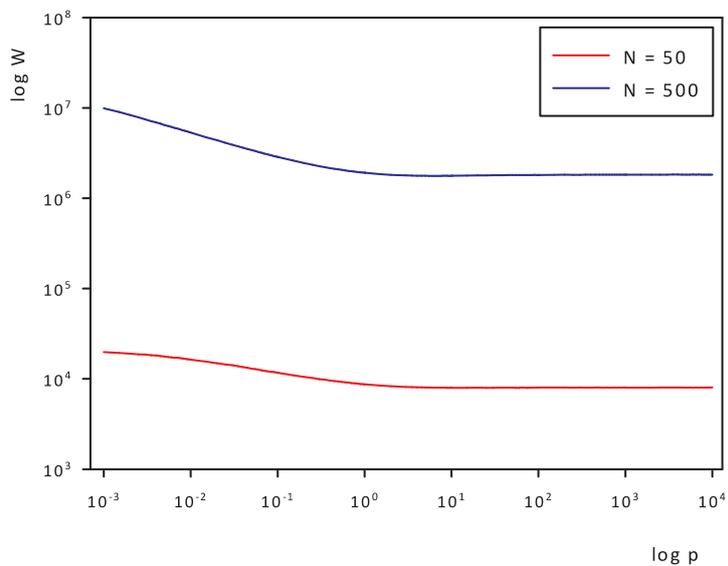


FIGURA 6.5: Representación de $\log W$ de hiperramificados con $N = 50$ y $N = 500$ frente a $\log p$.

Debido a la naturaleza de la ecuación, el error relativo de p calculado respecto a los datos de la simulaciones sólo es inferior al 10 % cuando $0.4 \leq DB \leq 0.75$ (correspondiente a $0.2 \leq p \leq 2$) por lo que los valores de p obtenidos de esta manera sólo son válidos, de manera orientativa, para la generación de topologías con DB similares a los que se emplean para calcular p . Por tanto, cualquier clase de conclusión obtenida a partir de los valores devueltos por esta ecuación no será excesivamente fiable, aunque sí produce resultados razonables si lo que se pretende es hacer una conversión de tipo $DB_{experimental} \rightarrow p \rightarrow DB_{generado}$.

6.5.1. Influencia de N en DB y W .

La tabla 6.3 presenta los valores promedio y sus desviaciones estándar de DB y W , calculados para $p=1$ y diferentes valores de N . Se ve que N apenas tiene influencia sobre DB , mientras que en W se produce un aumento de dos órdenes de magnitud al aumentar N un orden de magnitud. Dicho aumento no es lineal, sino potencial, aproximadamente $W = 0.97N^{2.34}$. Un fuerte aumento de W con N es de esperar debido a que el valor de W aumenta con el número de conectores de la cadena (ecuación 6.11).

N	$\langle DB \rangle \pm \sigma_{DB}$	$\langle W \rangle \pm \sigma_W$
50	0.64 ± 0.06	$(8.68 \pm 0.46) \times 10^3$
100	0.65 ± 0.04	$(4.66 \pm 0.22) \times 10^4$
150	0.66 ± 0.04	$(1.21 \pm 0.06) \times 10^5$
200	0.66 ± 0.03	$(2.37 \pm 0.10) \times 10^5$
250	0.66 ± 0.03	$(3.95 \pm 0.15) \times 10^5$
300	0.66 ± 0.02	$(6.00 \pm 0.23) \times 10^5$
350	0.66 ± 0.02	$(8.55 \pm 0.33) \times 10^5$
400	0.66 ± 0.02	$(1.16 \pm 0.04) \times 10^6$
450	0.66 ± 0.02	$(1.51 \pm 0.05) \times 10^6$
500	0.66 ± 0.02	$(1.91 \pm 0.07) \times 10^6$

TABLA 6.3: Valores de DB y W para diferentes valores de N y $p=1$.

6.6. Modelado, parametrización y simulación.

6.6.1. Generación de topologías.

Una vez estudiados los parámetros que afectan y definen la ramificación de los modelos, el siguiente paso consiste en generar los modelos que van a simularse, comenzando por sus topologías.

Dado que un polímero hiperramificado de un peso molecular concreto no posee una estructura única definida, las propiedades que se estudien deben promediarse no sólo respecto a una serie de conformaciones válidas, sino también respecto a una serie de diferentes estructuras o topologías, que deberán generarse y simularse de forma independiente.

La generación de las distintas topologías se hace utilizando un programa construido a tal efecto que implementa el algoritmo presentado en la sección 6.4. Como se ha descrito, el algoritmo requiere asignar un valor al parámetro p que da la probabilidad de que un nodo sea ramificado, y que en último término definirá el grado de ramificación de la estructura generada. Para generar las moléculas que se describen a continuación se ha utilizado el valor $p = 0.425$.

En una molécula de un polímero hiperramificado real sintetizado en una reacción de tipo *one-pot*, la topología es fruto del azar, aunque se puede suponer que habrá cierta influencia de la naturaleza química de los monómeros implicados, al menos en los aspectos que se refieren a su difusión en el medio de reacción y la accesibilidad a los centros de reacción. Sin embargo, cuantificar la influencia de estos aspectos en la topología final, y representarlos adecuadamente en nuestro proceso es algo muy complicado.

Una opción más sencilla sería la de generar modelos cuyos índices topológicos fuesen similares a los de los polímeros reales con cuyos datos experimentales se compararán los resultados obtenidos de nuestras simulaciones, sin embargo, en la bibliografía de la que se han extraído los datos experimentales, esta información es escasa y vaga en la mayoría de los casos. Por ejemplo, mientras que en [70] sí se especifica $DB = 0.5 \pm 0.1$ para todas las fracciones de pesos moleculares obtenidas, para el resto de datos experimentales no disponemos de ningún dato, ni siquiera orientativo, del grado de ramificación.

Por este motivo, generalizaremos el uso de un parámetro $p = 0.425$ para todos los tamaños de modelo estudiados, de forma que se obtengan los índices topológicos promedio sobre todas las topologías generadas ($\langle DB \rangle$ y $\langle W \rangle$) recogidos en la tabla 6.4.

N(Bolas)	50	100	200	400
n° modelos	640	640	640	640
p	0.425	0.425	0.425	0.425
$\langle DB \rangle \pm \sigma_{DB}$	0.53 ± 0.07	0.53 ± 0.05	0.53 ± 0.04	0.54 ± 0.02
$\langle W \rangle \pm \sigma_W$	9400 ± 700	$(5.1 \pm 0.3) \times 10^4$	$(2.6 \pm 0.2) \times 10^5$	$(1.3 \pm 0.1) \times 10^6$

TABLA 6.4: Parámetros que definen la generación y la topología de las moléculas simuladas

El uso de esta metodología y los parámetros indicados dan como resultado moléculas con topologías similares a la que se presenta en la figura 6.6, independientemente de cuál sea la naturaleza química exacta del producto. En la figura, para mejor visibilidad, se han representado únicamente los conectores del modelo, omitiendo las bolas, y fijando los ángulos entre los conectores a 120° . Además, cada rama del modelo se ha dibujado de un color diferente.

6.6.2. Modelado y parametrización.

Como en los demás trabajos que comprende esta Tesis, se emplearán modelos de bolas y muelles para realizar las simulaciones. Dados los buenos resultados obtenidos por el Grupo en su trabajo sobre simulación de dendrímeros [24], y dado que el objetivo es similar, la realización de simulaciones en unidades físicas para intentar reproducir los resultados experimentales obtenidos por otros autores, se optó seguir la misma estrategia de simulación multiescala que en el citado trabajo, empleando las mismas soluciones adoptadas en aquel trabajo, concretamente, la construcción de los modelos a simular con muelles tipo “Hard-FENE”, muelles angulares y un potencial de volumen excluido de “esferas duras”, todos ellos ya definidos en la sección 2.2.

Un punto común a todas las simulaciones de este Capítulo es que todos los polímeros aquí comprendidos pueden modelarse empleando un único tipo de bola para cada uno, por lo que sólo será necesario realizar la parametrización de un tipo de bola en cada caso.

En el caso de los muelles, dado que todos los monómeros son del tipo AB_2 , en los casos en los que haya ramificación, distinguiremos dos tipos de muelles angulares: los

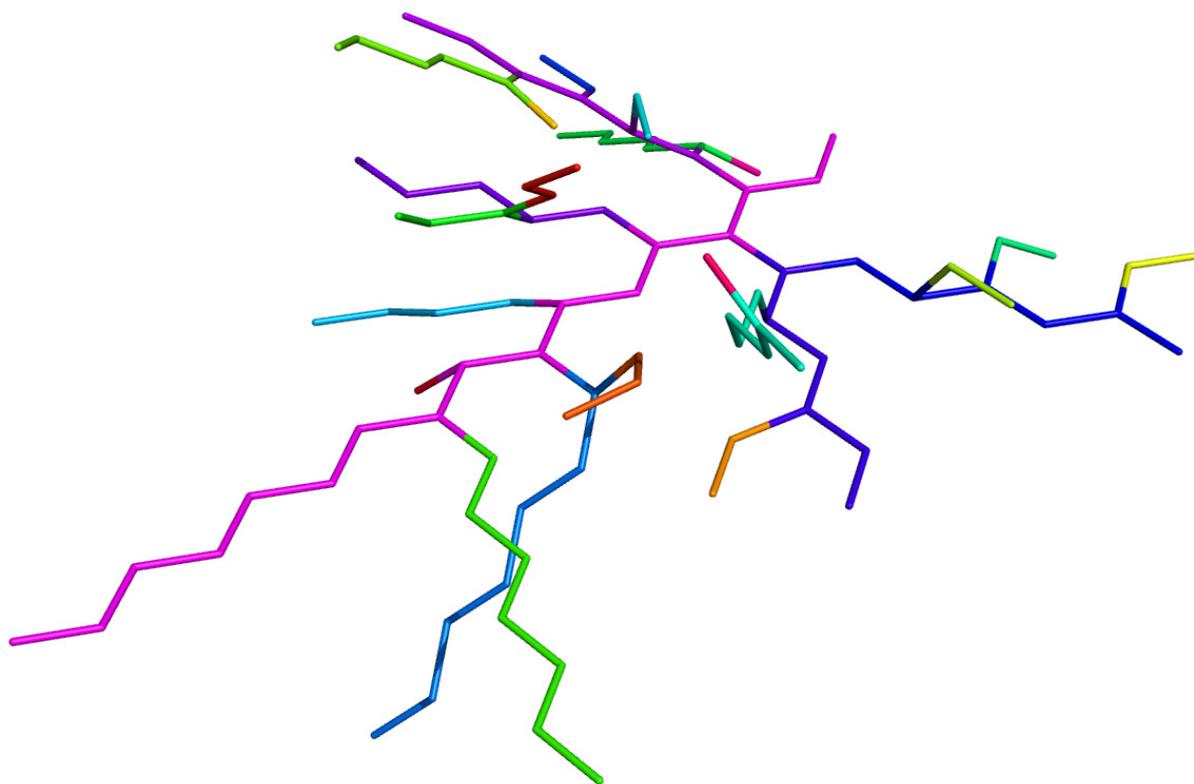


FIGURA 6.6: Ejemplo de molécula $N = 100$ generada con el algoritmo de topologías presentado en la sección 6.4 y $p = 0.425$.

“internos”, definidos entre las dos direcciones en las que se ramifica el monómero actual, y los “externos”, definidos entre cada una de las direcciones de ramificación y la dirección en la que el monómero actual se une a su antecesor inmediato. Y, como sólo existe un tipo de bola, para cada uno de los polímeros simulados sólo será necesario determinar los parámetros de un tipo de conector y de este par de tipos de ángulos.

Dado que algunos de los polímeros simulados en este Capítulo están compuestos por los mismos monómeros que ya se parametrizaron en el trabajo sobre dendrímeros, se han empleado directamente los datos obtenidos para aquél en la elaboración de los nuevos modelos hiperramificados de esas especies.

Para la parametrización del resto de los modelos se emplea la estrategia descrita en la sección 2.3, recurriendo a algunas adaptaciones específicas a cada caso. Los detalles concretos de cada uno de estos polímeros se resumen al principio de su correspondiente sección de resultados.

En cuanto a las interacciones de largo alcance, este procedimiento no está descrito en la sección 2.3, pues es específico para cada tipo de interacción. En todas las simulaciones de este Capítulo se emplea el potencial de esferas duras, por lo que una vez que se ha obtenido el radio de las bolas del modelo, es inmediato parametrizar las interacciones de largo alcance, pues el único dato necesario para aplicar este potencial es σ_{HS} , el tamaño de las esferas a las que se impide solaparse, que se corresponderá con radio de las bolas del modelo, $\sigma_{HS} = \sigma$.

6.7. Resultados.

En esta sección se presentan las parametrizaciones y los resultados numéricos de las simulaciones. En la siguiente sección se elaborarán y discutirán dichos resultados.

6.7.1. Polibenziléter (PBzE).

Un ejemplo de un fragmento de la estructura de un polímero de polibenziléter se presenta en la figura 6.7, mientras que el monómero a partir del que se sintetiza, el benziléter, se presenta en el recuadro de la misma figura.

Parametrización.

En el caso de este polímero, los valores de los parámetros del modelo ya fueron calculados en [24], por lo que se emplean directamente estos valores ya conocidos, recogidos en la tabla 6.5. Es destacable que, en este caso, los dos tipos de ángulos posibles (internos y externos) presentan los mismo parámetros, por lo que no se hace distinción alguna entre ellos.

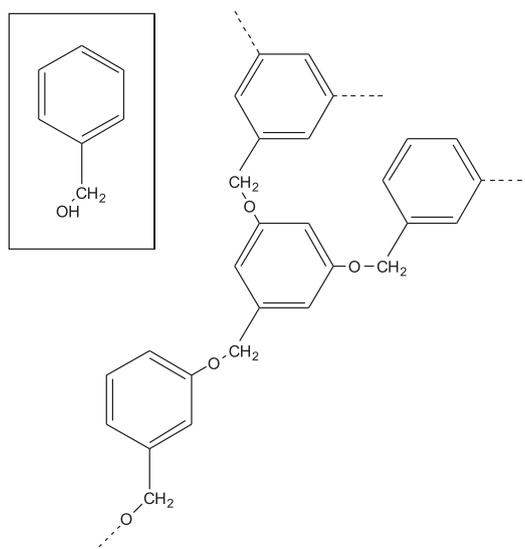


FIGURA 6.7: Monómero de PBzE y fragmento de la estructura de un polímero.

Elemento del modelo	Tipo	Parámetros y valores asociados
Radio de bola		$\sigma = 3.2 \text{ \AA}$
Conectores	Hard-FENE	$H = 404 \text{ erg/cm}^2$; $l_e = 5.24 \text{ \AA}$; $l_{max} = 7.84 \text{ \AA}$
Muelles angulares	Cuadrático	$Q = 1.22 \times 10^{-13} \text{ erg}$; $\alpha_0 = 1.055 \text{ rad}$
Volumen excluido	Esfera Dura	$\sigma_{HS} = 3.2 \text{ \AA}$; $V = 8.08 \times 10^{-12} \text{ erg}$

TABLA 6.5: Parámetros para los modelos de PBzE.

Simulación.

Con los datos de la parametrización, se construyen y simulan modelos compuestos por 50, 100, 200 y 400 bolas. Los parámetros con los que se lanzaron las simulaciones están resumidos en la tabla 6.6

Resultados.

Para cada uno de los tamaños moleculares se generan y simulan 640 topologías, obteniendo los dobles promedios, sobre conformaciones y topologías, de las propiedades presentadas en la tabla 6.7. También se presentan en dicha tabla los radios equivalentes calculados con dichas propiedades y las pertinentes relaciones entre ellos.

Número de bolas	50	100	200	400
M	5302	10602	21202	42402
Submuestras	5	5	5	5
Número de pasos	8×10^7	2×10^7	5×10^6	1.8×10^6
Δx	0.03	0.03	0.03	0.005
Tasa media de rechazo	15 %	24 %	37 %	15 %

TABLA 6.6: Parámetros de las simulaciones de Monte Carlo de PBzE.

Número de bolas	50	100	200	400
R_g (Å)	17.56 ± 0.05	22.69 ± 0.05	28.81 ± 0.08	35.17 ± 0.14
$D_t \times 10^7$ (cm²/s)	11.91 ± 0.01	8.79 ± 0.01	6.64 ± 0.01	5.19 ± 0.01
$[\eta]$ (cm³/g)	7.61 ± 0.02	9.12 ± 0.03	10.42 ± 0.04	11.09 ± 0.07
$(R_g)_{corr}$ (Å)	17.73 ± 0.05	22.83 ± 0.05	28.91 ± 0.08	35.26 ± 0.14
a_T (Å)	18.01 ± 0.02	24.42 ± 0.02	32.33 ± 0.04	41.30 ± 0.08
a_l (Å)	18.56 ± 0.02	24.84 ± 0.03	32.72 ± 0.04	42.09 ± 0.07
a_G (Å)	22.89 ± 0.06	29.47 ± 0.07	37.33 ± 0.10	45.52 ± 0.17
GI	1.23	1.19	1.14	1.08
GT	1.27	1.21	1.15	1.10
IT	1.03	1.02	1.01	1.02

TABLA 6.7: Promedios de las propiedades calculadas para cada tamaño de modelo simulado de PBzE.

6.7.2. Poliaminoamida (PAMAM).

Una representación de una parte de una molécula de poliaminoamida y su monómero, la N-(2-aminoetil)-propanamida, se presentan en la figura 6.8.

Parametrización.

También en el caso de este polímero se conoce el valor de los parámetros debido a trabajos anteriores. La tabla 6.8 presenta un resumen de estos datos. Hay que destacar que, en este caso, los ángulos externos son libres (es decir, no presentan restricciones, por lo que no se define ninguna ecuación de potencial para ellos), mientras que los ángulos internos sí presentan restricciones.

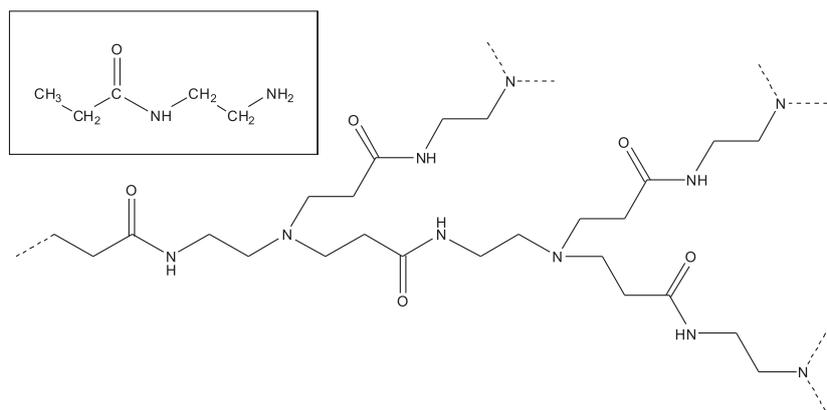


FIGURA 6.8: Monómero de N-(2-aminoetil)-propanamida y fragmento de la estructura de un polímero de PAMAM.

Elemento del modelo	Tipo	Parámetros y valores asociados
Radio de bola		$\sigma = 3.6 \text{ \AA}$
Conectores	Hard-FENE	$H = 161.6 \text{ erg/cm}^2$; $l_e = 6.45 \text{ \AA}$; $l_{max} = 10.0 \text{ \AA}$
Muelles angulares internos	Cuadrático	$Q = 1.35 \times 10^{-12} \text{ erg}$; $\alpha_0 = 2.2 \text{ rad}$
Volumen excluido	Esfera Dura	$\sigma_{HS} = 3.6 \text{ \AA}$; $V = 8.08 \times 10^{-12} \text{ erg}$

TABLA 6.8: Parámetros para los modelos de PAMAM.

Mientras el grueso de las simulaciones de PAMAM se realizaron con los datos antes mencionados, una vez vistos los resultados obtenidos, también se realizó una batería de simulaciones en las que se modificaron algunos de los parámetros del modelo de PAMAM, con el objetivo de estudiar cómo influyen los diferentes parámetros en las propiedades obtenidas.

En concreto, se estudiaron, sobre moléculas de PAMAM de 100 bolas, los efectos de las modificaciones consistentes en definir también restricciones en los ángulos externos (con los mismos valores que los internos), en no definir ningún potencial angular, en tomar un valor doble o mitad del valor inicialmente estimado como radio de la esfera dura en el volumen excluido, eliminar éste último, y finalmente, variaciones de un 20 % en el radio de las bolas empleadas (adaptando también el volumen excluido al cambio). Todas estas modificaciones se resumen en la tabla 6.9, donde se presentan los valores numéricos asociados a estos cambios.

Elemento modificado	Modificación	Parámetros utilizados
Restricciones angulares	Ángulos internos y externos Sin ángulos definidos	$Q = 1.35 \times 10^{-12}$ erg; $\alpha_0 = 2.2$ rad -
Volumen excluido	Sin V. E. $\sigma'_{HS} = 0.5\sigma_{HS}$ $\sigma'_{HS} = 2.0\sigma_{HS}$	- $\sigma'_{HS} = 1.8 \text{ \AA}$ $\sigma'_{HS} = 7.2 \text{ \AA}$
Radio de bola	$\sigma' = 0.8\sigma$ $\sigma' = 1.2\sigma$	$\sigma' = 2.88 \text{ \AA}$; $\sigma'_{HS} = 2.88 \text{ \AA}$ $\sigma' = 4.32 \text{ \AA}$; $\sigma'_{HS} = 4.32 \text{ \AA}$

TABLA 6.9: Parámetros modificados en las simulaciones de Monte Carlo de PAMAM para determinar su influencia en los resultados.

Simulación.

Con los datos de las parametrizaciones de PAMAM se generan una serie de topologías correspondientes a diferentes tamaños de moléculas. En la tabla 6.10 se recogen los datos empleados en la generación, así como también los parámetros utilizados en las simulaciones. La tabla 6.11 recoge los datos de tamaño de paso y tasa de rechazo correspondientes a las simulaciones de cadenas de 100 bolas con los parámetros modificados mencionados en la tabla 6.9.

Número de bolas	50	100	200	400
M	5702	11402	22802	45602
Submuestras	5	5	5	5
Número de pasos	8.0×10^7	2.0×10^7	5.0×10^6	1.0×10^6
Δx	0.05	0.02	0.005	0.001
Tasa media de rechazo	24 %	17 %	9 %	10 %

TABLA 6.10: Parámetros empleados en las simulaciones de Monte Carlo de modelos de PAMAM.

Resultados.

En la tabla 6.12 se recogen los resultados obtenidos de las simulaciones, así como los correspondientes radios equivalentes y las relaciones entre los mismos.

En lo que se refiere a los resultados obtenidos de las simulaciones con parámetros modificados, los resultados se presentan en las tablas 6.13 y 6.14.

Caso	Δx	Tasa media de rechazo
Ángulos internos y externos	0.01	15 %
Sin ángulos definidos	0.025	20 %
Sin volumen excluido	0.05	20 %
$\sigma'_{HS} = 0.5\sigma_{HS}$	0.03	21 %
$\sigma'_{HS} = 2.0\sigma_{HS}$	0.005	22 %
$\sigma' = 0.8\sigma$	0.03	22 %
$\sigma' = 1.2\sigma$	0.02	19 %

TABLA 6.11: Tamaño de paso e índice de rechazo de pasos en las simulaciones de Monte Carlo PAMAM con parámetros modificados.

Número de bolas	50	100	200	400
R_g (Å)	20.83 ± 0.04	26.29 ± 0.05	28.98 ± 0.06	31.08 ± 0.08
$D_t \times 10^7$ (cm ² /s)	9.85 ± 0.01	7.36 ± 0.01	6.11 ± 0.01	5.24 ± 0.01
$[\eta]$ (cm ³ /g)	12.13 ± 0.03	14.09 ± 0.04	12.33 ± 0.05	9.89 ± 0.05
$(R_g)_{corr}$ (Å)	21.01 ± 0.04	26.44 ± 0.05	29.12 ± 0.06	31.21 ± 0.08
a_T (Å)	21.78 ± 0.02	29.13 ± 0.02	35.11 ± 0.05	40.91 ± 0.07
a_l (Å)	22.22 ± 0.02	29.42 ± 0.03	35.46 ± 0.05	41.51 ± 0.07
a_G (Å)	27.13 ± 0.06	34.13 ± 0.06	37.59 ± 0.08	40.29 ± 0.10
GI	1.22	1.16	1.06	0.97
GT	1.25	1.17	1.07	0.98
IT	1.02	1.01	1.01	1.01

TABLA 6.12: Promedios de las propiedades calculadas para cada tamaño de modelo simulado de PAMAM.

6.7.3. Polimetilalilcarbosilano (PCS3).

A diferencia de los dos casos anteriores, en los que en el Grupo ya se habían realizado simulaciones sobre estos polímeros, y por tanto, ya se disponía de datos de parametrización, es la primera vez que se van a realizar simulaciones de modelos de esta especie, por lo que en primer lugar es necesario realizar la parametrización de los elementos de los modelos.

Caso	Ángulos internos y externos	Sin ángulos definidos	$\sigma' = 0.8\sigma$	$\sigma' = 1.2\sigma$
R_g (Å)	19.91 ± 0.04	26.70 ± 0.05	24.58 ± 0.05	28.12 ± 0.05
$D_t \times 10^7$ (cm ² /s)	9.17 ± 0.01	7.33 ± 0.01	8.09 ± 0.01	6.75 ± 0.01
$[\eta]$ (cm ³ /g)	7.61 ± 0.02	14.35 ± 0.04	10.52 ± 0.03	18.75 ± 0.05
$(R_g)_{corr}$ (Å)	20.11 ± 0.04	26.85 ± 0.05	24.68 ± 0.05	28.32 ± 0.05
a_T (Å)	23.39 ± 0.02	29.28 ± 0.02	26.52 ± 0.02	31.80 ± 0.03
a_I (Å)	23.96 ± 0.02	29.60 ± 0.03	26.69 ± 0.02	32.26 ± 0.03
a_G (Å)	25.96 ± 0.05	34.66 ± 0.07	31.86 ± 0.06	36.56 ± 0.07
GI	1.08	1.17	1.19	1.13
GT	1.11	1.18	1.20	1.15
IT	1.02	1.01	1.01	1.01

TABLA 6.13: Resultados de las simulaciones de Monte Carlo de PA-MAM con ángulos y tamaño de bolas modificados.

Caso	Sin V.E.	$\sigma'_{HS} = 0.5\sigma_{HS}$	$\sigma'_{HS} = 2.0\sigma_{HS}$
R_g (Å)	15.77 ± 0.02	21.30 ± 0.04	35.16 ± 0.06
$D_t \times 10^7$ (cm ² /s)	10.62 ± 0.01	8.70 ± 0.01	5.71 ± 0.01
$[\eta]$ (cm ³ /g)	5.29 ± 0.01	8.94 ± 0.02	29.24 ± 0.08
$(R_g)_{corr}$ (Å)	16.01 ± 0.02	21.48 ± 0.04	35.26 ± 0.06
a_T (Å)	20.20 ± 0.01	24.67 ± 0.02	37.55 ± 0.03
a_I (Å)	21.23 ± 0.01	25.28 ± 0.02	37.53 ± 0.03
a_G (Å)	20.67 ± 0.02	27.73 ± 0.05	45.52 ± 0.07
GI	0.97	1.10	1.21
GT	1.02	1.12	1.21
IT	1.05	1.02	1.00

TABLA 6.14: Resultados de las simulaciones de Monte Carlo de PA-MAM con volumen excluido modificado.

Parametrización.

Como se ha mencionado en la sección 2.3.1, la determinación del radio de las bolas del modelo se hace a partir de un pseudo-monómero del compuesto. La estructura correspondiente a la estructura pseudo-monomérica elegida para este caso se muestra en el detalle de la figura 6.9, en la que también se muestra un ejemplo de una estructura parcial de un

polímero de esta especie, el mismo empleado para la parametrización de los parámetros de los muelles conectores, los angulares, y el volumen excluido.

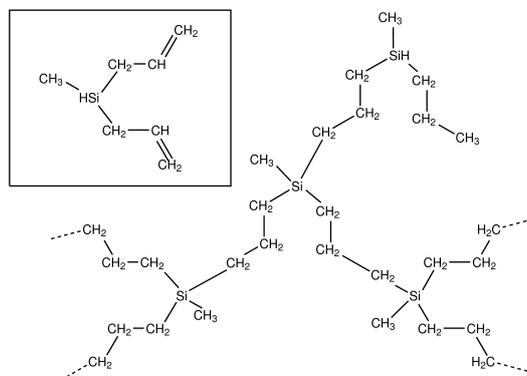


FIGURA 6.9: Monómero de PCS3 y estructura parcial del polímero.

Tal como se ha descrito, sobre el pseudo-monómero se realiza una simulación atomística mediante dinámica de Langevin, empleando para ello el programa HYPERCHEM, junto con los parámetros de simulación recogidos en la tabla 6.15.

Campo de fuerzas	AMBER
Duración de trayectoria	20000 ps (20 ns)
Tamaño de paso	0.002 ps
Factor de fricción	50 ps ⁻¹
Temperatura	300 K
Frecuencia de adquisición de datos	5000 pasos
Número de conformaciones adquiridas	2000

TABLA 6.15: Parámetros de la dinámica de Langevin para PCS3.

De los datos obtenidos de HYPERCHEM se extraen diez conformaciones equiespaciadas a lo largo de la trayectoria que se analizan mediante HYDROPRO, permitiendo al programa fijar automáticamente todos los parámetros necesarios (número de intervalos, tamaños de bolas, volumen específico, etc). De esta manera se obtienen los valores de la difusión traslacional (D_t) y radio de giro (R_g) correspondientes a las conformaciones extraídas de la trayectoria atomística. A partir de esos datos se calculan los respectivos radios equivalentes (a_T y a_G), y se calcula el valor promedio cada uno de ellos. Los resultados correspondientes a estos cálculos se presentan en la tabla 6.16.

Como se puede observar, los valores finales de ambos radios equivalentes son bastante próximos entre sí, por lo que no es disparatado aceptar el valor promedio de ambos,

(σ), como representativo del pseudo-monómero, y emplearlo como radio de las bolas del modelo.

Caso	$D_t \times 10^6$ (cm ² /s)	$R_g \times 10^8$ (cm)	a_T (Å)	a_G (Å)	$\langle \sigma \rangle$ (Å)
1	4.94	3.63	4.45	4.69	4.57
2	4.94	3.63	4.45	4.69	4.57
3	4.97	3.52	4.42	4.54	4.48
4	5.02	3.52	4.38	4.54	4.46
5	4.92	3.66	4.46	4.73	4.59
6	4.95	3.63	4.44	4.69	4.56
7	4.93	3.62	4.46	4.67	4.56
8	4.93	3.68	4.46	4.75	4.60
9	5.00	3.48	4.39	4.49	4.55
10	4.99	3.50	4.40	4.52	4.46
Promedio	4.96	3.59	4.43	4.63	4.53

TABLA 6.16: Cálculo del radio del pseudo-monómero de PCS3 a partir de radios equivalentes de conformaciones extraídas de HYPERCHEM.

Para la determinación de los parámetros de los muelles se emplea de nuevo una simulación mediante dinámica de Langevin, en esta ocasión aplicada al modelo de la unidad estructural mínima sobre la que se pueden definir conectores y ángulos entre bolas. Esta unidad mínima correspondería a la estructura que se mostraba en la figura 6.9, mientras que en la figura 6.10 se muestra un esquema equivalente a esa estructura y la correspondencia de la misma con las bolas del modelo que se emplea para calcular los parámetros de los muelles conectores (resaltados en azul) y los muelles angulares (destacados en rojo).

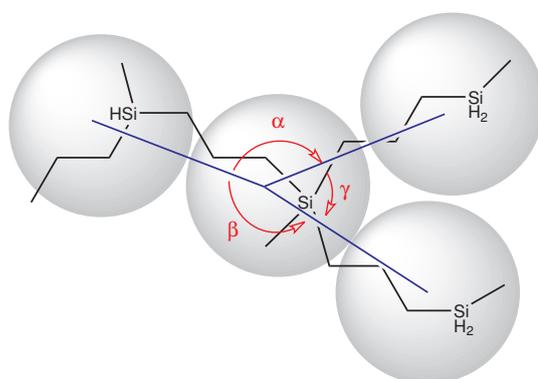


FIGURA 6.10: Esquema de la unidad estructural mínima de PCS3 con bolas, conectores y ángulos. Los ángulos α y β son “externos” y deberían ser aproximadamente equivalentes, mientras γ es interno.

Para realizar la simulación, se utilizan los mismos datos que para el caso de la determinación del tamaño de bola (tabla 6.15), con la salvedad de que la trayectoria generada debe ser más larga para permitir que la variabilidad conformacional barrida por la simulación sea lo suficientemente grande como para ser representativa; así pues, la duración de la trayectoria se prolonga hasta los 100 ns (100000 ps), y por lo tanto, el número de conformaciones adquiridas pasa a ser 10000.

Una vez terminada la simulación, se analiza la trayectoria obtenida. Para cada una de las conformaciones registradas, se calculan las posiciones de los centros de masas de cada una de las bolas, los cuales se toman como referencia para los cálculos. A continuación, se obtiene la distancia entre las bolas, que corresponde a la longitud instantánea de los conectores, y también los ángulos que se definen en torno a la bola central, tanto los dos que, atendiendo a la direccionalidad de la construcción, llamamos “externos” como el “interno”. Las distribuciones de las longitudes de los enlaces se representan en la figura 6.11, mientras que las distribuciones de los ángulos pueden observarse en la figura 6.12. Una vez obtenidos los datos, éstos se ajustan a las ecuaciones de los potenciales para obtener los valores de los parámetros de los distintos muelles.

Las figuras 6.13, 6.15 y 6.17 representan las funciones de distribución las longitudes de enlace y los ángulos obtenidos de las trayectorias calculadas mediante HYPERCHEM frente a las funciones de distribución obtenidas de los ajustes de los datos a las ecuaciones de los potenciales Hard-FENE y angular armónico, respectivamente, mediante el programa FUNFIT. En las figuras 6.14, 6.16 y 6.18 se presentan las curvas de nivel de la desviación cuadrática de las funciones de distribución ajustadas a los potenciales Hard-FENE y angular armónico respecto a los datos de HYPERCHEM en los rangos de valores de los parámetros entre los que se ha realizado la búsqueda de los valores que proporcionen el mejor ajuste.

Dado que los gráficos de las funciones de distribución y las curvas de nivel de desviación típica obtenidas en los ajustes y presentadas en las figuras anteriores no aportan información adicional relevante, se presentan en el caso del PCS3 a modo de ejemplo, pero en lo subsiguiente se omitirá en los casos de los demás polímeros que hubieron de ser parametrizados.

Los datos obtenidos en la parametrización se resumen en la tabla 6.17.

Como en el caso del PAMAM, además de las simulaciones realizadas con los datos obtenidos de la parametrización, y vistos los malos resultados obtenidos, se realizaron simulaciones adicionales para estudiar la influencia del valor de algunos de los parámetros

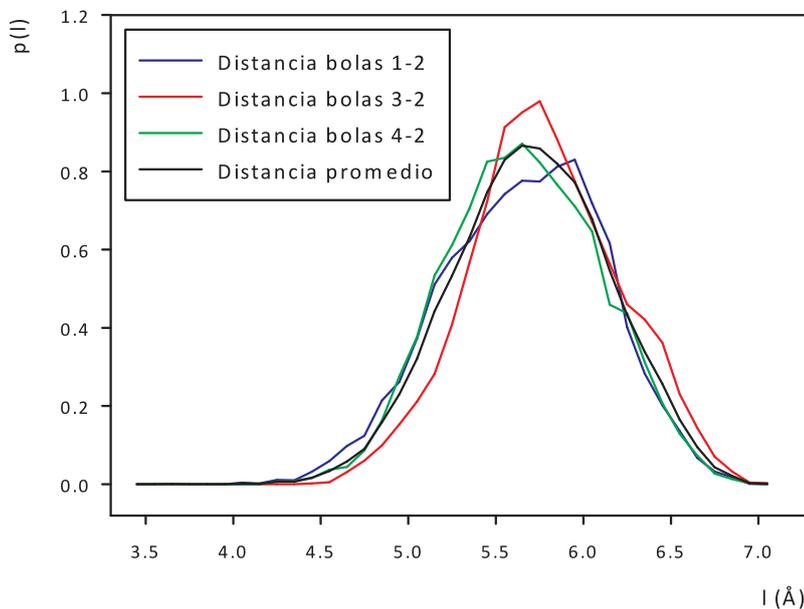


FIGURA 6.11: Funciones de distribución de la longitud de los conectores observados en PCS3.

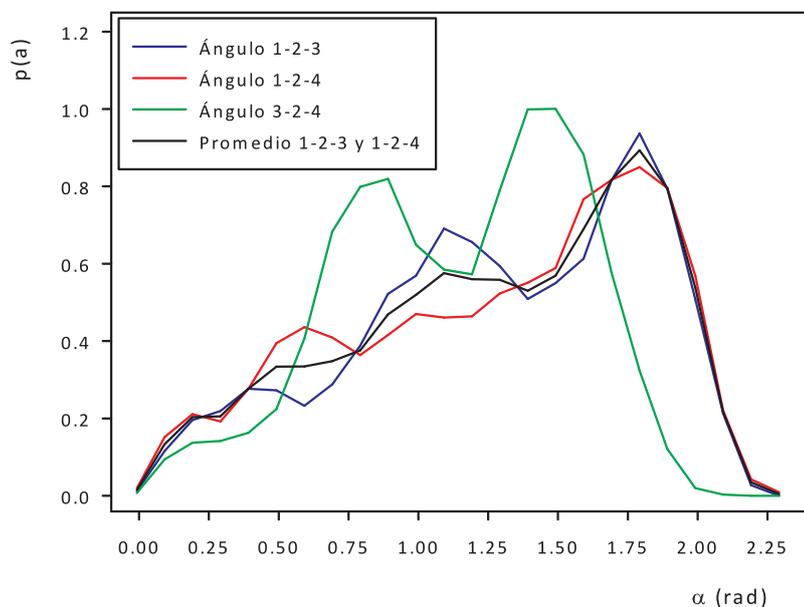


FIGURA 6.12: Funciones de distribución de los ángulos entre los conectores observados en PCS3.

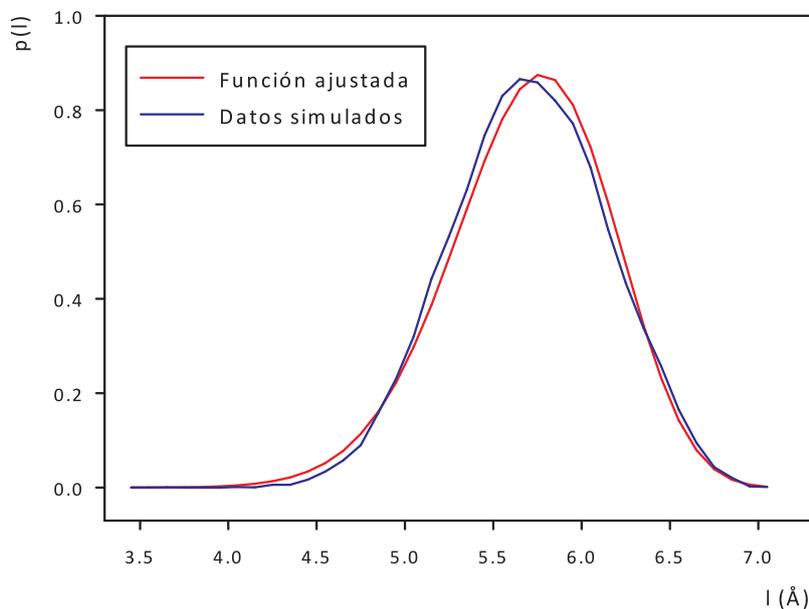


FIGURA 6.13: Ajuste de los parámetros de los conectores en PCS3.

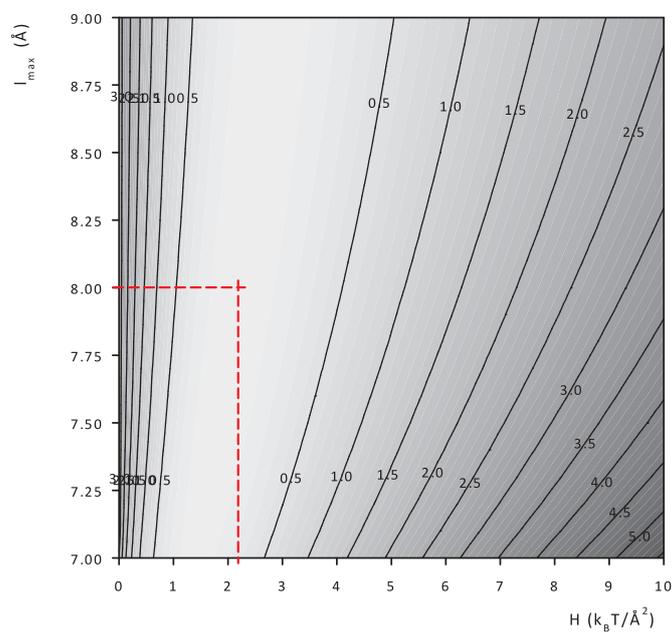


FIGURA 6.14: Desviaciones cuadráticas respecto a los parámetros de los conectores en PCS3. Las líneas discontinuas marcan los valores de H y l_{max} que hacen que la desviación cuadrática sea mínima, por lo que serán utilizados en la simulación.

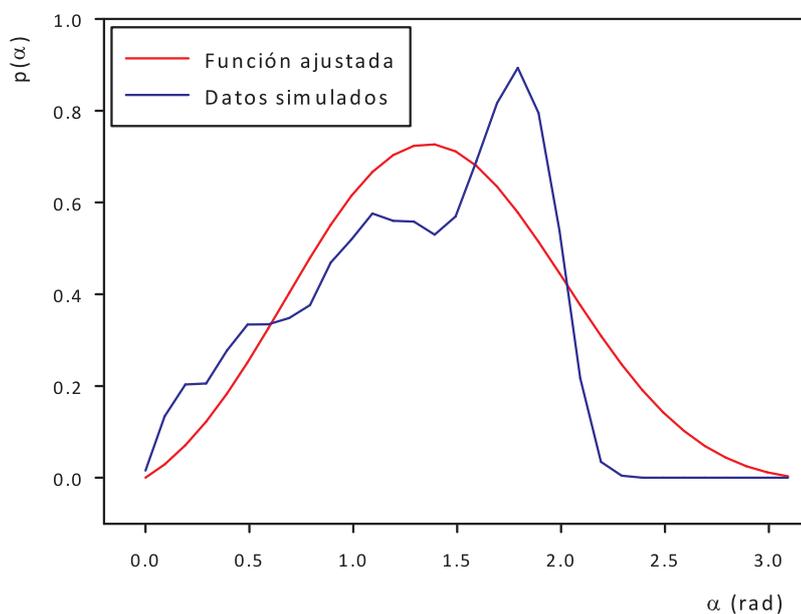


FIGURA 6.15: Ajuste de los parámetros de los muelles angulares externos en PCS3.

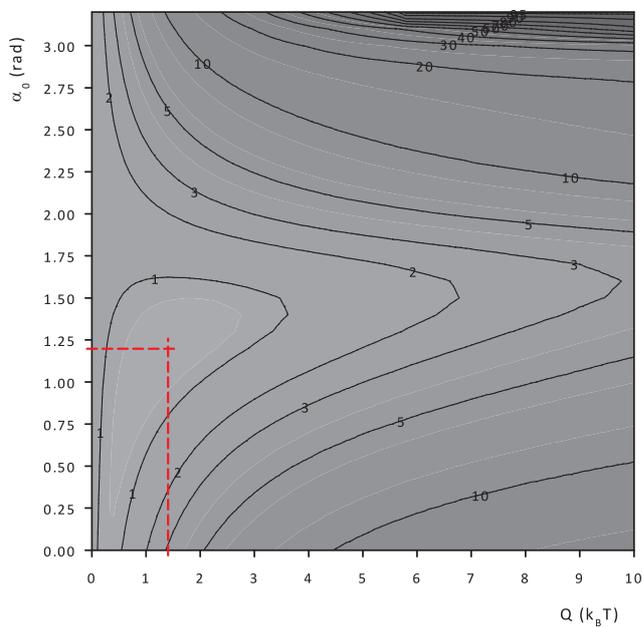


FIGURA 6.16: Desviaciones cuadráticas respecto a los parámetros de los muelles angulares externos en PCS3. Las líneas discontinuas marcan los valores de Q y α_e que hacen que la desviación cuadrática sea mínima, por lo que serán utilizados en la simulación.

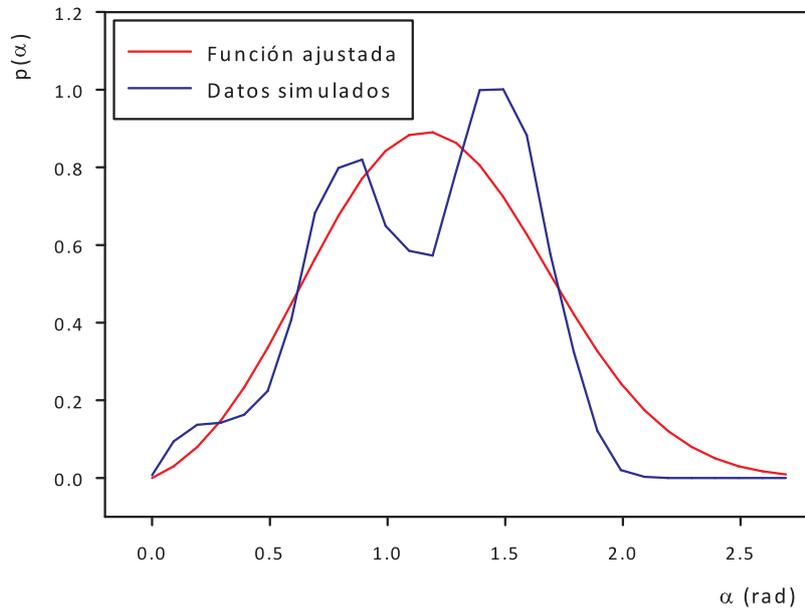


FIGURA 6.17: Ajuste de los parámetros de los muelles angulares internos en PCS3.

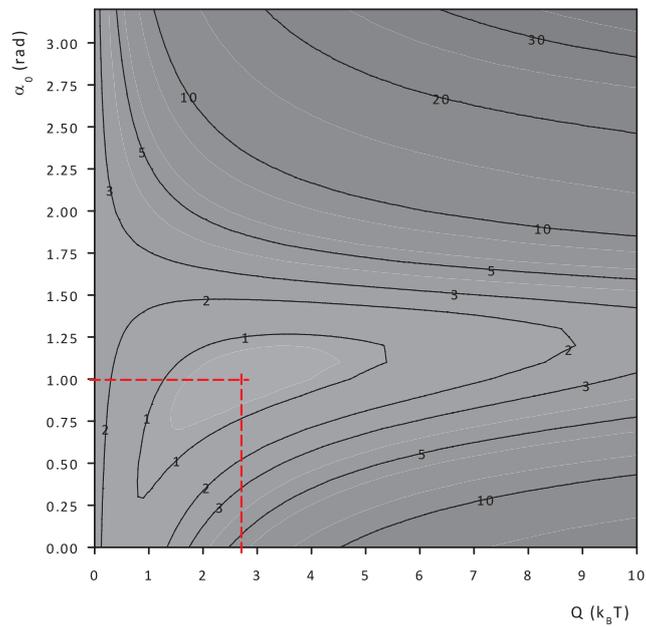


FIGURA 6.18: Desviaciones cuadráticas respecto a los parámetros de los muelles angulares internos en PCS3. Las líneas discontinuas marcan los valores de Q y α_e óptimos que serán utilizados en la simulación.

Elemento del modelo	Tipo	Parámetros y valores asociados
Radio de bola		$\sigma = 4.53 \text{ \AA}$
Conectores	Hard-FENE	$H = 1131 \text{ erg/cm}^2$; $l_e = 5.70 \text{ \AA}$; $l_{max} = 9.10 \text{ \AA}$
Muelles angulares externos	Cuadrático	$Q = 5.66 \times 10^{-14} \text{ erg}$; $\alpha_0 = 1.20 \text{ rad}$
Muelles angulares internos	Cuadrático	$Q = 1.09 \times 10^{-13} \text{ erg}$; $\alpha_0 = 1.00 \text{ rad}$
Volumen excluido	Esfera Dura	$\sigma_{HS} = 4.53 \text{ \AA}$; $V = 8.08 \times 10^{-12} \text{ erg}$

TABLA 6.17: Parámetros para los modelos de PCS3.

en los resultados de las simulaciones de PCS3. Los parámetros que se modificaron fueron los mismos, y en las mismas proporciones que en el caso del PAMAM. En concreto, se estudió sobre el PCS3 de 100 bolas la influencia de los parámetros angulares, primero, eliminando las restricciones en los ángulos externos (por ser los más restrictivos), y en segundo lugar, eliminando completamente las restricciones angulares; también se modificó el tamaño de esfera de volumen excluido, ensayando simulaciones sin volumen excluido, reduciendo el tamaño de la esfera inicial a la mitad, y finalmente, duplicándolo; y por último, se modificó el radio de las bolas del modelo, estudiando el efecto tanto de un aumento como de una reducción del 20 %, reflejando convenientemente la variación en el volumen excluido. La tabla 6.18 resume las variaciones que se hicieron y los datos que se modificaron respecto a los datos de la tabla 6.17.

Elemento modificado	Modificación	Parámetros utilizados
Restricciones angulares	Sólo ángulos internos Sin ángulos definidos	$Q = 1.09 \times 10^{-13} \text{ erg}$; $\alpha_0 = 1.00 \text{ rad}$ -
Volumen excluido	Sin V. E. $0.5\sigma_{HS}$ $2.0\sigma_{HS}$	- $\sigma'_{HS} = 2.26 \text{ \AA}$ $\sigma'_{HS} = 9.06 \text{ \AA}$
Radio de bola	0.8σ 1.2σ	$\sigma' = 3.62 \text{ \AA}$; $\sigma'_{HS} = 3.62 \text{ \AA}$ $\sigma' = 5.44 \text{ \AA}$; $\sigma'_{HS} = 5.44 \text{ \AA}$

TABLA 6.18: Parámetros modificados en las simulaciones de Monte Carlo de PCS3 para determinar su influencia en los resultados.

Simulación.

Los parámetros que se emplearon en las simulaciones se muestran en la tabla 6.19. En las simulaciones empleadas para determinar la influencia de los parámetros del modelo sobre los resultados, los parámetros de simulación son los mismos que los empleados para

las simulaciones originales de los modelos de 100 bolas recogidos en esta tabla, con la única excepción de los valores de Δx y la tasa de rechazo, cuyos valores para cada caso se detallan en la tabla 6.20.

Número de bolas	50	100	200	400
M	6402	12802	25602	51202
Submuestras	5	5	5	5
Número de pasos	7.0×10^7	2.0×10^7	5.0×10^6	1.2×10^6
Δx	0.03	0.015	0.008	0.0025
Tasa media de rechazo	22 %	18 %	17 %	28 %

TABLA 6.19: Parámetros de las simulaciones de Monte Carlo de PCS3.

Caso	Δx	Tasa media de rechazo
Sólo ángulos internos	0.01	19 %
Sin ángulos definidos	0.008	20 %
Sin volumen excluido	0.025	18 %
$\sigma'_{HS} = 0.5\sigma_{HS}$	0.025	23 %
$\sigma'_{HS} = 2.0\sigma_{HS}$	0.003	26 %
$\sigma' = 0.8\sigma$	0.025	26 %
$\sigma' = 1.2\sigma$	0.025	36 %

TABLA 6.20: Parámetros de las simulaciones de Monte Carlo de PCS3 con parámetros modificados.

Resultados.

En la tabla 6.21 se recogen los resultados obtenidos en las simulaciones, así como los correspondientes radios equivalentes y las relaciones entre los mismos.

Finalmente, en las tablas 6.22 y 6.23 se presentan los datos obtenidos a partir de las simulaciones con datos modificados.

6.7.4. Polimetilundecilcarbosilano (PCS11).

El siguiente paso es la simulación del polímero PCS11, muy similar al PCS3, con la diferencia de que en este caso los “brazos” del pseudo-monómero están compuestos por

Número de bolas	50	100	200	400
R_g (Å)	19.62 ± 0.05	25.46 ± 0.07	32.30 ± 0.11	38.35 ± 0.13
$D_t \times 10^7$ (cm ² /s)	9.94 ± 0.01	7.42 ± 0.01	5.69 ± 0.01	4.46 ± 0.01
$[\eta]$ (cm ³ /g)	11.19 ± 0.03	12.96 ± 0.04	14.36 ± 0.07	14.82 ± 0.08
$(R_g)_{corr}$ (Å)	19.93 ± 0.05	25.70 ± 0.07	32.48 ± 0.11	38.52 ± 0.13
a_T (Å)	21.58 ± 0.02	28.89 ± 0.03	37.72 ± 0.06	48.15 ± 0.08
a_I (Å)	22.48 ± 0.02	29.74 ± 0.03	38.77 ± 0.06	49.37 ± 0.06
a_G (Å)	25.74 ± 0.07	33.18 ± 0.09	41.93 ± 0.15	49.73 ± 0.16
GI	1.14	1.12	1.08	1.01
GT	1.19	1.15	1.11	1.03
IT	1.04	1.03	1.03	1.03

TABLA 6.21: Propiedades promediadas para cada tamaño de molécula de PCS3.

cadenas de once átomos de carbono en lugar de tres. En la figura 6.19 se presenta la estructura de uno de estos pseudo-monómeros.

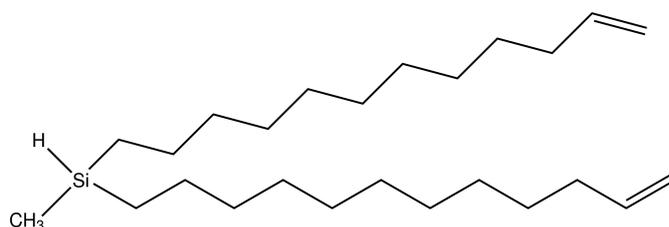


FIGURA 6.19: Monómero de PCS11.

Parametrización.

Se sigue el mismo procedimiento utilizado en la parametrización del PCS3, y empleando también los parámetros de simulación que se presentaron en la tabla 6.15 de la sección anterior.

En este caso, los resultados del análisis con HYDROPRO son los que se presentan en la tabla 6.24. Se puede observar que en esta ocasión la diferencia entre los valores de los promedios de a_T y a_G es bastante más elevada que en el caso del PCS3, sin embargo, aún a pesar de esto, se empleará el promedio de ambos valores como representativo del radio de las bolas de este modelo.

Caso	Sólo ángulos internos	Sin ángulos definidos	$\sigma' = 0.8\sigma$	$\sigma' = 1.2\sigma$
R_g (Å)	25.58 ± 0.05	25.36 ± 0.05	22.74 ± 0.04	25.99 ± 0.05
$D_t \times 10^7$ (cm ² /s)	7.21 ± 0.01	7.22 ± 0.01	8.35 ± 0.01	6.89 ± 0.01
$[\eta]$ (cm ³ /g)	13.86 ± 0.04	13.81 ± 0.04	8.83 ± 0.02	16.60 ± 0.04
$(R_g)_{corr}$ (Å)	25.82 ± 0.05	25.50 ± 0.05	22.92 ± 0.04	26.33 ± 0.05
a_T (Å)	29.76 ± 0.03	29.73 ± 0.03	25.70 ± 0.02	31.15 ± 0.03
a_I (Å)	30.41 ± 0.03	30.41 ± 0.03	26.17 ± 0.02	32.30 ± 0.03
a_G (Å)	33.34 ± 0.07	33.06 ± 0.03	29.58 ± 0.06	33.99 ± 0.06
GI	1.10	1.09	1.13	1.05
GT	1.12	1.11	1.15	1.09
IT	1.02	1.02	1.02	1.04

TABLA 6.22: Resultados de las simulaciones de Monte Carlo de PCS3 con ángulos y tamaño de bolas modificados.

Caso	Sin V.E.	$\sigma'_{HS} = 0.5\sigma_{HS}$	$\sigma'_{HS} = 2.0\sigma_{HS}$
R_g (Å)	14.04 ± 0.02	20.09 ± 0.04	32.06 ± 0.06
$D_t \times 10^7$ (cm ² /s)	11.08 ± 0.01	8.77 ± 0.01	6.03 ± 0.01
$[\eta]$ (cm ³ /g)	4.85 ± 0.01	8.35 ± 0.02	22.63 ± 0.07
$(R_g)_{corr}$ (Å)	14.47 ± 0.02	20.40 ± 0.04	32.25 ± 0.06
a_T (Å)	19.36 ± 0.01	24.45 ± 0.02	35.57 ± 0.04
a_I (Å)	21.43 ± 0.01	25.69 ± 0.02	35.81 ± 0.04
a_G (Å)	18.68 ± 0.03	26.33 ± 0.05	41.63 ± 0.08
GI	0.87	1.03	1.16
GT	0.96	1.08	1.17
IT	1.11	1.05	1.01

TABLA 6.23: Resultados de las simulaciones de Monte Carlo de PCS3 con volumen excluido modificado.

El resumen de los parámetros obtenidos para los muelles conectores y angulares se presenta en la tabla 6.25. Es importante destacar un dato: mientras en los casos anteriores la constante de fuerza de los muelles conectores era relativamente elevada y la longitud de equilibrio de los muelles era entre una y dos veces el radio de bola, en este caso la longitud de equilibrio es incluso inferior al valor del radio de las bolas del modelo, aunque la constante de fuerza es muy inferior a los casos anteriores, lo que quiere decir que en comparación con los modelos anteriores, los muelles conectores de este modelo serán muy flexibles debido a la mayor longitud del segmento de cadena que representan.

Caso	$D_t \times 10^6$ (cm ² /s)	$R_g \times 10^8$ (cm)	a_T (Å)	a_G (Å)	$\langle \sigma \rangle$ (Å)
1	3.39	5.45	6.48	7.04	6.76
2	3.42	5.40	6.42	6.97	6.70
3	3.17	6.72	6.93	8.68	7.80
4	3.05	7.54	7.20	9.73	8.47
5	3.06	6.84	7.18	8.83	8.00
6	3.42	5.27	6.42	6.80	6.61
7	3.11	6.70	7.06	8.65	7.86
8	3.14	6.75	6.99	8.71	7.85
9	3.45	5.30	6.37	6.84	6.60
10	3.51	5.13	6.26	6.62	6.44
Promedio	3.27	6.11	6.73	7.89	7.31

TABLA 6.24: Cálculo del radio del pseudo-monómero de PCS11 a partir de radios equivalentes de conformaciones extraídas de HYPERCHEM.

Elemento del modelo	Tipo	Parámetros y valores asociados
Radio de bola		$\sigma = 7.31 \text{ \AA}$
Conectores	Hard-FENE	$H = 61 \text{ erg/cm}^2$; $l_e = 5.72 \text{ \AA}$; $l_{max} = 25.00 \text{ \AA}$
Muelles angulares externos	Cuadrático	$Q = 3.07 \times 10^{-14} \text{ erg}$; $\alpha_0 = 3.14 \text{ rad}$
Muelles angulares internos	Cuadrático	$Q = 2.83 \times 10^{-14} \text{ erg}$; $\alpha_0 = 2.67 \text{ rad}$
Volumen excluido	Esfera Dura	$\sigma_{HS} = 7.31 \text{ \AA}$; $V = 8.08 \times 10^{-12} \text{ erg}$

TABLA 6.25: Parámetros para los modelos de PCS11.

Simulación.

Número de bolas	50	100	200	400
M	17639	35276	70550	141098
Submuestras	5	5	5	5
Número de pasos	7.0×10^7	2.0×10^7	5.0×10^6	1.2×10^6
Δx	0.03	0.015	0.005	0.0005
Tasa media de rechazo	23 %	24 %	12 %	8 %

TABLA 6.26: Parámetros de las simulaciones de Monte Carlo de PCS11.

Resultados.

En la tabla 6.27 se recogen los resultados obtenidos de las simulaciones, así como los correspondientes radios equivalentes y las relaciones entre los mismos.

Número de bolas	50	100	200	400
R_g (Å)	31.26 ± 0.07	40.29 ± 0.08	54.03 ± 0.20	69.62 ± 0.17
$D_t \times 10^7$ (cm ² /s)	6.29 ± 0.01	4.70 ± 0.01	3.44 ± 0.01	2.40 ± 0.01
$[\eta]$ (cm ³ /g)	16.19 ± 0.04	18.65 ± 0.05	23.48 ± 0.16	33.06 ± 0.15
$(R_g)_{corr}$ (Å)	31.76 ± 0.07	40.68 ± 0.08	54.32 ± 0.20	69.85 ± 0.17
a_T (Å)	34.11 ± 0.02	45.68 ± 0.04	62.32 ± 0.11	89.53 ± 0.14
a_I (Å)	35.64 ± 0.03	47.07 ± 0.04	64.04 ± 0.12	90.43 ± 0.14
a_G (Å)	41.01 ± 0.09	52.52 ± 0.11	70.13 ± 0.26	90.18 ± 0.22
GI	1.15	1.12	1.10	1.00
GT	1.20	1.15	1.13	1.01
IT	1.04	1.03	1.03	1.01

TABLA 6.27: Propiedades promediadas para cada tamaño de molécula de PCS11.

6.7.5. Polidimetil 5-(4-hidroxibutoxi)isofталato (PDHBI).

Una vez más desconocemos cuáles son los valores apropiados de los parámetros para representar las moléculas reales de PDHBI mediante nuestros modelos, por lo que es necesario repetir el mismo procedimiento de parametrización ya empleado en los casos de PCS3 y PCS11. También como en los casos anteriores, se presenta en primer lugar la figura 6.20, en la que se muestra la estructura de un oligómero de esta especie.

Parametrización.

También en este caso se emplean los parámetros de la tabla 6.15 para la parametrización. El resultado del análisis con HYDROPRO, realizado para la determinación del radio de bola se presenta en la tabla 6.28. Esta vez, los radios a_t y a_G son algo más cercanos que en el caso del PCS11, pero están aún muy lejos del grado de proximidad observado en el caso del PCS3; en cualquier caso, se empleará el promedio como representación del tamaño de bola de este polímero.

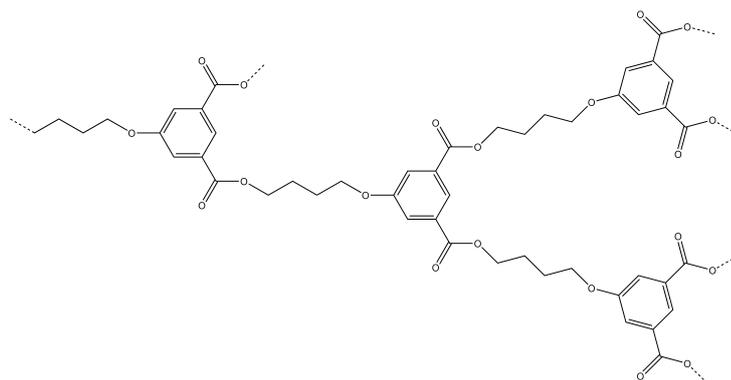


FIGURA 6.20: Ejemplo de estructura de PDHBI.

Caso	$D_t \times 10^6$ (cm ² /s)	$R_g \times 10^8$ (cm)	a_T (Å)	a_G (Å)	$\langle \sigma \rangle$ (Å)
1	5.17	4.42	4.25	5.71	4.98
2	5.26	4.15	4.18	5.36	4.77
3	5.17	4.48	4.25	5.78	5.02
4	5.18	4.21	4.24	5.44	4.84
5	5.27	4.12	4.17	5.32	4.74
6	5.35	3.72	4.11	4.80	4.45
7	5.66	3.21	3.88	4.14	4.01
8	5.61	3.24	3.92	4.18	4.05
9	5.17	4.46	4.25	5.76	5.00
10	5.31	3.84	4.14	4.96	4.55
Promedio	5.32	3.99	4.14	5.14	4.64

TABLA 6.28: Cálculo del radio del pseudo-monómero (bola) de PDHBI a partir de radios equivalentes de conformaciones extraídas de HYPERCHEM.

Hay que destacar que en este caso, en las trayectorias de 100 ns realizadas inicialmente para determinar los parámetros de los muelles, se observaba una serie de progresivos cambios en las conformaciones, que producían tendencias variables en los parámetros de conectores y muelles, por lo que fue necesario prolongar la duración de la simulación hasta los 1000 ns (10^6 ps) para conseguir alcanzar un estado de equilibrio en el que los cambios conformacionales observados fueran lo suficientemente representativos.

Los datos obtenidos en la parametrización de los muelles y ángulos se recogen en la tabla 6.29. En este caso de nuevo se observa un caso similar al anterior, en el que el radio de las bolas tiene prácticamente el mismo valor que la longitud de equilibrio de los muelles, aunque esta vez, el valor de la constante de fuerza es similar a los primeros

casos, tratándose de muelles bastante rígidos. Esto quiere decir que en este caso, existirá un grado de solapamiento importante.

Elemento del modelo	Tipo	Parámetros y valores asociados
Radio de bola		$\sigma = 4.64 \text{ \AA}$
Conectores	Hard-FENE	$H = 566 \text{ erg/cm}^2$; $l_{max} = 10.90 \text{ \AA}$; $l_e = 4.70 \text{ \AA}$
Muelles angulares externos	Cuadrático	$Q = 6.34 \times 10^{-14} \text{ erg}$; $\alpha_0 = 1.86 \text{ rad}$
Muelles angulares internos	Cuadrático	$Q = 3.35 \times 10^{-13} \text{ erg}$; $\alpha_0 = 1.20 \text{ rad}$
Volumen excluido	Esfera Dura	$\sigma_{HS} = 4.64 \text{ \AA}$; $V = 8.08 \times 10^{-12} \text{ erg}$

TABLA 6.29: Parámetros para los modelos de PDHBI.

Simulación.

Número de bolas	50	100	200	400
M	9613	19225	38448	76894
Submuestras	5	5	5	5
Número de pasos	7.5×10^7	2.0×10^7	5.0×10^6	1.2×10^6
Δx	0.03	0.015	0.005	0.0005
Tasa media de rechazo	27 %	25 %	17 %	10 %

TABLA 6.30: Parámetros de las simulaciones de Monte Carlo de PDHBI.

Resultados.

En este caso los parámetros empleados en las simulaciones de los modelos definitivos son los recogidos en la tabla 6.30, mientras que en la tabla 6.31 se recogen los resultados obtenidos de las simulaciones, así como los correspondientes radios equivalentes y las relaciones entre los mismos.

Número de bolas	50	100	200	400
R_g (Å)	17.48 ± 0.04	22.41 ± 0.05	33.75 ± 0.15	50.78 ± 0.14
$D_t \times 10^7$ (cm ² /s)	10.59 ± 0.01	8.04 ± 0.01	5.61 ± 0.01	3.35 ± 0.01
$[\eta]$ (cm ³ /g)	6.30 ± 0.02	6.98 ± 0.02	10.42 ± 0.06	21.87 ± 0.12
$(R_g)_{corr}$ (Å)	17.85 ± 0.04	22.69 ± 0.05	33.94 ± 0.14	50.91 ± 0.14
a_T (Å)	20.26 ± 0.02	26.68 ± 0.02	38.24 ± 0.07	64.01 ± 0.12
a_I (Å)	21.25 ± 0.02	27.71 ± 0.03	39.90 ± 0.08	64.36 ± 0.12
a_G (Å)	23.04 ± 0.05	29.30 ± 0.06	43.82 ± 0.19	65.73 ± 0.18
GI	1.08	1.06	1.10	1.02
GT	1.14	1.10	1.15	1.03
IT	1.05	1.04	1.04	1.01

TABLA 6.31: Propiedades promediadas para cada tamaño de molécula de PDHBI.

6.8. Discusión de resultados.

6.8.1. Variación de las propiedades con el peso molecular.

A partir de los resultados obtenidos para cada uno de los polímeros se ha intentado obtener leyes de escala que relacionen las propiedades estudiadas (radio de giro, coeficiente de difusión traslacional y viscosidad intrínseca) con el peso molecular (proporcional al tamaño del modelo). Los parámetros de estas leyes de escala se presentan en la tabla 6.32. Los parámetros cuyos valores se muestran en la tabla son los correspondientes a unas leyes de escala de tipo potencial

$$Propiedad = K \times M^\alpha$$

En las figuras 6.21, 6.22 y 6.23 se presentan en gráficas doble-logarítmicas las propiedades calculadas frente al peso molecular para obtener los parámetros K y α de las leyes de potencia. En el caso de estas propiedades, se encuentran leyes de potencia que ajustan bien algunos de los resultados obtenidos mediante simulación, principalmente en el caso del radio de giro. Sin embargo, los resultados correspondientes al PDHBI y el PAMAM

Polímero	Propiedad	α	K
PBzE	R_g (Å)	0.336	1.00
	D_t (cm ² /s)	-0.409	3.95×10^{-5}
	$[\eta]$ (cm ³ /g)	0.205	1.33
PAMAM	R_g (Å)	0.194	4.08
	D_t (cm ² /s)	-0.317	1.50×10^{-5}
	$[\eta]$ (cm ³ /g)	-0.054	20.7
PCS3	R_g (Å)	0.323	1.19
	D_t (cm ² /s)	-0.393	3.10×10^{-5}
	$[\eta]$ (cm ³ /g)	0.150	3.04
PCS11	R_g (Å)	0.382	0.75
	D_t (cm ² /s)	-0.449	5.13×10^{-5}
	$[\eta]$ (cm ³ /g)	0.327	0.64
PDHBI	R_g (Å)	0.508	0.16
	D_t (cm ² /s)	-0.494	10.1×10^{-5}
	$[\eta]$ (cm ³ /g)	0.542	0.04

TABLA 6.32: Parámetros de las leyes de escala calculadas para los diferentes polímeros estudiados.

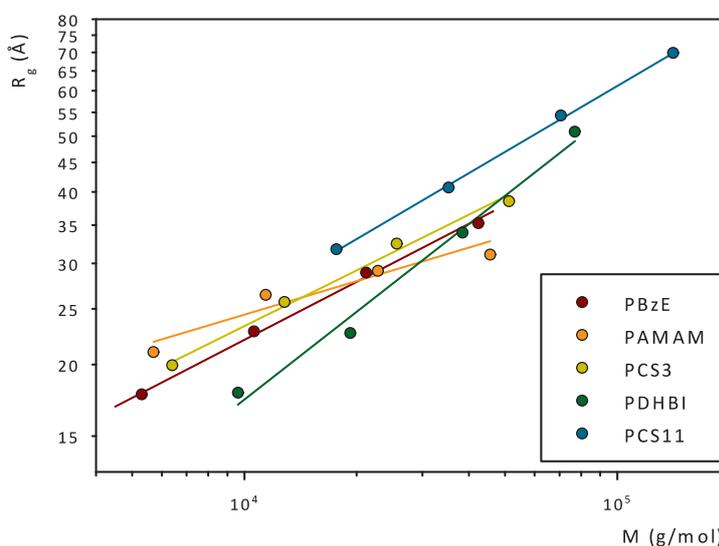


FIGURA 6.21: Representación en escala logarítmica de la ley de escala para el radio de giro.

claramente no se ajustan bien a una expresión como la propuesta. Además, también parece observarse cierta curvatura en los datos del coeficiente de difusión de PCS11 y en

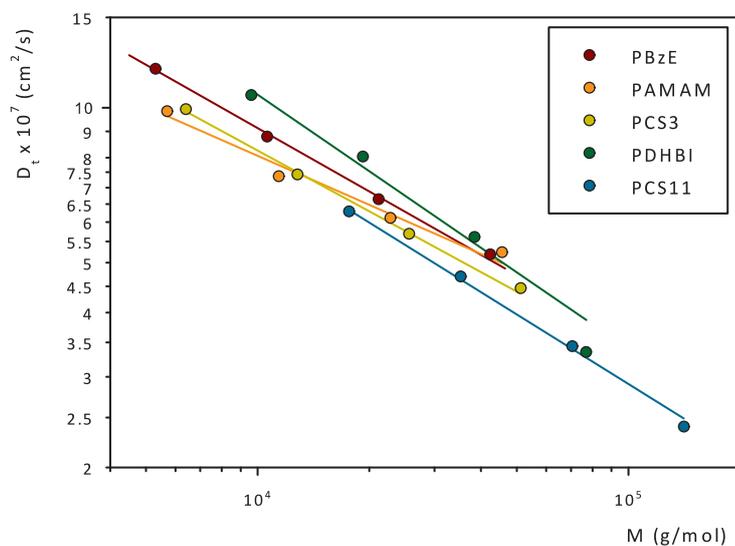


FIGURA 6.22: Representación en escala logarítmica de la ley de escala para difusión traslacional.

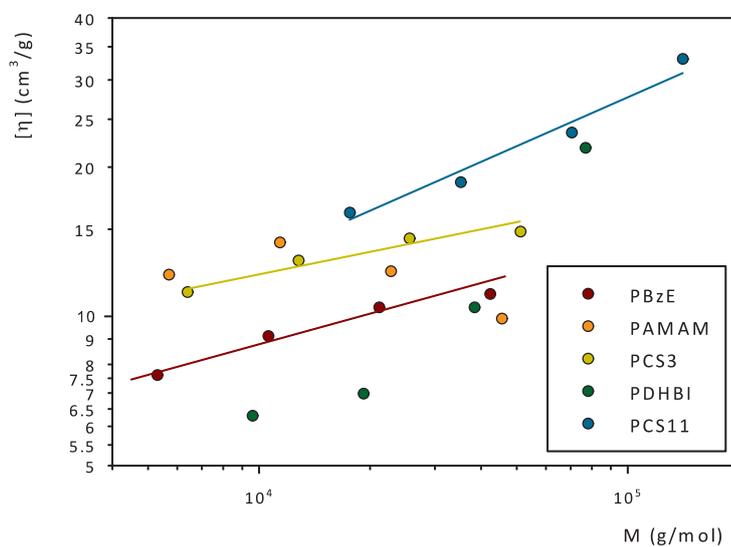


FIGURA 6.23: Representación en escala logarítmica de la viscosidad intrínseca.

los de radio de giro de PCS3, lo que indica que, aunque en el rango de pesos moleculares estudiados la desviación no es excesiva, en estos casos la variación de estas propiedades tampoco responde a las leyes de escala propuestas.

Finalmente, en el caso de la viscosidad intrínseca se aprecia claramente que en todos los casos existe un cierto grado de curvatura, por lo que ninguna de las leyes de escala calculadas sería estrictamente aplicable.

6.8.2. Función de distribución de radios de giro.

Para cada uno de los tipos de polímero estudiados, se ha analizado la distribución de los radios de giro de las diferentes topologías generadas para cada caso (entendiendo por “radio de giro de una topología” el valor promedio de este parámetro en la trayectoria simulada).

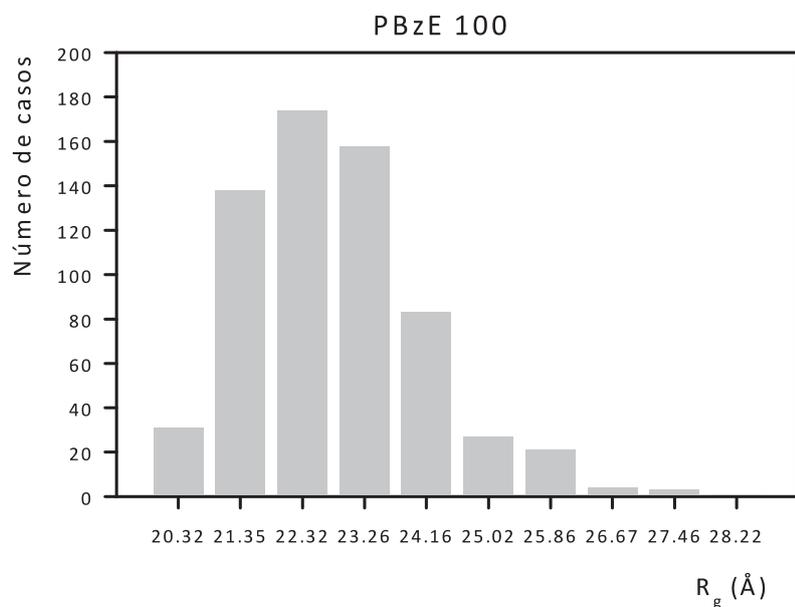


FIGURA 6.24: Función de distribución del radio de giro en diferentes topologías de PBzE-100.

A la vista de los valores presentados en las secciones anteriores, así como de sus respectivos errores, se esperaba que las funciones de distribución fueran bastante estrechas, con desviaciones estándar bastante bajas. Como se puede observar en la figura 6.24, presentada como ejemplo, y en la tabla 6.33, que recoge los datos de todas las simulaciones, se obtuvieron funciones de distribución con una desviación estándar inferior al 10% en prácticamente todos los casos, lo que confirma que nuestro algoritmo es capaz de generar cadenas con una variación pequeña en el grado de ramificación.

Polímero	N	$\langle R_g \rangle_{\text{topol}}$ (Å)	σ_{topol} (Å)	σ_{topol} (%)
PAMAM	50	21.01	1.12	5.33
	100	26.44	1.19	4.51
	200	29.11	1.63	5.58
	400	31.21	2.00	6.40
PBzE	50	17.73	1.18	6.63
	100	22.83	1.38	6.04
	200	28.91	2.04	7.06
	400	35.26	3.42	9.70
PCS3	50	19.93	1.31	6.55
	100	25.70	1.70	6.60
	200	32.48	2.87	8.83
	400	38.52	3.21	8.32
PCS11	50	31.77	1.72	5.43
	100	40.68	2.12	5.22
	200	54.33	5.01	9.22
	400	69.85	4.28	6.13
PDHBI	50	17.85	1.02	5.70
	100	22.69	1.17	5.15
	200	33.94	3.66	10.77
	400	50.91	3.59	7.04

TABLA 6.33: Promedios topológicos del radio de giro y desviaciones estándar.

6.8.3. Comparación con los datos experimentales.

PBzE.

No disponemos de datos experimentales de polímeros hiperramificados de este compuesto, por lo que no podemos hacer ninguna clase de comparación.

PAMAM.

De este tipo de polímeros hiperramificados sólo disponemos de los datos de viscosidad intrínseca de un polímero hiperramificado de PAMAM con *core* B_6 , obtenidos a partir del

trabajo de Hobson y col. [76]. Estos datos estaban disponibles únicamente en forma de gráfico, por lo que para la comparación hemos superpuesto al mismo los datos obtenidos en nuestras simulaciones. El resultado puede apreciarse en la figura 6.25.

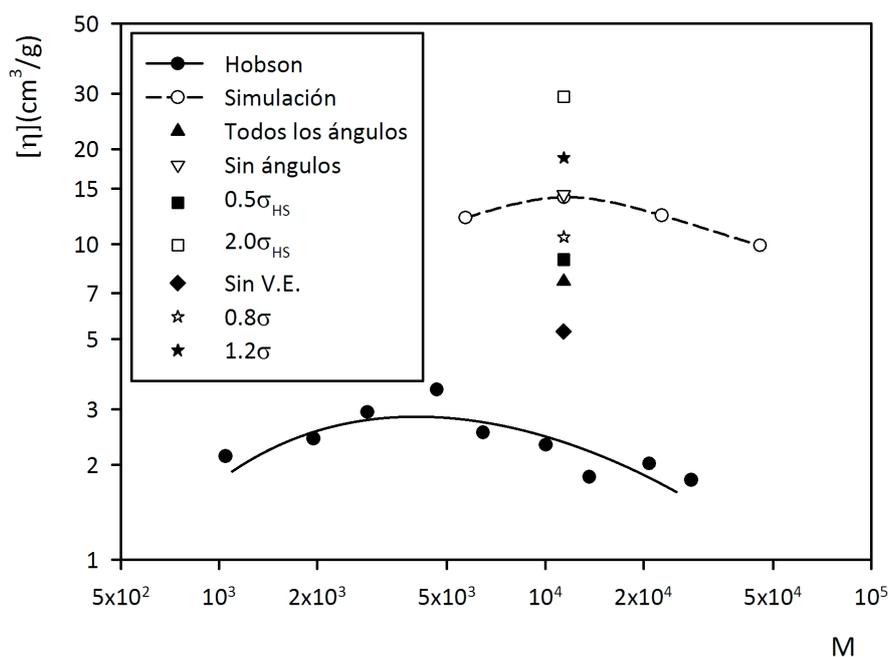


FIGURA 6.25: Comparativa de datos simulados y experimentales de PAMAM.

Ambos juegos de datos coinciden en la existencia de un máximo de viscosidad intrínseca, aunque en los resultados de las simulaciones éste se encuentra desplazado hacia pesos moleculares mayores. Tampoco el rango de valores de la viscosidad intrínseca obtenido coincide, dado que los resultados de la simulación son del orden de $10 \text{cm}^3/\text{g}$ superiores a los observados por Hobson y col.

Incluso los datos obtenidos a partir de las simulaciones con parámetros modificados, realizadas sobre un PAMAM de $N = 100$, continúan siendo bastante superiores a los experimentales, lo que hace pensar que las diferencias observadas no se deben únicamente a una parametrización inadecuada.

PCS3.

De este tipo de polímero encontramos datos experimentales en los artículos de Tarabukina y col., [77, 87], especialmente en el segundo. En él aparecen datos experimentales de viscosidad intrínseca y coeficientes de difusión traslacional, tanto de PCS3 como de PCS11. Sin embargo, dado que los ensayos están hechos utilizando hexano como disolvente y las simulaciones se han parametrizado con datos del agua, es conveniente ignorar los datos de difusividad y utilizar los de radio hidrodinámico, también recogidos en el artículo. Así, se elaboran los gráficos 6.26 y 6.27.

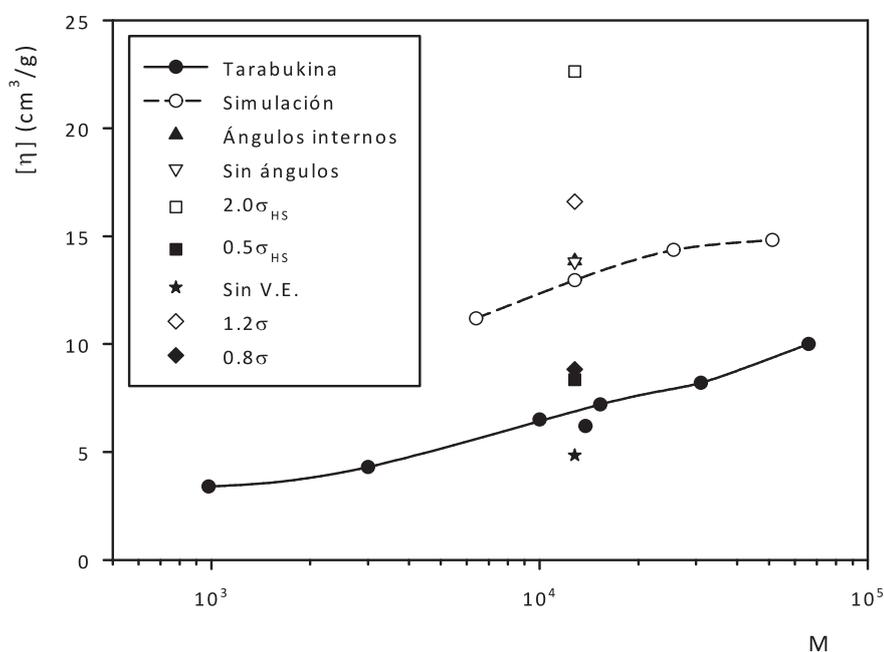


FIGURA 6.26: Comparativa de viscosidades intrínsecas simuladas y experimentales de PCS3.

También en este caso observamos lo mismo que en el caso del PAMAM: los datos, en todo su conjunto, son superiores a los experimentales, aunque en este caso, los valores obtenidos son más próximos entre sí, y en el caso de las simulaciones en las que se variaron los valores de parámetros del modelo, en algunos casos llegan a alcanzarse valores similares a los experimentales.

Sin embargo, aunque en la viscosidad intrínseca se observa una tendencia de aumento de esta propiedad frente al peso molecular similar a los datos de laboratorio, en el caso del radio hidrodinámico (del que sólo se dispone de tres datos experimentales), la tendencia

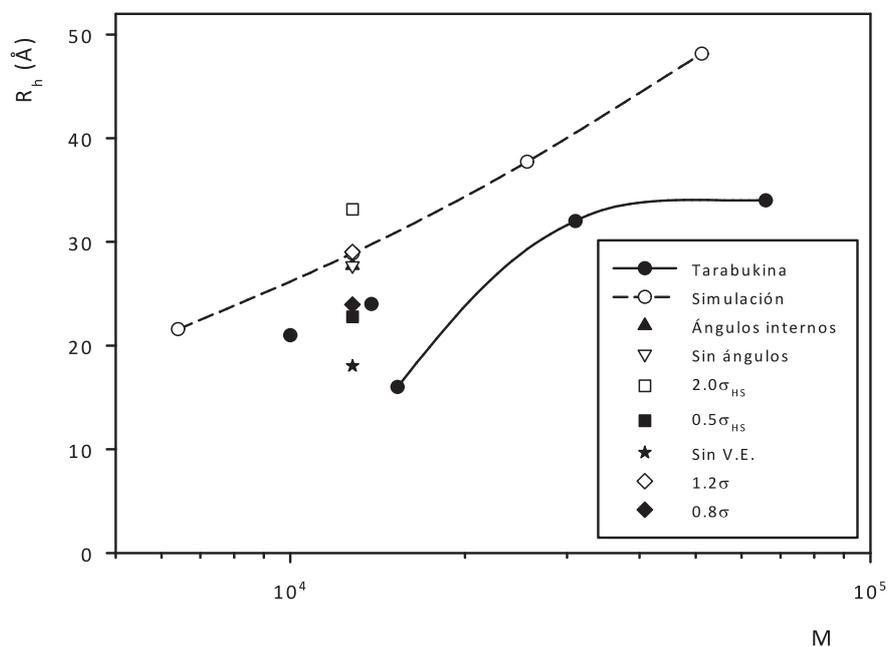


FIGURA 6.27: Comparativa de radios hidrodinámicos simulados y experimentales de PCS3.

observada diverge bastante de la de los datos experimentales, por lo que en conjunto, también en este caso debemos admitir que el resultado de las simulaciones no es bueno.

PCS11.

De nuevo volvemos a comparar nuestros datos experimentales con los publicados por Tarabukina y col. [87].

En esta ocasión los de viscosidad intrínseca presentados en la figura 6.28 suponen la mejor aproximación a los datos experimentales en cuantas simulaciones se han hecho: los datos correspondientes a los menores pesos moleculares son muy próximos a los experimentales, mientras que el dato obtenido para 400 bolas está algo más desviado, pero en general la correspondencia de las tendencias frente al peso molecular es aceptable.

Sin embargo, en el caso del radio hidrodinámico se observa, que, aunque las tendencias también mantienen un cierto parecido, el valor de las propiedades simuladas vuelve a discrepar hasta casi en un 50 % del valor experimental.

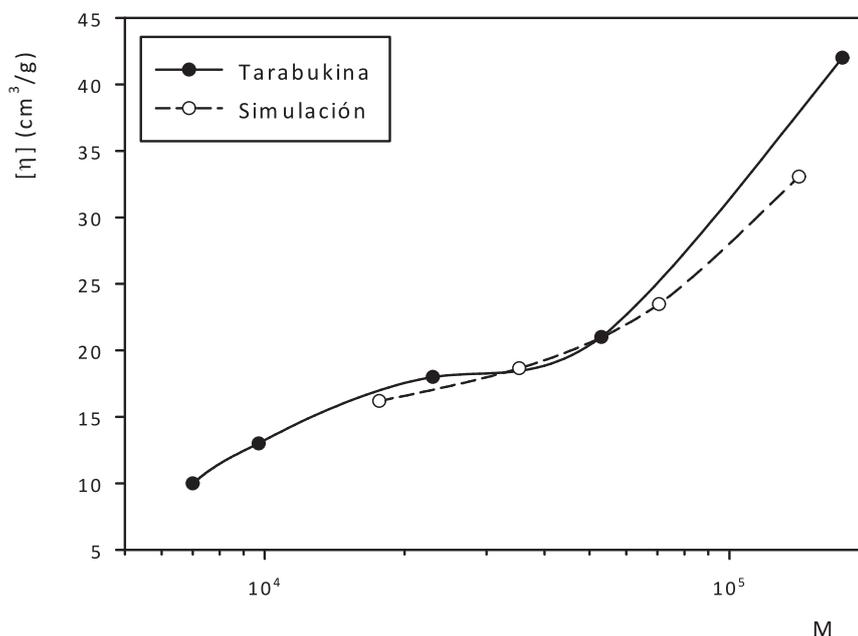


FIGURA 6.28: Comparativa de viscosidades intrínsecas simuladas y experimentales de PCS11.

PDHBI.

Obtenemos los datos experimentales de este polímero en los artículos de De Luca y col. [70] y [71]. En el primero disponemos de datos de las tres propiedades que calculamos en los polímeros simulados (radio de giro, coeficiente de difusión y viscosidad intrínseca), obtenidas a partir de disoluciones en cloroformo mediante cromatografía de exclusión y dispersión de luz, mientras que en el segundo sólo disponemos de dos aproximaciones diferentes al radio de giro obtenidas por dispersión de neutrones a partir de disoluciones en tetrahidrofurano.

El propio De Luca admite, en las conclusiones del segundo artículo, las diferencias en los radios de giro de moléculas del mismo peso molecular en uno y otro artículo, y lo atribuye a que el tetrahidrofurano no es tan buen disolvente como el cloroformo, por lo que, en concordancia con las condiciones empleadas para las simulaciones, preferiremos los resultados obtenidos en el primer artículo. Con ellos realizamos los gráficos de las figuras 6.30, 6.31 y 6.32, de nuevo comparando datos de radio hidrodinámico en lugar del coeficiente de difusión.

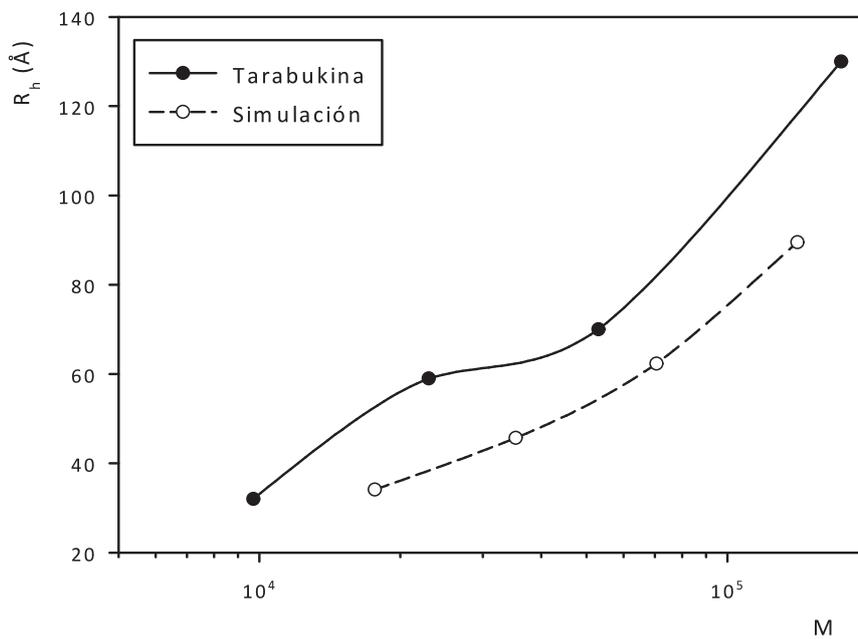


FIGURA 6.29: Comparativa de radios hidrodinámicos simulados y experimentales de PCS11.

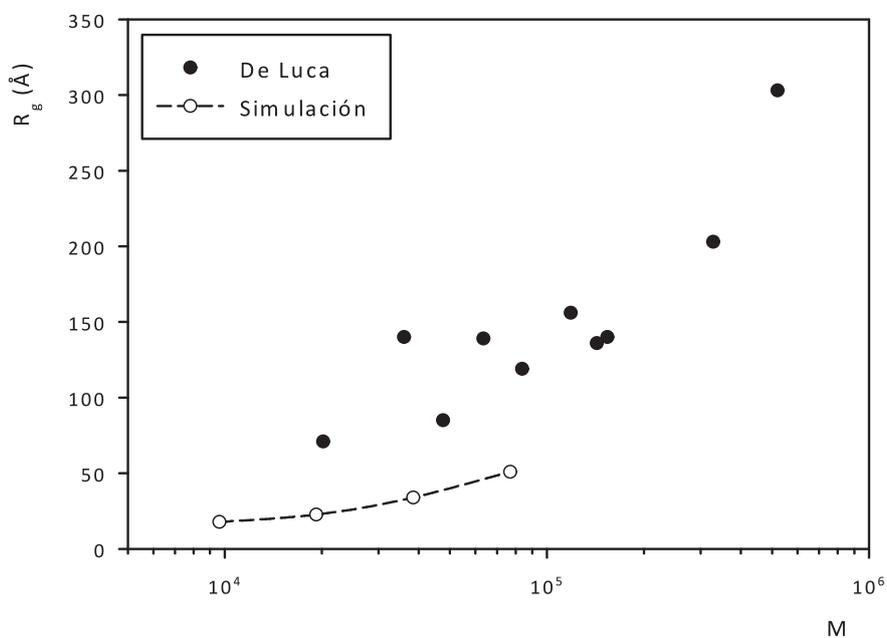


FIGURA 6.30: Comparativa de radios de giro simulados y experimentales de PDHBI.

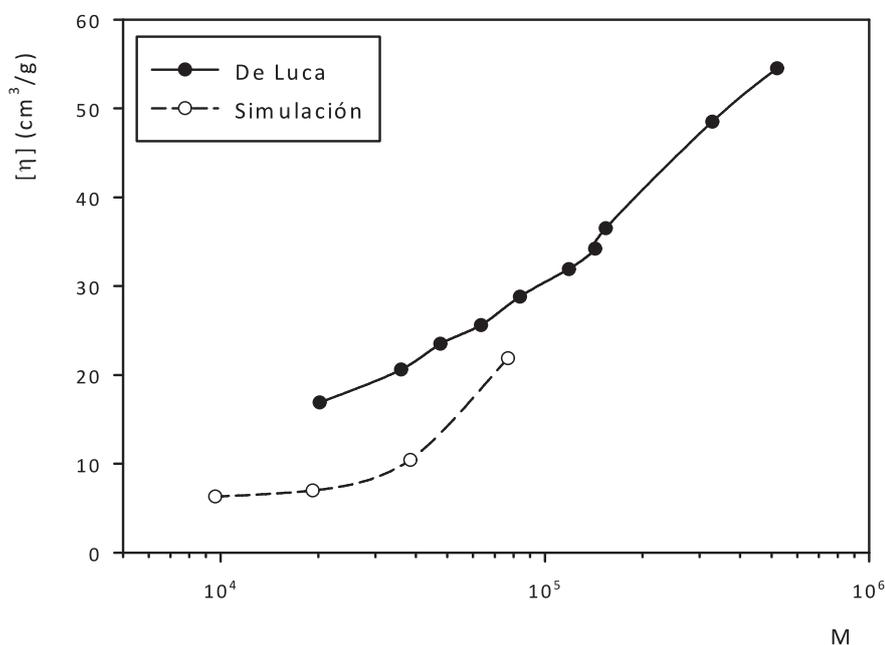


FIGURA 6.31: Comparativa de viscosidades intrínsecas simuladas y experimentales de PDHBI.

Es en este polímero donde se observan las mayores diferencias entre los valores simulados y los experimentales: se llega a un 70 % de error en el radio de giro, un 50 % en la viscosidad intrínseca, y casi un 60 % en el radio hidrodinámico, y aunque las tendencias de los radios de giro e hidrodinámicos siguen una tendencia similar a los datos experimentales, la de la viscosidad intrínseca ya presenta una forma claramente diferente, aunque ambas sean monótonamente crecientes.

6.9. Conclusiones del capítulo.

Vistas las importantes discrepancias entre los valores de las propiedades obtenidos a partir de las simulaciones y los datos experimentales publicados por otros autores, sólo podemos concluir que, en base a las suposiciones adoptadas para el desarrollo de este trabajo, y a pesar de los buenos resultados producidos por este procedimiento en su

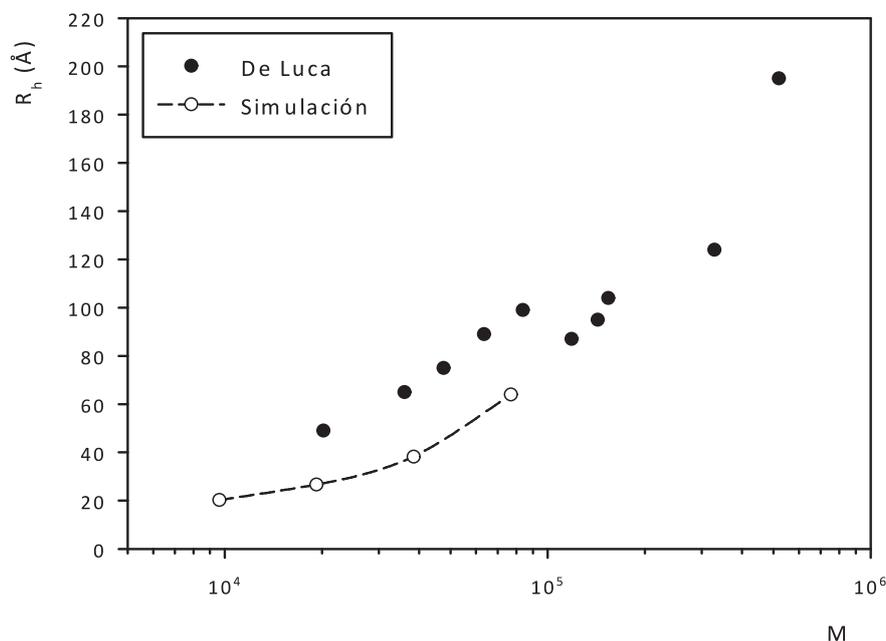


FIGURA 6.32: Comparativa de radios hidrodinámicos simulados y experimentales de PDHBI.

uso para la simulación de dendrímeros [24], el procedimiento de simulación multiescala empleado no es válido para la simulación de polímeros hiperramificados.

Sin embargo, a pesar de que los valores de las simulaciones y las de los resultados experimentales no coinciden, la tendencia general sí suele coincidir: en los casos en los que se observa un máximo en una propiedad experimental (como la viscosidad intrínseca del PAMAM), también se observa un máximo en los resultados de la simulación, y viceversa, aumentando o disminuyendo monótonamente el valor de las propiedades simuladas al hacerlo también las experimentales.

Durante el transcurso de este trabajo se han introducido variaciones en los parámetros de algunas de las simulaciones (PAMAM y PCS3 de $N = 100$) para intentar obtener resultados más próximos a los experimentales. Las variaciones introducidas respondían a cambios no excesivamente drásticos en los parámetros del tamaño de bola del modelo y el volumen excluido, y la consideración de diferentes combinaciones de las restricciones angulares, de manera que se ensayaban valores para estos parámetros que hubieran sido plausibles en caso de emplear diferentes criterios para la parametrización.

Los resultados de estas pruebas han sido dispares, pues mientras para el caso del PAMAM no se ha conseguido reproducir el mismo orden de magnitud de los datos experimentales de viscosidad intrínseca que se dispone, en el caso del PCS3 sí que se ha conseguido obtener algunos resultados similares a los experimentales.

Una observación importante es que, mientras que en los polímeros cuyos pseudo-monómeros son pequeños y relativamente rígidos (PAMAM y PCS3) los resultados de las simulaciones están por encima de los resultados experimentales, los resultados de aquellos en los que los pseudo-monómeros constan de brazos más largos, están por debajo de los valores experimentales.

Una posible opción que barajamos como causa de los malos resultados de las simulaciones es la baja representatividad estadística de las muestras generadas, pues a pesar de haber simulado un número elevado de topologías para cada tamaño de modelo de cada polímero, la gama de posibles topologías a las que se puede dar lugar en una reacción de tipo *one-pot*, como las que se utilizan para la síntesis de este tipo de polímeros, es muy amplia, y nuestra muestra de 640 topologías para cada tamaño de modelo no es lo suficientemente representativa.

Esta opción cobra más fuerza después de observar los buenos resultados obtenidos en el Capítulo 7, en el que se emplea el mismo esquema de simulación multiescala, aunque de nuevo, como en el caso de los dendrímeros, la topología de los modelos simulados está bien definida (se trata de cadenas lineales). Por este motivo podemos deducir que la variabilidad topológica supone un importante factor de error en este procedimiento.

Otro factor con cierta importancia es la discrepancia entre las topologías de los polímeros reales y los aquí generados. En las referencias de las que se han tomado los datos experimentales en los que se ha basado este trabajo no figuran datos concretos de los índices topológicos de las muestras, y las topologías generadas lo han sido sobre un valor intermedio del grado de ramificación (DB). Sin embargo, no hay ninguna garantía de que los polímeros a los que corresponden nuestras referencias tengan siquiera una topología similar, y ésta puede tener cierta influencia sobre las propiedades observadas, según la molécula se asemeje más a los casos límite que suponen una cadena lineal o un dendrímero perfectamente regular.

CAPÍTULO 7

SIMULACIÓN DE COPOLÍMEROS ANFIFÍLICOS TERMOSENSIBLES

7.1. Introducción.

Durante las últimas décadas, la investigación de copolímeros termosensibles compuestos por bloques anfifílicos han recibido bastante atención por parte de la comunidad científica [88–92]. Debido a que este tipo de sistemas posee una fuerte tendencia a producir estructuras con capacidad de autoensamblaje, dependiente de la temperatura, se ha encontrado una amplia variedad de aplicaciones en múltiples campos, especialmente, en el transporte de fármacos y las aplicaciones biomédicas [92, 93].

Típicamente estos sistemas están compuestos por un bloque termosensible con propiedades que responden a ciertos estímulos y un bloque menos sensible, o sin sensibilidad alguna a dichos estímulos. Habitualmente, ante un aumento en la temperatura, el bloque termosensible responde volviéndose más hidrofóbico, y desarrollando una tendencia a evitar el contacto con el agua. Como consecuencia de ello, en medio acuoso, con frecuencia se forman estructuras autoensambladas de tipo *core-shell*, en los que en el núcleo se localizan los bloques hidrofóbicos de las cadenas, mientras que los segmentos no sensibles forman la corona hidrofílica.

Este tipo de sistemas tienen utilidad en una multitud de campos [92–94], por lo que su estudio en detalle es interesante. Sin embargo, con frecuencia, si la concentración es

lo suficientemente alta, este proceso de autoensamblaje ya se produce incluso por debajo de la conocida como “temperatura crítica mínima de disolución” (o “LCST”, por las siglas en inglés de *lower critical solution temperature*). De esta manera, se entiende que realizar estudios experimentales del comportamiento de moléculas aisladas de este tipo de compuestos es algo complejo, a la vez que es algo fundamental para el diseño de nuevos tipos de copolímeros termosensibles.

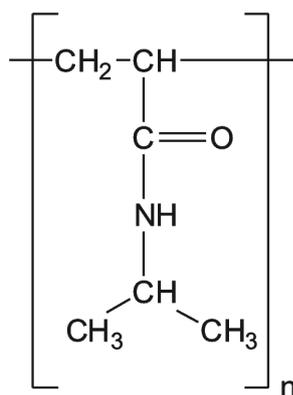


FIGURA 7.1: Estructura del PNIPAAm.

Un polímero termosensible cuyas disoluciones acuosas son ampliamente estudiadas debido a sus aplicaciones en nanotecnología, electrónica y transporte de fármacos es la poli(N-isopropilacrilamida) (PNIPAAm). Este material posee una LCST entre 30 y 35 °C [94], aunque este rango de temperaturas puede variar en función del peso molecular, la concentración (en el caso de polímeros de bajo peso molecular [95]), e incluso la velocidad a la que se produce el calentamiento [96, 97]. Por debajo de esta temperatura, el polímero es soluble en agua, pero por encima de la LCST, la hidrofobicidad de las moléculas aumenta y en el sistema tiende a producirse una separación de fases macroscópica. Para cadenas de PNIPAAm de alto peso molecular en disolución diluida, se ha observado que al aumentar la temperatura, el polímero pierde su carácter hidrofílico, la estructura colapsa, y pasa de una conformación en ovillo al azar a la formación de glóbulos [98].

Dada la dificultad de realizar estudios experimentales sobre el comportamiento de las moléculas individualizadas de copolímeros de este tipo, la simulación se presenta como una valiosa herramienta para ayudar a esclarecer dichos comportamientos, algo de suma importancia en el diseño de copolímeros de este tipo. Con este objetivo, se llevaron a cabo diferentes tipos de simulaciones.

En primer lugar, para determinar la relación entre el tamaño de cada uno de los bloques y las dimensiones y la contracción experimentada por los copolímeros de este tipo, se simularon polímeros dibloque de PNIPAAm y metoxipolietilenglicol (MPEG),

un polímero hidrofílico sin sensibilidad térmica aparente. En concreto, se modelaron dos series de polímeros: la primera, teniendo un bloque de PNIPAAm de un tamaño constante de 71 monómeros, y bloques de MPEG de tamaño variable, concretamente, 12, 23 o 114 monómeros, además de una simulación sin bloque de MPEG. Finalizado ese trabajo, se simuló también una segunda serie de modelos en la que era el bloque de MPEG el que tenía un tamaño constante de 53 monómeros, y era el bloque de PNIPAAm el que tenía tamaños variables de 19, 70 y 114 monómeros.

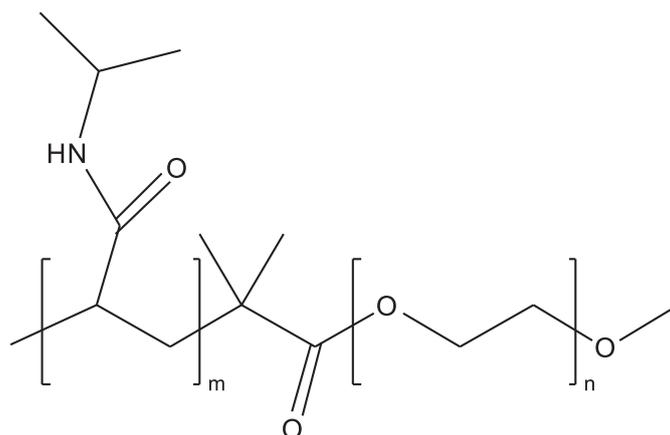


FIGURA 7.2: Estructura general de los copolímeros dibloque PNIPAAm-MPEG.

La elección de los tamaños de los bloques de los polímeros no es aleatoria, sino que se corresponden a los tamaños de los bloques de polímeros sintetizados en el laboratorio del profesor Nyström de la Universidad de Oslo, con quien se ha colaborado en el trabajo correspondiente a este Capítulo de la Tesis.

Sobre estas dos series de polímeros se realizaron simulaciones de Monte Carlo en un rango de temperaturas entre los 25 y los 32 °C, lo que permite establecer la influencia del tamaño de bloque de ambos tipos de monómero sobre la contracción inducida por la temperatura, por medio del estudio de las distancias extremo-extremo (r_i) de cada bloque del copolímero, el radio de giro, R_g , y el radio hidrodinámico, R_h , de la molécula global.

En segundo lugar, una vez estudiadas las propiedades de los copolímeros dibloque, se realizaron nuevas simulaciones, en esta ocasión, de copolímeros tribloque de tipo PNIPAAm-MPEG-PNIPAAm. En este caso, el objetivo consiste en el estudio de la influencia de la longitud del conector central de MPEG en el comportamiento asociativo inducido mediante la temperatura. Para este propósito, se construyeron nuevos modelos empleando los parámetros ya obtenidos para los copolímeros dibloque. Estos nuevos modelos consisten en una serie de moléculas en los que los bloques de PNIPAAm de los extremos

mantienen un tamaño aproximadamente constante, mientras que el conector central de MPEG se hace cada vez más largo. También en este caso, y de nuevo gracias a la colaboración con el grupo del profesor Nyström, se sintetizaron y caracterizaron algunos polímeros similares a los simulados, y se dispone de datos experimentales.

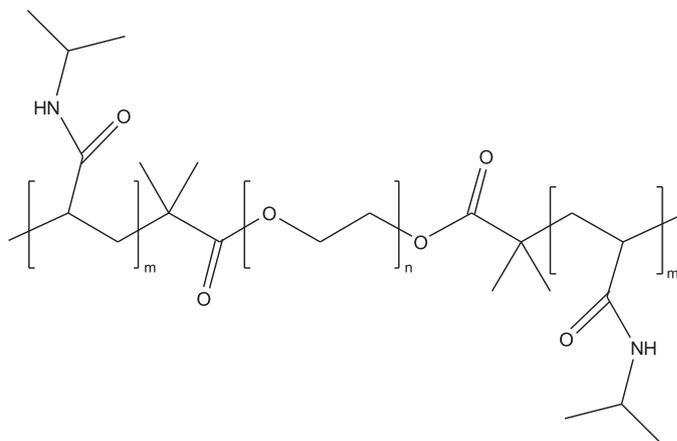


FIGURA 7.3: Estructura general de los copolímeros tribloque PNIPAAm-MPEG-PNIPAAm.

7.2. Modelo y parametrización.

Tanto los elementos del modelo como las metodologías de parametrización empleados en este caso son similares a los descritos en la sección 2.3, y que ya han sido empleados en el Capítulo dedicado a la simulación de polímeros hiperramificados (Capítulo 6), aunque en este caso, al tratarse de moléculas dibloque lineales, no se generarán topologías aleatorias. Dado que, tanto la técnica empleada como los elementos del modelo ya han sido descritos en el Capítulo 2, en esta sección sólo se describirán brevemente algunos de los detalles del proceso, haciendo hincapié únicamente en aquellos aspectos que se diferencien de los correspondientes a la parametrización de los polímeros hiperramificados. Además, dado que tanto los modelos de polímeros dibloque como los tribloque están formados por los mismos elementos, sólo será necesario realizar la parametrización una vez.

7.2.1. Parametrización de las bolas.

Una de las principales diferencias respecto a la simulación de estos copolímeros en bloques respecto a los hiperramificados consiste en que, a diferencia de los segundos, en el caso actual, al existir dos tipos de monómeros diferentes en los polímeros, será necesario modelar dos tipos de bolas diferentes. Además, teniendo en cuenta el tamaño de los monómeros implicados y la ausencia de ramificaciones, y considerando la premisa de que en las simulaciones es conveniente emplear el menor número de elementos posible (para reducir el tiempo de proceso necesario), se considerará que, para ambos tipos, una bola del modelo representará dos monómeros del polímero. Teniendo en cuenta este criterio, la equivalencia de cada uno de los tipos de bola sería similar a la que se presenta en la figura 7.4.

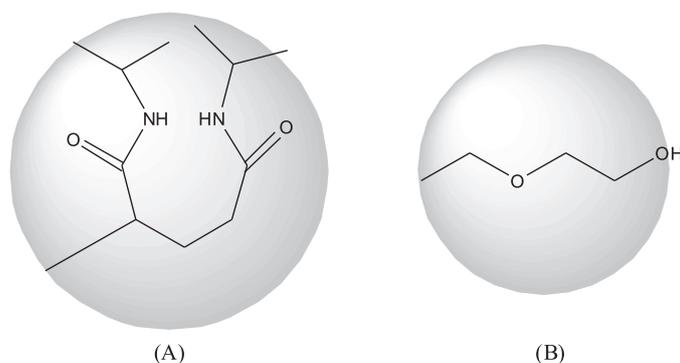


FIGURA 7.4: Estructura de los dímeros correspondientes a cada una de las bolas del modelo: (A) Bola de PNIPAA. (B) Bola de MPEG.

Al establecer la condición anterior, conviene tener en cuenta que podría darse el caso de que el número de monómeros en alguno (o en ambos) de los bloques de los copolímeros fuese impar; en este caso, al no poder representar un único monómero (habría que trabajar con “media bola”, o definir nuevos tipos de bolas), se añadirá una bola al modelo, representando un monómero de más (es decir, se redondea el número de bolas hacia arriba, hacia el siguiente número par). Sin embargo, como se detallará más adelante en la sección de resultados, teniendo en cuenta que el mínimo número de monómeros en un bloque será de 12 (que además es un número par y no presentará este problema), el hecho de trabajar en la simulación con un monómero de más no debe suponer una gran diferencia en las conclusiones que se obtengan de este trabajo.

Una vez tomada la decisión de a qué entidades representaremos con las bolas del modelo, se procede a parametrizar el radio de las bolas. Esta parametrización, como ya se describió en la sección 2.3.1, se realiza mediante simulaciones atomísticas, empleando

el programa HYPERCHEM, simulando por separado cada uno de los dímeros de PNIPAAm y MPEG. En esta ocasión, para generar las trayectorias tanto en el caso de las bolas de MPEG como las de PNIPAAm, se emplean los parámetros reflejados en la tabla 7.1.

Campo de fuerzas	AMBER
Duración de trayectoria	20000 ps (20 ns)
Tamaño de paso	0.002 ps
Factor de fricción	50 ps ⁻¹
Temperatura	298 K
Frecuencia de adquisición de datos	5000 pasos
Número de conformaciones adquiridas	2000

TABLA 7.1: Parámetros de la dinámica de Langevin para ambos tipos de bolas.

Como se hizo con los polímeros hiperramificados, las trayectorias obtenidas se analizan a su vez con el programa HYDROPRO, y se obtienen así para cada tipo de bola los valores promedio del radio de Stokes (a_T) y radio de giro (R_g), y a partir de este último, el correspondiente radio equivalente, a_G . Finalmente, ambos radios se promedian para obtener el valor que se empleará como radio de cada uno de los tipos de bola, $\langle\sigma\rangle$. Los valores de estos parámetros, así como los radios promedio obtenidos se reflejan en la tabla 7.2, y como puede observarse, para ambas especies, los valores de a_T y a_G son bastante próximos, por lo que la aproximación de usar los valores promedio es bastante acertada.

Caso	$D_t \times 10^6$ (cm ² /s)	$R_g \times 10^8$ (cm)	a_T (Å)	a_G (Å)	$\langle\sigma\rangle$ (Å)
Bola PNIPAAm	5.03	3.74	4.37	4.83	4.60
Bola MPEG	7.24	2.58	3.03	3.33	3.18

TABLA 7.2: Cálculo de los radios de bola a partir de los resultados de HYDROPRO.

7.2.2. Parametrización de los muelles.

Los muelles que conecten las bolas adyacentes representarán un número pequeño de átomos, por lo que se puede asumir que serán uniones relativamente rígidas, por lo que para la construcción de los modelos se emplearán muelles de tipo “hard-FENE”, apropiados para representar uniones rígidas y con una extensibilidad finita. La expresión para el potencial de un muelle de este tipo se presentaba en la sección 2.2.

También en este caso la presencia de dos especies químicas representa una diferencia respecto al caso de parametrización de los polímeros hiperramificados. Si para el caso de la parametrización de las bolas era necesario distinguir la parametrización de dos tipos de bolas diferentes, para el caso de los conectores será necesario distinguir entre tres tipos diferentes de muelles: aquellos que conectan dos bolas de PNIPAAM, los que conectan dos bolas de MPEG, y finalmente, los que conectan dos bolas de tipo diferente, es decir, la unión MPEG-PNIPAAM. Mientras que de los dos primeros tipos habrá varios muelles en cada modelo, del segundo tipo solamente habrá uno en el caso de los polímeros dibloque, en la unión entre los dos bloques, y dos en los polímeros tribloque, en la unión de cada extremo de PNIPAAM al bloque central de MPEG.

Otra diferencia importante en el modelado de los conectores respecto al que se hizo en los polímeros hiperramificados es que, debido a que las moléculas que se van a simular son moléculas flexibles lineales, no se emplearán restricciones angulares. De esta manera, no es necesario parametrizar los muelles angulares, y el modelo simulado con **HYPERCHEM** puede ser más simple, ya que no es necesario definir ángulos sobre el mismo; así, en lugar de la estructura de cuatro bolas en forma de “Y” que se empleaba en el caso de los hiperramificados, en el caso de los copolímeros termosensibles sólo es necesario modelar tres bolas (seis monómeros, recordemos que cada bola representa a dos monómeros) entre las que se definen dos conectores, como se representa en la figura 7.5, para el caso de conectores “puros” entre PNIPAAM y MPEG, y en la figura 7.6 para un conector mixto MPEG-PNIPAAM.

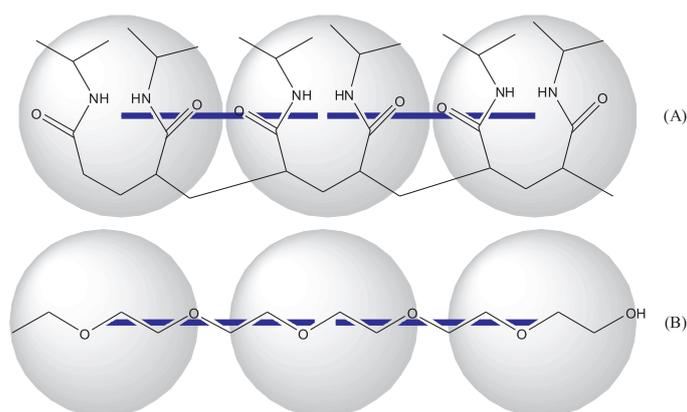


FIGURA 7.5: Correspondencia aproximada entre bolas, muelles y estructura simulada: (A) PNIPAAM. (B) MPEG.

En estos casos, las condiciones de la simulación en **HYPERCHEM** para los muelles de PNIPAAM y MPEG son las presentadas en la tabla 7.3. Los mismos datos se emplean también para el muelle que conecta los dos bloques con la diferencia de que en ese caso, la longi-

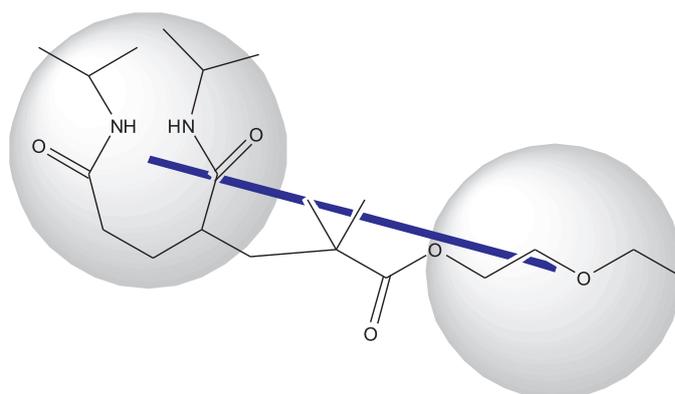


FIGURA 7.6: Estructura simulada para la unión entre los bloques del copolímero.

tud de la trayectoria es de un millón de pasos (1000 ns), y el número de conformaciones adquiridas es de 100000.

Campo de fuerzas	AMBER
Duración de trayectoria	100000 ps (100 ns)
Tamaño de paso	0.002 ps
Factor de fricción	50 ps ⁻¹
Temperatura	298 K
Frecuencia de adquisición de datos	5000 pasos
Número de conformaciones adquiridas	10000

TABLA 7.3: Parámetros de la dinámica de Langevin para los muelles de PNIPAAm y MPEG.

Finalmente (recordemos que no se definen restricciones angulares), y dado que la intención de este trabajo es la de estudiar los sistemas mencionados en cuanto a su respuesta frente a la temperatura, es necesario encontrar la manera de introducir la influencia de ésta en el modelo elaborado. Como se ha adelantado en la introducción de este Capítulo, el efecto de la temperatura sobre el PNIPAAm consiste en una transición desde el carácter hidrofílico a baja temperatura, hasta el carácter hidrofóbico a temperaturas más altas. Es decir, el cambio de comportamiento está relacionado con la afinidad por el disolvente, por lo que el efecto de volumen excluido puede ser una buena forma de introducir este efecto en las simulaciones.

Por este motivo, para el modelado de las interacciones de volumen excluido se empleará un potencial de tipo Lennard-Jones, de forma que el efecto del disolvente se introducirá

en la simulación mediante la variación del parámetro ϵ_{LJ} para las diferentes temperaturas simuladas.

En las siguientes secciones se detallan los procedimientos de parametrización, así como los resultados obtenidos, para cada uno de los tipos de muelles que se han descrito anteriormente, así como del volumen excluido.

Muelles de PNIPAAM.

El procedimiento para la parametrización de los muelles consiste, en primer lugar, en extraer de la trayectoria generada las longitudes instantáneas de los conectores definidos entre los centros de las bolas. En segundo lugar, se construye un histograma o distribución de frecuencias de longitudes instantáneas de los muelles, y finalmente, mediante un programa de ajuste por mínimos cuadrados, se ajustan las distribuciones obtenidas a la ecuación de los muelles hard-FENE, para encontrar los parámetros de la expresión que producen un perfil de longitudes lo más próximo posible a los histogramas obtenidos de la trayectoria.

Dado que los conectores se definen entre los centros de las bolas, el primer paso del procedimiento consistiría en el cálculo de los centros de masas de las bolas a lo largo de la trayectoria, para a continuación, encontrar las distancias entre los mismos. Sin embargo, existe una limitación del software empleado: el programa no está preparado para el cálculo de los centros de masas, y sólo es posible calcular la distancia entre las posiciones ocupadas por átomos de la molécula para la que se ha generado la trayectoria, y no entre puntos ficticios como los centros de masas. Por este motivo, se recurre a calcular las distancias entre los átomos de nitrógeno homólogos de cada una de las bolas, pues estos átomos están situados sobre las cadenas laterales de los monómeros, pudiendo tomar éstas diferentes orientaciones que afectarán a la posición de los centros de masas.

Al operar de esta forma, para cada conformación de las que componen la trayectoria se obtienen cuatro datos de distancia entre átomos de nitrógeno: dos correspondientes a las distancias entre los pares de átomos de nitrógeno de la primera y la segunda bola, y otras dos correspondientes a las distancias entre los pares de las bolas dos y tres. Para compensar el criterio asumido de calcular las distancias entre los pares de átomos de nitrógeno homólogos de cada monómero, se calculan los promedios de las distancias, obteniendo finalmente un dato de distancia promedio para cada conformación instantánea

de la trayectoria. La distribución de distancias, o longitudes de muelle, a lo largo de toda la trayectoria, así como la media de las cuatro medidas se presenta en la figura 7.7.

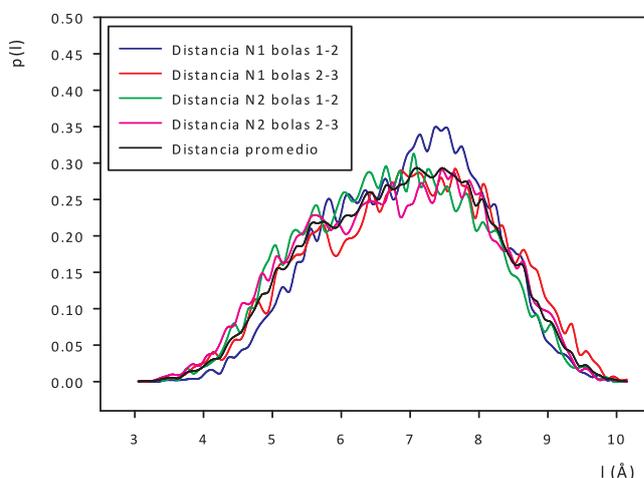


FIGURA 7.7: Distribución de longitudes entre bolas de PNIPAAM.

Como se observa en la figura, las distancias entre pares homólogos de átomos de nitrógeno son bastante similares entre sí, y la media es consistente, por lo que se acepta como bueno el tratamiento hecho, y se procede a ajustar los datos a la ecuación del potencial hard-FENE. Los parámetros resultantes del ajuste se presentan en la tabla 7.4, y en el gráfico la figura 7.8 se representa la curva obtenida empleando los parámetros del ajuste frente a la media de las distancias tomadas de la trayectoria. Como se puede observar, la curva se ajusta bastante bien a los datos de distancias promedio.

Constante del muelle	$H = 0.2 k_B T / \text{Å}^2$
Longitud de equilibrio	$l_e = 6.4 \text{ Å}$
Elongación máxima	$l_{max} = 10.5 \text{ Å}$

TABLA 7.4: Parámetros de los muelles hard-FENE de PNIPAAM.

En la figura 7.9 se representa el mapa de las desviaciones cuadráticas acumuladas obtenidas al barrer la gama de valores de la constante del muelle y la longitud máxima empleadas para el ajuste de los datos de distancias a la ecuación del potencial hard-FENE. Como se puede apreciar, mientras los cambios en la constante producen variaciones importantes en el valor de la desviación cuadrática acumulada, la longitud máxima produce cambios mucho menores, admitiendo una flexibilidad bastante grande en su valor. Esto indica que el comportamiento del muelle se ve poco influido por su longitud máxima, aproximándose más su comportamiento al de un muelle duro que a un muelle FENE.

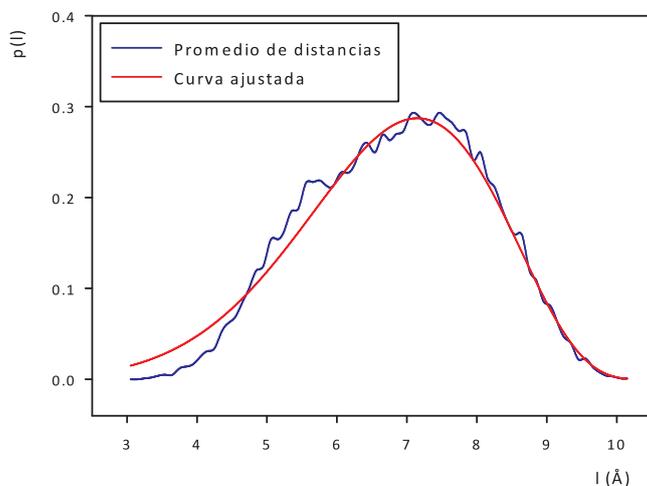


FIGURA 7.8: Comparación de la curva ajustada y la media de las distancias entre bolas de PNIPAAM.

Muelles de MPEG.

De nuevo nos vemos afectados por la limitación del software en cuanto al punto sobre el que se calculan las distancias entre las bolas, aunque en este caso los monómeros que componen las bolas no tienen cadenas laterales, por lo que los centros de masas de las bolas se encontrarán en todos los casos próximos a los átomos de oxígeno intermedios de cada bola. Por este motivo, sólo se analizan dos distancias, las que hay entre los átomos de oxígeno intermedios de las bolas primera y segunda, y segunda y tercera, respectivamente.

Análogamente a como se hizo en el caso del PNIPAAM, en la figura 7.10 se presentan las distribuciones de distancias entre los centros de las bolas y su promedio. También en este caso se observa que las distribuciones de distancias entre cada par de bolas son muy semejantes. Como en el caso anterior se hace el ajuste de los datos obtenidos a la ecuación de los muelles hard-FENE, obteniendo los valores de los parámetros de la tabla 7.5, con los que se construye la curva representada en la figura 7.11 junto al promedio de las distribuciones de distancias calculadas a partir de la simulación atomística. De nuevo, la correspondencia entre la curva ajustada y los datos provenientes de la simulación es muy buena.

Finalmente, en el mapa de desviaciones cuadráticas acumuladas de la figura 7.12 se vuelve a observar que el parámetro de dureza de los muelles tiene más influencia en las desviaciones acumuladas que la longitud máxima de los muelles, por lo que también en

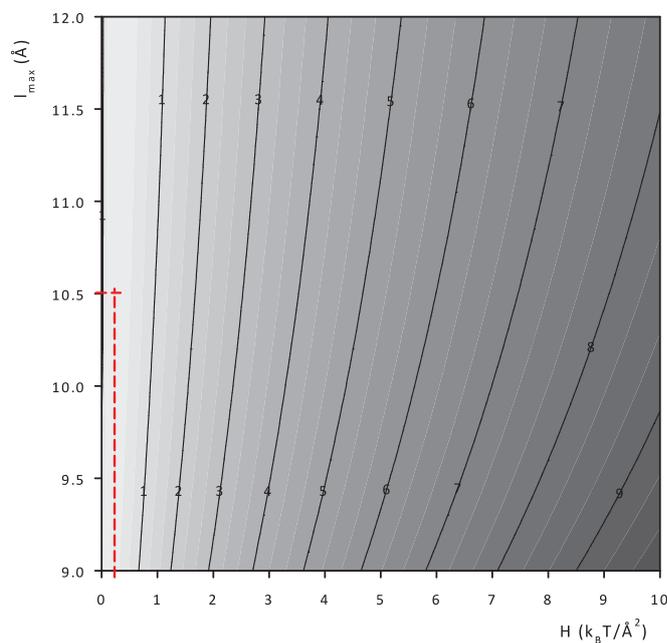


FIGURA 7.9: Desviaciones cuadráticas respecto a los parámetros de los conectores de PNIPAAM.

este caso los muelles tendrán un comportamiento más parecido al de un muelle duro que a uno FENE, aunque en este caso, el muelle será más duro y en general más corto que el de PNIPAAM.

Constante del muelle	$H = 0.7 k_B T / \text{Å}^2$
Longitud de equilibrio	$l_e = 4.9 \text{ Å}$
Elongación máxima	$l_{max} = 6.7 \text{ Å}$

TABLA 7.5: Parámetros de los muelles hard-FENE de MPEG.

Muelle del conector entre bloques.

Este tipo de muelle representa el punto de unión entre los bloques de distinta naturaleza, por lo que sólo aparece una vez en los polímeros dibloque, y dos veces no consecutivas en los tribloque. Se entiende por tanto que las características de este muelle no deben tener demasiada influencia en los resultados de la simulación.

Además, dado que no tiene sentido simular dos conectores de este tipo consecutivos, la estructura simulada con HYPERCHEM sólo contiene un conector del tipo que se está

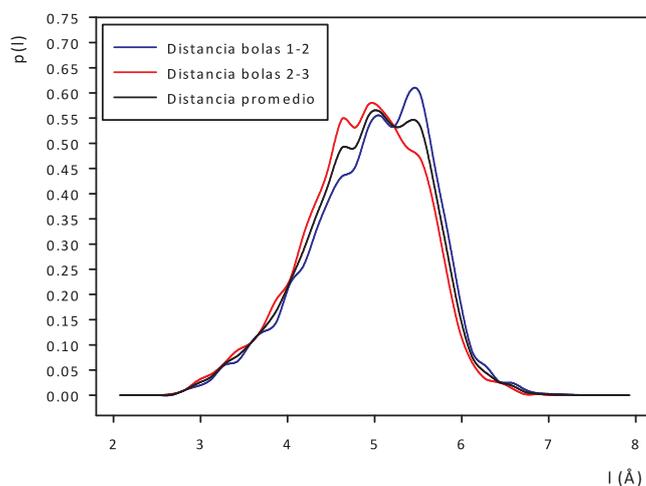


FIGURA 7.10: Distribución de longitudes entre bolas de MPEG.

parametrizando. Así, para que la gama de conformaciones incluidas en la trayectoria sea lo más amplia posible, se recurre a utilizar una trayectoria bastante más larga que en los casos anteriores, concretamente, la trayectoria es diez veces más larga que en los casos anteriores.

Debido a que en este caso no es posible promediar las longitudes instantáneas entre más de una instancia del conector, se decidió modificar la herramienta empleada para el cálculo de las longitudes de los muelles, dotándola de la capacidad que carecía de calcular los centros de masas de las bolas y poder así obtener la distancia real entre los centros de masas.

Como en los casos anteriores, una vez obtenida la distribución de longitudes del muelle, ésta se ajusta a la ecuación del muelle hard-FENE. Los resultados de este ajuste se presentan en la tabla 7.6. La figura 7.13 representa esta distribución junto a la curva que se obtiene al emplear los parámetros resultantes del ajuste. En cuanto al ajuste, de nuevo en este caso se puede observar que el parámetro de dureza del muelle tiene una mayor influencia en las diferencias cuadráticas acumuladas respecto a la distribución de distancias que la longitud máxima, y, tal como se puede apreciar a partir del valor de la dureza del muelle, y las curvas representadas, se trata también de un muelle bastante rígido, que mantiene su longitud próxima a la de equilibrio en la mayor parte de los casos.

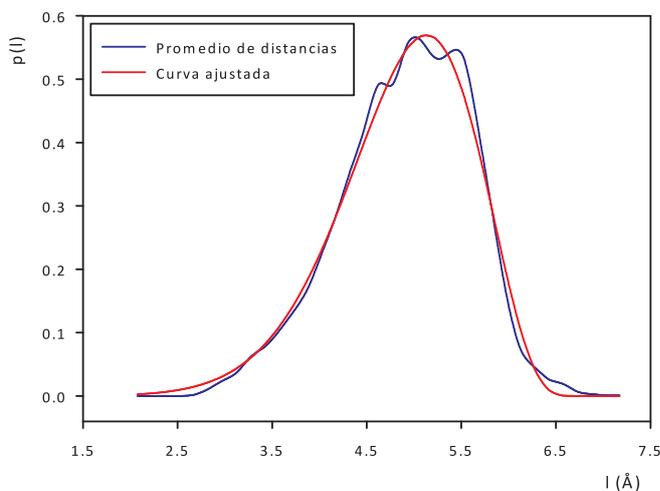


FIGURA 7.11: Comparación de la curva ajustada y la media de las distancias entre bolas de MPEG.

Constante del muelle	$H = 1.2 k_B T / \text{Å}^2$
Longitud de equilibrio	$l_e = 5.9 \text{ Å}$
Elongación máxima	$l_{max} = 13.5 \text{ Å}$

TABLA 7.6: Parámetros de los muelles hard-FENE del conector intermedio.

7.2.3. Parametrización del volumen excluido.

Para modelar las interacciones de largo alcance se ha elegido emplear un potencial de tipo Lennard-Jones, lo que implica que debemos determinar los parámetros σ_{LJ} , ϵ_{LJ} y $r_{cut-off}$ para cada uno de los tipos de interacción que se dan en los copolímeros simulados.

Dado que hay dos tipos distintos de bolas en el modelo, habrá tres tipos de interacciones: las que se den entre dos bolas de PNIPAAm, las que se den entre dos bolas de MPEG, y una interacción de tipo mixto entre una bola de PNIPAAm y una de MPEG.

Para la determinación del parámetro σ_{LJ} empleamos el siguiente criterio, generalmente aceptado [68, 99]:

$$\sigma_{LJ} = 0.8 l_{eq}, \quad (7.1)$$

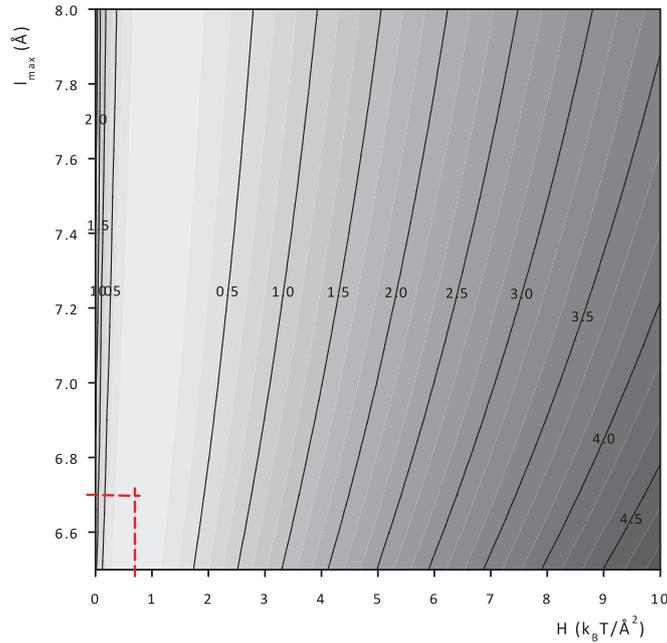


FIGURA 7.12: Desviaciones cuadráticas respecto a los parámetros de los conectores de MPEG.

donde l_{eq} es la longitud de equilibrio del muelle que uniría dos bolas del tipo para las que se desea determinar σ_{LJ} . Dado que l_{eq} es un dato que se obtiene de la parametrización de los muelles, la obtención de σ_{LJ} para la interacción de bolas del mismo tipo es trivial.

En lo que se refiere al MPEG, la parametrización del volumen excluido se hace considerando que su solubilidad en agua es siempre buena, independientemente de los cambios de temperatura del medio simulado, por lo que emplearemos un valor de $\epsilon_{LJ,MPEG}$ correspondiente a la transformación a unidades físicas de $\epsilon_{LJ}^* = 0.1$ habitualmente empleado en simulaciones adimensionales para representar condiciones de buen disolvente [68]. La transformación a unidades físicas se hace multiplicando por el valor apropiado de $K_B T$, la unidad de energía habitualmente empleada en nuestras simulaciones en unidades reducidas (ver sección 2.6).

De la misma manera, podemos considerar que el límite de las condiciones de buen disolvente vendrá dado por el valor de $\epsilon_{LJ}^* = 0.3$, correspondiente a condiciones theta, que se representa también en las figuras que contienen los resultados como referencia de la calidad del disolvente.

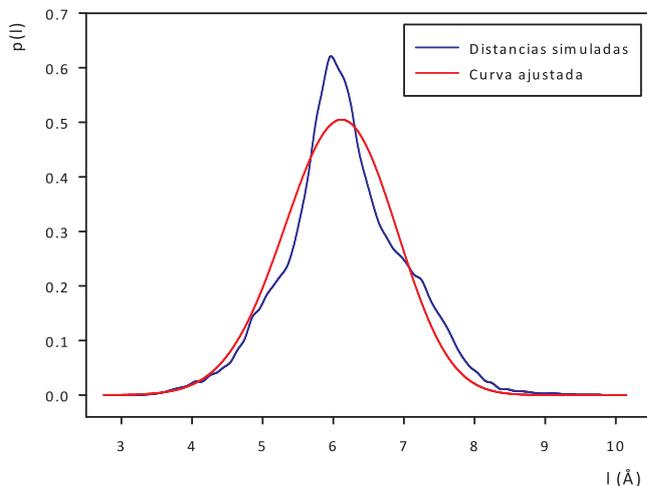


FIGURA 7.13: Comparación de la curva ajustada y la media de las distancias entre bolas de conector intermedio.

Para el caso del PNIPAAm la determinación es algo más elaborada. El procedimiento de la determinación de ϵ_{LJ} para las bolas de PNIPAAm se basa en un método de ensayo-error. Partiendo de los datos de un trabajo anterior, [98], se modela y parametriza una cadena de PNIPAAm de alto peso molecular similar a la presentada en dicho artículo, empleando el método de parametrización de cadenas FENE descrito en un trabajo anterior del Grupo, [100]. Para modelar la cadena se emplea un volumen excluido de Lennard-Jones, en el que se van ensayando diferentes valores de ϵ_{LJ} hasta conseguir que los datos de radio de giro procedentes de la simulación concuerden con los presentados en el artículo. De esta manera, repitiendo el ensayo a diferentes temperaturas, se obtiene una relación de datos de ϵ_{LJ} que imitarán el comportamiento hidrofílico o hidrofóbico del PNIPAAm al variar la temperatura.

Si bien en el caso de las interacciones entre dos bolas del mismo tipo la parametrización del volumen excluido es relativamente sencilla, para el caso de la interacción mixta entre una bola de MPEG y una bola de PNIPAAm es más complicada, pues debido a las dificultades que se comentaron en la introducción, no es posible disponer de datos experimentales.

Ni siquiera es posible obtener σ_{LJ} de la misma manera que en el caso de las interacciones “puras”, pues al parametrizar el conector entre bloques, se tiene en cuenta la unión química existente entre los bloques de la cadena, compuesta por átomos que no pertenecen a ninguna de las dos bolas, y de cierta rigidez, lo que puede hacer que el valor de l_{eq} del

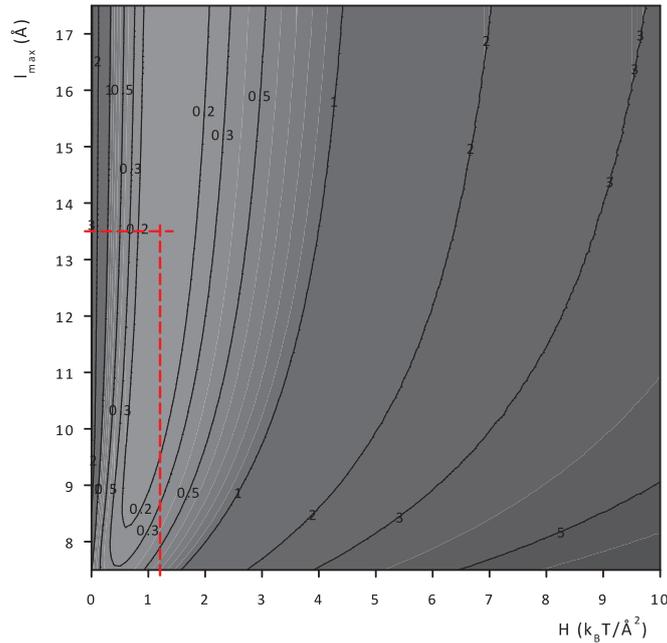


FIGURA 7.14: Desviaciones cuadráticas respecto a los parámetros de los conectores de conector intermedio.

conector entre bloques no corresponda adecuadamente con la magnitud apropiada para determinar σ_{LJ} de una interacción mixta.

Para obtener una aproximación a los parámetros de la interacción de volumen excluido mixta es necesario recurrir a la ley de combinación de Lorentz-Berthelot [101], según la cual el valor de los parámetros ϵ_{LJ} y σ_{LJ} corresponde a los promedios geométrico y aritmético, respectivamente, de los parámetros de las interacciones “puras” entre bolas de las dos especies a las que relaciona esta interacción mixta. Es decir,

$$\epsilon_{LJ,mixto} = \sqrt{\epsilon_{LJ,MPEG} \cdot \epsilon_{LJ,PNIPAAM}} \quad (7.2)$$

$$\sigma_{LJ,mixto} = \frac{\sigma_{LJ,MPEG} + \sigma_{LJ,PNIPAAM}}{2}. \quad (7.3)$$

De esta manera, estamos teniendo en cuenta la consideración de que la hidrofobicidad del bloque de PNIPAAM también afectará a las condiciones termodinámicas del bloque hidrófilo de MPEG [96, 102].

Finalmente, el tercer parámetro necesario para emplear el potencial de Lennard-Jones es el radio de corte, $r_{cut-off}$, que representa la distancia entre dos bolas para la cual se

considera que la interacción entre ellas es despreciable. El criterio general empleado para decidir el valor de este parámetro (tanto para las interacciones “puras”, como para la interacción mixta) es el de tomar el doble del valor del correspondiente σ_{LJ} , por lo que la determinación de este parámetro también es directa. Los valores resultantes de las parametrizaciones se recogen en las tablas 7.7 y 7.8.

Tipo de interacción	σ_{LJ} (Å)	$r_{\text{cut-off}}$ (Å)
PNIPAAM	5.12	10.24
MPEG	3.92	7.84
Mixta	4.48	9.96

TABLA 7.7: Parámetros de simulación no variables con la temperatura.

Temp (°C)	$\epsilon_{LJ,PNIPAAM} \times 10^{15}$ (erg)	$\epsilon_{LJ,MPEG} \times 10^{15}$ (erg)	$\epsilon_{LJ,mixto} \times 10^{15}$ (erg)
25	4.11	4.11	4.11
30	12.54	4.18	7.24
31.4	21.00	4.20	9.39
31.7	29.43	4.21	11.13
31.75	42.06	4.21	13.30
31.8	63.09	4.21	16.29
31.85	84.14	4.21	18.81
31.9	210.38	4.21	29.75

TABLA 7.8: Parámetros de simulación variables con la temperatura.

7.3. Resultados.

7.3.1. Copolímeros dibloque.

En primer lugar, es necesario establecer un criterio para referirse a los modelos simulados. Durante el trabajo, el criterio de nomenclatura adoptado consistió en identificar a cada modelo mediante dos números separados por un guión, de forma que el primer

número indicase la longitud del bloque de PNIPAAm, mientras el segundo número indica la longitud del bloque de MPEG.

Como se mencionaba en la introducción de este Capítulo, iban a simularse copolímeros de tamaño de bloque de PNIPAAm constante de 71 monómeros, con bloque de 0, 12, 23, y 114 monómeros de MPEG, cuya nomenclatura, recordando los criterios establecidos en la parametrización de las bolas (dos monómeros por bola, y redondeo hacia arriba), correspondería a 36-0, 36-6, 36-12 y 36-57. La segunda serie de copolímeros, consistente en el bloque de MPEG de 53 monómeros, con bloques de PNIPAAm de 0, 19, 70 y 114 monómeros, estaría compuesta por lo tanto por los modelos 0-27, 10-27, 35-27 y 57-27.

En las figuras 7.15, 7.16, 7.18 y 7.18 se presentan los resultados de las simulaciones de los copolímeros con un bloque de PNIPAAm de tamaño constante. Sobre los gráficos cabe destacar que, debido a una cuestión práctica que se desarrollará más adelante, resulta más conveniente representar los datos frente a los valores de $\epsilon_{LJ,PNIPAAm}$ que frente a las temperaturas a las que corresponden estos valores. La correspondencia entre los valores de $\epsilon_{LJ,PNIPAAm}$ y las temperaturas se recogía en la tabla 7.8.

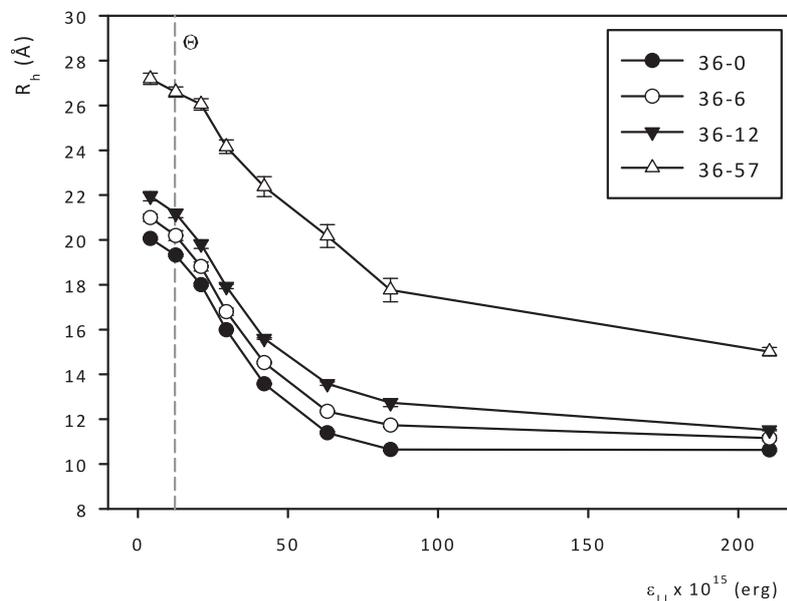


FIGURA 7.15: Radios hidrodinámicos de las cadenas con bloque de PNIPAAm de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

En términos generales, se puede concluir que, tanto en éstos como en los copolímeros con bloques de MPEG de tamaño constante se observa la tendencia esperada a un menor

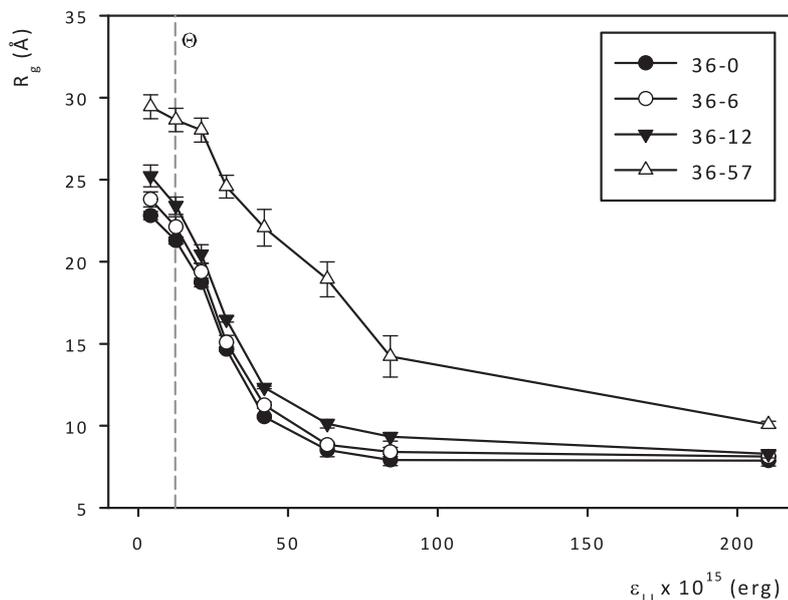


FIGURA 7.16: Radios de giro de las cadenas con bloque de PNIPAAAM de tamaño constante frente a $\epsilon_{LJ,PNIPAAAM}$.

tamaño global (R_h , R_g) cuanto mayor es $\epsilon_{LJ,PNIPAAAM}$ (y por tanto también la temperatura) dentro del rango estudiado.

Analizando los gráficos de las figuras 7.15 y 7.16 con mayor detalle, en primer lugar se observa que todos los copolímeros, salvo el que posee el bloque de MPEG de mayor longitud, tienen tamaños similares, de lo que se puede concluir que la adición de bloques de MPEG cortos tiene poco efecto sobre el tamaño global de la molécula.

Comparando las distancias extremo-extremo de los bloques de PNIPAAAM y MPEG, representadas respectivamente en las figuras 7.17 y 7.18 puede observarse que, aunque las cadenas cortas de MPEG no acusan ningún efecto termosensible, la cadena de mayor tamaño sí presenta cierta sensibilidad frente a la temperatura. Sin embargo, este efecto de contracción es mucho menos acusado que en el caso del PNIPAAAM (la distancia extremo-extremo sólo llega a hacerse la mitad, mientras que en el PNIPAAAM llega a reducirse hasta una sexta parte), y comienza a producirse a una temperatura más alta.

En cuanto al bloque de PNIPAAAM, se observa que incluso en condiciones theta ya existe un cierto grado de contracción del bloque respecto a las condiciones de buen disolvente. Además, este comportamiento parece ser independiente del tamaño del bloque de MPEG, aunque sí que se observa una pequeña diferencia en el comportamiento en el caso

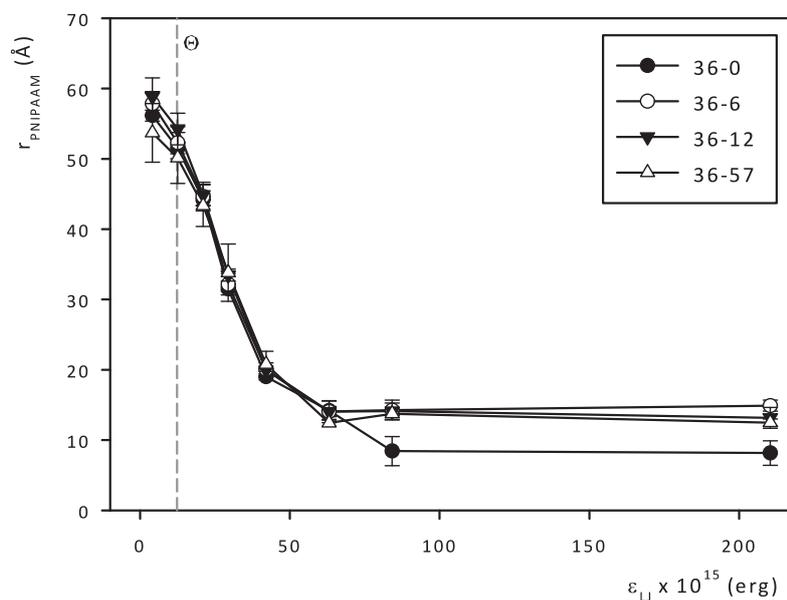


FIGURA 7.17: Distancia extremo-extremo del bloque de PNIPAAm de las cadenas con bloque de PNIPAAm de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

de no estar presente en bloque de MPEG, siendo en este caso aún mayor la contracción: la distancia extremo-extremo del bloque de PNIPAAm puro es de unos 8 Å, mientras que en los copolímeros que poseen un bloque de MPEG la distancia extremo-extremo del bloque de PNIPAAm oscila entre 13 y 15 Å.

A partir de este efecto, y de las imágenes obtenidas de los datos de la simulación, de las que se presenta un ejemplo en la figura 7.19, puede concluirse que el bloque de MPEG produce un efecto protector sobre el bloque de PNIPAAm, reduciendo el contacto de éste con el medio acuoso. Al aumentar la temperatura y empeorar las condiciones del disolvente, el bloque de PNIPAAm se contrae, y al alcanzar un grado suficiente de hidrofobicidad, el bloque de MPEG, hidrofílico, se envuelve alrededor del núcleo de PNIPAAm debido a la interacción entre ambos bloques.

Este efecto explica la contracción observada en el bloque de MPEG de mayor tamaño, a la vez que también explica la mayor contracción de la cadena de PNIPAAm que carece del bloque de MPEG; al carecer de esta corona protectora, o en los casos en que el bloque de MPEG no es lo suficientemente largo como para ofrecer una buena envoltura, el núcleo hidrofóbico se ve obligado a adoptar una conformación más compacta. Un ejemplo de esto puede observarse en la figura 7.20, en la que, al contrario de lo que se observaba en la

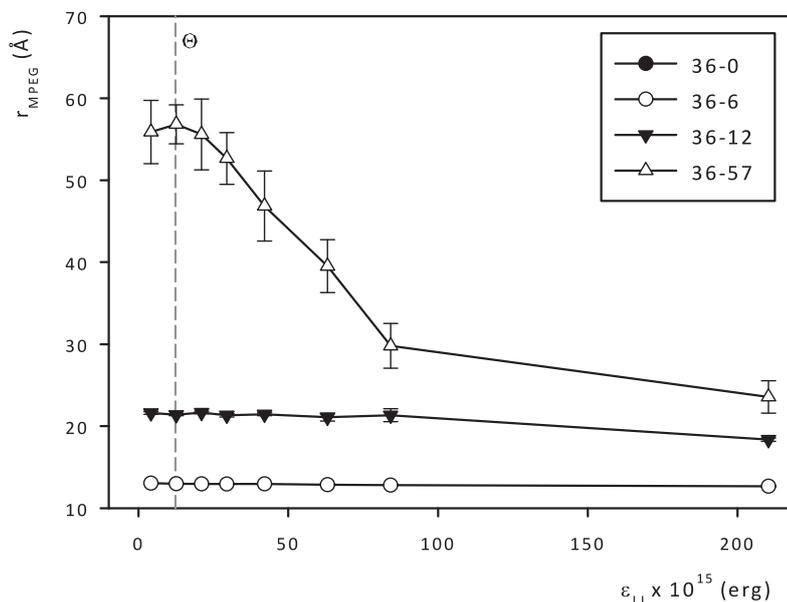


FIGURA 7.18: Distancia extremo-extremo del bloque de MPEG de las cadenas con bloque de PNIPAAm de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

figura 7.19, el bloque de MPEG es incapaz de envolver completamente al núcleo y éste adopta, como ya se ha mencionado, una conformación más compacta.

Por este motivo las disoluciones de mayor concentración de copolímeros de este tipo tienden a formar agregados con facilidad a temperaturas más bajas, pues como se argumenta en [103] en ausencia de una corona protectora, o si la cobertura cobertura es insuficiente, las moléculas tienden a agregarse asociando las partes expuestas de los núcleos hidrófobos.

Además de las conclusiones anteriores, uno de los puntos más interesantes de este trabajo es la posibilidad de comparar los resultados de las simulaciones con datos experimentales. Gracias a la colaboración con el profesor Nyström, en cuyo laboratorio se han sintetizado y caracterizado los polímeros que corresponden a los modelos simulados, disponemos de tales datos experimentales [103]. En concreto, el grupo del profesor Nyström ha obtenido datos de dispersión dinámica de luz (DLS) de disoluciones diluidas de los diferentes copolímeros en rangos de concentraciones entre el 0.5% y 1% en peso. Se emplearon estas concentraciones porque por debajo de ellas la intensidad de la señal no permitía la obtención de funciones de correlación viables.

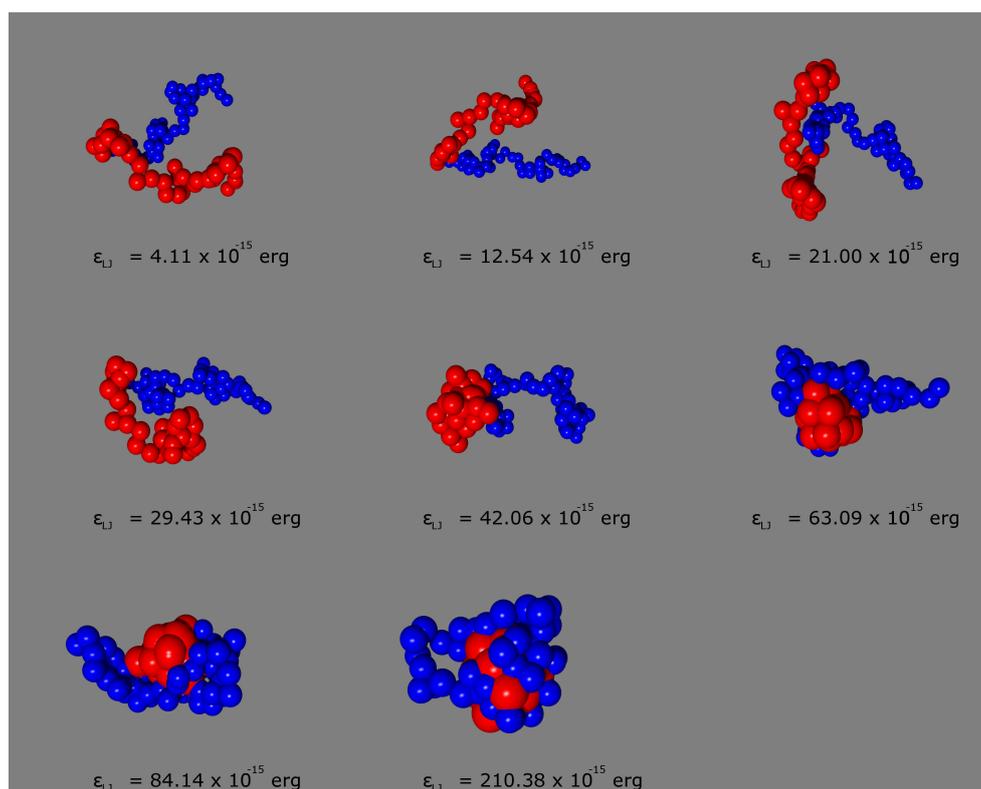


FIGURA 7.19: Conformaciones del copolímero 36-57 a diferentes valores de $\epsilon_{LJ,PNIPAAM}$. Cada bola roja representa dos monómeros de PNIPAAM, mientras cada bola azul representa dos monómeros de MPEG. La escala cambia de una imagen a otra.

Mediante estos experimentos se obtuvieron los datos de radios hidrodinámicos aparentes que se presentan en la figura 7.21. Como puede observarse en la figura, hay algunas cuestiones fundamentales que saltan a la vista: en primer lugar, las temperaturas a las que se producen las contracciones de los copolímeros son superiores a las que se habían empleado para la parametrización de las simulaciones. En segundo lugar, se observa que las temperaturas a las que se producen las contracciones de los copolímeros son diferentes entre los polímeros sintetizados, y en especial en el caso del polímero correspondiente al modelo 36-6 apenas llega a observarse una pequeña contracción antes del final del experimento.

Finalmente, también es de destacar que los tamaños observados de los copolímeros son superiores a los simulados en aproximadamente 5 \AA . Esta diferencia puede atribuirse a la dificultad de la determinación de R_h en los experimentos de DLS, pues las muestras de los copolímeros en disolución pueden contener tanto moléculas dispersas como agregados de mayor tamaño. Además, mientras en los resultados de la simulación no se tienen

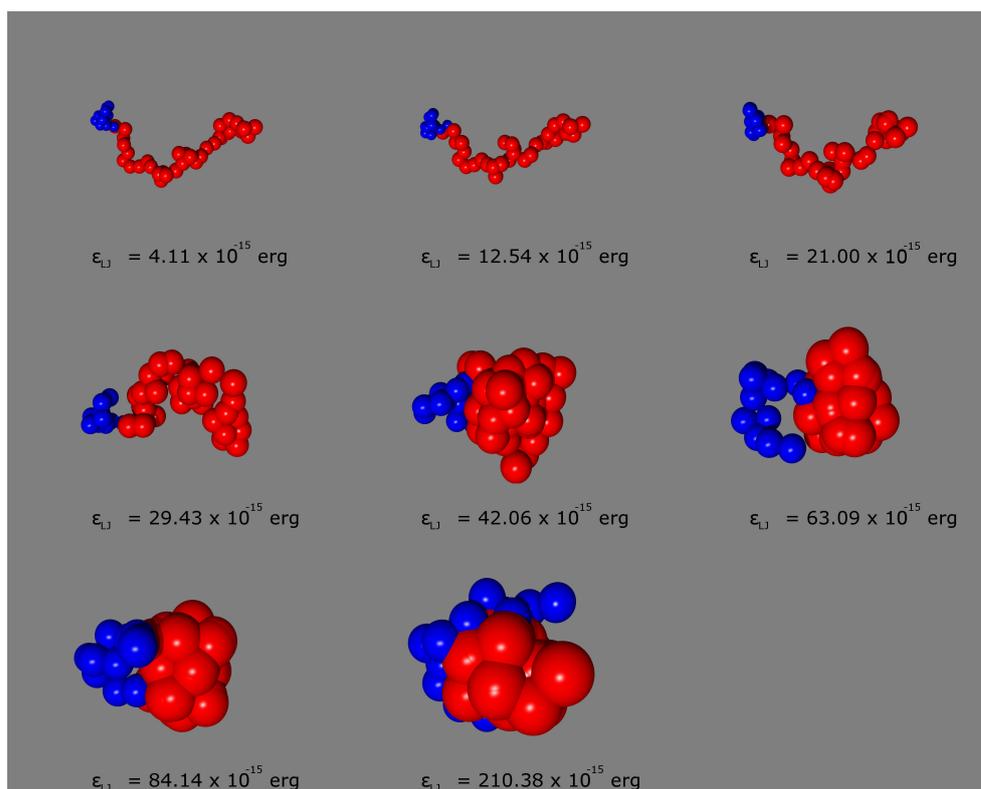


FIGURA 7.20: Conformaciones del copolímero 36-12 a diferentes valores de $\epsilon_{LJ,PNIPAAM}$. Cada bola roja representa dos monómeros de PNIPAAM, mientras cada bola azul representa dos monómeros de MPEG. La escala cambia de una imagen a otra.

en cuenta factores como la polidispersidad o las asociaciones intermoleculares, éstos sí podrían afectar a los valores experimentales.

En el gráfico se observa cómo el punto en el que disolvente “se empobrece” y se manifiesta la contracción de las cadenas de copolímero se desplaza hacia temperaturas cada vez mayores cuanto más largo es el bloque de MPEG del copolímero, lo que refuerza la hipótesis formulada anteriormente de que este bloque ejerce un efecto estabilizante sobre el núcleo hidrofóbico constituido por el PNIPAAM.

A pesar de las diferencias de temperaturas empleadas en la simulación y las observadas experimentalmente, es posible relacionar los datos obtenidos por ambas vías por medio de los valores de $\epsilon_{LJ,PNIPAAM}$, si se acepta que a pesar de la diferencia en el valor de la temperatura el comportamiento del copolímero es similar. Éste es el motivo por el que al analizar los resultados de las simulaciones se empleaba este parámetro en lugar de las temperaturas para las que se hizo la parametrización.

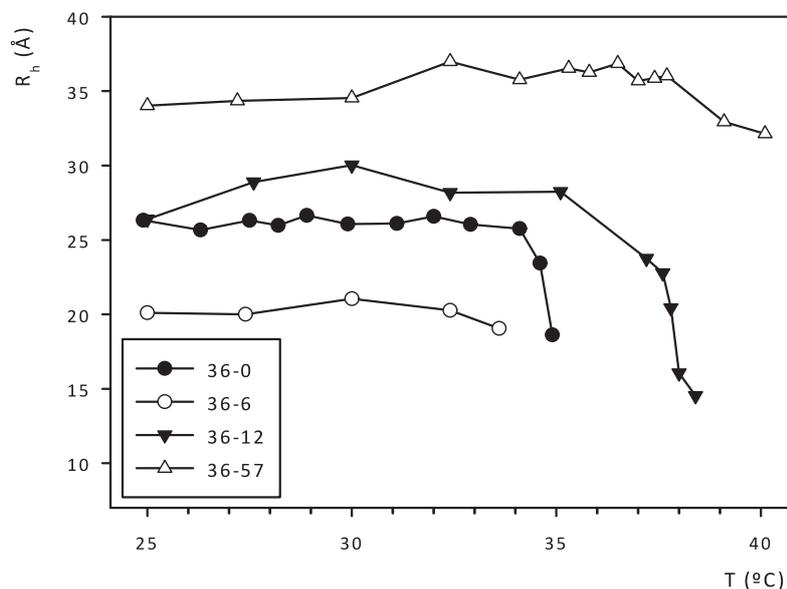


FIGURA 7.21: Radios hidrodinámicos aparentes de los copolímeros reales obtenidos mediante DLS.

Para relacionar los datos experimentales con los resultados de la simulación, en primer lugar es necesario normalizar los datos de los radios hidrodinámicos de cada copolímero respecto al valor de su radio hidrodinámico en condiciones de buen disolvente, tanto para los resultados de las simulaciones como de los datos experimentales. Una vez realizada la normalización, comparando el valor del radio normalizado se puede relacionar fácilmente la temperatura de las medidas experimentales con el valor de $\epsilon_{LJ,PNIPAAM}$ de las simulaciones, y construir un gráfico como el de la figura 7.22. El único caso en el que no pudo aplicarse este procedimiento (y por tanto no figura en el gráfico) es el del copolímero PNIPAAM₇₁-MPEG₁₂ (equivalente a la simulación 36-6), debido a que los datos experimentales no son lo suficientemente buenos como para apreciar una contracción significativa, aunque se puede asumir que por debajo de 34 °C se encuentra en condiciones de buen disolvente. Comparando la figura resultante con la 7.21 se puede apreciar que aunque las curvas no son idénticas, sí guardan una similitud cualitativa (e incluso cuantitativa) importante.

En cuanto a la serie de copolímeros con el bloque de MPEG de tamaño constante, los datos obtenidos de tamaño hidrodinámico, radio de giro y distancias extremo-extremo de los bloques de MPEG y PNIPAAM se recogen en las figuras 7.23, 7.24, 7.25, 7.26.

Como cabe esperar, dado que el MPEG no es termosensible, en estas figuras se observa que la cadena que sólo está compuesta por MPEG y carece de bloque de PNIPAAM,

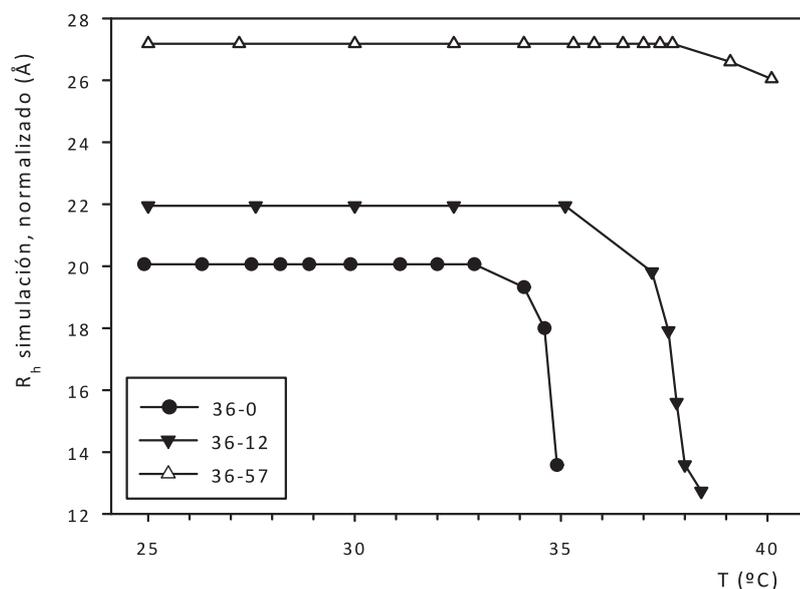


FIGURA 7.22: Radios hidrodinámicos de las simulaciones reescalados a las temperaturas experimentales.

0-27, conserva un tamaño aproximadamente constante y no se ve afectada por el cambio de temperaturas. En el resto de casos en los que sí existe el bloque de PNIPAAm, se observa el mismo fenómeno que en el caso de la cadena 36-57, en el que el bloque de MPEG también se contraía, siendo esta contracción más acusada cuanto más largo es el bloque de PNIPAAm asociado. Este fenómeno, como ya se ha comentado anteriormente, se asocia con la formación de una corona protectora alrededor del núcleo hidrofóbico de PNIPAAm.

Este efecto de formación de la corona protectora juega un papel importante en el comportamiento asociativo de estas moléculas en disolución: cuando el núcleo de PNIPAAm es lo suficientemente grande respecto a la cadena de MPEG, esta última no es capaz de cubrir completamente el núcleo hidrofóbico, lo que afecta a la estabilidad del copolímero a temperaturas elevadas. Si la corona protectora de MPEG sólo cubre parcialmente el núcleo, pueden producirse asociaciones intermoleculares que conduzcan a la formación de grandes estructuras intermicelares como las que se han encontrado en los trabajos experimentales sobre este tipo de copolímeros.

En cuanto al tamaño del bloque de PNIPAAm, se observa que al aumentar éste, la contracción manifestada se hace mayor. El caso más acusado es el del copolímero con el bloque de PNIPAAm de mayor tamaño, el 57-27, que llega a reducir su tamaño a una

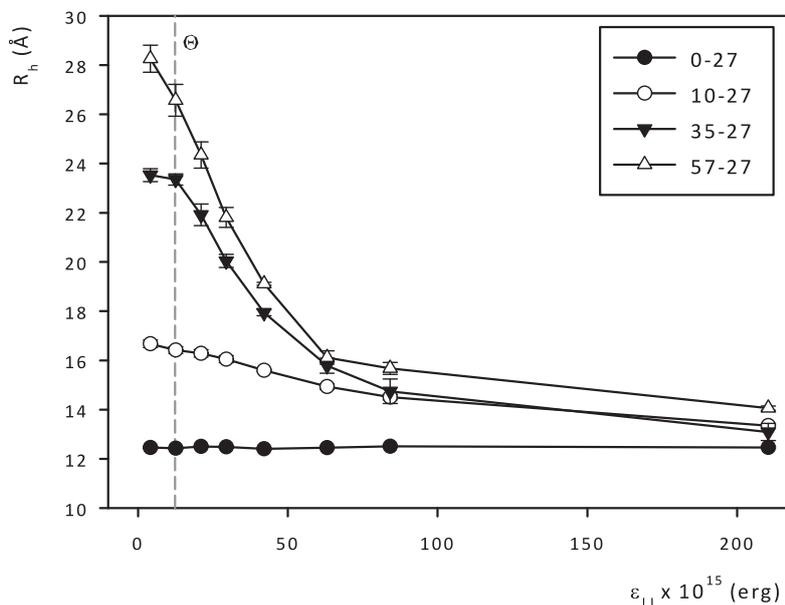


FIGURA 7.23: Radios hidrodinámicos de las cadenas con bloque de MPEG de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

séptima parte del que presentaba en condiciones de buen disolvente. De esta manera, se observa que, tal como se podía imaginar *a priori*, las dimensiones de las moléculas de copolímero pueden seleccionarse con facilidad mediante el control del peso molecular del bloque termosensible.

Al igual que en el caso de los copolímeros con bloque de PNIPAAm de tamaño constante, el grupo del profesor Nyström realizó la síntesis y caracterización mediante DLS de los copolímeros con bloque de MPEG de tamaño constante. También en este caso se recurrió a disoluciones diluidas con concentraciones entre el 0.5 % y el 1 % de concentración en peso, pues a concentraciones inferiores las intensidades de luz dispersada no eran suficientes para realizar análisis fiables.

Por desgracia, para esta serie de copolímeros no es posible repetir la re-parametrización de las temperaturas a partir de los datos experimentales, pues no fue posible observar la contracción de las moléculas individualizadas. Sin embargo, salvo para el caso del copolímero correspondiente al modelo 10-27, para el que sólo se pudo observar agregados, sí pudo determinarse el tamaño hidrodinámico del resto de copolímeros de la serie: para el correspondiente al modelo 35-27 se observó un radio hidrodinámico $R_h \approx 27 - 30 \text{ \AA}$, frente al valor de $R_h = 23.5 \text{ \AA}$ obtenido de las simulaciones, mientras para el copolímero 57-27, en el rango de 15 a 37 °C se observó un radio hidrodinámico de $R_h \approx 30 - 35 \text{ \AA}$,

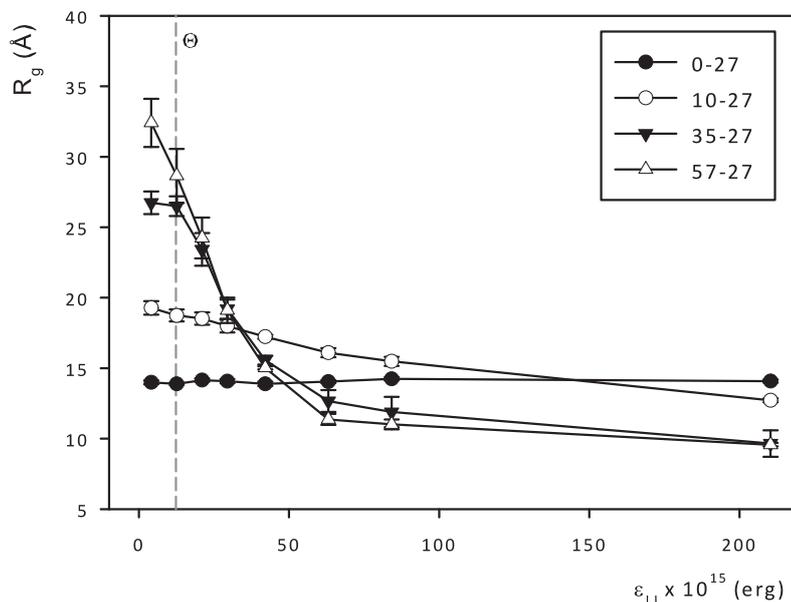


FIGURA 7.24: Radios de giro de las cadenas con bloque de MPEG de tamaño constante frente a $\epsilon_{LJ,PNIPAAM}$.

correspondiente al valor de simulación en condiciones de buen disolvente de $R_h = 28.3 \text{ \AA}$. En ambos casos se encuentra la misma desviación de aproximadamente 5 \AA entre valores experimentales y de simulación que se observaba también en los copolímeros con bloque de PNIPAAM de tamaño constante.

Para estudiar de forma comparativa la evolución de ambas series de copolímeros (la de bloque PNIPAAM constante, y la de bloque MPEG constante), a partir de las simulaciones pueden calcularse los radios hidrodinámicos normalizados respecto a las dimensiones en estado de buen disolvente. Dichos valores se representan para ambas series de copolímeros en las figuras 7.27 y 7.28.

A partir de estas figuras se observan dos tendencias bien definidas: en el caso del bloque de PNIPAAM de tamaño constante, al contraerse, todos los modelos tienden a alcanzar un tamaño similar. De aquí se puede concluir que efectivamente existe una tendencia a la compactación inducida por la temperatura, pero la longitud del bloque hidrofílico sólo tiene un impacto moderado en la contracción, lo que indica que la mayor parte del efecto de contracción se debe al tamaño del núcleo hidrofóbico.

Sin embargo, en el caso de la serie con el bloque de MPEG de tamaño constante, el tamaño alcanzado cuando se produce la contracción disminuye progresivamente según

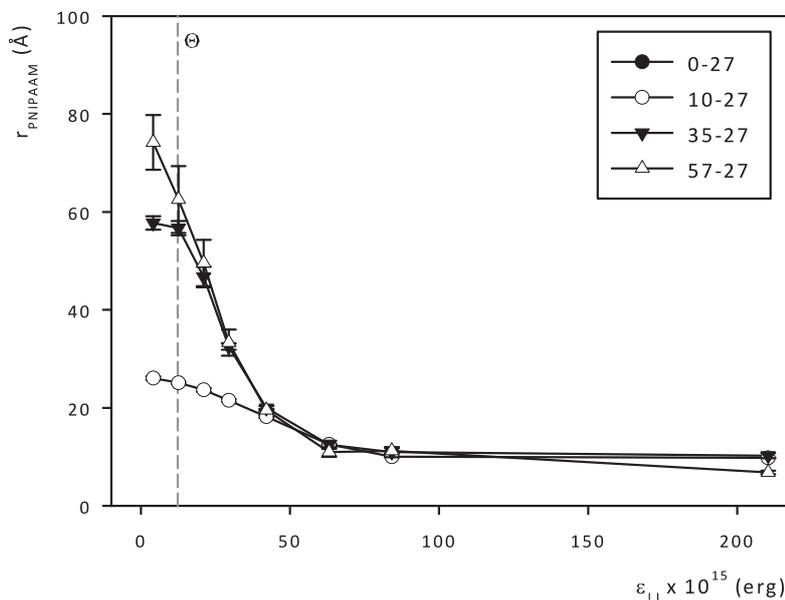


FIGURA 7.25: Distancia extremo-extremo del bloque de PNIPAAm de las cadenas con bloque de MPEG de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

aumenta el tamaño del bloque de PNIPAAm. Esto indica que el efecto de la contracción es más importante cuanto mayor es la longitud del bloque termosensible, de forma que copolímeros con un bloque de PNIPAAm de mayor longitud tendrán una capacidad de contracción mayor que moléculas similares con un bloque termosensible más corto.

Al aumentar la longitud del bloque PNIPAAm, aumenta la densidad de elementos hidrofóbicos en el núcleo, y en los casos en los que la cadena hidrofílica sólo es suficiente para cubrir parcialmente el núcleo, la compactación producida es aún mayor para evitar el contacto con el agua. Otra consecuencia de la incapacidad del bloque de MPEG de proteger completamente al núcleo es la tendencia a formar agregados intermoleculares en disoluciones de concentraciones finitas a altas temperaturas, como también se ha podido comprobar experimentalmente.

Además, tal como puede observarse en la figura 7.26, una mayor longitud del bloque termosensible tiende a desplazar la temperatura de transición hacia valores menores. Estos resultados revelan de forma clara que el tamaño del bloque de PNIPAAm es decisivo para la magnitud de la compresión de las moléculas.

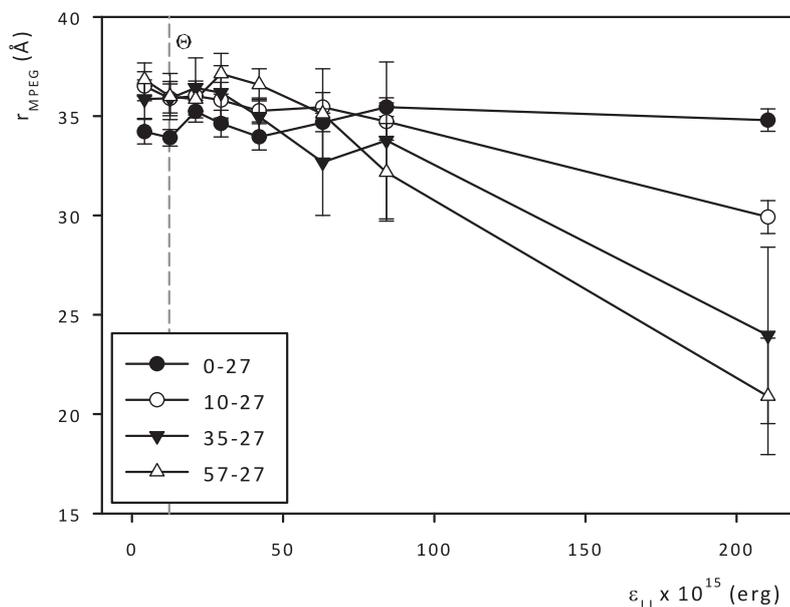


FIGURA 7.26: Distancia extremo-extremo del bloque de MPEG de las cadenas con bloque de MPEG de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

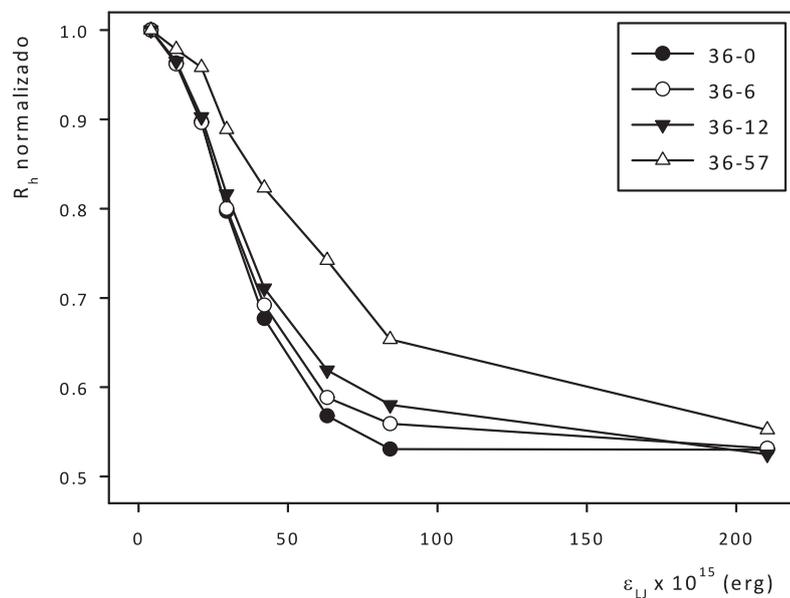


FIGURA 7.27: Radios hidrodinámicos normalizados de las cadenas con bloque de PNIPAAm de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

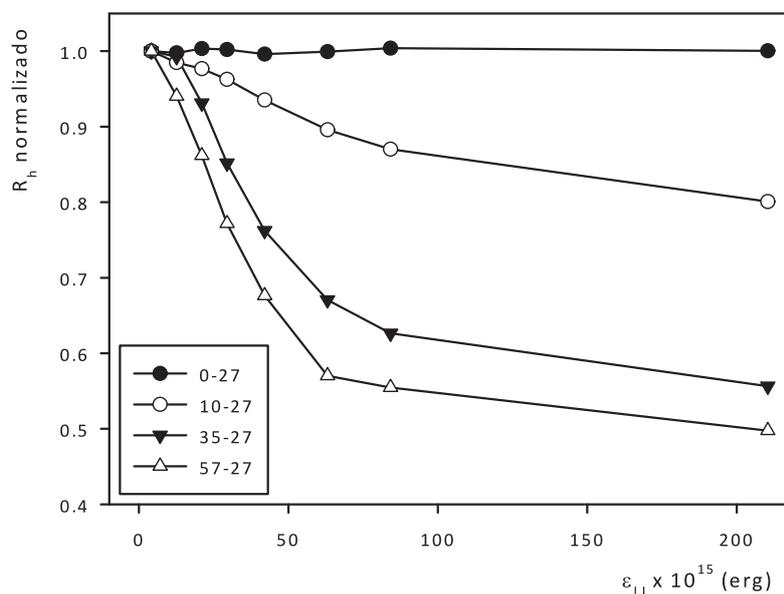


FIGURA 7.28: Radios hidrodinámicos normalizados de las cadenas con bloque de MPEG de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

La última consideración se refiere al cambio conformacional asociado al proceso de contracción. Un parámetro que se emplea con frecuencia para estudiar los cambios conformacionales inducidos por la temperatura en moléculas individuales de este tipo, es la relación R_g/R_h . Se ha calculado dicha relación a partir de los datos de la simulación, y se ha representado en las figuras 7.29 y 7.30 para ambas series de copolímeros.

Valores bajos de esta relación implican una conformación más esférica: para un ovillo al azar en condiciones de buen disolvente se obtienen valores de $R_g/R_h \approx 1.5$, mientras que para una esfera dura uniforme $R_g/R_h \approx 0.77$ [104–106]. Para los copolímeros de MPEG-PNIPAAm se observa una tendencia de R_g/R_h a disminuir desde ~ 1.15 hasta ~ 0.7 , lo que confirma la transición desde una conformación expandida a una globular. En la bibliografía pueden encontrarse valores de $R_g/R_h \approx 1.2$ en el estudio de nanofibras con comportamiento de varillas rígidas [107] y para estructuras de tipo vermiforme [108], por lo que se puede pensar que los copolímeros estudiados presentarían un comportamiento similar a tales estructuras rígidas a bajas temperaturas.

En el caso de los estudios experimentales llevados a cabo, no fue posible la determinación de R_g (y por tanto del parámetro R_g/R_h) debido a la coexistencia en el seno del disolvente de cadenas de copolímero aisladas y de micelas, lo que dificulta la determinación experimental. De esta manera, al no disponer de datos experimentales no es posible

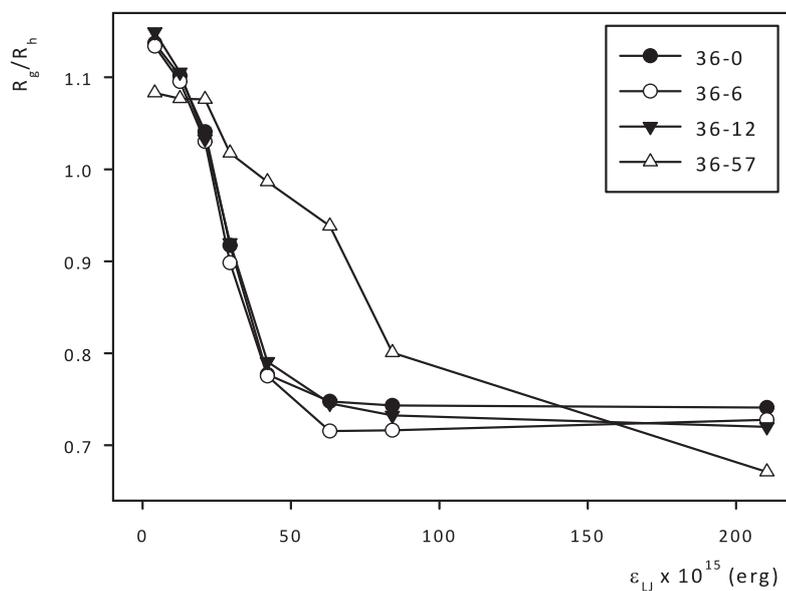


FIGURA 7.29: Relación R_g/R_h de las cadenas con bloque de PNIPAAm de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

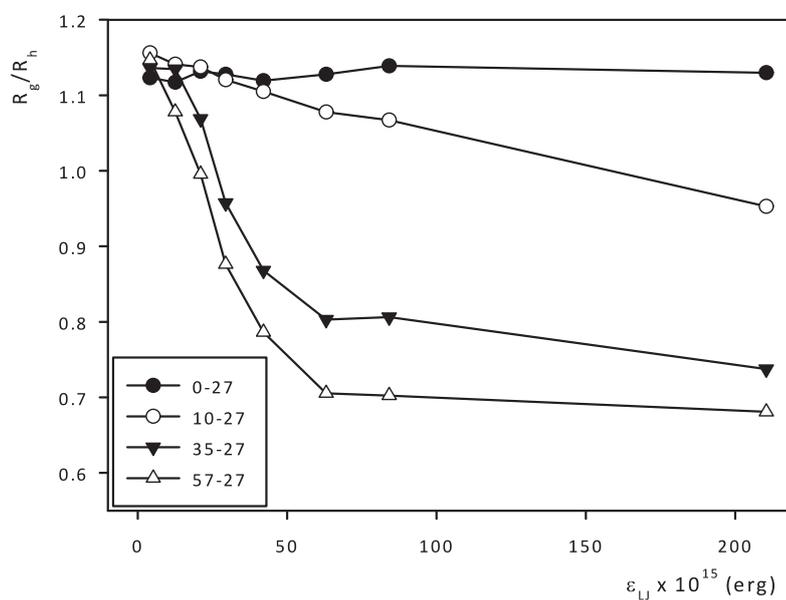


FIGURA 7.30: Relación R_g/R_h de las cadenas con bloque de MPEG de tamaño constante frente a $\epsilon_{LJ,PNIPAAm}$.

confirmar la hipótesis formulada acerca del comportamiento de estructura rígida de los copolímeros.

7.3.2. Copolímeros tribloque.

Para esta nueva serie de simulaciones es necesario establecer un nuevo criterio para la identificación de los modelos. Por similitud con el criterio empleado para los modelos de los copolímeros dibloque, y para distinguirlos de los anteriores, en este caso se emplearán tres cifras separadas por guiones; el primer y tercer números, iguales entre sí, se referirán al número de bolas de PNIPAAm (manteniendo el criterio adoptado en la parametrización de equivalencia de una bola a dos monómeros), mientras que el segundo número se refiere al número de bolas de MPEG en el conector central.

Identificados de esta manera, se construyeron seis nuevos modelos: 35-12-35, 32-17-32, 32-39-32, 32-50-320, 32-75-32 y 32-100-32. En este caso, de esta serie de modelos, únicamente los tres primeros tienen un equivalente físico sintetizado en el laboratorio del profesor Nyström, los cuales se corresponden, respectivamente, con los polímeros PNIPAAm₆₉-MPEG₂₃-PNIPAAm₆₉, PNIPAAm₆₇-MPEG₃₄-PNIPAAm₆₇ y PNIPAAm₆₇-MPEG₇₇-PNIPAAm₆₇. Además de éstos, en el laboratorio se sintetizó un polímero tribloque PNIPAAm₆₇-MPEG₁₆₅-PNIPAAm₆₇, que no corresponde exactamente con ninguno de los polímeros simulados, pero cuyo comportamiento debe ser similar al exhibido por los modelos 32-75-32 y 32-100-32.

Como ya se ha comentado en el caso de los copolímeros dibloque, al aumentar la temperatura, el aumento de la hidrofobicidad del bloque de PNIPAAm produce una transición desde el estado de ovillo al azar hasta una conformación globular, acompañada de una reducción del tamaño global de la molécula. En las figuras 7.31 y 7.32 se han representado los valores correspondientes al radio hidrodinámico y al radio de giro de los modelos de los tribloque frente al valor de $\epsilon_{LJ,PNIPAAm}$, tal como se hizo con los copolímeros dibloque.

Tal como cabe esperar, el efecto en los copolímeros tribloque es similar al observado en los dibloque, presentando en todos los casos una contracción similar, dado que los tamaños de los bloques hidrofóbicos son similares en todos los casos. Esto confirma definitivamente la conclusión alcanzada con los copolímeros dibloque de que el tamaño de la cadena hidrofílica no influye en gran medida en las dimensiones de estos copolímeros.

Para estudiar el efecto de la longitud del conector y poder comparar cuantitativamente el efecto de contracción, se recurre nuevamente a la normalización de los valores del radio hidrodinámico frente al valor en condiciones de buen disolvente (es decir, el valor de R_h a la mínima temperatura simulada). Los resultados de dicha normalización se presentan en

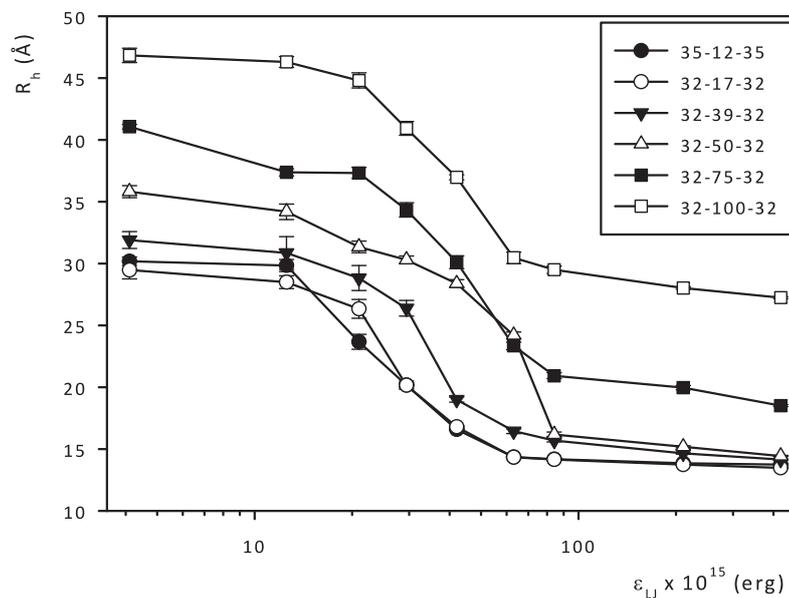


FIGURA 7.31: Evolución de los radios hidrodinámicos de los copolímeros tribloque frente a $\epsilon_{LJ,PNIPAAM}$.

la figura 7.33. En este gráfico se observa, en primer lugar, que el modelo con el conector más largo se desvía del comportamiento de los demás, pues no llega a alcanzarse el mismo grado de contracción que en el resto de casos. En segundo lugar, dado que un conector más largo significa una relación entre monómeros de MPEG y PNIPAAM más elevado, también se observa un desplazamiento de la temperatura de transición; es decir, la contracción comienza a observarse a un $\epsilon_{LJ,PNIPAAM}$ más bajo en los modelos con un conector más corto.

Finalmente, estudiando la contracción de esta serie de copolímeros desde el punto de vista conformacional a través de la relación R_g/R_h , recogida para los diferentes valores de $\epsilon_{LJ,PNIPAAM}$ en la figura 7.34, se observa algo parecido a lo que ya se había encontrado en los valores de R_h normalizados: el comportamiento del modelo 32-100-32 se desvía del resto.

Los copolímeros tribloque con los conectores de MPEG más cortos, 35-12-35, 32-17-32 y 32-39-32, presentan un comportamiento similar al que se había descrito también para los dibloque, una evolución desde $R_g/R_h \approx 1.2$ hasta $R_g/R_h \approx 0.7$, que indica el paso de una conformación similar a la de varillas rígidas hacia una conformación similar a esferas duras.

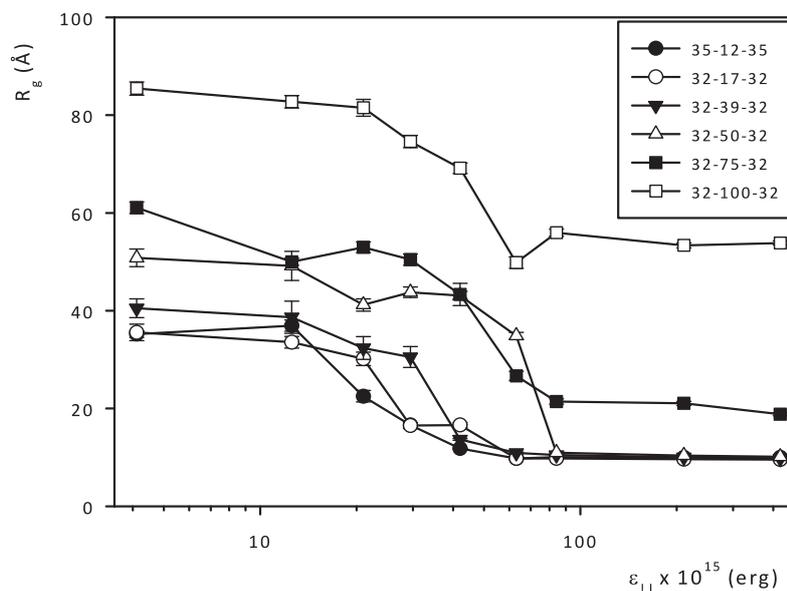


FIGURA 7.32: Evolución de los radios de giro de los copolímeros tribloque frente a $\epsilon_{LJ,PNIPAAM}$.

En cambio, los tribloque 32-50-32 y 32-75-32 presentan un valor inicial de $R_g/R_h \approx 1.5$, más próximo al de un ovillo al azar en condiciones de buen disolvente, probablemente debido al aumento de la flexibilidad global de la molécula debida al incremento de la longitud del espaciador central de MPEG. Sin embargo, entre estos dos modelos se observa una diferencia: mientras el 32-50-32 evoluciona de forma similar a los otros tribloque más cortos, colapsando al superar la temperatura de transición, y llegando a una conformación esférica con un $R_g/R_h \approx 0.7$, el copolímero 32-75-32 sólo llega a alcanzar un valor de $R_g/R_h \approx 1.0$, lo que indica que, aunque el polímero se contrae resultando en una conformación más compacta, ésta no llega a estar completamente plegada.

A diferencia de los otros modelos, el 32-100-32 presenta un valor de R_g/R_h casi constante de aproximadamente 1.7, a pesar de que, como se ha observado en la figura 7.33, también en este caso existe una contracción. Sin embargo, a pesar de contraerse, el copolímero conserva la conformación de ovillo al azar.

Finalmente, se puede recurrir al estudio de las distancias extremo-extremo de cada uno de los bloques que componen estos copolímeros. En los gráficos de la figura 7.35 se recogen las distancias extremo-extremo de los bloques de PNIPAAM terminales. Se puede observar que, aunque las dimensiones de ambos bloques no se comportan exactamente de la misma manera, no existe una diferencia significativa.

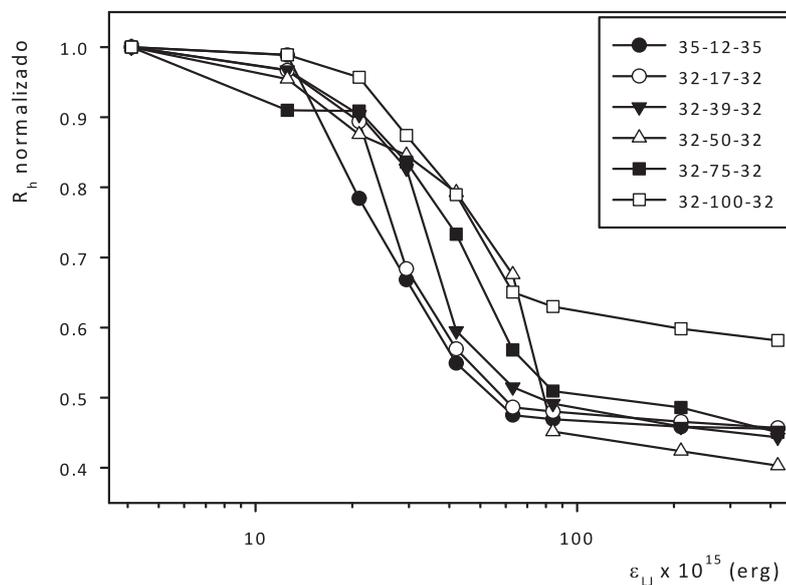


FIGURA 7.33: Evolución de los radios hidrodinámicos normalizados de los copolímeros tribloque frente a $\epsilon_{LJ,PNIPAAm}$.

En la figura 7.36 se recogen las distancias extremo-extremo de los bloques intermedios de MPEG para cada uno de los polímeros. Aquí, como cabe esperar, se observa un aumento en esta distancia conforme aumenta la longitud del conector. En general, se observa una tendencia a la contracción, aunque de nuevo se observan las excepciones de los casos 32-75-32 y 32-100-32.

En ambas figuras se observa que inicialmente, con $\epsilon_{LJ,PNIPAAm}$, los tres bloques se encuentran extendidos, y al aumentar la temperatura y la hidrofobicidad del bloque de PNIPAAm, se produce el colapso de estos bloques en los extremos de los copolímeros. En este punto, como puede observarse en las secuencias de imágenes que componen figura 7.37, se dan tres posibles situaciones.

En caso de que el bloque de MPEG sea lo suficientemente corto, ambos bloques de PNIPAAm sufren un colapso y se aglutinan, no siendo el bloque hidrofílico lo suficientemente largo como para rodearlos y proporcionar una protección adecuada. Esta situación es la que corresponde en la figura 7.37 a los modelos 35-12-35 y 35-17-35.

Se distingue un segundo caso para longitudes intermedias del bloque de MPEG. Es el caso de los modelos 35-39-35 y 35-50-35, en los que, como se puede apreciar en la figura, también se produce el colapso de los bloques hidrofóbicos en un núcleo único, pero

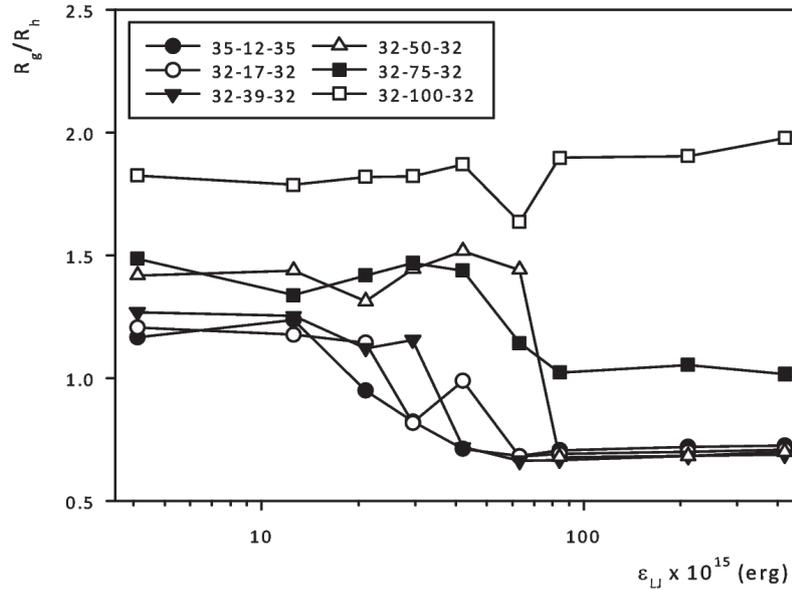


FIGURA 7.34: Factor R_g/R_h de los copolímeros tribloque frente a $\epsilon_{LJ,PNIPAAm}$.

a diferencia del caso anterior, el conector sí es lo suficientemente largo como para envolver el núcleo, formando una estructura de tipo *core-shell*.

Finalmente, en el caso de que el conector sea lo suficientemente largo, como en los casos 32-75-32 y 32-100-32, los bloques de PNIPAAm colapsan por separado, envolviéndose cada uno con parte del conector de MPEG, y no llegan a juntarse, formando una estructura con dos núcleos separados.

Al igual que en el caso de los copolímeros dibloque, en el laboratorio del grupo del profesor Nyström se sintetizaron y caracterizaron algunos de los copolímeros correspondientes a los modelos simulados. En concreto, se realizaron estudios de turbidimetría y de dispersión de luz dinámica (DLS). Sin embargo, ninguno de los resultados de estas técnicas es directamente comparable con los datos de las simulaciones, pues con las técnicas aplicadas no fue posible obtener resultados experimentales para cadenas aisladas de los polímeros estudiados, aunque algunos de los resultados obtenidos pueden explicarse mediante las conclusiones extraídas de los resultados de las simulaciones.

Los análisis de turbidez realizados producen resultados similares a los que se muestran en la figura 7.38. En el caso de la figura, se representan los análisis realizados a disoluciones al 0.1% en peso de los copolímeros PNIPAAm₆₉-MPEG₂₃-PNIPAAm₆₉,

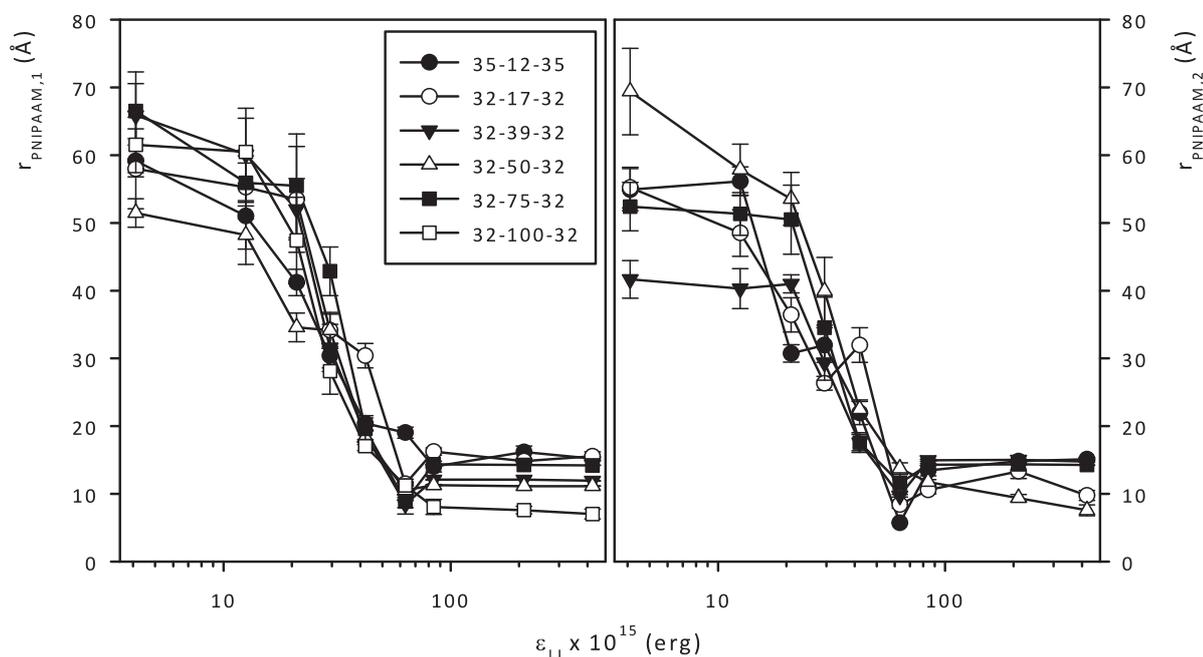


FIGURA 7.35: Distancias extremo-extremo de los bloques de PNIPAAAM en los copolímeros tribloque.

PNIPAAAM₆₇-MPEG₃₄-PNIPAAAM₆₇, PNIPAAAM₆₇-MPEG₇₇-PNIPAAAM₆₇ y PNIPAAAM₆₇-MPEG₁₆₅-PNIPAAAM₆₇.

En el caso de los tres primeros copolímeros se observa cómo al elevar la temperatura y aproximarse ésta a los 40 °C, se produce un súbito aumento de la turbidez de las disoluciones. Dicho fenómeno se atribuye a un aumento de la adherencia de las entidades hidrofóbicas, que daría lugar a la agregación intensiva y la formación de grandes complejos. Esto correspondería al comportamiento formulado para los modelos con la cadena de MPEG más corta, en el que los extremos de PNIPAAAM colapsan en un núcleo único que no puede ser apantallado por el bloque de MPEG, al ser éste demasiado corto para producir tal efecto.

Sin embargo, para el caso del copolímero PNIPAAAM₆₇-MPEG₁₆₅-PNIPAAAM₆₇ no se observa ningún aumento de la turbidez a temperaturas por debajo de 60 °C. La explicación de este comportamiento se asocia al comportamiento descrito anteriormente para los modelos 32-75-32 y 32-100-32, en los que la cadena de MPEG era lo suficientemente larga como para mantener separados los extremos de PNIPAAAM y protegerlos parcial o

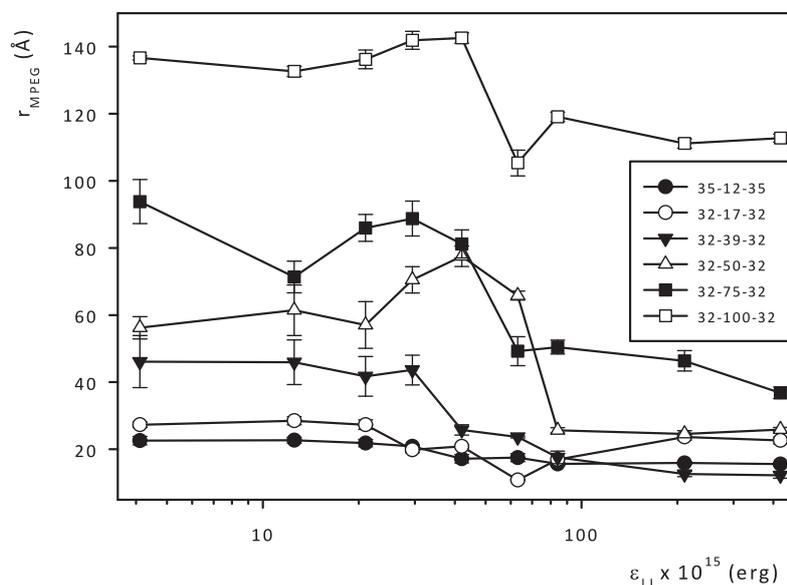


FIGURA 7.36: Distancia extremo-extremo del conector de MPEG en los copolímeros tribloque.

totalmente del contacto con el disolvente, y por tanto, reducir la probabilidad de formación de micelas.

En cuanto a los estudios mediante DLS, fueron realizados con concentraciones en peso del 1% (al igual que en el caso de los copolímeros dibloque, concentraciones menores no permitían obtener resultados fiables), y en todos los casos, a temperaturas bajas se observaron dos poblaciones con diferentes radios hidrodinámicos. Éstas se representan en la figura 7.39. La primera población corresponde al análisis de los tiempos de relajación rápidos, y está caracterizada por unos radios hidrodinámicos entre 4 y 6 nm, lo que podría corresponder al tamaño de las moléculas individuales.

La segunda población, correspondiente a los tiempos de relajación lentos, muestra tamaños mucho mayores (del orden de los cientos de nm), que se deben sin duda a grandes agregados formados por las asociaciones de las unidades PNIPAAm. La presencia de dichos agregados ya se ha confirmado con anterioridad en otros estudios de DLS [95], en los que se ha observado la formación de agregados incluso a temperaturas por debajo de la LCST.

En esta población de gran tamaño se observa que existe una tendencia a la disminución del tamaño de las micelas. Esto se explicaría mediante la contracción de los núcleos

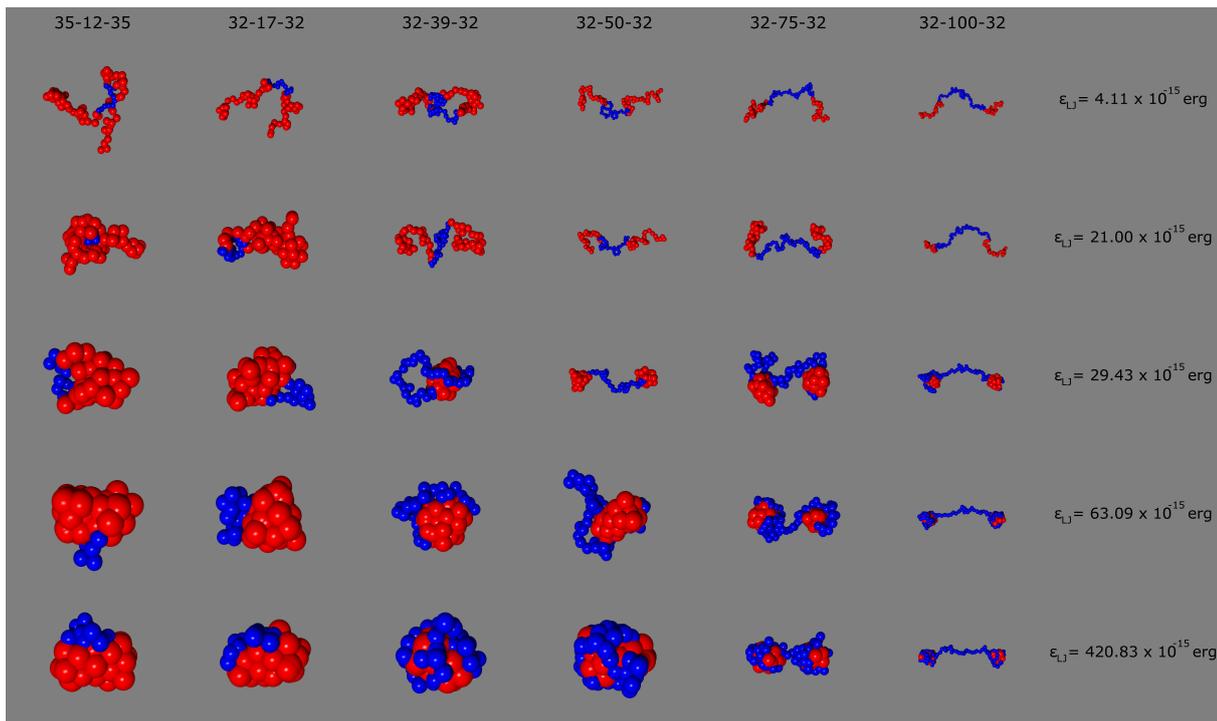


FIGURA 7.37: Evolución de los diferentes copolímeros tribloque frente a $\epsilon_{LJ,PNIPAAm}$.

hidrofóbicos, y como consecuencia, también disminuiría el radio hidrodinámico de los agregados. Sin embargo, para la población compuesta por unímeros no se observa ninguna contracción. Este comportamiento no coincidiría con el observado en las simulaciones, pero es necesario tener en cuenta que en éstas no se consideran los efectos de la interacción entre moléculas, ni la formación de micelas, mientras que en los ensayos experimentales existe en el seno de la disolución una coexistencia de unímeros, micelas y estructuras intermicelares.

Al aumentar la temperatura se observa cómo una de las poblaciones (la de menores radios hidrodinámicos) desaparece, probablemente debido al favorecimiento de los procesos de agregación que se produce al aumentar la temperatura, dando como resultado la pervivencia de una única población con radios hidrodinámicos elevados.

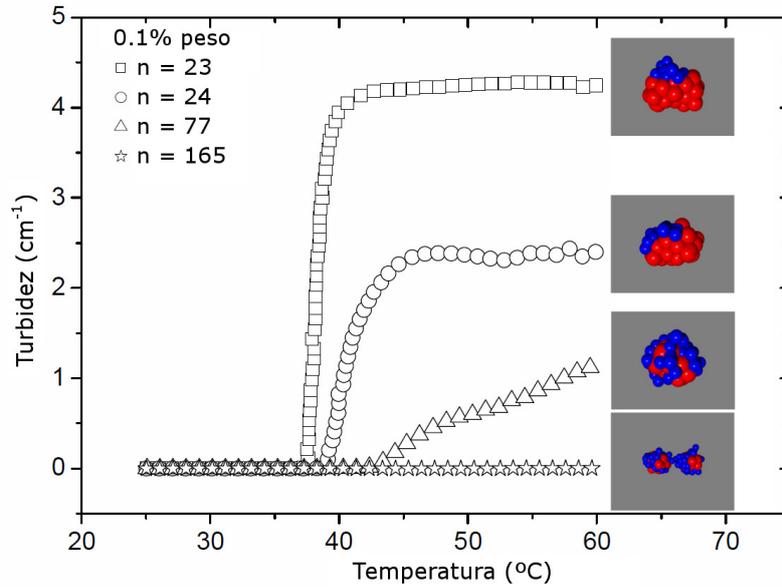


FIGURA 7.38: Turbidimetría en función de la temperatura. Las imágenes corresponden a los modelos de simulación que explicarían el efecto de la turbidez. El valor de n indica la longitud en monómeros del conector de MPEG.

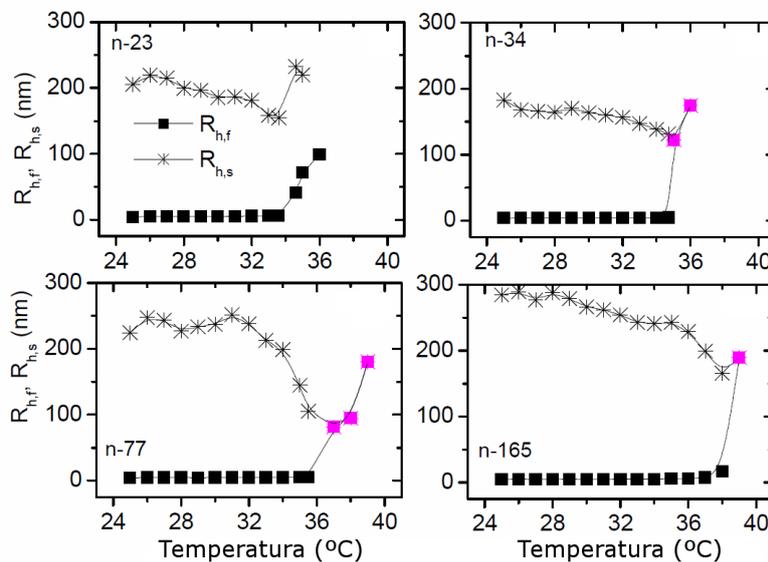


FIGURA 7.39: Radios hidrodinámicos promedio determinados a partir de los tiempos de relajación rápidos ($R_{h,f}$) y lentos ($R_{h,s}$). Los símbolos en color indican modos de relajación únicos. El valor de n indica la longitud en monómeros del conector de MPEG.

7.4. Conclusiones del Capítulo

A la vista de los resultados, y de la razonable coincidencia con los resultados experimentales, podemos decir que la aplicación de nuestro esquema de simulación multiescala a este problema, la simulación de cadenas lineales de composición heterogénea y condiciones de simulación heterogéneas, ha sido satisfactoria.

Se ha conseguido dar una explicación razonable a algunos de los fenómenos y dificultades observados en el trabajo de laboratorio, confirmando la hipótesis de la formación de estructuras de tipo core-shell que resultan en la formación de agregados, mediante la asociación de las partes expuestas del núcleo hidrofóbico, cuando la envoltura del núcleo por parte de la corona hidrofílica es insuficiente. Este efecto facilita la formación de agregados a temperaturas más bajas que en aquellos casos en los que la envoltura del núcleo es mejor.

El estudio de las simulaciones realizadas revela que, al aumentar la temperatura de una de éstas disoluciones, en primer lugar se produce un colapso del bloque hidrofóbico en un núcleo que se contrae progresivamente al aumentar la temperatura, mientras que el bloque hidrofílico forma una corona envolvente alrededor del núcleo hidrofóbico. Mientras se produce esta contracción, la molécula evoluciona desde una conformación estirada y relativamente rígida hacia una conformación esférica.

También a partir de las simulaciones se ha postulado que el tamaño del bloque hidrofílico tiene poca influencia en el grado de contracción alcanzado por el polímero a altas temperaturas. En cambio, la longitud del bloque de PNIPAAm tiene una gran importancia en el grado de contracción al que se somete la molécula al aumentar la temperatura, siendo éste el factor a tener en cuenta a la hora de diseñar moléculas que deban tener un tamaño determinado por encima de la LCST.

En lo que se refiere a los copolímeros tribloque, se observa que se produce aproximadamente el mismo comportamiento que en las moléculas dibloque en lo que se refiere a la formación de un núcleo hidrofóbico compacto al aumentar la temperatura, y por tanto también en la tendencia a la formación de agregados.

Las diferencias con los dibloque estriban en que en este caso hay dos extremos termosensibles, que, para el caso de conectores de MPEG de longitud corta e intermedia colapsan para formar un único núcleo hidrofóbico, que puede quedar mejor o peor recubierto por la cadena hidrofílica en función de la longitud de ésta. En cambio, a partir de un cierto grado de longitud del conector hidrófilo, los bloques hidrofóbicos colapsan individualmente, aislados el uno del otro, y se envuelven completamente en el conector. De esta manera, los núcleos están bastante apantallados y la tendencia a formar agregados a temperaturas más bajas se reduce. Para casos extremos en los que el conector sea muy largo, los extremos colapsados incluso pueden quedar separados una cierta distancia por un segmento de conector que no forma parte de la envoltura.

CAPÍTULO 8

CONCLUSIONES / CONCLUSIONS

Como se adelantaba en la introducción a esta Tesis, el presente Capítulo de conclusiones se ha redactado en inglés con el objetivo de cumplir con los requisitos para la Mención Internacional en el título de Doctor que se estipulan en el artículo 31 del Reglamento de las Enseñanzas Oficiales de Doctorado de la Universidad de Murcia, publicado en el Boletín oficial de la Región de Murcia 35/2012.

As it was stated in the foreword to this Thesis, the present Chapter has been written in English language to meet the requirements for the “International Mention” in the PhD title, as they are stipulated in article 31 of the Regulations for the Official Doctorate Studies in the University of Murcia, published in the Boletín Oficial de la Región de Murcia 35/2012.

From Chapter 4, on the study of the different methodologies for the calculation of the translational diffusion coefficient:

1. It has been confirmed that there is a relationship among the different approaches to the calculation of the translational diffusion coefficient, resulting in

$$D_K > D_L \approx D_E > D_{RB}. \quad (8.1)$$

This relationship is valid for any chain model which is not completely rigid (in that case, all approaches should yield the same value), so that D_K y D_{RB} are respectively an upper and a lower bound for the values of this coefficient, while D_L y D_E yield values very close to each other, and are intermediate to the previous ones.

2. From the studied approaches to the translational diffusion coefficient, D_{RB} is the most sensitive to the flexibility of the model, being the best approximation when the model has a high rigidity, and yielding slowly worsening results as the flexibility increases. In turn, the value of D_K , also based on the Kirkwood and Riseman theory poses as an always good approximation for the value of D_E , independently of the flexibility of the model.

From Chapter 5, on the comparison of algorithms for the Brownian Dynamics with Hydrodynamic Interaction:

3. In the conditions that were studied in the simulations in this work, among the various algorithms that were studied, the one presented by Geyer and Winter is the one that offers the best advantages: it is the fastest method, and is also the one which offers the best precision in its results. However, as it is an approximation-based method, it still implies a certain degree of error. The Ermak-McCammon procedure is the only studied method that produces exact solutions, which are valid in whichever conditions are used for the model and the simulation.
4. Both of the algorithms based on the Chebyshev expansion which have been evaluated offer the same degree of precision, but not the same speed: the method proposed by Kröger et al. is faster, but it requires that a pair of upper and a lower bounds for the eigenvalues for the diffusion tensor are provided prior to the simulation.
5. For the studied models, overlapping has a strong influence on the precision of the results produced by means of the algorithms based on the Chebyshev expansion,

so that these only produce an acceptable precision when the simulated models are guaranteed to avoid overlapping.

From Chapter 6, on the multiscale simulation of hyperbranched polymers:

6. In spite of the results in this chapter, we can only conclude that the method that our Group developed for the multiscale simulation of dendrimers is not directly applicable to the simulation of hyperbranched polymers. The most plausible causes are the lack of statistical representation of the generated samples respect to the ones from which the experimental data was extracted, and the possibility that the generated topologies are not representative of the real polymers whose experimental data we are comparing with our results. These hypothesis supporting that the averaging over multiple topologies is responsible of the not-matching results from simulations and bibliography is also supported by the good results attained in the last Chapter of this Thesis, in which the same multiscale method is used to simulate linear chains which do not require averaging over different topologies.
7. While the values of the properties found in the simulations and the experimental ones do not match well, the general tendency of a property in respect to the molecular weight do match. In spite of this, and the findings of Hobson y Feast [76] and other authors [73, 74] about the behaviour of the intrinsic viscosity in relation with the molecular weight, we can conclude that there is more than one possible behaviour, and the exact case is dependant on the nature of the studied hyperbranched polymer.

From Chapter 7, on the multiscale simulation of thermoresponsive amphiphilic copolymers:

8. Based on the plausibility of our results and the comparison with the experimental data available to us, we can conclude that our multiscale simulation method can successfully be applied to the simulation of this kind of systems.
9. Our results show that the studied diblock copolymers tend to assemble into *core-shell* structures when the temperature raises above the LCST. The formation of this kind of structures happens in two phases: in the first one, the hydrophobic block collapses and contracts, and then the hydrophilic block wraps around this core to shield it from the contact with the solvent. Along this transformation, the conformation of the molecule evolves into a more spherical shape.

10. From the point of view of the properties of these polymers above the LCST, there are two parameters of interest for their design: the size of the PNIPAAm block, which determines the magnitude of the contraction of the polymer when temperature raises, and the ratio between the lengths of hydrophobic and hydrophilic blocks, which determines the ability of the hydrophobic block to shield the hydrophobic core at higher temperatures. This factor is related to the tendency to form aggregates when approaching the LCST.
11. The behaviour of the triblock copolymers which are composed by two PNIPAAm blocks at the ends of a MPEG connector is similar to the shown by diblocks in that both kinds of polymers tend to assemble into *core-shell* structures. The difference with the diblocks consists in that the ending PNIPAAm blocks may collapse into a single hydrophobic core, which may have different degrees of wrapping and shielding depending on the length of the MPEG block. When the MPEG connector is long enough, each PNIPAAm block may collapse into individual cores (appropriately wrapped and shielded by the MPEG connector), both being separated a variable distance depending on the length of the MPEG connector.

GLOSARIO

Símbolos

Griegos y matemáticos

- $\langle \dots \rangle$: promedio del contenido del corchete.
- \mathbf{A}^T : Traspuesta de la matriz \mathbf{A} .
- \mathbf{a}^0 : valor de la variable \mathbf{a} previa a un paso de simulación.
- α : ángulo suplementario al ángulo instantáneo determinado por dos conectores de un modelo de bolas y muelles.
- α_e : ángulo suplementario al ángulo en el equilibrio determinado por dos conectores de un modelo de bolas y muelles.
- β : segundo parámetro arbitrario en un potencial de volumen excluido “suave”.
- $\beta_{ii} \equiv \beta, \beta_{ij} \equiv \beta'$: coeficientes correctores en el algoritmo de Geyer.
- δ_{ij} : función delta de Kronecker.
- Δ : máxima amplitud de un desplazamiento aleatorio en la técnica de Monte Carlo.
- Δt : intervalo de tiempo, generalmente, duración de un paso browniano.
- δ_i : i -ésimo número aleatorio, comprendido entre $-\Delta/2$ y $+\Delta/2$.
- ϵ : valor promedio de los elementos no diagonales del tensor de difusión.
- ϵ_{LJ} : parámetro energético en el potencial de Lennard-Jones.
- η_0 : viscosidad del disolvente.

- η_{rel} : viscosidad relativa.
- η_{sp} : viscosidad específica.
- η_{red} : viscosidad reducida.
- $[\eta]$: viscosidad específica.
- $\lambda_{min}, \lambda_{max}$: valores propios mínimo y máximo del tensor de difusión.
- $\Lambda_{min}, \Lambda_{max}$: cotas mínima y máxima de los valores propios del tensor de difusión.
- ζ_i : coeficiente de fricción de la esfera i .
- σ, σ_i : radio de la bola i en un modelo de bolas y muelles. En caso de figurar sin subíndice, se entiende que todas las bolas del modelo tienen igual radio.
- σ_{HS} : parámetro de distancia en el potencial de Lennard-Jones.
- σ_{LJ} : parámetro de distancia en el potencial de Lennard-Jones.

Latinos

- a_l ($l = 0 \dots L$): coeficientes del desarrollo de Chebyshev.
- a_G : radio de giro equivalente para la propiedad radio de giro.
- a_l : radio de giro equivalente para la propiedad viscosidad intrínseca.
- a_T : radio de giro equivalente para la propiedad coeficiente de difusión traslacional.
- A : primer parámetro arbitrario en un potencial de volumen excluido “suave”.
- b : distancia de equilibrio entre dos bolas consecutivas de un modelo.
- \mathbf{B} : matriz genérica resultante de la factorización de \mathbf{D} .
- c : concentración de soluto en una disolución.
- C_i : vectores de escala en el algoritmo de Geyer.
- C_l ($l = 0 \dots L$): variable tensorial del desarrollo de Chebyshev.
- \mathbf{C} : matriz triangular superior resultante de la factorización de \mathbf{D} .
- d : desplazamiento del centro de masas.

- dt : duración del paso browniano de simulación.
- $d\mathbf{w}$: vector aleatorio con distribución gaussiana $N(0, dt)$.
- D_t : coeficiente de difusión traslacional.
- D_E : en el Capítulo 4, el coeficiente de difusión traslacional (D_t) calculado mediante la expresión de Einstein.
- D_K : en el Capítulo 4, el coeficiente de difusión traslacional (D_t) calculado a partir de la teoría de Kirkwood y Riseman.
- D_L : en el Capítulo 4, el coeficiente de difusión traslacional (D_t) calculado mediante la expresión de Liu y Dünweg.
- D_{RB} : en el Capítulo 4, el coeficiente de difusión traslacional (D_t) calculado empleando el tratamiento del Cuerpo Rígido.
- \mathbf{D} : tensor de difusión.
- \mathbf{D}_{ij} : submatriz 3x3 del tensor de difusión que relaciona los elementos i y j .
- \mathbf{E} : tensor de difusión transformado para que sus valores propios estén en el intervalo (0,1).
- E_f : En el algoritmo de Jendrejack para la interacción hidrodinámica, error relativo de la aproximación de $\mathbf{B}d\mathbf{w}$.
- E_r : En el algoritmo de Kröger para la interacción hidrodinámica, error relativo de la aproximación de \mathbf{S} mediante \mathbf{S}_L .
- f : funcionalidad de una bola en un modelo de molécula hiperramificada.
- f_j : fuerzas aleatorias no correlacionadas en el algoritmo de Geyer.
- f_t : coeficiente de fricción traslacional.
- F : funcionalidad máxima de una bola en un modelo de molécula hiperramificada.
- \mathbf{F}_i : fuerza que actúa sobre el elemento i .
- H : constante de muelle de un conector de un modelo de bolas y muelles.
- \mathbf{I} : matriz identidad.
- k_1, k_2 : auxiliar para la transformación de \mathbf{D} en \mathbf{E} .

- k_B : constante de Boltzmann.
- \mathbf{K} : matriz de flujo con las velocidades del fluido en la posición de los elementos del modelo.
- l : longitud instantánea de un conector de un modelo de bolas y muelles.
- l_e : longitud de equilibrio de un conector de un modelo de bolas y muelles.
- l_{max} : longitud máxima de un conector de un modelo de bolas y muelles.
- L : número de términos en el desarrollo de Chebyshev.
- m_i : masa del elemento i .
- N_A : número de Avogadro.
- p : en la generación de hiperramificados, probabilidad de que un nuevo nodo entre sobre otro que dispone de un sólo “conector” libre.
- Q : constante de flexión de un conector angular de un modelo de bolas y muelles.
- r_i : ($i \in MPEG, PNIPAAM$) distancia extremo-extremo del bloque i de un copolímero termosensible.
- r_{ij} : distancia entre los elementos i y j de un modelo de bolas y muelles.
- \mathbf{r}_i : vector de posición del elemento i de un modelo de bolas y muelles.
- \mathbf{r}_{CM} : Desplazamiento cuadrático del centro de masas.
- R : parámetro de ramificación: número de ramificaciones en una molécula.
- R_g : radio de giro.
- R_{max} : parámetro de ramificación máxima posible en una molécula hiperramificada.
- \mathbf{R}_i : vector aleatorio con distribución gaussiana con media cero y covarianza $2\mathbf{D}_{ij}\Delta t$.
- s : radio de giro instantáneo.
- \mathbf{S} : matriz simétrica resultante de la factorización de \mathbf{D} .
- $\langle s^2 \rangle$: radio de giro cuadrático promedio.
- t_{CPU} : tiempo de CPU.
- T : temperatura absoluta.

- **T**: tensor de interacción hidrodinámica.
- U_i, U : (i -ésimo) número aleatorio comprendido en el intervalo $(0, 1)$.
- V : energía potencial.
- \mathbf{v}_i : velocidad del fluido en la posición del elemento i .
- \mathbf{y}, \mathbf{y}_L : desplazamiento estocástico; ídem, como resultado del sumatorio del desarrollo de Chebyshev.

Acrónimos

- DB: dinámica browniana. En hiperramificados, grado de ramificación, *degree of branching*.
- FENE: conector de extensibilidad finita y elongación no lineal (del inglés “*Finite Extensibility non-linear elongation*”).
- IH: interacción hidrodinámica.
- LJ: Lennard-Jones (generalmente, potencial de).
- MC: Monte Carlo.
- RB: tratamiento de cuerpo rígido (del inglés “*Rigid Body*”).
- RBMC: tratamiento de cuerpo rígido con muestreo de Monte Carlo.
- VE: volumen excluido.
- W: Índice de Wiener.

BIBLIOGRAFÍA

- [1] M. Rubinstein and R.H. Colby. *Polymer Physics*. Oxford University Press, New York, 2003.
(Citado en páginas 1 y 6.)
- [2] S.F. Sun. *Physical Chemistry of Macromolecules (2nd Ed.)*. John Wiley and Sons, New Jersey, 2004.
(Citado en páginas 1, 7, 11 y S-3.)
- [3] G. Patterson. *Physical Chemistry of Macromolecules*. CRC Press, Boca Raton, 2007.
(Citado en página 1.)
- [4] H.C. Öttinger. *Stochastic Processes in Polymer Fluids*. Springer, Berlin, 1996.
(Citado en páginas 2, 58, 69, 90 y S-7.)
- [5] M. Kröger. *Models for Polymeric and Anisotropic Liquids*. Springer, Heidelberg, 2005.
(Citado en página 2.)
- [6] D.L. Ermak and J.A. McCammon. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.*, 69:1352–1360, 1978.
(Citado en páginas 3, 25, 59, 70, 74, 89 y S-7.)
- [7] G. del Río Echenique. *Simulación multiescala de la conformación y dinámica de moléculas dendríticas en disolución*. PhD thesis, Universidad de Murcia, 2008.
(Citado en páginas 3 y 5.)
- [8] A. Horta. *Macromoléculas*, volume 1. UNED, Madrid, 1982.
(Citado en página 6.)
- [9] R.B. Bird, W.E. Stewart, and E.N. Lightfoot. *Fenómenos de transporte*. Ed. Reverté, Barcelona, 1993.
(Citado en página 7.)

- [10] K.E. van Holde, W. Johnson, and P. Ho. *Principles of Physical Biochemistry, 2nd edition*. Prentice Hall, Upper Saddle River, N.J., 1998.
(Citado en página 7.)
- [11] A. Einstein. On the movement of small particles suspended in a stationary liquid, demanded by the molecular-kinetic theory of heat. *Ann. Physik*, 17:549–560, 1905. (Traducción al inglés en “Investigation on the Theory of the Brownian Movement”, Dover, New York, 1956).
(Citado en páginas 7 y 24.)
- [12] R.F. Probstein. *Physicochemical Hydrodynamics. An Introduction (2nd Ed.)*. Wiley, New Jersey, 2003.
(Citado en páginas 7 y S-3.)
- [13] I.A. Katime and C. Cesteros. *Química Física Macromolecular II. Disoluciones y Estado Sólido*. Servicio Editorial Universidad del País Vasco, Zarautz, Guipuzkoa, 2002.
(Citado en página 7.)
- [14] H. Brenner. Coupling between the translational and rotational Brownian motions of rigid particles of arbitrary shape. II. General considerations. *J. Colloid Interface Sci.*, 23:407–436, 1967.
(Citado en página 8.)
- [15] J. García de la Torre and V.A. Bloomfield. Hydrodynamic properties of complex, rigid, biological macromolecules. Theory and applications. *Quart. Rev. Biophys.*, 14:81–139, 1981.
(Citado en páginas 8, 54 y 55.)
- [16] J. García de la Torre. Rotational diffusion coefficients. In S. Krause, editor, *Molecular Electro-Optics*, pages 75–103. Ed. Plenum Press, New York, 1981.
(Citado en página 8.)
- [17] A. Einstein. A new determination of molecular dimensions. *Ann. Physik*, 19:289–306, 1906. (Traducción al inglés en “Investigation on the Theory of the Brownian Movement”, Dover, New York, 1956).
(Citado en página 9.)
- [18] A. Einstein. A new determination of molecular dimensions. Corrections. *Ann. Physik*, 34:591–592, 1911. (Traducción al inglés en “Investigation on the Theory of the Brownian Movement”, Dover, New York, 1956).
(Citado en página 9.)

- [19] B. Carrasco and J. García de la Torre. Universal size-independent quantities for rigid and flexible macromolecules. *Prog. Colloid Polym. Sci.*, 113:81–86, 1999.
(Citado en página 10.)
- [20] A. Ortega and J. García de la Torre. Equivalent radii and ratios of radii from solution properties as indicators of macromolecular conformation, shape, and flexibility. *Biomacromolecules*, 8:2464–2475, 2007.
(Citado en páginas 10 y 11.)
- [21] J. García de la Torre, H.E. Pérez Sánchez, A. Ortega, J.G. Hernández Cifre, M.X. Fernandes, F.G. Díaz Baños, and M.C. López Martínez. Calculation of the solution properties of flexible macromolecules: methods and applications. *Eur. Biophys. J.*, 32:477–486, 2003.
(Citado en página 12.)
- [22] G.K. Fraenkel. Visco-elastic effect in solutions of simple particles. *J. Chem. Phys.*, 20:642–647, 1952.
(Citado en página 13.)
- [23] H.R. Warner. Kinetic-theory and rheology of dilute suspensions of finitely extendible dumbbells. *Ind. Eng. Chem. Fundam.*, 11:379–387, 1972.
(Citado en página 14.)
- [24] G. Del Río Echenique, R. Rodríguez Schmidt, J.J. Freire, J.G. Hernández Cifre, and J. García de la Torre. A multiscale scheme for the simulation of conformational and solution properties of different dendrimer molecules. *J. Am. Chem. Soc.*, 131:8548–8556, 2009.
(Citado en páginas 14, 17, 112, 128, 130, 163, S-13 y S-22.)
- [25] A. Rey, J.J. Freire, and J. García de la Torre. Translational diffusion coefficients, relaxation times and quasielastic light scattering of flexible chains. *Macromolecules*, 24:4666–4672, 1991.
(Citado en páginas 15, 53, 54, 62, 89 y S-3.)
- [26] G.S. Ayton and G.A. Voth. Simulation of biomolecular systems at multiple length and time scales. *Int. J. Mult. Comp. Eng.*, 2:291–311, 2004.
(Citado en página 17.)
- [27] A.N. Gorban, N.K. Kazantzis, I.G. Kevrekidis, H.C. Öttinger, and C. Theodoropoulos, editors. *Model Reduction and Coarse-Graining Approaches for Multiscale Phenomena*. Springer, Berlin, 2006.
(Citado en página 17.)

- [28] Hypercube Inc. <http://www.hyper.com/>.
(Citado en página 18.)
- [29] B. Carrasco and J. García de la Torre. Hydrodynamic properties of rigid particles. Comparison of different modelling and computational procedures. *Biophys. J.*, 76:3044–3057, 1999.
(Citado en página 19.)
- [30] J. García de la Torre, M.L. Huertas, and B. Carrasco. Calculation of hydrodynamic properties of globular proteins from their atomic-level structures. *Biophys. J.*, 78:719–730, 2000.
(Citado en página 19.)
- [31] P. Espinosa and J. García de la Torre. Theoretical prediction of translational diffusion coefficients of small rigid molecules. *J. Phys. Chem.*, 91:3612–3616, 1987.
(Citado en página 19.)
- [32] A. Bondi. Van der Waals volumes and radii. *J. Phys. Chem.*, 68:441–451, 1964.
(Citado en página 19.)
- [33] A. Iniesta and J. García de la Torre. A second-order algorithm for the simulation of the Brownian dynamics of macromolecular models. *J. Chem. Phys.*, 92:2015–2019, 1990.
(Citado en páginas 26, 27 y 70.)
- [34] G. Chirico and J. Langowski. Calculating the hydrodynamic properties of DNA through a second-order Brownian dynamics algorithm. *Macromolecules*, 25:769–775, 1992.
(Citado en páginas 27, 69 y 70.)
- [35] A.C. Branka and D.M. Heyes. Algorithms for Brownian dynamics computer simulation. multivariate case. *Phys. Rev. E.*, 60:2381–2387, 1999.
(Citado en página 27.)
- [36] J.G. Hernández Cifre and J. García de la Torre. Steady-state behavior of dilute polymers in elongational flow. Dependence of the critical elongational rate on chain length, hydrodynamic interaction and excluded volume. *J. Rheol.*, 43:339–358, 1999.
(Citado en páginas 28, 69, 70 y S-7.)
- [37] J. Rotne and S. Prager. Variational treatment of hydrodynamic interaction on polymers. *J. Chem. Phys.*, 50:4831–4837, 1969.
(Citado en páginas 30, 58, 70 y 73.)

- [38] H. Yamakawa. Transport properties of polymer chains in dilute solution: hydrodynamic interaction. *J. Chem. Phys.*, 53:436–443, 1970.
(Citado en páginas 30 y 58.)
- [39] J. García de la Torre and V.A. Bloomfield. Hydrodynamic properties of macromolecular complexes. I. Translation. *Biopolymers*, 16:1747–1763, 1977.
(Citado en página 31.)
- [40] N. Metropolis, A.W. Metropolis, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculation by fast computing machines. *J. Chem. Phys.*, 21:1087–1093, 1953.
(Citado en página 32.)
- [41] H.C. Öttinger. Generalized zimm model for dilute polymer solutions under theta conditions. *J. Chem. Phys.*, 86:3731–3749, 1987.
(Citado en página 35.)
- [42] J. García de la Torre, J.G. Hernández Cifre, A. Ortega, R. Rodríguez Schmidt, M.X. Fernandes, H. E. Pérez Sánchez, and R. Pamies. SIMUFLEX : Algorithms and tools for simulation of the conformation and dynamics of flexible molecules and nanoparticles in solution. *J. Chem. Theor. Comput.*, 5:2606–2618, 2009.
(Citado en páginas 44 y 46.)
- [43] J. García de la Torre, A. Ortega, H.E. Pérez Sánchez, and J.G. Hernández Cifre. MULTIHYDRO and MONTEHYDRO: Conformational search and Monte Carlo calculation of solution properties of rigid and flexible macromolecular models. *Biophys. Chem.*, 116:121–128, 2005.
(Citado en páginas 44 y 47.)
- [44] B.H. Zimm. Chain molecule hydrodynamics by the Monte-Carlo method and the validity of the Kirkwood-Riseman approximation. *Macromolecules*, 13:592–602, 1980.
(Citado en páginas 47, 54 y S-3.)
- [45] M. Doi and S.F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, Oxford, 1986.
(Citado en página 51.)
- [46] J.G. Kirkwood and J. Riseman. The intrinsic viscosities and diffusion constants of flexible macromolecules in solution. *J. Chem. Phys.*, 16:565–573, 1948.
(Citado en páginas 53 y S-3.)
- [47] J.G. Kirkwood. The general theory of irreversible processes in solutions of macromolecules. *J. Polym. Sci.*, 12:1–14, 1954.
(Citado en páginas 53 y S-3.)

- [48] M. Fixman. Variational bounds for polymer transport coefficients. *J. Chem. Phys.*, 78:1588–1593, 1983.
(Citado en páginas 54 y S-3.)
- [49] B. Liu and B. Dünweg. Translational diffusion of polymer chains with excluded volume and hydrodynamic interactions by Brownian dynamics simulation. *J. Chem. Phys.*, 118:8061–8072, 2003.
(Citado en páginas 54, 55, 56, 57, 62 y S-3.)
- [50] J. García de la Torre, S. Navarro, M.C. López Martínez, F.G. Díaz, and J.J. López Cascales. HYDRO: a computer software for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.*, 67:530–531, 1994.
(Citado en página 55.)
- [51] M. Fixman. Inclusion of hydrodynamic interaction in polymer dynamical simulations. *Macromolecules*, 14:1710–1717, 1981.
(Citado en páginas 55 y S-3.)
- [52] M.L. Huertas, S. Navarro, M.C. López Martínez, and J. García de la Torre. Simulation of the conformation and dynamics of a double-helical model for DNA. *Biophys. J.*, 73:3142–3153, 1997.
(Citado en páginas 59 y 74.)
- [53] D. Petera and M. Muthukumar. Brownian dynamics simulation of bead-rod chains under shear with hydrodynamic interaction. *J. Chem. Phys.*, 111:7614–7623, 1999.
(Citado en página 69.)
- [54] C.-C. Hsieh, L. Li, and R.G. Larson. Modeling hydrodynamic interaction in Brownian dynamics: simulations of extensional flows of dilute solutions of DNA and polystyrene. *J. Non-Newtonian Fluid Mech.*, 113:147–191, 2003.
(Citado en páginas 69 y 70.)
- [55] C.M. Schroeder, E.S.G. Shaqfeh, and S. Chu. Effect of hydrodynamic interactions on DNA dynamics in extensional flow: simulation and single molecule experiment. *Macromolecules*, 37:9242–9256, 2004.
(Citado en páginas 69 y 70.)
- [56] H. Jian, T. Schlick, and A. Vologodskii. Internal motion of supercoiled DNA: Brownian dynamics simulations of site juxtaposition. *J. Mol. Biol.*, 284:287–296, 1998.
(Citado en páginas 69 y 70.)

- [57] T. Frembgen-Kesner and A.H. Elcock. Striking effects of hydrodynamic interactions on the simulated diffusion and folding of proteins. *J. Chem. Theory Comput.*, 5:242–256, 2009.
(Citado en páginas 69 y 70.)
- [58] M. Fixman. Implicit algorithm for Brownian dynamics of polymers. *Macromolecules*, 19:1195–1204, 1986.
(Citado en páginas 70, 75, 81, 95, S-7 y S-8.)
- [59] M. Kröger, A. Alba Pérez, M. Laso, and H.C. Öttinger. Variance reduced Brownian simulation of a bead-spring chain under steady shear flow considering hydrodynamic interaction effects. *J. Chem. Phys.*, 113:4767–4773, 2000.
(Citado en páginas 70, 71, 80, 81, 89, 90 y S-7.)
- [60] R.M. Jendrejack, M.D. Graham, and J.J. de Pablo. Hydrodynamic interactions in long chain polymers: Application of the Chebyshev polynomial approximation in stochastic simulations. *J. Chem. Phys.*, 113:2894–2900, 2000.
(Citado en páginas 70, 71, 77, 89, 90, 92 y S-7.)
- [61] T. Geyer and U. Winter. A $O(N^2)$ approximation for hydrodynamic interactions in Brownian dynamics simulations. *J. Chem. Phys.*, 130:1149051–8, 2009.
(Citado en páginas 70, 71, 83, 89, 91, 101, S-7 y S-8.)
- [62] R. Rzehak, W. Kromen, T. Kawakatsu, and W. Zimmermann. Deformation of a tethered polymer in uniform flow. *Eur. Phys. J. E*, 2:3–30, 2000.
(Citado en página 70.)
- [63] J. Huang and T. Schlick. Macroscopic modeling and simulations of supercoiled DNA with bound proteins. *J. Chem. Phys.*, 117:8573–8586, 2002.
(Citado en página 70.)
- [64] D.M. Heyes. Mean-field-hydrodynamics Brownian dynamics simulations of viscosity and self-diffusion of near-hard-sphere colloidal liquids. *J. Phys.: Condens. Matter*, 7:8857–8865, 1995.
(Citado en página 70.)
- [65] A.J. Banchio and J.F. Brady. Accelerated Stokesian dynamics: Brownian motion. *J. Chem. Phys.*, 118:10323–1–10323–10, 2003.
(Citado en página 70.)
- [66] K. Klenin, H. Merlitz, and J. Langowski. A Brownian dynamics program for the simulation of linear and circular DNA and other wormlike chain polyelectrolytes. *Biophys. J.*, 74:780–788, 1998.
(Citado en página 74.)

- [67] W.H. Press, B.P. Flannery, S.A. Teukolsky, and M.T. Vetterling. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1986.
(Citado en página 76.)
- [68] J.J. Freire, J. Pla, A. Rey, and R. Prats. Monte Carlo calculations for linear and star polymers with intramolecular interactions. 1. Dimensions. *Macromolecules*, 19:452–457, 1986.
(Citado en páginas 90, 104, 178 y 179.)
- [69] H. Yamakawa. *Modern Theory of Polymer Solutions*. Harper and Row, New York, 1971.
(Citado en página 91.)
- [70] E. de Luca and R.W. Richards. Molecular characterization of a hyperbranched polyester. I. Dilute solution properties. *J. Pol. Sci. B*, 41:1339–1351, 2003.
(Citado en páginas 111, 127, 160, S-13 y S-14.)
- [71] E. de Luca, R.W. Richards, I. Grillo, and S.M. King. Molecular characterization of a hyperbranched polyester. II. Small-Angle Neutron Scattering. *J. Pol. Sci. B*, 41:1352–1361, 2003.
(Citado en páginas 111, 160, S-13 y S-14.)
- [72] C.J. Hawker, R. Lee, and J.M.J. Frechet. One-step synthesis of hyperbranched dendritic polyesters. *J. Am. Chem. Soc.*, 113(12):4583–4588, 1991.
(Citado en páginas 111, 114, 118 y S-13.)
- [73] D. Parker and W.J. Feast. Synthesis, structure, and properties of hyperbranched polyesters based on dimethyl 5-(2-hydroxyethoxy)isophthalate. *Macromolecules*, 34:2048–2059, 2001.
(Citado en páginas 111, 112, 211 y S-13.)
- [74] S.R. Turner, B.I. Voit, and T.H. Mourey. All-aromatic hyperbranched polyesters with phenol and acetate end groups: synthesis and characterization. *Macromolecules*, 26:4617–4623, 1993.
(Citado en páginas 111, 112, 211 y S-13.)
- [75] Y.H. Kim and O.W. Webster. Hyperbranched polyphenylenes. *Macromolecules*, 25:5561–5572, 1992.
(Citado en páginas 111 y S-13.)
- [76] L.J. Hobson and W.J. Feast. Poly(amidoamine) hyperbranched systems: synthesis, structure and characterization. *Polymer*, 40:1279–1297, 1999.
(Citado en páginas 111, 112, 157, 211, S-13 y S-14.)

- [77] E.B. Tarabukina, A.A. Shpyrkov, D.V. Potapova, A.P. Filippov, N.A. Shumilkina, and A.M. Muzafarov. Hydrodynamic and conformational properties of a hyperbranched polymethylallylcarbosilane in dilute solutions. *Polymer Science, Ser. A*, 48(9):974–980, 2006.
(Citado en páginas 111, 158, S-13 y S-14.)
- [78] J. Aerts. Prediction of intrinsic viscosities of dendritic, hyperbranched and branched polymers. *Comp. and Theor. Polymer Science*, 8:49–54, 1998.
(Citado en página 112.)
- [79] A.H. Widmann and G.R. Davies. Simulation of the intrinsic viscosity of hyperbranched polymers built by sequential addition. *Comp. Theor. Polym. Sci.*, 8(1/2):191–199, 1998.
(Citado en páginas 112 y 119.)
- [80] A.V. Lyulin, D.B. Adolf, and G.R. Davies. Computer simulations of hyperbranched polymers in shear flows. *Macromolecules*, 31(11):3783–3789, 2001.
(Citado en páginas 112 y 114.)
- [81] P.F. Sheridan, D.B. Adolf, A.V. Lyulin, I. Neelov, and G.R. Davies. Computer simulations of hyperbranched polymers: The influence of the Wiener index on the intrinsic viscosity and radius of gyration. *J. Chem. Phys.*, 117(16):7802–7812, 2002.
(Citado en páginas 112 y 119.)
- [82] G.K. Dalakoglou, K. Karatasos, S.V. Lyulin, and A.V. Lyulin. Brownian dynamics simulations of complexes of hyperbranched polymers with linear polyelectrolytes: Effects of the strength of electrostatic interactions on static properties. *Material Science and Engineering B*, 152:114–118, 2008.
(Citado en página 112.)
- [83] T. Mulder, A.V. Lyulin, P. van der Schoot, and M.A.J. Michels. Architecture and conformation of uncharged and charged hyperbranched polymers: Computer simulation and mean-field theory. *Macromolecules*, 38(3):996–1006, 2005.
(Citado en página 112.)
- [84] E. Rodríguez, J.J. Freire, G. del Río Echenique, J.G. Hernández Cifre, and J. García de la Torre. Improved simulation method for the calculation of the intrinsic viscosity of some dendrimer molecules. *Polymer*, 48:1155–1163, 2007.
(Citado en página 113.)
- [85] D. Hölder, A. Burgath, and H. Frey. Degree of branching in hyperbranched polymers. *Acta Polym.*, 48:30–35, 1997.
(Citado en páginas 115, 117, 118 y S-14.)

- [86] H. Wiener. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.*, 69(1):17–20, 1947.
(Citado en página 118.)
- [87] E.B. Tarabukina, A.A. Shpyrkov, E.V. Tarasova, A.I. Amirova, A.P. Filippov, N. A. Sheremet'eva, and A.M. Muzafarov. Effect of the length of branches on hydrodynamic and conformational properties of hyperbranched polycarbosilanes. *Polymer Science, Ser. A*, 51(2):150–160, 2009.
(Citado en páginas 158, 159 y S-14.)
- [88] M. Muthukumar, C.K. Ober, and E.L. Thomas. Competing interactions and levels of ordering in self-organizing polymeric materials. *Science*, 277:1225–1232, 1997.
(Citado en páginas 165 y S-21.)
- [89] X.P. Qiu and C. Wu. Study of the core-shell nanoparticle formed through the “coil-to-globule” transition of poly(N-isopropylacrylamide) grafted with poly(ethylene oxide). *Macromolecules*, 30:7921–7926, 1997.
(Citado en páginas 165 y S-21.)
- [90] S. Förster and M. Antonietti. Amphiphilic block copolymers in structure-controlled nanomaterial hybrids. *Adv. Mater.*, 10:195, 1998.
(Citado en páginas 165 y S-21.)
- [91] S. Minko, editor. *Responsive Polymer Materials: Design and Applications*. Blackwell Publishing, Ames, Iowa, 2006.
(Citado en páginas 165 y S-21.)
- [92] M. Lazzari, G. Liu, and S. Lecommandoux, editors. *Block Copolymers in Nanoscience*. Wiley-VCH, Weinheim, Germany, 2006.
(Citado en páginas 165 y S-21.)
- [93] M. Malmsten. *Surfactants and Polymers in Drug Delivery*. Marcel Drekker, New York, 2002.
(Citado en páginas 165 y S-21.)
- [94] H.G. Schild. Poly(N-isopropylacrylamide): experiment, theory and application. *Prog. Polym. Sci.*, 17:163–249, 1992.
(Citado en páginas 165 y 166.)
- [95] R. Pamies, K. Zhu, A.-L. Kjøniksen, and B. Nyström. Thermal response of low molecular weight poly-(N-isopropylacrylamide) polymers in aqueous solution. *Polym. Bull.*, 62:487–502, 2009.
(Citado en páginas 166 y 203.)

- [96] H.W. Chen, Q.J. Zhang, J.F. Li, Y.W. Ding, G.Z. Zhang, and C. Wu. Formation of mesoglobular phase of PNIPAM-g-PEO copolymer with a high PEO content in dilute solutions. *Macromolecules*, 38:8045–8050, 2005.
(Citado en páginas 166 y 181.)
- [97] K. Zhu, H. Jin, A.-L. Kjøniksen, and B. Nyström. Anomalous transition in aqueous solutions of a thermoresponsive amphiphilic diblock copolymer. *J. Phys. Chem. B*, 111:10862–10870, 2007.
(Citado en página 166.)
- [98] C. Wu and S.Q. Zhou. Laser-light scattering study of the phase-transition of poly(N-isopropylacrylamide) in water. 1. Single-chain. *Macromolecules*, 28:8381–8387, 1995.
(Citado en páginas 166, 180 y S-23.)
- [99] A. Rey, J.J. Freire, and J. García de la Torre. Monte Carlo calculations for linear and star polymers. 3. Dimensions and hydrodynamic properties in good solvents. *Macromolecules*, 20:342–346, 1987.
(Citado en página 178.)
- [100] R. Pamies, M.C. López Martínez, J.G. Hernández Cifre, and J. García de la Torre. Non-newtonian viscosity of dilute polymer solutions. *Macromolecules*, 38:1371–1377, 2005.
(Citado en página 180.)
- [101] T. Schnabel, J. Vrabec, and H. Hasse. Unlike Lennard-Jones parameters for vapor-liquid equilibria. *J. Mol. Liq.*, 135:170–178, 2007.
(Citado en página 181.)
- [102] J. Virtanen, S. Holappa, H. Lemmetyinen, and H. Tenhu. Aggregation in aqueous poly(N-isopropylacrylamide)-block-poly(ethylene oxide) solutions studied by fluorescence spectroscopy and light scattering. *Macromolecules*, 35:4763–4769, 2002.
(Citado en página 181.)
- [103] K. Zhu, R. Pamies, A.-L. Kjøniksen, and B. Nyström. Temperature-induced intermicellization of “hairy” and “crew-cut” micelles in an aqueous solution of a thermoresponsive copolymer. *Langmuir*, 24:14227–14233, 2008.
(Citado en páginas 186 y S-26.)
- [104] A.Z. Akcasu and C.C. Han. Molecular-weight and temperature-dependence of polymer dimensions in solution. *Macromolecules*, 12:276–280, 1979.
(Citado en páginas 195 y S-26.)

- [105] W. Burchard, M. Schmidt, and W.H. Stockmayer. Information on polydispersity and branching from combined quasi-elastic and integrated scattering. *Macromolecules*, 13:1265–1272, 1980.
(Citado en páginas 195 y S-26.)
- [106] J.F. Douglas, J. Roovers, and K.F. Freed. Characterization of branching architecture through universal ratios of polymer-solution properties. *Macromolecules*, 23:4168–4180, 1990.
(Citado en páginas 195 y S-26.)
- [107] C. Wu, M. Li, S.C.M. Kwan, and G.J. Liu. Laser light scattering characterization of a novel polymer nanofiber. *Macromolecules*, 31:7553–7554, 1998.
(Citado en páginas 195 y S-26.)
- [108] D.A. Xiong, L.Q. Shi, X.W. Jiang, Y.L. An, X. Chen, and J. Lu. Composite worm-like aggregates formed from a pair of block-copolymers containing hydrogen-bonding donor and acceptor. *Macromol. Rapid Commun.*, 28:194–199, 2007.
(Citado en páginas 195 y S-26.)
- [109] Hua Wei, Si-Xue Cheng, Xian-Zheng Zhang, and Ren-Xi Zhuo. Thermo-sensitive polymeric micelles based on poly(n-isopropylacrylamide) as drug carriers. *Progress in Polymer Science*, 34(9):893 – 910, 2009.
(Citado en página S-21.)

SUMMARY

This document contains a summary of the present PhD Thesis as required for the attainment of the “International Mention”, according to article 31 of the Regulations for the Official Doctorate Studies in the University of Murcia, published in the Boletín Oficial de la Región de Murcia 35/2012.

In this Thesis theoretical aspects of the numerical techniques employed to simulate macromolecules in solution are studied with the hope of improving the simulation methodology. There are also two chapters dedicated to the application of those simulation techniques to real macromolecules and the evaluation of the results with experimental data.

1. Comparative study of different approaches for the calculation of the translational diffusion coefficient D_t .

For the first Chapter of the Thesis we studied the relationship among four different approaches to the calculation of the translational diffusion coefficient of macromolecules. This gives rise to four coefficients: one obtained by means of the Einstein equation (D_E) [2, 12], another one obtained by approximating some equations of the theory of irreversible processes in solutions of macromolecules developed Kirkwood and Riseman (D_K) [46, 47], yet another by using the “Rigid Body treatment” (D_{RB}) [44] that is based on the previous theory but this time using exact solutions to the equations, and finally the expression for the long-time diffusion coefficient (D_L) defined by Fixman [48, 51] and further studied by Liu and Dünweg [49].

N	D_K^*	D_1^*	$D_L^* = D_K^* - D_1^*$	D_E^*	D_{RB}^*
11	0.2504 ± 0.0003	0.0043 ± 0.0001	0.2461 ± 0.0002	0.250 ± 0.002	0.2409 ± 0.0002
21	0.1785 ± 0.0002	0.0038 ± 0.0002	0.1747 ± 0.0003	0.174 ± 0.003	0.1662 ± 0.0003
50	0.1084 ± 0.0002	0.0027 ± 0.0003	0.1057 ± 0.0002	0.1054 ± 0.0003	0.1000 ± 0.0001
100	0.0727 ± 0.0001	0.0018 ± 0.0002	0.0709 ± 0.0003	0.0719 ± 0.0006	0.0668 ± 0.0003
200	0.0505 ± 0.0002	0.00145 ± 0.00005	0.0490 ± 0.0002	0.0495 ± 0.0002	0.0459 ± 0.0002

Table S.1: Translational diffusion coefficients for the linear flexible (pseudo-Gaussian) chains.

We simulated, on the one hand, chains with Gaussian springs and excluded volume (identical to those used in [25, 49]) and, on the other hand, semi-rigid wormlike chains (i.e. a quite flexible and a quite rigid model chain respectively). We found the results shown in tables S.1 and S.2.

From these results it is easy to see that the values of the different diffusion coefficients for semi-rigid models are much closer among them than the ones for flexible models. This agrees with the prediction that for a completely rigid molecule all approaches should yield the same value. But even in the case of the flexible models, the different results are so close that the differences are in the same order of magnitude than usual experimental error, so all values may be considered as good.

N	\mathbf{D}_K^*	\mathbf{D}_1^*	$\mathbf{D}_L^* = \mathbf{D}_K^* - \mathbf{D}_1^*$
11	0.184007 ± 0.000003	0.00221 ± 0.00002	0.18180 ± 0.00002
21	0.111575 ± 0.000001	0.00122 ± 0.00001	0.11035 ± 0.00001
50	0.055697 ± 0.000001	0.000568 ± 0.000006	0.05531 ± 0.00001
100	0.031492 ± 0.000001	0.00039 ± 0.00001	0.03110 ± 0.00001
200	0.017578 ± 0.000001	0.000137 ± 0.000004	0.01744 ± 0.00005

N	\mathbf{D}_E^*	\mathbf{D}_{RB}^*
11	0.181 ± 0.001	0.181673 ± 0.000003
21	0.1102 ± 0.0006	0.109966 ± 0.000001
50	0.0547 ± 0.0002	0.054716 ± 0.000001
100	0.0311 ± 0.0001	0.030793 ± 0.000002
200	0.01753 ± 0.00001	0.017176 ± 0.000001

Table S.2: Translational diffusion coefficients for the linear semi-rigid (wormlike) chains.

It can also be seen that there is a relationship among the different approaches to the calculation of the translational diffusion coefficient, such that

$$D_K > D_L \approx D_E > D_{RB}, \quad (\text{S.2})$$

As predicted by theory, D_K is an upper bound for the value of the diffusion coefficient, while D_{RB} is a lower bound. This last premise is a little bit harder to observe in the results for the smaller wormlike chains due to the proximity of the results and the uncertainty in the different approaches.

If considering D_E as the right value for the diffusion coefficient, we can calculate how much the other values deviate from it. This data are shown on table S.3. It can be seen that D_{RB} is quite sensitive to the flexibility of the model (that not only depends on the model, but also it grows with chain length), being the best approximation when the model has a high rigidity, and slowly worsening its results as the flexibility increases.

In turn, the value of D_K , based on the Kirkwood and Riseman theory, seems always a good approximation for the value of D_E , independently of the flexibility of the model. D_K is related to local phenomena with short relaxation times, which are independent of

N	Flexible model (type 1)			wormlike model (type 2)		
	D_K	D_L	D_{RB}	D_K	D_L	D_{RB}
11	2.9%	1.04%	3.1%	1.66%	0.44%	0.37%
21	2.7%	0.51%	4.3%	1.29%	0.18%	0.17%
50	2.9%	0.33%	5.1%	1.82%	0.79%	0.03%
100	1.08%	1.41%	7.2%	1.30%	0.04%	0.95%
200	2.02%	1.01%	7.3%	0.29%	0.50%	2.01%

Table S.3: Differences between the diffusion coefficients in tables [S.1](#) and [S.2](#) and the matching values of D_E , taken as reference.

the chain size. D_L , being related to long time scales is always a quite good approximation of D_E , and does not show any important trend related to flexibility.

2. A study of the algorithms for the treatment of Brownian Dynamics simulations with Hydrodynamic Interaction.

The Hydrodynamic Interaction (HI) [4] between elements of the simulated model is an important effect that must be included into the Brownian Dynamics (BD) simulations to reproduce the polymer's dynamics in dilute solution in a realistic way and reproduce experimental results [36]. At the same time that the inclusion of a fluctuating, conformation dependent, not pre-averaged HI is essential to produce realistic results, it usually leads to high computing power requirements and/or long simulation times.

For this reason, a number of methods have been developed alternative to that initially proposed by Ermak and McCammon in their pioneering work on a BD algorithm with HI [6] based on the Cholesky matrix factorization (which for shortness we will refer to from now on as the “Cholesky method”). In this work we compare the efficiency of this original method against three alternatives, two based on the Chebyshev-expansion series method developed by Fixman [58] (which will be altogether referred to as the “Chebyshev methods”), and the recently published “TEA”-method [61] (occasionally, also referred as the “Geyer-method”).

The two alternatives based on the Chebyshev-expansion series are the ones developed by Jendrejack et al. [60] (from now on referred as the “Jendrejack method”), based on an E_f parameter which measures the accuracy of the expansion series approximation, and the method by Kröger et al. [59] (the “Kröger method” from now on), which is based on an *a priori* estimation of an upper and lower bounds pair to the eigenvalues of the instantaneous diffusion tensors which are generated along the simulation.

The details of each of the algorithms can be found in the cited references, as they have been implemented in our simulation programs just as they were conceived, with the only exception of the Kröger method. Kröger et al. devised theoretical equations to estimate the values of the upper and lower bounds to the eigenvalues of the diffusion tensor along the simulation of Zimm models; this equations are not applicable to our models because of our use of an excluded volume (EV) potential. Instead, we estimate this bounds by calculating the eigenvalues of a pre-averaged diffusion tensor, and taking the half of the lower and the double of the higher eigenvalues.

The comparison of these methods is based on the simulation on flexible chains with Gaussian springs and Lennard-Jones (LJ) excluded volume potential with *theta*-solvent parameters. Those chains present a certain degree of overlap. Previous works are based on the simulation of chains which are not allowed to overlap at all [61]. However, as already concluded by Fixman [58], the overlap has a negative effect on the Chebyshev-method. Thus, the chosen model is a widely used one and, at the same time, a “worst case scenario” to put the different HI methods up to test. In addition, other simulations with stiffer chains were performed in order to test the influence of the overlap in the results.

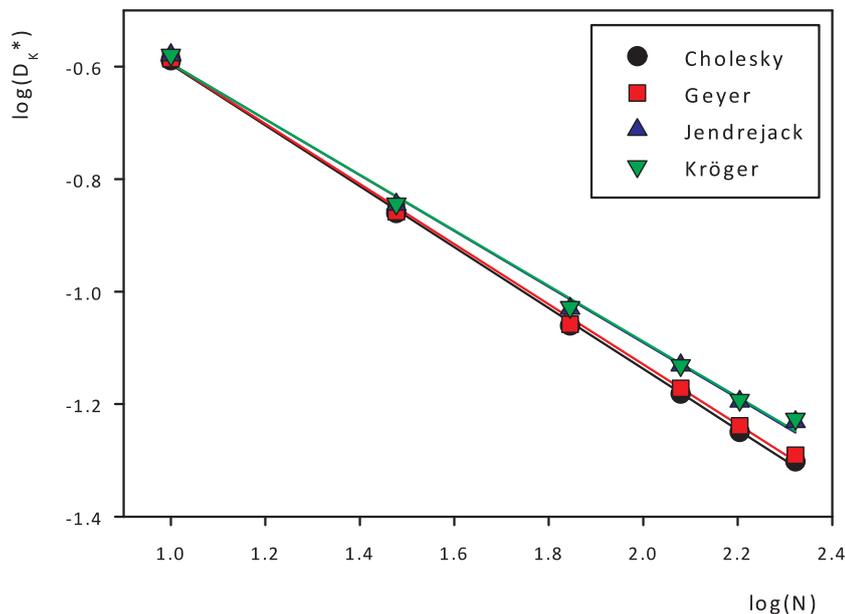


Figure S.1: Evolution of the calculated D_t^* over N for growing chain length.

Our results are summarized in figure S.1. In that plot we can see two different trends: the first one corresponds to data from the Cholesky and TEA methods, which are virtually identical, and the other one corresponds to the Chebyshev-based methods which yield practically coinciding results. From this second trend we can conclude that Kröger procedure that uses constant upper and lower bounds produces results as good as those obtained by the Jendrejack procedure consisting on recalculating eigenvalues along the trajectory and thus consuming time.

We suspected that the deviation in the Chebyshev methods was related to the overlapping, so we repeated some simulations by hardening the constants on the Gaussian connectors (using $H^* = 6$ instead of the $H^* = 3$). As we kept an equilibrium length of

$l_{eq} = 0$, this meant that the overlap in this second model would be somewhat higher than in the previous simulations. By doing this, we observed that the deviation between the results from the Chebyshev-based methods and the Cholesky-based method increased in an important way. To further confirm this hypothesis, we finally simulated stiff wormlike chains without overlapping. For that case, both the Cholesky method and the Chebyshev methods yielded almost the same results. Then, we concluded that the Chebyshev methods are quite sensitive to overlap, and therefore they are not suitable for simulations of models with an important amount of overlap.

N	Jendrejack	Kröger
	L_{mean} / L_{max}	L
10	14 / 75	5
30	18 / 84	7
70	24 / 114	8
120	27 / 112	9
160	29 / 115	10
210	32 / 126	11

Table S.4: Number of expansion terms in the Chebyshev expansion series in the Jendrejack and Kröger methods for the initial simulations.

We also studied the number of required terms in the Chebyshev expansion series (of the Jendrejack and the Kröger methods) to get a given accuracy. It is important to remember that the number of expansion terms L in the Kröger procedure is determined at the beginning of the simulation and kept constant during the calculation, while the Jendrejack method updates L each time the eigenvalue bounds are found not to be appropriated any more. The results for the model with Gaussian springs ($H^* = 3$) and LJ potential are shown in table S.4. As it can be seen, the mean number of terms in the Jendrejack simulations is generally close to three times the constant L values in the equivalent Kröger simulation.

When looking at the efficiency of these methods, we can also find two very different trends. The computing times of a simulation (t_{CPU}) of one million simulation steps for different model and methods are depicted in figure S.2. As we can see, both the Cholesky and Jendrejack methods tend to produce much higher simulation times than the Kröger or TEA methods.

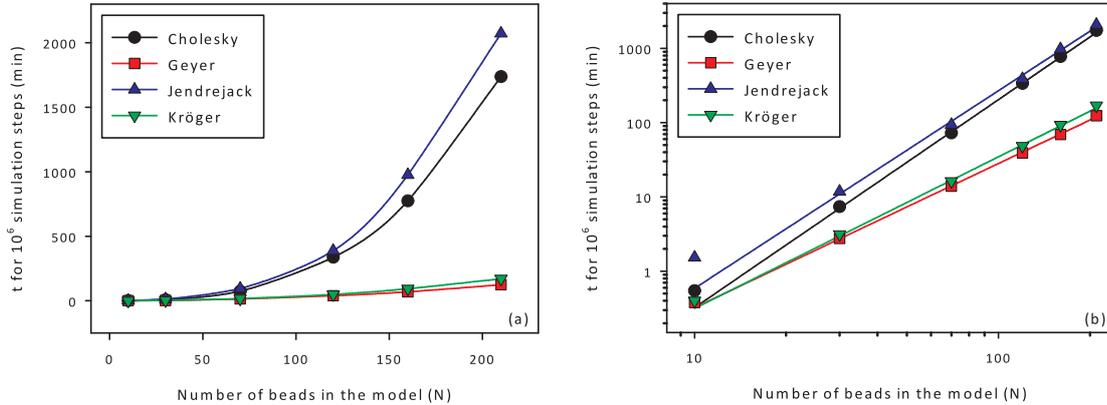


Figure S.2: Calculation times for a trajectory of 10^6 simulation steps using the Gaussian-LJ model, in natural scale (a) and logarithmic scale (b).

Method	$\log(k)$	α	α_T
Cholesky	-3.29	2.80	3
Geyer	-2.43	1.94	2
Jendrejack	-2.89	2.66	9/4
Kröger	-2.54	2.04	9/4

Table S.5: Scaling laws for the duration of a 10^6 step simulation with the different studied methods, $t_{CPU} = kN^\alpha$. α_T is the theoretically predicted-scaling law exponent

The parameters of the power laws relating t_{CPU} with the number of model elements, N , are shown in table S.5. It can be seen that the scaling law exponents for the Cholesky and Kröger methods are lower than those theoretically predicted (3 and 2.25 respectively), whereas the exponent obtained for the Jendrejack method is noticeably higher. Finally, TEA-method matches the theoretical prediction, scaling as N^2 . Theoretical exponents should be obtained in the case of very large chains, while our tests involve chains with relatively small N .

The case of the large exponent found for the Jendrejack method can be explained if we bear in mind that the duration of the simulation depends on the number of terms in the Chebyshev expansion series. Since the mean L values for the Jendrejack method are around three times the value of L in the Kröger method, it is expected that this simulations last much longer. Despite of this unexpectedly high scaling exponent, the

Jendrejack method still scales with time slower than the Cholesky one. Therefore, there should be a point from which on the Jendrejack method should be faster than the Cholesky method. We find that for the model used in this work, this point is around $N = 720$, which is a high number for many common simulations.

On the other hand, both the Kröger and TEA methods have scaling exponents close to two, and are both very fast simulation methods. There is only one consideration about this, which is that the Kröger method requires an estimation of the bounds of the eigenvalues of the diffusion tensor along the simulation. For the estimation of these values for a Zimm model, Kröger's expressions may be used, but for any other kind of model, a different approach must be considered. One valid option would be the determination of the eigenvalues from a pre-averaged tensor, but this requires some previous calculation and therefore time, which is a disadvantage for this method.

Summing up all previous observations, the best option for simulations would be the TEA-method, but it is important to keep in mind that it is an approximation, and may need a closer look for other kind of simulation conditions or models. Concerning this aspect, the most reliable approach is still the Cholesky method since it is applicable to any model.

3. **Multiscale simulation of hyperbranched polymers.**

In a previous work [24], our Group developed a a multi-scale approach for the simulation of dendrimer molecules. In this work, we try to extend this method to the simulation of hyperbranched polymers. This kind of polymers are similar to dendrimers in that they also present a high number of end-groups, but due to their much simpler synthesis, their branching is irregular and therefore they do not have the highly ordered layer structure of the dendrimers. This characteristic makes them a good, cheaper substitute for the dendrimers in some applications, and therefore have been studied for some time [70–77].

The study of hyperbranched polymers is not easy due to the fact that during their synthesis (often a *one-pot* reaction), a great number of different topologies are generated, even after the separation of the reaction products into fractions of narrow molecular weight intervals.

This variety of structures also implies a great variability in the properties, which is not always reproducible. To this respect, for example, there is a certain controversy about the behaviour of the intrinsic viscosity of these compounds: some authors state that it behaves like the intrinsic viscosity of dendrimers, having a maximum value for a certain molecular weight, above which the value of this property decreases [76], while other authors find that $[\eta]$ is monotonically rising with the molecular weight [73, 74].

Issues as the previous one may be studied from the point of view of computational simulations. To this purpose we want to extend the previously mentioned multi-scale simulation method for the simulation of dendrimers, to the simulation of hyperbranched polymers. This method consists in the generation of a coarse-grained bead-and-spring model, defined by hard-FENE springs, angular springs and hard sphere excluded volume potentials. The parameterization of these elements is made from atomic detailed simulations of small portions of the polymer that is going to be simulated. In this way, we avoid the use of fitting parameters and the need of previous experimental data. Both the parameterization procedure as well as the details of the constituents of the model may be reviewed in our previous work [24].

The biggest difference between the simulation of the dendrimers and the present work is due to the need of accounting for the conformational variability of the hyperbranched

polymers: while a dendrimer of a given molecular weight is well defined by a single structure, a hyperbranched polymer of a certain molecular mass may have a great number of different topologies. To take this into account, for each of the simulated molecular weights we simulate a wide spectrum of randomly generated topologies matching this molecular weight (precisely, 640 topologies will be considered for each molecular weight). These topologies are generated to match a *degree of branching* (DB) [85] topology index of $DB \approx 0.55$, which is an average value in between zero (for a completely linear polymer) and one (for a completely branched polymer). This value is chosen because we use experimental references where the topology of the samples is not well defined so that we hope that our results are not far from the references by using an average degree of branching.

The polymers that were simulated were polybenzileter (PBzE), polyamidoamine (PAMAM), polymethylallylcarbosilane (PCS3), polymethylundecenylcarbosilane (PCS11) y polydimethyl-5-(4-hydroxybutoxy)isophthalate (PDHBI). For PBzE there are no experimental data but we already had a good parameterization from our previous work with dendrimers; for PAMAM we use the experimental data by Hobson et al. [76] (despite the core there is hexafunctional). For PCS3 and PCS11 we have experimental data from Tarabukina et al. [77, 87], whereas for PDHBI we compare our results to the experimental ones coming from De Luca et al. [70, 71].

We have to admit that our results are not good: the data from the simulations are quite different from the experimental ones.

For the case of PAMAM, it can be seen in figure S.3 that the simulation results for the intrinsic viscosity are quite higher than the experimental ones. In both cases there is a $[\eta]$ maximum, but it appears at different molecular weights in our simulations and in the experimental results.

Both in that case as well as for PCS3 and PCS11, we modified the parameters values of an $N = 100$ chain to check how the changes influenced the results. As a conclusion, no change in the parameters achieved to yield an $[\eta]$ value in the range of the experimental data. For the PCS3 we can see similar results in figures S.4 and S.5. In this case, the simulation values are well above the experimental results. However, the effect of a 20% reduction in the hydrodynamic bead radius along with a 50% reduction in the hard-sphere radius (or even its omission), produce results which are in the same range as the experimental data.

For PCS11 (figures S.6 and S.7), we get the best results among all of the simulations: the intrinsic viscosity for the smaller chains correctly matches the experimental values,

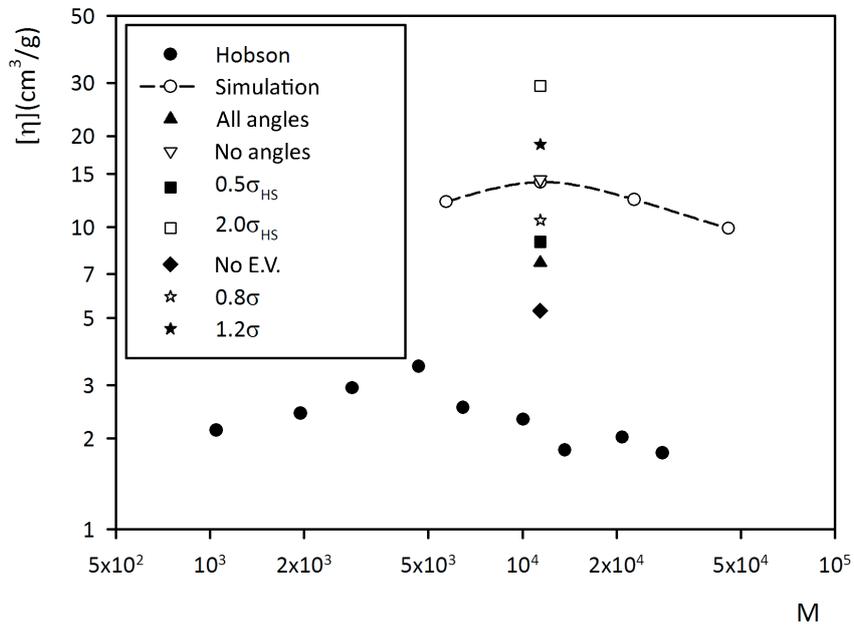


Figure S.3: PAMAM intrinsic viscosity data versus molecular weight both from simulations and from experimental data.

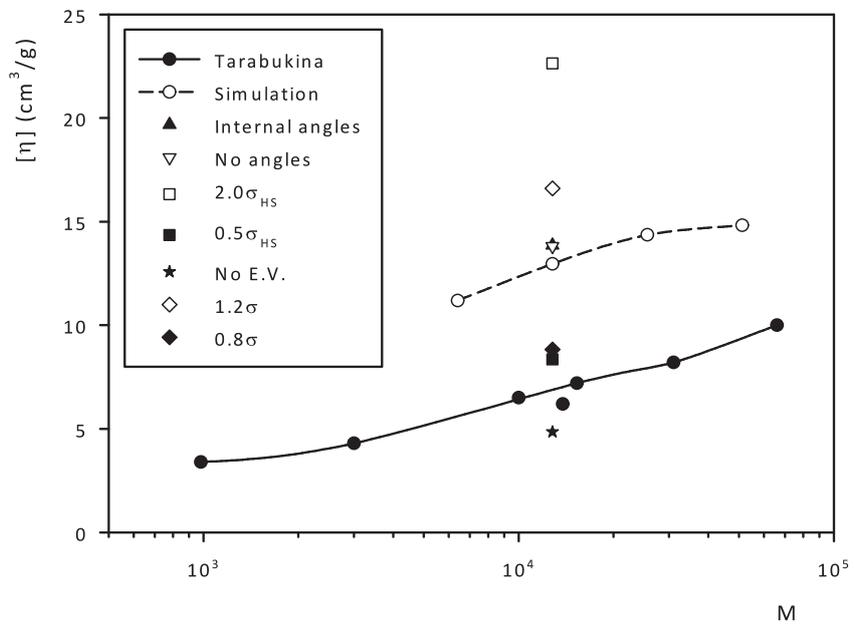


Figure S.4: PCS3 intrinsic viscosity data versus molecular weight both from simulations and from experimental data.

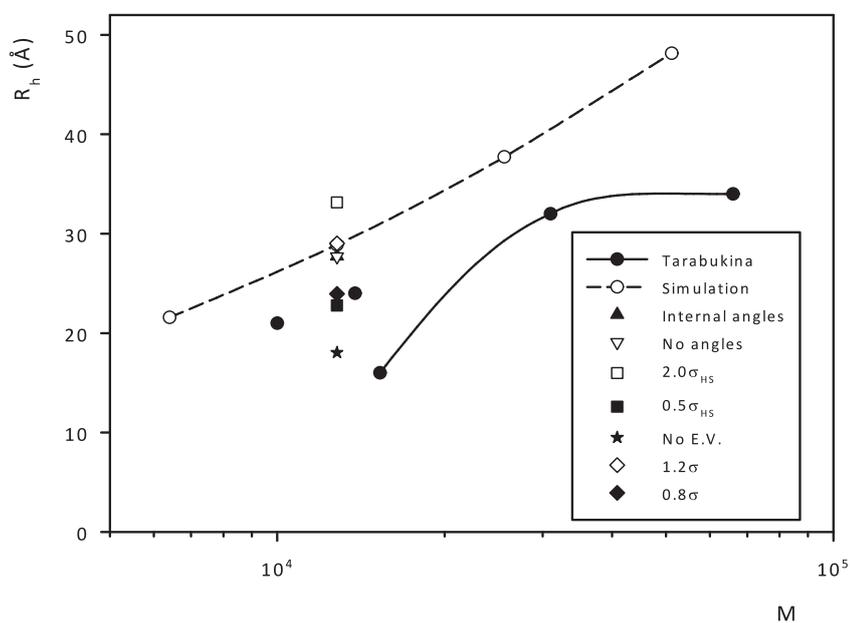


Figure S.5: PCS3 hydrodynamic radius versus molecular weight both from simulations and from experimental data.

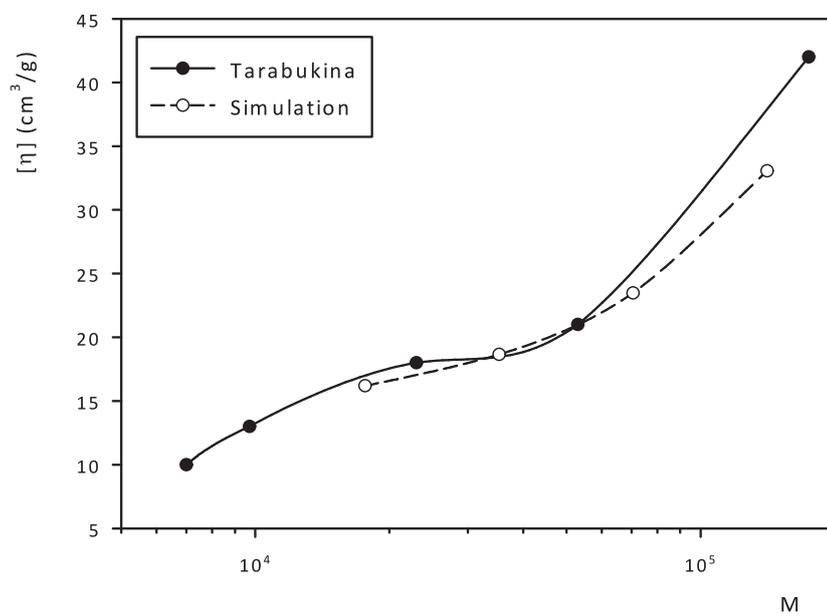


Figure S.6: PCS11 intrinsic viscosity versus molecular weight both from simulations and from experimental data.

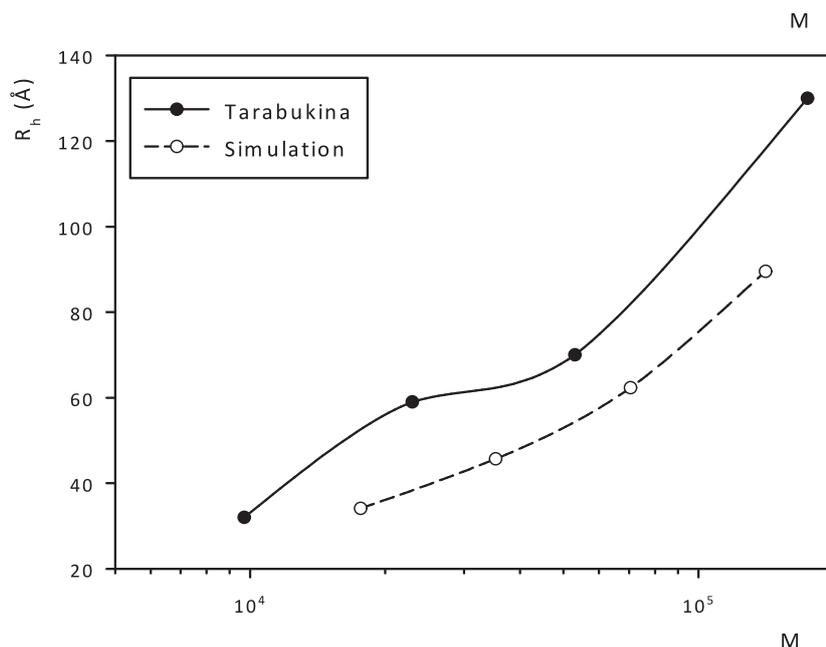


Figure S.7: PCS11 hydrodynamic radius versus molecular weight both from simulations and from experimental data.

but the hydrodynamic radii are still quite quite far from the experimental data. It can also be noted that for this polymer the tendency is opposite as in the two previous ones. Thus, the simulation data are below the experimental ones.

Finally, for the PDHBI (figures S.8, S.9 and S.10) we can see that, again, the simulation values are smaller than the experimental data.

The only conclusion that we can extract from the results in this chapter is that the procedure which gave good results for the simulation of dendritic molecules is not directly applicable to the simulation of hyperbranched polymers. One possible reason for this failure might be a lack of statistical representativeness of our samples (640 different topologies may not be enough to represent the huge number of topologies that are randomly created in a *one-pot* reaction). Also it is possible that our guess in generating topologies with $DB \approx 0.55$ for every case is not representative of the random topologies that are generated during such a *one-pot* synthesis.

These hypothesis about the topology being responsible for the bad results of this work is also supported by the good results in the last Chapter of this Thesis, where the same multi-scale simulation methodology is used, but in this case, as in dendrimers,

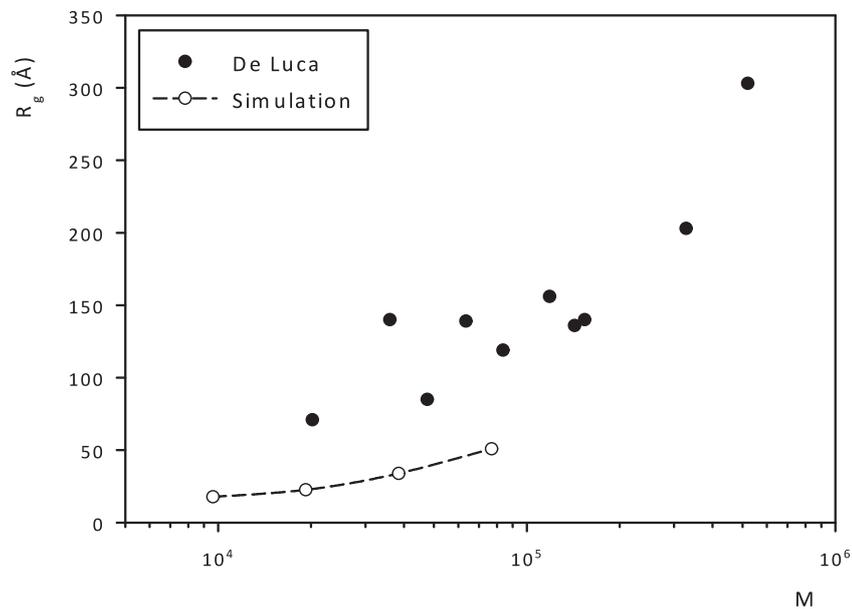


Figure S.8: PDHBI gyration radius versus molecular weight both from simulations and from experimental data.

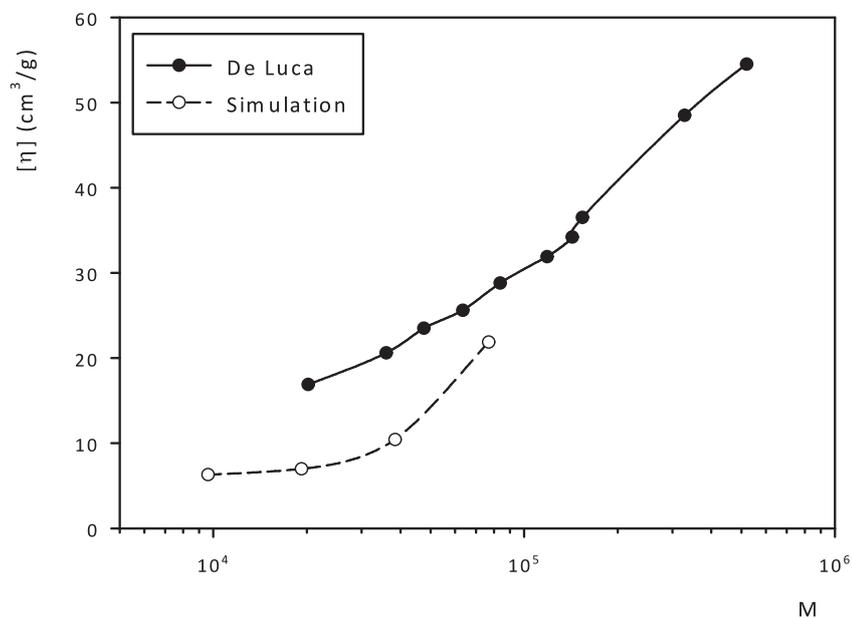


Figure S.9: PDHBI intrinsic viscosity versus molecular weight both from simulations and from experimental data.

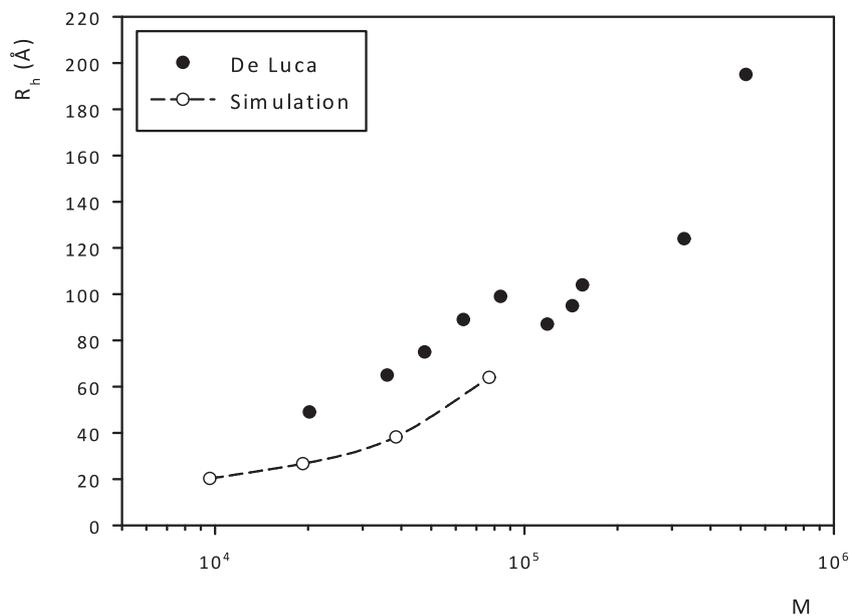


Figure S.10: PDHBI hydrodynamic radius versus molecular weight both from simulations and from experimental data.

the topology is well defined (the models are linear chains), and no topology averaging is needed.

4. Simulation of thermoresponsive amphiphilic copolymers.

During the last decades, there has been some research interest focused on thermoresponsive amphiphilic block copolymers [88–92] due to their tendency to self-assemble into *core-shell* structures when changing the temperature. This capability has some interesting applications in the biomedical field [92, 93].

That kind of systems typically consists of a thermoresponsive block like poly(N-isopropylacrylamide) or PNIPAAm in our case, and a second block which is less or not at all temperature responsive like methoxypoly(ethylene glycol) or MPEG in the case of this work. PNIPAAm is described to exhibit a LCST (lower critical solution temperature) of 32 °C, a temperature below which its hydrophobicity increases and collapses into the hydrophobic core. On the other hand, the MPEG block acts by enveloping and shielding the PNIPAAm core from contact with water. Besides, it is also known that PNIPAAm composites like these aggregate, even below the LCST [109]. This aggregation phenomenon makes difficult to study single molecule behaviour in laboratory experiments, so that simulations pose as a good alternative to study this behaviour.

In this work we have two objectives: 1) to study the influence of the ratio between the lengths of both blocks on the thermoresponsive behaviour of PNIPAAm-MPEG diblock copolymers, and 2) to study the influence of the MPEG connector length in the behaviour of PNIPAAm-MPEG-PNIPAAm triblock copolymers.

These studies were carried out in collaboration with the group of Professor Nyström from the University of Oslo. Our work consisted of modeling and simulating those diblock and triblock copolymers, while the group at Oslo synthesized and characterized experimentally those copolymers. The experimental results were compared with the ones coming from our simulations.

We employed bead-and-spring models to represent three cases: 1) a series of diblock copolymers with constant length for the PNIPAAm block and different lengths for the MPEG blocks; 2) a series of diblock copolymers with constant length for MPEG block and different lengths for the PNIPAAm blocks; 3) a series of triblocks copolymers with two identical PNIPAAm blocks at their ends and MPEG spacers with different length in the middle. The properties calculated from the simulations were the radius of gyration

(R_g), the hydrodynamic radius (R_h), and end-to-end distances of each of the individual block in each copolymer.

	Parameter	PNIPAAM	MPEG	Mixed
Bead	Bead size, σ (\AA)	4.60	3.18	-
Springs	Spring Constant, H ($k_B T / \text{\AA}^2$)	0.2	0.7	1.2
	Equilibrium length, l_e (\AA)	6.4	4.9	5.9
	Maximum length, l_{\max} (\AA)	10.5	6.7	13.5
E.V.	Potential well depth, ϵ_{LJ} (erg)	See table S.7		
	0-potential distance, σ_{LJ} (\AA)	5.12	3.92	4.48
	Cut-off radius, $r_{\text{cut-off}}$ (\AA)	10.24	7.84	9.96

Table S.6: Model parameters for PNIPAAM, MPEG and mixed interaction (PNIPAAM-MPEG).

The models and parameterization procedures were the same used for dendrimers [24] and hyperbranched polymers (see previous section). The difference is that in this case, we used no angular restrictions and a LJ excluded volume potential through which we are able to mimic the behaviour of PNIPAAM at different temperatures by tuning the ϵ_{LJ} parameter. In sum, we had to parameterize two kinds of beads, and three kinds of connector springs and EV potentials (those corresponding to MPEG-MPEG, PNIPAAM-PNIPAAM and MPEG-PNIPAAM interactions). Finally, we performed Monte Carlo simulations of those models. Table S.6 presents a summary of the parameters that were used to build the models.

Temp ($^{\circ}\text{C}$)	$\epsilon_{LJ,\text{PNIPAAM}} \times 10^{15}$ (erg)	$\epsilon_{LJ,\text{MPEG}} \times 10^{15}$ (erg)	$\epsilon_{LJ,\text{mixed}} \times 10^{15}$ (erg)
25	4.11	4.11	4.11
30	12.54	4.18	7.24
31.4	21.00	4.20	9.39
31.7	29.43	4.21	11.13
31.75	42.06	4.21	13.30
31.8	63.09	4.21	16.29
31.85	84.14	4.21	18.81
31.9	210.38	4.21	29.75

Table S.7: LJ potential well depths corresponding to the different simulation temperatures.

Table S.7 presents the different $\epsilon_{LJ,PNIPAAM}$ values that were used to mimic the behaviour at different temperatures. The values of that parameter were calculated by trial and error, simulating single PNIPAAM chains so that they could reproduce some experimental data [98] for a range of temperatures between 25 and 32 °C. We plot our results against the values employed for ϵ_{LJ} .

Block copolymers are identified by dash-separated numbers corresponding to the block lengths. Thus, diblock copolymers are identified by two numbers, the first one being the length (number of beads) of the PNIPAAM block, and the second one being the length of the MPEG block, whereas triblock copolymers are identified by three numbers (the one in the middle being the length of the MPEG block, and the two at the ends being the PNIPAAM block lengths). A bead of our model represents two monomers of the copolymers synthesized by the group of Professor Nyström.

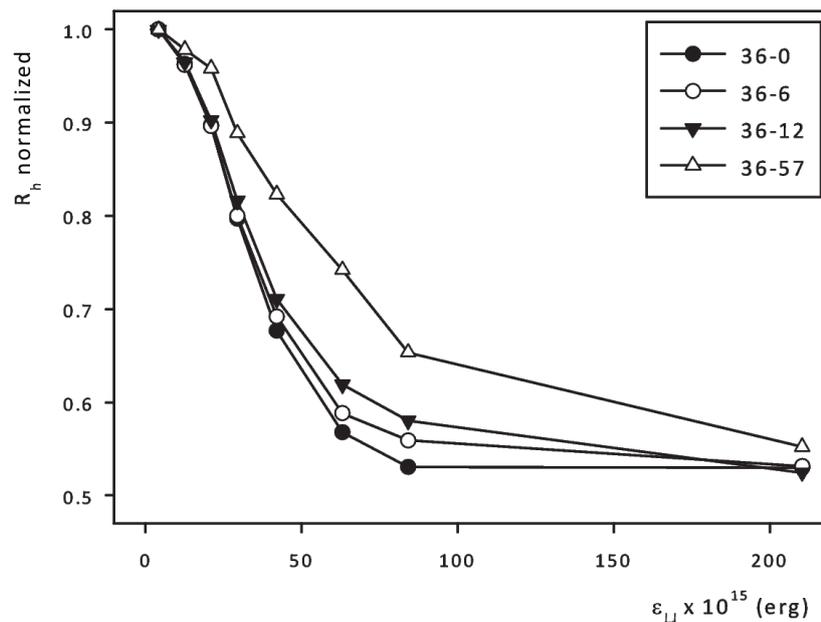


Figure S.11: Normalized hydrodynamic radii of the copolymers with constant PNIPAAM block size versus $\epsilon_{LJ,PNIPAAM}$.

Figures S.11 and S.12 show the dependence of the chain hydrodynamic radius with $\epsilon_{LJ,PNIPAAM}$. The hydrodynamic radius is normalized respect to its equilibrium value in good solvent conditions. From this figures we can figure out the contraction rates of each polymer. As it can be seen, the size of the PNIPAAM block controls the degree of contraction, with only a small influence of the MPEG block size.

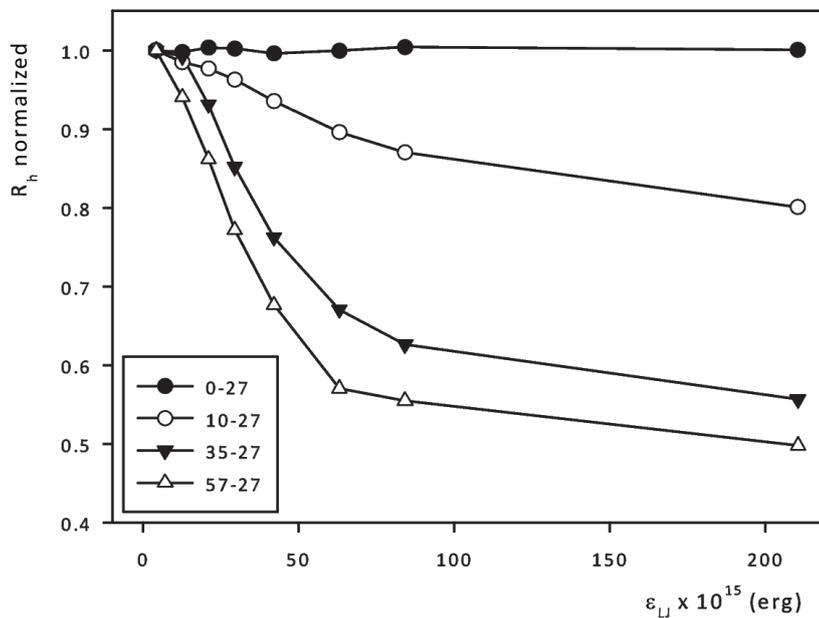


Figure S.12: Normalized hydrodynamic radii of the copolymers with constant MPEG block size versus $\epsilon_{LJ,PNIPAAm}$.

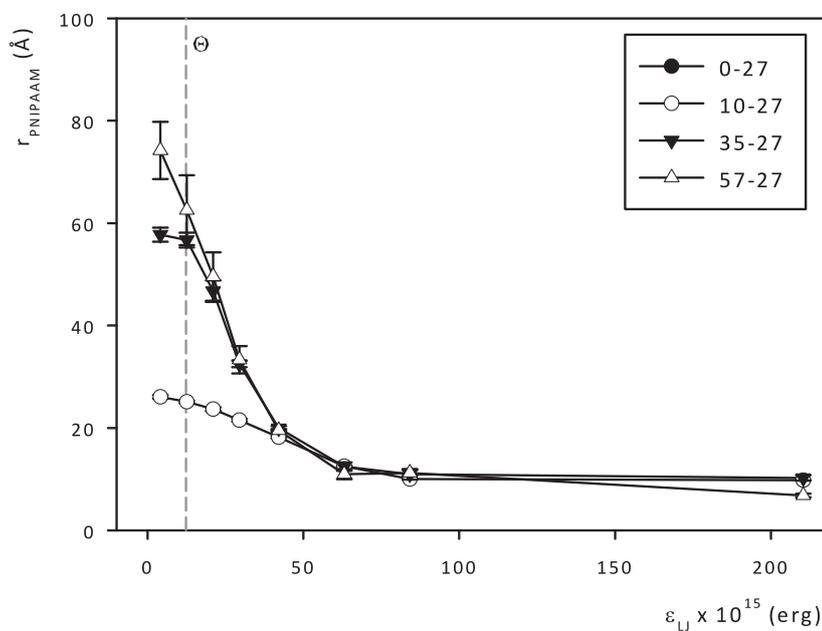


Figure S.13: End to end distance of the PNIPAAm block in the models with constant MPEG block size versus $\epsilon_{LJ,PNIPAAm}$.

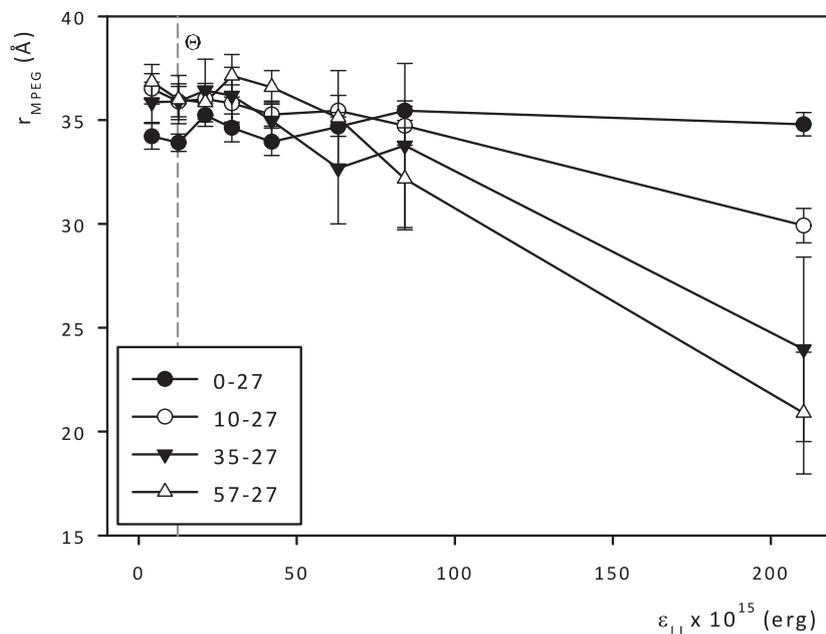


Figure S.14: End to end distance of the PNIPAAm block in the models with constant MPEG block size versus $\epsilon_{LJ,PNIPAAm}$.

Figures S.13 and S.14 show the dependence of end-to-end distances of the PNIPAAm and MPEG blocks on $\epsilon_{LJ,PNIPAAm}$. As it can be observed, the PNIPAAm blocks contract much more than the MPEG ones, so we can conclude that firstly the PNIPAAm collapses into a core that starts shrinking, and then the MPEG wraps around that core, shielding it from the solvent.

Finally, we can extract some images from the simulations to illustrate the conformational behaviour of the different copolymers depending on the temperature. Figures S.15 and S.16 show how the conformations of two types of PNIPAAm-MPEG copolymers evolve with temperature. Both figures show the process described before: the PNIPAAm collapsing (red beads) and the MPEG wrapping around the core (blue beads). The difference between both cases is that the MPEG block in 36-12 is too short, and does not provide a good shielding to the core, while the MPEG block in 36-57 is large enough to completely wrap around the core. In the 36-12 case, large parts of the core are exposed to the solvent and we think these parts try to avoid contact with the solvent by associating with hydrophobic regions of other molecules, giving rise to the presence of aggregates. As the degree of shielding of the hydrophobic core depends on the length of the MPEG block, there will be different aggregation tendencies depending on how long this MPEG

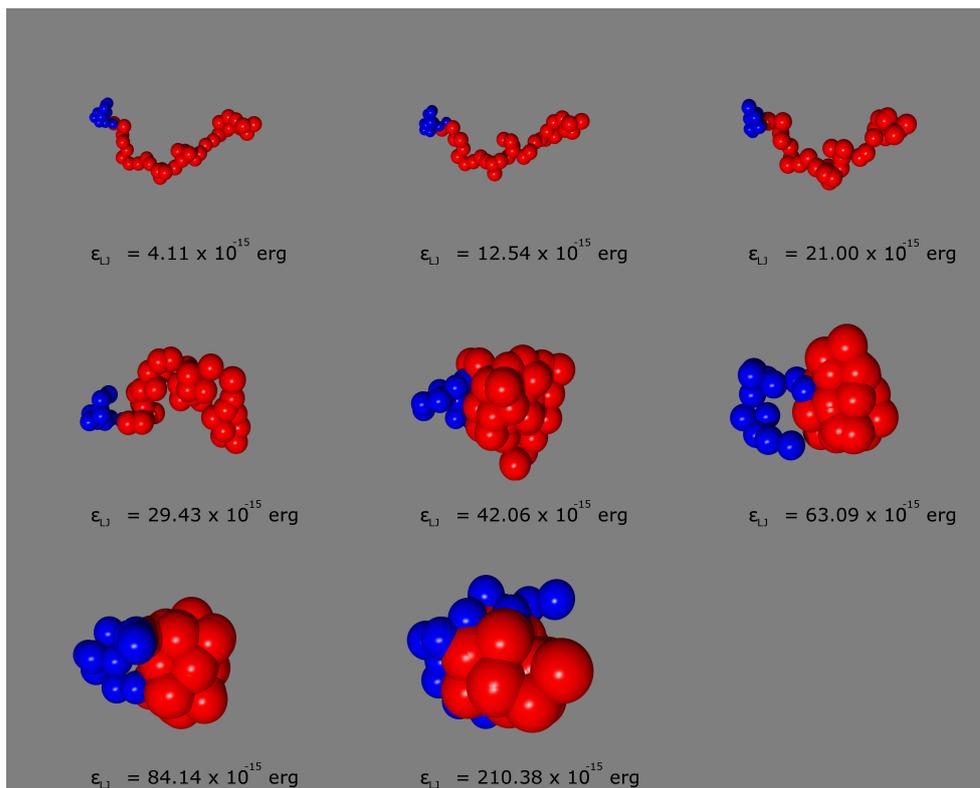


Figure S.15: Conformation image sequence for model 36-12 at different $\epsilon_{LJ,PNIPAAM}$. Red beads are PNIPAAM; blue ones are MPEG. Scaling may change from one picture to other.

block is, so that aggregates may form even below the LCST, as it was predicted in some previous work [103].

That work shows some experimental hydrodynamic radius data, which do not completely match our results (mean deviation in R_h is about 5 Å). The difference may be due to the difficulties to obtain reliable data because of the presence of large aggregates and polydispersity, effects that are not considered in the simulations.

The last quantity that was studied was the relation R_g/R_h , which gives a general idea of the shape of the molecule: values close to 0.77 are usually associated to uniform hard spheres, while values closer to 1.5 are normal for random coil conformations in good solvents [104–106]. Also, values of $R_g/R_h \approx 1.2$, which is close to those found for our copolymers in good solvent conditions, have been observed for rigid rods like nanofibers or wormlike structures [107, 108]. Therefore, we could think that at low temperatures our polymers act as these kinds of stiff structures. As we can see in figure S.17, when temperature increases, our polymers evolve into a more spherical conformation. Sadly,

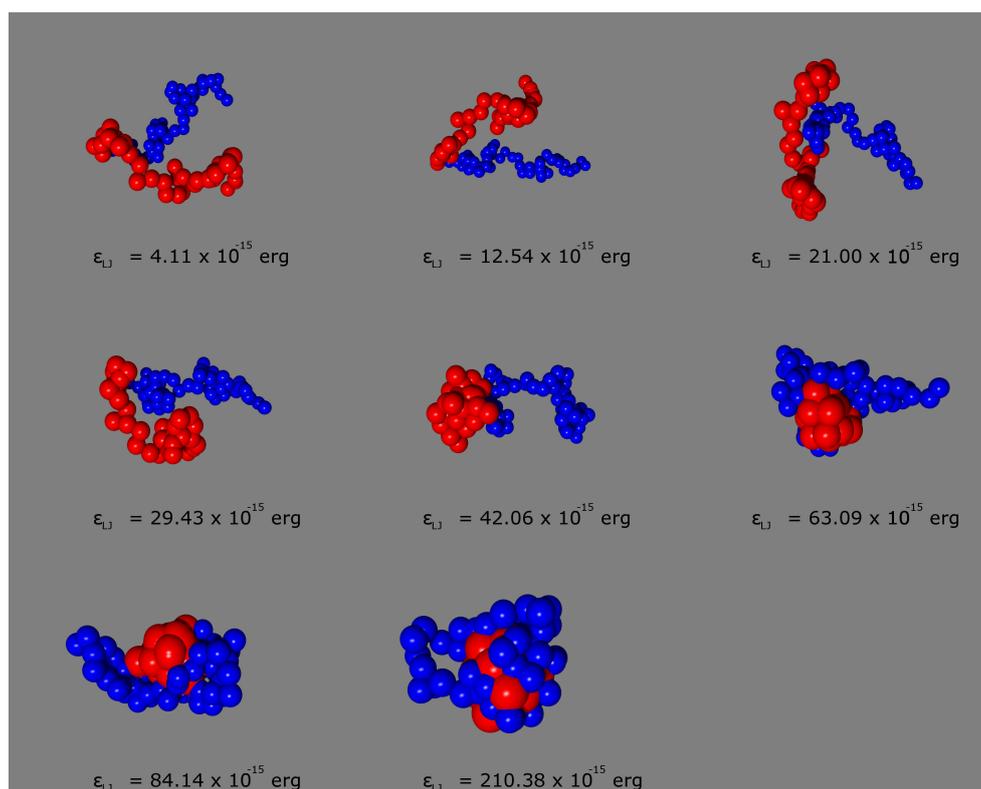


Figure S.16: Conformation image sequence for model 36-57 at different $\epsilon_{LJ,PNIPAAM}$. Red beads are PNIPAAM; blue ones are MPEG. Scaling may change from one picture to other.

this hypothesis could not be experimentally confirmed because it was not possible to get reliable R_g data.

For the triblock copolymers, we made similar simulations to the ones made for the diblock copolymers and obtained a similar behaviour.

Figure S.18 shows the normalized hydrodynamic radii for the triblock copolymers in a similar way as we presented the data for the diblock copolymers. The contraction that these triblock copolymers undergo are all of a similar order, except the one with the longer MPEG connector (case 32-100-32). This agrees with previous observation that the MPEG size has only a small effect on the contraction degree.

When we look at the ratio R_g/R_h , we can see that all but two models tend to evolve into spherical shapes at higher temperatures. The two exceptions are, again, the cases with the larger MPEG blocks: the case 32-100-32, which shows almost no change in its conformation, and the case 32-75-32, that shows a change in its shape but it does not reach such a low R_g/R_h value as the models with shorter MPEG spacers.

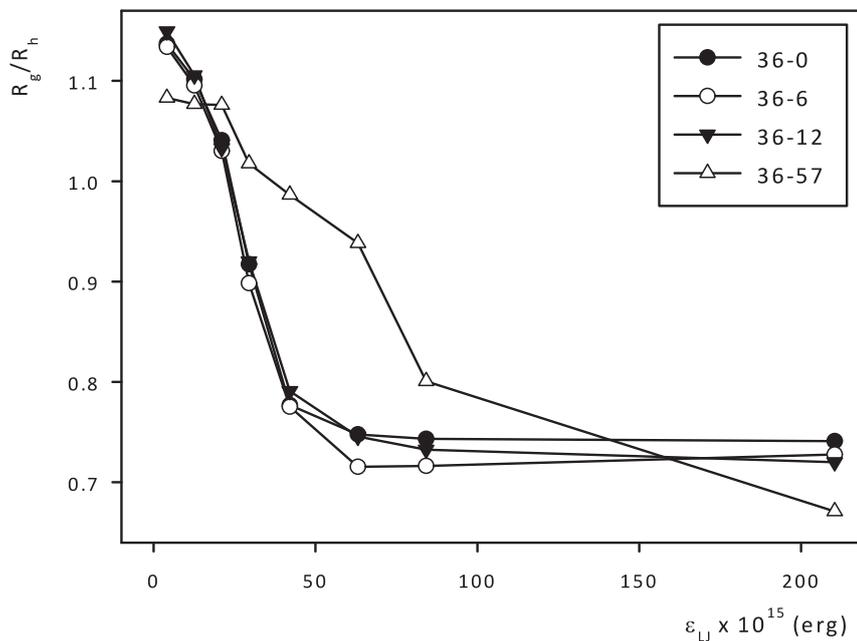


Figure S.17: R_g/R_h parameters of the chains with constant PNI-PAAM block size versus $\epsilon_{LJ,PNIPAAM}$.

If we study the end-to-end distances for each block of these copolymers along our temperature range, we find that the behaviour for the PNIPAAM block at each end of the chain although not identical is very similar, shrinking when raising the temperature. On the other hand, the behaviour of the MPEG spacer block is much more interesting. This data is presented in figure S.20. As it can be expected, at lower temperatures the length of this block increases with the number of beads in the block, but at temperatures over the LCST, when the PNIPAAM has already contracted, the sizes of the connectors of all the cases are again quite close. The contracted end-to-end distances of the five shorter connector blocks is similar, whereas the 32-100-32 copolymer presents a much smaller contraction. Also, for the 32-75-32 triblock copolymer, we can observe some difference since it undergoes an important contraction.

Finally, as we did before we can extract some image sequences from the simulations. A selection of these images is presented in figure S.21 where we can see three different behaviours (in agreement with the previous data):

a) For the copolymers with short MPEG spacer (like 35-12-35 and 32-17-32), both PNIPAAM blocks collapse into a single core. Because the MPEG connector is not long

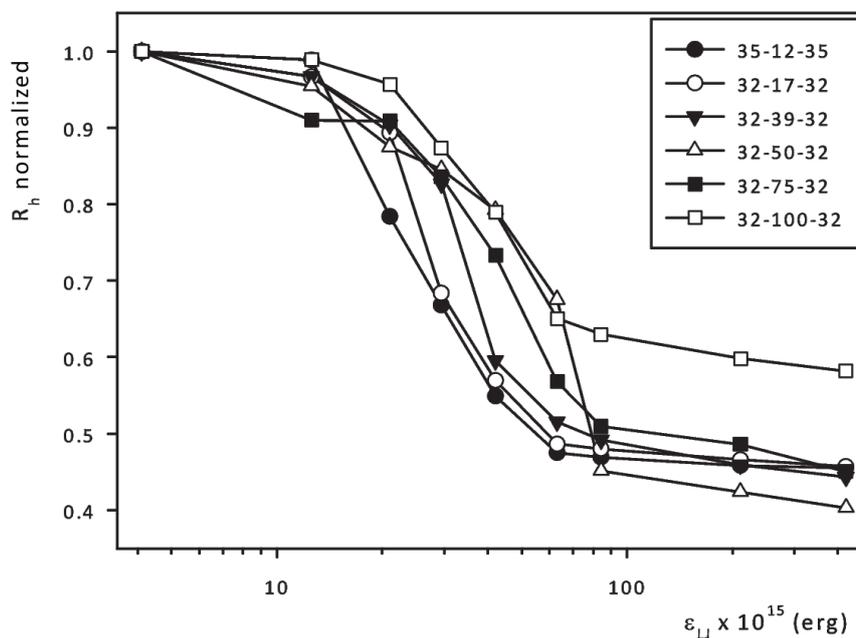


Figure S.18: Hydrodynamic radii of the triblock copolymers versus $\epsilon_{LJ,PNIPAAm}$.

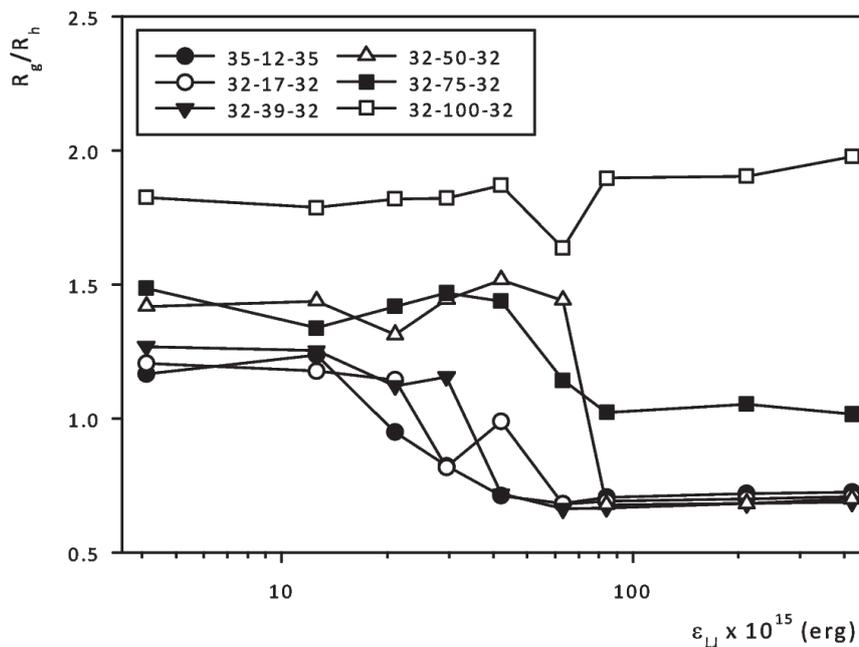


Figure S.19: R_g/R_h factor of triblock copolymers versus $\epsilon_{LJ,PNIPAAm}$.

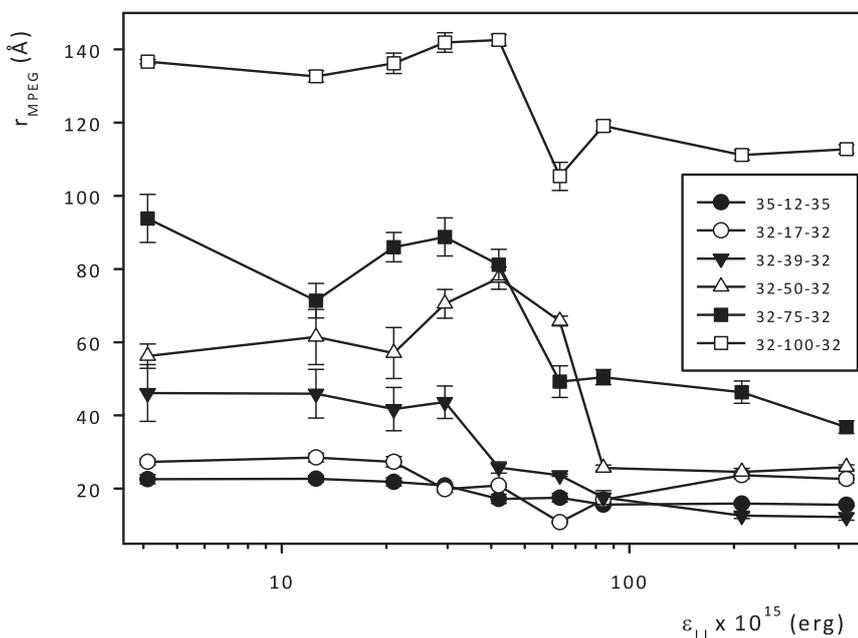


Figure S.20: End to end distance in the MPEG connectors of the triblock copolymers.

enough to provide a good shielding from the solvent, this polymers will have a high tendency to aggregate, even at low temperatures.

b) For the copolymers with medium MPEG block (like 35-39-35 and 32-50-32), both PNIPAAM blocks collapse also into a single core, but now the MPEG connector is large enough to protect the core from the solvent and the aggregation will not be so intense as in the previous case.

c) For the copolymers with large MPEG block, this wraps each PNIPAAM block and prevent them from collapsing into a single core so that two isolate PNIPAAM cores can be formed. Thus, a copolymer like 32-75-32 does not adopt a global spherical conformation, and a copolymer like 32-100-32 which has extremely long MPEG blocks, keeps a random coil conformation with the cores separated by a certain distance. In any case, PNIPAAM cores are well shielded and the copolymers will have a low tendency to aggregate.

The previous hypothesis about the formation of aggregates is supported by the experimental data provided by the group of Professor Nyström. Figure S.22 shows turbidimetry analysis made for the triblock copolymers that were synthesized by them. When raising the temperature, aggregation happens resulting in an increase of the turbidity of the so-

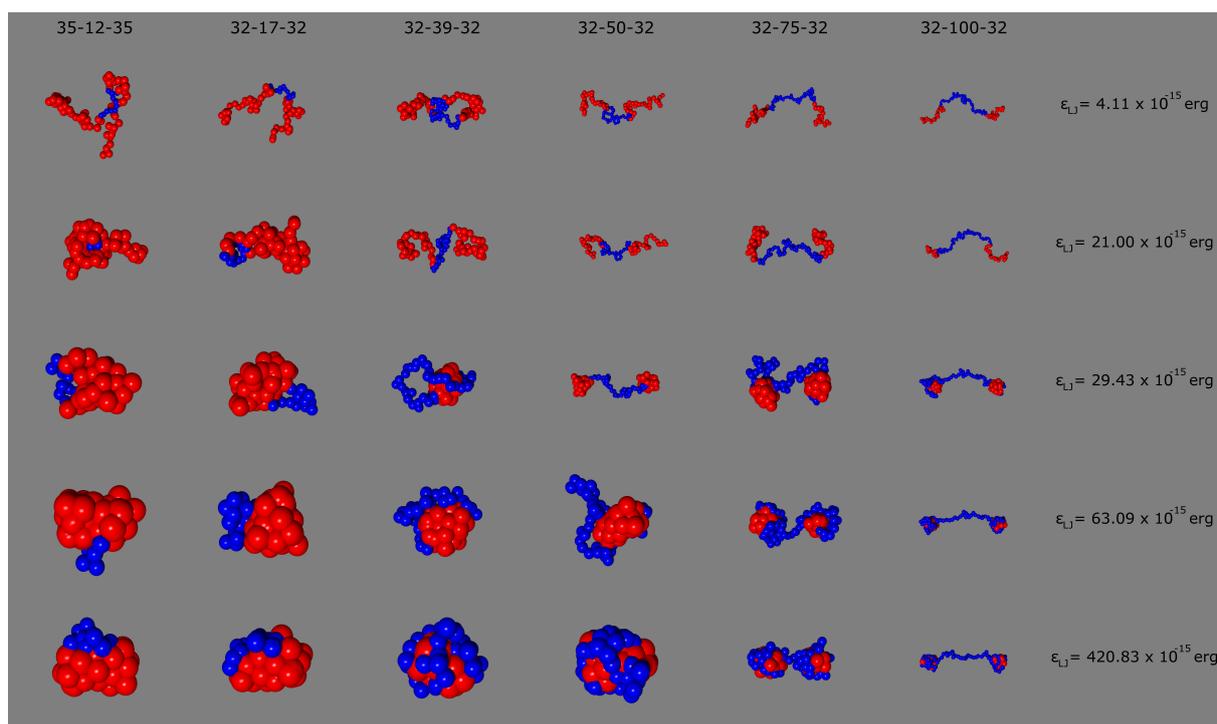


Figure S.21: Image sequences of the evolution of the different triblock chains when increasing $\epsilon_{LJ,PNIPAAM}$.

lution. As we can see in the figure, for the smallest triblock copolymers, an important increase in the turbidity is observed at a temperature below 40 °C. For some larger MPEG connectors (like $n = 77$ which corresponds to our model 32-39-32) turbidity increases slowly due to the better shielding of the single hydrophobic core. Finally, the copolymer with the longest MPEG spacer ($n = 165$ that could be well represented by our model 32-75-32), no increase in turbidity was observed in a temperature range up to 60 °C.

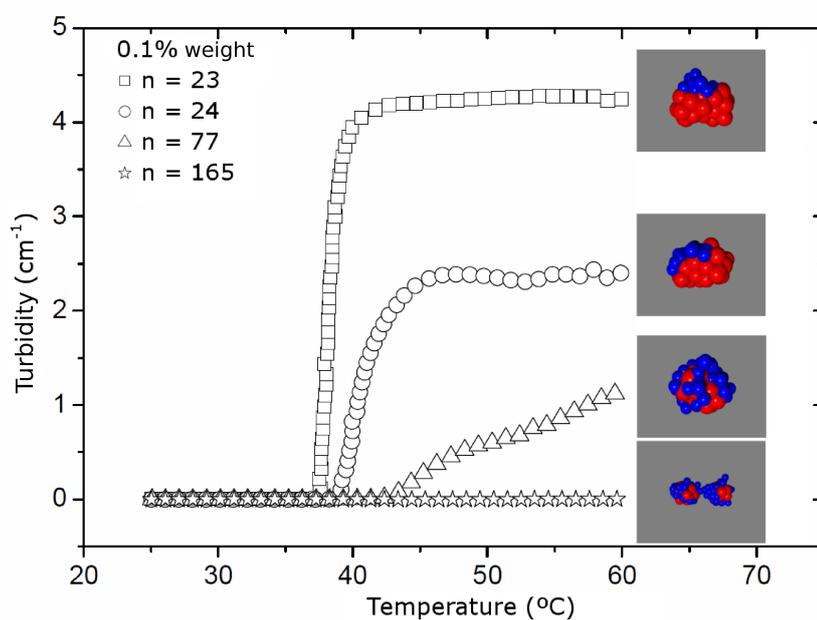


Figure S.22: Turbidimetry measurements provided by the group of Professor Nyström. The n value indicates the number of MPEG monomers in the spacer. As stated, two monomers are represented by one bead in the simulation models.

