

Superenrollamiento de ADN.

El ADN es una molécula enormemente larga que no es recta ni homogénea en toda su longitud. Muy al contrario, ya se ha visto que se puede curvar según las circunstancias pero, además, se puede enrollar o desenrollar respecto al modelo canónico definido por Watson y Crick. La variación en la helicidad se traduce en una tensión, tanto positiva como negativa, que se libera formando espirales de doble hélice en un sentido u otro.

Parámetros de enrollamiento.

Actualmente se puede definir una serie de parámetros para describir el estado de enrollamiento de una molécula o fragmento de ADN. Todos ellos se basan en el concepto de número de enlace.

Número de enlace.

Se define como el número de veces que una hebra se enrosca alrededor de la otra, y se representa por L (de *linking*, enlace en inglés). El número de enlace es, por su propia definición, constante en una molécula de ADN circular cerrada o en un fragmento que tenga extremos fijados a algún sitio, siempre que no se rompan las cadenas. Por otro lado, en ADN circular cerrado es necesariamente un número entero.

Este concepto es realmente difícil de imaginar sin conocer todavía los otros dos parámetros. Sin embargo, se hace intuitivo que en un ADN circular cerrado, el número de veces que una hebra da vueltas sobre la otra debe equivaler al número de vueltas totales de la doble hélice. De esto se puede extraer que en el caso de que no haya tensiones en una molécula cerrada, hay un número de enlace que llamaremos estable o ideal y denotaremos por L_0 y que, además, es también el número de vueltas que debería tener el ADN, es decir, el número de pares de bases dividido por 10,5 pares de bases por vuelta en el ADN B.

Número de torsión.

Se representa por T (de *twist*, torsión en inglés), y equivale al número real de vueltas de la doble hélice. ¿Y cuál es la diferencia con L ? La diferencia es que T sí puede cambiar según la forma que adquiera el ADN. Va a tender siempre a llegar al valor teórico del número de enlace estable L_0 , es decir, siempre procurará llegar a tener 10,5 pb/vuelta.

Número de contorsión.

Se representa por W (de *writhe*, contorsión en inglés), y es una medida del número de veces que se enrollan las dobles hélices una con otra. Da una idea de los superenrollamientos de la doble hélice. Ni este ni el número anterior tienen porqué ser enteros.

Los tres números se relacionan mediante la siguiente ecuación:

$$L = T + W$$

Partiendo del modelo del ADN circular cerrado, y suponiendo que no hay superenrollamientos ($W = 0$), entonces el número de vueltas y el número de enlace son iguales: $L = T$. Si tampoco hay tensiones, entonces los dos números anteriores son iguales al número de enlace ideal L_0 . Supongamos que se corta el ADN y uno de los extremos se gira una vuelta a derechas antes de volver a sellar los extremos, entonces se obtiene una molécula de ADN con una vuelta menos de la que le correspondería por el número de pares de bases, es decir, $L < L_0$. En este caso ya no es estable una molécula plana, porque el número de vueltas total es menor del que

correspondería, con lo que hay que aumentarlo en una unidad; es decir, con la molécula plana $T = L$, pero como hay una vuelta menos, hay más pares de bases por vuelta, con lo que hay que aumentar T en una unidad para volver al número ideal de pares de bases por vuelta o, lo que es lo mismo, hay que acercar T a L_0 para liberar la tensión producida por el desenrollamiento en una vuelta. Esto se consigue haciendo que el ADN se retuerza la vuelta que falta (regla mnemotécnica: si faltan vueltas en la doble hélice, que son a derechas, se meten superhélices a derechas). Como L , T y W están relacionados, y lo que buscamos es subir el valor de T , las vueltas a derechas de la superhélice se corresponderán con un decrecimiento de W ; es decir, partimos de que $L = L_0 - 1$ y, como con la molécula plana $L = T$, entonces $T < L_0$. Sabemos que L es invariable a menos que se rompan las hebras, con lo que si para acercarse al número ideal de vueltas hay que hacer $T + 1$, entonces W tendrá que bajar a $W - 1$ para que L siga constante. Se puede considerar que W contrarresta T para que L permanezca invariable. En este caso, como T aumenta y W disminuye, se dice que se produce un superenrollamiento negativo.

Conclusión: los superenrollamientos a derechas suponen decrecimientos de W , por lo que se llaman negativos, pero hacen que T aumente en tantas unidades como superenrollamientos.

El razonamiento análogo se puede considerar si aumentamos L en una unidad: se producirá un superenrollamiento a izquierdas (positivo) que reducirá el número de vueltas partiendo desde la molécula plana.

Todas las moléculas con igual número de pares de bases se llaman topoisómeros o isómeros topológicos, porque sólo se diferencian en la topología o estructura tridimensional de la doble hélice, es decir, sólo se diferencian en los valores de L , T y W .

Una manera distinta de relajar un ADN superenrollado negativamente es utilizar la energía provocada por la torsión en desenrollar las vueltas que sean necesarias para que la molécula recupere la estabilidad. Este fenómeno de interconversión entre superhélices negativas y desapareamiento de bases puntual es de extrema importancia en condiciones fisiológicas, ya que permite abrir la doble hélice en determinados sitios cuando se necesite, por ejemplo, en los procesos de replicación y transcripción para permitir la entrada de la polimerasa correspondiente. Además, y si la región desenrollada es un palíndromo, se podrían generar los cruciformes.

Según la forma que presente el ADN se consideran tres tipos:

- Tipo I: ADN superenrollado negativamente. Es el tipo normal en la célula.
- Tipo II: ADN circular. Se obtiene al cortar el ADN tipo I con una endonucleasa que corte sólo una de las dos cadenas.
- Tipo III: ADN lineal. Se obtiene al cortar ADN tipo I con una endonucleasa que corte ambas cadenas.

Los tres tipos son los que se pueden obtener de lisados con la mínima manipulación. El ADN superenrollado positivamente no es natural pero se puede fabricar.

Titulación de superenrollamientos.

Se puede determinar el número de superenrollamientos de una molécula mediante la utilización de agentes intercalantes. Éstos, como el bromuro de etidio, siempre distorsionan a doble hélice desenrollándola, es decir, disminuyendo el T al que una molécula determinada es estable. Con esto, y partiendo de que el ADN en medio fisiológico está superenrollado negativamente, al disminuir T se tiene que contrarrestar con un aumento equivalente de W , con lo que el efecto es que se eliminan superhélices negativas hasta llegar a no tener ninguna, y todavía se pueden provocar superhélices positivas al aumentar la concentración de agente intercalante. Se ha calculado que hacen falta 12,8 moléculas de agente intercalante para desenrollar una vuelta de hélice.

Se puede realizar una titulación del número de superenrollamiento mediante técnicas de sedimentación y de movilidad electroforética a distintas concentraciones del agente intercalante (Voet pág. 873). Las diferencias se aprecian en el valor absoluto de superenrollamiento, ya que un ADN con n o $-n$ superhélices tiene una forma equivalente. Teniendo en cuenta esto, cuanto mayor sea el número de supervueltas, indistintamente de su signo, mayor será el coeficiente de sedimentación y la movilidad electroforética.

Implicaciones biológicas del superenrollamiento.

En el cromosoma, tanto eucariótico como procariótico, el ADN no es circular cerrado, pero se organiza en dominios topológicos fijos y anclados a soportes proteicos, con lo que sí pueden presentarse fenómenos de superenrollamiento porque L no puede variar dentro de estos dominios. Concretamente, se calcula que en la célula hay una superhélice negativa cada 100 pb, aproximadamente. La densidad de superhélices (σ) es un parámetro topológico que se determina hallando el número de superhélices por cada diez pares de bases. De este modo la densidad normal de superhélices en la célula es de $\sigma = -0,1$. **Siempre que se habla de ADN desenrollado se refiere a que su número de enlace es menor del que correspondería, y no a que el número de vueltas real sea menor.**

Las implicaciones de que el ADN esté normalmente desenrollado son muy grandes porque, además de permitir la separación puntual e inmediata de las dos hebras, las superhélices negativas también se pueden relajar girando a izquierdas alrededor de un obstáculo (IMPORTANTE: Los superenrollamientos negativos pueden ser tanto superhélices a derechas de la propia doble hélice sobre sí misma, como giros a izquierdas de la doble hélice sobre un soporte, llamado superenrollamiento toroidal. Al final es la misma cosa, si lo lográis ver). Este fenómeno es el que permite la compactación del ADN y su organización en nucleosomas en el núcleo.

En los libros se cuenta que un sobrecruzamiento, es decir, una superhélice negativa arrollada a derechas, se corresponde con dos enrollamientos toroidales a izquierdas. Esto sólo es cierto para el caso de un ADN circular cerrado y sólo en el caso de que se considere el bucle externo como vuelta alrededor del nucleosoma. Es complicado de ver si no lo hace uno mismo con una cinta larga (puede servir un cinturón), y probando a mantenerla cerrada o abierta después y antes de darle las vueltas al nucleosoma. Realmente decir que un sobrecruzamiento equivale a dos vueltas toroidales es producto de una apreciación errónea del giro del ADN alrededor del soporte. Para resolver estas dudas viene muy bien la figura 28-36 del Voet, en la que se aprecia cómo no son dos vueltas alrededor de un soporte, sino una y lo que parece, pero no es, otra. Y todo esto pasa por circularizar el ADN.

Otra implicación importante del superenrollamiento negativo del ADN es durante la replicación y la transcripción. Ambos procesos necesitan una apertura inicial de la doble hélice, que se permite por el grado de desenrollamiento del ADN, pero al avanzar la polimerasa correspondiente por la doble hélice, se produce un apilamiento de vueltas delante y una falta de ellas detrás, lo que a razón de 100000 bases por minuto acabaría con el proceso en apenas un segundo (en realidad delante y detrás sigue habiendo el mismo número de vueltas, pero se distribuyen en menos y más bases, respectivamente). Esta acumulación se debe a varias causas, entre ellas a que la polimerasa no puede ir girando a esas velocidades alrededor del ADN, y a que el ADN tampoco puede girar libremente ni delante ni detrás de la enzima porque tiene sus extremos fijos (evidentemente, en un plásmido libre, la compresión por delante y la falta de vueltas detrás se anularían solos). La solución pasa por eliminar las vueltas de más que se generan delante mediante un conjunto de enzimas llamadas topoisomerasas, capaces de cambiar el número de enlace del ADN.

Topoisomerasas.

Son las enzimas encargadas de mantener el número y cantidad de superenrollamiento en una región, dominio o zona determinada del ADN según el índice y proceso metabólico que se dé en un momento dado. El mecanismo general de acción es cambiar el número de enlace de uno en uno o de dos en dos. Las

topoisomerasas son, por tanto, de dos tipos:

- Topo I: Aumentan de uno en uno el número de enlace, por lo que disminuyen o eliminan superhélices negativas. Esto lo consiguen rompiendo una de las dos hebras, fijando uno de los extremos a la enzima y haciendo pasar el otro por el otro lado de la hebra intacta. Como fisiológicamente el ADN está superenrollado negativamente, el proceso de eliminación de superhélices negativas es espontáneo. La energía necesaria para resellar la hebra viene de que la energía de la hidrólisis se conserva en un enlace fosfodiéster entre el extremo 5' y un residuo de Tyr. En las topo I eucarióticas se enlaza el extremo 3' de manera similar, y además pueden relajar ADN en cualquier sentido *in vitro*.
- Topo II: Disminuyen de dos en dos el número de enlace, por lo que aumentan el superenrollamiento negativo o, también, eliminan superenrollamientos positivos. El mecanismo de acción también pasa por la rotura pero, en este caso, de las dos hebras. Un mecanismo propuesto es que el ADN superenrollado positivamente gire alrededor de la enzima con una superhélice toroidal a derechas. Este sería el sustrato para que la enzima actuara cortando las dos hebras y pasando cada extremo por un lado distinto de la hélice intacta, volviendo a sellar en el otro lado. El producto es un ADN con un superenrollamiento toroidal a izquierdas, con lo que el número de enlace ha bajado en 2 en ese punto (+1 a -1). Cuando se parte de un ADN superenrollado positivamente el proceso es espontáneo, pero como en la célula siempre se parte de ADN superenrollado negativamente, introducir otra supervuelta negativa requiere energía. Sin ATP sólo puede relajar superhélices de cualquier signo, pero muy lentamente. La topo II eucariótica ni utiliza ATP ni introduce superenrollamientos negativos, sino sólo relaja las supervueltas de cualquier signo. Ambas unen transitoriamente el ADN por ambos extremos 5' mediante un residuo de fosfotirosina.

Los superenrollamientos se pueden ahora titular partiendo de un ADN superenrollado negativamente y un posterior tratamiento con topo I que, según el tiempo de reacción, se obtendrían bandas diferenciadas de cada grado individual de superenrollamiento en un gel de electroforesis. A su vez, se puede titular la actividad topo II al partir de un ADN totalmente relajado y en presencia de la enzima y ATP, con lo que se distinguirían bandas producidas por la incorporación de superhélices negativas hasta el máximo posible que cupiera en la molécula.

Estabilización de superhélices.

En el cromosoma bacteriano el ADN se presenta prácticamente libre de proteínas asociadas y los dominios son simples bucles de ácido nucleico relativamente libre y cuyo superenrollamiento adquiere conformaciones más o menos al azar, transitorias e inestables.

Los cromosomas eucarióticos, por su parte, estabilizan cada grado de desenrollamiento mediante la formación de estructuras supramoleculares cuya estructura se tratará más adelante pero que, esencialmente, se basan en la estabilización de los superenrollamientos negativos en forma de superhélices toroidales a izquierdas alrededor de una estructura proteica que hace de una especie de cilindro. Esta asociación de ADN alrededor de una estructura proteica se llama nucleosoma. Estas estructuras permiten estabilizar o proteger las superhélices, restringiendo el libre giro de una hebra sobre la otra cuando se produce un corte en una de ellas, de modo que no se pierda la superhelicidad del dominio entero en caso de que haya un pequeño corte de una sola de las cadenas. Se ha determinado mediante ensayos de digestión parcial de ADN que hay unos 200 pares de bases girados en dos vueltas alrededor del núcleo proteico de un nucleosoma. Una vez más, la apreciación no es correcta del todo, porque de esas casi dos vueltas toroidales a izquierdas, realmente es sólo una la que está distorsionada, y el resto son ramas de entrada y salida del ADN en forma B. En esos 200 pb hay, aproximadamente, un solo superenrollamiento negativo.

Superenrollamiento y la replicación y la transcripción del ADN.

Retomamos el problema planteado anteriormente de cómo se resuelve en el medio celular la compresión del

ADN delante de la polimerasa correspondiente. Ahora se puede decir qué topoisomerasa se necesita por delante, donde se acumulan las vueltas pero se queda sin bases suficientes para todas ellas, es decir, se crean superhélices positivas ($L > L_0$). La enzima bacteriana girasa es una topoisomerasa de tipo II, que recorre la doble hélice delante de la horquilla de replicación y elimina superenrollamientos positivos (baja L) a razón de 10000 por minuto, para que la ADN polimerasa pueda mantener el ritmo de 100000 bases por minuto. Durante la transcripción ocurre algo parecido pero, además, se genera demasiado desenrollamiento por detrás, ya que no se sintetiza ADN nuevo. Esto se debe a que, dado que la polimerasa no puede girar sobre sí misma porque tiene el ARNm y los ribosomas pegados a ella, además de que es demasiado grande, provoca realmente la división del dominio en otros dos, cada uno con sus características topológicas. Esto ha dado lugar a la teoría de los dominios gemelos, en la que una topo II actuaría por delante, y una topo I por detrás durante el proceso de transcripción. Teniendo en cuenta la organización del cromosoma bacteriano, en el que un gen puede estar transcribiéndose simultáneamente por tantas ARN polimerasas como quepan en su longitud, se hace necesario un estricto control para de la superhelicidad del ADN, aunque ya sólo sea para no romperlo. En eucariotas las superhélices, tanto positivas como negativas, se podrían relajar con cualquiera de los dos tipos de topoisomerasas.

La transcripción presenta, además, el fenómeno de la regulación. En procariotas la ARN polimerasa se ancla selectivamente a las secuencias -10 y -35 de la región del operador, antes de los genes estructurales del operón. La presencia de la enzima, unido a la de ADN curvo en esa zona, provoca un curvamiento y una unión fuerte y específica de la polimerasa cuyo efecto sería el de desenrollar el ADN (generar algo de superenrollamiento negativo) para poder provocar la transición hacia la separación de las dos hebras en la región de más debilidad, es decir, en la caja TATA (-10). El modelo de regulación mediante la superhelicidad del ADN prevé que la función de activadores e inhibidores que interaccionen tanto con el ADN como con la propia polimerasa provocarían, respectivamente, un mayor o menor superenrollamiento negativo, de modo que sea más o menos probable la transición al desapareamiento de bases. Todo esto sería necesario porque comenzar a transcribir el ADN hace falta que la ARN polimerasa se interponga entre las bases de una o dos vueltas.

Estructura cuaternaria de ADN.

Se considera estructura cuaternaria de ADN a las interacciones entre éste y las proteínas accesorias.

Estructura de la cromatina.

Es una cuestión evidente que para almacenar todo el ADN de un organismo, incluido un virus, hay que empaquetarlo hasta niveles que pueden llegar a ser de 5 ó 6 órdenes de magnitud. Este empaquetamiento se consigue sólo mediante la interacción con proteínas específicas. El estudio de la estructura cuaternaria de los ácidos nucleicos es necesariamente destructivo, con lo que se pierde mucha información en el abordaje experimental. Lo primero que hay que hacer es romper la cápsula o célula en la que esté incluido el ácido nucleico, con lo que se esparcen y pierden multitud de proteínas (estructurales, con funciones metabólicas, reguladoras) y ARN que pudieran aportar más información que la que realmente se obtiene.

Estructura en virus y bacterias.

Los virus generalmente tienen espacio suficiente en su cápsida para albergar todo su genoma sin más compresión que la generada al hacer un ovillo. De todos modos, las proteínas de la cápsida se unen específicamente a secuencias del ADN, y el tamaño es ligeramente ajustable según la cantidad que entre. Otro tipo de empaquetamiento es que la cápsida se enrolle alrededor de una molécula del ácido nucleico que se trate.

Las bacterias tienen su cromosoma en una estructura llamada nucleoide, y que suele ocupar 1/3 del volumen total de la bacteria. Dentro de un nucleoide se considera que puede haber entre 50 y 100 dominios de

topología y condiciones metabólicas diferentes, separados no se sabe cómo, y cada uno entre 50 y 100 Kpb. Dentro de ellos se observa un superenrollamiento negativo cada 200 pb, pero se considera que es sólo la mitad de los que realmente hay *in vivo*. Se sabe muy poco acerca de la estructura de este nucleóide, excepto las proporciones relativas de ADN (80%) y ARN y proteínas (20%). Al comparar estos datos con los del núcleo eucariota (33% de ADN, proteínas y ARN) se observa que en las bacterias hay muy pocas proteínas, con lo que el ADN está prácticamente desnudo, probablemente para maximizar el proceso de transcripción.

El ADN bacteriano se asocia a proteínas básicas de algún modo todavía no determinado, pero bastante débilmente, porque se pierden irremisiblemente en cualquier proceso de análisis estructural hasta la fecha. Aún así, se han detectado proteínas llamadas HU (HU1 y HU2) de entre 18 y 20 KDa, que se especula que forman un tetrámero al estilo de las histonas en el nucleosoma (ver más adelante), pero la interacción es tan débil que no se puede medir su tiempo medio de disociación del ADN. Esta unión tan laxa también hace hincapié en la idea de la maximización de la eficacia de los recursos y los procesos, tanto de transcripción como de replicación. Se ha comprobado que estas proteínas básicas inducen curvatura en el ADN en fragmentos tan pequeños como 80 pb, y que estabilizan superenrollamientos negativos porque mutaciones en sus genes *hupA* y *hupB* (en los dos) cursan con una disminución del número de superenrollamientos. La mutación ha de ser en ambos genes, ya que si no, no se observa cambio alguno, es decir, parece que ambos productos génicos son intercambiables o equivalentes, o también que hay otras que pueden realizar la función de una u otra. Por ahora no se conoce ninguna secuencia de unión específica para estas proteínas, por lo que la unión sería inespecífica.

Además de estas proteínas HU existen otras como la HNS (proteína parecida a H1; *H1-like*), que tiene afección por fragmentos de ADN curvo, y la proteína P (positiva), que no se sabe para qué sirve.

Estructura en eucariotas.

El conocimiento de la estructura de la cromatina eucariota es mayor que el de la bacteriana. Se conocen mejor las proteínas básicas asociadas al ADN, llamadas histonas, porque forman estructuras estables llamadas nucleosomas, presentes a lo largo de todo el ADN, formando de una fibra de 11 nm. Éstas se enrollan sobre sí mismas en una estructura de supersolenóide para formar fibras de 30 nm, cada vuelta del solenoide formada por seis nucleosomas consecutivos. Después se sabe muy poco hasta llegar a la estructura de máxima compactación que es el cromosoma metafásico, bastante estable. En él cada cromátida es una molécula de ADN comprimida hasta ser un cilindro de 800 nm de diámetro y algunas micras de largo.

La composición de la cromatina eucariota es un tercio de ADN, otro de histonas y otro de otras proteínas estructurales y con función metabólica y ARNhn. Las histonas se asocian al ADN sin especificidad de secuencia pero con una regularidad asombrosa y forman nucleosomas del modo que ya se ha descrito someramente, al estabilizar los superenrollamientos negativos. Una hilera de nucleosomas sobre la misma molécula de ADN forma una fibra visible al microscopio electrónico como una especie de collar de cuentas o rosario, en determinadas condiciones de baja sal, que tiene una anchura de 11 nm. En condiciones de salinidad alta se observa cómo los nucleosomas ya no están separados, sino que se pegan unos a otros sin poder ver espacios entre ellos. Esta agregación es posible porque se une al núcleo del nucleosoma otra proteína básica, llamada histona H1, que hace de una especie de grapa para que el ADN no se desenrolle. Cada nucleosoma completo puede asimilarse a un cilindro que, gracias a la mediación de H1, se puede llegar a asociar con el siguiente de modo que se genere un solenoide de nucleosomas, exactamente con seis por vuelta de solenoide y con las H1 en el centro. Estos solenoides serían las fibras de 30 nm observables en el microscopio electrónico en determinadas condiciones. Después de esto se sabe muy poco acerca de cómo se organizan estos solenoides, probablemente sobre una matriz o un armazón (*scaffold*) proteico no histónico. Se especula que pueden formar bucles de diferentes longitudes que para formar el cromosoma metafásico se apilan de determinada manera para formar cada cromátida.

En citología se ha distinguido desde hace mucho tiempo entre heterocromatina densa y eucromatina más

difusa según el grado diferencial de tinción, que se corresponde con la mayor o menor unión a proteínas. Generalmente se pueden asimilar a cromatina inactiva y activa, respectivamente, pero esto no es siempre del todo cierto, porque hay genes transcripcionalmente activos en la heterocromatina, y otros inactivos en la eucromatina. Además, hay heterocromatina constitutiva (permanente) y otra facultativa, que según del tejido y el estado metabólico de la célula se presenta o no.

No se sabe, hilando con lo anterior, si las zonas de unión de las fibras de 30 nm al armazón son estables o si se pueden cambiar según las necesidades. Actualmente se investigan las llamadas MAR (*matrix attachment regions*: regiones de unión a la matriz), para ver si son fijas o no. Las pocas que se han podido aislar no presentan ninguna homología secuencial asimilable a una misma secuencia consenso (apenas riqueza en C y A) y, además, se ha detectado que algunas de estas MAR son regiones de regulación en cis (contiguo al gen) de genes transcripcionalmente activos en uno de los bucles adyacentes. Por todos estos datos, parece que la unión a la matriz o al armazón es dinámica durante la interfase. En la metafase la unión de la cromatina al armazón para formar cromosomas es más estable pero, a la vista de estos resultados, parece que, además, es inespecífica.

Nucleosomas.

Los nucleosomas son las partículas estructurales básicas de la cromatina eucariótica. Ya se ha avanzado que consisten en un núcleo proteico alrededor del cual se arrollan dos superhélices negativas de ADN de manera toroidal y levógira. Al tratar suavemente la cromatina con ADNasa micrococcal, que corta ADNds, se pueden obtener oligómeros de nucleosomas que, tras un análisis electroforético, se observa que abarcan siempre múltiplos de 200 pb, con lo que se puede concluir que cada nucleosoma contiene unos 200 pb. La separación en oligómeros también indica que hay tramos de ADN desnudo y muy expuesto, pero que el 90% de esos 200 pb está protegido por el núcleo proteico de histonas. Esto permite diferenciar entre el ADN del núcleo del nucleosoma, que será básicamente constante, y el ADN de unión, al que se deberán las posibles variaciones en la longitud total de cada nucleosoma.

Esta última apreciación se demuestra tratando los nucleosomas aislados con nucleasas. En principio se obtienen fragmentos de 166 pb, correspondientes al ADN completo que da dos vueltas enteras alrededor del núcleo proteico y está todavía asegurado por la histona H1, que se encuentra por fuera. Esta disposición obliga a la doble hélice a entrar y salir por el mismo sitio, obligando a los nucleosomas a empaquetarse unos junto a otros cuando están unidos. Si se aumenta el tiempo de digestión, o si se disminuye la fuerza iónica, la histona H1 se libera y entonces se obtiene un fragmento de ADN de 146 pb unido al resto de proteínas del núcleo del nucleosoma: las histonas H2A, H2B, H3 y H4. Como regla general, la histona H1 interacciona con unos 22 pb del ADN de unión.

El núcleo del nucleosoma está constituido por dos unidades de cada histona, ensambladas como el tetrámero [H4·H3·H3·H4] al que se unen dos dímeros [H2A·H2B]. El conjunto se asemeja a un cilindro de 6 nm de alto por 6,4 nm de diámetro y de un peso molecular de 108000 Da (H2A: 14000; H2B: 14000; H3: 15000; H4: 11000). Sobre este núcleo proteico (denominado octámero) se enrollan de promedio unos 200 pb de ADN, cuyo peso es de 130000 Da, y se asegura con la histona H1 que tiene, también de promedio, unos 24000 Da. El resultado es un cilindro tan alto como el anterior, pero con 10,4 nm de diámetro, una estructura formada a partes iguales por ADN y proteínas y estabilizada mediante interacciones electrostáticas entre los fosfatos del ADN y los aminoácidos básicos (His, Lys, Arg) de las histonas.

Las histonas que componen el octámero están enormemente conservadas en todos los eucariotas. De ellas la más constante es la H4, que es exactamente igual en todos los animales estudiados hasta la fecha, y sólo se han encontrado cambios conservadores en dos o tres aminoácidos entre los de animales y algunas plantas superiores. La segunda más conservada es la H3, y posteriormente las H2A y H2B, con más variabilidad dentro de grupos filogenéticamente emparentados. La extremada conservación de H3 y H4 se debe a que se ha comprobado que ellas solas son capaces de formar estructuras parecidas a los nucleosomas, uniéndose al ADN

sin la participación de H2A y H2B. El caso de H1 es distinto, porque su misión principal no es formar nucleosomas, sino unirlos o empaquetarlos de diferente manera, con lo que presenta una variabilidad muy grande incluso dependiente del momento del desarrollo o del tejido que se considere. De hecho, en los eritrocitos de pollo la H1 es tan diferente a las demás que hay quien la llama histona H5.

Esta gran diversidad de la histona H1 es, probablemente, la causa de que el ADN de unión no tenga una longitud constante, sino que pueda variar entre 8 y 114 pb (54 pb de media). En cualquier caso, siempre habrá 146 pb alrededor del octámero, que se corresponden con 13 vueltas de hélice, según una digestión con ADNasa I, que provoca cortes en una sola de las cadenas de ADNds. El grado de contorsión de estas 1,7 vueltas de la doble hélice no es constante, sino que se mantiene bastante estable, salvo en sitios donde se produce un giro brusco. Estos curvamientos se deben principalmente al efecto de las proteínas (ya se ha dicho que la unión es inespecífica de secuencia), pero también a pares AA y GC en el sitio. Estas distorsiones son muy acusadas, y provocan cambios importantes en la anchura de los surcos en el ADN que se reflejan en cambios de la contorsión, que pasa de 10 pb/vuelta en los segmentos de entrada y salida a 10,7 pb/vuelta en los centrales. Es en estos segmentos contorsionados donde se generan entre 0,7 y 0,8 superenrollamientos negativos, porque son éstos los que recorren el núcleo proteico oblicuamente para poder salir por abajo y entrar por arriba.

Periodicidad de los nucleosomas.

Ya se ha avanzado que la unión del ADN a los nucleosomas es independiente de la secuencia pero, sin embargo, los nucleosomas se forman con una regularidad asombrosa. Esto llevó a desarrollar experimentos en los que se pondría de manifiesto cuál sería el patrón de formación. Se selecciona una secuencia con sitios de restricción específicos y se marca radioactivamente. Después se ensamblan las histonas y se corta el ADN en nucleosomas. Luego se digiere con la enzima de restricción de uno de los extremos y se somete a electroforesis y a autorradiografía. Si fuera la secuencia la que determinara el lugar del nucleosoma, sólo se obtendría una banda nítida (estarían los nucleosomas en fase); si, por el contrario, no hubiera absolutamente ninguna determinación por parte de la secuencia, se observaría una mancha difusa y uniforme (los nucleosomas no estarían en fase). El resultado obtenido es que la situación parece intermedia. Actualmente se postula que debe haber una secuencia fija sobre la que se forma rápidamente el primer nucleosoma, mientras que los demás, sencillamente, se forman después sobre su referencia y a distancias marcadas por la histona H1.

Replicación y transcripción con nucleosomas.

Durante la replicación los nucleosomas se desorganizan y reorganizan inmediatamente, con lo que no hay grandes espacios sin proteger al menos en la hebra conductora. Cabe la pregunta de si la replicación de las histonas es también semiconservativa o sigue un patrón distinto al del ADN. En principio parece que cada octámero es una mezcla de proteínas viejas y nuevas, pero la hebra conductora adquiere rápidamente sus octámeros, mientras que la hebra retardada no, probablemente debido a la presencia de fragmentos de Okazaki y a las proteínas SSB (*single stranded binding proteins*) que se unen a fragmentos de ADNss.

In vivo las histonas no se encuentran libres en el nucleoplasma, porque precipitarían debido a las fuertes interacciones que se dan entre ellas. Para evitar esto existen las proteínas N1 y la nucleoplasmina, muy ácidas, que transportan y protegen (o al revés) tetrámeros H3H4 y dímeros H2A2B. Además de estas proteínas transportadoras, que se pueden considerar chaperonas moleculares, debe haber alguna otra (que a lo mejor no lo es, llámalo X) que indique los sitios que corresponderán al ADN de unión, probablemente una proteína del grupo HMG (*high mobility group*: grupo de alta movilidad [electroforética]) y que algunos se aventuran a señalar como el dímero 14/17. Evidentemente, esto es un modelo y debe considerarse como algo transitorio.

Durante la transcripción hace falta eliminar los nucleosomas porque la ARN polimerasa es demasiado grande como para que pudieran permanecer (recordar que lleva consigo, además, una hebra de ARN_h y un montón

de proteínas accesorias). Esta necesidad llevada al extremo se observa en sitios con una intensa actividad transcripcional, como la de los genes ARNr, donde el ADN aparece totalmente desprovisto de nucleosomas. Esto se puede comprobar midiendo la longitud del ARN naciente y comparándola con la del ADN que lo codifica. La mayor o menor estructuración de los nucleosomas se puede relacionar con una menor o mayor tasa de transcripción, hecho que se observa en los genes activados por calor de *Drosophila*. En esta aproximación se aprecia que cuando el gen no está inducido, la digestión con ADNasa micrococcal genera el patrón de bandas característico de la presencia de nucleosomas, pero al subir la temperatura ya no se ven bandas, sino una mancha difusa, lo que indica que la nucleasa corta en cualquier sitio y no debe haber nucleosomas. Esta correlación entre la mayor o menor organización en nucleosomas y la tasa de transcripción puede utilizarse para la determinación de la actividad de un gen en determinados estados del desarrollo o en determinados tejidos. Para ello se trata el ADN con la nucleasa micrococcal a distintas concentraciones, y luego se digiere con una enzima de restricción que se sabe que proporciona un fragmento que hibrida con una sonda. Si el gen no es transcripcionalmente activo, siempre aparecerá una banda en cada pocillo, independientemente de la cantidad de nucleasa que se hubiera puesto, pero si la transcripción es activa, cuanta más ADNasa hubiera, menores serían los fragmentos recogidos, con lo que quedarían menos fragmentos capaces de hibridar con la sonda después del tratamiento con la enzima de restricción y, por tanto, la intensidad de la banda disminuiría hasta desaparecer.

Regulación del empaquetamiento mediante modificación covalente.

Las histonas, sobre todo la H1, reciben numerosas modificaciones posttraduccionales y sufren regulación covalente en su unión a ADN. Así, se observan acetilaciones de Lys y fosforilaciones de Ser, principalmente, que influyen distintamente sobre la interacción con ADN.

Las histonas H3 y H4 pueden acetilarse para evitar una interacción demasiado fuerte con el ADN, y funcionalmente se ha relacionado con el hecho de que se favorece la transcripción cuando esto ocurre. La histona H1 se fosforila para provocar un efecto parecido pero, sin embargo, se ha comprobado que se necesita esta fosforilación para formar cromosomas metafásicos, que representan el grado máximo de empaquetamiento de la cromatina, lo que parece una paradoja. Otras proteínas que se fosforilan son N1 y nucleoplasmina, que unen así histona más fuertemente y, quizás sirva para formar más nucleosomas y compactar así el ADN, pero esto también parece una contradicción con lo anterior.

Interacciones del ADN con otras proteínas.

Las histonas se han descartado como elementos de regulación fina porque su unión al ADN es inespecífica de la secuencia, con lo que la unión de cualquiera de estas proteínas al ADN debe ser mediante el reconocimiento específico de una secuencia concreta. Aunque hay muchas acciones mediadas por la interacción entre ADN y proteínas (desenrollamiento local, reparación, recombinación, regulación de otros reguladores, metilación de ADN), aquí sólo se hablará de los factores de transcripción y regulación, que son proteínas de tamaño medio o pequeño, generalmente dimericas y simétricas respecto a un eje, capaces así de unirse a una molécula en forma de hélice bicatenaria antiparalela. En general se pueden clasificar estas proteínas en cuatro tipos según su unión al ADN:

- Mediante una hélice α (tipo hélice–bucle–hélice).
- Mediante dos hélices α (tipo cremallera de leucina).
- Mediante una lámina β antiparalela.
- Mediante dedos de cinc.

Proteína tipo hélice–bucle–hélice: el represor del fago 434.

Esta proteína es sintetizada en la propia bacteria y se une a los genes líticos del fago cuando está infectada para impedir su transcripción. Es una proteína dimerica con dos subunidades iguales, cada una con cinco

hélices δ , de las que la tercera es la que interacciona con el surco mayor del ADN con bases de cada cadena, debido al giro de la doble hélice. La otra hélice del motivo (2) se encarga de mantener en su lugar la otra y, además, interacciona con los fosfatos del esqueleto. La estructura doblemente simétrica de la molécula se corresponde en el ADN en un palíndromo que abarca casi una vuelta y media (14 pb) de la doble hélice, pero en el que sólo están implicados en la interacción cuatro nucleótidos de cada extremo. En este caso, la proteína contacta con el ADN por un lado y en el mismo sitio de dos vueltas consecutivas de la doble hélice.

Las proteínas de los factores de transcripción del tipo homeobox (regulan el patrón de desarrollo durante la embriogénesis) también se unen a ADN mediante motivos hélice–bucle–hélice.

Tanto en este caso como en todos los demás tipos, se hace necesaria una interacción topológica, es decir, la curvatura, superhelicidad y torsión local del ADN debe ser la que encaje en la proteína y viceversa. Del mismo modo, la interacción puede modificar la topología del ADN como las interacciones que se dan entre enzimas y sus sustratos correspondientes. También común a la mayoría de proteínas de unión al ADN, las interacciones se pueden clasificar como puentes de hidrógeno entre aminoácidos y el esqueleto de azúcar–fosfato, con las bases, e interacciones hidrofóbicas entre aminoácidos apolares y las bases del ADN.

Proteína tipo cremallera de leucina.

Estos factores suelen ser activadores de genes, generalmente implicados en el desarrollo del ciclo celular (*myc*, *fos*, *jun*), por lo que se estudian intensamente en el campo de la Oncología. Destaca el hecho de que la mayoría pueden formar heterodímeros, pero en cualquier caso se componen de un tallo que es la cremallera y dos brazos que interaccionan con el surco mayor del ADN en una secuencia de unos 8 pb que, en el caso de ser un homodímero, es un palíndromo. Esto quiere decir que interacciona por ambos lados de una sola vuelta. También en este caso y debido al giro del ADN, la interacción se realiza por cada lado con bases de ambas cadenas.

Proteína tipo lámina δ antiparalela: el represor del operón Met.

Vuelve a ser un dímero en el que una lámina δ formada por un grupo lineal de cada subunidad encaja en el surco mayor y establece interacciones con unas pocas bases de cada cadena, en un palíndromo que abarca sólo 6 pb.

Proteína del tipo dedo de cinc.

Las proteínas con dedos de cinc no son dímeros, sino monómeros y, por tanto, no interaccionan con secuencias palindrómicas, sino lineales, generalmente sólo sobre una de las hebras, a la que van recorriendo por el surco mayor.

Un dedo de cinc es una estructura básica que se repite modularmente de dos a 16 veces en este tipo de proteínas, constituida por dos grupos lineales δ en horquilla seguidos de una hélice δ de tres vueltas y colocada sobre la lámina que forman los anteriores. Esto define un hueco, cerca del extremo más abierto, en el que cabe un átomo de cinc, que establece cuatro enlaces de coordinación con sendos aminoácidos, dos His en las vueltas 2ª y 3ª de la hélice y una Cys en cada grupo lineal que, además, mantiene reducidas. Además, en la base de la unión entre la lámina y la hélice se establecen interacciones hidrofóbicas entre tres aminoácidos, uno por cada elemento lineal: Tyr en el extremo N–terminal del primer grupo lineal δ , Phe en el extremo C–terminal del segundo grupo lineal, y Leu en la primera vuelta de la hélice δ .

No hay un patrón definido de interacción de los dedos de cinc con el ADN, pero se observa una abundancia de interacciones entre Arg y G, con tres interacciones por dedo de cinc.

Proteínas de unión al surco menor.

Es bastante más extraño que una proteína se una al surco menor, principalmente porque no hay sitio, y también porque en ese lado las bases no presentan casi información. Sin embargo, cada día se descubren factores nuevos que se unen específicamente a secuencias de regulación por el lado del surco menor, como el TBP que ya se ha comentado y que, además, provoca una gran distorsión en la topología del ADN. La interacción se establece entre una región hidrófoba en la proteína y la caja TATA.

Conclusiones.

De toda la cantidad de datos puntuales recopilados hasta la fecha, no se puede extraer ninguna idea general, salvo que la interacción es extremadamente específica de secuencia y la topología, y de la interacción misma. La existencia de una simetría palindrómica o no depende de si la proteína es ella misma simétrica y, además, la interacción se establece por, pero no depende sólo de los aminoácidos implicados en una sola de las hélices o de las láminas. Generalmente son aminoácidos polares los que se encargan de establecer las interacciones, pero no hay un código de reconocimiento establecido o hallado hasta la fecha, con lo que no es posible todavía diseñar un factor a la carta y, además, porque no se conoce por qué funcionan estos factores ni cómo regulan la expresión de un gen. Este es un campo muy importante de investigación hoy día.