



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

“MODELACIÓN ESTOCÁSTICA DE
EXPERIMENTOS DE ESTIRAMIENTO DE
MOLÉCULAS DE ARN”.

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

PRESENTA:

MARIO HERNÁN HIDALGO SORIA

DIRECTOR DE TESIS:

IVÁN SANTAMARÍA HOLEK



2014



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

II

A mi padre, madre, hermanos y amigos. Por su apoyo incondicional.

Datos del jurado:

1. Datos del alumno

Hidalgo

Soria

Mario Hernán

55 34 11 97 36

Universidad Nacional Autónoma de México

Facultad de Ciencias

Física

410018699

2. Datos del tutor Dr.

Iván

Santamaría

Holek

3. Datos del sinodal 1

Dr.

Rosalio Fernando

Rodríguez

Zepeda

4. Datos del sinodal 2

Dr.

Carlos Ignacio

Mendoza

Ruiz

5. Datos del sinodal 3

Dra.

Patricia

Goldstein

Menache

6. Datos del sinodal 4

Dr.

Gerardo

García

Naumis

7. Datos del trabajo escrito

Modelación estocástica de experimentos de estiramiento de moléculas de ARN

76p.

2014

Notación:

ARN	Ácido ribonucleico
LOT	Pinzas ópticas láser (<i>laser optical tweezers</i>)
z_{sys}	Extensión del sistema Asa-ARN-Asa
z_m	Extensión de la horquilla de ARN
Δz	Diferencia en la extensión del ARN con respecto al estado plegado y extendido
SOP	Polímero auto-organizado (<i>self organized polymer</i>)
GRM	Modelo generalizado de Rouse (<i>generalized Rouse model</i>)
NBA	cuenca de atracción para el estado plegado (<i>native basin of attraction</i>)
UBA	cuenca de atracción para el estado extendido (<i>unfolded basin of attraction</i>)
f	Fuerza externa en los experimentos usando LOT
F_{ef}	Energía libre efectiva
τ_{ef}	Tensión efectiva
L	Longitud de contorno de la asa
l_p	Longitud de persistencia de la asa
k	Constante de Boltzmann
T	Temperatura del baño térmico
γ	Coefficiente de fricción de Stokes
κ_b	Constante de restitución asociada a las LOT
z_b	Posición de la cabeza de poliestireno manipulada por LOT
P5GA	Tipo de horquilla para el modelo SOP
z_F	Valor de la extensión en el estado plegado
z_U	Valor de la extensión en el estado extendido
S	Función de estado para la entropía
ρ_j	Concentración del constituyente j
$\Delta_i S$	Cambio de entropía debida a los procesos irreversibles
$\tau_{NBA \rightarrow UBA}$	Tiempo medio de transición entre el estado plegado y extendido
$\tau_{UBA \rightarrow NBA}$	Tiempo medio de transición entre el estado extendido y plegado
$K_{NBA \rightarrow UBA}^+$	Tasa de transición entre el estado plegado y extendido
$K_{UBA \rightarrow NBA}^-$	Tasa de transición entre el estado extendido y plegado
$p_{st}(z)$	Densidad estacionaria de probabilidad para la extensión

Índice general

1. Antecedentes	1
1.1. Experimentos con moléculas individuales	1
1.2. Dinámica de moléculas individuales	3
1.3. Modelos y paisaje energético de moléculas de ARN	4
2. Herramientas del modelo	15
2.1. Termodinámica irreversible	15
2.2. Procesos estocásticos	20
2.2.1. Probabilidad y procesos estocásticos	20
2.2.2. Procesos de Markov Estacionarios y Ecuación Maestra	23
2.2.3. La expansión Kramers-Moyal y Ecuación de Fokker-Planck	24
2.2.4. Ecuación de Langevin	25
2.2.5. Derivación de la ecuación Fokker-Planck a partir de la ecuación de Langevin	27
2.2.6. Solución estacionaria para la ecuación Fokker-Planck	29
2.2.7. Ecuación de Fokker-Planck para ecuación de Langevin con ruido multiplica- tivo de color	30
2.2.8. Ecuaciones de Langevin y Fokker-Planck para régimen sobre-amortiguado	32
2.3. Teoría de Kramers	33
2.4. Simulación Estocástica	37
2.5. Apéndice: Deducción de la solución para la ecuación de Langevin	39
3. Desarrollo y Resultados	41
3.1. Formulación del modelo	41
3.1.1. Termodinámica irreversible del modelo	42
3.1.2. Ecuación de Langevin con ruido aditivo blanco	44
3.1.3. Ecuación de Fokker-Planck para (3.16) y su solución estacionaria	46
3.1.4. Tiempos medios y tasas de transición entre estados plegado y extendido . . .	47
3.2. Resultados de las simulaciones	48
3.3. Ecuación de Langevin con ruido de color aditivo	61
3.3.1. Resultados numéricos para el caso con ruido de color aditivo	63
4. Conclusiones	67

5. Apéndice: Códigos empleados

69

Resumen

Este trabajo presenta un estudio estocástico y termodinámico de la dinámica de plegamiento para moléculas de ácido ribonucleico (ARN) sujetas a la influencia de fuerzas externas. En concreto se introduce un modelo que describe de manera aceptable dicha dinámica de plegamiento. A *grosso modo*, nuestra aproximación al problema consiste en que a partir de los perfiles energéticos “experimentales” asociados al plegamiento de la molécula de ARN sujeta a una fuerza externa e inmersa en un baño térmico, el cual funciona como un ruido “externo”, se puede deducir una expresión analítica para la energía libre del sistema estudiado, y por ende se puede obtener la tensión efectiva que actúa sobre el sistema. Teniendo esto, se encuentran las ecuaciones de evolución de las variables que caracterizan el estado del sistema, en este caso la elongación de la molécula, las cuales consisten en ecuaciones diferenciales estocásticas, cuyas soluciones pueden ser simuladas por medio de un método de aproximación adecuado. Una vez obtenidas dichas simulaciones, es posible deducir las correspondientes distribuciones estacionarias y la producción de entropía asociada al proceso de plegamiento y extensión.

El capítulo 1 presenta los antecedentes experimentales y teóricos relacionados con el estudio de la dinámica de macromoléculas individuales, haciendo énfasis en aquellos para los cuales se han utilizado experimentos con pinzas ópticas láser. Se discute de forma breve la estructura de las moléculas de ARN y su correspondiente dinámica de plegamiento. Por último, en este capítulo se introduce sin detalle el modelo propuesto.

En el capítulo 2 se discuten las herramientas de análisis utilizadas en la formulación del modelo, en particular se exponen elementos de la teoría de procesos irreversibles, los conceptos básicos necesarios en el desarrollo de procesos estocásticos, sus ecuaciones de evolución, una discusión breve de la teoría de Kramers y también el esquema de aproximación utilizado para simular las ecuaciones de evolución mencionadas.

Posteriormente, en el capítulo 3, se aplican dichas herramientas a la dinámica de plegamiento de horquillas de ARN sujetas a una tensión externa y se muestran los resultados obtenidos a partir de las simulaciones realizadas para 11 configuraciones distintas del sistema estudiado. También se presenta una modificación al modelo propuesto, en la cual las fluctuaciones “internas” del sistema se modelan a través de un proceso de Ornstein-Uhlenbeck. Las simulaciones y los resultados obtenidos se discuten para tres casos representativos. Se presentan de igual forma las correspondientes simulaciones y sus correspondientes resultados para tres casos representativos. Finalmente, en el capítulo 4 se presentan las conclusiones obtenidas a partir del análisis realizado.

En el Apéndice de este trabajo se presentan los respectivos códigos empleados en las simulaciones presentadas.

CAPÍTULO 1

Antecedentes

Se describen los experimentos con moléculas individuales, comentando su importancia, diversidad y mencionando algunas técnicas utilizadas en el campo de la biofísica. Se discute el uso de procesos estocásticos en el modelaje de moléculas individuales. Se muestra y explica el sistema físico empleado para obtener los paisajes energéticos y las tasas de salto dependientes de una fuerza para horquillas de ARN de manera experimental. Se presentan los modelos principales para describir la dinámica de plegamiento de horquillas de ARN sujetas a una fuerza externa constante. Se expone un método alternativo para obtener la dinámica de las horquillas de ARN, vía la implementación de las ecuaciones de evolución para los procesos estocásticos inmiscuidos, deducidas usando un análisis de termodinámica irreversible.

1.1. Experimentos con moléculas individuales

El estudio de moléculas biológicas individuales como ADN¹ y ARN², es determinante en el desarrollo de la bio-tecnología y el entendimiento del funcionamiento biológico de los seres vivos. Dada la importancia de estos dos grandes temas para nuestra sociedad en general, diversos experimentos y sus respectivos estudios, han sido llevados a cabo en aras de lograr un mayor conocimiento acerca de estas dos importantes tareas.

En los últimos, años el desarrollo de nuevos métodos de manipulación de moléculas individuales, ha permitido realizar mediciones directas de las fuerzas generadas en las reacciones químicas asociadas a dichas moléculas. Así también, estos nuevos métodos han permitido modificar la naturaleza de las reacciones químicas en cuestión, al aplicar fuerzas externas sobre las moléculas [1].

Una amplia variedad de fenómenos han sido analizados vía dichos métodos de manipulación, desde el plegamiento³ de proteínas y ARN's, pasando por la separación de cadenas de ADN, hasta la dinámica de motores celulares. Nuestro trabajo se concentra en el primero de estos fenómenos para el caso de horquillas de ARN.

El ARN es una macromolécula formada por varias bases nucleótidas, en concreto de cuatro bases

¹Ácido desoxirribonucleico.

²Ácido ribonucleico.

³El término plegamiento indica el proceso por el cual una estructura proteínica asume su forma o conformación funcional.

nitrogenadas: adenina, citosina, guanina y uracilo, una ribosa y un grupo fosfato[2]. El ARN es una hebra, la cual presenta a lo largo pequeñas regiones locales de secuencias complementarias, las cuales pueden ser unidas por enlaces de hidrógeno. Este apareamiento entre dichas regiones pliega la hebra de ARN generando sobre ésta una estructura de horquilla (ver Figura 1.1).

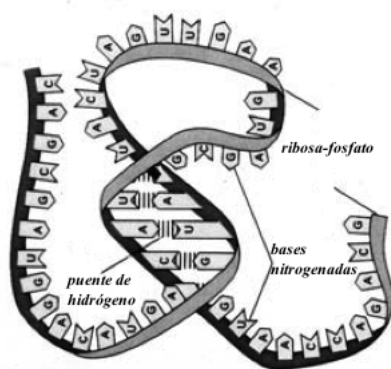


Figura 1.1: Estructura del ARN.

El hecho de que el ARN esté plegado significa que éste posee una estructura funcional, la cual le permite realizar diversas tareas dentro de la célula, como: catalizar reacciones biológicas, recibir o enviar respuestas a señales; además de controlar el proceso por el cual la información de un gen⁴ del ARN, es usada en la síntesis de un producto génico⁵ funcional (expresión génica). Dado esto, el estudio de la dinámica de plegamiento del ARN ante estímulos externos puede permitir visualizar para que configuraciones, éste es más propenso o no a realizar las mencionadas tareas dentro de la célula.

Siguiendo la línea de los experimentos con moléculas individuales, en este caso la aplicación de una fuerza externa modifica el proceso de plegamiento o no plegamiento, para la estructura proteínica estudiada (cadenas de ADN, horquillas de ARN, etc.).

Los experimentos consisten en aislar a la molécula de interés, preparar un arreglo experimental en el cual sea posible afectar al sistema por medio de una fuerza externa, la cual puede ser controlada a discreción. Una vez teniendo esto se determina la coordenada de reacción pertinente por medio de la cual se realizan las respectivas mediciones y se recogen los datos del experimento.

⁴Secuencia de nucleótidos en el ADN o ARN que contiene la información necesaria para la síntesis de una macromolécula con función celular específica.

⁵Material bioquímico resultante de la expresión de un gen, dicho material puede ser un ARN o una proteína.

Diversas técnicas han sido usadas en el desarrollo de los experimentos de moléculas individuales [3], como lo son el uso de confinamiento por medio de fuerzas magnéticas, microscopía por medio de fuerza atómica, fluorescencia y el uso de pinzas ópticas láser.⁶

En concreto, para el estudio del plegamiento de estructuras proteínicas las técnicas basadas en LOT han hecho posible el monitoreo de las transiciones entre los estados plegado y extendido. Aquí se obtienen histogramas y sus respectivas funciones de densidad de probabilidad vía las mediciones realizadas en la coordenada de reacción del sistema (que usualmente es la extensión del bio-polímero estudiado). Con lo anterior ya es posible determinar los estados más probables para el polímero. La obtención de las funciones de densidad de probabilidad permite, a través de los métodos de la física estadística tradicional, la obtención de las variables termodinámicas pertinentes: energía libre, entropía, etc.

En general, el análisis descrito arriba es el más usado en los experimentos LOT asociados al plegamiento de estructuras proteínicas. En el presente trabajo se presenta un método novedoso para el análisis de la información obtenida a partir de dichos experimentos. Dicho método toma en cuenta los efectos disipativos del proceso de extensión-plegamiento. En síntesis, los métodos usados en nuestro trabajo son los procesos estocásticos y sus respectivas ecuaciones de evolución para la coordenada de reacción (ecuación de Langevin) y para la densidad de dicha coordenada (ecuación de Fokker-Planck), así como la termodinámica irreversible, la cual caracteriza a todo proceso disipativo, aquí la entropía producida del proceso será la principal característica a tomar en cuenta.

Dichos procesos y ecuaciones están asociados al paisaje energético de plegamiento del bio-polímero. Nuestro método nos permite introducir la naturaleza estocástica del plegamiento de bio-polímeros en nuestras herramientas de análisis. Al hacerlo no sólo se procede a hacer una descripción del fenómeno estudiado sino que también por medio de las técnicas de simulación estocástica se pueden hacer inferencias del mismo.

1.2. Dinámica de moléculas individuales

Si desde un punto de vista general, en la naturaleza tienen lugar fenómenos en los cuales alguna cantidad de interés varía en el tiempo de manera irregular ó azarosa, y si, además, recordamos que la definición más simple de un proceso estocástico es que éste es una sucesión “temporal” de variables aleatorias, que en términos físicos significa que dicho proceso estocástico se considera como un *ensemble*⁷ de variables aleatorias [4]. Podemos decir que las características de estas definiciones y las de los fenómenos mencionados coinciden, por lo que el uso de procesos estocásticos se hace pertinente en el modelaje de los fenómenos descritos.

El ejemplo clásico de un fenómeno con magnitudes aleatorias modelado por procesos estocásticos es el movimiento Browniano, que describe la dinámica de una partícula en un medio fluctuante. Dicho modelo de movimiento Browniano se ha aplicado de igual manera, al estudio de polímeros inmersos en soluciones, entre muchos otros ejemplos. Debido a que éstos presentan incesantes cambios aleatorios en su forma y posición. Dichos cambios se deben a la agitación térmica asociada al medio en el cual están inmersos [5].

⁶Laser Optical Tweezers (LOT) en inglés.

⁷Por *ensemble* nos referimos a una colección ideal de un gran número de copias virtuales del sistema, cada una representando un estado real en el cual el sistema puede estar.

Desde una perspectiva fenomenológica, el movimiento Browniano de las moléculas individuales puede modelarse matemáticamente como un proceso estocástico. Por medio de esta suposición es posible construir una ecuación que describa a dicho movimiento y que esté basada en leyes macroscópicas conocidas teniendo en consideración las escalas temporales y dimensionales asociadas al fenómeno.

Aquí las ecuaciones de movimiento tienen que tomar en cuenta el efecto de las fluctuaciones sobre las moléculas individuales. Es común idealizar a dichas fluctuaciones como un ruido, según la clasificación de Hänggi y Jung [6], cuando un subsistema está inmerso en un medio ambiente fluctuante, se habla de que el ruido es “externo”, este tipo de ruidos es asociado a un ruido blanco, puesto que en éste las fluctuaciones no están correlacionadas para tiempos distintos. En el caso de que las fluctuaciones representen efectos acumulados de acoplamientos del medio con el subsistema, el ruido se considera “interno” y se asocia a un ruido de color, debido a que en éste las fluctuaciones tienen cierta correlación⁸.

Como se mencionó, el proceso de plegamiento de bio-polímeros como las cadenas de ADN, horquillas de ARN y proteínas en general, está sujeto a fluctuaciones provenientes del medio ambiente en que están inmersos, por lo que el uso de un proceso estocástico, el cual describa a la variable de reacción adecuada asociada al bio-polímero, es inmediato.

El uso de procesos estocásticos y sus respectivas ecuaciones de evolución nos permite introducir leyes probabilísticas (en lugar de deterministas) para describir el comportamiento del bio-polímero al ser estimulado por una fuerza externa. Esto permite asignar probabilidades a los distintos estados de configuración del mismo, y observar la dinámica de transición entre dichas configuraciones. Así, obtenemos información que puede ser usada para explorar las propiedades físicas y químicas de la molécula.

A continuación se muestra un ejemplo concreto de como se han aplicado los procesos estocásticos en la modelación de bio-polímeros, en concreto se muestra el caso de las transiciones inducidas por medio del uso de una fuerza externa entre los estados “plegado” y “extendido” para una horquilla de ARN.

1.3. Modelos y paisaje energético de moléculas de ARN

Las moléculas de ARN son bio-polímeros importantes en el funcionamiento de las células. Entre algunas de sus características más importantes se encuentra el almacenar y transferir información, y funcionar como catalizadores de reacciones bioquímicas. Es conocido que diversos procesos biológicos, como la división celular y la síntesis de proteínas, dependen de la capacidad de la célula para plegar moléculas de ARN, plegamiento que involucra la acción de fuerzas mecánicas[2]. En la Figura 1.2 se muestra la configuración de una horquilla de ARN en los estados plegado y extendido bajo la acción de una fuerza externa f , la cual tira el extremo derecho de la horquilla de ARN. La magnitud Δz denota el cambio en la extensión de la horquilla de ARN.

⁸Cabe mencionar que ésta clasificación de ruidos no es única y existe cierta ambigüedad para considerar que es un ruido “interno” y uno “externo”. Por ejemplo para algunos un ruido es “interno” si existe una relación de fluctuación-disipación, siguiendo esta definición, existen ruidos blancos y de color “internos”. Nosotros utilizaremos la clasificación de Hänggi y Jung.

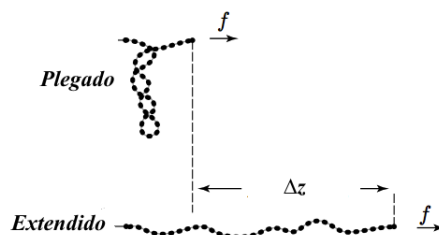


Figura 1.2: Configuración de estados para la horquilla de ARN ante la influencia de una fuerza externa f .

Como antecedente teórico de modelos utilizados para describir la dinámica de plegamiento de horquillas de ARN bajo la influencia de una tensión externa constante, tenemos los trabajos que utilizan a la ecuación maestra para deducir la probabilidad de permanecer en el estado plegado y la del extendido [7] y [8], en estos trabajos se logra esto definiendo a las probabilidades de transición entre estados y se resuelve la ecuación maestra estacionaria. Teniendo así una expresión para la distribución de probabilidad de estar en el estado plegado, pudiendo encontrar por medio de la física estadística las variables termodinámicas pertinentes. Este tipo de modelos encuentra de manera analítica por medio del principio de balance detallado las probabilidades de estar en el estado plegado y las de estar en el extendido, ambas en el caso estacionario y considerando que el sistema es Markoviano. Si bien los modelos que utilizan a la ecuación maestra son congruentes con los experimentos de estiramiento de horquillas de ARN [8], el hecho de considerar a los sistemas como Markovianos, presupone que el ruido presente en el sistema no está correlacionado para distintos tiempos, lo que implica que la naturaleza del ruido, es tal que, éste no afecta de una manera sistemática a la dinámica del sistema. Lo anterior puede ser un supuesto muy simplificador en los experimentos de estiramiento de horquillas de ARN, como se observará más adelante⁹.

En las secciones anteriores se ha comentado que el uso de técnicas utilizadas en los experimentos de moléculas individuales pueden ser aplicadas para reproducir dichas fuerzas y por ende dichos plegamientos.

En los diversos trabajos reportados en la literatura, con la finalidad de comprender como las moléculas de ARN responden a la influencia de fuerzas mecánicas, se procede en términos generales a fijar los extremos de la molécula de ARN, entonces se tira de ellos y se observa como se doblan, voltean y desdoblan, bajo el efecto de la aplicación de la fuerza externa.

Las técnicas más comunes usan pinzas ópticas láser (LOT), las cuales permiten manipular objetos de escala nanométrica al ejercer fuerzas extremadamente pequeñas¹⁰. Puesto que las horquillas de ARN son muy pequeñas para ser manipuladas por ellas mismas, sus extremos se unen a objetos los cuales permitan dicha manipulación. Los objetos manipulados por las LOT en los experimentos asociados a las horquillas de ARN son esferas de poliestireno cuyas dimensiones son del orden de micrómetros. El poliestireno es un polímero que tiene las propiedades de ser transparente, dieléctrico y rígido, lo que permite su manipulación vía LOT.

Los extremos de las horquillas de ARN se unen a las cabezas de poliestireno por medio de asas

⁹Ver sección 3.3.

¹⁰Usualmente de magnitudes de pico Newtons (pN).

hechas de ADN, las cuales son químicamente modificadas para lograr este propósito.

Aún cuando dichas cabezas son transparentes, éstas doblan la luz de los rayos asociados a las LOT, los fotones de los rayos son absorbidos o dispersados por las partículas de poliestireno que conforman dicha esfera, por lo que los fotones transfieren parte de su momento hacia las partículas de poliestireno, lo cual ejerce una fuerza sobre éstas últimas.

Para mantener fija alguna cabeza de poliestireno en algún lugar deseado sólo tiene que colocarse un par de láseres opuestos y se dirigen hacia un foco común. De igual forma, si se desea mover a la cabeza de poliestireno se colocan en la misma configuración los láseres, éstos tienen que irse moviendo al unísono de tal forma que el foco que apunta sobre la cabeza, sólo se vaya trasladando en la dirección deseada.

Esta técnica permite diseñar un arreglo experimental el cual permita estudiar las secuencias de plegamiento y estiramiento de la horquilla de ARN. Se fija un extremo del ARN, manteniendo quieta la esfera de poliestireno correspondiente, mientras que en el otro extremo se induce una fuerza al desplazar el foco de las pinzas ópticas. Este desplazamiento induce una fuerza de tensión en la molécula de ARN, teniendo como resultado secuencias de plegamiento y estiramiento en el ARN.

En la Figura 1.3 se muestra el arreglo experimental descrito, en éste no se han usado las escalas exactas de los componentes del experimento con la finalidad de que se aprecie mejor. Más adelante se muestra un esquema con las dimensiones precisas.

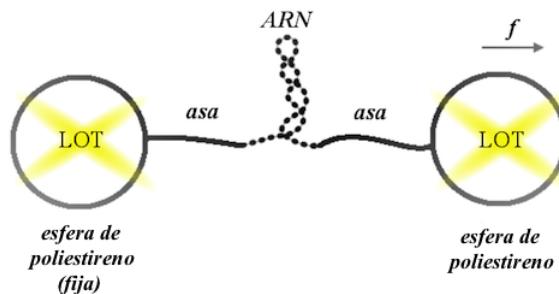


Figura 1.3: Arreglo experimental

El mecanismo por el cual se logran las secuencias de plegado y extendido consiste, en primera instancia, en las situaciones en las que cuando se tensa el ARN por medio de las asas, la fuerza asociada a dicha tensión aumenta suavemente con respecto a la extensión del sistema (Asa-ARN-Asa). Si se tensa el asa con suficiente fuerza, la horquilla de ARN se desenvuelve y por lo tanto se encuentra en el estado extendido. Por otro lado cuando se disminuye la fuerza externa, el ARN se contrae y se pliega, encontrándose así en el estado plegado.

Sin embargo cuando se llega a un valor específico de la fuerza, el cual está asociado a un cambio en la extensión del sistema del tamaño de la horquilla de ARN (Δz en la Figura 1.2), los experimentos muestran que existe una discontinuidad en la fuerza, disminuyendo ésta con respecto al aumento en la extensión del sistema, para después de un pequeño intervalo seguir aumentando. En este intervalo en donde la fuerza decrece y posteriormente aumenta con respecto al aumento de la extensión, se presenta un mismo valor de la fuerza externa f para dos valores distintos de la extensión del ARN (ver inciso c) de la Figura 1.4), los cuales están asociados al estado plegado para el primer valor de la extensión y de extendido para el segundo. Se dice así que para un valor determinado de f existen oscilaciones espontáneas entre estados plegado y extendido en la configuración de la horquilla de ARN.

Estas características de la dinámica del sistema son consistentes con la idea de que la horquilla de ARN, bajo la influencia de una fuerza externa, es un sistema de dos estados: plegado y extendido. Las horquillas de ARN son uno de los sistemas modelados más utilizados para ejemplificar la complejidad en el plegamiento molecular [1], [9] y [10]. En las referencias citadas se expone, que las trayectorias de plegamiento generadas en experimentos donde se usa una fuerza de sujeción en polímeros semiflexibles unidos al final de las horquillas pueden ser usadas para inferir sus respectivos paisajes energéticos. La Figura 1.4 muestra un esquema más específico de un montaje experimental típico.

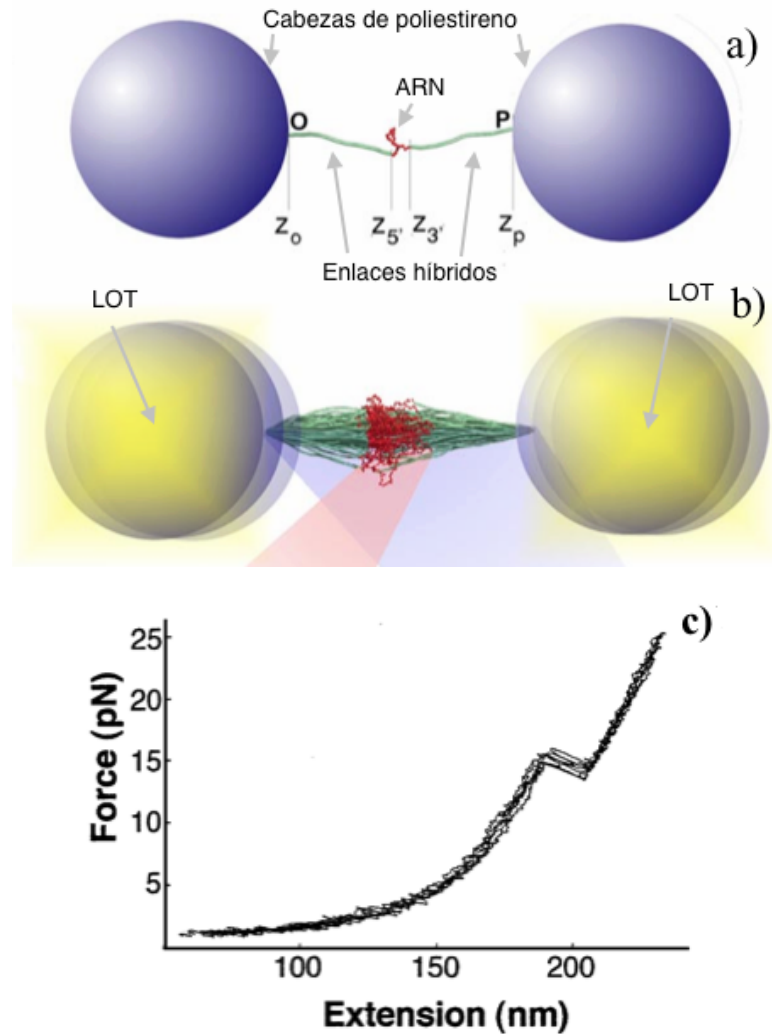


Figura 1.4: Montaje Experimental: a) sistema sin perturbar y b) sistema en presencia de LOT. Figura tomada de [9]. c) Tensión contra extensión, gráfica tomada de [1].

La Figura 1.4 muestra una horquilla de ARN de color rojo, a la cual se unen en sus extremos 5' y 3' enlaces híbridos de ARN/ADN de color verde, los cuales funcionan como asas. Cada extremo libre de los enlaces híbridos está unido a una cabeza de poliestireno de color morado, lo suficientemente grande para poder ser manipulada por LOT. En la parte b) se muestra como por medio de las LOT se fija una cabeza de poliestireno y se manipula a la otra para producir una fuerza de tensión constante f sobre la horquilla.

Es pertinente mencionar que dichos trabajos se centran en la obtención de series de tiempo para las variables z_{sys} y z_m con el afán de obtener los paisajes energéticos asociados.

Aquí, z_{sys} es la distancia entre las cabezas de poliestireno, que representa la extensión del sistema Asa-ARN-Asa¹¹, tal que $z_{sys} = z_0 - z_P$ (ver Figura 1.4).

Por otro lado, z_m es la distancia asociada a la extensión molecular de la horquilla de ARN (estructura de color rojo en la Figura 1.4).

Sólo para visualizar la dinámica de nuestro sistema, presentamos en la Figura 1.5 ejemplos de las series de tiempo obtenidas en la literatura y por nuestro trabajo. En el inciso a) se muestra la serie de tiempo para la extensión obtenida experimentalmente via LOT y presentada en [1]. En el inciso b) se muestra la serie de tiempo obtenida de forma teórica y presentada en [9]. Seguido en el inciso c) se muestra una serie de tiempo para la extensión obtenida a partir del primer modelo propuesto en esta tesis¹². Por último en el inciso d) se muestra la serie de tiempo para la extensión obtenida a partir del segundo modelo propuesto en este trabajo¹³.

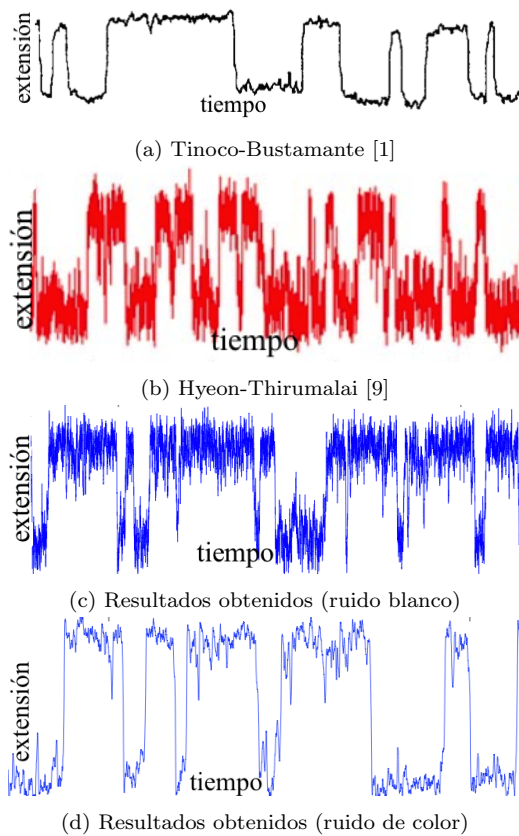


Figura 1.5: Ejemplos de series de tiempo para la extensión del sistema asa-ARN-asa, (extensión contra tiempo), obtenidas experimentalmente (a) y por medio de simulaciones (b), (c) y (d).

¹¹Handle-RNA-Handle en inglés.

¹²Ver capítulo 3 ecuación 3.16.

¹³Ver capítulo 3 ecuación 3.34.

En [1] y [9], las series de tiempo se obtienen de dos formas distintas. Por medio de experimentos de LOT como los descritos arriba y por medio de simulaciones de modelos de polímeros auto-organizados (SOP) ¹⁴. Así también, se usan argumentos teóricos para validar los resultados obtenidos, usando modelos generalizados de Rouse (GRM) en donde se encuentran valores teóricos para z_m y z_{sys} , deduciendo los paisajes energéticos asociados y las tasas de salto entre los estados plegado y extendido para la horquilla del ARN. Dichos modelos se explican más adelante.

En los experimentos LOT y en las simulaciones de [9], se infiere z_m a partir de z_{sys} , y en ambos casos se trabaja con un sistema Asa-ARN-Asa, encontrando los paisajes de energía asociados $F_{eq}(z_m)$ y $F_{eq}(z_{sys})$. Sin embargo, puesto que se desea conocer la dinámica de la horquilla del ARN, en la simulación SOP también se utiliza un modelo sin asas en el cual se obtiene únicamente z_m y su paisaje de energía $F_{eq}^o(z_m)$, así como las tasas de salto entre pozos de energía asociados a las configuraciones del sistema.

Los perfiles obtenidos en [9] se muestran en la Figura 1.6, en particular se observa el perfil energético obtenido para la simulación SOP del sistema sin asas $F_{eq}^o(z_m)$. En dicho perfil se observan los estados de plegado ó NBA (*native basin of attraction*) y extendido ó UBA (*unfolded basin of attraction*), así como la configuración física para la horquilla de ARN en cada caso.

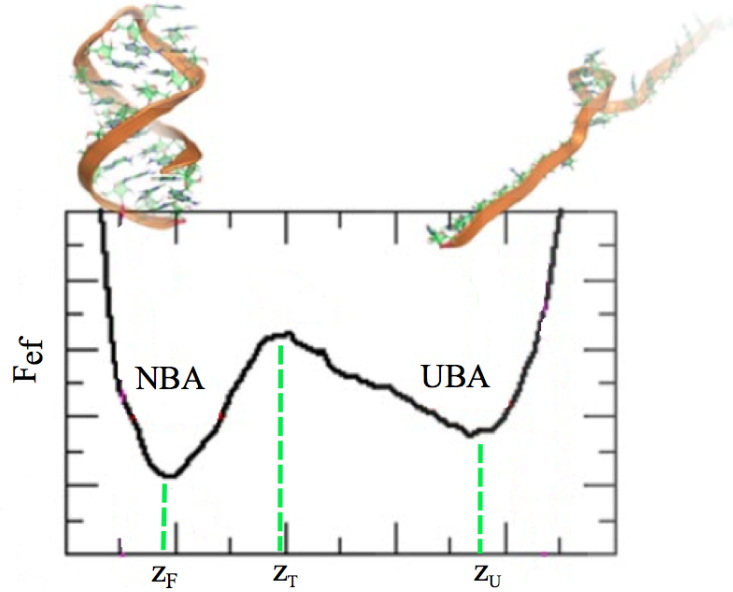


Figura 1.6: Energía libre contra extensión. Paisaje energético y configuración física de la horquilla de ARN, z_F extensión en el estado plegado y z_U extensión en el estado extendido.

Con este antecedente, los trabajos mencionados se centran en el estudio del efecto de la dinámica de las “asas” unidas a la horquilla de ARN sobre los perfiles de energía libre del ARN sin asas ($F_{eq}^o(z_m)$).

¹⁴Self-Organized Polymer, en inglés. El SOP es una estructura basada en un modelo de grano-grueso para una cadena polipeptídica, el SOP describe la evolución temporal de la cadena mediante la ecuación de Langevin [3].

Con el afán de obtener medidas más precisas de $F_{eq}^o(z_m)$, se definen las cantidades L y l_p . Aquí, L es la longitud de contorno del asa que une a los polímeros con la horquilla de ARN, y l_p es la longitud de persistencia¹⁵ del polímero asociado a la asa. Usando polímeros rígidos con un cociente $\frac{L}{l_p}$ pequeño, se obtienen dichas medidas.

Tomando en cuenta esto, los trabajos citados tratan de mostrar que cuando el perfil de energía libre de equilibrio interactúa con una fuerza externa de tensión crítica igual a $f_m \approx 15.4 pN$, se tiene que los estados plegado y extendido son igualmente probables.

Haciendo uso de la teoría de Kramers, proveen estimaciones precisas de las tasas de salto dependientes de la fuerza, en ausencia de asas y para valores arbitrarios de la fuerza externa f , implicando que z_m es una buena coordenada de reacción para las horquillas de ARN bajo tensión.

La forma de proceder en [9] para construir la forma del paisaje energético es usando trayectorias dinámicas, las cuales son generadas al aplicar una fuerza f sobre los extremos de los ARN, es tomar en cuenta el teorema ergódico, que plantea que para un observable X y para tiempos muy grandes existe una igualdad entre los promedios temporales y espaciales¹⁶.

Se supone que la distribución del observable denotada por $P(X)$ converge a su distribución estacionaria $P_{eq}(X)$ ¹⁷. Llamando a nuestro observable $X = z_{3'} - z_{5'} \sim z_m$, donde z_m es la distancia de extremo a extremo de la horquilla tal que dicha coordenada de reacción es conjugada a la fuerza f .

Utilizando el *ensemble* canónico, dicha distribución estacionaria nos permite encontrar el perfil de energía libre en el equilibrio $F_{eq}(z)$ por medio de la relación[9]:

$$\beta F_{eq}(z_m) = -\log P_{eq}(z_m) \quad (1.1)$$

Aquí se supone que las fluctuaciones en las asas perturban de manera mínima a la dinámica de la horquilla.

Aunque las fluctuaciones en las distancias L y l_p pueden distorsionar la señal proveniente de la horquilla, y también pueden afectar la dinámica de la horquilla, dicha distorsión suele atribuirse al proceso de medición. El efecto sobre la dinámica de la horquilla se debe a la existencia de un acoplamiento entre ésta y los instrumentos.

Las tasas de plegamiento y estiramiento dependientes de f son muy cercanas a sus valores ideales, es decir, aquellos que se obtuvieron por medio de simulaciones de modelos sin asas, sólo cuando las asas son flexibles. La dinámica de las fluctuaciones de la distancia relativa entre las esferas de poliestireno se puede describir por medio de la ecuación de Langevin:

$$\gamma \frac{dz_p}{dt} = -\kappa_b z_p + \xi(t), \quad (1.2)$$

¹⁵La longitud de persistencia se define como la longitud máxima promedio de la cadena polimérica que permanece recta. Una mayor longitud de persistencia implica que el polímero es rígido [5].

¹⁶ $\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{k=0}^{t-1} X_k = \frac{1}{\mu} \int X d\mu$, donde μ es una medida de probabilidad. En el contexto de la física estadística, en donde se obtienen valores para macroestados, promediándose sobre la colección de microestados posibles que componen el *ensemble* y en particular considerando los diferentes tipos de *ensembles*, una medida de probabilidad esta asociada al volumen fase en el *ensemble* microcanónico y la medida de Gibbs en el canónico .

¹⁷Esta notación denota que la distribución es una de equilibrio. Sin embargo, más adelante veremos que esto no es cierto, ya que las distribuciones encontradas son estacionarias, pero reflejan un proceso de no equilibrio subyacente. Por lo que el subíndice utilizado por los autores citados [9], es un abuso de notación.

donde κ_b es la constante de restitución asociada a la fuerza restauradora correspondiente a las pinzas ópticas y al sistema Asa-ARN-Asa. $\xi(t)$ es la fuerza aleatoria tal que $\langle \xi(t) \rangle = 0$ y $\langle \xi(t)\xi(t') \rangle = 2\gamma kT\delta(t-t')$, cumpliendo así con las propiedades de un ruido térmico (blanco).

La esfera relaja a su posición de equilibrio en una escala de tiempo $\tau_r = \frac{\gamma}{\kappa_b} \leq 1ms$, donde $\kappa_b = \kappa_p + \kappa_m$, con κ_p la rigidez de la trampa para las LOT y κ_m la rigidez asociada al sistema Asa-RNA-Asa. La rigidez total κ_b es la suma ya que, para el caso de las LOT la esfera de poliestireno está confinada dentro de un campo electromagnético el cual hace funcionar a ésta como un oscilador armónico con una constante de rigidez κ_p , y desde que el sistema Asa-ARN-Asa está unido a la cabeza de poliestireno, éste funciona como otro oscilador armónico con constante κ_m , el cual esta conectado en paralelo al oscilador representado por las LOT. Por lo que la rigidez total es la suma de las correspondientes constantes.

Aquí, el coeficiente de fricción asociado a la ley de Stokes es $\gamma = 6\pi\eta a$, con $a = 1\mu m$ el radio de las cabezas de poliestireno, $\eta \approx 1cP$ la viscosidad del medio, $\kappa_p \approx 0,01\frac{pN}{nm}$ y $\kappa_m \approx 0,1\frac{pN}{nm}$.

Como se había anticipado, en los experimentos de LOT, cuando la fuerza de sujeción f es tal que es igual a un valor crítico $f \approx f_m$, se obtiene que los estados NBA y UBA son igualmente probables.

Con la finalidad de corroborar las mediciones de z_m y z_{sys} realizadas en los experimentos LOT para los trabajos mencionados en la referencia [9], en la misma, se simuló un modelo de polímero auto organizado (SOP) en la horquilla P5GA y aplicaron una fuerza externa de $f = f_m \approx 15,4pN$. La fuerza es ejercida en el extremo del asa unida al final del 3' del ARN (punto P, ver Figura 1.3), mientras el otro extremo esta fijo, punto O.

La simulaciones del P5GA con asas de longitud $L = 25nm$ y $l_p = 70nm$ mostraron que la extensión del sistema entero, denotado por $z_{sys} = z_p - z_o$, fluctúa entre dos límites centrados alrededor de $z_{sys} \approx 50nm$ y $z_{sys} \approx 56nm$. Dichas fluctuaciones corresponden a los saltos del ARN entre los estados NBA y UBA. La variable z_{sys} refleja las transiciones en z_m al realizar la descomposición: $z_{sys} = z_{H5'} + z_m + z_{H3'}$ (longitud del sistema asas, ARN), con $z_{H5'} = z_{5'} - z_o$ y $z_{H3'} = z_p - z_{3'}$ (ver Figura 1.4). Por lo que por medio de esta descomposición se puede obtener la dinámica de z_m .

A partir de las series de tiempo obtenidas, se pueden inferir las distribuciones estacionarias¹⁸ que en [9] son llamadas de equilibrio y son denotadas por: $P_{eq}(X)$ para $X = z_{sys}, z_{H5'}, z_m, z_{H3'}$.

Para investigar el efecto de las asas en el paisaje energético, se repitieron simulaciones con la misma fuerza de sujeción para $L = 5 - 100nm$ y l_p de $0,6nm$ y $70nm$. Encontraron que la discrepancia entre las energías libres $F_{eq}(z_{sys})$ y $F_{eq}(z_m)$ aumenta entre más flexible y largas sean las asas. Para l_p pequeña y L grande, los pozos de atracción en $F_{eq}(z_m)$ no están bien definidos. La mayor desviación entre las energías libres se encontró cuando $l_p = 0,6nm$ y $L = 25nm$ ($\frac{L}{l_p} \approx 40$). En cambio, la mayor coincidencia entre las energías libres se dió en el caso $l_p = 70nm$ y $L = 5nm$ ($\frac{L}{l_p} \approx 0,07$).

Con la meta de confirmar que el uso de asas rígidas es necesario para obtener mediciones precisas del paisaje energético, se introdujo el modelo GRM (*Generalized Rouse Model*) para representar el ARN y las asas en términos de una cadena Gaussiana con N_0 monómeros, cada uno de longitud a (longitud de Khun). Los extremos del ARN se modelan atrapados en un potencial armónico con rigidez κ y con una distancia de corte de $c = 4nm$. Se unen dos asas a los extremos del ARN, cada una con N_h monómeros y longitud de Khun b . Se fija un extremo y al otro se le aplica una fuerza

¹⁸No toda distribución estacionaria es una distribución de equilibrio.

$f_m \approx 15,4pN$. En este modelo, las energías libres en función de z_m y z_{sys} se pueden resolver de manera exacta para una representación continua de ambas.

Usando los valores de $\kappa \approx 0,54 \frac{pN}{nm}$ y $N_0 = 20$, y se encontró que una menor longitud para las asas Gaussianas en el modelo GRM imita asas rígidas en el sistema Asa-ARN-Asa. Las energías libres que se encontraron en este esquema son consistentes con los resultados de las simulaciones de [9]. Los perfiles de energía libre se desvían de F_{eq}^o cuando la rigidez de las asas disminuye, i.e., N_h se incrementa. Se plantea que el modelo GRM confirma la medición del paisaje energético, el cual requería asas rígidas al usar z_{sys} .

Así también, en [9], por medio de la simulación de P5GA sin asas, se estimaron las tasas de salto para este caso. Se recogieron de las series de tiempo para z_m y se obtuvo su respectivo histograma, por medio del cual se determinaron los estados NBA y UBA, y, por medio de (1.1) se encontró que los mínimos en la la energía libre para NBA y UBA fueron $z_F = 1,9nm$ y $z_U = 7,4nm$, respectivamente. Dichos trabajos afirman que el modelo GRM es consistente con las simulaciones P5GA.

Como se mencionó anteriormente, en estos trabajos para los experimentos usando LOT se encontró que se necesitan asas rígidas para obtener una $F_{eq}(z_{sys})$ que se asemeje a $F_{eq}^o(z_m)$. Mientras que cuando se usaron asas flexibles, se produjeron tasas de salto cercanas a las obtenidas sin asas. Lo que conlleva a un resultado contradictorio. A pesar de esto, se argumenta que si z_m es una buena coordenada de reacción, entonces, es posible extraer las tasas de salto usando una medida exacta de $F_{eq}(z_{sys}) \approx F_{eq}(z_m) \approx F_{eq}^o(z_m)$ en $f \approx f_m$, para asas con un cociente $\frac{L}{l_p}$ pequeño. Obteniendo los tiempos de plegamiento y estiramiento usando el tiempo medio de primer paso con sus condiciones de frontera apropiadas (expresión de tasa de Kramers).

Los correspondientes cálculos para energías libres y tiempos de plegamiento y estiramiento desprendidos de la teoría de Kramers concuerdan con las simulaciones del P5GA [9].

Por lo anterior, se afirma que mediciones precisas de la sucesión de plegamientos de ARN son suficientes para obtener buenas estimaciones de las tasas de salto dependientes de f en la ausencia de asas.

Ahora, este enfoque (usado en [9] y explicado arriba) para describir la dinámica de plegamiento de la horquilla de ARN, implica hacer uso de la relación (1.1). Dicho artificio implica que nuestras series de tiempo z_{sys} y z_m cumplen la hipótesis ergódica. Dicha hipótesis implica que el sistema modelado eventualmente llegará a una distribución estacionaria. Usualmente y de forma errónea se asocia a las distribuciones estacionarias a distribuciones de equilibrio. Esto no es necesariamente cierto¹⁹, el principal resultado del teorema ergódico es que la dinámica del sistema es tal que para tiempos muy grandes, todos los estados posibles asociados al sistema son visitados, cumpliéndose con esto la igualdad entre promedios temporales y espaciales. Esto permite que, a partir de las series de tiempo para las coordenadas de reacción pertinentes²⁰ recogidas de nuestros experimentos ó simulaciones, podamos encontrar un distribución temporal de la serie y se asocie a una distribución espacial, la cual pueda ser manipulada por medio del formalismo de la física estadística: encontrar función de partición energía libre, entropía, etc.

El problema que surge es que la física estadística tradicional implica el uso de las distribuciones

¹⁹Como se observará en los siguientes capítulos en donde se resuelven las ecuaciones de evolución para el sistema estudiado y se encuentran las correspondientes distribuciones estacionarias.

²⁰ z_{sys} y z_m en nuestro caso.

de probabilidad convencionales²¹ aplicadas a situaciones de equilibrio. Si las series de tiempo encontradas no son de equilibrio, dicho esquema en primera instancia deja de ser útil. Una respuesta a dicho dilema es la modificación de las distribuciones de probabilidad a unas de no equilibrio, las cuales sean más próximas a las de las series de tiempo encontradas²² y, una vez conseguido esto insertarlas en el formalismo de la física estadística, como se hace sin explicarlo en [9]. Aquí se malinterpreta a la distribución estacionaria encontrada como una de equilibrio.

Es claro que desde que se está tratando con un sistema bioquímico susceptible a situaciones fuera del equilibrio, la idea de que dicho sistema puede acceder a una distribución de equilibrio es muy optimista. De hecho, las distribuciones temporales de las series de tiempo encontradas son de no equilibrio, desde que presentan fluctuaciones entre dos estados distintos.

Para evitar malinterpretaciones o abuso de los resultados del teorema ergódico, en esta tesis se propone un método alternativo de análisis, el cual explote la naturaleza de no equilibrio de los procesos estocásticos asociados a la dinámica de las horquillas de ARN, y utilice la hipótesis ergódica lo menos posible. En una primera aproximación sin el uso de la hipótesis ergódica, se utilizan los datos obtenidos a partir de las simulaciones de los modelos SOP P5GA dado que son una medida precisa de la dinámica del ARN[9]. En concreto, se utilizan los resultados para los paisajes energéticos del sistema Asa-ARN-Asa, y a partir de éstos se obtiene la fuerza que interviene en la dinámica del sistema. Teniendo ésta, es posible obtener, en primera instancia los valores de los procesos z_{sys} ó z_m por medio de la respectiva ecuación de Langevin. Dado esto, también es posible calcular la densidad de probabilidad que caracteriza a nuestro sistema, la cual es la solución de la ecuación de Fokker-Planck asociada a la ecuación de Langevin utilizada. Posteriormente, se muestra que aún usando la hipótesis ergódica se obtienen distribuciones estacionarias de no equilibrio para la solución de la ecuación de Fokker-Planck estacionaria. Dicha forma de proceder es más que pertinente para nuestro caso, ya que desde que los coeficientes de deriva son no lineales, la solución por métodos numéricos es mucho más complicada que para otras ecuaciones diferenciales parabólicas con coeficientes lineales.

Como colofón, la estructura de nuestro modelo nos permitirá analizar sistemas con ruido de color, en los que existe un acoplamiento de las LOT con la dinámica de plegamiento del sistema.

El método propuesto se presenta en el capítulo 3. Antes se exponen de manera más amplia las herramientas utilizadas en el desarrollo del presente trabajo.

²¹Medida de Gibbs asociada al *ensemble* canónico.

²²Un ejemplo de dichas distribuciones de no equilibrio son las llamadas Medidas de Sinai, Ruelle y Bowen [11].

CAPÍTULO 2

Herramientas del modelo

Se introducen las herramientas físicas y matemáticas utilizadas. En concreto se presenta el esquema de termodinámica irreversible lineal y cinética química, haciendo énfasis en la producción de entropía para esta clase de sistemas. Después se construyen, para diversos casos, las ecuaciones de Fokker-Planck y Langevin a partir de procesos estocásticos. Enseguida, se expone el esquema desarrollado por Kramers para describir las tasas de salto para un sistema de dos estados. Y por último se muestra el esquema de simulación estocástica empleado en nuestro modelo.

2.1. Termodinámica irreversible

Dada la naturaleza de los sistemas a estudiar (dinámica de horquillas de ARN), introduciremos la Termodinámica Irreversible Lineal, para los casos de reacciones químicas y procesos de difusión. A continuación se desarrolla la ecuación de balance de entropía a partir de los principios básicos de la termodinámica. A partir de dicha ecuación de balance se puede caracterizar a las reacciones químicas y a los procesos de difusión que intervienen en el comportamiento de las estructuras biológicas a tratar. Para nuestro caso, dicho proceso de difusión es la dinámica de plegamiento de la horquilla de ARN y no hay reacciones químicas. La obtención de la ecuación de balance de entropía nos permite identificar los flujos y fuerzas que intervienen en los procesos mencionados. A continuación presentamos por completez el caso general para obtener la ecuación de balance de entropía y encontrar a la producción de entropía en términos de fuerzas y flujos generalizados, cabe mencionar, que en este caso general, los flujos y las fuerzs pueden estar asociados a procesos de difusión, reacciones químicas, etc. Pero para nuestro modelo sólo se tomarán en cuenta el flujo y la fuerza generalizada asociados al proceso de difusión de la horquilla de ARN.

Para encontrar la ecuación de balance de entropía partimos de dos principios básicos de la termodinámica: la conservación de masa, y, el principio de entropía creciente para un sistema cerrado. Para un sistema abierto, el cambio en la entropía durante un instante de tiempo dt se puede escribir de la siguiente manera[12], [13]

$$\frac{dS}{dt} = \frac{d_e S}{dt} + \frac{d_i S}{dt}, \quad (2.1)$$

donde $\frac{d_e S}{dt}$ es el flujo de entropía debido a intercambios de energía o materia con el ambiente y $\frac{d_i S}{dt}$ es la producción de entropía debida a procesos irreversibles dentro del sistema, como: difusión, conducción de calor y reacciones químicas.

Si el sistema tratado se divide en r partes y se utiliza la hipótesis de equilibrio local, al aplicar la segunda ley de la termodinámica tenemos:

$$d_i S = d_i S^1 + d_i S^2 + \dots + d_i S^r; d_i S^k \geq 0, \forall k, \quad (2.2)$$

donde cada $d_i S^k$ es el cambio de entropía debida a procesos irreversibles para la k -ésima parte que compone el sistema.

Sabemos que la ecuación de estado para la entropía debe de estar en términos de las variables extensivas del sistema, en nuestro caso supondremos que $S = S(U, V, X_1, \dots, X_n)$, donde U es la energía interna, V es el volumen, X_i es el número de moléculas para cada constituyente químico del sistema¹.

Si, además, las variables extensivas dependen de la posición y del tiempo, podemos decir que la entropía también es función de estos, i.e. $S = S(U, V, X_i, \dots, X_n, \vec{r}, t)$. Tomando en cuenta lo anterior, el aumento local de la entropía se puede definir utilizando la densidad de entropía por unidad de volumen $s = s(u, v, \rho_1, \dots, \rho_n, \vec{r}, t)$, donde u es densidad de energía interna por unidad de volumen, v es el volumen específico y $\rho_1 \dots \rho_n$ son las concentraciones de los n constituyentes químicos de nuestro sistema $(X_1, \dots, X_n)^2$.

Por lo tanto se define a la densidad de entropía s como:

$$S(U, V, X_i, \dots, X_n, \vec{r}, t) = \int s(\vec{r}, t) dV; s = s(u, v, \rho_1, \dots, \rho_n, \vec{r}, t), \quad (2.3)$$

Teniendo como resultado que la densidad de entropía se convierte en una función implícita del tiempo.

Al tratar con las concentraciones y su dependencia en el tiempo, es imprescindible introducir el principio de conservación de masa.

El principio de conservación de masa puede ser escrito como:

$$\frac{dm}{dt} \equiv \frac{d_e m}{dt}, \quad (2.4)$$

en donde el cambio en la masa en el tiempo en un volumen dado del espacio debe de ser igual a la masa que fluye a través de una superficie arbitraria Σ que limita el volumen, sin embargo, esta relación nos dice que no existe producción de masa dentro del volumen V . La situación cambia si uno se fija por separado en la masa de cada uno de los constituyentes químicos del sistema, haciéndolo se tiene:

$$\frac{dm_j}{dt} = \frac{d_e m_j}{dt} + \frac{d_i m_j}{dt}; j = 1, \dots, n, \quad (2.5)$$

¹Como mencionamos esta deducción de la ecuación de balance de entropía es para el caso general en el que el sistema esta compuesto por n constituyentes, para nuestro caso particular solo existe un constituyente en el sistema, pero las ecuaciones encontradas para la producción de entropía siguen siendo válidas.

²De nuevo, en nuestro caso de un sistema con un solo constituyente, la concentración ρ_j se referirá a la concentración de masa, que es la masa del constituyente j en cierto estado, dividido entre el volumen.

donde $\frac{d_i m_j}{dt}$ =producción de X_j resultante de las reacciones químicas. Si se parte de que el constituyente X_j participa en r reacciones químicas y designamos como $W_k (k = 1, \dots, r)$ a las tasas de reacción individuales, entonces la producción de los constituyentes se puede expresar como

$$\frac{d_i m_j}{dt} = \sum_{k=1}^r \nu_{jk} W_k, \quad (2.6)$$

donde ν_{jk} es el coeficiente estequiométrico de la reacción química presente entre los constituyentes X_j y X_k . Típicamente, las ecuaciones (2.6) determinan la cinética de la reacción y pueden escribirse la ley de acción de masas[12]. Utilizando a la concentración de masa ρ_j como una densidad de masa y la reacción por unidad de volumen w_k tal que $W_k = \int w_k dV$, la ecuación (2.5) es ahora

$$\frac{d}{dt} \int \rho_j dV = \frac{d_e m_j}{dt} + \sum_{k=1}^r \nu_{jk} \int w_k dV. \quad (2.7)$$

En general, para un sistema químico espacialmente inhomogéneo, el término de flujo $\frac{d_e m_j}{dt}$ describe como penetra o sale la materia dentro del volumen V a través de la superficie Σ . Si dicho flujo lo denominamos como \vec{j}_j^Σ se tiene

$$\frac{d_e m_j}{dt} = - \int_{\Sigma} \vec{n} \cdot \vec{j}_j^\Sigma d\Sigma. \quad (2.8)$$

La anterior relación simplemente muestra que el flujo que entra al sistema es igual al que sale. Aplicando el teorema de la divergencia al término de la derecha para la última expresión encontrada, obtenemos

$$\int_{\Sigma} \vec{n} \cdot \vec{j}_j^\Sigma d\Sigma = \int_V \nabla \cdot \vec{j}_j dV. \quad (2.9)$$

Combinando este último resultado con (2.7) y (2.8), uno puede obtener una ecuación local para ρ_j , pidiendo que se cumpla la igualdad de los integrandos para cada punto de V y dado que $\frac{d}{dt}$ conmuta con la integral. Obtenemos la ecuación de balance para las concentraciones de masa

$$\frac{\partial \rho_j}{\partial t} = -\nabla \cdot \vec{j}_j + \sum_k \nu_{jk} w_k. \quad (2.10)$$

Esta ecuación muestra el comportamiento de cada uno de los constituyentes del sistema, el cual está determinado por cada flujo de difusión de cada componente. Dichos flujos se denotan como \vec{j}_i . Así también, el comportamiento de los constituyentes esta influenciado por cada tasa de reacción por unidad de volumen entre ellos w_k .

Para obtener la ecuación de balance de entropía derivamos (2.3) con respecto al tiempo, obteniendo

$$\frac{\partial s}{\partial t} = \sum_i \left[\frac{\partial s}{\partial \rho_i} \right] \frac{\partial \rho_i}{\partial t}. \quad (2.11)$$

En el análisis de equilibrio local al tomar la diferencial total a $s(u, v, \rho_i)$ se cumple

$$\frac{\partial s}{\partial \rho_i} = -\frac{\mu_i}{T}, \quad (2.12)$$

donde μ_i es el potencial químico por unidad de masa del constituyente X_i y T es la temperatura. Sustituyendo esto en (2.11) se consigue

$$\frac{\partial s}{\partial t} = - \sum_i \frac{\mu_i}{T} \frac{\partial \rho_i}{\partial t}. \quad (2.13)$$

La relación (2.13) es un caso particular de la ecuación de Gibbs, la cual en general es

$$TdS = dU + pdV - \sum_i \mu_i dN_i + \sum_j F_j dx_j, \quad (2.14)$$

donde U es la energía interna, p es la presión, N el número de constituyentes químicos y F_j una fuerza conjugada a la coordenada de reacción x_j .

Ahora, introduciendo la ecuación de balance de masa (2.10) en (2.13) llegamos a

$$\frac{\partial s}{\partial t} = - \sum_i \sum_k \frac{\mu_i}{T} \nu_{ik} w_k + \sum_i \frac{\mu_i}{T} \nabla \cdot \vec{j}_i \quad (2.15)$$

$$= \sum_k \left\{ - \sum_i \frac{\mu_i}{T} \nu_{ik} \right\} w_k + \nabla \cdot \sum_i \frac{\mu_i}{T} \vec{j}_i - \sum_i \vec{j}_i \cdot \nabla \frac{\mu_i}{T}. \quad (2.16)$$

En el segundo término del lado derecho (2.16) se utilizó la igualdad para la divergencia del producto de un campo escalar por uno vectorial de la forma: $\nabla \cdot (\phi \vec{A}) = (\nabla \phi) \cdot \vec{A} + \phi \nabla \cdot \vec{A}$.

En nuestro caso, el campo escalar es el potencial químico y el campo vectorial es el flujo de difusión.

En la ecuación (2.16) reconocemos que el término que está entre las llaves es la afinidad de la k -ésima reacción $\implies \mathcal{A}_k = - \sum_i \mu_i \nu_{ik}$, que mide la desviación de la reacción química con respecto a su estado de equilibrio. Teniendo esto en cuenta podemos reescribir (2.16) como

$$\frac{\partial s}{\partial t} = \nabla \cdot \sum_i \frac{\mu_i}{T} \vec{j}_i - \sum_i \vec{j}_i \cdot \nabla \frac{\mu_i}{T} + \sum_k \frac{\mathcal{A}_k}{T} w_k. \quad (2.17)$$

El lado derecho de la ecuación (2.17) puede ser descompuesto en dos partes

$$\nabla \cdot \vec{J}_s \equiv - \nabla \cdot \sum_i \frac{\mu_i}{T} \vec{j}_i, \quad (2.18)$$

$$\frac{d_i s}{dt} = - \sum_i \vec{j}_i \cdot \nabla \frac{\mu_i}{T} + \sum_k \frac{\mathcal{A}_k}{T} w_k. \quad (2.19)$$

Aquí \vec{J}_s es el flujo de entropía asociado al sistema y $\frac{d_i s}{dt} \geq 0$ es la producción de entropía debida a procesos irreversibles. La relación (2.17) es una ecuación de balance de entropía, la cual se escribe como

$$\frac{\partial s}{\partial t} = - \nabla \cdot \vec{J}_s + \frac{d_i s}{dt}. \quad (2.20)$$

Se observa de (2.19) que la producción de entropía es una forma bilineal de las cantidades \vec{j}_i y de w_k . Éstas pueden ser pensadas como flujos \vec{J}_l los cuales están asociados a los “ l ” procesos irreversibles existentes. Las cantidades $\nabla \frac{\mu_i}{T}$ y $\frac{\mathcal{A}_k}{T}$, pueden ser pensadas como las fuerzas generalizadas \vec{F}_l , que

producen dichos flujos.

Lo anterior nos lleva a identificar la siguiente estructura

$$\frac{d_i s}{dt} = \sum_l \vec{F}_l \vec{J}_l, \quad (2.21)$$

en donde para el caso de difusión debida a la inhomogeneidad espacial se identifican los respectivos flujos y fuerzas como

$$\vec{J}_i^{diff} = \vec{j}_i; \quad \vec{F}_i^{diff} = -\nabla \frac{\mu_i}{T}. \quad (2.22)$$

Para el caso de reacciones químicas tenemos

$$\vec{J}_k^{react} = w_k; \quad \vec{F}_k^{react} = \frac{a_k}{T}. \quad (2.23)$$

En los casos donde existe conducción de calor y conducción eléctrica, los flujos y las fuerzas son respectivamente

$$\vec{J}_u^{heat} = \vec{j}_u; \quad \vec{F}_u^{heat} = \nabla \frac{1}{T}, \quad (2.24)$$

$$\vec{J}_l^{elect} = \vec{I}_l; \quad \vec{F}_l^{elect} = -\frac{\nabla \varphi}{T}. \quad (2.25)$$

En general, para sistemas como los analizados, las fuerzas causan los flujos. En el equilibrio, todas las fuerzas y los correspondientes flujos desaparecen. Esto implica que los flujos son función de las fuerzas. Por lo que para una pequeña desviación de las fuerzas de su estado de equilibrio, los flujos pueden representarse como una función lineal de las fuerzas[14]

$$\vec{J}_l = \sum_j L_{lj} F_j. \quad (2.26)$$

Aquí L_{kj} son los coeficientes fenomenológicos de Onsager. Sustituyendo (2.26) en (2.21), la producción de entropía es

$$\frac{d_i s}{dt} = \sum_{jl} L_{jl} F_j F_l > 0. \quad (2.27)$$

Las relaciones (2.27) y (2.26) son importantes, ya que permiten introducir en nuestro esquema parámetros fenomenológicos. En el caso de reacciones químicas los coeficientes de Onsager están asociados a las tasas de reacción pertinentes. Para procesos de difusión, los coeficiente de Onsager se relacionan con sus respectivos coeficientes de difusión. La ecuación (2.27) nos permitirá más adelante encontrar la producción de entropía para la dinámica de plegamiento de las horquillas de ARN estudiadas.

Estas relaciones pertenecientes a la Termodinámica Irreversible Lineal pueden agruparse en una ecuación de balance de entropía global, es decir, independiente de la posición, la cual caracteriza el comportamiento temporal de nuestro sistema.

Al integrar (2.20) sobre el volumen ocupado obtenemos³

$$\frac{\partial S}{\partial t} = - \int \nabla \cdot \vec{J}_s dV + \int \frac{d_i s}{dt} dV = - \int \vec{n} \cdot \vec{J}_s d\Sigma + \int \frac{d_i s}{dt} dV \quad (2.28)$$

³Aquí se aplica el Teorema de la divergencia en el primer término de la derecha integrando sobre el elemento de superficie $d\Sigma$.

$$= \frac{\partial_e S}{\partial t} + \frac{\partial_i S}{\partial t} \quad (2.29)$$

en donde se ha llegado a una relación de la forma de (2.1), en concreto con esto se muestra por qué el cambio en la entropía de un sistema se puede representar como la suma de un flujo y un término de producción de entropía. Los componentes de la producción de entropía representados como flujos y fuerzas, caracterizan de buena forma a las reacciones químicas y procesos de difusión que protagonizan los sistemas biológicos estudiados.

Por último, en esta sección mostramos el principal vínculo entre la termodinámica irreversible y la teoría de procesos estocásticos que será expuesta en el siguiente apartado. Como se observó a partir del principio de conservación de masa, llegamos a una ecuación de balance para las concentraciones (2.10), dichas concentraciones ρ_j pueden considerarse como densidades de masa⁴, y en un sentido amplio como densidades de probabilidad⁵. Si bien la ecuación de balance (2.10) es algo distinta a la ecuaciones de evolución para las funciones de densidad de probabilidad usuales⁶, al momento de plantear la primera ley de Fick

$$\vec{j}_j = -D_j \nabla \rho_j, \quad (2.30)$$

y aplicarla a dicha ecuación de balance de concentración, tenemos la ecuación de reacción-difusión siguiente

$$\frac{\partial \rho_j}{\partial t} = \sum_k \nu_{jk} w_k + D_j \nabla^2 \rho_j. \quad (2.31)$$

Aquí D_j es un coeficiente de difusión, la ecuación (2.31) posee la estructura de una ecuación de evolución para la función de densidad de una variable estocástica.

2.2. Procesos estocásticos

Como se comentó en el apartado anterior, la dinámica que presentan las moléculas individuales como las horquillas de ARN es no determinista. Por esta razón su modelación debe hacerse usando procesos estocásticos y sus respectivas ecuaciones de evolución, para la variable y su función de densidad. A continuación se presentan los conceptos básicos para poder construir dichas ecuaciones, para después obtenerlas y mostrar la relación que existe entre ellas.

2.2.1. Probabilidad y procesos estocásticos

La definición más sencilla de un proceso estocástico es la siguiente[16]:

Definición 1: Un proceso estocástico es una sucesión “temporal” $\{X_t\}_{t \in I}$ de variables aleatorias.

El conjunto de índices I puede ser discreto ($I = \mathbb{N}$) lo que nos denotaría un proceso estocástico en tiempo discreto. Por otro lado si I es continuo ($I = \mathbb{R}$) tenemos un proceso estocástico en tiempo continuo. Un proceso estocástico en tiempo continuo se representa como: $X : \mathbb{R} \times \Omega \rightarrow \mathbb{R}^d$,

⁴Puesto tienen las mismas unidades.

⁵Al multiplicar por un factor que normalice la densidad de masa.

⁶Ver ecuación de Fokker-Planck en los apartados 2.2.3 y 2.2.5.

$(t, \omega) \mapsto X(t, \omega)$; donde Ω es el espacio muestral de nuestro sistema. Dicho espacio es el conjunto de todos los posibles resultados individuales del experimento aleatorio estudiado, denotado cada uno por ω , tal que $\omega \in \Omega$.

Para un tiempo fijo, cada X_t es una variable estocástica, la cual puede tomar un valor dado en $x(t)$, de modo que puede definirse su función de distribución de probabilidad como ⁷

$$\begin{aligned} p(x, t) &= \int_{\Omega} \delta(x - X_t(\omega)) d\mu(\omega) \\ p(x, t) &= \int_{\Omega} \delta(x - X_t(\omega)) p(X_t(\omega)) dX_t(\omega). \end{aligned} \quad (2.32)$$

Aquí μ es una medida de probabilidad definida sobre la σ -álgebra \mathcal{F} (conjunto potencia del espacio muestral Ω). En concreto se utiliza la estructura de espacio de probabilidad $(\Omega, \mathcal{F}, \mu)$. Es así como se puede generar una distribución de probabilidad dependiente del tiempo sobre \mathbb{R}^d . Generalizando podemos definir la distribución de probabilidad conjunta para las variables aleatorias X_1, X_2, \dots, X_n como

$$p_n(x_1, t_1; \dots; x_n, t_n) = \int_{\Omega} \delta(x_1 - X_{t_1}) \cdots \delta(x_n - X_{t_n}) d\mu(\omega). \quad (2.33)$$

Con la ayuda de esta función de distribución conjunta para n puntos, se pueden calcular los momentos que dependen del tiempo para el proceso estocástico estudiado, y esto se hace de la siguiente manera (momento n -ésimo)

$$\langle X(t_1) \cdots X(t_n) \rangle = \int_{\mathbb{R}^d} dx_1 \cdots dx_n x_1 \cdots x_n p_n(x_1, t_1, \dots, x_n, t_n). \quad (2.34)$$

Así mismo, la matriz de covarianza para dos procesos estocásticos dependientes del tiempo tiene por elementos a

$$\begin{aligned} Cov(X_i(t_1), X_j(t_2)) &= \sigma_{ij}^2(t_1, t_2) \\ &= \langle X_i(t_1), X_j(t_2) \rangle - \langle X_i(t_1) \rangle \langle X_j(t_2) \rangle, \end{aligned} \quad (2.35)$$

donde los elementos diagonales de esta matriz son las funciones de autocorrelación y los que no estan en la diagonal son las funciones de correlación cruzada.

La jerarquía de las funciones de distribución p_n caracteriza de manera completa a un proceso estocástico si p_n satisface:

- Positividad, $p_n \geq 0$.
- Simetría,

$$p_n(\dots; x_k, t_k; \dots; x_l, t_l; \dots) = p_n(\dots; x_l, t_l; \dots; x_k, t_k; \dots).$$

- Completez,

⁷La función delta de Dirac se define como $\delta(x - a) = \{\infty \text{ si } x = a ; 0 \text{ si } x \neq a\}$, y además cumple con $\int f(x)\delta(x - a)dx = f(a)$.

$$\int p_n(x_1, t_1; \dots; x_n, t_n) dx_n = p_{n-1}(x_1, t_1; \dots; x_{n-1}, t_{n-1}).$$

- Medida de probabilidad (normalización), $\int p_1(x, t) dx = 1$.

Cada jerarquía de funciones de distribución que satisface las condiciones anteriores definen de manera única al proceso estocástico $\{X_t\}_{t \in I}$.

La condición de simetría está asociada al principio de balance detallado. Dicho principio sostiene que para un sistema el cual puede ser subdividido en procesos temporales elementales, en el equilibrio, cada proceso elemental puede ser equilibrado por su proceso inverso.

Un proceso estocástico es llamado estacionario si para todo n se cumple:

$$p_n(x_1, t_1 + \Delta t; x_2, t_2 + \Delta t; \dots; x_n, t_n + \Delta t) = p_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n). \quad (2.36)$$

Si se tienen X_1 y X_2 variables aleatorias la función de densidad de probabilidad condicional para X_1 dado X_2 es

$$p_{1|1}(x_2 | x_1) = \frac{p_2(x_1, x_2)}{p_1(x_1)}. \quad (2.37)$$

Aquí $p_2(x_1, x_2)$ es la función de distribución conjunta para X_1 y para X_2 ; $p_1(x_1)$ es la densidad de probabilidad marginal⁸ tal que

$$\begin{aligned} p_1(x_1) &= \int p_2(x_1, x_2) dx_2 \\ p_1(x_1) &= \int p_{1|1}(x_1 | x_2) p_1(x_2) dx_2. \end{aligned} \quad (2.38)$$

El concepto de probabilidad condicional puede ser usado para tener información acerca de *subensembles* del proceso estocástico en cuestión. Si suponemos que el proceso estocástico llega a un resultado en $(x_1, x_1 + dx_1)$ al tiempo t_1 , a $(x_2, x_2 + dx_2)$ en t_2 , ..., $(x_k, x_k + dx_k)$ en el tiempo t_k , ¿cuál es la probabilidad de encontrar un valor en $(x_{k+1}, x_{k+1} + dx_{k+1})$, al tiempo t_{k+1} , ..., $(x_{k+l}, x_{k+l} + dx_{k+l})$ al tiempo t_{k+l} ?

Para tener obtener dicha respuesta se define a la probabilidad condicional para los eventos en $k+1$ a $k+l$ dados los eventos de 1 a k como:

$$\begin{aligned} p_{k+l}(x_1, t_1; \dots; x_{k+l}, t_{k+l}) &= \\ p_k(x_1, t_k; \dots; x_k, t_k) p_{l|k}(x_{k+1}, t_{k+1}; \dots; x_{k+l}, t_{k+l} | x_1, t_k; \dots; x_k, t_k), \end{aligned} \quad (2.39)$$

de donde se deduce

$$p_{l|k}(x_{k+1}, t_{k+1}; \dots; x_{k+l}, t_{k+l} | x_1, t_k; \dots; x_k, t_k) = \frac{p_{k+l}(x_1, t_1; \dots; x_{k+l}, t_{k+l})}{p_k(x_1, t_k; \dots; x_k, t_k)}. \quad (2.40)$$

⁸Si la variable aleatoria es continua, en el caso discreto se habla de distribuciones marginales de probabilidad.

Definición 2: Para un proceso de Markov, se tiene que para todo n y todos $t_1 < t_2 < \dots < t_n$ se cumple [16]:

$$p_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = p_{1|1}(x_n, t_n; x_{n-1}, t_{n-1}). \quad (2.41)$$

Por lo tanto, sólo es necesario el conocimiento del actual estado del sistema (x_{n-1}, t_{n-1}) , con objeto de calcular la probabilidad de ocurrencia de (x_n, t_n) . Se dice que el proceso tiene memoria corta. Iterativamente encontramos

$$p_n(x_1, t_1; \dots; x_n, t_n) = \prod_{l=2}^n p_{1|1}(x_l, t_l | x_{l-1}, t_{l-1}) p_1(x_1, t_1), \quad (2.42)$$

por lo que el conocimiento de $p_{1|1}$ y p_1 es suficiente para describir de manera completa el proceso.

Definición 3: Para $X_{t_1}, X_{t_2}, X_{t_3}$, con $t_1 < t_2 < t_3$, la relación

$$p_{1|1}(x_3, t_3 | x_1, t_1) = \int p_{1|1}(x_3, t_3 | x_2, t_2) p_{1|1}(x_2, t_2 | x_1, t_1) dx_2, \quad (2.43)$$

se llama ecuación de Chapman-Kolmogorov.

2.2.2. Procesos de Markov Estacionarios y Ecuación Maestra

Para un proceso de Markov estacionario se tiene: $p_1(x, t) = p_{est}(x)$, la distribución de probabilidad de nuestra variable aleatoria es igual a una distribución dada. Además $p_{1|1}(x_2, t_2 | x_1, t_1) = p_t(x_2 | x_1)$; $t_2 - t_1 = t$. Aquí p_t denota la probabilidad de transición de x_1 a x_2 en el intervalo de tiempo t .

Usando la ecuación de Chapman-Kolmogorov para p_t tenemos[16]

$$p_{t+t'}(x_3 | x_1) = \int p_{t'}(x_3 | x_2) p_t(x_2 | x_1) dx_2. \quad (2.44)$$

Ahora derivemos la forma diferencial de la ecuación de Chapman- Kolmogorov para un proceso de Markov estacionario. Consideremos que los intervalos t' son muy pequeños, esto nos permite tomar a las probabilidades de transición y expandirlas en una serie de Taylor alrededor de cero

$$p_{t'}(x_3 | x_2) = (1 - w_{tot}(x_2)t')\delta(x_3 - x_2) + t'w(x_3 | x_2) + o(t'), \quad (2.45)$$

esta ecuación define $w(x_3 | x_2)$ como la tasa de transición (probabilidad de transición por unidad de tiempo) del estado x_2 al x_3 , con $(1 - w_{tot}(x_2)t')$ como la probabilidad de permanecer en x_2 por un tiempo t' tal que

$$w_{tot}(x_2) = \int w(x_3 | x_2) dx_3. \quad (2.46)$$

En (2.45) la función delta expresa que la probabilidad de estar en el mismo estado (x_2) después del tiempo t es $1 - w_{tot}(x_2)t'$. Aquí se ha tomado a $w(\cdot | \cdot)$ como la derivada de la probabilidad de

transición por unidad de tiempo.

Sustituyendo (2.45) en la ecuación de Chapman-Kolmogorov (2.44) tenemos

$$\begin{aligned} p_{t+t'}(x_3 | x_1) &= (1 - w_{tot}(x_3)t')p_t(x_3 | x_1) + t' \int w(x_3 | x_2)p_t(x_2 | x_1)dx_2 \\ \frac{p_{t+t'}(x_3 | x_1) - p_t(x_3 | x_1)}{t'} &= \int w(x_3 | x_2)p_t(x_2 | x_1)dx_2 - \int w(x_2 | x_3)p_t(x_3 | x_1)dx_2. \end{aligned} \quad (2.47)$$

En el límite $t' \rightarrow 0$ de (2.47) obtenemos la ecuación maestra:

$$\frac{\partial}{\partial t}p_t(x_3 | x_1) = \int w(x_3 | x_2)p_t(x_2 | x_1)dx_2 - \int w(x_2 | x_3)p_t(x_3 | x_1)dx_2. \quad (2.48)$$

Asumiendo que el proceso no es estacionario ($p_1(x_1, t) \neq p_{est}(x)$), pero manteniendo que existe homogeneidad en el tiempo para las probabilidades de transición (estas sólo dependen de diferencias en el tiempo). Si multiplicamos (2.48) por $p_1(x_1, t)$ e integrando sobre x_1 obtenemos la ecuación maestra para la densidad de probabilidad marginal o reducida $p_1(x_3, t)$. Renombrando variables obtenemos la ecuación maestra:

$$\begin{aligned} \frac{\partial}{\partial t}p_1(x_3, t) &= \int w(x_3 | x_2)p_1(x_2)dx_2 - \int w(x_2 | x_3)p_t(x_3)dx_2, \\ p_1 &\rightarrow p; \quad x_3 \rightarrow x; \quad x_2 \rightarrow x', \\ \frac{\partial}{\partial t}p(x, t) &= \int w(x | x')p(x', t)dx' - \int w(x' | x)p(x, t)dx'. \end{aligned} \quad (2.49)$$

La Ecuación Maestra (2.49) tiene una gran importancia en física, química y biología, que consite en aplicaciones en diversos ramos de estas áreas del saber como la mecánica cuántica, reacciones químicas y biología moelcular. En concreto la Ecuación Maestra nos permite estudiar en términos de probabilidades de transición los flujos que entran y salen a determinado sistema.

2.2.3. La expansión Kramers-Moyal y Ecuación de Fokker-Planck

La expansión de Kramers-Moyal es un método sistemático para derivar ecuaciones diferenciales parciales de evolución para la probabilidad condicional de interés. Habitualmente cuando los fenómenos descritos pueden asociarse con procesos de Markov, esta expansión da como resultado la ecuación de Fokker-Planck.

Llamemos $w(x | x') = w(x - r; r)$, con $r = x - x'$, supongamos que los saltos pequeños son más probables que los grandes. Aquí las tasas de transición por unidad de tiempo están en función del tamaño del salto dado de un estado a otro. Y asumiendo que w y p son funciones suaves en ambos argumentos, podemos escribir (2.49) como [16]

$$\frac{\partial}{\partial t}p(x, t) = \int w(x - r; r)p(x - r, t)dr - p(x, t) \int w(x; -r)dr. \quad (2.50)$$

Desde que se ha supuesto la preponderancia de saltos pequeños y que w y p varían suavemente en función de sus argumentos, es posible tratar con el paso de x a $x - r$. Tomando una expansión de Taylor en $x - r$ alrededor de $r = 0$ en la primera integral del lado derecho de (2.50), obtenemos:

$$\begin{aligned} \frac{\partial}{\partial t}p(x, t) &= p(x, t) \int w(x; r)dr - p(x, t) \int w(x; -r)dr - \\ &\int r \frac{\partial}{\partial x}(w(x; r)p(x, t))dr + \frac{1}{2} \int r^2 \frac{\partial^2}{\partial x^2}(w(x; r)p(x, t))dr \mp \dots \end{aligned} \quad (2.51)$$

De donde se sigue que

$$\frac{\partial}{\partial t} p(x, t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \{a_n(x) p(x, t)\}, \quad (2.52)$$

tras haber definido los momentos como

$$a_n(x) = \int r^n w(x; r) dr. \quad (2.53)$$

La ecuación (2.52) es la expansión Kramers-Moyal de la ecuación maestra.

La suposición de que $w(x, r)$ y $p(x, t)$ son lo suficientemente suaves permite truncar la expansión de Kramers-Moyal (2.52) a un cierto orden de derivadas. Un corte hasta el segundo orden lleva a la ecuación Fokker-Planck

$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} \{a_1(x) p(x, t)\} + \frac{1}{2} \frac{\partial^2}{\partial x^2} \{a_2(x) p(x, t)\}. \quad (2.54)$$

Aquí $a_1(x)$ es el coeficiente de deriva (primer momento de la distribución de probabilidad: la media) y $a_2(x)$ es el coeficiente de difusión (segundo momento de la distribución de probabilidad: la varianza).

La ecuación de Fokker-Planck (2.54) es importante en física y en muchas de sus ramas (física nuclear, biofísica, etc.) ya que permite obtener información de cómo evoluciona la densidad de probabilidad de la variable estocástica de interés.

2.2.4. Ecuación de Langevin

Teniendo en cuenta lo anterior, podemos caracterizar la dinámica de una partícula inmersa en un medio fluctuante, en concreto aplicar los procesos estocásticos para una partícula Browniana. En este caso el proceso estocástico estudiado es la velocidad de dicha partícula, denotado por $\{V_t\}_{t \in \mathbb{R}}$, aquí la variable aleatoria V_t toma un valor $v(t)$ para un determinado tiempo.

El modelo mecanicista que permite hacer una representación de su dinámica en tiempo continuo es la ecuación de Langevin [17]. Esta ecuación es del tipo diferencial estocástica, y modela la velocidad ($v(t)$) de una partícula (Browniana) microscópica inmersa en un medio fluido cuyo movimiento es aleatorio, dicho medio es visto a la vez como continuo y compuesto de moléculas. En el caso más común y tratado, la partícula se encuentra en presencia de una fuerza de fricción y una fuerza fluctuante, dicha ecuación se presenta a continuación.

Si en general el balance de fuerzas para un sistema físico sigue la 2da Ley de Newton: $ma = F$

$$m \frac{dv}{dt} = -\gamma v + \xi(t), \quad (2.55)$$

donde $\gamma > 0$ es un coeficiente de fricción el cual esta dado por la Ley de Stokes $\gamma = 6\pi\eta a$, donde η es la viscosidad dinámica, a es el radio de las partículas, y ξ es un ruido blanco⁹ asociado a la fuerza fluctuante tal que [17]:

$$\begin{aligned} \langle \xi(t) \rangle &= 0 \\ \langle \xi(t) \xi(t') \rangle &= 2B\delta(t - t'). \end{aligned} \quad (2.56)$$

⁹Un ruido blanco es un proceso estocástico para el cual su densidad espectral de potencia es constante.

Aquí B es una medida de la fuerza fluctuante y está relacionada con el coeficiente de fricción γ por el teorema de fluctuación disipación:

$$B = \gamma kT, \quad (2.57)$$

donde k es la constante de Boltzmann y T es la temperatura. Dicho teorema expresa un balance entre la fuerza de fricción la cual tiende a llevar al sistema a un punto de equilibrio estático y la fuerza fluctuante asociada al ruido la cual perturba al sistema.

De (2.56) se implica que la fuerza fluctuante tiene una distribución gaussiana; reescribiendo (2.55) y tomando $m = 1$ tenemos:

$$dv = -\gamma v dt + dW; v(0) = v_0 \quad (2.58)$$

En la que abusando de la notación se representa al ruido blanco como

$$\frac{dW}{dt} \approx \xi(t), \quad (2.59)$$

donde $W(t)$ es un movimiento Browniano. La solución de la ecuación diferencial presentada en (2.58) es¹⁰

$$\begin{aligned} v(t) &= v_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-s)} dW \\ v(t) &= v_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-s)} \xi(s) ds. \end{aligned} \quad (2.60)$$

La integral del segundo término del lado derecho para la primera expresión, puede tratarse conforme el Lema de Itô ó por la integral de Stratonovich, puesto que $W(t)$ es un movimiento Browniano¹¹ y la integral se realiza sobre el mismo.

Aun así, podemos mostrar que para tiempos muy grandes $t \rightarrow \infty$, la velocidad de la partícula Browniana se acerca a una distribución $N(0, \frac{B}{\gamma})$, por lo que la velocidad para tiempos muy grandes tiende a una distribución de Maxwell-Boltzmann. Esto se consigue tomando el promedio en la velocidad y el cuadrado de esta, como se observa a continuación

$$\langle v(t) \rangle = \langle v_0 \rangle e^{-\gamma t}. \quad (2.61)$$

¹⁰El cálculo de dicha solución se muestra en el Apéndice de este capítulo.

¹¹Y éste es no diferenciable en todo punto del proceso.

$$\begin{aligned}
\langle v(t)^2 \rangle &= \langle v_0^2 e^{-2\gamma t} + 2e^{-\gamma t} v_0 \int_0^t e^{-\gamma(t-s)} \xi(s) ds + \left(\int_0^t e^{-\gamma(t-s)} \xi(s) ds \right)^2 \rangle \\
&= e^{-2\gamma t} \langle v_0^2 \rangle + 2e^{-\gamma t} \langle v_0 \rangle \langle \int_0^t e^{-\gamma(t-s)} \xi(s) ds \rangle + \langle \int_0^t e^{-\gamma(t-s)} \xi(s) ds \int_0^t e^{-\gamma(t-s')} \xi(s') ds' \rangle \\
&= e^{-2\gamma t} \langle v_0 \rangle + \int_0^t e^{-\gamma(t-s)} ds \int_0^t e^{-\gamma(t-s')} 2B \delta(s-s') ds' \\
&= e^{-2\gamma t} \langle v_0 \rangle + 2B \int_0^t e^{-2\gamma(t-s)} ds \\
&= e^{-2\gamma t} \langle v_0 \rangle + \frac{B}{\gamma} (1 - e^{-2\gamma t}).
\end{aligned} \tag{2.62}$$

Tomando los límites para $\langle v(t) \rangle$ y para $\langle v(t)^2 \rangle$ cuando $t \rightarrow \infty$, tenemos

$$\langle v(t) \rangle \rightarrow 0; \quad \langle v(t)^2 \rangle \rightarrow \frac{B}{\gamma}, \tag{2.63}$$

por lo que la solución de la ecuación de Langevin para tiempos grandes tiende a una distribución Gaussiana con media 0 y varianza $\frac{B}{\gamma}$, la cual representa un balance entre la fuerza fluctuante ξ y la fuerza de amortiguamiento (de fricción) $-\gamma v$. En el equilibrio térmico, una condición para la velocidad cuadrática media es que cumpla con en el teorema de equipartición $\langle v(t)^2 \rangle_{eq} = kT$, puesto hemos tomado la masa unitaria. Al igualar la velocidad cuadrática media para tiempos grandes (2.63) con la asociada al teorema de equipartición obtenemos $B = \gamma kT$, corroborando el teorema de fluctuación disipación mencionado en (2.57).

2.2.5. Derivación de la ecuación Fokker-Planck a partir de la ecuación de Langevin

Ahora veamos la relación existente entre la ecuación de Langevin y la ecuación de Fokker-Planck[18]. Aquí consideramos que la variable aleatoria v sigue un proceso de Markov Gaussiano y la dinámica de ésta cumple con la ecuación de Langevin ¹²

$$\begin{aligned}
\frac{dv}{dt} &= a_1(v) + \xi(t), \\
\langle \xi(t) \rangle &= 0; \quad \langle \xi(t)\xi(t') \rangle = 2B\delta(t-t'),
\end{aligned} \tag{2.64}$$

donde $a_1(v)$ es el coeficiente de deriva y $\xi(t)$ sigue siendo un ruido blanco.

Si nos centramos en la distribución de probabilidad $p(v, t)$, tal que se le pida que ésta es una cantidad conservada, i.e., $\int p(v, t) dv = 1, \forall t$; por lo que se espera que se cumpla la ecuación de continuidad

$$\frac{\partial p(v, t)}{\partial t} + \frac{\partial}{\partial v} \cdot \left[\frac{\partial v}{\partial t} p(v, t) \right] = 0. \tag{2.65}$$

¹²De manera más general que (2.26).

Sustituyendo la derivada de v con respecto al tiempo por el lado derecho de la ecuación de Langevin, obtenemos

$$\frac{\partial p(v, t)}{\partial t} = -\frac{\partial}{\partial v} \cdot [a_1(v)p(v, t) + \xi(t)p(v, t)]. \quad (2.66)$$

Utilizando un operador del tipo Liouville¹³ definido (en este caso) como:

$$L\Phi \equiv \frac{\partial}{\partial v} \cdot (a_1(v)\Phi), \quad (2.67)$$

para cualquier función Φ y tomando en cuenta el problema sin ruido

$$\frac{\partial p(v, t)}{\partial t} = -Lp(v, t), \quad (2.68)$$

se obtiene la solución $p(v, t) = e^{-tL}p(v, 0)$.

Tomando en cuenta el ruido y sustituyendo la solución anterior obtenemos

$$\begin{aligned} \frac{\partial p(v, t)}{\partial t} &= -Lp(v, t) - \frac{\partial}{\partial v} [\xi(t)p(v, t)] \\ &= -Le^{-tL}p(v, 0) - \frac{\partial}{\partial v} [\xi(t)e^{-tL}p(v, 0)]. \end{aligned} \quad (2.69)$$

Integrando sobre el tiempo (2.69), tenemos

$$p(v, t) = e^{-tL}p(v, 0) - \int_0^t e^{-(t-s)L} \frac{\partial}{\partial v} \cdot [\xi(s)p(v, s)] ds, \quad (2.70)$$

que es la solución formal para el caso con ruido.

Aquí $p(v, t)$ depende del ruido $\xi(s)$ sólo para tiempos s anteriores a t . Sustituyendo (2.70) en el segundo término de la derecha de (2.69) obtenemos

$$\begin{aligned} \frac{\partial p(v, t)}{\partial t} &= -Lp(v, t) - \frac{\partial}{\partial v} [\xi(t)e^{-tL}p(v, 0)] \\ &+ \frac{\partial}{\partial v} \cdot \xi(t) \int_0^t e^{-(t-s)L} \frac{\partial}{\partial v} \cdot [\xi(s)p(v, s)] ds. \end{aligned} \quad (2.71)$$

Tomando el valor esperado en ambos lados, el segundo término de la derecha desaparece por (2.64)¹⁴ y reorganizando tenemos

$$\begin{aligned} \frac{\partial \langle p(v, t) \rangle}{\partial t} &= -L \langle p(v, t) \rangle \\ &+ \frac{\partial}{\partial v} \int_0^t \langle \xi(t)\xi(s) \rangle \langle e^{-(t-s)L} \rangle ds \frac{\partial}{\partial v} (\langle p(v, t) \rangle). \end{aligned} \quad (2.72)$$

¹³Puesto v no es una variable definida en el espacio fase.

¹⁴Y además desde que $\xi(t)$ y e^{-tL} son independientes $\langle \xi(t)e^{-tL} \rangle = \langle \xi(t) \rangle \langle e^{-tL} \rangle$.

Por las relaciones (2.34) y (2.32) tenemos que $\langle p(v, t) \rangle = \int dp p(v, t) \int \delta(v - V_t) p(V_t) dV_t = p(v, t)$, ya que $p(v, t)$ esta normalizada por lo que sustituyendo esta y L , al hacer

$$a_2(v) = \int_0^t \langle \xi(t) \xi(s) \rangle \langle e^{-(t-s)L} \rangle ds. \quad (2.73)$$

llegamos a la ecuación Fokker-Planck

$$\frac{\partial p(v, t)}{\partial t} = -\frac{\partial}{\partial v} \cdot a_1(v) p(v, t) + \frac{\partial}{\partial v} \cdot a_2(v) \cdot \frac{\partial}{\partial v} (p(v, t)). \quad (2.74)$$

Los términos $a_1(v)$ y $a_2(v)$ son llamados coeficientes de deriva y difusión respectivamente, para el caso de una partícula browniana como el discutido en (2.55), el coeficiente de deriva es $-\gamma v$ y el coeficiente de difusión es γkT para este caso.

2.2.6. Solución estacionaria para la ecuación Fokker-Planck

La solución de la ecuación (2.74) al resolverla nos muestra cómo evoluciona la densidad de probabilidad del proceso estocástico analizado. La accesibilidad para resolver dicha ecuación depende de la naturaleza de los coeficientes de deriva y difusión del proceso. En nuestro caso, como se verá más adelante, el de deriva es de carácter no lineal, por lo que existen dificultades para encontrar una solución analítica de (2.74). Es por esto que suele recurrirse a la hipótesis ergódica para encontrar una solución estacionaria. Como se discutió brevemente en la sección 1.3, el principal resultado de ésta, es que el sistema evoluciona por un tiempo suficientemente largo, tal que, visita todos los estados accesibles. Como consecuencia los promedios temporales son equivalentes a los espaciales. Esto permite encontrar una función de densidad de probabilidad independiente del tiempo para el proceso estudiado. A continuación se muestra la estructura de dicha función de densidad. Si definimos al flujo de probabilidad como [19]

$$j = \left[a_1(v) - a_2(v) \frac{\partial}{\partial v} \right] p(v, t), \quad (2.75)$$

podemos escribir la ecuación de Fokker-Planck (2.74) como

$$\frac{\partial p(v, t)}{\partial t} + \frac{\partial j}{\partial v} = 0. \quad (2.76)$$

En el caso estacionario $\frac{\partial p(v, t)}{\partial t} = 0$, por lo que simplemente queda resolver

$$\begin{aligned} \frac{\partial}{\partial v} \left\{ [a_1(v) - a_2(v) \frac{\partial}{\partial v}] p(v, t) \right\} &= 0 \\ \implies \frac{a_1(v)}{a_2(v)} dv &= \frac{dp(v, t)}{p(v, t)}. \end{aligned} \quad (2.77)$$

Al integrar y despejar $p(v, t)$ obtenemos la solución estacionaria de (2.74)

$$p(v, t) = C e^{\int \frac{a_1(v) dv}{a_2(v)}}, \quad (2.78)$$

donde C es una constante de normalización y deben conocerse $a_1(v)$ y $a_2(v)$ para determinar (2.78) de manera explícita.

2.2.7. Ecuación de Fokker-Planck para ecuación de Langevin con ruido multiplicativo de color

Por ser necesario para los fines de la tesis, consideramos ahora un caso todavía más general que los anteriores, tomemos una ecuación de Langevin general no-lineal [20]

$$\frac{dv(t)}{dt} = a(v(t), t) + g(v(t), t)\xi(t). \quad (2.79)$$

Aquí, $\xi(t)$ es tal que su media es cero y posee una función de correlación arbitraria $\langle \xi(t)\xi(t') \rangle = \Gamma(t, t')$, $\xi(t)$ es multiplicado por una función que depende de v y t por lo que se considera que éste producto induce un ruido de color¹⁵. Para cada realización del ruido se consideran colecciones de sistemas que obedecen (2.79). Si $p(v(t), t)$ es una cantidad conservada, ésta cumple con la ecuación de continuidad

$$\frac{\partial p(v(t), t)}{\partial t} = -\frac{\partial}{\partial v} \{ [a(v(t), t) + g(v(t), t)\xi(t)] p(v(t), t) \}. \quad (2.80)$$

La solución $v(t)$ para cada realización de $\xi(t)$ es una funcional de $\xi(t)$ y de $v(0)$, es decir, $v(t) = v(\xi(t), v(0), t)$, recordando la relación (2.32) tenemos

$$p(v, t) = \int \delta(V_t - v) p(v, t) dv = \langle \delta(V_t - v) \rangle. \quad (2.81)$$

Aplicando la ecuación de continuidad (2.80) a (2.81) obtenemos

$$\frac{\partial p(v, t)}{\partial t} = -\frac{\partial}{\partial v} [a(v, t)p(v, t)] - \frac{\partial}{\partial v} [g(v, t)\langle \xi(t)\delta(v(t) - v) \rangle]. \quad (2.82)$$

Puesto que $\delta(V(t) - v)$ es una funcional de $\xi(t)$, y el teorema de Novikov¹⁶ implica que se cumple:

$$\langle \xi(t)\delta(V_t - v) \rangle = \int_0^t \Gamma(t, t') \langle \frac{\delta[\delta(V_t - v)]}{\delta\xi(t')} \rangle dt', \quad (2.83)$$

entonces sustituyendo (2.83) en (2.82) obtenemos

$$\frac{\partial p(v, t)}{\partial t} = -\frac{\partial}{\partial v} [a(v, t)p(v, t)] + \frac{\partial}{\partial v} [g(v, t) \int_0^t \Gamma(t, t') \langle \frac{\delta[\delta(V_t - v)]}{\delta\xi(t')} \rangle dt']. \quad (2.84)$$

Sólo si no existen términos de orden superior al segundo, la ecuación (2.84) se puede considerar como una generalización de la ecuación Fokker-Planck para procesos con ruido de color tales que cumplen la ecuación de Langevin (2.79).

¹⁵Un ruido de color es un proceso estocástico para el cual su densidad espectral de potencia varía con la frecuencia.

¹⁶El teorema de Novikov [21] sostiene que para un proceso gaussiano $\xi(t)$ con una función de correlación arbitraria $\Gamma(t, t')$ y $R(\xi)$ una funcional arbitraria sobre $\xi(t)$ se cumple $\implies \langle \xi(t)R(\xi) \rangle = \int \Gamma(t, t') \langle \frac{\delta R(\xi)}{\delta\xi(t')} \rangle dt'$, aquí $\frac{\delta R(\xi)}{\delta\xi(t')}$ es la derivada variacional de $R(\xi)$ sobre $\xi(t')$ definida como: $\frac{\delta R(\xi)}{\delta\xi(t')} = \lim_{\Delta t' \rightarrow 0} \frac{R(\xi_1) - R(\xi)}{\int_{\Delta t'} h(t) dt}$, para la variación $\xi_1(t) = \xi(t) + h(t)$.

La naturaleza de la solución de (2.84) esta determinada por las funciones $a(\cdot)$, $g(\cdot)$ y $\Gamma(\cdot)$. Por lo que se deben encontrar expresiones analíticas para dichas funciones. Esto conlleva a aproximar por medio de algún método el término que contiene a la derivada variacional. Un caso de interés donde se lleva a cabo lo anterior es el siguiente:

La meta es encontrar una expresión para la integral de (2.84), desde que $\Gamma(t, t')$ es conocida sólo basta con dar una expresión para el término asociado al teorema de Novikov, el cual cumple con [6]

$$\frac{\delta[\delta(V_t - v)]}{\delta\xi(s)} = \left[-\frac{\partial}{\partial v}\delta(V_t - v)\right]\frac{\delta v}{\delta\xi}. \quad (2.85)$$

Al haber factorizado el término asociado al teorema de Novikov, para poder trabajar con éste dentro de la integral en (2.84), sólo basta expresar el término correspondiente a la derivada variacional $\frac{\delta v}{\delta\xi}$. Dicho término se puede aproximar como [6]

$$\frac{\delta v}{\delta\xi} = \Theta(t - t')\{g + [(\frac{dg}{dv})a - g(\frac{da}{dv})](t - t')\}, \quad (2.86)$$

donde $\Theta(\cdot)$ es la función de Heaviside¹⁷. En el caso particular de que la función de autocorrelación para el ruido sea de la forma (ver apartado 3.3):

$$\Gamma(t, t') = \langle \xi(t), \xi(t') \rangle = \left(\frac{D}{\tau_n}\right)e^{-\frac{|t-t'|}{\tau_n}}, \quad (2.87)$$

al sustituir (2.87), (2.86), (2.85) en (2.84) obtenemos

$$\begin{aligned} \frac{\partial p(v, t)}{\partial t} &= -\frac{\partial}{\partial v}[a(v, t)p(v, t)] + \\ &\frac{\partial}{\partial v}[g(v, t) \int_0^t \left(\frac{De^{-\frac{|t-t'|}{\tau_n}}}{\tau_n}\right) \left[-\frac{\partial}{\partial v}\delta(v - V_t)\right] [\Theta(t - t')\{g + [(\frac{dg}{dv})a - g(\frac{da}{dv})](t - t')\}] dt']. \end{aligned} \quad (2.88)$$

Donde al realizar la integral del segundo miembro se obtiene

$$\begin{aligned} \frac{\partial p(v, t)}{\partial t} &= -\frac{\partial}{\partial v}[a(v, t)p(v, t)] + \\ &D\frac{\partial}{\partial v}g(v)\frac{\partial}{\partial v}[g(v)h(v, t)p(v, t)], \end{aligned} \quad (2.89)$$

con

$$h(v, t) = [1 - e^{-\frac{t}{\tau_n}}] + \tau_n g(v) \frac{d}{dv} \left[\frac{a(v)}{g(v)} \right] \left\{ 1 - e^{-\frac{t}{\tau_n}} - \frac{t}{\tau_n} e^{-\frac{t}{\tau_n}} \right\}. \quad (2.90)$$

En particular, si $\tau_n \rightarrow 0$ se tiene

$$\begin{aligned} \frac{\partial p(v, t)}{\partial t} &= -\frac{\partial}{\partial v}[a(v, t)p(v, t)] + \\ &D\left\{\frac{\partial}{\partial v}g(v)\frac{\partial}{\partial v}g(v)[1 + \tau_n g(v)\frac{d}{dv}\left(\frac{a(v)}{g(v)}\right)]\right\}p(v, t). \end{aligned} \quad (2.91)$$

¹⁷ $\Theta(t - t') = 0$ si $t < t'$ y $\Theta(t - t') = 1$ si $t \geq t'$.

Para el caso estacionario se tiene $\frac{\partial p(v,t)}{\partial t} = 0$, por lo que al sustituir y reorganizar (2.91) tenemos

$$\frac{a(v)p_{st}}{Dg(v)} = \frac{\partial}{\partial v} \left\{ g(v) \left[1 + \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right] \right\} p_{st}. \quad (2.92)$$

La última ecuación se puede integrar separando variables haciendo el cambio de variable $u = \left\{ g(v) \left[1 + \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right] \right\} p_{st}$, por lo que se tiene ahora

$$\frac{a(v)dv}{Dg^2(v) \left[1 + \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right]} = \frac{du}{u}, \quad (2.93)$$

que al integrar se obtiene finalmente

$$p_{st} = \frac{C e^{\int^v \frac{a(v')dv'}{Dg^2(v') \left[1 + \tau_n g(v') \frac{d}{dv'} \left(\frac{a(v')}{g(v')} \right) \right]}}}{g(v) \left[1 + \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right]}, \quad (2.94)$$

donde C es una constante de normalización. La solución explícita de (2.94) depende de la naturaleza de la integral en la exponencial. En [6] se propone expresar el término $\left[1 + \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right]^{-1}$ como el primer término de una serie geométrica

$$\left[1 + \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right]^{-1} = 1 - \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right). \quad (2.95)$$

Sustituyendo (2.95) en (2.94) y tomando solo el primer término en la exponencial ya que τ_n es pequeño, obtenemos

$$p(v,t) = \frac{C \left[1 - \tau_n g(v) \frac{d}{dv} \left(\frac{a(v)}{g(v)} \right) \right]}{g(v)} e^{\int^v \frac{a(v')dv'}{Dg^2(v')}}. \quad (2.96)$$

Así, la solución estacionaria particular para un ruido de color con función de autocorrelación exponencial de la forma (2.87) y con tiempo de correlación pequeño, esta dada por la ecuación (2.96).

2.2.8. Ecuaciones de Langevin y Fokker-Planck para régimen sobre-amortiguado

Como mencionamos, la dinámica del movimiento Browniano para una partícula en un medio fluctuante queda determinada por la ecuación (2.64). Podemos expresar esta ecuación en términos de la posición de la partícula, la cual esta determinada por un proceso estocástico $\{X_t\}_{t \in \mathbb{R}}$, el cual toma valores iguales a $x(t)$ para los respectivos tiempos [6]. Esto se logra simplemente utilizando las definiciones de velocidad y aceleración

$$\begin{aligned} \frac{dx}{dt} &= v, \\ \frac{d^2x}{dt^2} &= \frac{dv}{dt}. \end{aligned} \quad (2.97)$$

Expresando el coeficiente de deriva presente en (2.64) como la suma de dos términos, el usual asociado a una fuerza de fricción y otro término asociado a otra fuerza debida a un campo conservativo U , se obtiene

$$a(v) = -\gamma \frac{dx}{dt} - \frac{dU}{dx}. \quad (2.98)$$

Al sustituir (2.98) y (2.97) en (2.64), obtenemos la ecuación de Langevin

$$\frac{d^2x}{dt^2} = -\gamma \frac{dx}{dt} - \frac{dU}{dx} + \xi(t). \quad (2.99)$$

En el régimen sobre-amortiguado $\gamma \rightarrow \infty$, si dividimos (2.99) entre γ y en el límite no se toma en cuenta el término inercial, obtenemos la ecuación de Langevin sobre-amortiguada

$$\frac{dx}{dt} = -\frac{1}{\gamma} \frac{dU}{dx} + \xi(t). \quad (2.100)$$

La estructura de la ecuación (2.100) es similar a (2.64), por lo que se pueden determinar las propiedades anteriormente discutidas. Las propiedades del ruido blanco las cuales son

$$\begin{aligned} \langle \xi(t) \rangle &= 0, \\ \langle \xi(t), \xi(t') \rangle &= 2 \frac{kT}{\gamma} \delta(t - t'), \end{aligned} \quad (2.101)$$

por lo que el correspondiente teorema de fluctuación disipación es

$$D = \frac{kT}{\gamma}. \quad (2.102)$$

Para este caso la correspondiente ecuación de Fokker-Planck es

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left[\left(-\frac{dU}{dx} \right) p(x, t) \right] + \frac{\partial}{\partial x} \frac{kT}{2\gamma} \frac{\partial}{\partial x} (p(x, t)). \quad (2.103)$$

Las ecuaciones (2.100) y (2.103) son las utilizadas en la descripción de la dinámica de plegamiento de horquillas de ARN, sólo que la variable de reacción descrita por éstas, es la extensión del biopolímero.

2.3. Teoría de Kramers

Como se mencionó en el apartado 1.3., uno de los principales resultados obtenidos de los experimentos y simulaciones del plegamiento de horquillas de ARN es la obtención de las tasas de salto entre los estados de plegado (NBA) y extendido (UBA). Para lograr esto se hace uso de la Teoría de Kramers, la cual se expone a continuación[16].

Se parte del supuesto de que el sistema posee dos estados de equilibrio, por lo que su energía potencial $U(x)$ tiene dos pozos, uno más profundo que el otro. Esta situación se muestra en la Figura 2.1 a continuación.

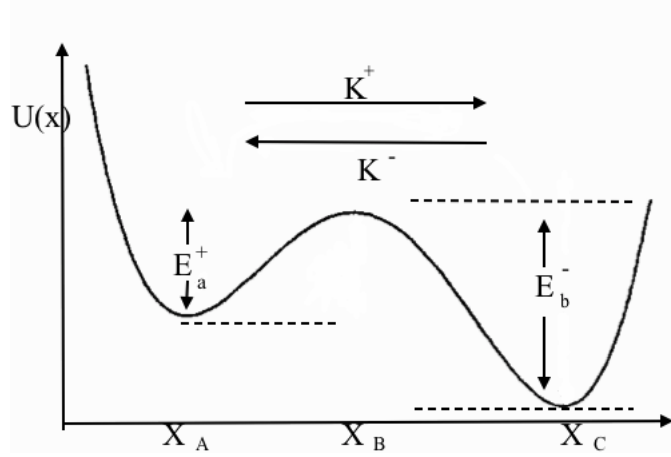


Figura 2.1: Energía potencial del sistema como función de una coordenada de reacción.

El estado asociado a x_C es el de energía más baja, mientras que el estado asociado a x_A es un mínimo local de $U(x)$. Kramers plantea esta situación en concreto sobre las reacciones químicas de ciertos productos. El estado A corresponde a los reactivos y el estado C corresponde a los productos. Es por esto que se llama a x coordenada de reacción. Para que la reacción ocurra, el sistema debe de cruzar la barrera de energía en x_B con diferencias de energía E_a^+ en comparación con el estado metaestable A y E_b^- en comparación con el estado estable C . Llamamos K^+ la tasa en que el sistema cruza de A a C y K^- a la tasa en que el sistema cruza la barrera de C a A . En el caso expuesto tenemos que $E_b^- > E_a^+$, por lo que se espera que K^- se más pequeña que K^+ . Aquí se plantea que la dinámica del producto es como la de una partícula Browniana modelada por la ecuación de Langevin (2.100) y su respectiva ecuación de Fokker-Planck de la forma (2.103).

Dada esta estructura se pretende encontrar el tiempo promedio que necesita la partícula Browniana para llegar de A a C por vez primera.

En concreto se trata de un proceso de difusión en presencia de una barrera energética, en el cual la partícula cumple con las condiciones de frontera

- Frontera absorbente en $x = x_C$

$$p(x_C, t) = 0, \quad (2.104)$$

- Frontera reflejante para $x \rightarrow -\infty$

$$\frac{\partial}{\partial x} p(x, t) \rightarrow 0. \quad (2.105)$$

Así la posición de la partícula comienza en $x \in (-\infty, x_C)$, para $t = 0$. Por lo que la probabilidad de que dicha partícula se mantenga en este intervalo para un tiempo t se expresa

$$G(x, t) = \int_{-\infty}^{x_C} p(x', t | x, 0) dx'. \quad (2.106)$$

Aplicando la ecuación de Fokker-Planck (2.103) a las densidades de (2.106), integrando y considerando que las probabilidades de transición son homogéneas

$$p(x', t | x, 0) = p(x', 0 | x, -y), \quad (2.107)$$

la ecuación de Fokker-Planck para este caso es

$$\frac{\partial}{\partial t} G(x, t) = a_1(x) \frac{\partial}{\partial x} G(x, t) + \frac{1}{2} a_2(x) \frac{\partial^2}{\partial x^2} G(x, t), \quad (2.108)$$

donde $a_1(x)$ y $a_2(x)$ son los respectivos coeficientes de deriva y difusión. $G(x, t)$ es el complemento de una distribución de probabilidad acumulada para los tiempos de primer paso, para un punto inicial x . La diferencia $G(x, t) - G(x, t + dt)$ es el número de puntos iniciales que no han dejado el punto inicial antes del tiempo t , pero que lo han hecho durante dt . Por lo que dicha diferencia determina la distribución de los tiempos de primer paso para su respectiva función de densidad llamada $f(x, t)$, como

$$\begin{aligned} G(x, t) - G(x, t + dt) &= f(x, t)dt, \\ f(t, x) &= -\frac{dG(x, t)}{dt}. \end{aligned} \quad (2.109)$$

Dado esto podemos escribir el tiempo medio de primer paso desde x hasta x_C , como

$$\tau_{x \rightarrow C} = -\int_0^\infty t \frac{\partial}{\partial t} G(x, t) dt, \quad (2.110)$$

por lo que al integrar (2.108) con respecto a t y usar (2.110) obtenemos

$$-1 = a_1(x) \frac{\partial}{\partial x} \tau_{x \rightarrow C} + \frac{1}{2} a_2(x) \frac{\partial^2}{\partial x^2} \tau_{x \rightarrow C}. \quad (2.111)$$

Usando las condiciones de frontera $\tau_{x \rightarrow C}(x_C) = 0$, $\lim_{x \rightarrow -\infty} \frac{\partial}{\partial x} \tau_{x \rightarrow C} = 0$. La solución de la parte homogénea de (2.111) es

$$\frac{\partial}{\partial x} \tau_{x \rightarrow C}^h = \tau_0 e^{-\int \frac{2a_1(x')}{a_2(x')} dx'}. \quad (2.112)$$

Sustituyendo (2.112) en (2.111) obtenemos

$$\frac{\partial}{\partial x} \tau_0 e^{-\int \frac{2a_1(x')}{a_2(x')} dx'} = \frac{-2}{a_2(x)}. \quad (2.113)$$

Integrando se tiene

$$\tau_0(x) = -2 \int \frac{dx'}{a_2(x')} e^{-\int \frac{2a_1(x'')}{a_2(x'')} dx''}. \quad (2.114)$$

A partir de (2.114) y (2.112), la solución de (2.111) se puede escribir como

$$\frac{\partial}{\partial x} \tau_{x \rightarrow C} = \tau_0(x) e^{-\int \frac{2a_1(x')}{a_2(x')} dx'} - \left[2 \int \frac{dx'}{a_2(x')} e^{-\int \frac{2a_1(x'')}{a_2(x'')} dx''} \right] \left[e^{-\int \frac{2a_1(x')}{a_2(x')} dx'} \right]. \quad (2.115)$$

De las condiciones de frontera (2.104) y (2.105), la parte asociada a la solución homogénea desaparece teniendo

$$\tau_{x \rightarrow C} = 2 \int_x^{x_C} \frac{dx'}{\phi(x')} \int_{-\infty}^{x'} \frac{dx''}{a_2(x'')} \phi(x''), \quad (2.116)$$

donde $\phi(x) = e^{-\int \frac{2a_1(x')}{a_2(x')} dx'}$.

Ahora en el problema de Kramers los coeficientes de deriva y difusión son

$$\begin{aligned} a_1(x) &= -\frac{1}{\gamma} \frac{dU}{dx}, \\ a_2(x) &= \frac{2kT}{\gamma}. \end{aligned} \quad (2.117)$$

Puesto que $kT \ll U(x_B)$, la exponencial $e^{\frac{U}{kT}}$ tiene un pico alrededor de x_B por lo que en las respectivas integrales solo necesitamos considerar para $x' \approx x_B$, sustituyendo dichos coeficientes en (2.116) tenemos

$$\tau_{x \rightarrow C} = \frac{\gamma}{kT} \int_x^{x_C} dx' e^{\frac{U(x')}{kT}} \int_{-\infty}^{x_B} dx'' e^{-\frac{U(x'')}{kT}}. \quad (2.118)$$

En la primera integral de (2.118) expandimos el potencial alrededor de $x = x_B$ y en $x = x_A$ para la segunda, como

$$\begin{aligned} U(x') &= U(x_B) - \frac{1}{2} \omega_B^2 (x' - x_B)^2, \\ U(x'') &= U(x_A) - \frac{1}{2} \omega_A^2 (x'' - x_A)^2. \end{aligned} \quad (2.119)$$

Donde $\omega_B^2 = \frac{d^2U}{dx^2} |_{x_B}$ y $\omega_A^2 = \frac{d^2U}{dx^2} |_{x_A}$. Para ser consistentes con las expansiones a segundo orden, el límite de integración en ambas integrales es reemplazado por $+\infty$, y el límite inferior por $-\infty$ para la primera, obteniendo

$$\tau_{x \rightarrow C} \approx \frac{\gamma}{kT} \int_{-\infty}^{+\infty} dx' e^{[\frac{U(x_B) - \frac{1}{2} \omega_B^2 (x' - x_B)^2}{kT}]} \int_{-\infty}^{+\infty} dx'' e^{-[\frac{U(x_A) - \frac{1}{2} \omega_A^2 (x'' - x_A)^2}{kT}]}. \quad (2.120)$$

Estas integrales son gaussianas y al resolverlas tenemos para la transición de A a C

$$\tau_{A \rightarrow C} \approx \frac{\gamma}{kT} \left[e^{\frac{U_{x_B}}{kT}} \sqrt{\frac{2\pi kT}{\omega_B^2}} \right] \left[e^{-\frac{U_{x_A}}{kT}} \sqrt{\frac{2\pi kT}{\omega_A^2}} \right], \quad (2.121)$$

que reorganizando nos da el tiempo medio de primer paso el cual es

$$\begin{aligned} \tau_{A \rightarrow C} &\approx \frac{2\pi\gamma}{\omega_B\omega_A} e^{[\frac{U_{x_B} - U_{x_A}}{kT}]}, \\ \tau_{A \rightarrow C} &\approx \frac{2\pi\gamma}{\omega_B\omega_A} e^{\frac{E_a^+}{kT}}. \end{aligned} \quad (2.122)$$

Aquí $E_a^+ = U_{x_B} - U_{x_A}$, y la tasa de transición del estado A al C es

$$K_{A \rightarrow C}^+ = \frac{1}{\tau_{A \rightarrow C}} = \frac{\omega_B\omega_A}{2\pi\gamma} e^{-\frac{E_a^+}{kT}}. \quad (2.123)$$

De manera similar se obtienen para la transición del estado C al A, el tiempo medio de primer paso y su respectiva tasa de transición, las cuales son:

$$\tau_{C \rightarrow A} \approx \frac{2\pi\gamma}{\omega_B\omega_C} e^{\frac{E_b^-}{kT}}, \quad (2.124)$$

$$K_{C \rightarrow A}^- = \frac{1}{\tau_{C \rightarrow A}} = \frac{\omega_B\omega_C}{2\pi\gamma} e^{-\frac{E_b^-}{kT}}. \quad (2.125)$$

Con $E_b^- = U_{x_B} - U_{x_C}$. En el caso de la dinámica de plegamiento de las horquillas de ARN las ecuaciones (2.122), (2.123) y (2.124), (2.125) representan respectivamente a los tiempos medios de primer paso y las tasas de transición del estado plegado (NBA) al extendido (UBA) en el primer caso y los tiempos y tasas para la transición de UBA a NBA en el segundo.

2.4. Simulación Estocástica

En seguida se presenta el método utilizado para encontrar una solución aproximada para la ecuación de Langevin (3.16) p. 46, dicho esquema nos permitirá obtener series de tiempo para las coordenadas de reacción (z_{sys}) y (z_m), las cuales son necesarias para encontrar las respectivas producciones de entropía y densidades de probabilidad.

Como se mencionó en el apartado anterior, las soluciones para la ecuación de Langevin asociada a nuestro sistema Asa-ARN-Asa, se simularon a partir de la aproximación de Euler-Maruyama. En dicha aproximación una ecuación diferencial estocástica de la forma[22]

$$\frac{dX(t)}{dt} = a[X(t), t] + \sigma[X(t), t]\xi(t), \quad (2.126)$$

tal que $\xi(t) \simeq \frac{dW}{dt}$ ¹⁸, con W movimiento Browniano, puede aproximarse como

$$X_{i+1} = X_i + h_i a(X_i) + (W_{t_{n+1}} - W_{t_n})\sigma(X_i). \quad (2.127)$$

Para nuestro caso, a partir de (3.16) se despeja el respectivo coeficiente friccional de Stokes $\gamma = 6\pi\eta a$, por lo que nuestro coeficiente de deriva asociado es

$$a[Z(t)] = \frac{1}{\gamma}\tau_{ef}(Z), \quad (2.128)$$

considerando para cada caso los respectivos coeficientes de la tensión (b_i con $i = 0, 1, 2, 3$), en este caso la aproximación de Euler-Maruyama (2.127) es

$$\begin{aligned} z_{i+1} &= z_i + \frac{1}{\gamma} \{h_i \tau_{ef}(z_i) + (W_{t_{i+1}} - W_{t_i})\} \\ &= z_i + \frac{1}{\gamma} \{h_i (-b_3 z_i^3 + b_2 z_i^2 - b_1 z_i + b_0) + (W_{t_{i+1}} - W_{t_i})\}, \end{aligned} \quad (2.129)$$

donde $h_i = t_{i+1} - t_i$ es el tamaño de paso y $(W_{t_{i+1}} - W_{t_i})$ son los incrementos del movimiento Browniano. Por hipótesis, cada incremento se distribuye de manera normal con media cero y varianza

¹⁸Un ruido blanco puede aproximarse como el cambio en el tiempo de un movimiento Browniano.

$h_i = t_{i+1} - t_i$, i.e.¹⁹

$$(W_{t_{i+1}} - W_{t_i}) \sim N(0, h_i) \sim \sqrt{h_i}N(0, 1), \quad (2.130)$$

por lo que cada incremento puede ser simulado al tomar un número de una distribución normal (estándar) y multiplicarlo por su desviación.

Del apartado 3.1.2 conocemos los valores de $\frac{1}{\gamma} = \frac{1}{6\pi(1 \times 10^{-6}) \frac{pNs}{nm}} = 53051,64 \frac{nm}{pNs}$, ya que $\gamma = 6\pi\eta a$ y $a = 1\mu m$, $\eta \approx 1cP$.

Para todas la simulaciones (en cada caso) se han sustituido los valores de los coeficientes b_i dados en las Tablas 3-14 . Las respectivas gráficas de dichas simulaciones se presentan en las Figuras 3.2-3.12. La simulación de la solución de la ecuación (3.16) por medio del esquema Euler-Maruyama se realizó en Octave/Matlab y en C. Para cada simulación se utilizó un cambio de variable $t = \gamma t^*$, obteniendo la ecuación

$$\frac{dz}{dt^*} = -b_3 z^3 + b_2 z^2 - b_1 z + b_0 + \xi(t^*), \quad (2.131)$$

con un intervalo de tiempo de $t^* \in [0, 2500]$, y un número de pasos $N = 10000$. Los últimos están relacionados con el tamaño de paso como $h = \frac{T}{N}$. Para regresar al tiempo del sistema físico se divide entre γ .

La respectiva aproximación de Euler-Maruyama usando el cambio de variable $t = \gamma t^*$ es

$$z_{i+1} = z_i + h_i(-b_3 z_i^3 + b_2 z_i^2 - b_1 z_i + b_0) + (W_{t_{i+1}^*} - W_{t_i^*}). \quad (2.132)$$

En el Apéndice de este trabajo se muestra el algoritmo utilizado para un caso de los 11 procesos de la ecuación de Langevin con ruido blanco (2.132), ya que todos son iguales en estructura y solo cambian los parámetros asociados a los coeficientes de la tensión para cada caso.

El mismo esquema fue utilizado para simular el proceso con ruido de color representado por la ecuación (3.34) p.62. De hecho se utiliza la aproximación de Euler-Maruyama en dos ocasiones, primero en la ecuación (3.33), para generar el ruido asociado a las cabezas de poliestireno z_b el cual se simula por

$$z_{b_{i+1}} = z_{b_i} - h_i \kappa_b z_{b_i} + (W_{t_{i+1}^*} - W_{t_i^*}), \quad (2.133)$$

después se vuelve a usar la aproximación de Euler -Maruyama pero ahora sobre (3.34) para generar z por medio de

$$z_{i+1} = z_i + h_i(-b_3 z_i^3 + b_2 z_i^2 - b_1 z_i + b_0) + \kappa_b(z_{b_{t_{i+1}^*}} - z_{b_{t_i^*}}), \quad (2.134)$$

donde se utiliza la serie de tiempo de z_b obtenida anteriormente. De igual forma se usan en ambos $t = \gamma t^*$. El respectivo código en OCTAVE/MATLAB y en C, para simular la ecuación de Langevin con ruido de color (2.134), se presenta en el Apéndice.

Por último en esta sección mostramos el método para simular la entropía producida en la dinámica de plegamiento de las horquillas de ARN, la cual esta representada por la ecuación (3.12) p.43 de la sección 3.1.1. Una vez simulada la serie de tiempo de z dada por (2.132) o por (2.134), se

¹⁹La siguiente expresión se desprende del proceso de estandarización de variables aleatorias normales, en el que una variable aleatoria que se distribuye $X \sim N(\mu, \sigma^2)$, puede ser expresada por una variable aleatoria normal estándar $Z \sim N(0, 1)$, por medio de la relación $Z = \frac{X-\mu}{\sigma}$, donde μ y σ^2 son la media y la varianza de X .

escogen los estados en los cuales se produce entropía, es decir se excluyen los estados de equilibrio NBA y UBA. Dado ésto, simplemente se sustituyen los valores de la nueva serie $z(t_i)$ en la ecuación (3.12). El respectivo código se muestra en el Apéndice.

2.5. Apéndice: Deducción de la solución para la ecuación de Langevin

Reescribiendo la ecuación de Langevin (2.58) [17]

$$\xi = \frac{dv}{dt} + \gamma v \simeq \frac{dW}{dt}, \quad (2.135)$$

e integrando²⁰ de t_0 a t

$$W(t) - W(t_0) = (v(t) - v(t_0)) + \gamma \int_{t_0}^t v(s) ds, \quad (2.136)$$

la ecuación anterior esta bien definida, i.e., $W(t)$ es continuo. Reescribiendo (2.136) obtenemos

$$v(t) + \gamma \int_{t_0}^t v(s) ds = v(t_0) - W(t_0) + W(t), \quad (2.137)$$

multiplicando ambos lados por un factor integrante $e^{\gamma t}$

$$\begin{aligned} e^{\gamma t} v(t) + \gamma e^{\gamma t} \int_{t_0}^t v(s) ds &= (v(t_0) - W(t_0)) e^{\gamma t} + W(t) e^{\gamma t} \\ \frac{d}{dt} (e^{\gamma t} \int_{t_0}^t v(s) ds) &= (v(t_0) - W(t_0)) e^{\gamma t} + W(t) e^{\gamma t}. \end{aligned} \quad (2.138)$$

Integrando ambos lados de t a t_0 y despejando $e^{\gamma t}$

$$\int_{t_0}^t v(s) ds = \left(\frac{v(t_0) - W(t_0)}{\gamma} \right) (1 - e^{\gamma(t_0-t)}) + \int_{t_0}^t e^{\gamma(s-t)} W(s) ds. \quad (2.139)$$

Diferenciando ahora con respecto a t

$$v(t) = (v(t_0) - W(t_0)) e^{\gamma(t_0-t)} + W(t) - \gamma \int_{t_0}^t e^{\gamma(s-t)} W(s) ds. \quad (2.140)$$

Recordando que para un movimiento Browniano $W(t)$ y una función C^1 , en el intervalo cerrado la integral de f sobre dW se aproxima por la expresión

$$\int_a^b f(t) dW \approx f(b)W(b) - f(a)W(a) - \int_a^b f'(t)W(t) dt, \quad (2.141)$$

²⁰Concretamente esta no es una integración en el sentido estricto, ya que $W(\cdot)$ es un movimiento Browniano el cual no es diferenciable, lo único que se hace aquí es dar una representación para dicho movimiento.

utilizando este resultado para $f(s) = e^{\gamma(s-t)}$, podemos sustituir el último término de la expresión anterior, lo que conduce al resultado

$$\begin{aligned} v(t) = & v(t_0)e^{-\gamma(t-t_0)} - W(t_0)e^{-\gamma(t-t_0)} + W(t) - \\ & W(t) + W(t_0)e^{-\gamma(t-t_0)} + \int_{t_0}^t e^{-\gamma(t-s)} dW(s), \end{aligned} \quad (2.142)$$

y eliminando términos semejantes obtenemos

$$v(t) = v(t_0)e^{-\gamma(t-t_0)} + \int_{t_0}^t e^{-\gamma(t-s)} dW(s). \quad (2.143)$$

CAPÍTULO 3

Desarrollo y Resultados

En este apartado se emplean las herramientas mencionadas previamente en la formulación de un modelo que permita describir cualitativa y cuantitativamente la dinámica de plegamiento de moléculas de ARN sujetas a fuerzas externas. El modelo cinético se construye partiendo de consideraciones pertenecientes a la termodinámica irreversible y se generaliza teniendo en cuenta el efecto de las fluctuaciones o el ruido estocástico presente en el sistema.

Se presentan dos aproximaciones distintas para modelar el sistema, la primera consiste en atribuir el ruido a las fluctuaciones térmicas del sistema, siendo el ruido blanco. La segunda consiste en atribuir las fluctuaciones a la dinámica intrínseca del sistema, por lo que en este caso se presenta un modelo con ruido de color.

En ambos casos el modelo consiste en las ecuaciones de Langevin y Fokker-Planck, asociadas a sus respectivos ruidos. Para cada caso se realizan simulaciones de las respectivas ecuaciones de Langevin y de la entropía producida. Así también se muestran soluciones aproximadas para las ecuaciones de Fokker-Planck correspondientes.

3.1. Formulación del modelo

En la sección 1.3. se discutió una forma alternativa para describir la dinámica de plegamiento para la horquilla de ARN. A partir del paisaje energético presentado en [9], es posible hacer inferencias sobre la evolución de la dinámica en cuestión. Con sólo la información del perfil de energía, es posible utilizar las ecuaciones de evolución para las extensiones de la horquilla de ARN (ecuación de Langevin) y su respectiva ecuación de evolución para la densidad de probabilidad pertinente (ecuación de Fokker-Planck), para obtener información sobre la dinámica de la horquilla de ARN. A continuación se presenta dicho modelo.

Partiendo del hecho de que la ecuación de Langevin asocia fuerzas a la dinámica del sistema y que éstas son causa de los flujos difusivos existentes dentro del mismo, podemos decir que dichas fuerzas están relacionadas con la producción de entropía del sistema estudiado, por lo que siguiendo el marco de la termodinámica irreversible lineal, es posible calcular la producción de entropía de la dinámica de plegamiento de las horquillas de ARN.

3.1.1. Termodinámica irreversible del modelo

En la sección 2.1. hemos partido de la hipótesis de que para determinado sistema el cambio en la entropía del sistema se puede escribir como (2.29). En concreto, nuestro sistema consiste en una horquilla de ARN sujeta a una fuerza externa τ y un medio fluctuante asociado a un baño térmico con temperatura T . Aplicando los métodos descritos en 2.1. se pueden determinar la entropía producida en la dinámica de plegamiento para dichos bio-polímeros, así como las ecuaciones cinéticas que determinan la evolución temporal de la extensión de la horquilla de ARN. Integrando (2.29) sobre un intervalo finito de tiempo, el cambio total de la entropía del sistema es

$$\Delta S = \Delta_e S + \Delta_i S. \quad (3.1)$$

Desde que hemos supuesto que el sistema esta inmerso en un baño térmico y nuestro sistema es cerrado¹, podemos usar la formulación de Helmholtz para la energía libre del sistema $F = F(T, V, N)$, en ésta, la energía libre esta dada por

$$F = U - TS. \quad (3.2)$$

Tomando el cambio en la energía libre y recordando que el cambio en la energía interna del sistema es $\Delta U = \Delta Q + dW$, tenemos

$$\Delta F = \Delta Q + \Delta W - T\Delta S. \quad (3.3)$$

Si el trabajo realizado sobre nuestro sistema es $\Delta W = \tau\Delta z$ y, además, como mencionamos nuestro sistema es cerrado y se encuentra a temperatura y volumen constante. Por lo anterior y por (3.1), la ecuación (3.3) es ahora

$$\Delta F = \Delta Q + \tau\Delta z - T\Delta_e S - T\Delta_i S. \quad (3.4)$$

Desde que el sistema tratado es cerrado se cumple $\Delta Q = T\Delta_e S$, por lo que al sustituir este resultado en (3.4) y despejar obtenemos [23]

$$T\Delta_i S = -\Delta F + \tau\Delta z. \quad (3.5)$$

La ecuación (3.5) relaciona el cambio de entropía con las demás variables extensivas e intensivas del sistema y lo hace para el caso particular del cambio de entropía asociado a los procesos irreversibles de la dinámica de plegamiento de la horquilla de ARN ($\Delta_i S$). Esta ecuación es un caso particular de la ecuación de Gibbs (2.14) presentada en la sección 2.1.

Al dividir (3.5) por Δt y tomar el límite para $\Delta t \rightarrow 0$, obtenemos la producción de entropía por unidad de tiempo

$$T \frac{d_i S}{dt} = \left[-\frac{\partial F}{\partial z} + \tau \right] \frac{dz}{dt}. \quad (3.6)$$

Como en (2.21), ésta es el producto de una fuerza generalizada $[-\frac{\partial F}{\partial z} + \tau]$ por su flujo conjugado $\frac{dz}{dt}$, y este flujo cumple con la relación (2.26) de la forma

$$\frac{dz}{dt} = \frac{1}{\gamma} \left[-\frac{\partial F}{\partial z} + \tau \right], \quad (3.7)$$

donde $\frac{1}{\gamma} > 0$ es el respectivo coeficiente de Onsager.

¹Ensemble canónico.

Entropía producida

De la Figura 1.5, p.9, presentada en la sección 1.3, observamos que la extensión contra el tiempo mostrada en experimentos y simulaciones es tal que su dinámica se caracteriza porque el valor de la extensión permanece cierto intervalo de tiempo en un estado (plegado o extendido) para después saltar al otro estado de equilibrio en donde permanece por otro intervalo de tiempo, por lo que, si suponemos cierta periodicidad en estas series de tiempo ², una forma matemática simple de modelar dicha dinámica de transiciones entre los estados de plegado y extendido es a través de [23]

$$z(t) \simeq \sum_i^{imax} \Delta z \Theta(t - t_i). \quad (3.8)$$

Donde Δz es la distancia entre los estados plegado y extendido, $\Theta(\cdot)$ es la función de Heaviside. La expresión (3.8) reproduce la serie de saltos entre estados para la extensión.

Llamando a la fuerza generalizada como una tensión efectiva³ $\tau_{ef} = -\frac{\partial F_{ef}}{\partial z} = -\frac{\partial F}{\partial z} + \tau$ y considerando a la velocidad como la derivada de (3.8) con respecto al tiempo tenemos

$$\frac{dz}{dt} = \Delta z \sum_{i=0}^{imax} \delta(t - t_i). \quad (3.9)$$

Al sustituir en (3.6) se tiene

$$T \frac{d_i S}{dt} = -\frac{\partial F_{ef}}{\partial z} \left[\Delta z \sum_{i=0}^{imax} \delta(t - t_i) \right], \quad (3.10)$$

e integrando ambos lados con respecto al tiempo

$$T \Delta_i S(t) = \Delta z \sum_{i=0}^{imax} -\frac{\partial F_{ef}}{\partial z}(t_i), \quad (3.11)$$

donde $-\frac{\partial F_{ef}}{\partial z}$ es la tensión efectiva τ_{ef} . Por lo tanto, el cambio en la entropía debido a los procesos irreversibles en la dinámica de plegamiento del ARN es

$$\Delta_i S(t) = \frac{\Delta z}{T} \sum_{i=0}^{imax} \tau_{ef}(t_i). \quad (3.12)$$

La tensión $\tau_{ef}(t_i)$ es aquella evaluada sólo en los tiempos i -ésimos para los cuales se considera que el proceso genera entropía, es decir, los saltos entre los estados plegado y extendido excluyendo las fluctuaciones térmicas dentro de cada pozo de energía libre. Esto se logra evaluando la tensión con z_{t_i} que es la solución de la ecuación de Langevin (3.16) o (3.34), tal que cumple dicha condición. En el apartado de resultados se muestra la simulación correspondiente para la entropía producida descrita por (3.12) y en el Apéndice el código en OCTAVE/MATLAB y en C utilizado.

²En aras de simplificar las operaciones a realizar.

³Por tensión efectiva entendemos al balance de fuerzas no aleatorias que actúan sobre el sistema Asa-ARN-Asa. Aquí τ es la tensión constante aplicada por las LOT sobre la esfera de poliestireno y $-\frac{\partial F}{\partial z}$ es la tensión interna que aplica la horquilla de ARN sobre la esfera de poliestireno.

3.1.2. Ecuación de Langevin con ruido aditivo blanco

El análisis de termodinámica irreversible realizado, nos permitió reconocer las fuerzas y flujos que intervienen en el proceso de plegamiento de la horquilla de ARN. En este caso el flujo encontrado es el cambio de la extensión con respecto al tiempo dado por (3.7). Esta ecuación no contempla el efecto de las fluctuaciones en el sistema, por lo que se considera necesaria la implementación de procesos estocásticos para completar esta descripción. Esto se consigue modelando las fluctuaciones del sistema de diversas formas, la más sencilla es por medio del método de Onsager, el cual consiste en sumar un ruido blanco a (3.7). Claro está que dicho método es equivalente a plantear que la dinámica de la horquilla de ARN satisface la ecuación de Langevin

$$\begin{aligned} \gamma \frac{dz}{dt} &= -\frac{\partial F_{ef}}{\partial z} + \xi(t) \\ &= \tau_{ef} + \xi(t). \end{aligned} \quad (3.13)$$

Aquí γ es el coeficiente de fricción de Stokes, F_{ef} es la energía libre del sistema y τ_{ef} es la tensión efectiva aplicada al sistema. El ruido blanco está representado por $\xi(t)$ (ver ecuación (2.101)). Para obtener información sobre la dinámica de la extensión de la horquilla de ARN durante el proceso de plegamiento, es preciso simular por un método de aproximación adecuado la ecuación de Langevin (3.13). Por lo que son necesarias expresiones analíticas para la energía libre y la tensión efectivas. La obtención de dichas expresiones se explica a continuación.

En la Figura 3.1 se muestran los perfiles de energía obtenidos en las simulaciones SOP P5GA para un modelo con asas para los casos de $F_{eq}(z_{sys})$ (línea azul discontinua) y $F_{eq}(z_m)$ (línea roja continua), y para un modelo sin asas para $F_{eq}^o(z_m)$ (línea negra continua), dichos perfiles son extraídos de [9].

Para nuestro análisis se utilizaron los casos que se muestran en las tablas 1 y 2. Para sistemas con asas rígidas con cociente $\frac{L}{l_p} < 1$ y para sistemas con asas flexibles tal que $\frac{L}{l_p} > 1$, respectivamente.

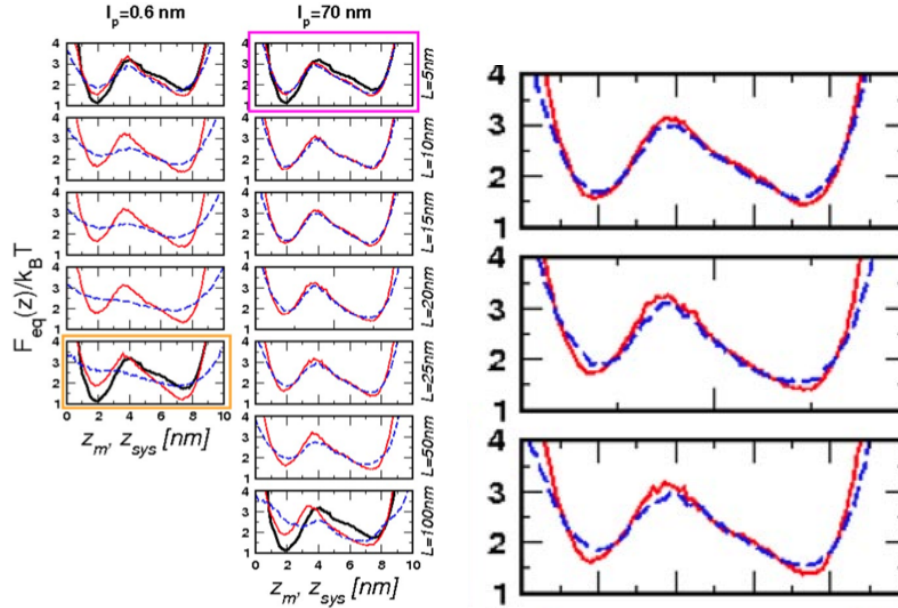


Figura 3.1: A la izquierda se muestran los perfiles de energía para $F_{eq}(z_{sys})$ (línea azul discontinua), $F_{eq}(z_m)$ (línea roja continua) y $F_{eq}^o(z_m)$ (línea negra continua), figuras tomadas de [9]. En el lado derecho se muestra una ampliación de dichos perfiles.

Tabla 1: Casos con asas rígidas

Caso	Coordenada de reacción	l_p	L
1	z_{sys}	70nm	5nm
2	z_{sys}	70nm	10nm
3	z_{sys}	70nm	15nm
4	z_{sys}	70nm	20nm
5	z_{sys}	70nm	25nm
6	z_{sys}	70nm	50nm

Tabla 2: Casos con asas flexibles

Caso	Coordenada de reacción	l_p	L
7	z_{sys}	70nm	100nm
8	z_{sys}	0,6nm	5nm
9	z_{sys}	0,6nm	25nm
10	z_m	0,6nm	5nm
11	z_m	0,6nm	25nm

En concreto, se utilizaron nueve casos para z_{sys} y dos para z_m , debido a que el interés se centra en la dinámica del sistema Asa-ARN-Asa, ya que éste está asociado a los experimentos LOT. Asimismo, los casos mencionados muestran a primera vista una potencial existencia de saltos entre estados NBA y UBA.

Primero se encontró una expresión para el polinomio asociado a la energía libre en cuestión. La cual representa el desdoblamiento de una horquilla de ARN en presencia de un sistema de LOT ó sin este. El ajuste obtenido fue de la forma:⁴

$$F_{ef} = \frac{F}{kT} = a_4 z^4 - a_3 z^3 + a_2 z^2 - a_1 z + a_0, \quad (3.14)$$

donde hemos llamado a F_{ef} a la energía libre en cuestión dividida entre la constante de Boltzmann y la temperatura del baño.

Ahora, teniendo en cuenta que la fuerza de tensión asociada a esta energía libre es $\tau_{ef} = -\frac{dF_{ef}}{dz}$, dicha fuerza es de la forma

$$\tau_{ef} = \tau(kT) = -b_3 z^3 + b_2 z^2 - b_1 z + b_0. \quad (3.15)$$

Introduciendo la fuerza encontrada en la ecuación de Langevin (3.13), se obtiene

$$\gamma \frac{dz}{dt} = -b_3 z^3 + b_2 z^2 - b_1 z + b_0 + \xi(t), \quad (3.16)$$

donde se usarán los valores reportados en [9] para $\gamma = 6\pi\eta a$, con $a = 1\mu m$ y $\eta \approx 1cP$, por lo que se tiene $\gamma = 6\pi(1cP)(1\mu m)$, si $1cP = 0,001\frac{Ns}{m^2}$, tenemos que ⁵ $\gamma = 6\pi(1 \times 10^{-9}\frac{pNs}{nm^2})(1000nm)$, que al simplificar es igual a $\gamma = 6\pi(1 \times 10^{-6})\frac{pNs}{nm}$.

Dado esto, a partir de la ecuación (3.15), ya que ésta es un balance de fuerzas, se pueden obtener las unidades para los coeficientes de la tensión las cuales son: pN para b_0 , $\frac{pN}{nm}$ para b_1 , $\frac{pN}{nm^2}$ para b_2 y $\frac{pN}{nm^3}$ para b_3 .

De igual forma, dado que en la ecuación (3.14) la energía libre (con unidades J) está dividida por la temperatura (con unidades en K) y la constante de Boltzmann $1,38 \times 10^{-23}\frac{J}{K}$, el lado izquierdo es adimensional. Por lo que se deduce que las unidades para los coeficientes son $\frac{1}{nm^4}$ para a_4 , $\frac{1}{nm^3}$ para a_3 , $\frac{1}{nm^2}$ para a_2 , $\frac{1}{nm}$ para a_1 y a_0 es adimensional.

Como se ha manejado, la variable z representa, para el caso del sistema Asa-ARN-Asa, a la distancia entre las esferas de poliestireno unidas por la horquilla de ARN y las asas, y se denota z_{sys} . Para el otro caso, la variable z representa a la extensión molecular de la horquilla de ARN para un sistema sin asas, siendo denotada por z_m . El sistema cumple con el respectivo teorema de fluctuación-disipación (2.102), tal que $\xi(t)$ es la fuerza fluctuante asociada al medio y esta cumple con las condiciones de ruido blanco (2.101).

La solución de la ecuación de Langevin (3.16) se simula a partir de la aproximación de Euler-Maruyama[22] presentada en la sección 2.4.

3.1.3. Ecuación de Fokker-Planck para (3.16) y su solución estacionaria

La ecuación (3.16) que hemos utilizado para modelar a nuestra sistema es de la forma

$$dz(t) = a[z(t)]dt + \sigma[z(t)]dW, \quad (3.17)$$

⁴Se usó el programa Plot Digitizer para obtener los datos de las gráficas y se hizo un ajuste polinomial en Octave/Matlab para obtener los coeficientes.

⁵Si la distancia para la extensión es medida en nano metros (nm) y la fuerza es medida en pico Newtons (pN).

donde $a[z(t)]$ es el coeficiente de deriva, $\sigma[z(t)]$ es la desviación estándar del proceso y W_t es un movimiento Browniano. En general, la función de densidad de probabilidad para el proceso $z(t)$, el cual es descrito por (3.17), sigue una ley de evolución temporal y espacial dada por la ecuación de Fokker-Planck

$$\frac{\partial p(z, t)}{\partial t} = -\frac{\partial}{\partial z}\{a[z(t)]p(z, t)\} + \frac{\partial}{\partial z}\{D[z(t)]\frac{\partial}{\partial z}[p(z, t)]\}. \quad (3.18)$$

Aquí $D = \frac{\sigma^2}{2}$, para nuestro caso se ha tomado al coeficiente de difusión como constante (igual a uno). El coeficiente de deriva es de la forma

$$a(z) = -\frac{dF_{ef}(z)}{dz}, \quad (3.19)$$

por lo que la ecuación (3.17) es

$$dz(t) = -\frac{dF_{ef}(z)}{dz}dt + \sigma dW, \quad (3.20)$$

y la respectiva ecuación de Fokker-Planck es [24]

$$\frac{\partial p(z, t)}{\partial t} = \frac{\partial}{\partial z}\left\{\frac{dF_{ef}(z)}{dz}p(z, t)\right\} + \frac{\sigma^2}{2}\frac{\partial^2 p(z, t)}{\partial z^2}. \quad (3.21)$$

Recordando la ecuación (2.78) de la sección 2.2.6, la solución estacionaria de la ecuación de Fokker-Planck (3.21) viene dada por

$$p_{st}(z) = Ce^{-\frac{2F_{ef}(z)}{\sigma^2}}. \quad (3.22)$$

Aquí C es una constante de normalización tal que cumple

$$\int_{\mathcal{D}} p_0(z)dz = 1, \quad (3.23)$$

donde \mathcal{D} es el dominio para nuestra variable de estado. En nuestro caso $\mathcal{D} = [0nm, 10nm]$, por lo que usando (3.23) en (3.22) podemos depejar C y calcular su valor. Teniendo

$$C = \frac{1}{\int_0^{10} e^{-\frac{2F_{ef}(z)}{\sigma^2}} dz}. \quad (3.24)$$

Por lo tanto, usando las ecuaciones (3.24), (3.22) y (3.14), así como los datos para los coeficientes de la energía libre presentados en las tablas 3-13 de la sección 3.2, podemos encontrar la solución estacionaria de la ecuación Fokker-Planck (3.21) para cada caso. Las gráficas para dichas soluciones se presentan en las Figuras del apartado 3.2.

3.1.4. Tiempos medios y tasas de transición entre estados plegado y extendido

La teoría de Kramers expuesta en el apartado 2.3, se puede aplicar al modelo descrito por (3.16), (3.15) y (3.14), dado que este posee dos estados de equilibrio. Dichos estados son⁶, en

⁶Ver Figura 1.6

primera instancia, el de plegado (NBA) y en segunda el de extendido (UBA). Para conocer los tiempos medios de transición entre estos estados y sus respectivas tasas de transición simplemente debemos de utilizar la ecuaciones (2.122),(2.123), (2.124) y (2.125) de la sección 2.3.

De la Figura 1.6 de la sección 1.3, llamamos z_F al valor de la extensión de la horquilla de ARN, para el cual se tiene un mínimo local de energía libre F_{ef} , asociado al estado NBA. Por otro lado, llamamos z_U al valor de la extensión para el cual se tiene un mínimo local de F_{ef} , correspondiente al estado UBA. Por último definimos como z_T al valor de la extensión para el cual F_{ef} tiene un máximo local entre los estados NBA y UBA.

Dada esta notación, en el caso de la transición del estado plegado al extendido, el respectivo tiempo medio y su tasa de transición son

$$\tau_{NBA \rightarrow UBA} \approx \frac{2\pi\gamma}{\omega_T\omega_F} e^{[F_{ef}(z_T) - F_{ef}(z_F)]}, \quad (3.25)$$

$$K_{NBA \rightarrow UBA}^+ = \frac{\omega_T\omega_F}{2\pi\gamma} e^{-[F_{ef}(z_T) - F_{ef}(z_F)]}, \quad (3.26)$$

donde $\omega_F = \sqrt{\frac{d^2 F_{ef}}{dz^2}}|_{z_F}$ y $\omega_T = \sqrt{\frac{d^2 F_{ef}}{dz^2}}|_{z_T}$. Para la transición del estado extendido al plegado, de manera análoga se tiene

$$\tau_{UBA \rightarrow NBA} \approx \frac{2\pi\gamma}{\omega_T\omega_U} e^{[F_{ef}(z_T) - F_{ef}(z_U)]}, \quad (3.27)$$

$$K_{UBA \rightarrow NBA}^- = \frac{\omega_T\omega_U}{2\pi\gamma} e^{-[F_{ef}(z_T) - F_{ef}(z_U)]}, \quad (3.28)$$

con $\omega_U = \sqrt{\frac{d^2 F_{ef}}{dz^2}}|_{z_U}$. Por lo que para determinar dichos tiempos y tasas, sólo debemos conocer los valores de z_F , z_U y z_T , usarlos en la respectiva ecuación de energía libre dada por (3.14), en su segunda derivada y sustituir éstas en las ecuaciones (3.25), (3.26), (3.27) y (3.28). Los resultados obtenidos se muestran en la tabla 14 del siguiente apartado.

3.2. Resultados de las simulaciones

Los valores obtenidos para los coeficientes de la energía libre F_{ef} y la tensión τ_{ef} , para cada caso, se presentan en las tablas 3-11. Las respectivas gráficas para la energía libre a partir de [9]⁷, el ajuste de dicha energía libre, la fuerza de tensión asociada, las correspondientes simulaciones para la ecuación de Langevin, entropías producidas y las soluciones estacionarias de la ecuación de Fokker-Planck se muestran en las Figuras que acompañan a cada tabla. Primero se muestran los casos con asas rígidas y después los casos con asas flexibles.

⁷Para los datos obtenidos de las simulaciones SOP P5GA.

Tabla 3: Caso 1, $z_{sys}, l_p = 70nm, L = 5nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.021	0.39	2.39	5.45	6.02
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.084	1.17	4.78	5.45

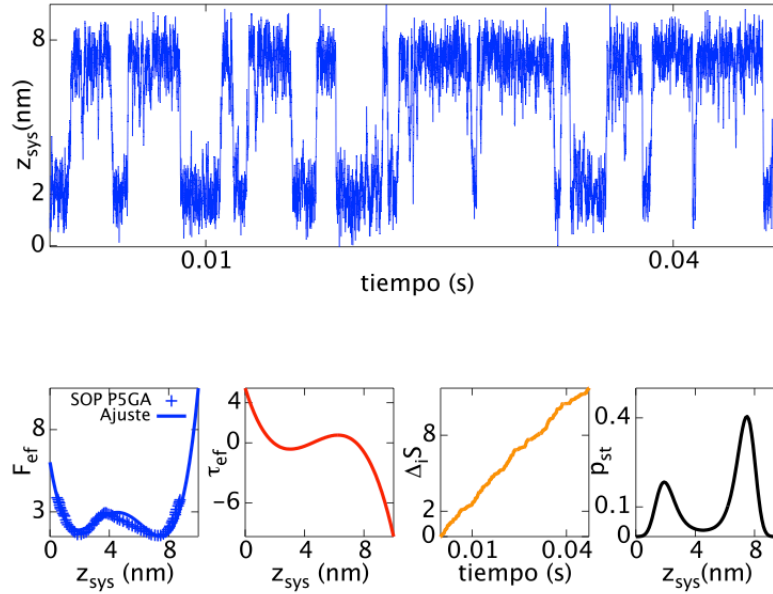


Figura 3.2: $z_{sys}; l_p = 70nm; L = 5nm$.

Tabla 4: Caso 2, z_{sys} , $l_p = 70nm$, $L = 10nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.021	0.38	2.29	5.22	5.82
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.084	1.14	4.58	5.22

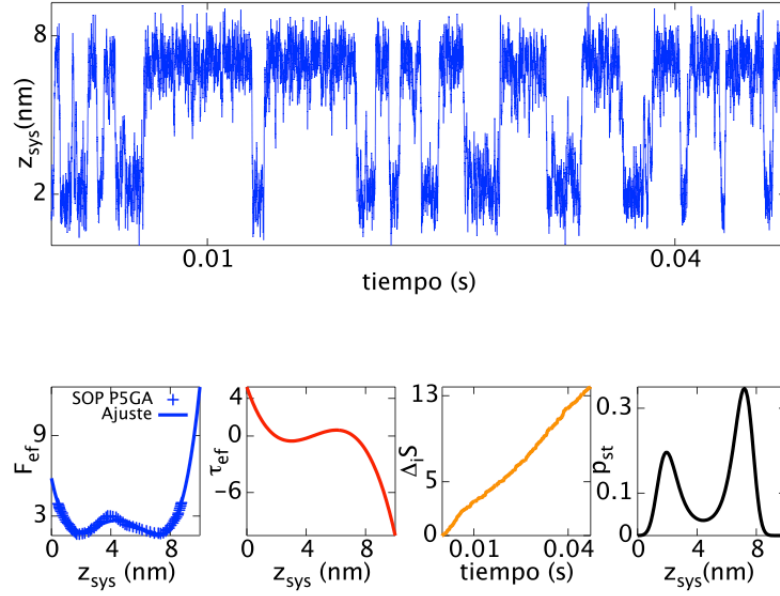


Figura 3.3: z_{sys} ; $l_p = 70nm$; $L = 10nm$.

Tabla 5: Caso 3, z_{sys} , $l_p = 70nm$, $L = 15nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.019	0.3515	2.14	4.77	5.42
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.076	1.0515	4.28	4.77

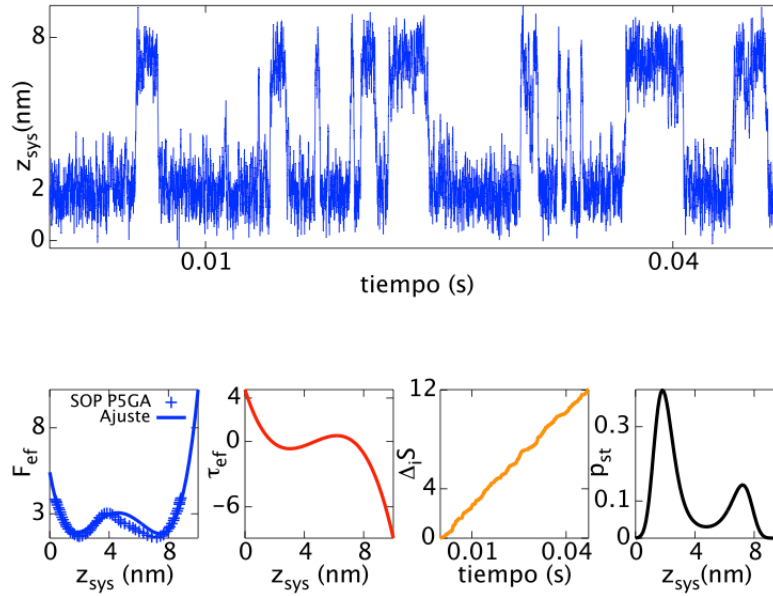


Figura 3.4: z_{sys} ; $l_p = 70nm$; $L = 15nm$.

Tabla 6: Caso 4, $z_{sys}, l_p = 70nm, L = 20nm$; **coeficientes de**
 $F_{ef} = a_4 z^4 - a_3 z^3 + a_2 z^2 - a_1 z + a_0$ **y** $\tau_{ef} = -b_3 z^3 + b_2 z^2 - b_1 z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.04	0.75	4.86	12.88	14.17
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.16	2.25	9.72	12.88

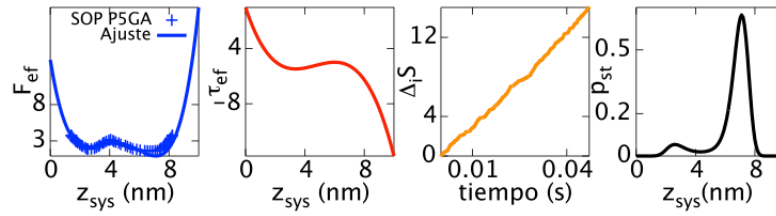
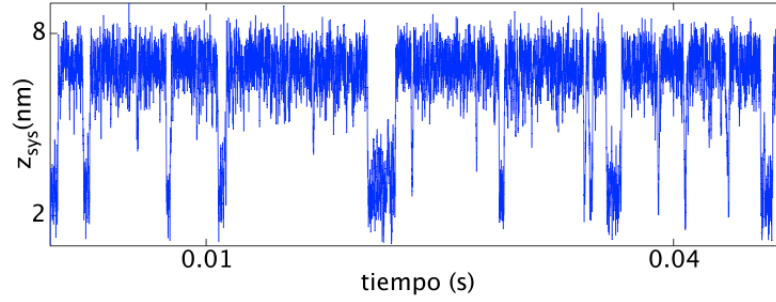


Figura 3.5: $z_{sys}; l_p = 70nm; L = 20nm$.

Tabla 7: Caso 5, $z_{sys}, l_p = 70nm$, $L = 25nm$; **coeficientes de**
 $F_{ef} = a_4 z^4 - a_3 z^3 + a_2 z^2 - a_1 z + a_0$ y $\tau_{ef} = -b_3 z^3 + b_2 z^2 - b_1 z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.017	0.31	1.86	4.17	5.09
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.068	0.93	3.72	4.17

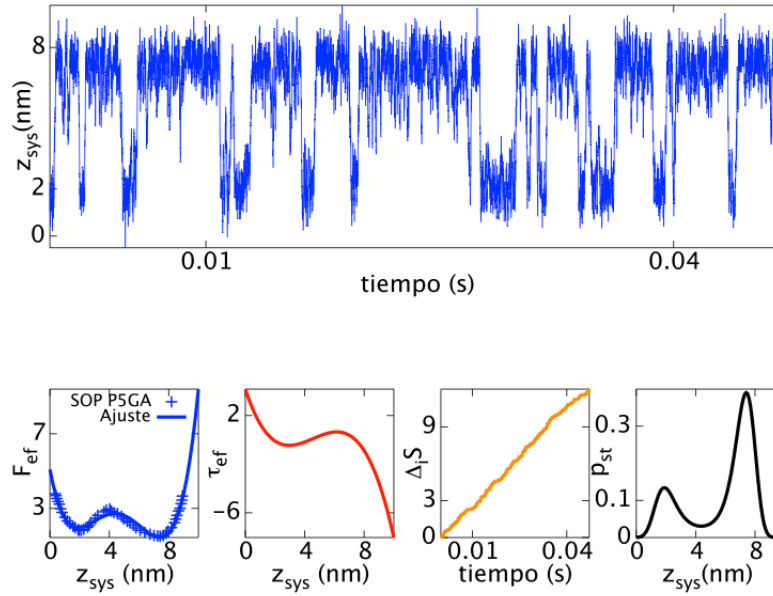


Figura 3.6: $z_{sys}; l_p = 70nm; L = 25nm$.

Tabla 8: Caso 6, $z_{sys}, l_p = 70nm, L = 50nm$; **coeficientes de**
 $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ **y** $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.0122	0.214	1.23	2.62	3.93
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.0488	0.642	2.46	2.62

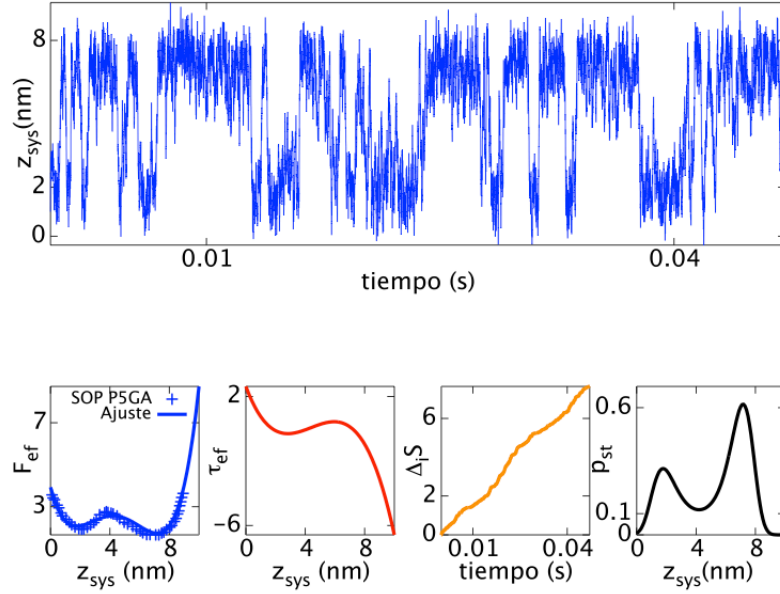


Figura 3.7: $z_{sys}; l_p = 70nm; L = 50nm$.

Tabla 9: Caso 7, $z_{sys}, l_p = 70nm, L = 100nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.0037	0.061	0.365	1.14	3.99
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.0148	0.183	0.73	1.14

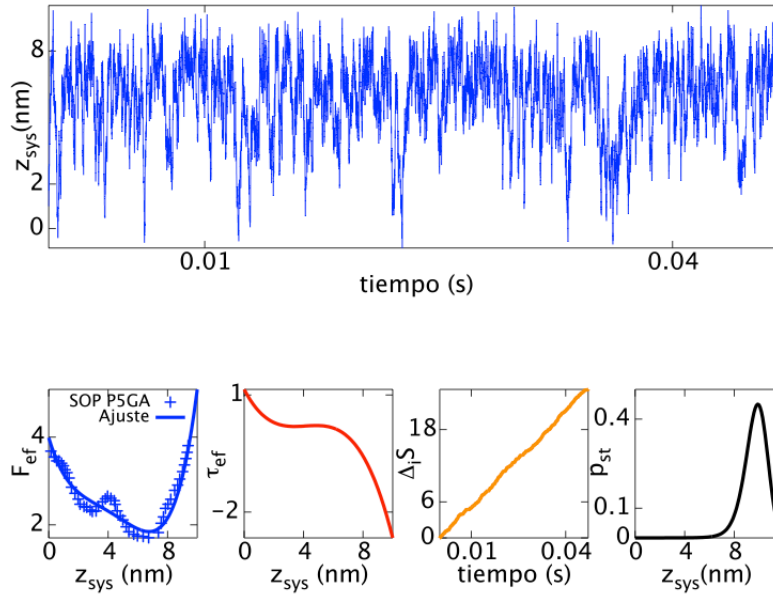


Figura 3.8: $z_{sys}; l_p = 70nm; L = 100nm$.

Tabla 10: Caso 8, $z_{sys}, l_p = 0,6nm, L = 5nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_{sys}, l_p = 0,6nm, L = 5nm$	0.0121	0.221	1.32	2.85	4.10
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.048	0.66	2.64	2.85

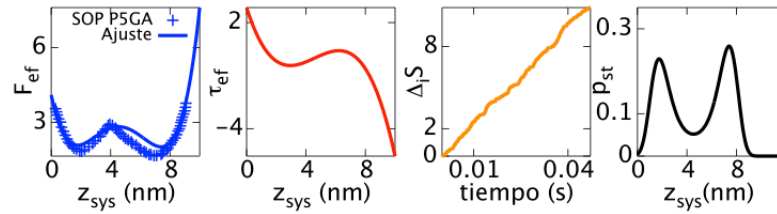
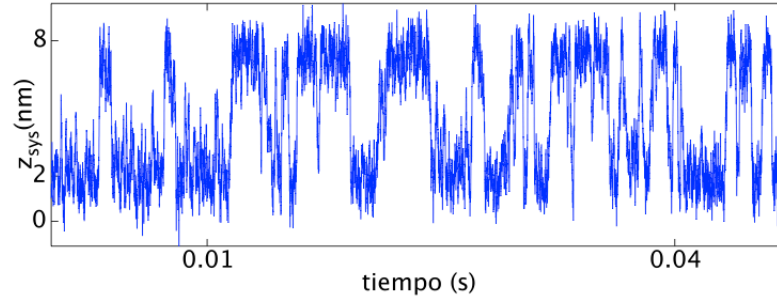


Figura 3.9: $z_{sys}; l_p = 0,6nm; L = 5nm$.

Tabla 11: Caso 9, $z_{sys}, l_p = 0,6nm, L = 25nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_{sys}, l_p = 0,6nm, L = 25nm$	0.003	0.0495	0.28	0.81	3.50
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.012	0.1485	0.56	0.81

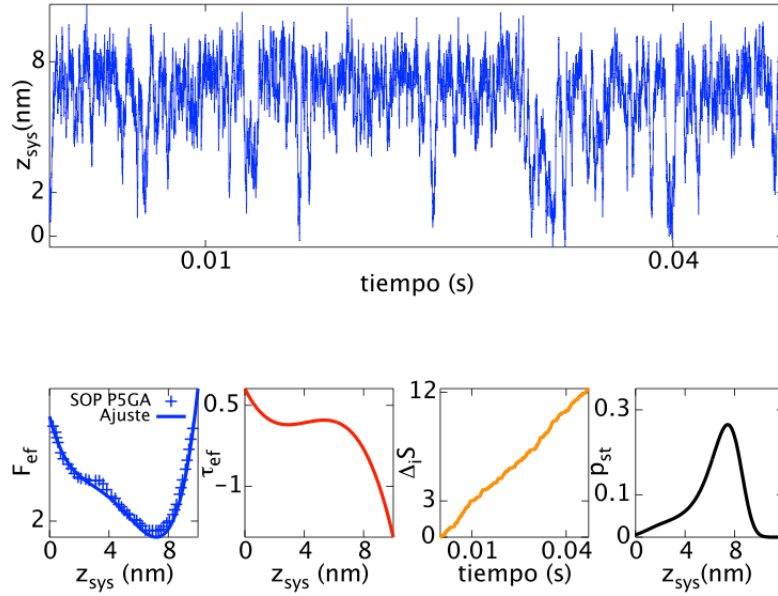


Figura 3.10: $z_{sys}; l_p = 0,6nm; L = 25nm$.

Tabla 12: Caso 10, $z_m, l_p = 0,6nm, L = 5nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 5nm$	0.0274	0.49	2.88	6.27	6.33
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.1096	1.47	5.76	6.27

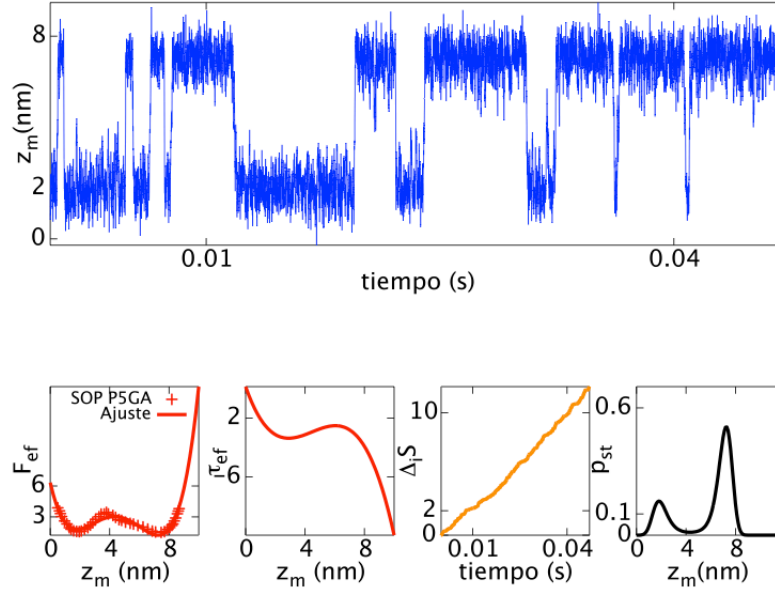


Figura 3.11: $z_m; l_p = 0,6nm; L = 5nm$.

Tabla 13: Caso 11, $z_m, l_p = 0,6nm, L = 25nm$; coeficientes de $F_{ef} = a_4z^4 - a_3z^3 + a_2z^2 - a_1z + a_0$ y $\tau_{ef} = -b_3z^3 + b_2z^2 - b_1z + b_0$

Caso	$a_4(\frac{1}{nm^4})$	$a_3(\frac{1}{nm^3})$	$a_2(\frac{1}{nm^2})$	$a_1(\frac{1}{nm})$	a_0
$z_m, l_p = 0,6nm, L = 25nm$	0.027	0.48	2.84	6.33	6.93
		$b_3(\frac{pN}{nm^3})$	$b_2(\frac{pN}{nm^2})$	$b_1(\frac{pN}{nm})$	$b_0(pN)$
		0.108	1.44	5.68	6.33

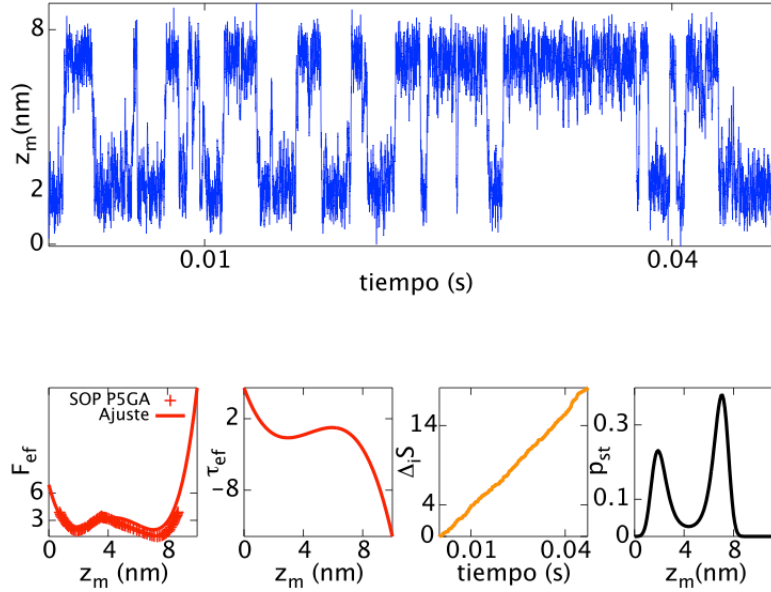


Figura 3.12: $z_m; l_p = 0,6nm; L = 25nm$.

A partir de las gráficas anteriores 3.2-3.12, observamos en primera instancia las simulaciones para la extensión de las horquillas de ARN. Como se puede apreciar en todas ellas existen transiciones entre los estados plegado y extendido, dichas transiciones son mayores para unos casos que para otros. Esta diferencia entre el número de transiciones entre estados está asociada al paisaje energético del sistema (F_{ef}). Entre más profundo sea un pozo de energía asociado a un estado determinado, en comparación con el otro restante, el sistema “preferirá” permanecer más tiempo en el primero, provocando que se reduzca el número de transiciones entre estados. Dicha situación se plasma en las respectivas densidades estacionarias de probabilidad asociadas, mostrando que es más probable encontrar el sistema en una configuración tal que su pozo de energía es más profundo con respecto al restante.

Dado esto, en términos de la clasificación por asas rígidas o flexibles para el sistema simulado, la preferencia por estar en determinado estado accesible “plegado o extendido”, no está determinada por la naturaleza de las asas.

Los casos en los que la dinámica del sistema es tal que este tiene en mayor o menor medida preferencia por el estado plegado (NBA) solo es el caso 3. En cambio los casos para los cuales el sistema

tiene preferencia por el estado extendido (UBA) son: 1, 2, 4, 5, 6, 7, 8, 9, 10 y 11. Por lo que, de los casos estudiados el sistema para una arbitraria configuración de asas, ante el estímulo de cierta fuerza externa constante, tiene una tendencia a ocupar el estado extendido.

El hecho de que las transiciones entre estados sean mayores para unos casos con respecto a otros, desemboca también en que algunas densidades estacionarias de probabilidad sean monomodales o bimodales. Para pocas transiciones entre estados tenemos densidades monomodales que están representadas por los casos 7 y 9. Dichas funciones de densidad monomodales claramente están sesgadas y poseen cierta curtosis. En cambio todos los demás casos presentan funciones de densidad estacionarias bimodales.

Es pertinente mencionar que desde que las densidades estacionarias son monomodales (con sesgo y curtosis) y bimodales, dichas densidades no son de equilibrio.

En los 11 casos el ajuste polinomial (línea continua) de la energía libre F_{ef} es consistente con los datos obtenidos de la literatura ⁸ (puntos denotados por cruces). A su vez, para todos los casos la entropía asociada a los procesos irreversibles realizados por el sistema es positiva y crece con respecto del tiempo, por lo que la producción de entropía es positiva.

A continuación, en la Tabla 14 se presentan los tiempos medios y tasas de transición entre estados, obtenidos a partir de la teoría de primer paso de Kramers presentada en el apartado 3.1.4 y usando nuestra aproximación polinomial de la energía libre.

Tabla 14: Tiempos medios y tasas de transición entre estados, para cada caso.

Caso	$z_F [nm]$	$z_U [nm]$	$z_T [nm]$	$\tau_{NBA \rightarrow UBA} [s]$	$K_{NBA \rightarrow UBA}^+ [\frac{1}{s}]$	$\tau_{UBA \rightarrow NBA} [s]$	$K_{UBA \rightarrow NBA}^- [\frac{1}{s}]$
1	1,93	7,28	3,71	$3,9 \times 10^{-4}$	2564,10	$5,9 \times 10^{-4}$	1694,91
2	1,84	7,20	3,85	$3,5 \times 10^{-4}$	2857,14	$4,4 \times 10^{-4}$	2272,72
3	1,96	7,11	3,92	$4,6 \times 10^{-4}$	2173,91	$4,5 \times 10^{-4}$	2222,2
4	2,13	6,29	3,6	$1,7 \times 10^{-4}$	5882,35	11×10^{-4}	909,09
5	2,00	7,41	3,91	$3,9 \times 10^{-4}$	2564,10	$5,8 \times 10^{-4}$	1724,13
6	2,04	6,90	3,89	$4,9 \times 10^{-4}$	2040,81	$6,1 \times 10^{-4}$	1639,34
7	2,92	6,91	3,97	$3,3 \times 10^{-3}$	297,61	$2,2 \times 10^{-3}$	296,87
8	1,90	7,18	3,99	$5,03 \times 10^{-4}$	1988,07	$5,5 \times 10^{-4}$	1818,18
9	1,77	7,02	3,69	$1,09 \times 10^{-3}$	917,43	$2,1 \times 10^{-3}$	476,19
10	1,86	7,30	3,97	$2,3 \times 10^{-4}$	4347,82	$5,9 \times 10^{-4}$	1694,91
11	1,80	7,32	3,67	$3,2 \times 10^{-4}$	3125	$3,5 \times 10^{-4}$	2857,14

Los datos obtenidos de esta tabla corroboran el hecho de que las transiciones entre estados son menores para los casos que presentaron densidades estacionarias monomodales, como se observa para los casos 7 y 9 ambas tasas de transición $K_{NBA \rightarrow UBA}^+$ y $K_{UBA \rightarrow NBA}^-$ son menores en un orden de magnitud en comparación con los demás casos bimodales. Asimismo, para estos casos 7 y 9 los tiempos medios entre transiciones $\tau_{NBA \rightarrow UBA}$ y $\tau_{UBA \rightarrow NBA}$ son mayores en un orden de magnitud comparadas con los casos bimodales.

Las tasas de transición encontradas son acordes a los perfiles de energía y densidades estacionarias, puesto que para los casos en que el sistema es tal que prefiere estar en el estado NBA, caso 3, las tasas de transición del estado extendido al plegado $K_{UBA \rightarrow NBA}^-$ son mayores que las tasas de transición del estado plegado al extendido $K_{NBA \rightarrow UBA}^+$. Para todos los demás casos en los que el sistema prefiere permanecer en el estado extendido, las tasas de transición del estado plegado al extendido $K_{NBA \rightarrow UBA}^+$ son mayores a las de transición de extendido a plegado $K_{UBA \rightarrow NBA}^-$.

⁸Modelo SOP P5GA [9].

Por último, es importante enfatizar que los tiempos medios encontrados son del mismo orden de magnitud que los reportados en la literatura [9].

3.3. Ecuación de Langevin con ruido de color aditivo

En el modelo representado por (3.16) se ha supuesto que la fuerza fluctuante $\xi(t)$ cumple con las propiedades de un ruido blanco, puesto que dicha fuerza estaba solamente asociada al baño térmico en el que está insertado el sistema. En el modelo anterior no se considero que la tensión aplicada por las LOT a través de las esferas de poliestireno también está sujeta a fluctuaciones, ya que las esferas realizan, en general un movimiento aleatorio confinado a un potencial armónico⁹. En este caso particular el ruido es considerado interno ya que no se atribuye a cualidades externas como las fluctuaciones térmicas del baño [25].

En otros términos, la dinámica del sistema físico descrito, Asa-ARN-Asa, esta acoplada con las fluctuaciones de éste, y dichas fluctuaciones son atribuidas a la dinámica de las esferas de poliestireno unidas a los extremos de las asas, las cuales son manipuladas por LOT. Para esta situación se dice que la fuerza fluctuante es un ruido de color, o en otras palabras que dicha fuerza esta correlacionada para distintos tiempos de evolución del sistema. Esto implica que el modelo presentado en la ecuación (3.16) cambia. Si las fluctuaciones de las esferas de poliestireno decimos que son armónicas como se planteó en la sección 1.3 (ver ecuación (1.2)) decimos entonces que la fuerza fluctuante sigue un proceso de Ornstein-Uhlenbeck y llamamos a dicha fuerza fluctuante $\xi_{OU}(t)$. Por lo que nuestra ecuación de Langevin es ahora:

$$\gamma \frac{dz}{dt} = -b_3 z^3 + b_2 z^2 - b_1 z + b_0 + \xi_{OU}(t). \quad (3.29)$$

Aquí la función de correlación para $\xi_{OU}(t)$ es[6]

$$\Gamma(t, t') = \langle \xi_{OU}(t) \xi_{OU}(t') \rangle = \left(\frac{D}{\tau_n} \right) e^{-\frac{|t-t'|}{\tau_n}}, \quad (3.30)$$

donde τ_n es el tiempo de correlación para el ruido ¹⁰.

Para poder realizar la simulación para la solución de la ecuación (3.29) por medio del esquema de Euler-Maruyama, es necesario generar el proceso Ornstein-Uhlenbeck para $\xi_{OU}(t)$. En términos de una ecuación diferencial estocástica dicho proceso se escribe como

$$d\xi_{OU} = \theta(\xi_{OU} - \mu)dt + \sigma dW, \quad (3.31)$$

donde θ es una constante de proporcionalidad, μ es la media y σ representa la volatilidad o intensidad del ruido, estos son parámetros a determinar dadas las características del sistema. Dado lo anterior, el proceso Ornstein-Uhlenbeck también puede ser obtenido por la aproximación de Euler-Maruyama¹¹. Al hacerlo, la serie obtenida es utilizada para simular mediante el mismo esquema la solución de (3.29).

Como comentamos, en nuestro caso el ruido de color está asociado a la dinámica de las esferas de poliestireno que estan presentes en los experimentos con LOT, definiendo $\xi_{OU} := z_b$, donde z_b

⁹Dicho potencial armónico está asociado a la influencia de las LOT sobre la esfera.

¹⁰Como se verá adelante en nuestro caso particular dicho tiempo de correlación es igual a γ

¹¹Ver apartado 2.4.

es la posición de la esfera de poliestireno que es manipulada por medio de LOT, en concreto z_b es un proceso de Ornstein-Uhlenbeck de la forma de (3.31) y está representado por

$$dz_b = -\frac{\kappa_b}{\gamma} z_b dt + \frac{1}{\gamma} dW, \quad (3.32)$$

la cual puede ser reescrita como

$$\frac{dz_b}{dt} = -\frac{\kappa_b}{\gamma} z_b + \frac{1}{\gamma} \xi_b(t), \quad (3.33)$$

donde κ_b es una constante de restitución y su valor es $\kappa_b \approx 0,11 \frac{pN}{nm}$, γ es el respectivo coeficiente asociado a la ley de Stokes, donde se toma el valor usado en (3.29) $\gamma = 6\pi(1 \times 10^{-6}) \frac{pNs}{nm}$, puesto que la dinámica de la esfera tiene lugar en un entorno similar al del sistema Asa-ARN-Asa. Por último $\xi_b(t)$ es un ruido blanco asociado al ruido térmico que afecta a la cabeza de poliestireno.

Con estas consideraciones, la ecuación de Langevin con ruido de color es

$$\gamma \frac{dz}{dt} = -b_3 z^3 + b_2 z^2 - b_1 z + b_0 + \kappa_b z_b(t). \quad (3.34)$$

Siguiendo la sección 2.2.7, en donde se encuentra la respectiva ecuación de Fokker-Planck para una ecuación de Langevin con ruido de color representada por (2.79), usando este esquema podemos encontrar la respectiva ecuación de Fokker-Planck para (3.34). En nuestro caso a partir de (3.33) y (3.34), podemos reconocer los respectivos coeficientes de deriva y difusión: $a(\cdot) = \tau_{ef}$, $g(\cdot) = \kappa_b$, $D = 1$. En este caso el tiempo de correlación es muy pequeño, ya que $\tau_n = \gamma$. Siendo pertinente utilizar la ecuación de Fokker-Planck (2.91). Sustituyendo, la correspondiente ecuación de Fokker-Planck es

$$\frac{\partial p(z, t)}{\partial t} = -\frac{\partial}{\partial z} [\tau_{ef} p(z, t)] + \left\{ \frac{\partial}{\partial z} \kappa_b \frac{\partial}{\partial z} \kappa_b [1 + \gamma \kappa_b \frac{d}{dz} (\frac{\tau_{ef}}{\kappa_b})] \right\} p(z, t). \quad (3.35)$$

En la sección 2.2.7, se encontró una expresión para la solución estacionaria para este tipo de ecuación de Fokker-Planck, dicha expresión es la ecuación (2.96), por lo que sustituyendo los respectivos coeficientes de deriva y difusión, para nuestro caso tenemos

$$p_{st}(z) = \frac{C [1 - \gamma \frac{d}{dz} (\tau_{ef})]}{\kappa_b} e^{\int^z \frac{\tau_{ef} dz'}{\kappa_b^2}}, \quad (3.36)$$

que en términos de la energía libre efectiva se tiene

$$p_{st}(z) = \frac{C [1 - \gamma \frac{d}{dz} (\tau_{ef})]}{\kappa_b} e^{-\frac{F_{ef}}{\kappa_b^2}}. \quad (3.37)$$

De igual manera que en la sección 3.1.3, el valor de la constante de normalización viene dado por

$$C = \frac{\kappa_b}{\int_0^{10} [1 - \gamma \frac{d}{dz} (\tau_{ef})] e^{-\frac{F_{ef}(z)}{\kappa_b^2}} dz}. \quad (3.38)$$

Por lo que a partir de la ecuaciones (3.37) y (3.38) se pueden obtener las funciones de densidad estacionarias que satisfacen la ecuación de Fokker-Planck (3.35) y los procesos (3.34) y (3.33).

3.3.1. Resultados numéricos para el caso con ruido de color aditivo

A continuación se presentan simulaciones de la ecuación de Langevin (3.34) usando la serie de tiempo obtenida del proceso Ornstein-Uhlenbeck $z_b(t)$. Así también, como en 3.2, dichas simulaciones se presentan junto con sus respectivas gráficas para la energía libre, tensión, producción de entropía y solución estacionaria para la correspondiente ecuación de Fokker-Planck.

Dichas simulaciones se muestran en las Figuras 3.13, 3.14 y 3.15 a continuación, en específico se muestran los casos representativos: caso 2 z_{sys} , $l_p = 70nm$, $L = 10nm$; caso 5 z_{sys} , $l_p = 70nm$, $L = 25nm$ y caso 8 z_{sys} , $l_p = 0,6nm$, $L = 5nm$.

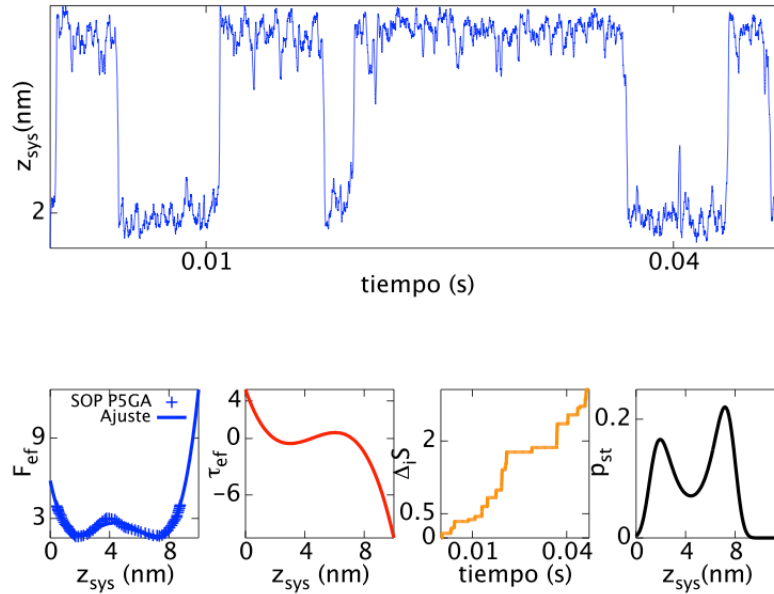


Figura 3.13: z_{sys} ; $l_p = 70nm$; $L = 10nm$, con ruido de color.

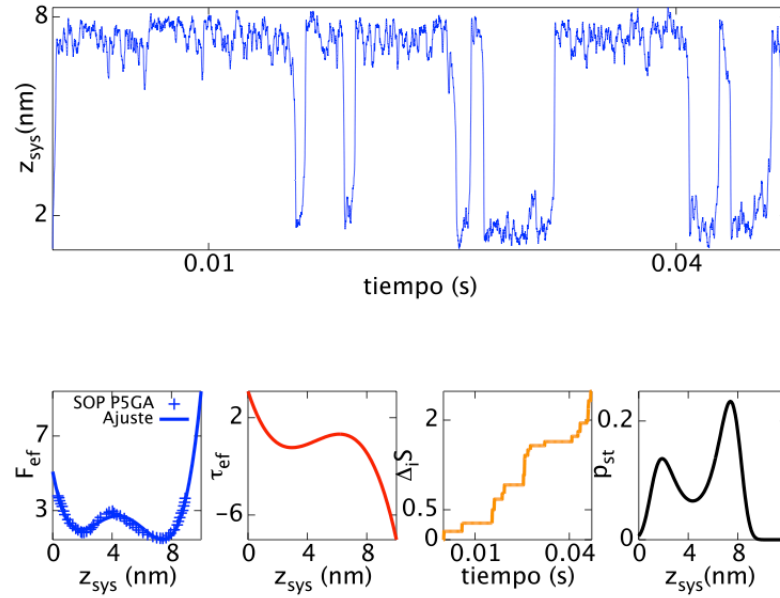


Figura 3.14: z_{sys} ; $l_p = 70nm$; $L = 25nm$, con ruido de color.

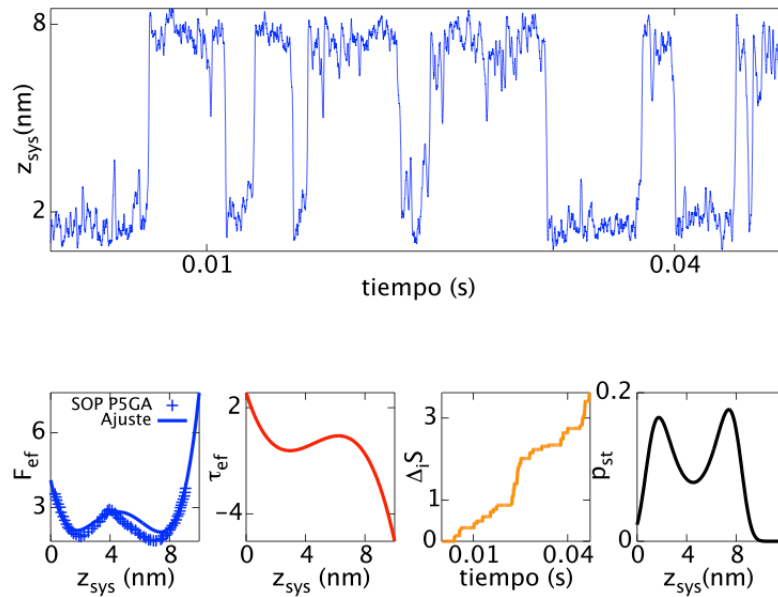


Figura 3.15: z_{sys} ; $l_p = 0,6nm$; $L = 5nm$, con ruido de color.

Las simulaciones del proceso con ruido de color presentan en primera instancia menos fluctuaciones alrededor de los estados de equilibrio plegado ó extendido, ya que si comparamos la gráficas con las respectivas del problema con ruido blanco, las primeras son más densas en la cercanía de dichas regiones. Aún así, los saltos entre estados son más notorios para los procesos con ruido de color para cada caso simulado, en contraste con los desarrollados con ruido blanco.

Se siguen conservando las preferencias de cada sistema por permanecer en el estado con un pozo de potencial más profundo, puesto que en las gráficas para los tres casos, se observa que el sistema tiende a permanecer más veces en el estado extendido, en comparación con el estado plegado.

Al comparar las series de tiempo para la extensión obtenidas con las obtenidas de manera experimental en [1], se puede apreciar que nuestros resultados en éste caso de ruido de color aditivo, se asemejan más a los resultados experimentales¹².

Las gráficas de cambio de entropía interna con respecto al tiempo siguen teniendo características similares para los casos con ruido blanco. Para los tres casos, el cambio de entropía es siempre positivo y este crece con respecto del tiempo, por lo que la producción de entropía asociada al proceso de plegamiento de la horquilla de ARN es positiva, esto demuestra que los procesos de plegamiento conllevan disipación de energía. Cabe mencionar que las gráficas de entropía contra tiempo toman ahora una forma de escalera, la cual es más evidente para los casos 2 y 5. Dicha forma de escalera puede atribirse al hecho de que ahora los saltos entre los estados tardan más en ocurrir, como se observa en la gráfica de entropía contra tiempo para el caso 2.

En lo que respecta a las densidades estacionarias de probabilidad, éstas en estructura son similares en comparación con las de los casos con ruido blanco.

¹²Sólo basta con comparar las series de tiempo para la extensión mostradas en las Figuras 3.13, 3.14 y 3.15 con la serie de tiempo para la extensión mostrada en la Figura 1.5a.

CAPÍTULO 4

Conclusiones

A partir de los resultados expuestos en el capítulo anterior, podemos decir que los modelos propuestos para representar la dinámica de plegamiento de una horquilla de ARN sujeta a una fuerza externa y en particular para un sistema Asa-ARN-Asa manipulado por LOT, son una buena aproximación a este fenómeno. Puesto que para el primer modelo representado por las ecuaciones (3.16), (3.15) y (3.14), los ajustes polinomiales de la energía libre F_{ef} coinciden de buena manera con los datos de las simulaciones de los modelos SOP PGA5 presentados en [9].

A su vez, las simulaciones realizadas con la aproximación de Euler-Maruyama para la extensión del sistema Asa-ARN-Asa, son consistentes con los respectivos paisajes energéticos.

Este hecho se corroboró al momento de encontrar las densidades estacionarias de probabilidad para las correspondientes ecuaciones de Fokker-Planck, mostrando que la mayor probabilidad ó menor para cierta configuración de la extensión del sistema, estaba determinada por su paisaje energético. Obteniendo así, funciones de densidad monomodales para energías libres en las que la diferencia de profundidad entre los pozos de energía para los estados plegado y extendido era muy grande (casos 7 y 9), y resultando funciones de densidad bimodales para los casos en que esta diferencia de profundidad entre pozos era menor (casos 1, 2, 3, 4, 5, 6, 8, 10 y 11).

Además de que dichas densidades estacionarias favorecieron al estado extendido, ya que la forma de ésta indicaba en mayor o menor medida preferencia hacia esta configuración para todos los casos excepto para el tres en la que la función de densidad favorecía al estado plegado. Como se comenta, esto fue producto de las características de los perfiles energéticos, ya que para los mencionados en primera instancia el paisaje energético tenía una inclinación hacia el estado extendido, teniendo lo opuesto para el caso 3.

Todo lo anterior es corroborado con los resultados desprendidos de la aplicación de la teoría de primer paso de Kramers al modelo, ya que para los casos con funciones de densidad monomodales (casos 7 y 9) las tasas de transición entre estados son menores en un orden de magnitud con respecto a los casos con densidades bimodales, además de que los tiempos medios de transición entre estados para los casos monomodales son mayores en un orden de magnitud con respecto a los otros.

Para todos los casos exceptuando el 3, las tasas de transición del estado plegado al extendido son mayores con respecto a las tasas de transición del estado extendido al plegado. Teniendo que los tiempos medios de transición del estado plegado al extendido son menores a los tiempos medios de transición del estado extendido al plegado. Ocurrendo lo contrario para el caso 3, por lo que los resultados mostrados son congruentes con los de las simulaciones realizadas para la extensión del sistema Asa-ARN-Asa.

La entropía producida (3.12) para la dinámica de plegamiento para el sistema estudiado, cumplió para todos los casos, con ser positiva y con que ésta se incrementaba conforme el tiempo avanzaba, lo que muestra la naturaleza irreversible del fenómeno estudiado.

El segundo modelo representado por (3.34) y (3.33), propuesto para describir el fenómeno de plegamiento del sistema Asa-ARN-Asa ante la manipulación de LOT y tomando en cuenta fluctuaciones intrínsecas al sistema, como lo es la perturbación asociada a una de las esferas de poliestireno manipulada por LOT y sujeta a una asa, presentó coincidencias de resultados con el primer modelo. Dichas coincidencias son, en el sentido de que las simulaciones realizadas bajo la respectiva aproximación de Euler-Maruyama mostraron que eran consistentes con los perfiles energéticos asociados a la correspondiente configuración del sistema, por otro lado, éstas se asemejaron más a los resultados experimentales presentados en [1]. Así como también fueron consistentes con dichos perfiles, las funciones de densidad estacionarias encontradas para este modelo. Igualmente que en el modelo anterior, la entropía producida para todos los casos fue positiva y crecía con respecto al tiempo. En las gráficas de ésta, se observó un ligero cambio, en el sentido de que éstas tomaron una forma de escalera ascendente, lo cual se atribuyó al hecho de que para las simulaciones de la extensión del sistema Asa-ARN-Asa, en éstas los saltos entre los estados tardan más en ocurrir.

El esquema desarrollado para describir la dinámica de plegamiento de horquillas de ARN bajo la influencia de una tensión externa, nos permitió introducir situaciones en las que los instrumentos de manipulación están acoplados con la dinámica del sistema, en nuestro caso representadas por un ruido de color aditivo. Esta libertad de poder introducir distintos tipos de ruidos¹ nos permitió encontrar series de tiempo para la extensión del sistema que empatan de una mejor manera con la evidencia experimental registrada en la literatura. Por lo que nuestro análisis presenta una ventaja con respecto a los que utilizan formalismos como la ecuación maestra [7] y [8]. Puesto que estos últimos sólo permiten analizar sistemas Markovianos y no ofrecen un método para encontrar las series de tiempo para la extensión del sistema.

Dado todo esto, podemos decir que los dos modelos presentados y sus respectivos esquemas de simulación, son consistentes con los perfiles energéticos desprendidos de la literatura [9] y representan una buena aproximación de la dinámica de plegamiento de moléculas de ARN manipuladas por LOT.

¹No sólo ruidos blancos asociados a las fluctuaciones del baño térmico.

CAPÍTULO 5

Apéndice: Códigos empleados

```
N=10000; % Numero de pasos a tomar
T=2500; % Tiempo máximo de la simulación
h=T/N; % Tamaño del paso
ta=(0:h:T); % ta es el vector [0 1h 2h 3h ... Nh]

% Arreglo donde se guardarán los valores de las z's
z=zeros(size(ta));

% coeficientes de la tensión
% ejemplo para zsys,lp=70nm y L=5nm
b0=5.45;
b1=4.78;
b2=1.17;
b3=0.084;

s=1; % Coeficiente de difusión

z(1)=1; % Condición inicial para z

% Iteración para la aproximación de Euler-Maruyama
for i=1:N
z(i+1)=z(i)+b0*h-b1*z(i)*h+b2*(z(i)^2)*h-b3*(z(i)^3)*h+s*sqrt(h)*randn;
end;

% tiempo del sistema físico
t=ta/53051.64;

% gráfica para la extensión en contra del tiempo
plot(t,z)
```

Figura 5.1: Código en OCTAVE/MATLAB para la aproximación de Euler-Maruyama de la ecuación de Langevin(3.16).

```

%Primero: se genera el ruido de color zb

N=10000; % Numero de pasos a tomar
T=2500; % Tiempo máximo de la simulación
h=T/N; % Tamaño del paso
ta=(0:h:T); % ta es el vector [0 1h 2h 3h ... Nh]
t=ta/53051.64; % tiempo del sistema físico

% Arreglo donde se guardarán los valores de las zb's (ruido de color)
zb=zeros(size(ta));

s=1; % Coeficiente de difusión
k=0.11; % Coeficiente de rigidez
zb(1)=1; % Condición inicial para zb
% Iteración para la aproximación de Euler-Maruyama
for i=1:N
zb(i+1)=zb(i)-k*zb(i)*h+s*sqrt(h)*randn;
end;

%Segundo: se sustituye el ruido de color en la ecuación
de Langevin general y se genera z

% Arreglo donde se guardarán los valores de las z's
z=zeros(size(ta));

% coeficientes de la tensión
% ejemplo para zsys, lp=70nm y L=5nm
b0=5.45;
b1=4.78;
b2=1.17;
b3=0.084;

z(1)=1; % Condición inicial para z
% Iteración para la aproximación de Euler-Maruyama
for i=1:N
z(i+1)=z(i)+b0*h-b1*z(i)*h+b2*(z(i)^2)*h-b3*(z(i)^3)*h+k*sqrt(h)*zb(i);
end;

% gráfica para la extensión en contra del tiempo
plot(t,z)

```

Figura 5.2: Código en OCTAVE/MATLAB para la aproximación de Euler-Maruyama de la ecuación de Langevin con ruido de color (3.34) y (3.33).

```

%Primero se genera el proceso z mediante cualquiera
de los dos códigos anteriores

%Se excluyen los estados NBA y UBA del proceso z
zz=zeros(size(z));
for i=1:length(z)
if 3 > z(i)
zz(i) = 0;
else
zz(i)=z(i);
end
end

zzz=zeros(size(z));
for j=1:length(z)
if zz(j) > 6
zzz(j) = 0;
else
zzz(j)=zz(j);
end
end

%Se sustituye el nuevo proceso en la tensión efectiva
sin los estados NBA, UBA

tau=zeros(size(zzz));
tau=b0-b1*zzz+b2*(zzz.^2)-b3*(zzz.^3);

%Se hace cero la fuerza cuando esta es igual a b0, para
tenerla evaluada en los puntos deseados

tau1=zeros(size(tau));
for ii=1:length(tau)
if tau(ii) == b0
tau1(ii) = 0;
else
tau1(ii)=tau(ii);
end
end
tau2=abs(tau1);

%distancia entre estados NBA y UBA depende del caso para z
% ejemplo para zsys, lp=70nm y L=5nm
Dz=5.35;
%entropía producida
S=(Dz*cumsum(tau2))/293.15;

% gráfica para la entropía producida en contra del tiempo
plot(t,S)

```

Figura 5.3: Código en OCTAVE/MATLAB para la simulación de la entropía producida (3.12).


```

#include<stdio.h>
#include<stdlib.h>
#include<math.h>

#define GNUPLOT "gnuplot -persist"

double randn(double,double);
int main()
{ /*Los datos producidos en la simulación son guardados en Langevin.dat*/
FILE *cam=fopen("Langevin.dat", "w");
float h,t,k,b0=5.45,b1=4.78,b2=1.17,b3=0.084,z[10001],sig=1; /*ejemplo para zsys, lp=70nm y L=5nm*/
double mu=0,sigma=1;
int i,N;
z[0]=1; /*condición inicial para z*/
h=0.25; /*tamaño del paso*/
t=2500; /*tiempo máximo de la simulación*/
N=10001; /*numero de pasos a tomar*/
for(i=2;i<N;i++) /*iteración para la aproximación de Euler-Maruyama*/
{
z[i]=z[i-1]+h*b0-h*b1*z[i-1]+h*b2*pow(z[i-1],2)-h*b3*pow(z[i-1],3)+sig*sqrt(h)*randn(mu,sigma);
fprintf(cam, "%f\t%f\n",i*h,z[i]);
}
fclose(cam);
FILE *gp; /*los datos guardados en Langevin.dat se grafican en GNU plot*/
gp = popen(GNUPLOT,"w"); /* 'gp' is the pipe descriptor */
if (gp==NULL)
{
printf("Error abriendo GNU plot");
exit(1);
}

fprintf(gp, "set title 'Langevin Ruido Blanco'\n");
fprintf(gp, "set xlabel 't'\n");
fprintf(gp, "set ylabel 'z'\n");
fprintf(gp, "set xrange [0:2500]\n");
fprintf(gp, "set yrange [0:10]\n");

fprintf(gp, "plot 'Langevin.dat' t 'z' w lines\n");
fprintf(gp, "set terminal png\n");
fprintf(gp, "set output 'Langevin.png'\n");
fprintf(gp, "replot\n");
fflush(gp);
fclose(gp);

return 0;
}

/* generador de numeros aleatorios con distribución normal*/
double randn (double mu, double sigma) /*http://phoxis.org*/
{
double U1, U2, W, mult;
static double X1, X2;
static int call = 0;

if (call == 1)
{
call = !call;
return (mu + sigma * (double) X2);
}

do
{
U1 = -1 + ((double) rand () / RAND_MAX) * 2;
U2 = -1 + ((double) rand () / RAND_MAX) * 2;
W = pow (U1, 2) + pow (U2, 2);
}
while (W >= 1 || W == 0);

mult = sqrt ((-2 * log (W)) / W);
X1 = U1 * mult;
X2 = U2 * mult;

call = !call;

return (mu + sigma * (double) X1);
}

```

Figura 5.4: Código en C para la aproximación de Euler-Maruyama de la ecuación de Langevin(3.16).

```

#include<stdio.h>
#include<stdlib.h>
#include<math.h>

#define GNUPLOT "gnuplot -persist"

double randn(double,double);
int main()
{
    /*Los datos producidos en la simulación son guardados en LanOU.dat*/
    FILE *cam=fopen("LanOU.dat", "w"); /*ejemplo para zsys,lp=70nm y L=5nm*/
    float h,t,k,b0=5.45,b1=4.78,b2=1.17,b3=0.084,z[10001],zb[10001],kb=0.11;
    double mu=0,sigma=1;
    int i,N;
    zb[0]=1; /*condición inicial para zb*/
    z[0]=1; /*condición inicial para z*/
    h=0.25; /*tamaño del paso*/
    t=2500; /*tiempo máximo de la simulación*/
    N=10001; /*numero de pasos a tomar*/
    for(i=2;i<N;i++) /*iteración para la aproximación de Euler-Maruyama*/
    {
        zb[i]=zb[i-1]-kb*zb[i-1]*h+sqrt(h)*randn(mu,sigma);
        z[i]=z[i-1]+h*b0-h*b1*z[i-1]+h*b2*pow(z[i-1],2)-h*b3*pow(z[i-1],3)+kb*sqrt(h)*zb[i-1];
        fprintf(cam, "%f\t%f\n",i*h,z[i]);
    }

    fclose(cam);
    FILE *gp; /*los datos guardados en LanOU.dat se grafican en GNU plot*/
    gp = popen(GNUPLOT,"w"); /* 'gp' is the pipe descriptor */
    if (gp==NULL)
    {
        printf("Error abriendo GNU plot");
        exit(1);
    }

    fprintf(gp, "set title 'Langevin Ruido de Color'\n");
    fprintf(gp, "set xlabel 't'\n");
    fprintf(gp, "set ylabel 'z'\n");
    fprintf(gp, "set xrange [0:2500]\n");
    fprintf(gp, "set yrange [0:10]\n");

    fprintf(gp, "plot 'LanOU.dat' t 'z' w lines\n");
    fprintf(gp, "set terminal png\n");
    fprintf(gp, "set output 'LanOU.png'\n");
    fprintf(gp, "replot\n");
    fflush(gp);
    fclose(gp);

    return 0;
}

/* generador de numeros aleatorios con distribución normal*/
double randn (double mu, double sigma) /*http://phoxis.org*/
{
    double U1, U2, W, mult;
    static double X1, X2;
    static int call = 0;

    if (call == 1)
    {
        call = !call;
        return (mu + sigma * (double) X2);
    }

    do
    {
        U1 = -1 + ((double) rand () / RAND_MAX) * 2;
        U2 = -1 + ((double) rand () / RAND_MAX) * 2;
        W = pow (U1, 2) + pow (U2, 2);
    }
    while (W >= 1 || W == 0);

    mult = sqrt ((-2 * log (W)) / W);
    X1 = U1 * mult;
    X2 = U2 * mult;

    call = !call;

    return (mu + sigma * (double) X1);
}

```

Figura 5.5: Código en C para la aproximación de Euler-Maruyama de la ecuación de Langevin con ruido de color (3.34) y (3.33).

```

#include<stdio.h>
#include<stdlib.h>
#include<math.h>

#define GNUPLOT "gnuplot -persist"

double randn(double,double);
int main()
{
    FILE *cam=fopen("entropy.dat", "w"); /*Los datos producidos en la simulación son guardados en entropy.dat*/
    float h,t,k,b0=5.45,b1=4.78,b2=1.17,b3=0.084,z[10001],z1[10001],tau[10001],tau1[10001],tau2[10001],S[10001],sum[10001],sig=1,dz=5.35;
    double mu=0,sigma=1;
    int i,j,N;
    z[0]=z1[0]=1; /*condición inicial para z y z1*/
    tau[0]=b0-b1*z1[0]+b2*pow(z1[0],2)-b3*pow(z1[0],3); /*condición inicial para la tensión*/
    tau1[0]=tau[0]-tau2[0];
    S[0]=0; /*condición inicial para S*/
    sum[0]=tau1[0];
    h=0.25; /*tamaño del paso*/
    t=2500; /*tiempo máximo de la simulación*/
    N=10001; /*numero de pasos a tomar*/
    for(i=2;i<N;i++)
    {
        /*se genera el proceso z mediante la aproximación de Euler-Maruyama*/
        z[i]=z[i-1]+h*b0-h*b1*z[i-1]+h*b2*pow(z[i-1],2)-h*b3*pow(z[i-1],3)+sig*sqrt(h)*randn(mu,sigma);
        if (z[i] < 3 || z[i] > 6){ /*Se excluyen los estados NBA y UBA del proceso z*/
            z[i]=0;
        } else {
            z1[i]=z[i];
        } /*Se sustituye en la tensión efectiva el nuevo proceso sin los estados NBA, UBA*/
        tau[i]=b0-b1*z1[i]+b2*pow(z1[i],2)-b3*pow(z1[i],3);
        if(tau[i] == b0){ /*Se hace cero la fuerza cuando esta es igual a b0*/
            tau1[i]=0;
        } else {
            tau1[i]=tau[i];
        }
        tau2[i]=fabs(tau1[i]);
        sum[i]=sum[i-1]+tau2[i];
        S[i]=(dz* sum[i])/293.15; /*entropía producida*/
        fprintf(cam, "%f\t%f\n",i,h,S[i]);
    }
    fclose(cam);
    FILE *gp; /*los datos guardados en entropy.dat se grafican en GNU plot*/
    gp = popen(GNUPLOT,"w"); /* 'gp' is the pipe descriptor */
    if (gp==NULL)
    {
        printf("Error abriendo GNU plot");
        exit(1);
    }

    fprintf(gp, "set title 'Entropía'\n");
    fprintf(gp, "set xlabel 't'\n");
    fprintf(gp, "set ylabel 'S'\n");
    fprintf(gp, "set xrange [0:2500]\n");
    fprintf(gp, "set yrange [0:8]\n");

    fprintf(gp, "plot 'entropy.dat' t 'S' w lines\n");
    fprintf(gp, "set terminal png\n");
    fprintf(gp, "set output 'entropy.png'\n");
    fprintf(gp, "replot\n");
    fflush(gp);
    fclose(gp);

    return 0;
}
/* generador de numeros aleatorios con distribución normal*/
double randn (double mu, double sigma) /*http://phoxis.org*/
{
    double U1, U2, W, mult;
    static double X1, X2;
    static int call = 0;

    if (call == 1)
    {
        call = !call;
        return (mu + sigma * (double) X2);
    }

    do
    {
        U1 = -1 + ((double) rand () / RAND_MAX) * 2;
        U2 = -1 + ((double) rand () / RAND_MAX) * 2;
        W = pow (U1, 2) + pow (U2, 2);
    }
    while (W >= 1 || W == 0);

    mult = sqrt ((-2 * log (W)) / W);
    X1 = U1 * mult;
    X2 = U2 * mult;

    call = !call;

    return (mu + sigma * (double) X1);
}

```

Figura 5.6: Código en C para la simulación de la entropía producida (3.12).

Bibliografía

- [1] TINOCO, I., BUSTAMANTE, C., *The effect of force on thermodynamics and kinetics of single molecule reactions*, Biophysical Chemistry, 101-102 (2002), pp. 513-533.
- [2] NELSON, P., *Biological Physics: Energy, Information, Life*, W. H. Freeman, 2013.
- [3] HYON, C., *Exploring the Energy Landscape of Biopolymers Using Single-Molecule Force, Spectroscopy and Molecular Simulations*, Simulations in Nanobiotechnology, edited by Kilho Eom, CRC Press, 2012.
- [4] VAN KAMPEN, N. G., *Stochastic Processes in Physics and Chemistry*, Elsevier, 2007.
- [5] DOI, M., EDWARDS, S., *The Theory of Polymer Dynamics*, Clarendon Press Oxford, 1986.
- [6] HÄNGGI, P., JUNG, P., *Colored noise in dynamical systems*, Advances in Chemical Physics, Volume 89, John Wiley & Sons, 1995.
- [7] RUBI, J. M., BEDEAUX, D., KJELSTRUP, S., *Unifying Thermodynamic and Kinetic Descriptions of Single-Molecule Processes: RNA Unfolding under Tension*, J. Phys. Chem., B 2007, 111, pp. 9598-9602.
- [8] COCCO, S., MARKO, J. F., MONASSON, R., *Slow nucleic acid unzipping kinetics from sequence-defined barriers*, Eur. Phys. J. E. 10, 2003, pp. 153-161.
- [9] HYEON, C., MORRISON, G., THIRUMALAI, D., *Force-dependent hopping rates of RNA hairpins can be estimated from accurate measurement of the folding landscapes*, 9604-9609, PNAS, July 15, 2008, vol 105, no. 28.
- [10] NASPREDA, M., REGUERA, D., PÉREZ-MADRID, A., RUBÍ, J.M., *Glassy dynamics: effective temperatures and intermittencies from a two state model*, Physica A 351 (2005), pp. 14-21.
- [11] GALLAVOTTI, G., BONETTO, F., GENTILE, G., *Aspects of Ergodic Qualitative and Statistical Theory of Motion*, Springer-Verlag, 2004.
- [12] NICOLIS, G., PRIGOGINE, I., *Self-Organization in Nonequilibrium Systems*, John Wiley & Sons Inc., 1977.
- [13] SANTAMARÍA-HOLEK, I., *Termodinámica Moderna*, Editorial Trillas, 2014.
- [14] KONDEPUDI, D., PRIGOGINE, I., *Modern Thermodynamics*, John Wiley & Sons Ltd., 1998.

- [15] SERRA, R., ANDRETTA, M., ZANARINI, G., COMPIANI, M., *Introduction to the Physics of Complex Systems: The mesoscopic approach to fluctuations, non linearity and self-organization*, University of Bangalore Press, 1986.
- [16] PAUL, W., BASCHNAGEL, J., *Stochastic Processes*, Springer, 1999.
- [17] HOEL, P.G., PORT, S.C., STONE, C.J., *Introduction to Stochastic Processes*, Houghton Mifflin Company, 1972.
- [18] ZWANZIG, R., *Nonequilibrium Statistical Mechanics*, Oxford University Press, 2001.
- [19] RISKEN, H., *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer-Verlag, 1989.
- [20] SAN MIGUEL, M., SANCHO, J. M., *A Colored Noise Approach to Brownian Motion in Position Space. Corrections to the Smoluchowski Equation*, Journal of Statistical Physics, Vol. 22, No. 5, 1980.
- [21] KONOTOP, V., VÁZQUEZ, L., *Nonlinear Random Waves*, World Scientific, 1998.
- [22] GRAHAM, C., TALAY, D., *Stochastic Simulation and Monte Carlo Methods: Mathematical Foundations Of Stochastic Simulation*, Springer-Verlag, 2013.
- [23] SANTAMARÍA-HOLEK, I., LÓPEZ-ALAMILLA, N., J., HIDALGO-SORIA, M., PÉREZ-MADRID, A., *Non-linear irreversible thermodynamics of small systems*, “enviado”.
- [24] LAPEYRE, B., PARDOUX, E., SENTIS, R., *Introduction to Monte Carlo Methods for Transport and Diffusion Equations*, Springer-Verlag, 1998.
- [25] HIDALGO-SORIA, M., SANTAMARÍA-HOLEK, I., RODRÍGUEZ, R., F., PÉREZ-MADRID, A., *Langevin equation with colored noise for RNA hairpins stressed by LOT*, “en proceso”.