



UNIVERSIDAD NACIONAL AUTONOMA DE
MÉXICO

FACULTAD DE CIENCIAS

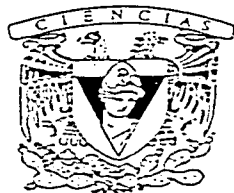
**“UN MODELO MUY SIMPLE DE
DESNATURALIZACIÓN DE ADN”**

T E S I S
QUE PARA OBTENER EL TITULO DE
F I S I C O
P R E S E N T A
ANSELMO GARCIA CANTÚ ROS

DIRECTOR DE TESIS:
DR. HERNÁN LARRALDE RIDAURA

TESIS CON
FALLA DE ORIGEN

2002



FACULTAD DE CIENCIAS
SECCION ESCOLAR



Universidad Nacional
Autónoma de México



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.



24

M. EN C. ELENA DE OTEYZA DE OTEYZA

Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunico a usted que hemos revisado el trabajo escrito:
"Un modelo muy simple de desnaturalización de ADN"

realizado por GARCIA CANTU ROS ANSELMO

con número de cuenta 9561046-7 , quién cubrió los créditos de la carrera de Física.

Dicho trabajo cuenta con nuestro voto aprobatorio

Atentamente

Director de Tesis

Propietario

DR. HERNAN LARRALDE RIDAURA

Propietario

DR. GERMINAL COCHO GIL

Propietario

M. EN B. BENITO ESTRADA MERR

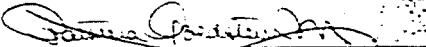
Suplente

DR. GERARDO GARCIA NALMIS

Suplente

DR. RAUL MAURICIO RECHTMAN SCHRENZEL

Consejo Departamental de Física


DRA. PATRICIA GOLOSTEIN MENSEGNE
Coordinadora de Licenciatura

Al ADN de todos los seres.

A mis padres, por lo obvio e inexpressable.

A Nina y al cálido y tierno abrazo que encuentro en el fondo de su existencia.

A Teshin y sus enseñanzas

A Güicha

Agradecimientos

Al terminar este trabajo me he puesto a pensar en la UNAM, más allá de la institución, como un fenómeno tal que, al participar algunos años en él y él en mí, me convertí en un ser un poco más humano y que además ya sabe algo de Física. Gracias a mis compañeros: profesores y alumnos; gracias a todos por compartir conmigo este proceso de mutuo enriquecimiento.

Gracias a mi madre, porque con su apoyo esta experiencia duró lo suficiente. A mi padre, por su invaluable compañía y amistad.

Al profesor Hernán Larralde Ridaura, mi director de tesis, de quien aprendo mucho y quien me ha apoyado y motivado para continuar aprendiendo física, de manera más rigurosa y aterrizada.

Le agradezco también a los sinodales sus sugerencias para mejorar esta tesis.

Gracias al proyecto "Mecánica estadística de procesos sencillos clásicos y cuánticos" de DGAPA por la beca que me otorgó para la realización de esta investigación.

A mi paisano el amigo José Lizardi, gracias por todos estos años de labor en equipo. A Erick Luna, quien me ha ayudado siempre a resolver problemas de cómputo.

Muchas gracias a los que no menciono, por favor inclúyanse todos.

Índice

1	INTRODUCCION	3
1.1	De qué trata esta tesis:	3
1.2	EL ADN Y EL ARN, SU COMPOSICIÓN Y GEOMETRÍA.	6
1.2.1	Estructura del ADN	6
1.2.2	Estructura del ARN.	8
1.2.3	Geometría del ADN: La doble hélice.	9
1.3	LOS POLINUCLEÓTIDOS Y SU FUNCIÓN EN LOS SERES VIVOS.	12
1.3.1	Funciones biológicas del ADN y ARN	12
1.4	PROPIEDADES DEL ADN EN DISOLUCION.	18
2	TERMODINAMICA CLASICA Y MECANICA ESTADISTICA.	22
2.1	Termodinámica	22
2.2	Mecánica estadística	24
2.3	Transiciones de fase.	28
3	MODELADO	32
3.1	Por qué y cómo modelar	32
3.2	Modelos físicos del ADN	35
3.2.1	Modelo de Poland y Scheraga (PS) [9]	37
3.2.2	Variantes del modelo (PS)	39

4	DESCRIPCIÓN DEL MODELO.	44
4.1	Consideraciones del modelo.	44
4.2	Función de partición.	47
4.2.1	Desorden dinámico	48
4.3	Resultados	52
4.4	Discusión	61
5	CONCLUSIONES Y PERSPECTIVAS DEL MODELO	69
6	APENDICE I	71
7	APENDICE II	74
8	APENDICE III	76

Capítulo 1

INTRODUCCION

Con la misma actitud del que quiere subir una montaña sin más instrumento que un mapa, el modelo que aquí se construye, empleando herramienta básica de Mecánica Estadística, pretende hacer un recorrido accesible y educativo por la fenomenología del ADN. En este terreno, recorrer veredas ya trazadas, con la intención de llegar más alto para observar un poco mejor paisajes ya muy fotografiados, hoy en día, implicaría el manejo de herramienta pesada. Entonces, en este trabajo se explora una ruta algo distinta, tomando un rumbo intermedio entre caminos que llevan a contemplar panoramas cualitativos o cuantitativos, delimitada por modelos previos, que aluden a lo uno o a otro, teniendo como meta ver lo antes visto, pero desde un ángulo distinto.

El mecanismo de desnaturalización propuesto es el más simple posible. No obstante, el modelo considera la heterogeneidad de las secuencias de ADN. Esta consideración no se incluye en la mayoría de los modelos de desnaturalización hasta ahora trabajados.

1.1 De qué trata esta tesis:

Las bien conocidas funciones del ADN están estrechamente ligadas a la posibilidad que éste tiene de separarse en sus dos cadenas constituyentes, lo cual permite que la información contenida, en forma de secuencias de bases nitrogenadas, sea interpretada y

reproducida a través de la interacción de éstas con complejos moleculares de proteínas y de enzimas. Por esta razón, un primer paso en la comprensión de los complejos procesos involucrados en la fisiología del ADN consiste en el estudio de los procesos de asociación y disociación del par de cadenas, que éste sufre al ser expuesto a cambios de temperatura y pH, lo cual se conoce como "desnaturalización" [1].

En lo concerniente a sus aplicaciones en la medicina, el estudio de la desnaturalización de ADN despierta interés ya que es empleado en técnicas de diagnóstico tales como la reacción en cadena de la polimerasa (PCR) [2].

Desde un punto de vista físico, la desnaturalización presenta dos características por las cuales resulta de gran interés:

1) Resulta natural abordarlo desde la perspectiva de la Termodinámica Estadística.

Y lo más importante:

2) Es susceptible de ser vista como una transición de fase orden-desorden, presente en algunos sistemas ampliamente estudiados en este contexto.

En primer lugar se estudian los aspectos fisiológicos y estructurales primarios del ADN, para una mejor comprensión de su relación mutua. Posteriormente se revisan modelos ya trabajados, identificando las perspectivas fundamentales con las que se ha abordado el estudio de la desnaturalización, los parámetros involucrados, la herramienta teórica empleada y los diferentes tratamientos propuestos.

En general, es posible reconocer dos actitudes diferentes en el modelado de la desnaturalización que aquí nos ocupa:

algunos modelos hacen énfasis en los aspectos cualitativos del fenómeno y se concentran en elucidar el orden de la transición, sin atender siquiera a los valores de temperatura para los que ésta ocurre.

Por otra parte, otros modelos buscan reproducir, con la mayor precisión posible, el comportamiento observado en experimentos, para lo cual consideran un gran número de parámetros, de los cuales, algunos no expresan un significado físico evidente.

A partir de estas consideraciones se propone un modelo bastante simple en comparación con cualquiera de los previamente revisados.

El estudio trata segmentos cortos de ADN, como los utilizados en las PCR, mismos que se asumen como formados por un par de cadenas totalmente flexibles, con la posibilidad de autointersectarse, pero con un número limitado de ángulos de configuración para cada par de bases; de acuerdo con la teoría de cadenas libres.

En lo concerniente a las interacciones entre nucleótidos, se consideran únicamente a los enlaces de hidrógeno que forman los pares complementarios. Se distinguen dos tipos de enlaces, A-T y C-G, permitiendo que sus concentraciones relativas fluctúen a lo largo de la cadena.

El mecanismo de desnaturalizado se supone ordenado y unidireccional, i.e., debido a la ruptura consecutiva de los enlaces a partir de un sólo extremo. Esto permite evaluar el peso específico del cambio en la entropía, que resulta de la disociación externa de las cadenas, en segmentos cortos de ADN.

Para obtener los parámetros de orden que caracterizan la transición, se calcula el logaritmo de la función de partición, con base en la aproximación de desorden dinámico o recocido (annealed disorder) [3]. La aproximación de desorden dinámico, en este caso, es muy fácil de resolver y adecuada.

Los resultados obtenidos son muy cercanos a los que resultan de algoritmos avanzados de modelado de desnaturalización de ADN, tales como el algoritmo de Steger [4].

Se concluye que, para cadenas cortas, la entropía que domina la transición se podría deber a la disociación de las cadenas a partir de sus enlaces.

Se omiten parámetros tales como la interacción a primeros vecinos [5], la rigidez de las cadenas, [6] y su autoevitancia [7] y [8]; sin embargo, los resultados son suficientemente buenos y exhiben una transición de primer orden que concuerda con el modelo de Poland [9].

1.2 EL ADN Y EL ARN, SU COMPOSICIÓN Y GEOMETRÍA.

En el año de 1953, Francis Crick junto con James Watson, reconstruyó la estructura de doble hélice de la macromolécula del ADN mediante la técnica de refracción de rayos X. En el proceso, lograron deducir la manera en que la información genética se almacena y se duplica [10]. Ya que la relación entre estructura, geometría y función de los ácidos nucleicos, ADN y ARN, es de enorme importancia [11] para una comprensión completa de estas enigmáticas moléculas, y que en dicha relación se ubica la importancia del estudio del fenómeno de desnaturalización, a continuación doy un resumen de su estructura y geometría, así como de algunas de las funciones que desempeñan en los seres vivos.

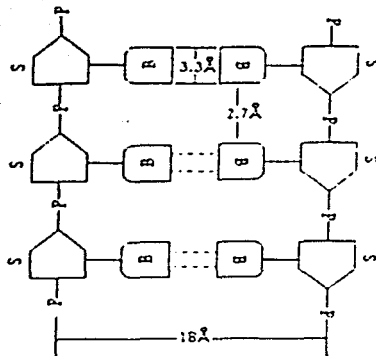
1.2.1 Estructura del ADN

Estructura secundaria

A partir de átomos de carbón, fósforo, hidrógeno, nitrógeno y oxígeno se forman tres agrupaciones. Un azúcar de 5 carbonos (desoxirribosa), un grupo fosfato y cuatro tipos de compuestos de nitrógeno con propiedades base, figura 1.2.1 a), que se clasifican como Purinas y Pirimidinas.

El compuesto que se forma al unirse estos tres grupos, como se indica en la figura 1.2.1 b), se llama nucleótido. La unión sucesiva, mediante enlaces covalentes de grupos fosfato en el carbón denotado, en la figura 1.2.1 b), por 5' de ciertos nucleótidos con los hidroxilos (OH) en el carbón 3' de otros, forma cadenas flexibles 5' - 3' con longitudes macroscópicas (milímetros en el caso de bacterias) llamadas polinucleótidos [12], ver figura 1.2.2.

La estructura de los polinucleótidos es tal que éstos pueden enlazarse entre sí a través de las bases nitrogenadas, figura 1.2.3, formando pares de cadenas que recuerdan a las



**TESIS CON
FALLA DE ORIGEN**

Ver figura 1.2.2

Figura 1.2.1 a) Bases nitrogenadas: Adenina A y Guanina G (Purinas), Citosina C y Timina T (Pirimidinas) [1]

vías de un ferrocarril figura 1.2.3, cuya geometría estable corresponde a una doble hélice [12].

El apareamiento de bases nitrogenadas se logra a partir de la posibilidad química de estas moléculas para formar enlaces de hidrógeno entre sí. Sin embargo, no cualquier tipo de combinación de bases da como resultado un apareamiento apropiado. Las uniones estables son exclusivamente las complementarias que a continuación se definen [13], junto con sus configuraciones simétricas:

A-T.- Formado por dos enlaces de hidrógeno, uno entre los grupos keto y amina ($O:::H-N$) de la purina y pirimidina respectivamente y el otro entre los átomos de nitrógeno ($N-H:::N$) de los anillos.

G-C.- Tres enlaces, dos entre los grupos keto y amina y uno entre los átomos de nitrógeno de los anillos.

En cuanto a los enlaces de hidrógeno ($O:::H-N$) y ($N-H:::N$), éstos tienen las siguientes características: lineales en su arreglo y pueden considerarse covalentes (guión) por un lado y Coulombianos por el otro (punteado). La importancia de este tipo de enlaces es que son relativamente débiles, comparados con los covalentes que unen a los nucleótidos,

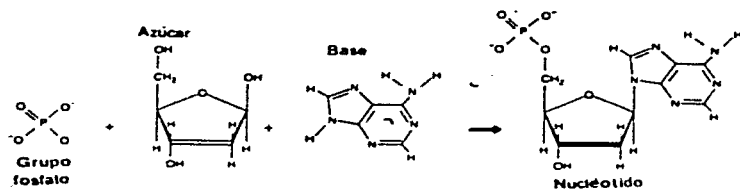


Figura 1.2.1 b) La constitución de un nucleótido [1].

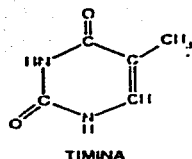
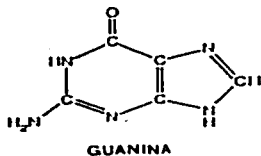
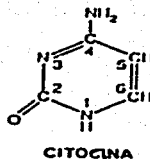
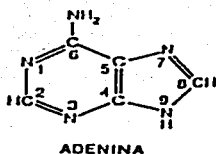
lo que permite que se rompan y se vuelvan a formar a temperatura ambiente, y esto, como se verá más adelante, es una condición fundamental para que se realicen las funciones del ADN. Un análisis detallado de éstos se encuentra en [13].

Los pares A-T y C-G no se presentan en proporciones iguales [14]; de hecho

$[A+T] = K[C+G]$, donde el cociente K de estas concentraciones es característico de cada organismo y, como veremos más adelante, esta relación entre concentraciones de bases complementarias es muy importante en el estudio de la desnaturalización [1]. Por otra parte, como consecuencia del carácter complementario de las bases se cumplen las semi igualdades $A/T \approx G/C \approx 1$.

1.2.2 Estructura del ARN.

La composición del ARN [1], en general, es casi idéntica a la del ADN: si se sustituye el azúcar desoxirribosa por una ribosa (igual a la desoxirribosa pero con un grupo hidroxilo en el vértice inferior derecho) y en lugar de bases timina se colocan bases uracilo, se obtiene un ribonucleótido, con la misma posibilidad de un desoxirribonucleótido para formar enlaces covalentes 5'-3' con otros de su misma especie, dando lugar a polinucleóti-



→ Ver figura 1.2.1.

Figura 1.2.2 Esbozo de cadena de 3 nucleótidos, longitudes y direcciones entre pares; P, S y B, fósforo, Azúcar y Base.

dos de ARN. Lo peculiar del ARN es que está formado por una cadena simple, pero cuya composición permite que sus bases complementarias se enlacen entre sí, dando al ARN geometrías relacionadas con aspectos fisiológicos.

1.2.3 Geometría del ADN: La doble hélice.

Una característica importante del par de cadenas es que poseen direcciones antiparalelas; esto es, una en la dirección 5' - 3' y la otra en la 3' - 5', figuras 1.2.2 y 1.2.4. Un arreglo distinto de enlaces entre pares de bases, que daría lugar a direcciones paralelas entre cadenas, propuesto por Hoogsteen [12], sólo sería estable bajo condiciones de pH ácido y el interior de las células de los organismos vivos consiste en un medio acuoso neutro.

En 1954 Watson y Crick encontraron que la forma para el ADN, en mayor concordancia con las características observadas en experimentos de difracción, era un par de cadenas antiparalelas con geometría helicoidal [12], envolviéndose una a la otra formando dos surcos: uno corto de 3.4nm que separa bases vecinas y uno largo de 3.4nm, que corresponde a un giro completo de cada una de las cadenas (10 pares de bases), con un

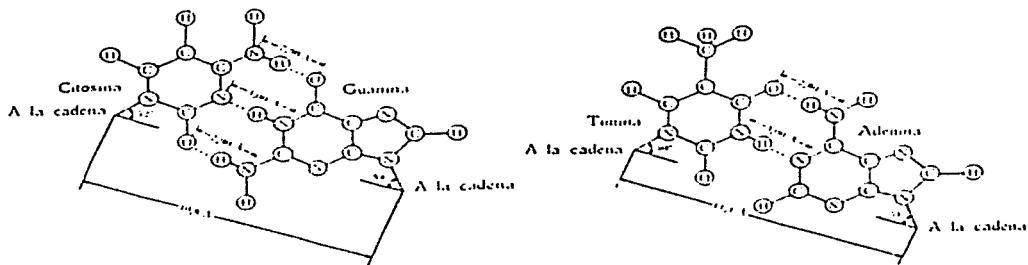


Figura 1.2.3 Apareamiento de bases complementarias, tomada de [1].

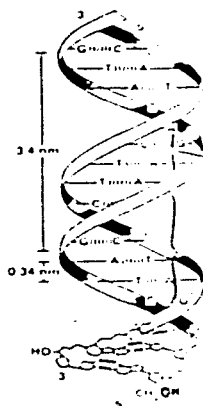
diámetro externo de 18 μ m. como se muestra en la figura 1.2.4.

Esta forma ha sido ampliamente verificada. Pero si buscamos el por qué de esta peculiar geometría, deberíamos atender a la insolubilidad o hidrofobicidad de las bases nitrogenadas.

Para que una estructura molecular hidrofóbica sea estable en una solución acuosa con pH neutro, es necesario que se encuentre en el centro de una estructura envolvente que evite el contacto con el agua [12]. Esto sucede en la conformación de las uniones entre las cadenas del ADN. Las bases se encuentran en el centro de una estructura formada por los azúcares que son hidrofílicos. Ahora, si observamos la figura 1.2.2, encontraremos un hueco de ancho 2.7 Å entre bases, que por lo antes expuesto, no será ocupado por agua, lo cual determina, únicamente, dos configuraciones como posibles [12]: una consiste en que las cadenas se inclinen hacia un lado, como si fuera una escalera, reduciendo al máximo el hueco entre bases y la otra en que las cadenas giren dentro de una superficie cilíndrica, como realmente sucede.

La explicación detallada se puede ver en [12] y de ella se deducen los ángulos y las distancias medidos en experimentos.

La descripción anterior de la forma doble helicoidal del ADN se conoce como forma B; sin embargo no es la única, ya que el ángulo entre base y desoxirribosa puede cambiar;



TESIS CON
FALLA DE ORIGEN

Figura 1.2.4 Doble hélice y sus dimensiones [1].

además, los eslabones fosfato-desoxirribosa son lo suficientemente flexibles como para permitir configuraciones alternativas [1], ver figura 1.2.5.

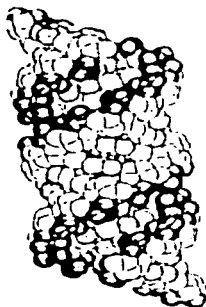
La forma A existe únicamente en estados deshidratados. El plano en el que se encuentran los pares de bases forma un ángulo de 20° con la perpendicular al eje de la hélice, lo que disminuye la distancia entre pares a 0.29nm , incrementándose a 11 o 12 pares de bases por vuelta. Hasta ahora no se le atribuye ninguna función biológica a esta forma.

La forma B posee quiralidad diestra. Bajo ciertas condiciones se observa que algunos segmentos de ADN, en los cuales purinas y pirimidinas se alternan, presentan una conformación levógiro con espaciamiento entre bases de 0.77nm y 12 pares de bases por vuelta. A esta conformación se le conoce como hélice Z. Las condiciones naturales que favorecen este tipo de conformación son todavía objeto de estudio [1].

El ADN también se presenta como una molécula circular. El ADN en plásmidos y cromosomas bacteriales, en cloroplastos, varios ADN mitocondriales y genomas de virus de mamíferos, consiste en una molécula simple de ADN circular unida covalentemente[1].



B-DNA



A-DNA



Z-DNA

Figura 1.2.5 Tipos de doble hélice [1].

1.3 LOS POLINUCLEÓTIDOS Y SU FUNCIÓN EN LOS SERES VIVOS.

Desde el punto de vista biológico, el estudio del proceso de desnaturalización del ADN es un primer paso en la comprensión de los procesos, enormemente más complejos, implicados en la ejecución de las funciones vitales realizadas por los polinucleótidos dentro de los seres vivos [1]; entonces, antes de explicar en qué consiste dicho fenómeno, conviene mencionar las funciones que con él se relacionan.

1.3.1 Funciones biológicas del ADN y ARN

El ADN como archivo de información

EL ADN, presente en toda célula, almacena en código toda la información necesaria para la realización de los procesos [1], que dan lugar a la formación de estructuras ascendentes

en la escala de complejidad y tamaño (células, tejidos, órganos, aparatos, sistemas). De hecho, el ADN y sus interacciones con otras moléculas, tales como el ARN y proteínas, constituyen una unidad funcional con las propiedades de autoreplicamiento y autoregulación, que entre otras, intuitivamente reconocemos como aquéllas que definen a un organismo vivo.

Una forma de interpretar cómo se encuentra guardada la información en el ADN es imaginándola como un texto unidimensional: un solo renglón escrito con un alfabeto de cuatro letras: A, T, C, G. A la unidad mínima de significado o "palabra" se le llama gen, y a las subestructuras de estas palabras se les llaman codones, que son secuencias de tres "letras" o nucleótidos [1].

No en todos los organismos las cadenas completas contienen pura información, sólo en los unicelulares o Procariotes. En los Eucariotes u organismos pluricelulares, sólo una fracción pequeña de la secuencia total contiene información codificada (en el humano únicamente el 3%) [15]. Dentro de las regiones codificadas, llamadas Exones, se puede observar una enorme irregularidad en las secuencias. Lo que tal hecho resalta es, según la teoría de la información, la gran complejidad del código genético [16], pues la cantidad de información que se requeriría para definir una versión abreviada de cualquier secuencia genética sería casi igual a la de la secuencia misma; i.e., las secuencias de genes son irresumibles. Esto contrasta con el caso de los intrones, o regiones no codificadas, donde las secuencias de bases presentan regularidades. El análisis de regularidades en las secuencias genéticas se ha podido realizar mediante la aplicación de las técnicas de la Física estadística que ofrecen una metodología para el estudio de secuencias en general [17], [18] y [19].

El ARN y la transcripción e interpretación del código genético [1]

En cuanto a la decodificación de la información del ADN, ésta se da mediante la construcción de una versión legible o interpretable del ADN: el ARN.

El ARN, con la información de alguna secuencia de ADN, se construye a partir de la

acción de la Holoenzima ARN Polimerasa, que abreviaremos como H-ARN-P¹.

Ahora bien, cuando la H-ARN-P localiza, o mejor dicho, interacciona con un codón específico de ADN, cuya información indica el comienzo de una transcripción, ésta rompe los enlaces de Hidrógeno a partir de 35 nucleótidos anteriores al que comienza la transcripción. A continuación, al quedar así expuesto el interior de la cadena con la información a transcribir y el de su complementaria, la ARN Polimerasa cataliza la reacción en la que, haciendo las veces de plantilla, el primer nucleótido de la secuencia complementaria se enlaza con un ribonucleótido-trifosfato con base nitrogenada idéntica a la de la secuencia a reproducir (excepto en el caso de Timina ya que es reemplazada por Uracilo). La ARN Polimerasa se desplaza, mediante mecanismos aún no bien entendidos^[1] en la dirección de la cadena en transcripción, y con ella el proceso de separación de cadenas y secuenciamiento de ribonucleótidos, cuyo orden es el de la cadena leída en el ADN. A medida que la ARN Polimerasa continúa en su desplazamiento, las cadenas disociadas de ADN se reenlazan y se recupera la forma original de doble hélice.

El eslabonamiento sucesivo de ribonucleótidos se da paso a paso, al unirse el grupo OH del carbón 3' de cada ribonucleótido con el grupo fosfato en el carbón 5' del ribonucleótido-trifosfato vecino. Los dos grupos fosfatos restantes se disocian.

El ARN en el que se transcribe el código del ADN recibe el nombre de mensajero o m-ARN.

El responsable de la lectura y materialización del mensaje es otro ARN. El tipo de ARN encargado de dicha labor se conoce como traductor o t-ARN; una cadena que, enlazándose consigo misma en uno de sus extremos, da lugar a una conformación de trébol, figura 1.3.1, con 3 extremidades terminando en un lazo. En el otro extremo, algún tipo de aminoácido² se conecta con la terminación 3' del ARN. El tipo de aminoácido, para cada t-ARN, queda determinado por el anticodón presente en el lazo del extremo

¹Una enzima es una estructura catalítica cuya geometría es tal que un determinado proceso, al llevarse a cabo sobre la superficie de ésta, es restringido en los grados de libertad de sus componentes, de lo cual resulta un aumento en la probabilidad de que las reacciones constituyentes de dicho proceso se lleven a cabo.

²Compuesto formado por un grupo carboxilo y un amino de estructura distintiva.

opuesto. Los anticodones del t-ARN se acoplan con los codones del m-ARN, figura 1.3.1; esto traduce secuencias de codones del m-ARN a secuencias de aminoácidos que, eslabonándose entre sí, forman cadenas llamadas proteínas, cuyos plegamientos dan lugar a estructuras de múltiples geometrías.

Los procesos de traducción de m-ARN a proteínas se catalizan dentro de una macroestructura llamada Ribosoma, formada por 50 tipos de proteínas y por 3 de ARN.

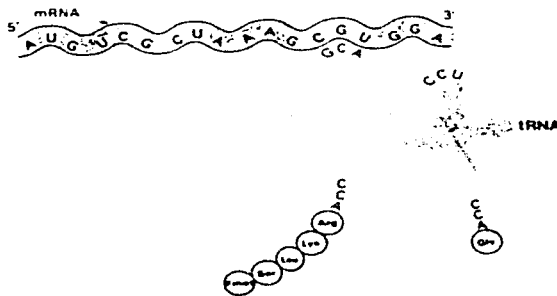


Figura 1.3.1, los codones del m-ARN son leídos de uno en uno en la dirección 5'-3', a través de su enlace con el t-ARN, partiendo del codón iniciador, AUG [1].

Podemos decir que los procesos de lectura y traducción a proteínas del ADN son autoregulados, tomando en cuenta que dichos procesos son a su vez controlados por proteínas que, a través de sus peculiares geometrías, interactúan con la doble hélice. Estas, al adherirse a ciertas secuencias, impiden que puedan ser leídas por obstaculizar el desplazamiento de la polimerasa, además de controlar la longitud de las cadenas del ARN en formación. Tal tipo de proteínas se conocen como represoras. Por otra parte, la presencia de ciertas proteínas sobre secuencias específicas puede favorecer su lectura, en cuyo caso se habla de proteínas activadoras.

Siguiendo la analogía con la lectura de un texto, podemos observar que lo que da lugar al significado completo de una oración, no es únicamente la suma de palabras, sino su mutua interacción, que obedeciendo a reglas fijas, determinará, según la posición relativa de las palabras, la función que cada una desempeña dentro de la oración. En el caso del ADN, las interconexiones semánticas entre genes, no son fijas y se establecen y varían a medida que las sucesivas lecturas se realizan.

El ADN se replica

En 1953 Watson y Crick escribieron lo siguiente: "Si el orden de las bases sobre una del par de cadenas fuera dado, uno podría dar el orden exacto de las bases sobre la otra, debido al apareamiento específico de las bases. Una cadena es el complemento de la otra, y ésta es la propiedad que sugiere como el ADN podría duplicarse a sí mismo".

De manera muy similar a como sucede en la construcción del ARN, al romperse los enlaces de hidrógeno que mantienen juntas al par de cadenas del ADN, la doble hélice se despliega y las secuencias de bases quedan expuestas, con posibilidad de servir como moldes para la formación de secuencias idénticas a sus complementarias; de esta manera, las secuencias son sintetizadas a partir de la adición, sobre secuencias de ADN específicas (orígenes de replicación), de unidades 5'-desoxinucleótido de Desoxinucleósido-trifosfatos al extremo 3'-hidroxilo de un segmento de cadena ya existente (*iniciador*). Consecuentemente, las cadenas de ADN crecen en la dirección 5'-3' a lo largo de la cadena molde con orientación 3'-5'; por lo tanto, los procesos de replicamiento del par de cadenas se desplazan en sentidos opuestos. En la síntesis de cadenas se observa que el replicamiento de una de las cadenas es continua y la otra es segmentada, figura 1.3.2; este proceso es catalizado por la ADN-polimerasa.

Estructura terciaria del ADN

El resumen anterior de los procesos de transcripción y replicación del ADN sobresimplifica lo que ocurre al interior de los seres vivos, ya que el objeto de esta tesis no es estudiar

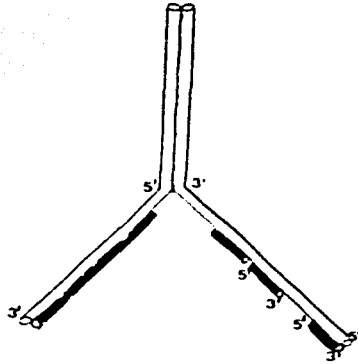


Figura 1.3.2 Replicamiento de cadenas de ADN [1].

estos aspectos; sin embargo los hemos mencionado con la finalidad de situar, de manera más amplia, la importancia que tiene, desde un punto de vista funcional, la capacidad del ADN para separar al par de cadenas constituyentes. Aún cuando el nivel de descripción del proceso de desnaturalización, que aquí hacemos, se basa en la estructura secundaria del ADN, por completéz y con la intención de que se capte que tan complejos son en realidad estos procesos de transcripción y replicamiento del ADN, a continuación doy una muy resumida descripción de la estructura terciaria.

La enormemente larga estructura doble helicoidal del ADN, contenida en el interior del núcleo de la célula presenta una conformación que minimiza el volumen ocupado por el ADN dentro del núcleo y que permite que las funciones biológicas del ADN se realicen. A esta estructura global del ADN se le conoce como cromosoma.

El plegamiento del ADN se lleva a cabo, en primer lugar, en su interacción con proteínas de forma cilíndrica de diámetro 11nm, conocidas como histonas octaméricas, sobre cuyas superficies el ADN se enrolla dos veces, dando lugar a una unidad llamada nucleosoma. Esta unidad, repitiéndose cada 200 pb a lo largo del ADN, forma una estructura lineal llamada cromatina, que recuerda a un hilo con cuentas. A su vez

la cromatina se pliega en forma de hélice cuyas vueltas, de 30 nm de diámetro, están constituidas por seis nucleosomas.

La fibra de cromatina helicoidal forma lazos, con un perímetro máximo de 600nm, que agrupan de 20000 a 100000 pb. El Conjunto de lazos forma, una vez más, una hélice de 700nm de diámetro: cuya estructura global se conoce como cromosoma. El proceso de plegamiento del ADN es esquematizado en la figura 1.3.3.

Recapitulando: la estructura microscópica del ADN le confiere la forma de doble hélice, con la propiedad de alternar, a temperatura ambiente, entre esta conformación y un estado de disociación en el que las cadenas se comportan como simples e independientes. Sin la posibilidad de llevar a cabo tal transición (desnaturalización) no hay entendimiento plausible de cómo el ADN podría realizar sus principales funciones.

1.4 PROPIEDADES DEL ADN EN DISOLUCION.

Los grupos fosfato que forman los puentes entre nucleótidos adyacentes, tienden poco a disociarse y para pH mayor a cuatro se ionizan totalmente, lo que significa que el ADN es muy ácido. Por otra parte, los grupos fosfato, se combinan fuertemente con cationes divalentes tales como Mg^{2+} y Ca^{2+} o con aminas policationicas, éstas al enlazarse con el interior de los surcos mayores estabilizan a la molécula, lo que se traduce en un incremento de la flexibilidad.

La propiedad de las bases de enlazar hidrógeno depende de su forma iónica y por lo tanto del pH. La estabilidad del ADN es función del pH de la solución; la estabilidad máxima se alcanza dentro de los límites fisiológicos de pH (entre 4.0 y 11.0): fuera de estos límites el ADN es inestable y sus polinucleótidos se separan. Esto mismo sucede en presencia de temperaturas mayores a 90°C, compuestos orgánicos como urea o formaldehído y disminución de la constante dieléctrica, por la incorporación de alcoholes y cetonas, en el medio acuoso. Este fenómeno se conoce como desnaturalización y si el cambio en los parámetros de control, T y pH, es suficientemente suave el proceso será

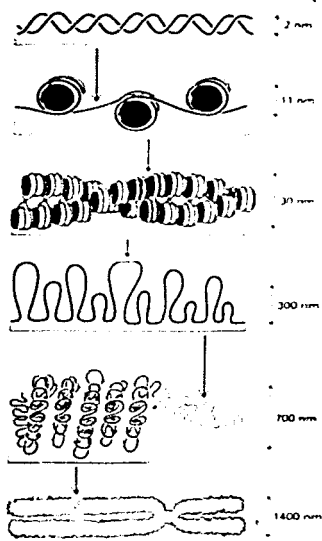


Figura 1.3.3 Esquema de los ordenes de empaquetamiento de la cromatina, que dan lugar a un cromosoma.

reversible. Al proceso inverso se le conoce como renaturalización.

Al aumentar los parámetros de control se alcanza un punto T_m o pH_m , conocido como punto de fusión (melting), para el cual el estado asociado de las cadenas de ADN deja de ser estable. Se considera que las cadenas han alcanzado dicho punto, cuando la mitad de sus nucleótidos constituyentes se encuentra disociada. El punto de fusión es en cierta manera característico del contenido de cada secuencia; esto es claro si se considera la diferencia en la energía de enlace entre pares AT y CG, pues en tanto el primero está formado por un par de puentes de hidrógeno, el segundo posee tres, lo que

lo hace más estable al aumento de temperatura. Así, el punto de fusión de una secuencia será más alto en la medida en que lo sea su contenido de pares CG, de donde se sigue que una determinación precisa del punto de fusión da información de la composición de la secuencia. La forma más usual de aproximar el valor del punto de fusión para una secuencia es ubicar la temperatura correspondiente al punto medio de la gráfica del número de pares disociados (o asociados) v.s. temperatura; este tipo de gráfica recibe el nombre de curva de fusión. La importancia de tal tipo de curvas es que éstas permiten inferir la distribución de pares de bases. En las curvas de fusión, para secuencias heterogéneas, se observan cambios abruptos en la pendiente, correspondientes a regiones con alto contenido en AT, mediados por secciones de pendiente cercana a cero, a su vez debidas a concentraciones dominantes de CG, mostrando un perfil escalonado. De esta forma, la distribución de escalones revela la distribución de cambios en la densidad promedio de ambos tipos de pares.

El método experimental comúnmente usado en la obtención de curvas de fusión, se basa en el efecto de hipercromicidad:

Al hacer incidir un haz de luz, con longitud de onda de 260nm, a través de una solución de ADN, se observan incrementos en la absorción de luz, como función del aumento en la temperatura. La capacidad de los nucleótidos para absorber longitudes de onda de 260nm, es 40% menor si éstos se encuentran apareados. Entonces, las curvas de fusión se obtienen al graficar el porcentaje de incremento de absorción lumínica como función de la temperatura, aumentando a medida que las cadenas se desnaturalizan.

El estudio del fenómeno de desnaturalización es de interés por sus aplicaciones dentro de la genética médica.

Un ejemplo de aplicación es la Reacción en Cadena de la Polimerasa (PCR):

Esta técnica tiene como finalidad producir copias, de manera controlada, de una secuencia dada.

La PCR consiste en ciclos de inducción artificial del proceso de replicamiento antes expuesto.

A grandes rasgos, cada ciclo se compone de tres pasos:

Por simplicidad supongamos una sola cadena en solución acuosa, a la que se le agregan concentraciones de los cuatro tipos de nucleótidos, de *iniciadores*, y de enzimas ADN-polimerasa.

Paso 1. Se aumenta la temperatura de la muestra a más de 94°C , con lo que la secuencia se desnaturaliza.

Paso 2. Se disminuye gradualmente la temperatura a alrededor de 56°C ; esto da lugar a la unión de un *iniciador* con cada una de las cadenas ya separadas.

Paso 3. Se aumenta la temperatura a 72°C , con esto, una ADN-polimerasa comienza ensamblando un nucleótido con la base adecuada al extremo $5'$ del *iniciador*. El proceso es continuado en dirección $3'$ sobre el par de cadenas, hasta que finalmente la cadena original es reconstruida, y además, se obtiene una réplica exacta de la secuencia original.

El proceso continúa con un segundo ciclo, que al finalizar habrá producido dos réplicas. Así, tras la ejecución de N ciclos, la secuencia habrá sido reproducida 2^{N-1} .

Este método es usado en la detección de subsecuencias que pueden estar relacionadas con patologías [2].

Capítulo 2

TERMODINAMICA CLASICA Y MECANICA ESTADISTICA.

A continuación, se hace un breve resumen de los conceptos de Termodinámica y Mecánica Estadística que serán utilizados en capítulos posteriores. Este resumen se basa en lecturas hechas en [21], [22], [23] y [24].

2.1 Termodinámica

La Termodinámica Clásica describe fenómenos macroscópicos de sistemas cerrados, i.e., sistemas cuya interacción con el medio es controlable y por lo tanto las variables que definen su estado pueden ser medidas por el observador. En sentido estricto, hablar del estado de un sistema se refiere a la especificación del estado de cada una de las partículas que lo componen. Aún en sistemas simples, en el límite cuando el número de partículas tiende a infinito o "límite termodinámico", los microestados evolucionan con tal rapidez y complejidad que el registro del microestado de una sola partícula resulta imposible. De aquí que la Termodinámica Clásica, exclusivamente, describa los estados macroscópicos estáticos de un sistema, cuyas propiedades no dependan de las fuerzas externas que hayan actuado sobre éste.

Los estados estáticos se conocen como estados de equilibrio y la Termodinámica Clásica postula que cualquier cambio de estos estados puede ser totalmente descrito en términos de su energía interna, que llamaremos U , función de las variables de estado: V volumen, número de partículas N y una variable S que más abajo definiremos. $U(S, V, N)$.

Para producir un cambio en el estado de un sistema termodinámico es necesario considerarlo en contacto con un conjunto de sistemas fuente o sistemas cuyas variables de estado el observador mantiene constantes. Los sistemas baño, aislados entre sí, interactúan, intercambiando energía y/o materia con el sistema en estudio, al retirar los aislantes físicos que los separan.

Gracias a que en los sistemas termodinámicos existe una tendencia natural a evolucionar hacia estados de equilibrio, todo cambio de estado, una vez alcanzado el equilibrio, se puede cuantificar evaluando el cambio en la energía interna. Formalmente, tal tendencia se representa postulando la existencia de una función S llamada entropía, función de variables extensivas o dependientes de la escala a la cual se observa el sistema, tales como U , V , y N y con las propiedades de:

a) Aditividad: la entropía total S de un sistema formado por i subsistemas, es la suma de las entropías S_i de los subsistemas

$$S = \sum_i S_i \quad (2.1)$$

b) Continuidad y diferenciabilidad.

c) Homogeneidad, y monótona creciente con respecto a la energía interna:

$$\lambda S = S(\lambda U, \lambda V, \lambda N) \quad (2.2a)$$

$$\frac{\partial S(U, V, N)}{\partial U} \geq 0 \quad (2.3)$$

Las condiciones exigidas a S implican su invertibilidad en U , i.e., $U(S, V, N)$ existe y las derivadas

$$\frac{\partial U}{\partial S} = T \geq 0 \quad (2.4)$$

$$\frac{\partial U}{\partial V} = -P \quad (2.5a)$$

se definen como temperatura y presión, respectivamente, donde

$$\Delta Q = T \Delta S \quad (2.6)$$

es el cambio de energía en forma de calor debido a un cambio en la entropía y

$$\Delta W = P \Delta V \quad (2.7a)$$

es el cambio en la energía en forma de trabajo que se obtiene o se realiza sobre el sistema al modificar su volumen.

Para todo proceso, la Termodinámica afirma que:

I) La energía se conserva:

$$dU = dQ + dW, \quad (2.8a)$$

i.e., el cambio en la energía interna se debe a flujos de calor y al trabajo.

II) Los cambios en las variables U , V y N suceden en forma tal que, en el estado de equilibrio final, la entropía alcanza un valor máximo.

2.2 Mecánica estadística

Desde un punto vista microscópico, tenemos que a cada macro estado de equilibrio le corresponde un número Ω de micro configuraciones y la Mecánica Estadística postula

que el sistema transita con igual probabilidad entre tales microestados.

Supongamos un sistema en equilibrio dividido por una pared que mantiene a dos subsistemas a y b con distintas cantidades de partículas y energía, y por lo tanto en dos macroestados diferentes U_a y U_b , a los que les corresponden un número Ω_a y Ω_b de microestados, respectivamente. Si se elimina instantáneamente la pared que los divide, el sistema evolucionará hasta alcanzar un nuevo estado de equilibrio, con microconfiguraciones antes no accesibles y por lo tanto con una Ω mayor. Esto revela cierta relación entre el número de microestados posibles y la entropía del sistema, ya que un aumento en Ω se manifiesta a escala macroscópica como un aumento en S . De hecho, L. Boltzmann, uno de los padres de esta teoría, propuso que

$$S = k_B \ln \Omega, \quad (2.9)$$

con k_B la constante de Boltzmann. Considerando que

$$\Omega = \Omega_a \Omega_b, \quad (2.10a)$$

entonces

$$S = k_B \ln \Omega_a \Omega_b = S_a + S_b, \quad (2.11)$$

lo cual concuerda con el postulado de aditividad de la Termodinámica que se expresa en 2.1.

Por otra parte, un estado de equilibrio corresponde al conjunto más probable de microestados equivalentes. Para construir la probabilidad de un microestado supongamos un sistema baño, al cual le mantenemos fija la temperatura. Si observáramos un elemento de volumen, delimitado por paredes que impidan el flujo de partículas, pero no de calor, encontraríamos que dicho subsistema intercambia energía con sus alrededores. Estas fluctuaciones impiden determinar la energía de nuestro subsistema, así que definir su estado requiere hacerlo en términos de probabilidades.

La probabilidad f_j de que el subsistema se encuentre en el microestado j (con energía E_j)

$$f_j = \frac{1}{\Omega_j},$$

de 2.10a, vemos que es la probabilidad de que el sistema baño tenga energía $E_{tot} - E_j$:

$$f_j = \frac{\Omega_{baño}(E_{tot} - E_j)}{\Omega_{tot}(E_{tot})}. \quad (2.12)$$

donde E_{tot} es la energía del sistema compuesto baño + subsistema, con un número total Ω_{tot} de estados y $\Omega_{baño}(E_{tot} - E_j)$ el número de estados accesibles para el sistema fuente cuando el subsistema está en el estado j .

Por 2.9, tenemos que podemos escribir a 2.12 como

$$f_j = \frac{\exp\{k_B^{-1} S_{baño}(E_{tot} - E_j)\}}{\exp\{k_B^{-1} S_{tot}(E_{tot})\}}. \quad (2.13)$$

Si definimos a U como la energía promedio del subsistema sobre todas las configuraciones j

$$U = \langle E_j \rangle, \quad (2.14)$$

entonces, considerando la aditividad de la entropía

$$S_{tot}(E_{tot}) = S(U) + S_{baño}(E_{tot} - E_j), \quad (2.15)$$

que, expandiendo a $S_{baño}(E_{tot} - E_j) = S_{baño}(E_{tot} - U + U - E_j)$ alrededor del punto de equilibrio $E_{tot} - U$, resulta

$$S_{baño}(E_{tot} - E_j) = S_{baño}(E_{tot} - U) + \frac{(U - E_j)}{T}. \quad (2.16)$$

De introducir 2.15 y 2.16 en 2.13, tenemos que

$$f_j = e^{\frac{1}{k_B T} (U - TS(U))} e^{-\frac{1}{k_B T} E_j} \quad (2.17)$$

donde

$$F(T, V, N) = U - TS(U) \quad (2.18)$$

es una función conocida como energía libre que, conteniendo toda la información del sistema, permite una formulación de la termodinámica alternativa a la de la función U^1 , con

$$dF = -SdT - PdV \quad (2.19)$$

equivalente a 2.8a.

Entonces, sustituyendo 2.18 en 2.17, finalmente tenemos una expresión para la probabilidad de que el subsistema se encuentre en el estado j :

$$f_j = e^{\beta F} e^{-\beta E_j} \quad (2.20a)$$

con $\beta = \frac{1}{k_B T}$, donde $e^{\beta F}$ resulta ser un factor de normalización independiente del estado:

$$\sum_j f_j = e^{\beta F} Z = 1, \quad (2.21)$$

con

$$Z = \sum_j e^{-\beta E_j}. \quad (2.22)$$

A Z se le llama función de partición canónica y su importancia radica en que si contamos con un modelo microscópico del subsistema del cual podamos derivar una expresión

¹ La deducción termodinámica formal de F se puede estudiar en [24].

para E_j , a partir de ella obtendremos la termodinámica del subsistema. ya que todas las cantidades termodinámicas se obtienen a partir de ella como sigue:
energía libre

$$F = -k_B T \ln Z, \quad (2.23)$$

ya que de 2.21

$$Z = e^{-\beta F}. \quad (2.24a)$$

la energía interna

$$U = \langle E_j \rangle = \frac{\sum E_j e^{-\beta E_j}}{Z} = -\frac{\partial \ln Z}{\partial \beta}, \quad (2.25)$$

de derivar 2.18 la entropía

$$S = -\frac{\partial F}{\partial T}. \quad (2.26a)$$

y el calor específico

$$C_v = \frac{\partial U}{\partial T} = -T \frac{\partial S}{\partial T}. \quad (2.27a)$$

Debido a que en su gran mayoría se trabaja con modelos mecánicos, en general, la energía de los microestados se expresa como la función de Hamilton H .

2.3 Transiciones de fase.

Hemos dicho que la evolución de los estados de un sistema aislado está determinada por su tendencia a incrementar la entropía; así que podemos entender que los estados más estables serán aquellos en los que la entropía es máxima; en cambio, para sistemas que intercambian energía con sus alrededores, se presenta una competencia entre los

términos U y S de 2.18. que, para una temperatura dada, serán tales que en los estados de equilibrio estable la energía libre será la mínima². De acuerdo con esta idea intuitiva de estabilidad, en el equilibrio, las propiedades estructurales microscópicas que minimizan a la función F , determinan a las propiedades termodinámicas del sistema a través de las derivadas de F , como se observa en 2.25, 2.26a y 2.27a. En termodinámica, al conjunto de propiedades de un sistema se le conoce como fase y por lo anterior, podemos afirmar que para toda fase las derivadas de F existen y son continuas; sin embargo, en general, esto no necesariamente se cumple para todo valor de las variables de estado, pues en el límite termodinámico, para ciertos valores se pueden presentar discontinuidades en las derivadas de orden n de F , dando lugar a divergencias en las de orden $m > n$. Desde el punto de vista físico, la presencia de tales discontinuidades significa que, para un mismo valor de energía libre, el sistema puede manifestar, con misma probabilidad, más de un conjunto de propiedades. A tales cambios cualitativos en las propiedades de un sistema se le conoce como transiciones de fase.

Las transiciones de fase se clasifican de acuerdo con el orden de las derivadas de F para el cual se presentan discontinuidades: así, si la discontinuidad se da en una derivada de orden $n-1$, hablamos de una transición de primer orden o discontinua, y para $n > 2$, de una transición continua.

A partir de 2.26a podemos observar que una discontinuidad en la primera derivada de F , con respecto a la temperatura, implica que para una cierta temperatura de transición T_c , existen dos valores distintos para la entropía, correspondientes a dos fases distintas con un mismo valor de F , y por lo tanto, por 2.27a, una divergencia en el calor específico. Un ejemplo de esto es la transición *sólido* \rightarrow *líquido* en la cual, para todo un intervalo de temperaturas y volúmenes coexisten dos valores distintos de entropía y presión, correspondientes a fases distintas.

² Este criterio de estabilidad recuerda al de los sistemas mecánicos en los que la estabilidad se alcanza en los mínimos de la energía potencial, de ahí que las funciones G y S se conoce como potenciales termodinámicos.

La entropía suele ser vista como una medida del desorden del sistema [25]. Bajo esta interpretación de entropía, tenemos que una transición de fase constituye una transición *orden* \rightarrow *desorden*. Así, en la transición *sólido* \rightarrow *líquido*, vemos que el líquido presenta un mayor número de microestados posibles, y por lo tanto mayor desorden.

Macroscópicamente, podemos caracterizar las transiciones en términos de las simetrías que presentan las propiedades del sistema, en donde al estado desordenado se le asocia un grado mayor de simetría [26]. Podemos confirmar esta equivalencia en el caso del agua: si subdividimos un volumen con agua y permutamos cualquiera de sus elementos, para el estado líquido o desordenado, vemos que el aspecto del sistema no cambia, lo cual no sucede para el estado sólido. Los parámetros que cuantifican el orden o asimetría de una fase reciben el nombre de parámetros de orden y los parámetros que el observador puede manipular para cambiar el estado del sistema se conocen como parámetros de control.

En el caso de transiciones continuas, a los puntos en los que ocurren las discontinuidades se les llama puntos críticos: en éstos, los cambios de una fase a otra ocurren en forma continua: un ejemplo es el modelo de un ferromagneto en el que para valores mayores que T_c del parámetro de control T , el sistema se encuentra desordenado; ésto es, los momentos magnéticos (espines) no se orientan en ninguna dirección preferencial. Tal simetría se pierde al disminuir la temperatura más allá de T_c , punto para el cual la energía de acoplamiento entre los espines comienza a ser considerable y el estado de cualquier espín se ve afectado por el de todos los demás (longitud de correlación infinita). Así, la presencia de una fluctuación puede definir una orientación hacia la cual todos los espines tenderán a alinearse con probabilidad uno, a medida que la temperatura continúa bajando, dando lugar a la emergencia del parámetro de orden, en este caso la magnetización promedio, que rompe la simetría del sistema. Esta ruptura de simetría es evidente si consideramos que la interacción entre imanes no será la misma si a uno de ellos lo rotamos 180° .

Una propiedad muy interesante de los sistemas en el punto crítico, es que presentan propiedades de universalidad; esto es, los parámetros de orden crecen como leyes poten-

cias como función de la distancia del valor del parámetro de control al punto crítico: los exponentes que caracterizan estas leyes de potencia son llamados exponentes críticos y son esencialmente independientes de las características microscópicas del sistema.

Capítulo 3

MODELADO

3.1 Por qué y cómo modelar

En la introducción he mencionado que el objeto de esta tesis es el estudio de la desnaturalización del ADN, y que para tal fin se elaboró un modelo. Así que, antes de comenzar la descripción de algunos de los modelos que en buena medida han inspirado y delimitado la elaboración del que aquí se trabaja, dedico las siguientes líneas a intentar hacer más clara la idea de lo que es un modelo y la finalidad detrás de un trabajo de construcción de este tipo.

¿Qué es un modelo?

A continuación, cito algunos intentos de respuesta a esta cuestión.

Según Ulises Moulines [27], una teoría empírica consiste en:

Un conjunto de funciones (reales, complejas, vectoriales, etc.). Así mismo, de uno o más enunciados relacionando a tales funciones, expresando principios fundamentales o leyes. Un ejemplo es la primera ley de la Termodinámica representada por la ecuación 2.8a. Moulines [27] afirma que una teoría empírica no tendría validez si no existiera la posibilidad de aplicar e interpretar semánticamente tales enunciados sobre un universo empírico.

Un modelo, según Moulines [27], es precisamente una aplicación en la que se satisfacen

tales enunciados dentro de un fragmento de la realidad a la que pretende describir la teoría.

Desde un punto de vista práctico, de los modelos y su construcción se puede decir lo siguiente:

cuando se busca una explicación para fenómenos que hasta ese momento no la tienen, resulta útil construir relaciones entre éstos y aquéllos para los cuales se ha logrado dar una explicación sólida: un ejemplo de tal identificación entre modelos es la imagen del átomo propuesta por Bohr, la cual le confiere al átomo una estructura planetaria, y con ésto, asumiendo algunas consideraciones extra, consigue dar una explicación para los espectros atómicos.

Así, al mapeo de las estructuras que subyacen en un fenómeno, a representaciones abstractas de estructuras y formas ya estudiadas se le llama modelo. Para N. R. Hanson [28] "los modelos sugieren ámbitos que posibilitan explicaciones y trayectos hacia lo conocido" y afirma: "en la medida en que se resalta la estructura de un modelo como sobrepuesta a los datos, y se tiene cuidado de no atribuir al fenómeno estudiado los rasgos propios del modelo que son extraños a éste, en esa medida el modelo será útil". Hanson subraya la importancia de las diferencias existentes entre las estructuras de un modelo y las presentes en el fenómeno a describir: "Un modelo, que se pretende que presenta una estructura de ideas como posible esquema de los enlaces de las descripciones de una materia dada, debe diferir de la materia. Si no fuera diferente, la estructura original sería observacionalmente obvia para todo aquel que se enfrente con las descripciones, o al menos tan obvia como en el modelo. O bien no sería obvia para nadie, ni siquiera para el constructor del modelo".

En cuanto al grado de simplificación con el que los modelos presentan las posibles estructuras contenidas en los objetos que describen, se pueden reconocer dos maneras o tendencias distintas de proceder dentro de la comunidad científica. Estas diferencias responden en buena medida a las distintas motivaciones para modelar un fenómeno. Algunos científicos se interesan en reproducir con la mayor precisión posible la fenome-

nología del sistema en estudio, con la finalidad de emplear la información obtenida para realizar predicciones o desarrollar tecnología, y por lo tanto buscan minimizar de un modo sistemático y seguro las discrepancias entre los fenómenos originales y sus modelos teóricos. Esto los lleva a contemplar el mayor número posible de parámetros, muchos de ellos libres, y a dejar a un lado la labor de encontrar interrelaciones entre las estructuras del modelo, en términos diferentes de los directamente aplicables al sistema estudiado.

Por otra parte, para quienes buscan obtener un conjunto de relaciones entre estructuras que permita el acercamiento a la identificación de aquellas inmersas en el fenómeno en estudio, su análisis se centra en los aspectos cualitativos y sus modelos buscarán ser sencillos, restringiéndose a incluir dentro del modelo sólo aquellos aspectos que resulten sustanciales en la caracterización del fenómeno. Siguiendo la línea de Hanson, para quienes adoptan esta visión del modelado, la eliminación completa de todas las discrepancias entre el modelo y el fenómeno original, termina destruyendo lo que el modelo pretendía conseguir: dar una representación de la estructura no explícita en la confrontación con el fenómeno original.

Para ilustrar el riesgo que se corre si se pretende modelar buscando eliminar las discrepancias entre la realidad y el modelo que se haga de ésta, Hanson recurre a un ejemplo similar al siguiente:

un modelo de 5cm en plástico de un coche Ferrari es menos fiel al original que el de 50cm, con carrocería de metal, que camina y que además puede ser controlado. Ambos son a su vez menos fieles que una versión con todos los componentes de un Ferrari real pero cuyas dimensiones son tres cuartas partes de las originales. Esta discrepancia puede también ser eliminada en un modelo que iguale en todo al original. El resultado, sin embargo no sería un modelo del Ferrari, sino un Ferrari más y cualquiera que haya sido la motivación para construir un modelo de un Ferrari no será satisfecha si el resultado es simplemente otro Ferrari original.

Al respecto N.R. Hanson opina [28]: "...La comprensión de los fenómenos sorprendentes exige atender a lo que les hace andar. Dentro de la terrible variedad de modos de

dirigir la atención a rasgos especiales de materias complejas, una cosa es común a todas ellas: debe haber siempre algunas diferencias entre 1) la representación (el modelo, la teoría, la ciencia), y 2) las grandes perplejidades fenoménicas que llevan a los humanos a tratar en primer lugar de entenderlas. Las fotografías detalladas de un embrollado rompecabezas de piezas irregulares son tan desconcertantes como lo fotografiado: no son lo suficientemente diferentes”.

En los modelos físicos de desnaturalización que presento adelante se pueden reconocer actitudes distintas para modelar. Mientras en algunos modelos se incluye un gran número de parámetros, en otros, el análisis es casi meramente cualitativo. Mi intención, para el tratamiento del fenómeno de desnaturalización, es describirlo de la manera más sencilla posible, buscando resaltar el papel que desempeña un sólo aspecto del complejo proceso; esto es, la entropía de las cadenas libres para un extremo del ADN: la desnaturalización vista como proceso unidireccional, dejando como parámetros libres a las energías de interacción entre bases complementarias y al número de configuraciones posibles de un nucleótido.

3.2 Modelos físicos del ADN

En el caso de la desnaturalización del ADN, resulta conveniente trazar su descripción en términos termodinámicos. Para el sistema formado por dos cadenas de ADN arregladas como doble hélice, el proceso de desnaturalización constituye un desordenamiento del sistema, ya que la configuración espacial de las cadenas disociadas está sujeta a una menor restricción. Por la forma en la que ocurre, el desordenamiento del ADN, como función de la temperatura, se puede considerar como una transición de fase. De acuerdo con lo expuesto en el capítulo 2, si se pretende estudiar la desnaturalización desde el punto de vista de transiciones de fase, resulta deseable conocer una expresión para la entropía del par de cadenas que forman al ADN. Con dicha finalidad, desde un comienzo, los modelos para la desnaturalización han representado al ADN como una secuencia Ξ_N

de N elementos σ_i , cada uno de los cuales hace referencia a la unidad formada por un par de bases complementarias. Todo σ_i se encuentra en uno de los estados $\{O, D\}$, ordenado o desordenado, o unido o disociado, respectivamente.

A temperaturas bajas los σ_i tienden al estado O , de mínima energía; sin embargo, fluctuaciones térmicas pueden romper los enlaces energéticos que unen al par, dando lugar a micro transiciones $\sigma_i : O \rightarrow D$ que, si suponemos interacción entre los σ_i , a su vez, afectarán la estabilidad de pares vecinos. Así, como función del cambio en la temperatura, la secuencia Ξ_N puede descomponerse en un número de i y j subsecuencias Ξ_i^D y Ξ_j^O de estados puros, cada una con ℓ_i elementos desordenados y n_j ordenados, donde claramente

$$N = \sum_i \ell_i + \sum_j n_j. \quad (3.28)$$

En estos términos, la desnaturalización consiste en la transición $\Xi_N : O \rightarrow D$, donde a partir de cierta temperatura T_m ,

$n_j \rightarrow 0 \forall j$, quedando ambas cadenas completamente disociadas.

Como se dijo en el capítulo 2, la entropía S de este sistema consiste en la suma de los logaritmos del número de configuraciones espaciales posibles $\Omega_i(\ell_i)$ de cada una de las regiones Ξ_i^D :

$$S = k_B \sum_i \ln \Omega_i(\ell_i) \quad (3.29)$$

Los modelos de desnaturalización varían, en la forma de calcular Ω , en la longitud N del sistema, en el tipo de interacción entre los pares σ , y en la concentración relativa de ambos tipos de pares de bases (AT y CG), en la contracción iónica del solvente y en la rigidez de las cadenas.

Debido a que en los modelos que a continuación se presentan, el cálculo de la entropía emplea herramientas matemáticas especializadas para su presentación, omitimos todo detalle técnico.

3.2.1 Modelo de Poland y Scheraga (PS) [9]

El modelo de Poland y Scheraga es, de los primeros modelos de desnaturalización, el que ha aportado una base más sólida para la aplicación de la mecánica estadística en modelos posteriores. En (PS) se estudian los aspectos cualitativos de la transición de fase como función de la temperatura, que a partir de ellos se conoce como de hélice a cadena (*Helix-Coil*), para cadenas de biopolímeros en general, desde proteínas hasta ácidos nucleicos. A continuación expongo su modelo del ADN.

EL modelo (PS) considera un par de cadenas con $N \rightarrow \infty$.

Para los estados ordenados toma en cuenta:

- 1) las interacciones entre bases complementarias que dan lugar a un par σ_i , asignándoles un peso estadístico t .
- 2) las interacciones entre pares σ_i adyacentes, asignándoles un peso estadístico τ , y para el peso estadístico de una secuencia asociada j con n_j elementos:

$$r_{n_j} = t(t\tau)^{n_j} \equiv tw^{n_j} \quad (3.30)$$

donde $w \equiv t\tau$.

Por simplicidad, suponen al segmento de ADN compuesto por un sólo tipo de par de bases, con energía de enlace $-E$.

A cada secuencia i desordenada se le asigna un peso estadístico u_i .

Como el sistema es infinito, todo segmento disociado tiene la forma de un lazo: entonces, calcular la entropía de un segmento disociado de longitud l , equivale a conocer el número de configuraciones de un lazo de perímetro $2l$. Con tal fin, se mapea el segmento de ADN a una malla de ángulos rectos, como se muestra en la figura 3.1, en la cual cada unidad de la secuencia puede ocupar con igual probabilidad una de las aristas que constituyen una celda de la malla (caminata aleatoria). Con estas consideraciones, según Poland y Scheraga, la entropía resulta ser:

$$S_i/k_B = \ln u_i = a\ell - c \ln \ell - b \quad (3.31)$$

para ℓ grande, donde a , b y c , son parámetros libres.

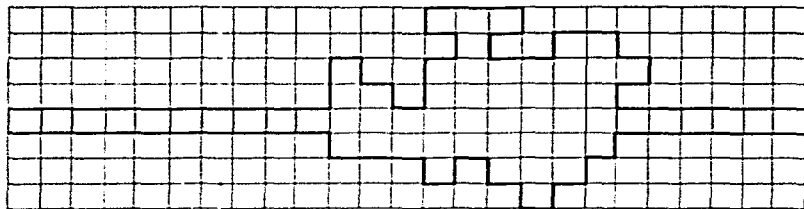


Figura 3.1 Malla cuadrada a la cual se restringen los grados de libertad espaciales de las cadenas disociadas.

Asumiendo que las sucesivas secuencias, ordenadas y desordenadas, no interactúan entre sí, Poland y Scheraga mostraron que el orden de la transición varía con el valor de c , ver la tabla 1

Valores de c	Orden de la transición
$c \leq 1$	0 (no hay transición)
$1 \leq c \leq 2$	1 (transición continua)
$2 \leq c$	2 (transición discontinua)

Tabla 1

Poland y Scheraga encontraron que los valores de c , a su vez dependen de la dimensión de la malla, así, para una malla cuadrada $c = 1$, para una malla cúbica $c = \frac{3}{2}$ y en general $c = \frac{1}{2}d$. De esta observación, lo que en primer lugar concluyeron fue que en cadenas de ácido nucléico reales, tridimensionales, la transición que ocurra no será de primer orden.

3.2.2 Variantes del modelo (PS)

Modelo de Fisher (MF) [7]

El trabajo de M. Fisher consiste en un modelo con idénticas consideraciones al (PS), salvo en la construcción de la entropía. En (MF) también se estudian las configuraciones posibles de las secuencias disociadas, visto como una caminata aleatoria, pero a diferencia de (PS), toma en cuenta que en realidad las cadenas de ADN no se auto-intersecan, lo cual reduce el volumen o número de configuraciones de cada cadena del par que forma el lazo. A esto se le llamo volumen excluido. El objetivo del trabajo de Fisher es estudiar el efecto de volúmenes excluidos en la transición de fase.

Con una expresión aproximada de 3.31, en (MF), se encuentra que para dimensión dos $c \approx 1.46$ y para dimensión tres $c \approx 1.75$, que claramente difieren de los valores $c = 1$ y $c = \frac{3}{2}$ para cadenas que se autointersecan. Sin embargo, la conclusión de (PS) [9], al considerar los efectos de volúmenes excluidos acerca del orden de la transición, se satisface, ya que para dimensiones dos y tres $c \leq 2$, y por lo tanto la transición es de segundo orden.

Modelo de Kafri, Mukamel y Peliti (KMP) [8]

Una versión muy reciente del modelo (PS) [9], extensión del modelo de Fisher [7], es el (KMP). Su objetivo es refinar el estudio del grado de la transición, buscando resultados analíticos que resulten de considerar los efectos de las interacciones de volumen excluido entre cadenas, i.e., suponiendo que las cadenas de ADN no se intersectan mutuamente, ni a sí mismas.

Al igual que en (MF), en (KMP) se argumenta que la forma general de la degeneración de un lazo de longitud n :

$$A \frac{z^n}{c^n} \tag{3.32}$$

no cambia al considerar volúmenes excluidos y sus interacciones, sino que, lo único

que se modifica es el exponente c .

Para evaluar el exponente c , en (KMP) se aplican los resultados de Duplantier [36] para la entropía general de arreglos de lazos en polímeros. Este resultado propone un método para calcular el número de configuraciones del polímero, a partir de la topología de cada uno de los lazos presentes en el sistema.

En (KMP) primero se calcula el caso de un sólo lazo, en el límite cuando la razón de las longitudes del lazo y del polímero tiende a cero y encuentra que:

$$c = 2 + \frac{13}{32} \text{ para } d = 2 \quad (3.33)$$

$$\approx 2.115 \text{ para } d = 3 \quad (3.34)$$

En el segundo caso, el arreglo está formado por tres lazos adyacentes, con el lazo intermedio mucho menor, de donde resulta

$$c = 2 + \frac{11}{16} \text{ para } d = 2 \quad (3.35)$$

$$c \approx 2.22 \text{ para } d \lesssim 4 \quad (3.36)$$

Ya que para dimensiones $d=2,3$ el exponente es mayor a dos, en (KMP) se concluye que la transición es de primer orden para $d \geq 2$.

Este resultado, según (KMP), no cambia al permitir que la desnaturalización ocurra también en uno de los extremos.

Modelo de Wartell y Benight (WB) [29]

A diferencia de los modelos antes presentados, el modelo (WB) pretende reproducir con precisión las curvas experimentales de desnaturalización, para lo cual considera un gran número de parámetros.

En (WB) se estudia un conjunto de segmentos de longitud N , cada uno con k configuraciones posibles con un peso estadístico W_k . Por ser un sistema finito, (WB) considera que la disociación ocurre en forma de lazos, y en ambos extremos de los segmentos de ADN.

W_k es propuesto como un producto de parámetros.

Parámetros del modelo (WB)

a) Parámetro de estabilidad:

A cada par de bases en estado ordenado, se le asigna un valor Φ asociado a la constante de equilibrio de la reacción de asociar un par de bases en el extremo de un subsegmento desordenado. S esta dado por:

$$\Phi_\ell = e^{-\frac{1}{k_B T} (G_\ell - G_{\ell-1}^\circ)} \quad (3.37)$$

Con G_ℓ la diferencia de energía libre del par ℓ entre estados abierto y cerrado, tomando en cuenta la energía de los puentes de hidrógeno, el efecto del solvente, el cambio de energía de interacción de primeros vecinos y la contribución entrópica por apareamiento de bases, $G_{\ell-1}^\circ$ es la diferencia de energía libre considerando el estado del par $\ell - 1$ y la contribución entrópica de asociar el par ℓ . Φ_ℓ toma en cuenta la naturaleza del par, AT o CG.

Parámetro de cooperatividad.

A cada par de bases dentro de un subsegmento desordenado se le asigna un valor φ correspondiente al cambio de energía libre que resulta de un cambio de estado en un par de bases vecino, esto es:

$$\varphi = e^{-\frac{1}{k_B T} (G_{\ell-1}^\circ - G_{\ell-2}^\circ)} \quad (3.38)$$

Parámetro de Entropía.

A un subsegmento desordenado de longitud ℓ se le asigna la probabilidad $f(\ell)$ de que las cadenas que forman este lazo se encuentren en una configuración tal que sea posible

la asociación de sus pares extremos. Usan la forma:

$$f(\ell) = \frac{1}{(\ell + d)^k} \quad (3.39)$$

con d y k parámetros libres, propuesta por Coroh y Y. Tagashira [30].

Así, a un segmento con la configuración

$$\{\sigma_1^{ord} \sigma_2^{ord} \sigma_3^{des} \sigma_4^{des} \sigma_5^{ord} \sigma_6^{ord}\}$$

se le asignaría un peso estadístico:

$$W = S_1 S_2 \varphi_3 \varphi_4 f(2) S_5 S_6 \quad (3.40)$$

Una vez hecho lo anterior calculan el parámetro de orden θ , conocido como helicidad del conjunto de segmentos

$$\theta = \theta_{ext} \theta_{int} \quad (3.41)$$

Para una secuencia $\{\sigma_1 \sigma_2 \sigma_3 \dots \sigma_N\}$ con:

$$\sigma_i = \{1 \text{ si orden, } -1 \text{ si desorden}\}$$

$\theta_{int} \equiv$ fracción de pares de bases en estado ordenado dentro de un segmento de longitud N , con al menos un par de bases asociado, en función de la temperatura, dada por:

$$\theta_{int} = \frac{1}{N} \sum_{\{k\}} \frac{\ell_k W_k}{Z_{int}} \quad (3.42)$$

con ℓ_k el número de pares asociados en la configuración k y Z_{int} la función de partición de cada segmento con al menos un par ordenado:

$$Z_{int} = \sum_{\{\sigma_i\}} \prod_{i=1}^N P(\sigma_i, \sigma_{i-1}) \prod_{\ell} f(\ell) \quad (3.43)$$

con $P(\sigma_i, \sigma_{i-1})$ la probabilidad conjunta de los estados σ_i y σ_{i-1} .

$\theta_{ext} \equiv$ fracción de cadenas completamente disociadas dentro del conjunto de secuencias. esto debido al proceso

$$C_2 = 2C_1 \quad (3.44)$$

con C_2 la concentración de cadenas aparejadas y C_1 la concentración de cadenas simples.

K_D la constante de la reacción 3.44 (32):

$$K_D = 2 \frac{C_1^2}{C_2} \quad (3.45)$$

Partiendo de 3.45 y de

$$C_{Total} = C_1 + 2C_2 \quad (3.46)$$

se obtiene:

$$\theta_{ext} = \frac{2C_2}{C_T} = \frac{\frac{4C_1^2}{K_D} + 1 - [1 + \frac{4C_1^2}{K_D}]^{1/2}}{\frac{4C_1^2}{K_D}} \quad (3.47)$$

Se evalúa y gráfica a 3.41 como función de la temperatura para diversos valores de N . Las curvas obtenidas son muy próximas a las experimentales [30]. Por otra parte, (WB) no retoma la discusión acerca del orden de la transición, pero señala que el aspecto de las curvas de fusión experimentales, sugiere una transición de primer orden.

Capítulo 4

DESCRIPCIÓN DEL MODELO.

4.1 Consideraciones del modelo.

A continuación se enumeran las consideraciones mencionadas en 1.2. que pretenden ser las mínimas necesarias para reproducir las principales características de la desnaturalización de segmentos cortos de ADN, dada su distribución de bases.

Consideraciones:

I) Al igual que en (WB), se consideran cadenas finitas.

II) Las cadenas reales son flexibles sólo en cierta escala; esto es, para longitudes mayores que L_p o longitud de persistencia. Las cadenas con $L_p = 0$ se conocen como cadenas ideales. En este modelo se consideran como cadenas totalmente flexibles o ideales a los segmentos en estado desordenado, y totalmente rígidas para el estado ordenado.

III) En los polímeros se encuentra que para cada monómero existe un número z de mínimos energéticos, asociados cada uno con un ángulo de conformación [31]; i.e., cada monómero cuenta con z grados de libertad [31], ver figura 3.2. En este modelo, z queda como parámetro libre.

IV) A diferencia de los modelos antes expuestos, el mecanismo de desnaturalización aquí propuesto es un mecanismo de zipper, en el que la ruptura es ordenada y unidireccio-

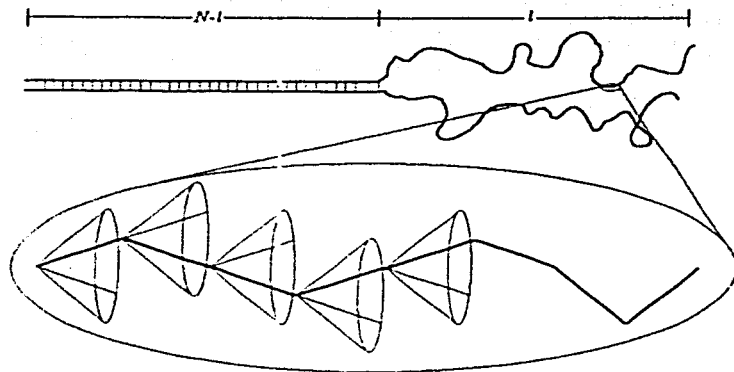


Figura 3.2 Esbozo de segmento de ADN, desnaturalizado como un *zipper*, en donde se muestran los z ángulos de configuración.

nal; i.e., no se forman lazos¹ y se da a partir de un sólo extremo del segmento, ver figura 3.2. De esta manera, al aumentar temperatura, se forma un subsegmento disociado por ruptura térmica, que crecerá en longitud y con ello la entropía del sistema, hasta alcanzar un punto para el cual el cambio en la energía libre es nulo y el estado ordenado se vuelve inestable, ocurriendo la disociación completa de las cadenas, ya que en este estado se da el grado máximo de degeneración del sistema.

El modelo no toma en cuenta la contribución entrópica de cadenas separadas, así que se restringe a estudiar cadenas con al menos un par asociado.

V) Como se menciona en 1.4, las curvas experimentales exhiben un comportamiento escalonado. El modelo pretende reproducir cualitativamente tal propiedad, y para ello se toman en cuenta a ambos tipos de energía de enlace, para, AT y CG, cuya densidad fluctúa a lo largo de la secuencia.

En realidad, la estabilidad de segmentos de ADN varía, no sólo dependiendo de la concentración relativa de pares, sino según la especificidad de las secuencias, ya que la energía de enlace de cada par de bases es influenciada por el tipo de interacción que establezca con sus pares vecinos [5]. Esto da lugar a diferencias de estabilidad locales, a las que hacen referencia los parámetros de cooperatividad y estabilidad de (WB); tales diferencias de estabilidad se pueden ver reflejadas en la estructura secundaria del ADN, y pueden ser la base estructural para mecanismos de control de expresión de genes [5].

En este modelo, a diferencia de (WB) [29], no se consideran las interacciones entre pares vecinos, ni interacción con el solvente, ni efectos cooperativos de estabilidad; únicamente se toman en cuenta las energías de enlace que forman a los dos tipos pares.

La consideración V, fue primero estudiada por Poland y Scheraga (Theory of helix coil transition) [9] y por E.W. Montroll y N.S. Goel [32], y más recientemente por D. Cule y T. Hwa [33]. En este último trabajo se plantea que la heterogeneidad de la secuencias no es suficiente para dar lugar a escalonamiento y se afirma que es necesario que el efecto sea

¹Esta consideración es consistente con I ya que en segmentos cortos la disociación en los extremos es más probable que al interior del segmento.

amplificado al considerar la rigidez de las cadenas, para lo cual proponen el Hamiltoniano

$$3H = \frac{1}{T} \sum_{n=1}^N \left\{ \frac{k}{2} (y_{n-1} - y_n)^2 + V_n(y_n) \right\}, \quad (4.1)$$

donde y_n es la distancia entre las bases del n -ésimo par en la dirección de los puentes de hidrógeno, determinada por una caminata aleatoria en una dimensión; el término cuadrático, de acoplamiento, describe la rigidez de las cadenas y $V_n(y_n)$; el potencial de interacción, incluye la especificidad del tipo de enlace, AT o CG.

El Hamiltoniano 4.1 fue introducido por M. Peyrard y A.R. Bishop [34], y ha sido muy empleado en estudios recientes de desnaturalización de ADN.

Con la consideración de heterogeneidad de cadenas en V se pretende, al igual que en el estudio de Cule y Hwa [33], obtener cualitativamente perfiles escalonados en las curvas de fusión.

4.2 Función de partición.

Para construir la función de partición del sistema de acuerdo con las consideraciones anteriores, se propone, para una temperatura dada, un segmento de ADN visto como una cadena de longitud N con un subsegmento desnaturalizado de longitud ℓ , y por lo tanto uno ordenado de $N-\ell$. *VERFIG*

Ahora bien, como cada unidad de la secuencia desordenada, de acuerdo con III), tiene z configuraciones, el número total de configuraciones posibles será z^ℓ para el subsegmento desnaturalizado y entonces para el par de cadenas

$$\Omega(\ell) = z^{2\ell}. \quad (4.2)$$

Como sólo se considera el intercambio de calor del solvente con el fragmento de ADN estudiado, y no los efectos químicos sobre el mismo, el solvente hace las veces de un baño térmico para nuestro subsistema en cuestión, cuya energía fluctuará alrededor de un valor

promedio de $k_B T$. Tomando en cuenta que la longitud es fija, se sigue que es posible dar una descripción adecuada a partir de la construcción de una función de partición para el sistema basada en los elementos planteados en el capítulo 2.

La función de partición completa, como se expone en (WB) se escribiría así:

$$Z = Z_{ext} + Z_{int}. \quad (4.3)$$

con Z_{ext} la función de partición para el par de cadenas completamente disociadas y Z_{int} la función de partición para el par de cadenas con al menos un par de bases enlazado. Por la consideración IV, únicamente nos interesa Z_{int} , en la cual, l está restringida al intervalo $[0, N-1]$. En adelante, Z_{int} será denotado simplemente por Z .

Como muestran los modelos antes mencionados, resulta conveniente tratar al ADN como una secuencia de elementos, cada uno con dos estados posibles. En este modelo procedemos de igual forma; sin embargo, considerar a ambos tipos de enlace, AT y CG, distribuidos aleatoriamente a lo largo de la secuencia, introduce un desorden *intrínseco*, no correlacionado, que dificulta obtener en forma usual la termodinámica del sistema, ya que el Hamiltoniano es diferente para la posición ocupada por un par AT, que para una con CG. Un ejemplo típico de sistema con esta característica de desorden espacial estático es el de magnetos con impurezas [3]. Este tipo de desorden es conocido como desorden fijo o "quenched disorder", que *no* debe confundirse con el desorden de estados de cada elemento del sistema, debido a la posibilidad que éstos tienen de ocupar, con igual probabilidad, uno de los dos estados permitidos. La entropía del sistema es debida a esto último.

4.2.1 Desorden dinámico

La forma de proceder en el caso de un sistema con desorden espacial fijo es a) suponer un conjunto de sistemas, cada uno con distinta configuración de desorden b) obtener la termodinámica para cada uno de dichos sistemas y c) promediar los parámetros ter-

modinámicos del sistema original sobre los parámetros de todo el conjunto de sistemas [3].

Para el tratamiento de este modelo, se supone al desorden fluctuando en el tiempo, y se obtiene la termodinámica promediando sobre las configuraciones internas k de las cadenas [3]; i.e., se pretende obtener la termodinámica a partir de $\ln\langle Z_k \rangle$ en lugar de $\langle \ln Z_k \rangle$, con Z_k la función de partición para la k -ésima configuración de desorden espacial.

Aún cuando este procedimiento no es equivalente al de desorden fijo, los resultados llegan a ser semejantes, como veremos más adelante.

Considerando una distribución de pares CG dada, introducimos las energías de interacción $U_{i,k}$ entre ambos tipos de pares, AT y CG, en el Hamiltoniano H_k :

$$H_k = \sum_{i=1}^{N-1} U_{i,k}, \quad (4.4)$$

donde i indica la posición del i ésimo par dentro de la cadena y k se refiere a la k ésima configuración de desorden espacial, con

$$U_{i,k} = \begin{cases} U_{AT} & \text{si A-T} \\ U_{CG} & \text{si C-G} \end{cases}, \quad (4.5)$$

donde U_{AT} y U_{CG} son las energías para cada tipo de par.

Cada una de estas configuraciones se puede realizar de $\Omega(\ell)$ maneras distintas determinadas por 4.2; así que, de acuerdo con 2.22, la función de partición de nuestro sistema es

$$Z_k = \sum_{\ell=0}^{N-1} \{ \Omega(\ell) e^{-\ell H_k} \} \quad (4.6)$$

Si definimos una función $s_{i,k}$ representando los dos valores posibles para cada par, como

$$s_{i,k} = \begin{cases} 0 & \text{si A-T} \\ 1 & \text{si C-G} \end{cases} \quad (4.7)$$

H_k se escribe

$$H_k = \sum_{\ell=1}^{N-\ell} [U_{AT} + s_{\ell,k}(U_{CG} - U_{AT})]. \quad (4.8)$$

Entonces, por la expresión 4.2, 4.6 queda

$$Z_k = \sum_{\ell=0}^{N-1} [z^{2\ell} e^{-\beta U_{AT}(N-\ell)} e^{-\beta \Delta U S_k(\ell)}] \quad (4.9)$$

donde

$$\Delta U = (U_{CG} - U_{AT}) \quad (4.10)$$

y

$$S_k(\ell) = \sum_{i=1}^{N-\ell} s_{i,k} \quad (4.11)$$

es el número de pares C-G en la región ordenada.

De acuerdo con la aproximación de desorden dinámico, suponemos a la configuración espacial de pares variando en el tiempo, así que obtenemos el promedio de la función de partición 4.9, sobre todas las k configuraciones posibles. El cálculo de dicho promedio lo separamos en dos casos, fáciles de resolver:

1. Aproximación de campo promedio.

Este caso considera al promedio sobre las k configuraciones de pares CG igual al promedio de la secuencia original.

$$\rho = \langle s_i \rangle_k = \langle s_i \rangle, \quad (4.12)$$

de esta forma

$$S(\ell) = \rho(N - \ell), \quad (4.13)$$

entonces

$$\langle Z_k \rangle = \sum_{\ell=0}^{N-1} [\Omega(\ell) e^{-\beta U_{AT}(N-\ell)} e^{-\beta \Delta U \rho(N-\ell)}] \quad (4.14)$$

o

$$\langle Z_k \rangle = e^{-\beta N(U_{AT} - \rho \Delta U)} \sum_{\ell=0}^{N-1} [z e^{\beta(U_{AT} - \rho \Delta U)}]^\ell \quad (4.15)$$

que con

$$V = U_{AT} + \rho \Delta U \quad (4.16)$$

y

$$r_o = e^{\beta V} \quad (4.17)$$

resulta claro que la expresión 4.15 es una serie geométrica y por lo tanto

$$\langle Z_k \rangle = r_o^{-N} \left[\frac{1 - [z r_o]^{N-1}}{1 - z r_o} \right] \quad (4.18)$$

2. Secuencias de Exones

En este caso se suponen secuencias de bases sin correlación: exones [19].

Por 4.7 vemos que

$$e^{-\beta M^N(S)} = \prod_{i=1}^{N-1} (1 + s_i (e^{-\beta \Delta U} - 1))$$

entonces 4.9 se puede escribir como

$$\langle Z_k \rangle = e^{-\beta U_{AT} N} \sum_{\ell=0}^{N-1} [z e^{\beta U_{AT}}]^\ell \prod_{i=1}^{N-\ell} (1 + s_i (e^{-\beta \Delta U} - 1)) \quad (4.19)$$

Si consideramos secuencias de exones, en las que las probabilidades de encontrar s_i y s_j son independientes ó equivalentemente

$$\langle s_i s_j \rangle = \langle s_i \rangle \langle s_j \rangle, [22] \quad (4.20)$$

entonces en 4.19

$\prod_{i=1}^{N-\ell} (1 + s_i(e^{-\beta\Delta U} - 1)) = (1 + \rho(e^{-\beta\Delta U} - 1))^{N-\ell}$, lo que en este caso permite escribir a $\langle Z_k \rangle$ como

$$\langle Z_k \rangle = e^{-\beta U_{Ar} N} (1 + \rho(e^{-\beta\Delta U} - 1))^N \sum_{\ell=0}^{N-1} \left[\frac{z e^{-\beta U_{Ar}}}{(1 + \rho(e^{-\beta\Delta U} - 1))} \right]^\ell \quad (4.21)$$

con

$$r_1 = \frac{z e^{-\beta U_{Ar}}}{(1 + \rho(e^{-\beta\Delta U} - 1))}. \quad (4.22)$$

Podemos observar que 4.15 y 4.21 tienen la misma estructura de serie geométrica, por lo que el comportamiento de los parámetros termodinámicos será cualitativamente el mismo.

4.3 Resultados

En esta sección obtenemos los parámetros termodinámicos que describen la transición de fase en nuestro modelo.

Densidad homogénea de pares CG

Al igual que en (WB), esperamos obtener resultados que puedan ser comparados con datos experimentales, así que calculamos la helicidad θ para poder observar las curvas de desnaturalización del modelo.

El parámetro de orden en este sistema es la fracción promedio de bases en estado ordenado, esto es:

$$\theta = \frac{N - \langle \ell \rangle}{N} \quad (4.23)$$

donde

$$\langle \ell \rangle = \frac{\sum_{\ell=0}^{N-1} \ell [zr_0]^\ell}{\sum_{\ell=0}^{N-1} [zr_0]^\ell} \quad (4.24)$$

$$= z \frac{\partial \langle Z_k \rangle_k}{\partial z} \quad (4.25)$$

es la longitud promedio de la región desnaturalizada. Así, la helicidad resulta ser

$$\theta = \frac{1}{1 - (zr)^N} - \frac{zr}{N(1 - zr)} \quad (4.26)$$

En este punto es posible reconocer que 4.6 es consistente con lo esperado del este modelo, ya que si inspeccionamos 4.26, para ambas aproximaciones, r igual a r_0 o r_1 , vemos que cuando $T \rightarrow 0$, el parámetro de orden $\theta \rightarrow 1$, y para $kT < U_{AT}$, $r \rightarrow \infty$ y $\theta \rightarrow 0$; tal comportamiento en el límite termodinámico ($N \rightarrow \infty$), como hemos dicho en el capítulo 2, se puede interpretar como una transición de fase.

La gráfica 1 muestra las curvas de fusión para la aproximación de exones, con $N=50$, 100, 200, 400 y 1000, que resultan de 4.26 al sustituir los valores teóricos de U_{AT} y U_{CC} empleados en [6], ver tabla 2. Observamos que las curvas de fusión para exones y para campo promedio son prácticamente indistinguibles.

Por otra parte, encontramos que la variación de los parámetros (ρ, z) no modifica la forma de las curvas de fusión, únicamente la temperatura de transición T_m :

variación de ρ :

La temperatura de transición T_m aumenta linealmente con ρ , mostrando que a mayor concentración de enlaces fuertes (CG), es necesaria mayor cantidad de calor para desestabilizar la molécula.

Variación de z :

el aumento en el número de configuraciones posibles de cada unidad disociada, implica mayor incremento de entropía y por lo tanto, para una T_m menor se alcanza la cantidad

de calor requerida para que la fase ordenada deje de ser estable.

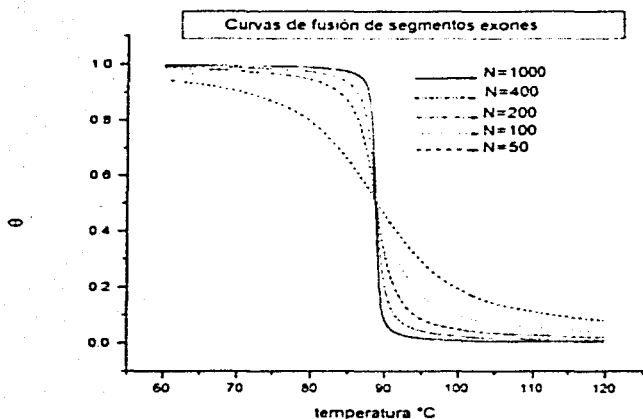


Figura 4-2: Gráfica 1. Curvas de fusión de segmentos exones con distintas longitudes.

Transición de fase Con la finalidad de conocer a que tipo de transición da lugar nuestro modelo, a partir de 2.23, 2.25, 2.18 y 2.27a y de acuerdo con la aproximación de desorden dinámico, calculamos la energía libre, la energía interna, la entropía y el calor específico del sistema para r_0 , pues los cálculos en la aproximación de campo promedio resultan más sencillos y, como hemos mencionado, la diferencia con el caso de correlación nula es despreciable, así:

aplicando 2.23, la energía libre resulta ser

$$F \approx -k_B T \ln \langle Z_k \rangle \quad (4.27)$$

$$= NV - \frac{1}{\beta} \ln \left[\frac{1 - [zr_0]^{N-1}}{1 - zr_0} \right], \quad (4.28)$$

y de 2.25 la energía interna

$$U = \frac{(N+1)(U_{AT} + \rho\Delta U)}{1 - [ze^{\beta U}]^{N-1}} - \frac{U_{AT} + \rho\Delta U}{1 - ze^{\beta U}}, \quad (4.29)$$

con la entropía

$$S = - \frac{\partial F}{\partial T} \quad (4.30)$$

y el calor específico

$$C_v = \frac{\partial U}{\partial T} \quad (4.31)$$

Las gráficas 5a, 5b, 5c y 5d muestran los resultados obtenidos, usando el algoritmo que se muestra en el apéndice III.

Densidad de pares CG uniforme a tramos

Para reproducir el escalonamiento de las curvas de fusión observado en las curvas experimentales [29], [30], definimos por tramos a ρ , la densidad de pares CG, a lo largo de un segmento de 300pb.

Primero, el segmento fue dividido en dos regiones de igual longitud, cada una con ρ distinta, y a partir de una estimación numérica de 4.21, ver Apéndice I con valores de $z=9$ y U_{AT} y U_{CG} tomados de [6], ver tabla 2, obtuvimos las curvas de fusión del modelo, en las que se observó lo siguiente:

Las curvas de fusión presentan un escalón en el caso en que el zipper atraviesa de una región de menor a otra de mayor densidad. En el caso opuesto, para $T < T_m$, θ se mantiene prácticamente constante y en T_m cae a cero bruscamente.

Aparte, para segmentos con distintas densidades promedio $\bar{\rho}$, se registraron los valores de los cambios en la densidad $\Delta\rho$ de una región a otra, suficientes para producir

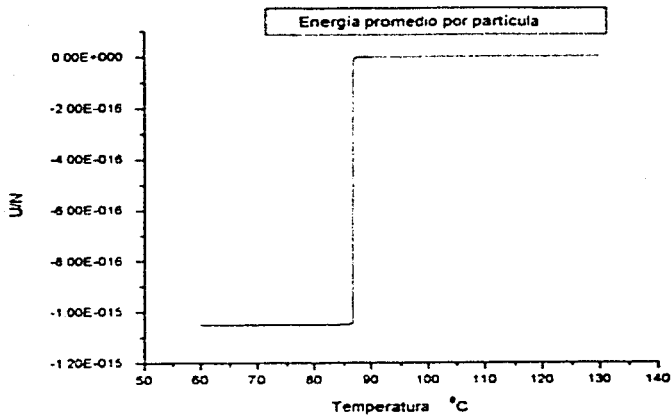


Figura 5a Energía promedio por partícula, con $N=1 \times 10^7$.

escalonamiento. Observamos que secuencias con concentraciones promedio de AT altas son más sensibles a cambios en la densidad y $\Delta\rho$ pequeñas producen escalonamiento, así, para $\bar{\rho} = .15$ una $\Delta\rho \geq .019$ es suficiente para producir un escalón, en tanto que con $\bar{\rho} = .8$ es necesaria una $\Delta\rho \geq .9$. Un siguiente paso fue observar estas variaciones para un número mayor de divisiones del segmento.

Con la idea de evaluar que tan realistas son estos resultados, para una cierta distribución de densidad, comparamos las curvas de fusión del modelo y las que resultan del algoritmo de G. Steger (AS) [4]. Este algoritmo es empleado para reproducir en forma precisa las curvas de desnaturalización por cambio en la absorción de luz con longitud de onda de 260nm, que, como ya hemos señalado en 1.4, se observarían experimentalmente para una secuencia específica. El cambio en la absorción lumínica ΔA_{260nm} guarda una

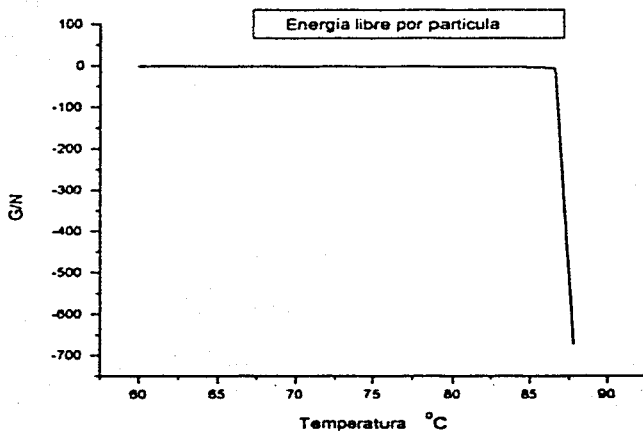


Figura 5b Energía libre por partícula, con $N=1 \times 10^7$.

relación lineal con la concentración relativa de pares disociados [29]; i.e:

$$1 - \Delta A_{2GO_{im}} \propto \theta.$$

esto permite que los resultados de (AS) y los de nuestro modelo sean comparables. Elegimos como referencia a las curvas de fusión de (AS), ya que este algoritmo toma en cuenta el cuadro de consideraciones de (WB), que es el más detallado, además de la rigidez de las cadenas. (AS) cuenta con conjuntos de combinaciones de nueve parámetros libres que optimizan la simulación de curvas de fusión experimentales para distintas concentraciones iónicas en el solvente.

Comenzando con los valores ya mencionados de U_{AT} y U_{CG} y para distintos valores

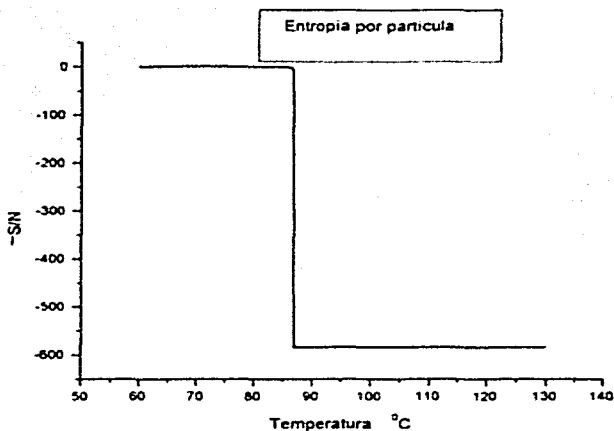


Figura 5c Entropía por partícula, con $N=1 \times 10^7$.

de ρ , constantes a lo largo de toda la cadena obtuvimos las curvas de fusión.

Las curvas de nuestro modelo fueron comparadas con las que resultan de (AS) para secuencias con densidades promedio iguales a las distintas ρ y con los parámetros convencionales de (AS).

Tanto para el modelo como para (AS), graficamos las curvas de fusión en el intervalo de temperaturas de $60^\circ C$ a $110^\circ C$, con $\Delta T = .5$; esto permitió superponer ambas gráficas y ajustar el parámetro z para hacer coincidir sus T_m . Los diferentes valores ajustados de z , que resultan de ρ distintas, fueron promediados.

Para el valor promedio \bar{z} , fijamos al azar cuatro valores de ρ :

$\rho_A = .52$, $\rho_B = .61$, $\rho_C = .78$ y $\rho_D = .8$ y con ellos formamos dos distribuciones

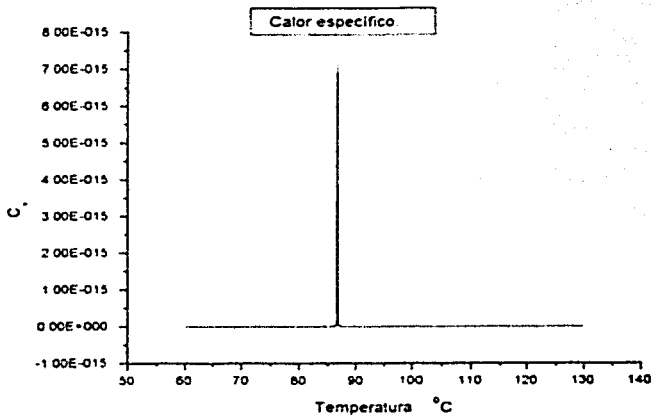


Figura 5d Calor específico, con $N=1 \times 10^7$.

$$PAPBPCCPD \quad (4.32)$$

$$PBAPCCPD \quad (4.33)$$

que dividían en partes iguales al segmento. Por otro lado, con la ayuda del algoritmo que se muestra en apéndice II, generamos secuencias aleatorias, cuyos promedios coinciden con las distribuciones 4.32 y 4.33.

Primero, para longitudes de $N=100$, 200 y 300, formamos dos conjuntos de 15 secuencias aleatorias, cada uno de ellos asociado con una distribución. Con estos conjuntos de secuencias alimentamos a (AS) y promediamos a las 15 curvas de fusión de longitud N que resultan de cada conjunto.

Debido a que resulta más fácil hacer coincidir "Picos" que pendientes, de cada una de las curvas promedio construimos las gráficas

$$\frac{d(1 - \Delta A_{200nm})}{dt} \text{ c.s. } T \quad (4.34)$$

que fueron comparadas con las

$$\frac{d\theta}{dT} \text{ c.s. } T \quad (4.35)$$

del modelo.

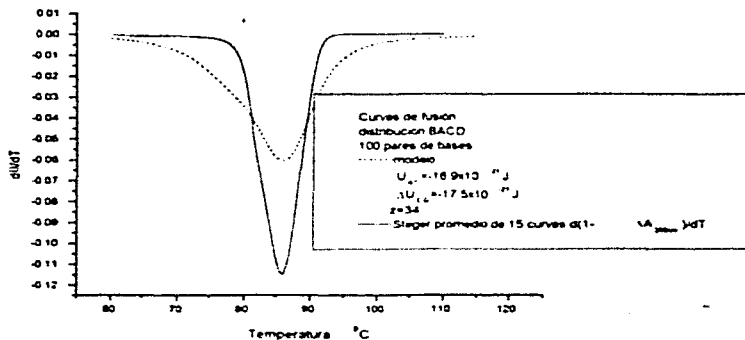
Para cada longitud N considerada, y para ambas distribuciones simultáneamente, modificamos el valor de U_{AT} hasta hacer coincidir los picos principales de las gráficas 4.34 y 4.35, seguido de los cambios en ΔU que ajustaron a los picos secundarios. Encontramos que para reproducir el número, forma y distancia entre picos, simultáneamente para ambas distribuciones, eran necesarios valores de U_{AT} y ΔU más altos de los originalmente empleados. Esto aumentó considerablemente la temperatura de transición del modelo, lo cual tuvo que ser compensado con un valor de z mayor a \hat{z} , capaz de recorrer el conjunto de picos hasta la región en que tienen lugar los de la gráfica promedio de (AS). Pudimos observar que la forma de las subtransiciones es independiente del valor de z . Los valores óptimos de nuestros tres parámetros libres, se agrupan en la tabla 3.1. Las gráficas 2 y 3 muestran las curvas de fusión promedio de las 15 secuencias relativas a cada distribución, 4.33 y 4.32, para las distintas N empleadas.

El procedimiento de ajuste antes mencionado se realizó para curvas promedio de cincuenta secuencias de longitud $N=200$, cuyos resultados mostramos en gráficas 4.

Como se aprecia en las gráficas 4, la diferencia entre las curvas del modelo y las de (AS) disminuyó notablemente, comparado con gráficas 2b y 3c, con un cambio muy pequeño en los parámetros libres, tabla 2

ABCD/BACD	N=100	N=200	N=300	50 prom.	teóricos**
Z	34	34	31	34	
$U_{AI}(x10^{-21}J)$	-16.9	-17/*	-17.11	17	8/6.4
$U_{CG}(x10^{-21}J)$	-17.5	-17.91/*	-17.91	17.78	12

Tabla 2 Resultados de los ajustes del modelo. *resultados del ajuste con curvas promediadas de 50 secuencias. **Los valores teóricos son los reportados en [6] y [34].

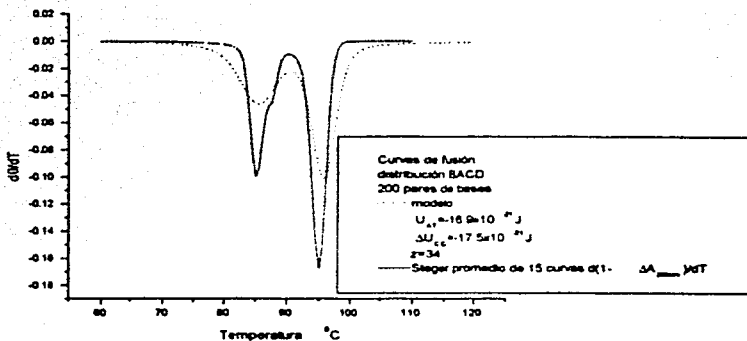


Gráfica 2a Curvas de fusión de la distribución BACD con 100 pares de bases: modelo y (AS) promedio de 15 curvas.

4.4 Discusión

Aspectos cualitativos:

El primer punto que quisiera tocar es el orden de la transición de fase, pues aún cuando su estudio se remonta a varias décadas atrás, la naturaleza del orden de la transición continúa siendo un tema controvertido. En tanto los modelos pioneros en este campo [9],



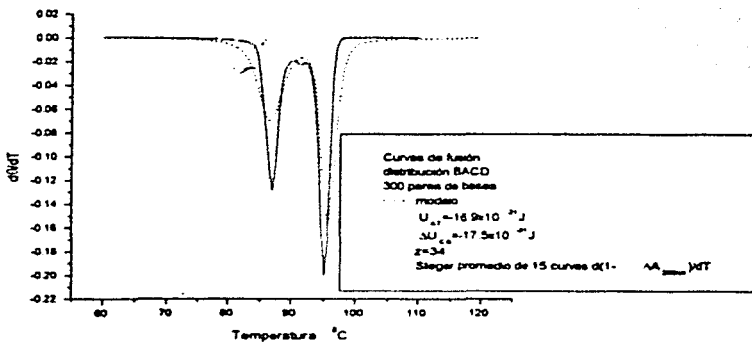
Gráfica 2b Curvas de fusión de la distribución BACD con 200 pares de bases: modelo y (AS) promedio de 15 curvas.

[7] concluyen que la transición de fase, para tres dimensiones es continua, las curvas experimentales sugieren que se trata de una transición de primer orden [29], [30]. Aún cuando algunos tratamientos numéricos han dado lugar a transiciones que se asemejan bastante a las de primer orden [34], [33], no es sino recientemente que, por considerar interacciones de volumen excluido y empleando resultados recientes de propiedades de escalamiento en polímeros [36], (KMP) ofrece una solución analítica que presenta transición de fase de primer orden.

En nuestro modelo, a partir de los resultados 4.27 y 4.29, para la helicidad, la energía promedio, la entropía y el calor específico, encontramos que en el límite termodinámico (cadenas muy largas):

$$\frac{\partial \theta}{\partial T} \propto \frac{\partial U}{\partial T} \propto \frac{\partial S}{\partial T} \propto C_v \propto \delta(T - T_m), \quad (4.36)$$

con $\delta(T - T_m)$ delta de Dirac. De acuerdo con lo expuesto en el capítulo 2, de 4.36



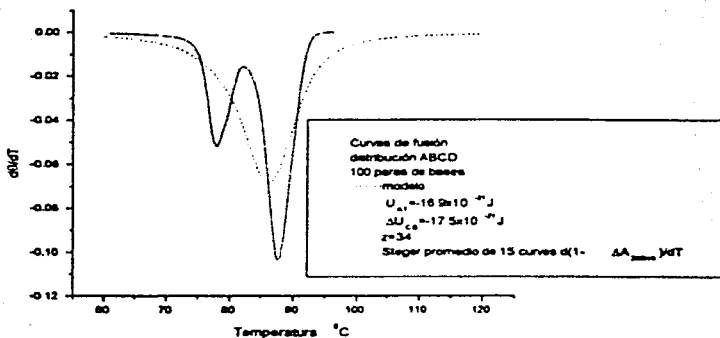
Gráfica 2c Curvas de fusión de la distribución BACD con 300 pares de bases: modelo y (AS) promedio de 15 curvas.

podemos afirmar que nuestro modelo de desnaturalización presenta una transición de fase de primer orden.

Algunos modelos de desnaturalización involucran secuencias heterogéneas (consideración V) [6] y [33]. En nuestro modelo encontramos que la consideración V da lugar a los siguientes efectos:

I) Como ya hemos dicho, al graficar T_m vs $\bar{\rho}$, encontramos que T_m crece linealmente con $\bar{\rho}$. Este comportamiento concuerda con observaciones experimentales [20], como una manifestación de la diferencia que existe entre las energías de enlace que forman a los pares, AT y CG donde la energía de disociación del par AT es aproximadamente un 30% menos intensa que la del CG. Por lo que a mayor concentración de pares CG, mayor es la estabilidad de la conformación helicoidal.

II) Otro aspecto que reproduce el modelo es el hecho de que las curvas de fusión experimentales presenten escalonamiento: en [33] se emplea un modelo de zipper de



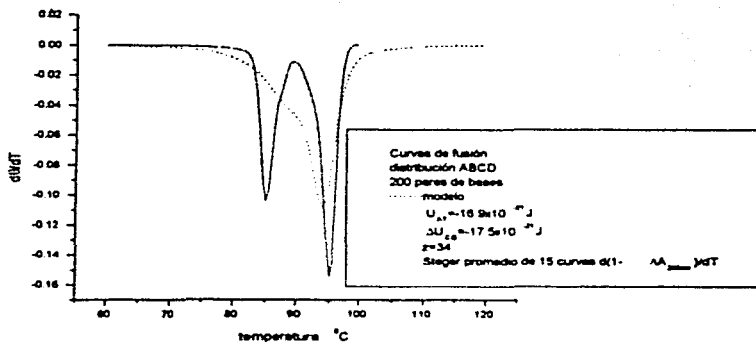
Gráfica 3a Curvas de fusión de la distribución ABCD con 100 pares de bases: modelo y (AS) promedio de 15 curvas.

cadena libre similar al nuestro y se concluye que la heterogeneidad de las secuencias no es suficiente para producir escalonamiento en las curvas de fusión y que para ello es necesario incluir un parámetro de rigidez para las cadenas. Nuestro modelo, permitiendo que la densidad de pares CG fluctúe a lo largo de la secuencia, da lugar a escalones.

Aspectos cuantitativos

Hasta aquí podemos decir que la descripción que nuestro modelo hace de la desnaturalización es satisfactoria, pues reproduce las principales características del fenómeno: sin embargo, si deseamos saber que tan precisa es, deberíamos hacer comparaciones con datos experimentales.

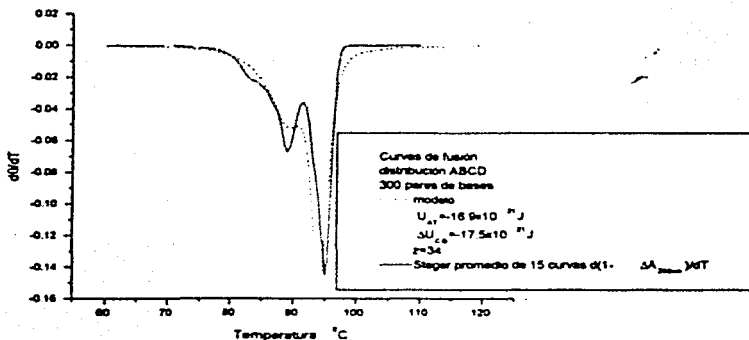
Ya que no fue fácil encontrar curvas de fusión experimentales, y mucho menos producirlas en el laboratorio, recurrimos al algoritmo de Steger. Hoy en día, gracias al gran avance en la capacidad de cómputo con que se dispone, algunos experimentos han caído en desuso, siendo sustituidos por simulaciones precisas como son las de (AS): así que



Gráfica 3b Curvas de fusión de la distribución ABCD con 200 pares de bases: modelo y (AS) promedio de 15 curvas.

podemos confiar en que la información que obtendríamos de comparar nuestros resultados con los de experimentos, no sería considerablemente distinta. Con valores de U_{AT} y U_{CC} del orden de los reportados en la literatura, tabla, no sólo es posible reproducir el número y la forma de las subtransiciones (picos), sino también los valores aproximados de temperatura para los cuales éstos ocurren. El que los valores de energía resultantes de nuestro ajuste resulten algo mayores, es el precio que el modelo debe pagar por omitir interacciones a primeros vecinos, que aumentan la estabilidad de los microestados ordenados [29] y [5], aunado al efecto en el valor de T_m que resulta de un aumento en la entropía limitado por la rigidez de las cadenas, y a interacciones con el solventes, ambos efectos considerados en (AS).

Como podemos observar en las gráficas 2.3 y 4, las curvas promedio de (AS) se asemejan más a las de nuestro modelo en la medida en que el número de secuencias promediadas es más grande, pues al promediar un número más grande de curvas, los



Gráfica 3c Curvas de fusión de la distribución ABCD con 300 pares de bases: modelo y (AS) promedio de 15 curvas.

picos promedio tienden a ser más gruesos y a disminuir en tamaño.

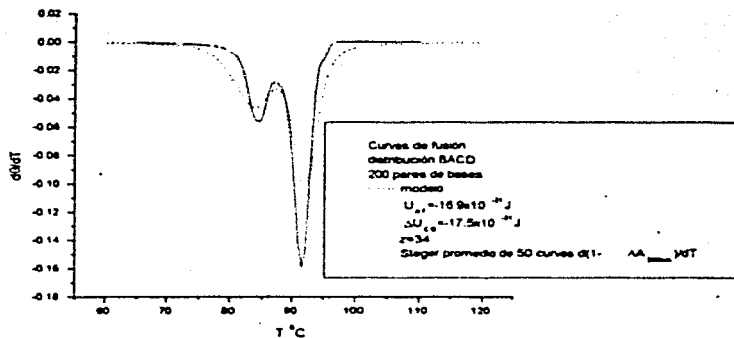
El promedio de 50 secuencias introduce un error del 14%². Creo que una comparación de gráficas, con error despreciable, requeriría promediar del orden de 10^4 secuencias.

La versión del (AS) que empleamos se encuentra en el sitio de internet

(www.biophys.uni-dusseldorf.de/local/POLAND/poland.html). Algunas veces, el procesamiento de datos, para una sola secuencia, tardó alrededor de 10 minutos, así que por razones técnicas no promediamos más de 50 curvas.

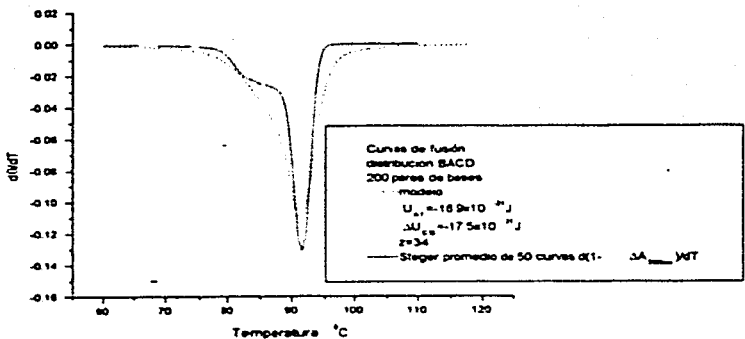
Debemos recordar que cierta diferencia entre las curvas del modelo y las de (AS) es inevitable, pues la helicidad y el cambio de absorción lumínica difieren por un factor de proporcionalidad.

²El error = $\frac{1}{\sqrt{k}}$, con k el número de secuencias.



Gráfica 4a Curvas de fusión de la distribución BACD con 200 pares de bases: modelo y (AS) promedio de 50 curvas.

TESIS CON
FALLA DE ORIGEN



Gráfica 4b Curvas de fusión de la distribución ABCD con 200 pares de bases: modelo y (AS) promedio de 50 curvas.

**TESIS CON
FALLA DE ORIGEN**

Capítulo 5

CONCLUSIONES Y PERSPECTIVAS DEL MODELO

El modelo aquí presentado, siendo la versión simple de un modelo de zipper¹, ofrece resultados analíticos que concuerdan con los de modelos basados en consideraciones realistas como es (KMP) y su interacción de volúmenes excluidos, y reproduce parte de la fenomenología de la desnaturalización, como es la sensibilidad a la concentración de pares CG, ó el escalonamiento de las curvas experimentales, sin necesidad de considerar rigidez en las cadenas, como se afirma en [33].

Otro aspecto importante de éste modelo es que logra resultados muy próximos a los de simulaciones precisas, pero con un número mucho menor de parámetros libres; por lo que podemos concluir que las consideraciones propuestas para el modelo son relevantes en la comprensión de la desnaturalización. Para el caso de cadenas cortas, a partir del modelo propuesto, resulta evidente que la disociación en los extremos de las cadenas domina en el proceso de desnaturalización.

Un incremento en el desorden del sistema, capaz de producir una transición discontinua, no es en modo alguno despreciable, por lo que podemos concluir que no es del todo

¹En el contexto de la desnaturalización de ADN, existen modelos de zipper más complicados que el nuestro, un ejemplo se encuentra en [35].

correcto omitir la disociación en los extremos de cadenas largas, si pretendemos estudiar su desnaturalización.

Por otra parte, este modelo es un ejemplo de sistema para el cual el desorden espacial puede ser aproximado por desorden temporal, dando buenos resultados.

Consideramos que para una mejor comprensión del papel que juega la disociación en los extremos de cadenas de ADN en el incremento de entropía, responsable de la transición, será conveniente estudiar el caso en el que la disociación ocurre en ambos extremos. También sería interesante tratar de resolver el caso de desorden fijo: para lo cual, un posible tratamiento es representar a las secuencias de energías entre pares como un potencial aleatorio formado por una caminata aleatoria, que asocia a cada par, AT y CG, pendientes de signo contrario, para, finalmente interpretar la dinámica del zipper a lo largo de la secuencia, como el movimiento de una partícula a través del potencial.

El efecto de considerar disociación en ambos extremos de las cadenas seguramente disminuiría los valores de los parámetros libres U_{AT} y U_{CG} , además de posiblemente mejorar, en cuanto a precisión, la forma de las curvas de fusión.

Para terminar, quisiera decir que la experiencia que tuve al realizar este trabajo de tesis, fue satisfactoria y motivante, pues muestra que a nivel licenciatura es posible plantear y desarrollar modelos que permiten un acercamiento a la comprensión de fenómenos, aún más allá del dominio estricto de la física, empleando únicamente herramientas del salón de clase.

APENDICE I

program melstep

implicit none

INTEGER T,I,P,ENE,N

REAL(S) TETA0,TETA1,TETA2,RO,RO0,DELTAT,TE,U,Z,K,DELTAU

REAL(S) ZETA,R0,R1,R2,R3,nPROM,ZETAa,A,B,C,D

PARAMETER(P=1000,N=1000)

DIMENSIONTE(P),ZETA(0:N-1),ZETAa(0:N+1),R0(P),R1(P),R2(P)

DIMENSION R3(P),TETA0(P),TETA1(P),TETA2(P),nPROM(N+1)

K= 1.381D-23 ! constante de Boltzmann

DELTAU=-.94-21

T=60.D0 ! longitud del intervalo de temperatura

Z=34.D0 ! número de coordinación

DELTAT= T/1000.D0 ! tamaño de incremento

TE(1)=333.15D0 ! temperatura inicial

!concentraciones de CG

A=.52d0

B=.61d0

C=.78d0

D=.8d0

ZETA(0)=1.D0

ZETA_n(0)=1.D0

! LECTURA DE VALORES

112 write(*,*) 'DAME u' ! energia de par AT

read *,u

write(*,*) 'DAME DELTAU' !diferencia de energías

read *,DELTAU

write(*,*) 'DAME z' !número de coordinación

read *,z

RO0 = (A+B+C+D)/4.D0

open(2,file='DENAT.dat', STATUS='UNKNOWN')

open(3,file='DENAT1.dat', STATUS='UNKNOWN')

open(8,file='meltep0.dat', STATUS='UNKNOWN')

DO I=1,P-1

TE(I+1)= TE(I) + DELTAT

RO(I)=EXP((1.D0/(K*TE(I)))*(U+DELTAU*RO0)) !c. prom.

R1(I)=EXP((1.D0/(K*TE(I))*U)

R2(I)=R1(I)/(1+RO0*(EXP(-(1.D0/(K*TE(I)))*DELTAU)-1.D0)) !s/correl.

! DENSIDAD CONSTANTE

TETA0(I)=1.D0/(1.D0-(Z*RO(I))*N)-(Z*RO(I))/(N*(1.D0-Z*RO(I))) !helicidad c. prom.

TETA1(I)=1.D0/(1.D0-(Z*R2(I))*N)-(Z*R2(I))/(N*(1.D0-Z*R2(I))) !helicidad exo-

nes.

! DENSIDAD VARIABLE:

DO ENE=1,N

SELECT CASE (ENE)

CASE(1:100)

RO=a

```

CASE(101:120)
RO =b
CASE(121:175)
RO = c
CASE DEFAULT
RO =d
END SELECT
R3(I)=Z*R1(I)/(1.D0+RO*(EXP(-(1.D0/(K*TE(I))*DELTAU)-1.D0)) !d. var.
ZETA(ENE)-R3(I)**FLOAT(ENE)+ZETA(ENE-1)
ZETAn(ENE)-FLOAT(ENE)*R3(I)**FLOAT(ENE)+ZETAn(ENE-1)
nPROM(ENE)-ZETAn(ENE)/ZETA(ENE)
END DO
TETA2(I)=1.D0-nPROM(N-1)/N !helicidad escalones
write(8,*) TE(1)-273.15d0, TETA2(I)
write(2,*) TE(1)-273.15d0, TETA0(I)
write(3,*) TE(1)-273.15d0, TETA1(I)
END DO
close(2)
close(3)
close(8)
write(*,*) 'densidad promedio=',RO0
go to 112
END

```

APENDICE II

program secuencias aleatorias

```
USE DFPORT
implicit none
INTEGER(4) I,P,ENE,N,CG
REAL(8) NBASE, NCG, RO
REAL(8) X
CHARACTER(1) LETRA
PARAMETER(P=2,N=75)
DIMENSION NBASE(1000),NCG(1000),RO(1000)
CG=30.D0
NBASE(N)=FLOAT(N)
NCG(N)=CG
open(7,file='SEQ.TXT', STATUS='UNKNOWN')
DO I=1,P-1
DO ENE=1,N
RO(ENE)=NCG(N+1-ENE)/(NBASE(N+1-ENE))
X = rand()
WRITE(*,*) RO(1)
```

```

IF(X<.5*RO(ENE))THEN
LETRA = 'n'
ELSE IF(RO(ENE)/2.D0<X<RO(ENE)) THEN
LETRA = 't'
ELSE IF(RO(ENE) < X <0.5D0*1) THEN
LETRA = 'c'
ELSE
LETRA = 'g'
END IF
IF (X<RO(ENE)) THEN
NCG(N+1-ENE)=NCG(N+1-ENE)-1
END IF
NBASE(N+1-ENE)= NBASE(N+1-ENE-1)
write(7,*) X.LETRA
END DO
write(".*") X.LETRA
END DO
close(7)
end

```

APENDICE III

```
program CALORESPEC
  implicit none
  INTEGER I,P,T,N
  REAL(8) CE,RO,ROO,DELTAT,TE,U,Z,K,DELTAU,B,V,PSI,E,F
  REAL(8) R0
  PARAMETER(P=1000)
  DIMENSION TE(P),R0(P),CE(P),B(P),PSI(P),E(P),F(P)
  K= 1.381D-23 !Cte. de boltzmann
  U=-17.d-21 !Energía de AT
  DELTAU=-.9d-21 !diferencia de energías
  T=70.D0 !longitud del intervalo de temperatura
  Z=34.D0 ! coordinación
  ROO=.58D0 !densidad de CG promedio
  DELTAT= T/1000.D0 !tamaño de incremento
  TE(1)= 333.15d0 !temp. inicial
  N=900 !longitud de las cadenas
  V=(U+DELTAU*R0O)
  open(9,file='enerint.dat', STATUS='UNKNOWN')
```

```

open(10,file='enerlib.dat', STATUS='UNKNOWN')
DO I=1,P-1
B(I)=1.D0/(K*TE(I))
TE(I+1)= TE(I) + DELTAT
R0(I)=z*EXP(B(I)*V)
F(I)= N*V - LOG((1-(R0(I))**(N+1))/(1-R0(I)))
E(I)= (N+1)*V/((1-(R0(I))**(N+1))) - V/((1-R0(I)))
write(9,*) TE(I)-273.15d0,(E(I))/n
write(10,*) TE(I)-273.15d0,F(I)
END DO
close(9)
close(10)
end program CALORESPEC

```


Bibliografía

- [1] Genes & Genomes. M Singer y P. Berg. University Science Books. Blackwell Scientific Publications (1991).
- [2] I.E. Alcamo. DNA Technology, Hartcourt/ Academic Press. 2ª Edición 1999.
- [3] Ziman. Models of Disorder: The theoretical physics of homogeneously disordered systems. Cambridge University Press 1979.
- [4] G. Steger. Nucleic Acids Research 22. 14 (1994).
- [5] K. J. Breslauer. R. Frank. H. Blöcker y L. A. Marky. Proc. Natl. Acad. Sci. USA 83 (1986) Biochemistry.
- [6] A. Campa. A Giasanti. Phys Rev. E 58. 3 (1998).
- [7] M.E. Fisher. J. Ch. Phys. 45. 1469 (1966).
- [8] Y. Kafri. D. Mukamel y L. Peliti. Phys Rev. Lett. 85. 23 (2000).
- [9] D. Poland y H. Scheraga. J. Ch. Phys. 45. 1456 (1966); 45. 1464 (1966); Theory of Helix-Coil Transition in Biopolymers (Academic Press. New York. 1970).
- [10] S.J. Gould. S.T. Kaufmann. J.M. Smith. et al. La Biología del Futuro, Tus Quets, Metatemas 58 (1999).
- [11] H. Flyberg. J.Hertz. M.H. Jensen y G. Mouritsen. Physics of Biological Systems. Springer Verlag (1997).

- [12] C.R. Calladine y H.R. Drew, *Understanding DNA*, Academic Press, segunda edición (1997).
- [13] L. Pauling, *The Nature of the Chemical Bond*, Oxford University Press, cuarta edición (1963).
- [14] Genes V, B. Lewin. Oxford University Press (1994).
- [15] S.V. Buldyrev, N.V. Dokholyan, C.K. Peng, A.L. Goldberger, S. Havlin, G.M. Viswanathan, H.E. Stanley. *Physica A* 249 (1998).
- [16] C. Adami y N.J. Cerf. Elsevier Preprint, 7 Abril 1999.
- [17] J. Barral, A. Hasmy, J. Jiménez, A. Marcano, *Phys Rev. E* 61, 2 (2000).
- [18] W. Lee y L. Luo, *Phys Rev. E* 56, 1 (1997).
- [19] C.K. Peng, S.V. Buldyrev, A.L. Goldberger, S. Havlin, F. Sciortino, M. Simons y H.E. Stanley. *Nature*, 356 (1992).
- [20] Bioquímica, A.L. Lehninger, Ediciones Omega, décimotercera reimpresión, (1989).
- [21] L.E. Reichl. *A Modern Course in Statistical Physics*, University of Texas Press, 4ª reimpresión 1991.
- [22] E. Stanley, *Introduction to phase transitions and Critical Phenomena*, Oxford University Press 1987.
- [23] *Statistical Mechanics: Fundamentals and Modern Applications*, R.E. Wilde y S. Singh, John Wiley and Sons.
- [24] H.B. Callen, *Thermodynamics and Introduction to Thermostatistics*, John Wiley and Sons 2ª Ed..

- [25] C.E. Shannon y W Weaver. *The Mathematical Theory of Communications*. University of Illinois Press 1960.
- [26] J. Rosen. *Symmetry in Science*. Springer Verlag 1995.
- [27] U. Moulines. *Exploraciones Metacientíficas*. Alianza Universidad 38 (1982).
- [28] N.R. Hanson. *Patrones de Descubrimiento Observación y Explicación*. Alianza Universidad 117 (1971).
- [29] R.M. Wartell y A.S. Benight . *Phys. Rep.* 126. 67 (1985).
- [30] O. Gotoh y Y. Tagashira, *Biopolymers* 20 (1981), 1033; O. Gotoh, *Adv. Biophys.* 16 (1983) 1.
- [31] P.G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, 4ª edición 1993.
- [32] E.W. Montroll y N.S. Goel, *Biopolymers* 4. 855 (1966); R.M. Wartell y E.W. Montroll. *Adv. Chem. Phys.* 22. 129 (1972).
- [33] D. Cule. T. Hwa. *Phys Rev. Lett.* 79. 12 (1997).
- [34] T. Dauxois. M Peyrard y A. R. Bishop. *Phys Rev. E* 47. 1 (1993); M Peyrard y A. R. Bishop. B. Duplantier. *Phys Rev. Lett.* 62. 23 (1989).
- [35] O. Resendis y L.S. García Colín, *Application of the theory of Stochastic processes to the configuration of biological systems*. *Physica A* 290 (2001).
- [36] B. Duplantier. *Phys. Rev. Lett.* 57. 941 (1984); *J. Stat. Phys.* 54. 581 (1989).

TESIS CON
FALLA DE ORIGEN