

DOBLE HÉLICE, GENES Y CROMOSOMAS

LUIS FRANCO VERA *

* Universidad de Valencia. Departamento de Bioquímica. Facultad de Biológicas. Burjasoto 46100 Valencia

INTRODUCCIÓN

El título de la presente contribución sigue un orden lógico. Tendremos ocasión de ver cómo el DNA, con su conocida estructura en *doble hélice*, es el componente fundamental de los *genes* y cómo éstos se encuentran localizados en los *cromosomas*. Parece lógico, a simple vista, que un discurso intelectual se remonte desde lo más sencillo a lo más complejo. Pero no coincide el orden lógico con el cronológico: históricamente se elaboró primero el concepto de gen aún dentro del siglo XIX, luego se descubrieron experimentalmente los cromosomas en los albores del siglo XX y la propuesta de la doble hélice como modelo estructural del DNA data de sólo unos 50 años. Por otro lado, como es frecuente en el desarrollo de una ciencia experimental, el avance en el conocimiento de los objetos de nuestra consideración, genes, cromosomas y doble hélice, ha sufrido innumerables idas y venidas hasta llegar a su estado actual. Por ese motivo, aunque el orden de este capítulo corresponda más o menos con el cronológico, será preciso dar numerosos saltos, hacia atrás o hacia adelante en el tiempo, para llegar a un compromiso entre la lógica y la cronología histórica.

No deja de tener interés ese esfuerzo por comprender la lógica de un proceso biológico, especialmente lo que Lehninger llamaba la *lógica molecular de la vida*. No hay que ver en esta actitud un intento de juzgar a la naturaleza reduciéndola a unas categorías antropocéntricas. Pero no cabe duda de que, de esa manera, se logra un conocimiento más profundo de los

fenómenos biológicos. Por otra parte, esa *dimensión lógica* que observamos en la naturaleza favorece la capacidad de asombro que, lejos de constituir una actitud acientífica, proporciona al investigador un sustrato que, con frecuencia, dispara su creatividad y dota a toda persona culta interesada por la naturaleza de unas armas sumamente útiles para adentrarse en su comprensión.

GENES

Es imposible hablar de los genes sin mencionar a Gregor Mendel (figura 1). Era un fraile agustino, que pasó casi toda su vida en el monasterio de Brno, en la actual república Checa. Tenía una gran afición por las Ciencias Naturales e intentó varias veces, sin éxito, dedicarse a la docencia universitaria. Tuvo que conformarse con experimentar en un rincón del jardín de su monasterio. Pero esas circunstancias resultaron providenciales. Dotado de una singular capacidad de observación, de un rigor poco frecuente y de una insaciable curiosidad, logró que el tranquilo ambiente que le rodeaba, lejos del ajetreo de las ciudades universitarias de su época, se convirtiera en el caldo de cultivo ideal para uno de los grandes descubrimientos de la ciencia moderna. Mendel se dedicó a observar cómo se heredaban los caracteres, color y aspecto de las semillas, color y forma de las flores, de plantas de guisante. Llegó así a establecer, en un prodigio de rigor intelectual las leyes de la herencia biológica que llevan su nombre. Pero, en el contexto del presente capítulo, lo que interesa es que a Mendel se debe la primera for-

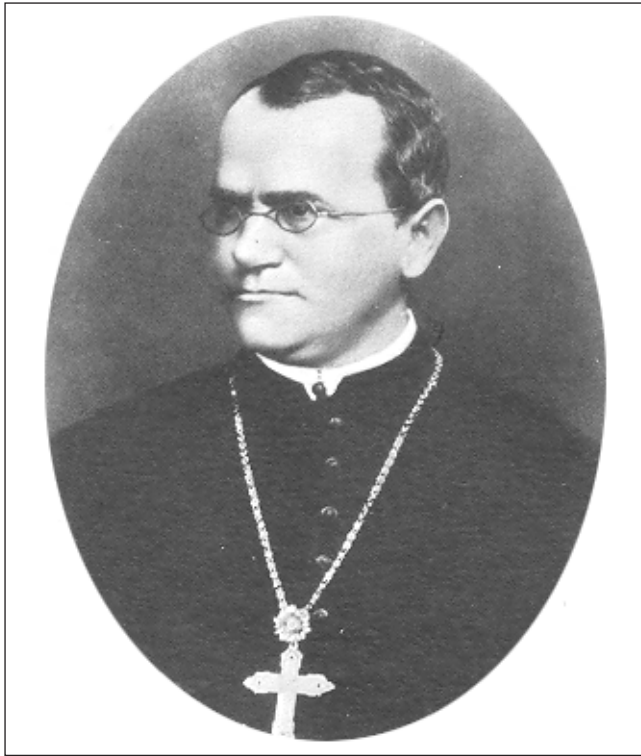


Figura 1. Gregor Mendel (1822-1884).

mulación del concepto de gen, como factor particulado responsable de la herencia de un carácter. Según esa primitiva concepción, un gen sería responsable del color amarillo de las semillas de guisante, otro del color verde, otro de su aspecto rugoso, etc. En esa definición originaria, de 1866, hay un detalle que resulta de la máxima importancia. Un gen es un “factor particulado”, es decir, no se trata de una abstracción, ni de algo ideal que sólo tenga cabida en la mente del investigador. No; los genes son partículas materiales y, por lo tanto, susceptibles de ser estudiados por los métodos de las ciencias empíricas. Podemos llegar a conocer la composición química de los genes, podemos analizarlos, podemos sintetizarlos...

Admitiendo la definición de Mendel, esas metas no eran sueños irrealizables. No obstante, hubo que esperar. Hubo que esperar, primero porque los resultados de Mendel se publicaron en una oscura revista y pasaron inadvertidos para sus contemporáneos; en segundo término, porque estaban fuera del alcance de la ciencia de la época. El redescubrimiento de las investigaciones de Mendel se produce en los umbrales del siglo XX, en el mismo año, 1900, en el que Sutton

y Boveri postulan la teoría cromosómica de la herencia. Estos autores encontraron que los cromosomas, partículas visibles al microscopio en células en división, poseían unas propiedades que los hacían aptos para ser portadores de la información genética que, precisamente al dividirse una célula, debía pasar a las células hijas.

Naturalmente, la comparación de los resultados de Sutton y Boveri con los de Mendel inmediatamente sugería que los genes, esas partículas materiales postuladas por el fraile agustino, estaban contenidas en los cromosomas. Era el primer avance en el esfuerzo multidisciplinario que ha conseguido hacer de la investigación en esta línea tema uno de los temas de mayor interés actual. En el caso presente, se dio la conjunción de los aportaciones genéticas con las citológicas y quedaba abierta la puerta para la entrada de las aportaciones bioquímicas que contemplaremos en el siguiente apartado.

Pero antes es preciso, con la idea expuesta en la introducción, hacer un apunte *lógico*. Los genes, si son responsables de la herencia biológica, han de contener unas *instrucciones* que, al ser ejecutadas, den lugar a los diversos caracteres fenotípicos, u observables. Pero es un hecho de experiencia primaria que esos caracteres son propios de las especies y que se transmiten a la descendencia de acuerdo con unas leyes precisas, las leyes de Mendel. Naturalmente, cuando una célula se divide en dos, cada una de las células resultantes debe heredar la misma información. Eso significa que el material que constituya los genes debe dividirse, replicarse diremos dentro de poco en el lenguaje técnico, de modo que donde había una *instrucción* pase a haber dos copias de la misma que puedan repartirse entre las células hijas. Es preciso, además, que la célula haga la copia de la información con la mayor fidelidad posible. De otro modo, si se admitieran errores, podría resultar bien información sin sentido, incapaz de ser ejecutada por la célula, bien información alterada, que diera lugar a la ejecución de una orden distinta de la presente en el gen de la célula madre. Estas alteraciones —mutaciones— en la información genética se dan de hecho. Aunque sea un fenómeno que ocurra con escasa frecuencia, tienen gran importancia. Es cierto que la mayor parte de las veces son perjudiciales, incluso letales, para las células, pero también es cierto que son las mutaciones

las que permiten la variación del contenido informativo de los organismos, un proceso clave para hacer posible la evolución biológica. En resumen, pues, la lógica nos dice que el material del que estén compuestos los genes ha de ser capaz de replicarse con fidelidad, pero al mismo tiempo debe admitir las mutaciones, los errores en la replicación.

LOS GENES Y SUS PRODUCTOS

Los genes contienen instrucciones que han de ejecutarse. Al fin y al cabo esto es algo a lo que estamos acostumbrados en la vida ordinaria. ¿No es cierto que todos los días vemos numerosísimas instrucciones? Por ejemplo, ¿quién no ha recibido uno de esos sobres preparados para una apertura fácil, en los que aparece la instrucción: “cortar por la línea de puntos”? Pero está claro que la instrucción no basta. Es preciso leerla, entenderla. Y es preciso ejecutarla, bien sea rasgando a mano, bien empleando unas tijeras. Una vez más, la lógica nos dice que no basta con que existan los genes; se precisa una maquinaria capaz de entender sus instrucciones y de ejecutar las órdenes que contienen. Los biólogos pronto se dieron cuenta de estos requerimientos, de modo que en esa conjunción de Genética y Citología que permitió localizar los genes en los cromosomas, la Bioquímica tenía que aportar datos para resolver dos cuestiones. No bastaba con contestar la pregunta “¿cuál es la naturaleza molecular de los genes?”, sino que era menester despejar también el interrogante: ¿qué naturaleza molecular tienen los productos de los genes?

Durante los primeros años del siglo XX se fueron acumulando datos que ponían de manifiesto la importancia de las proteínas en todos los procesos celulares. En efecto, se fue comprobando que las proteínas no sólo desempeñan funciones estructurales o de reserva, sino que participan en todos los aspectos de la dinámica celular. Proteínas son las enzimas¹, los anticuerpos, muchos reguladores, receptores, transportadores de moléculas dentro de la célula o entre distintas partes de un organismo, etc. Por ello, la identificación

de los productos de los genes como proteínas fue poco a poco ganando terreno. En paralelo, los bioquímicos fueron avanzando en el desciframiento de la estructura de las proteínas. A finales del siglo XIX se sabía que las proteínas estaban formadas por aminoácidos y se habían caracterizado casi todos ellos, aunque la estructura de la metionina, el último de los aminoácidos proteicos en ser descubierto, no se describiría hasta 1935. Ya en 1902 Fischer y Hofmeister habían propuesto que los aminoácidos proteicos, que poseen en su estructura un grupo carboxilo y un grupo amino, se unen entre sí a través de enlaces peptídicos, amidas secundarias, que pueden dar lugar a polipéptidos, cadenas que resultan de la unión de un número grande de aminoácidos. Las proteínas son, pues, polipéptidos, polímeros lineales de aminoácidos, cuyo orden de encadenamiento constituye la denominada estructura primaria.

A mediados del siglo XX, la aplicación de métodos físicos, especialmente difracción de rayos X, al estudio de la estructura de proteínas permitió estudiar el modo con que las cadenas polipeptídicas se organizan en el espacio. Así pues, y propiciados por el método de la Biología Molecular que nació con vocación integradora y estructuralista, aparecieron en escena los conceptos de estructuras secundaria y terciaria², que añadían al puramente químico de estructura primaria una dimensión espacial.

Es precisamente en esa época cuando surge una nueva definición de gen, como consecuencia del avance sobre la estructura y función de proteínas. Concretamente, ese nuevo avance se data en 1945, pero había sido precedido de importante descubrimientos en la década anterior. En el transcurso de estudios de genética bioquímica —otra vez aparece en escena la multidisciplinariedad— Ephrussi y sus colaboradores estudiaron la síntesis de los pigmentos oculares de *Drosophila melanogaster*, que tiene lugar a partir del triptófano. Algunos mutantes de *Drosophila* se caracterizan por tener ojos con una pigmentación diferente de la normal (figura 2). Parecía evidente que en dichos mutantes fallaba de alguna manera la ruta de

¹ Desde hace años se sabe que algunas enzimas, las ribozimas, están constituidas por RNA pero, aunque en los primeros estadios de la evolución biológica las ribozimas fueran muy importantes, en la actualidad representan una ínfima parte de las enzimas.

² El lector interesado puede encontrar las descripciones de estos niveles estructurales de las cadenas polipeptídicas, así como del de estructura cuaternaria al que se aludirá más adelante, en cualquiera de los excelentes textos de Bioquímica que existen en la actualidad.

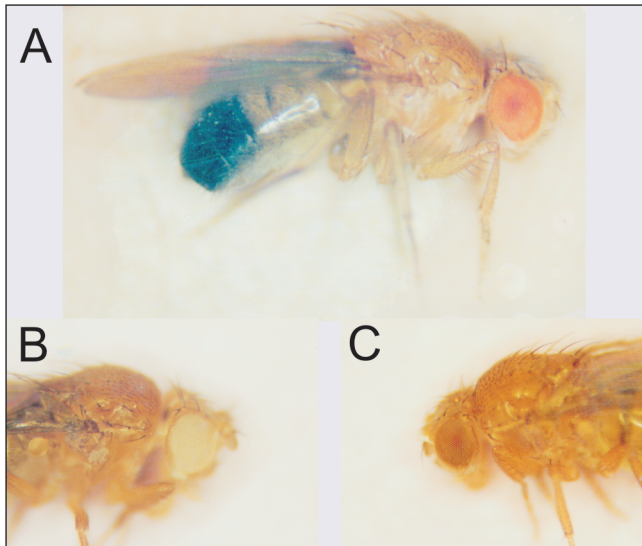


Figura 2. Ejemplares de *Drosophila melanogaster*. A: Cepa salvaje (Oregon R); B: Mutante *white*; C: Mutante *sepia*. Cortesía del Departamento de Genética de la Universidad de Valencia (Estudi General).

biosíntesis que conduce al pigmento normal y, en su lugar, se acumulaba otro diferente. Con estos antecedentes, y teniendo en cuenta que cada etapa de una ruta metabólica está catalizada por una enzima, Beadle y Tatum concluyeron en 1945 que las enzimas eran los productos de los genes, de modo que la diversidad de enzimas de una célula, que es responsable en gran medida de la diversidad de sus funciones, sería consecuencia de la diversidad de genes. La hipótesis de Beadle y Tatum se plasmó en una célebre frase: un gen, una enzima. Con ella se quería significar que la información contenida en los genes se traducía, fundamentalmente, en la síntesis de enzimas. Pero hay otras proteínas que no son de naturaleza enzimática y tienen también un papel decisivo en la dinámica celular, por lo que el adverbio “fundamentalmente” que aparece más arriba deja abierta la función de los genes de modo que, sin faltar al espíritu de Beadle y Tatum, su hipótesis se podría enunciar: un gen, una proteína.

Durante los años 40, continúa avanzando la Bioquímica en el estudio de las proteínas y se van acumulando datos que indican que muchas proteínas están

compuestas por más de un tipo de cadenas polipeptídicas. La hemoglobina, que se encarga del transporte de oxígeno en los vertebrados (figura 3) constituye un caso paradigmático de estas proteínas. En su estructura cuaternaria, participan dos tipos de cadenas, α y β , que difieren en su estructura primaria y son producto de dos genes diferentes. Así pues, es menester precisar aún más la hipótesis de Beadle y Tatum, de manera que quede formulada como: un gen, una cadena polipeptídica.

El conocimiento de la naturaleza de sus productos nos ha permitido avanzar en el concepto de gen. Pero aún no se ha comentado nada sobre la naturaleza de los genes. En realidad, así se encontraba históricamente la ciencia a mediados de la década de 1940. Una cosa era evidente: si los genes están contenidos en los cromosomas, su naturaleza molecular debía corresponder con alguno de los componentes de los cromosomas. O, si se quiere, de la cromatina, nombre que recibe el material genético en el núcleo en interfase³.

DOBLE HÉLICE

EL DNA, MATERIAL GENÉTICO

Los primeros análisis bioquímicos pusieron de manifiesto que la cromatina está constituida por proteínas y un ácido nucleico, el ácido desoxirribonucleico (DNA). Los ácidos nucleicos son polímeros formados por nucleótidos, que, a su vez, están constituidos por una pentosa —ribosa en el ácido ribonucleico (RNA) y dextrosirribosa en el DNA—, una base nitrogenada y fosfato. Hay dos tipos de bases nitrogenadas, las púricas y las pirimidínicas. Son purinas la adenina (A) y la guanina (G), que se encuentran tanto en el RNA como en el DNA. Las pirimidinas son: la citosina (C), presente en ambos ácidos nucleicos, el uracilo (U), característico del RNA y la timina (T), que típicamente se halla en el DNA.

La estructura covalente de los ácidos nucleicos (figura 4) se estaba investigando ya en los años 40⁴.

³ En el ciclo celular se pueden definir varias fases. La mitosis, que coincide con el momento de división celular, se caracteriza por la condensación del material genético en forma de cromosomas. El resto del ciclo celular, que media entre dos divisiones sucesivas y ocupa la mayor parte del tiempo, se denomina genéricamente interfase. Se subdivide en varias etapas y es el periodo de máxima actividad metabólica y genética de la célula.

⁴ Fue Todd quien, en 1951, describió la estructura covalente de las cadenas de polidesoxirribonucleótidos que se recoge en la figura 4.

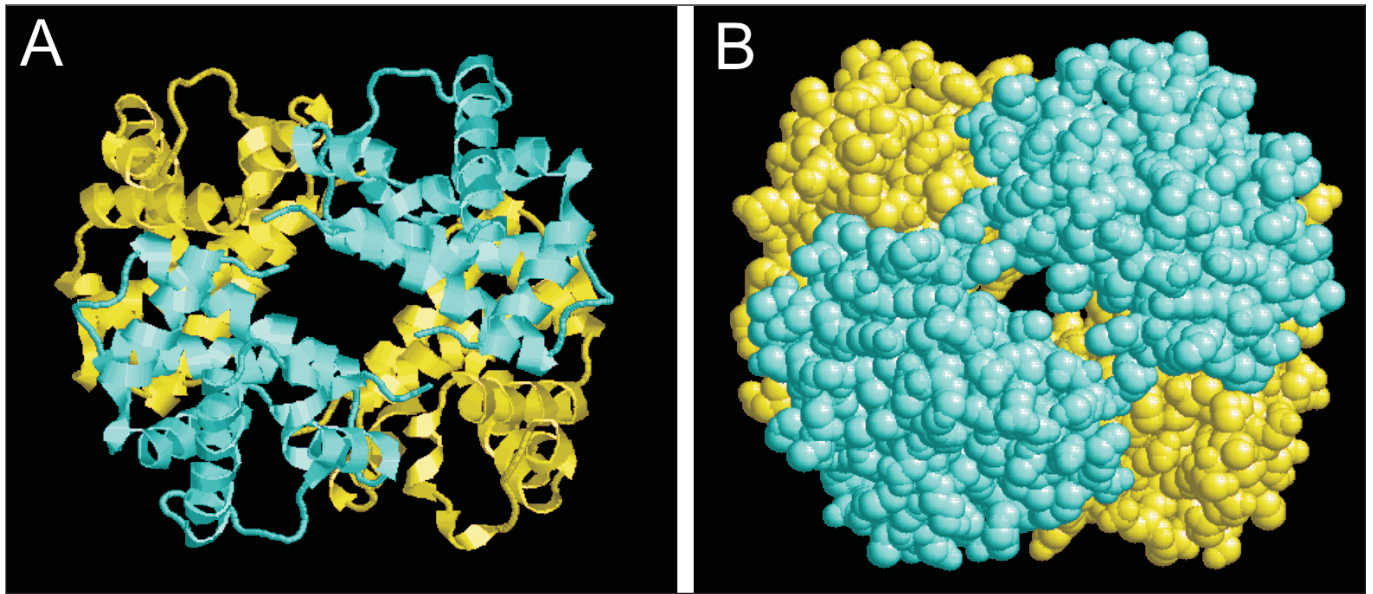


Figura 3. Estructura de la hemoglobina humana, mostrando su estructura cuaternaria, formada por dos subunidades α (azul) y otras dos β (amarillo). En la parte A sólo se representa esquemáticamente el esqueleto polipeptídico para mostrar la estructura secundaria (hélices α). En la parte B se representan todos los átomos de las cadenas. Ambas figuras están construidas con el programa RasMol.

Pero existía la idea generalizada de que el DNA era un mero componente estructural de la cromatina y de los cromosomas, de modo que serían las proteínas cromosomales las portadoras de la información genética y, por tanto, las constitutivas de los genes. Esa idea se basaba en una consideración apriorística: el hecho de que sólo hubiera 4 desoxirribonucleótidos y, sin embargo, 20 aminoácidos, daba la impresión de que las proteínas admitirían un mayor grado de variabilidad estructural que el DNA, algo que parecía importante para explicar la variabilidad de los genes.

Sin embargo, en 1943 ocurre un acontecimiento trascendental. Avery, McLeod y McCarthy, estudiando la transformación genética de *Streptococcus pneumoniae* comenzaron a sospechar que era el DNA el portador de la información. Así se lo contaba el propio Avery a su hermano en una carta fechada el 13 de mayo de 1943:

“...But at last perhaps we have it. The active substance is not digested by crystalline trypsin or chymotrypsin, it does not lose activity when treated with crystalline ribonuclease... polysaccharides can be removed... Lipids can be extracted... without impairing biological activity (...) When extracts (...) are further fractionated by the dropwise addition of absolute ethyl alcohol an interesting thing occurs (...) There separates out a fibrous substance (which

on elementary analysis conforms very closely to the theoretical values of pure deoxyribose nucleic acid (...). If we are right, then it means that nucleic acids are not merely structurally important but functionally active substances in determining the biochemical activities and specific characteristics of the cells (...). But today it takes a lot of well documented evidence to convince anyone that the sodium salt of deoxyribonucleic acid, protein free, could possibly be endowed with such biologically active and specific properties and that is the evidence we are now trying to get. It is lots of fun to blow bubbles but it is wiser to prick them yourself before someone else tries to”

EL DESCUBRIMIENTO DE LA ESTRUCTURA DEL DNA

Avery era consciente de las dificultades que iban a encontrar para convencer a la comunidad científica de su descubrimiento, pero lograron acumular las pruebas suficientes y un año más tarde publicaron sus resultados. Con todo, fueron recibidos con escepticismo y pocos científicos aceptaron la revolucionaria idea. Entre los convencidos estaban todos los investigadores que algo más tarde se embarcaron en la aventura de dilucidar la estructura espacial del DNA. Cupo a James Watson y a Francis Crick la gloria de llevar a buen puerto esa aventura, inaugurando la era de la

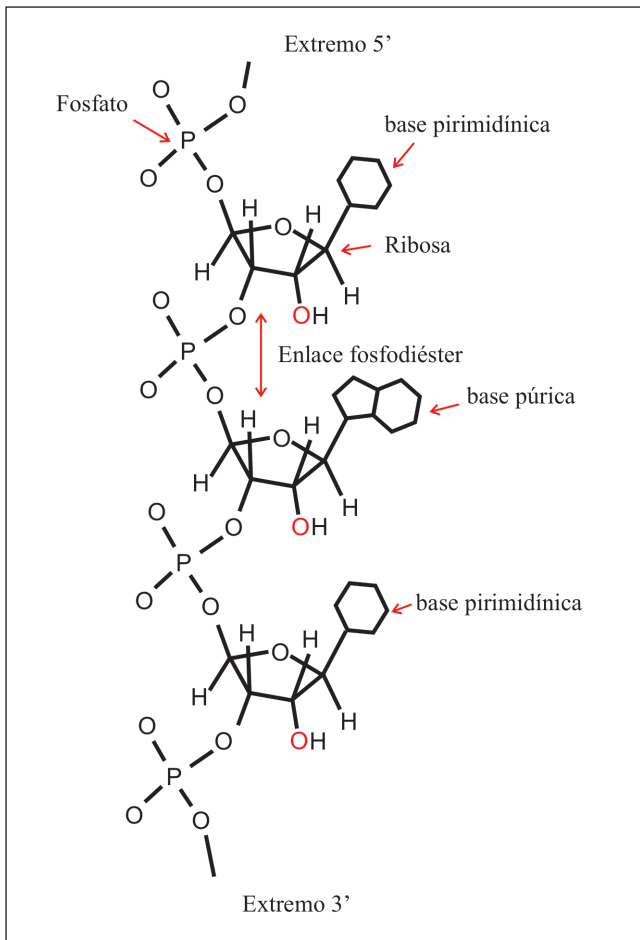


Figura 4. Estructura covalente de una cadena de ácido nucleico. La cadena representada corresponde a un segmento de RNA, ya que la pentosa es ribosa. Las cadenas de DNA tienen una organización covalente similar, pero la ribosa está sustituida por desoxirribosa, cuya estructura resulta al eliminar el átomo de oxígeno representado en rojo.

Biología moderna. Watson y Crick no hicieron experimentos; se basaron en los resultados de otros. Pero tuvieron el mérito indiscutible de ver en esos resultados lo que sus propios autores no habían visto.

Fueron decisivos, por ejemplo, los resultados obtenidos por Chargaff, que había determinado que en el DNA la proporción de adenina era aproximadamente igual a la de timina y la de guanina igual a la de citosina. También fueron fundamentales los datos logrados con difracción de rayos X de fibras de DNA en el laboratorio de Wilkins. La interpretación de esos

datos indicaba claramente que el DNA poseía una estructura helicoidal, lo que llevó a algunos autores a proponer una estructura en la que tres cadenas de polidesoxirribonucleótidos se enrollaban entre sí, de modo que los fosfatos quedarían próximos al eje de la estructura y las bases se orientarían hacia fuera. Otros investigadores habían propuesto también diferentes modelos basados en la organización helicoidal. Pero ninguno de ellos era satisfactorio. Watson y Crick, a base de construir modelos moleculares, encontraron una posible explicación para los resultados de Chargaff. El anillo de timina podía unirse con el de adenina formando puentes de hidrógeno. Lo mismo ocurría con las bases guanina y citosina (figura 5). Si eso ocurría en el DNA, y dos cadenas se entrelazaban de modo que se unieran mediante esos puentes de hidrógeno entre las bases, quedaban justificadas las “reglas de Chargaff”. Más aún, el par A-T tenía unas dimensiones prácticamente idénticas que las del par G-C. Esto era de la máxima importancia. Si el DNA estuviera formado por dos cadenas, la separación entre los esqueletos desoxirribosa-fosfato de ambas sería constante, independiente de la naturaleza del par de bases. Watson y Crick, trabajando otra vez con modelos moleculares, llegaron a una estructura que satisfacía todos los requisitos estructurales y era compatible con todos los datos experimentales. Acababa de nacer la doble hélice (figura 6).

Watson y Crick publicaron esquemáticamente la estructura propuesta en un breve artículo, que ocupaba una sola página, pero que abrió un horizonte insospechado para la Biología y aún para toda la ciencia moderna. La estructura postulada en ese artículo corresponde a lo que actualmente se denomina estructura B del DNA, y es en realidad una estructura arquetípica, ya que, en realidad, el DNA, manteniendo su organización helicoidal bicatenaria, presenta una cierta microheterogeneidad estructural, de forma que parámetros tales como paso de rosca, inclinación de los pares de bases con respecto al eje de la hélice, ángulo de giro de un par de bases con respecto al siguiente, etc. no son constantes a lo largo de la doble hélice, sino que dependen de la secuencia de bases concreta de cada tramo⁵.

⁵ En general, se puede decir que la estructura del DNA oscila entre la del la forma B y la de la forma A, descrita inicialmente como característica del DNA en medios de menor polaridad que el acuoso. Caso especial es el de la estructura Z, presente sólo en regiones muy concretas del genoma, que difiere en la configuración de algunos enlaces glicosídicos y en el sentido de giro de las cadenas.

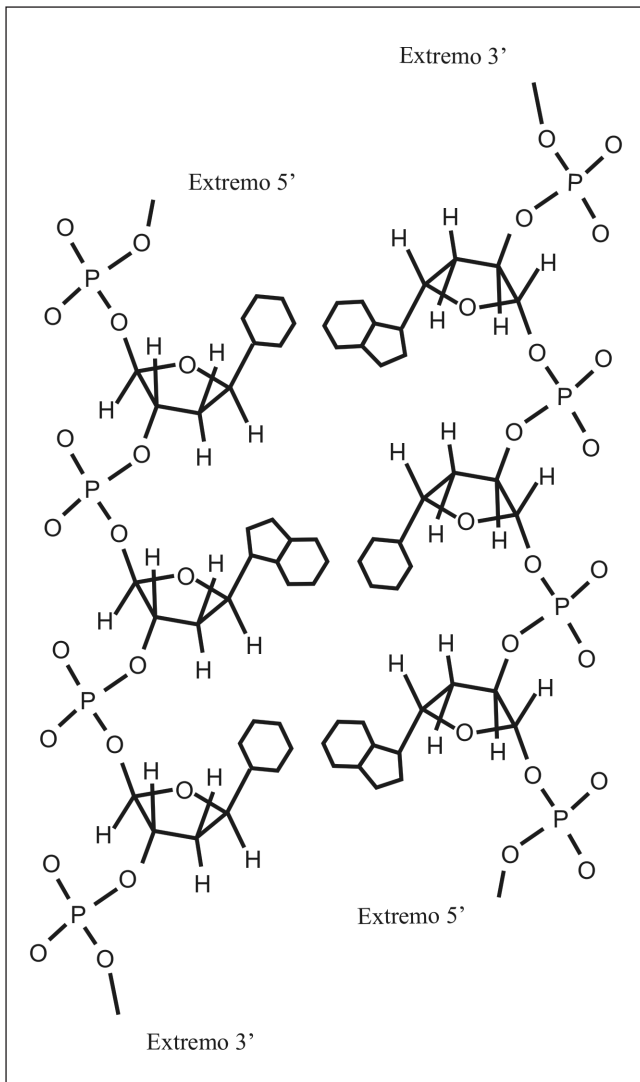


Figura 5. Dos cadenas antiparalelas de polidesoxirribonucleótidos pueden interactuar entre sí por puentes de hidrógeno (no mostrados) entre adenina y timina o entre guanina y citosina. Advértase que las dimensiones del par A-T son iguales a las del par G-C, lo que permite que los esqueletos de desoxirribosa-fosfato (en el exterior), presenten una separación uniforme.

LA FUNCIONALIDAD DE LA DOBLE HÉLICE: UN NUEVO CONCEPTO DE GEN

Se quedaría muy corto quien pensara que el artículo de Watson y Crick sólo abordaba una cuestión estructural. Cerca del final del artículo, los autores decían textualmente:

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Y es que la Biología Molecular, emergente aún en aquellos años, había nacido con esa mentalidad que contempla la estructura a la luz de la función y viceversa. Lo que quizá faltó a los predecesores de Watson y Crick fue tener en cuenta que la estructura que se postulara para el DNA tenía que dar cuenta de su función que, como se ha comentado, implica la replicación. Efectivamente, poco después de la publicación del artículo que venimos comentando, Watson y Crick, en la misma revista, *Nature*, publicaron otro con el clarificador título “Genetic implications of the structure of deoxyribonucleic acid”, en el que postulaban que, al ser las dos cadenas del DNA complementarias en su secuencia, si se separaban, podría construirse sobre el molde de cada una de las cadenas antiguas, otra nueva, por yuxtaposición de nucleótidos complementarios. De este modo resultarían dos moléculas hijas, cada una con una cadena antigua y otra complementaria recién construida, que serían idénticas a la molécula parental. No es necesario añadir —ya es de dominio público— que el tiempo dio la razón a Watson y Crick.

La estructura del DNA permite, pues, la replicación y, con ella, la constancia hereditaria en la división celular. Pero no hay que olvidar que no basta que la información se transmita a las células hijas, puesto que

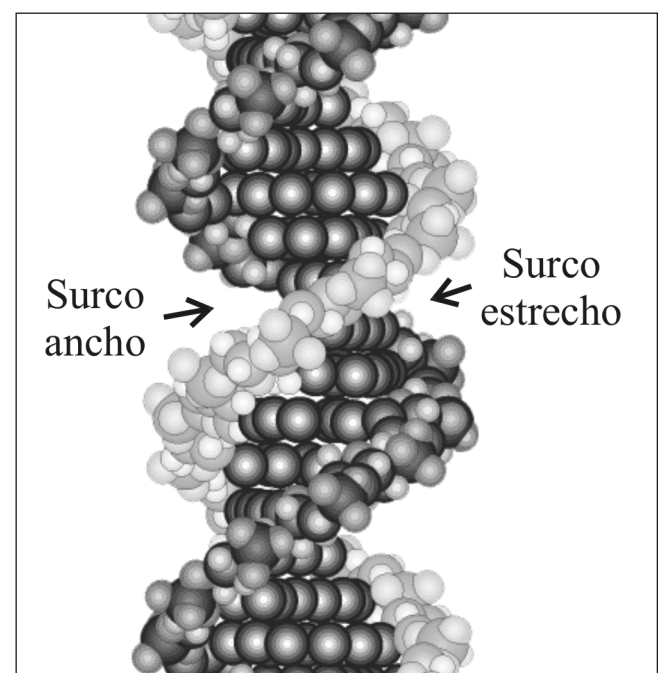


Figura 6. La doble hélice, con los parámetros propuestos por Watson y Crick, que corresponden a la estructura B del DNA.

su función es la de expresarse para dar lugar a las proteínas, los productos génicos y ejecutores de la información. Nuevamente la consideración de la función biológica permitió a Crick postular que en el flujo de información desde el DNA a las proteínas el RNA desempeñaba un papel intermediario. La organización covalente del RNA es similar a la del DNA, pero en su estructura se dan algunas importantes diferencias (figura 4). En primer lugar, la pentosa es ribosa, aunque el hidroxilo extra en 2' no participa en la concatenación de nucleótidos. En segundo lugar, salvo excepciones, no existe timina en el RNA y esa base está sustituida por el uracilo. Y, finalmente, el rasgo estructural más sobresaliente es que el RNA es monocatenario. El hecho de que el uracilo pueda formar enlaces de hidrógeno específicos con la adenina, sugiere inmediatamente un mecanismo para la transmisión de la información desde el DNA al RNA. Crick postuló que la información contenida en un gen, podía emplearse para determinar la secuencia de un RNA si una de las cadenas del DNA actuaba como molde. El ensamblamiento de nucleótidos ocurriría de modo parecido al de la replicación, con dos salvedades: la ya indicada que sólo se copia a RNA la información de una de las cadenas del DNA y el hecho de que en los ribonucleótidos a incorporar a la cadena el uracilo ocuparía el lugar de la timina. Este proceso recibió el nombre gráfico de transcripción.

Los idiomas que emplean un alfabeto distinto del nuestro presentan una dificultad adicional para comprender un texto escrito. Cuando, por ejemplo, vemos la palabra rusa *Спасибо*, si no conocemos el alfabeto cirílico, ni siquiera sabremos qué letras contiene. Transcribir esa palabra es volverla a escribir utilizando el equivalente latino de cada signo cirílico. Así resulta el vocablo *Spasibo*, pero es necesario traducirlo para comprender su significado: "gracias". Por eso, se propuso el nombre de traducción para el proceso mediante el cual la información pasaba del RNA a las proteínas. Quedaba así completo el esquema del flujo de información en las células (figura 7) que constituye lo que se pasó a denominar el "dogma central" de la Biología Molecular. Aunque años más tarde se demostraría que el dogma admite excepciones⁶, en líneas

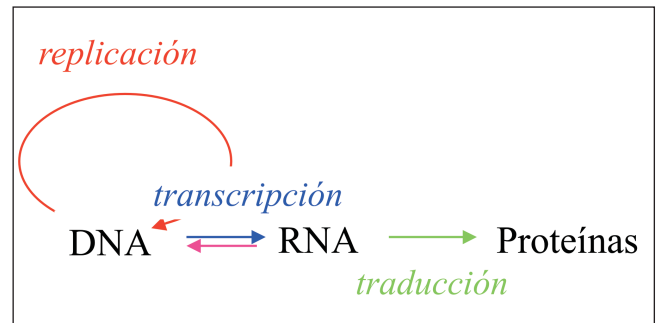


Figura 7. Esquema del metabolismo informativo, que muestra cómo la información genética fluye del DNA a las proteínas a través del RNA. La flecha rosa, de RNA a DNA, corresponde al proceso de *retrotranscripción*, que constituye una excepción al "dogma central".

generales sigue siendo, por supuesto, válido. El esclarecimiento de este flujo de información permitió redefinir el concepto de gen. Antes de formarse una proteína, la información contenida en su gen debe transcribirse para dar lugar a un RNA mensajero (mRNA), es decir, el producto inmediato de los genes es un RNA. Más aún, existen algunos tipos de RNA, como el ribosomal (rRNA) o el de transferencia (tRNA), que desempeñan funciones auxiliares en la traducción, aunque ellos mismos no se traduzcan. También pueden llamarse genes las unidades informativas del DNA que dan lugar a esas moléculas de RNA, con lo que se echa de ver que la definición de un gen como la unidad de información que dirige la síntesis de una cadena polipeptídica, se queda corta. Después de la elaboración del "dogma central", el aforismo: un gen, una cadena polipeptídica, pasó a ser: un gen, un RNA.

Los sucesivos avances en la comprensión del mecanismo de la transcripción pusieron de manifiesto que la mayoría de los genes eucarióticos poseen secuencias internas, denominadas intrones, que, aunque se transcriben, no se traducen. La transcripción, en estos numerosos casos, da lugar a una molécula primaria que ha de procesarse para llegar al mRNA maduro. El procesamiento implica algunas modificaciones en los extremos del transcrito primario, que hacen posible la exportación del mRNA desde el núcleo al citoplasma, pero fundamentalmente supone la eliminación de los

⁶ Los retrovirus, como el causante del SIDA, contienen RNA como material genético. Al introducirse en la célula huésped, el RNA se copia para formar DNA, que se integra posteriormente en el genoma de la célula infectada. Este proceso de flujo de información de RNA a DNA se denomina retrotranscripción.

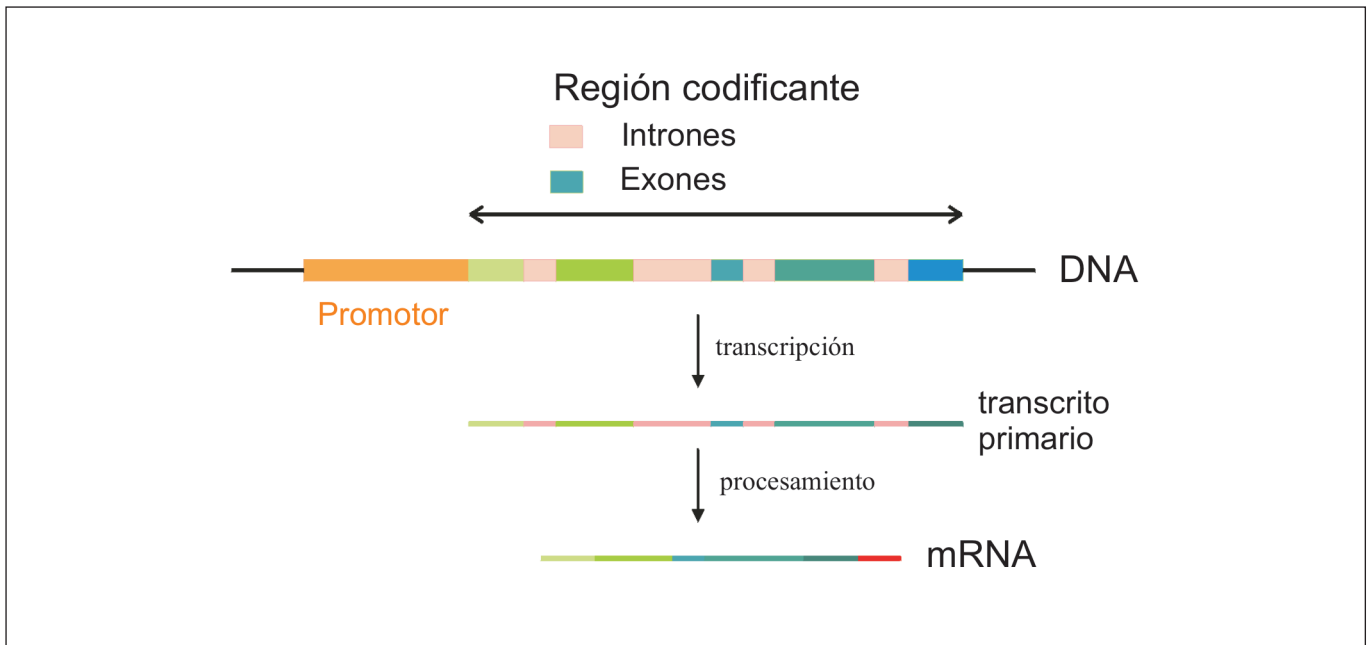


Figura 8. Esquema de un gen de un eucariota superior, que muestra la presencia de intrones dentro de la región codificante y el procesamiento del transcrito primario de DNA para eliminar dichos intrones y añadir otras secuencias antes de producir el mRNA maduro.

intrones (figura 8). Y es en este contexto donde surge una complicación adicional, que obligará, una vez más, a reconsiderar la definición de gen. En algunos casos, pueden darse formas alternativas de procesamiento, con lo que de un transcrito primario pueden surgir diferentes mensajeros y, por tanto, diferentes proteínas. Llegados a este punto, se puede ver que, aunque se haya precisado notablemente el concepto de gen, resulta realmente difícil expresar ese concepto en un aforismo breve.

CROMOSOMAS

LAS HISTONAS

Por aproximaciones sucesivas, se ha revisado hasta este momento el concepto de gen. Se ha contemplado la estructura de la doble hélice, lo que nos ha permitido, en función del principio antes aludido, esbozar su función. También se ha comentado que el DNA, el material con contenido informativo de los cromosomas, está acompañado en ellos por proteínas. Puede ser el momento de preguntarse, ¿qué naturaleza tienen esas proteínas?, ¿qué función desempeñan? La respuesta a la primera pregunta se conocía desde el siglo

XIX. Aunque con las imperfecciones propias de la época, Miescher había conseguido aislar en 1884 unas abundantes proteínas nucleares que denominó histonas. Pasado el tiempo se comprobaría que las histonas contienen una elevada proporción de los aminoácidos básicos lisina y arginina, mientras que su contenido en aminoácidos ácidos es sumamente bajo.

En 1950 Edgar y Ellen Stedman formularon una sugerente hipótesis sobre el papel de las histonas. Puesto que ya se sabía, tras los experimentos de Avery, que el DNA era el portador de la información genética, los Stedman propusieron que las histonas, proteínas básicas y, por tanto, capaces de unirse fuertemente al DNA, serían represores específicos de los genes. La hipótesis tenía el atractivo de que ofrecía una explicación para la diferenciación que se observa en organismos pluricelulares. En efecto, aunque todas las células somáticas de un organismo poseen el mismo DNA —y, por tanto, los mismos genes— el conjunto de genes expresados por cada línea celular es diferente. Si las histonas, que pueden interaccionar con gran afinidad con el DNA, se unieran específicamente a ciertos genes, éstos quedarían reprimidos, mientras que los desprovistos de histonas podrían expresar su función. Naturalmente, esto implicaba que las his-

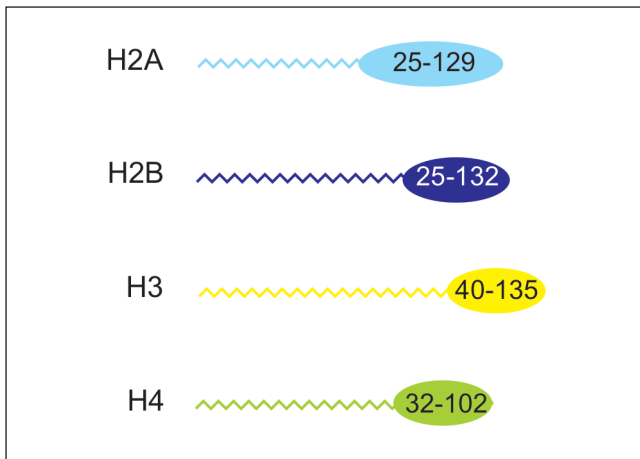


Figura 9. Estructura esquemática de las "histonas internas", que forman parte de la partícula núcleo. En los cuatro casos, las regiones N-terminales, desestructuradas en disolución, se representan como líneas en zigzag y el resto de las moléculas se esquematiza como una elipse. Los números que aparecen en ella indican la extensión (en residuos de aminoácidos) de las regiones estructuradas en las histonas humanas, aunque no se trate estrictamente de "regiones globulares" (véase el texto).

tonas, al unirse en unas células con determinados genes y en otras con otros diferentes, debían ser específicas de especie y de tejido.

La hipótesis resultó ser falsa, pero tuvo la virtud de orientar el interés de los investigadores hacia unas proteínas que, aunque conocidas desde hacía más de sesenta años, habían recibido escasa atención. De esa manera, en poco más de una década se consiguieron caracterizar las histonas. En todos los eucariotas sólo existen cinco clases de histonas denominadas H1, H2A, H2B, H3 y H4⁷. Por otro lado, las histonas constituyen un caso llamativo de conservatividad evolutiva. La estructura primaria de las histonas H4 aisladas de embrión de guisante y de timo de ternera se determinó en 1968 y, para asombro de los científicos, se comprobó que ambas poseen 102 aminoácidos de los que 100 son idénticos y sólo hay variaciones — además conservativas⁸— en los dos restantes. Evidentemente, esa conservatividad —que más tarde se confirmaría, en mayor o menor grado, para el resto de las histonas— implicó el definitivo abandono de la

hipótesis de los Stedman. Las histonas pasaron a considerarse como proteínas estructurales de la cromatina pero con una importante precisión: si estaban tan conservadas sería porque su papel en la organización de la cromatina era crítico.

De esta manera se inició una apasionante carrera para descubrir la estructura de la cromatina. El problema no era fácil. Para hacerse una idea de su complejidad, basta con tener en cuenta que los 6 pg de DNA de una célula somática humana tienen una longitud cercana a 2 m. Estas asombrosas dimensiones son consecuencia de los parámetros estructurales de la doble hélice: un diámetro aproximado de 2 nm y una separación entre cada par de bases de unos 0,34 nm a lo largo del eje de la hélice. Era evidente que el DNA tenía que enrollarse y superenrollarse para caber dentro de un núcleo, que puede tener un diámetro de unos 10 μm .

Los primeros datos vinieron de la mano de la microscopía electrónica. Al observar cromatina a baja fuerza iónica, se apreciaban unas fibras de un diámetro aproximado de 10 nm. Al ser un diámetro superior al de la doble hélice, parecía obvio que el DNA, al interactuar con las histonas, adoptaría una organización más compacta. La difracción de rayos X de cromatina en las mismas condiciones en que se observaban las fibras de 10 nm indicó que existía una organización helicoidal superpuesta a la del DNA. Un error de interpretación hizo suponer que la cromatina se organizaba como una superhélice continua con un paso de rosca de 11 nm y un diámetro de 10 nm. Este modelo, planteado en 1968, tuvo una aceptación general y rápida.

El papel de las histonas en el mantenimiento de esta estructura nunca llegó a definirse con precisión. Con todo, también había consenso en una cuestión. Al analizar la estructura primaria de las histonas H2A, H2B, H3 y H4 se observaba un rasgo común: mientras que el tercio N-terminal posee una notable abundancia de aminoácidos básicos y una casi total ausencia de aminoácidos apolares, los dos tercios restantes de las

⁷ En algunos tipos celulares, como los eritrocitos nucleados de aves, la histona H1 está parcialmente sustituida por la H5, que presenta unas características muy similares. Por otro lado, aunque el número de clases de histonas sea tan reducido, en casi todas las clases existen diferentes variantes, que difieren muy poco entre sí.

⁸ Variaciones conservativas son aquéllas en las que un aminoácido está sustituido por otro de similares propiedades químicas. En el caso presente las dos variaciones son: lisina por arginina y valina por isoleucina.

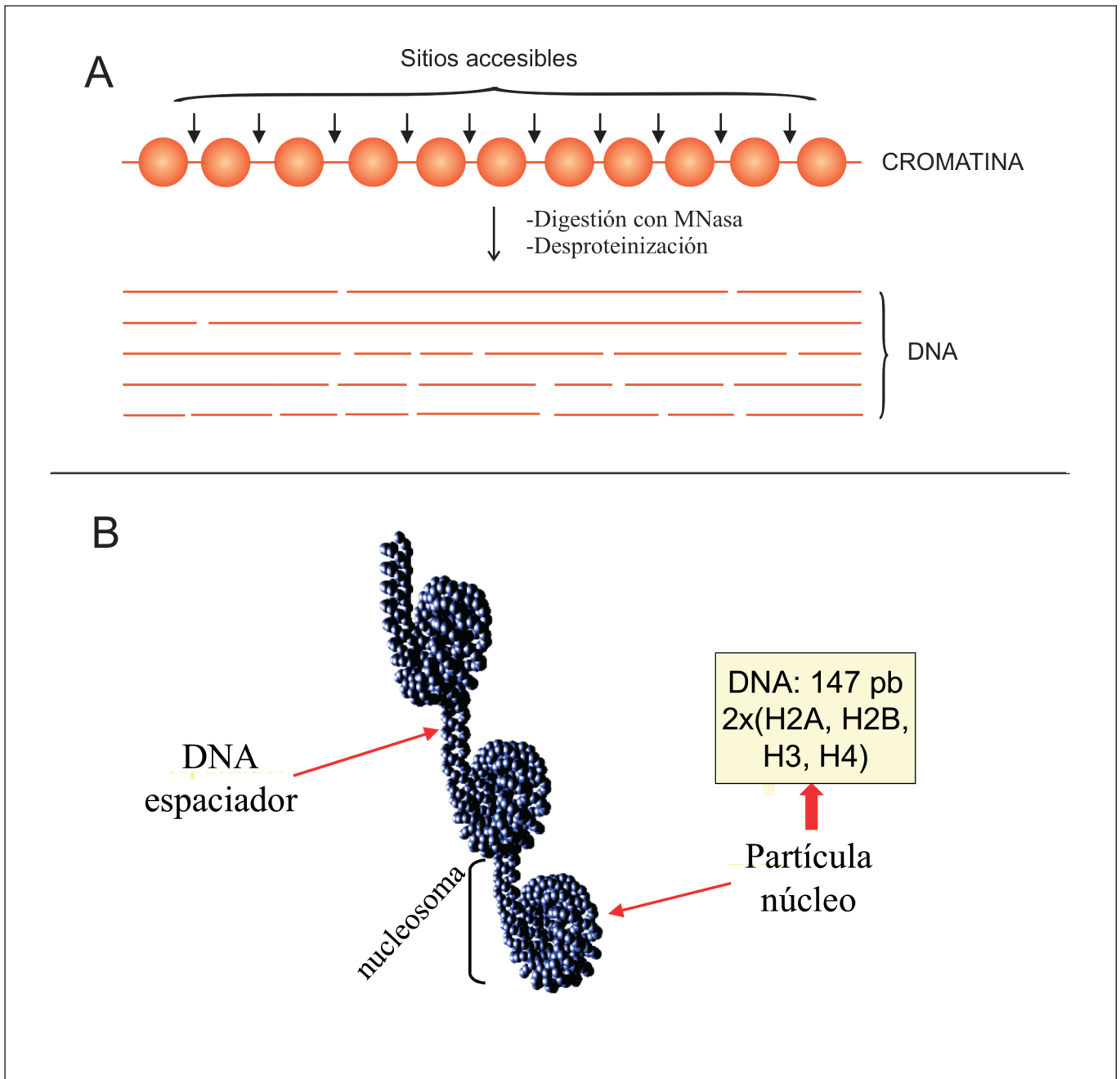


Figura 10. Representación esquemática de un filamento de nucleosomas (inicialmente denominados cuerpos v). En la parte A se muestra que los únicos sitios accesibles a la nucleasa de micrococo (MNasa) son los espaciadores internucleosomales. Como consecuencia, la digestión de cromatina con esa enzima y la posterior desproteización, da lugar a fragmentos de DNA cuyos tamaños, aproximadamente, son múltiplos del correspondiente a un mononucleosoma. En la parte B se observa con más detalle la organización nucleosomal y se describe su estequiometría.

moléculas contienen aminoácidos polares y apolares en la proporción comúnmente encontrada en las proteínas globulares. Parecía evidente que las regiones N-terminales serían incapaces de adquirir estructura organizada, mientras que el resto de la molécula podría adoptar una estructura terciaria típica de las proteínas

globulares. La hipótesis, también admitida sin discusión, se refleja en la figura 9. Y aunque, como queda dicho, el papel estructural de las histonas en el mantenimiento de la superhélice no se definiera claramente, algo parecía claro: serían las colas N-terminales las regiones prioritarias de unión al DNA, mientras que

los dominios globulares podrían originar interacciones histona-histona, que afianzarían la organización de la cromatina.

EL DESCUBRIMIENTO DEL NUCLEOSOMA

Así las cosas, en 1973 ocurre un hecho trascendental. Donald y Ada Olins volvieron a observar fibras de cromatina a baja fuerza iónica empleando procedimientos experimentales que minimizaban la alteración de la muestra. En esas condiciones detectaron la presencia de unas partículas globulares de unos 10 nm de diámetro, separadas por un material filamentosamente más fino. Era algo así como un collar de cuentas, y de ese modo gráfico lo designaron al presentar sus resultados en un congreso de Biología Celular, que tuvo lugar en Miami en 1973. Con un juego de palabras sólo comprensible en el original inglés, denominaron a las *cuentas del collar* cuerpos ν , porque “eran nuevos en el campo de las nucleohistonas”⁹. Los resultados, en su forma definitiva, se publicaron en 1974¹⁰. Ese mismo año, otros dos investigadores, Hewish y Burgoyne, fragmentaron cromatina de hígado de rata por digestión con una nucleasa endógena. Al separar por electroforesis el DNA obtenido de los fragmentos aparecían unas moléculas discretas, cuyos tamaños eran múltiplos del de la más pequeña. Los resultados cuadraban perfectamente con los de la observación microscópica. Si la cromatina estaba constituida por una serie regular de cuentas, los cuerpos ν en los que se encontrarían presentes las histonas, separadas por el *hilo del collar*, el DNA desnudo, sólo éste sería accesible a las nucleasas. Dependiendo del sitio de corte, se obtendrían fragmentos de cromatina formados por un número entero de cuerpos ν (figura 10).

A pesar de la resistencia inicial¹¹, la aplastante evidencia experimental acabó por convencer a los más escépticos y se produjo de esa manera la primera gran revolución en el estudio de la estructura de la cromatina. Los cuerpos ν fueron objeto de numerosísimas investigaciones en los años subsiguientes. Gracias a

ellas se llegó a conocer su composición y a tener una idea global de su estructura. Hubo también un cambio de denominación y se acuñó el término nucleosoma, que se ha conservado hasta nuestros días, para designar a esas partículas básicas en la organización de la cromatina. Un nucleosoma consta de una partícula núcleo y un DNA espaciador que, más o menos, coincidirían, respectivamente, con las cuentas del collar y el hilo que las separa. Mientras que el DNA espaciador es de longitud variable, la partícula núcleo tiene una absoluta constancia en todas las especies. Consta de un octámero formado por dos copias de cada una de las histonas H2A, H2B, H3 y H4 y 147 pares de bases de DNA (figura 10). El DNA se enrolla alrededor del octámero, formando casi dos vueltas de superhélice a izquierdas. Este era el motivo superhelicoidal detectado en los primitivos experimentos de difracción de rayos X, cuyo error consistió en pensar que el espaciado de 11 nm era el paso de rosca de la superhélice, cuando en realidad corresponde al diámetro de la partícula núcleo, ya que el paso de rosca medio del DNA a su alrededor es de sólo 2,8 nm.

El grupo de Aaron Klug en Cambridge, heredero de la tradición estructuralista del famoso laboratorio Cavendish que determinó los parámetros de la doble hélice, consiguió pronto cristalizar partículas núcleo. Pero no pudieron estudiar su estructura con resolución aceptable porque los cristales eran inestables y se destruían al someterlos a radiación X. Se llegó así a 1991 sin conseguir bajar la resolución de 7 Å. Mientras que este nivel de resolución permitía determinar las dimensiones de la partícula núcleo y trazar algunos detalles de la organización del DNA y de las histonas, era absolutamente insuficiente para decidir la estructura terciaria exacta de las histonas y, menos aún, para localizar en el espacio todos sus aminoácidos. Con todo, fue decisivo para que Klug recibiera el premio Nobel de Química en 1982.

Pero en 1991 se inicia una nueva revolución en el estudio de la estructura de la cromatina. Unos años antes, el grupo de Moudrianakis había conseguido aislar y cristalizar octámeros de histonas a partir de

⁹ La pronunciación en inglés del nombre de la letra griega ν coincide con la de la palabra *new* y con la de la primera sílaba de la palabra *nucleohistone*.

¹⁰ Olins, A. L. y Olins, D. E. (1974) *Science* **183**, 330-332.

¹¹ La historia del descubrimiento de los cuerpos ν está excelentemente recogida en el libro de van Holde que se cita en la bibliografía.

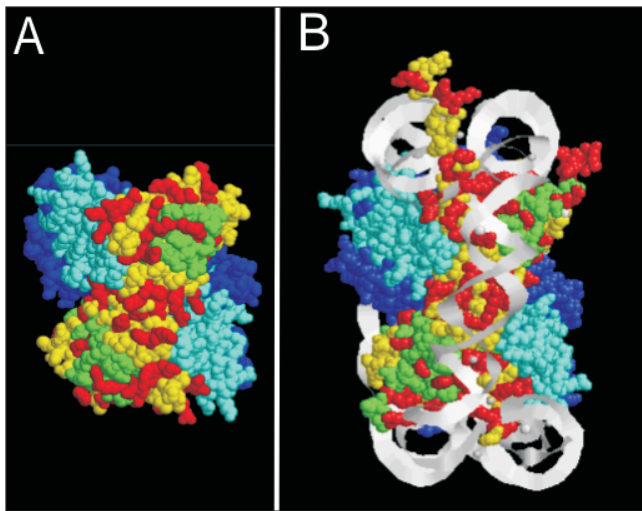


Figura 11. A: estructura de un octámero de histonas. El color de los átomos de las histonas internas sigue el mismo código que en la figura 9. En el tetrámero central (H3-H4)₂, las cadenas laterales de lisina y arginina, cuyos extremos poseen carga positiva a pH fisiológico, se han dibujado en rojo. La figura está construida a partir de los datos de Arents *et al.* (1991). B: estructura de una partícula núcleo, dibujada a partir de los datos de Luger *et al.* (1997). Se observa que la organización del octámero de histonas es similar cuando está aislado y cuando está incorporado a la partícula núcleo y que los aminoácidos básicos, especialmente en el tetrámero central, marcan la trayectoria del DNA. En las dos partes de la figura la orientación es la misma, con el eje pseudobinario de simetría perpendicular al plano del papel y en el centro de las partículas. El eje de la superhélice de DNA está en el plano del papel, de izquierda a derecha. La figura está construida con el programa RasMol.

partículas núcleo. Estos cristales sí eran suficientemente estables y se pudo determinar su estructura con una resolución de 3,1 Å. Evidentemente, al hacer del octámero el objeto del estudio, la observación directa quedaba restringida a las histonas y, además, no a la totalidad de sus moléculas. Las colas N-terminales son, como se había supuesto correctamente, demasiado flexibles como para que adopten una estructura organizada y esto implica que esas regiones no ocupen posiciones invariables en los cristales, o lo que es lo mismo, que la difracción de rayos X no aporte ninguna información sobre ellas.

Pero, a pesar de esta limitación, la información obtenida por el grupo de Moudrianakis fue extraordinariamente valiosa. En primer lugar, se observó que el octámero presenta una organización tripartita, en la que un tetrámero (H3-H4)₂ ocupa una posición central, flanqueado por dos dímeros H2A-H2B (figura 11A). Más llamativa fue la observación de que, en contra del

modelo entonces vigente para la estructura de las histonas (figura 9), los dominios C-terminales de las histonas no eran, estrictamente hablando, globulares. Se organizaban con un motivo estructural, desconocido hasta entonces y que se comenzó a denominar “*histone fold*” (figura 12). Lo esencial de este motivo es la presencia de una hélice α larga flanqueada por dos cortas, que están separadas la central por dos lazos que contienen algunos residuos en estructura β . Pero una histona aislada no es capaz de adoptar el *histone fold*. Se requiere siempre, como ocurre en el octámero, la presencia de otra histona *complementaria*: H2A se puede complementar con H2B y H3 con H4. Surge entonces la organización del tetrámero (H3-H4)₂ o de los dímeros H2A-H2B por interdigitación de los correspondientes *histone folds*, formando un motivo que Moudrianakis, de una manera sumamente gráfica, denominó *apretón de manos* (figura 12). Los aminoácidos en estructura β son esenciales para la interacción por puentes de hidrógeno entre las dos histonas de cada par.

Por otro lado, aunque el DNA no estuviera presente en la muestra objeto de la investigación —no olvidemos que se estaba llevando a cabo con octámeros aislados—, no era difícil vislumbrar su posible trayec-

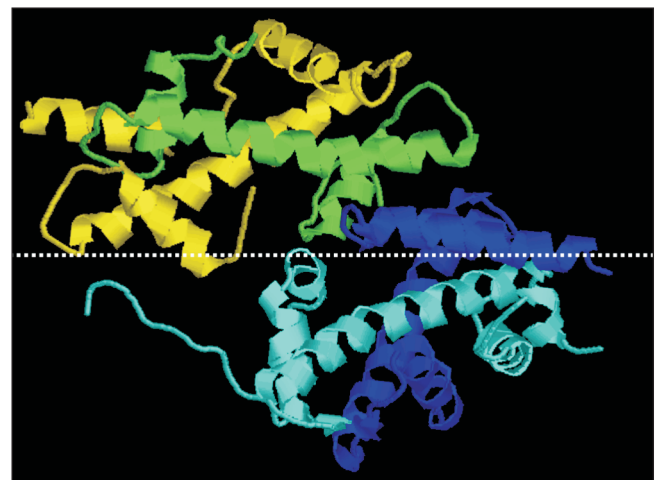


Figura 12. Estructura de las histonas internas en el octámero, según los datos de Arents *et al.* (1991). Para mayor claridad, sólo se ha representado la mitad de un octámero, en la que aparece una copia de cada una de las histonas internas, con el mismo código de colores que en la figura 9. Se observa el *histone fold* y el motivo del “apretón de manos” (véase el texto). La figura está construida con el programa RasMol a partir de los datos de Arents *et al.* (1991). La línea horizontal de puntos indica el recorrido del eje pseudobinario de simetría.

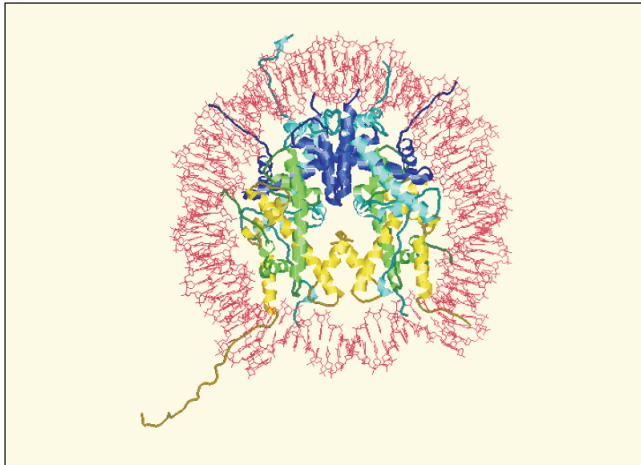


Figura 13. Estructura de la partícula núcleo. Se muestra sólo el esqueleto de las histonas internas, para poner de manifiesto que se conservan los motivos estructurales presentes en el octámero aislado (figura 12). El DNA se representa en color rosa. La orientación es tal que el eje de su superhélice está centrado y es perpendicular al plano del papel. El eje pseudobinario de simetría está contenido en el plano, de abajo hacia arriba. La figura está construida con el programa RasMol a partir de los datos de Luger *et al.* (1997).

toria alrededor de las histonas. En la figura 11A se puede observar una especie de rampa helicoidal que recorre la superficie del octámero, con unas dimensiones más o menos concordantes con las predichas por las observaciones de la partícula núcleo a baja resolución. Más aún, la disposición de aminoácidos básicos en esas rampas parece seguir la impronta de los fosfatos de las dos cadenas de una doble hélice y Moudrianakis se atrevió a predecir su recorrido.

La figura 11 posee una gran belleza intrínseca, no sólo por razones de índole estética, sino porque muestra una sorprendente y maravillosa organización. La conservatividad de las histonas está en función de su papel estructural, que no es otro que el de servir de base a la estructura del nucleosoma. Pero, para ello, se requiere que los residuos que forman el *histone fold*, sin necesidad de ser idénticos, exhiban propiedades que permitan la adquisición de ese apretón de manos al que antes se aludía. Y, por otro lado, se precisa que los residuos básicos queden dispuestos —por supuesto en la superficie del octámero en virtud del efecto hidrofóbico¹²— no en cualquier lugar, sino en los sitios con-

cretos por donde *van a pasar* los fosfatos del DNA cuando éste se ensamble sobre las histonas.

En 1997, el grupo de Richmond, un autor que había trabajado previamente en el equipo de Klug, consiguió el éxito que tanto había tardado en llegar: describir la estructura de la partícula núcleo por difracción de rayos X con una resolución a nivel atómico. Para lograr ese éxito tuvieron que concurrir varios factores determinantes. En primer lugar, no emplearon nucleosomas naturales, sino reconstituidos *in vitro*. Hay que advertir que, como se sabía desde tiempo atrás, el proceso de reconstitución mantiene fielmente la estructura del nucleosoma. Hoy conocemos la causa: la precisa organización del octámero descubierta por Moudrianakis hace que las histonas, tras mezclarlas en las condiciones adecuadas, adopten espontáneamente la estructura correcta y que al añadir el DNA éste se *limite* a recorrer la trayectoria que aquéllas le marcan. En los experimentos de Richmond, las histonas empleadas en la reconstitución eran histonas recombinantes, producidas por técnicas de Biología Molecular en una bacteria. De esta manera, se evitaba la pequeña, pero importante, heterogeneidad que presentan las histonas en los nucleosomas naturales, en virtud de sus modificaciones post-traduccionales (véase más adelante) y de la limitada variabilidad de las histonas pertenecientes a una misma clase. Por otro lado, el DNA era también artificial, de secuencia única, precisamente de una que se adapta especialmente bien a la formación de la superhélice que el nucleosoma exige. Finalmente, desde un punto de vista instrumental, el grupo de Richmond no utilizó radiación X de una fuente convencional, sino la procedente de un sincrotrón. La mayor energía de los rayos X asociados al sincrotrón permitió obtener los datos suficientes para llegar a una resolución a 2.7 Å.

El examen de las imágenes conseguidas permite obtener conclusiones importantes: la primera, es que la organización de las histonas descrita para el octámero aislado, sigue siendo válida en la partícula núcleo (figura 13). En segundo lugar, el recorrido del DNA queda perfectamente esclarecido. Como se sospechaba desde los resultados obtenidos a resoluciones más

¹² Debido a que la mayor parte de las proteínas celulares se encuentran en un entorno acuoso, los aminoácidos con cadena lateral apolar tienden a situarse en la región interna de las proteínas globulares, mientras que los polares quedan en la superficie interactuando con el agua. Este principio de organización, gobernado por razones de tipo entrópico, recibe el nombre de “efecto hidrofóbico”.

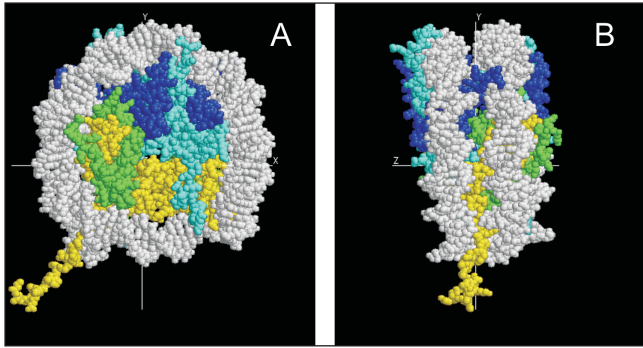


Figura 14. Estructura de la partícula núcleo en la que se muestran todas las posiciones atómicas determinadas por Luger *et al.* (1997). La parte A de la figura es equivalente a la figura 13 en orientación y permite la observación de la cara plana del nucleosoma. En la parte B los ejes de la superhélice (z) y pseudobinario (y) están en el plano del papel. Las figuras permiten observar cómo las colas N-terminales de las histonas salen de la partícula núcleo por los huecos que dejan los surcos del DNA (representado en blanco).

bajas, la superhélice no es regular. Finalmente, los nuevos resultados aportaron datos valiosos para aclarar algunas cuestiones referentes a las histonas, que por el análisis del octámero habían quedado sin respuesta. La presencia del DNA permite que queden más residuos de las histonas en posiciones fijas y, por tanto, que se pueda dilucidar la localización de más del 80% de los átomos de las histonas. La cuestión es singularmente importante en el caso de las colas N-terminales. Aunque sólo se puedan observar parcialmente, se puede ver, en algunos casos, que salen del octámero a través de los surcos del DNA, pero no están interactuando con éste, sino que se proyectan hacia fuera (figura 14). Por otro lado, las colas de H4 interactúan electrostáticamente con la superficie de la partícula núcleo vecina en el cristal. Al posible significado de esta observación se hará referencia más tarde.

Tanto los resultados del grupo de Moudrianakis como los de Richmond ponen de manifiesto que, para la organización del nucleosoma, no son esenciales las colas N-terminales de las histonas, ya que sus lugares de interacción con el DNA se localizan en los dominios estructurados de los dos tercios C-terminales. Aunque mediante otras técnicas se ha avanzado

en el estudio de la disposición de estas colas, sigue habiendo numerosos interrogantes sobre su disposición estructural en la cromatina. También siguen existiendo interrogantes sobre la localización exacta de la histona H1, que interacciona con 20 pares de bases del DNA espaciador, aunque se han planteado algunas posibilidades en los últimos años.

DEL FILAMENTO DE NUCLEOSOMAS AL CROMOSOMA

Con las incertidumbres que se acaban de mencionar, se puede decir que el nivel primario¹³ de organización de la cromatina, el filamento de nucleosomas, se conoce con bastante precisión. Pero es evidente que la cromatina ha de adoptar una conformación mucho más compacta si tiene que empaquetarse en un núcleo. La investigación de los niveles superiores de organización de la cromatina se inició casi en paralelo al estudio de la estructura nucleosomal, pero es forzoso reconocer que los avances ha sido mucho menos satisfactorios. Los primeros datos se obtuvieron también por observaciones al microscopio electrónico. Si a baja fuerza iónica la cromatina adopta el aspecto de un collar de cuentas, cuando la concentración salina supera un cierto nivel (del orden de 80 mM) se pueden apreciar fibras de unos 30 nm de diámetro. El primer modelo para explicar la estructura de esa fibra data de 1977 y es el denominado *solenoides*, según el cual los nucleosomas se dispondrían de forma que los ejes de su superhélice adoptarían, a su vez, una trayectoria helicoidal. Los nucleosomas, entre 6 y 8 por vuelta de solenoide, quedarían con sus caras planas más o menos paralelas al eje de la fibra. Pero como ha advertido recientemente Daban, la compactación que se lograría de esa manera es insuficiente para justificar la que debe tener lugar en el interior del núcleo. En efecto, la concentración de DNA en un núcleo eucariótico oscila entre 0,12 y 0,20 g/ml y la que se puede lograr con los modelos clásicos para la fibra de 30 nm varía entre 0,04 y 0,19 g/ml, insuficiente para explicar la compactación real del DNA. Por ese motivo, este autor ha propuesto un nuevo modelo, el solenoide interdigitado, que permite una organización más compacta. En este

¹³ Por analogía con los niveles estructurales de las proteínas, se ha propuesto recientemente designar a la organización de los nucleosomas a lo largo del DNA como nivel estructural primario de la cromatina. El nivel secundario estaría constituido por la fibra de 30 nm y el nivel terciario por las fibras más gruesas que se observan en los núcleos.

modelo, el contacto entre nucleosomas también se produce a través de sus caras planas. Es a este nivel donde la observación de que las colas de H4 pueden interaccionar con las superficies de los nucleosomas cobra especial interés. Si esa interacción se da realmente en la fibra de 30 nm, la adquisición de esta estructura —y, por tanto, la compactación de la cromatina— exigiría que los extremos N-terminales de H4 mantuvieran inalterada su carga positiva, situación que, como se verá más adelante, puede perderse en determinadas circunstancias. Finalmente, hay que tener en cuenta que la fibra de 30 nm aún debe compactarse más, pero lo que se conoce de esos niveles superiores apenas sobrepasa el nivel de las conjeturas.

FUNCIONALIDAD DE LOS GENES

Hasta aquí, la descripción de la estructura de la cromatina. El panorama puede parecer un tanto estático; ¿acaso hemos olvidado ese principio, varias veces repetido, de que la estructura ha de contemplarse en relación con la función? No; lo que ocurre es que, históricamente se procedió más o menos del modo aquí relatado y hubo momentos en que, en cierto modo, se oscureció ese principio clarificador. Un principio que, como se ha visto, facilitó a Watson y a Crick el establecimiento de la estructura de la doble hélice y que, cuando se soslaya, aunque sólo sea metodológicamente, obliga a que la investigación dé innecesarios rodeos. Y la cromatina no es simplemente *una estructura*, sino una estructura que debe permitir la actividad nuclear: replicación, transcripción, reparación del DNA, recombinación, etc.

Por limitarse a uno de esos procesos, que es por otro lado el que más está clarificando las relaciones estructura-función en la cromatina, vale la pena centrarse en la transcripción. En eucariotas existen tres tipos de RNA polimerasas, las enzimas que catalizan la síntesis de RNA ensamblando ribonucleótidos sobre la cadena molde. De ellas, la RNA polimerasa II se encarga de transcribir los genes estructurales, es decir, aquellos que codifican proteínas y, por tanto, esta polimerasa se encarga de formar el mRNA, mientras que las polimerasas I y III se ocupan de la síntesis de rRNA y tRNA. La transcripción catalizada por la RNA polimerasa II es, sin duda, la que más atención ha recibido y a ella se circunscriben los comentarios finales de este artículo.

En la figura 8 se indica que antes de la región transcrita de un gen se encuentra el promotor. En él se ensambla la RNA polimerasa II junto con otras numerosas proteínas que forman el denominado complejo de preiniciación, como paso previo a la transcripción. Pero el promotor contiene también numerosas secuencias que resultan esenciales para la regulación de la transcripción. Y es que la transcripción no ocurre de una forma indiscriminada, sino que para la expresión de cada gen existen *un aquí y un ahora* precisos. Con estas palabras se quiere indicar que existe un doble nivel de regulación en la transcripción eucariótica, especialmente en los organismos pluricelulares. Por un lado, cada línea celular tiene una característica batería de genes que se expresan y otros muchos genes que permanecen silenciados durante toda la vida del organismo. Como se ha comentado anteriormente, esta expresión génica diferencial en cada línea celular, el aquí que se acaba de mencionar, hace posible la diferenciación celular. Pero, además, dentro de la serie de genes que se pueden expresar en una célula dada, cada uno tiene un momento, un ahora, en el que su producto es necesario.

Comentar, aunque sólo fuera en líneas generales, los intrincados mecanismos de regulación transcripcional excedería en mucho los límites tolerables para esta exposición. Para seguir su hilo argumental basta con mencionar que los precisos circuitos de regulación exigen la presencia de numerosas proteínas —factores transcripcionales se suelen denominar genéricamente— que, o bien reconocen esas secuencias específicas del promotor, o bien interaccionan con otros factores o con la propia maquinaria basal de transcripción: la RNA polimerasa y otras proteínas asociadas. Dentro de todos esos factores existen activadores, represores, coactivadores, correpresores, adaptadores..., cuyo mismo nombre da idea bastante precisa de su función.

Si se considera la enorme diversidad de factores requeridos para la transcripción y la amplia distribución de sus secuencias diana en el promotor y aún en otras regiones distales del DNA (figura 15) parece evidente que la organización de las histonas en nucleosomas y en estructuras de orden superior supone una seria dificultad para el ensamblamiento de todos esos complejos. Efectivamente, hasta hace unos 10 años, de esa manera se solían considerar las histonas: simple-

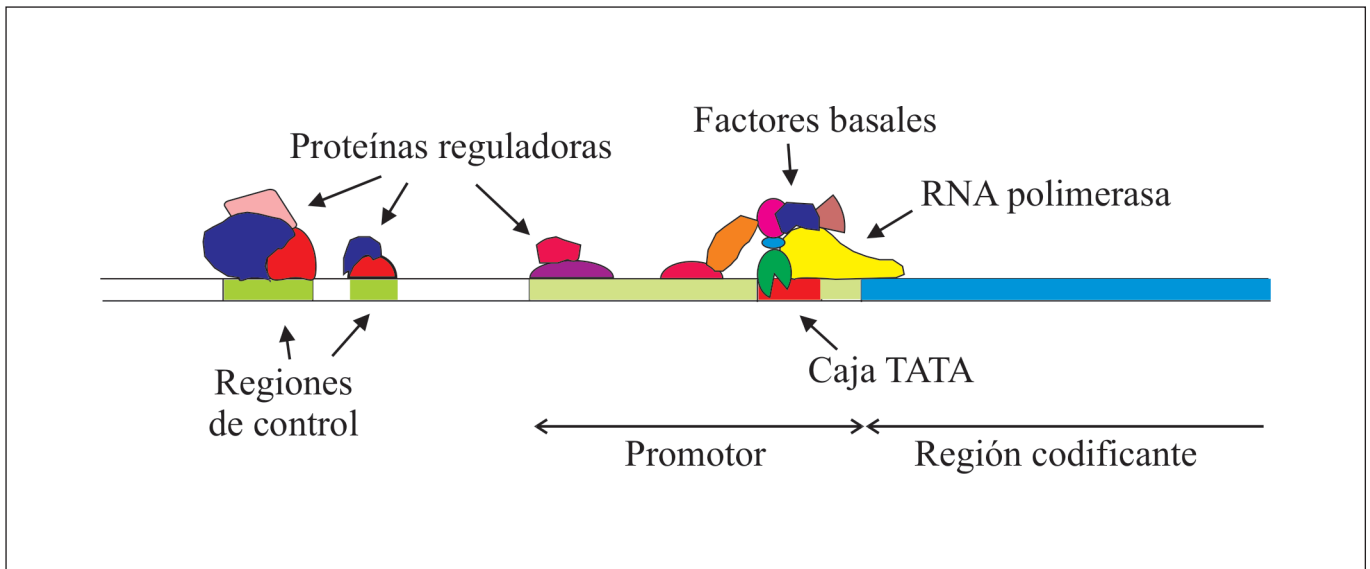


Figura 15. Representación esquemática de la interrelación de diversos factores en la transcripción regulada en eucariotas.

mente como un obstáculo para las funciones nucleares, como la transcripción. Aproximadamente desde 1992 comenzó a imponerse otra forma de enfocar la cuestión. No es que la organización de la cromatina ya no se considere un obstáculo, sino que el papel de las histonas se contempla de una forma mucho más dinámica. Que la cromatina obstaculiza la transcripción resulta evidente, por ejemplo, al observar la figura 16. En ella se representa la interacción de la RNA polimerasa II de levadura, cuya estructura tridimensional se ha resuelto en 2001, con el DNA en el complejo de preiniciación. Se puede ver cómo el DNA se introduce por una especie de túnel formado por varias de las cadenas polipeptídicas de la enzima. Pero el túnel tiene las dimensiones justas para que quepa la doble hélice; si en ese lugar existiera un nucleosoma, habida cuenta de sus dimensiones, la unión de la RNA polimerasa sería imposible. Entre otras razones, ésta es la causa de un hecho constatado ya en 1987: en un nucleosoma no se puede iniciar la transcripción.

LAS HISTONAS: MUCHO MÁS QUE PROTEÍNAS ESTRUCTURALES

Éste es el momento de presentar el papel dinámico de las histonas. Como se ha visto al considerar la estructura del nucleosoma, las colas N-terminales de las histonas se proyectan hacia fuera. Esas colas son el lugar de numerosas reacciones de modificación cova-

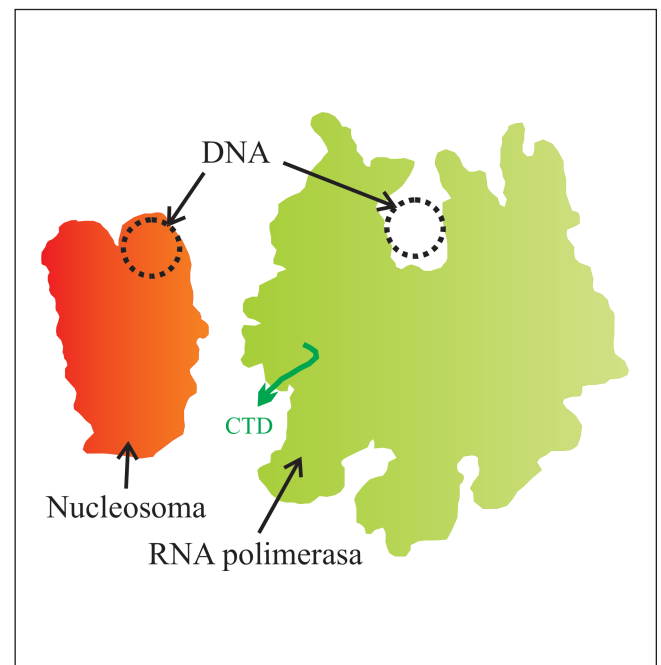


Figura 16. Tamaños relativos de un nucleosoma y de la RNA polimerasa. Como ejemplo de ésta última, se ha tomado la RNA polimerasa II de levadura. Se representa sólo la estructura global de la molécula, sin detallar la posición de sus subunidades, y el inicio del dominio C-terminal (CTD). El contorno del nucleosoma se representa en una orientación como la de la figura 14 B. Los círculos de trazos señalados como DNA indican las dimensiones de una sección transversal de la doble hélice. En el nucleosoma, está dibujado por la trayectoria de una de las vueltas de superhélice; en la polimerasa indica la posición en que quedaría el DNA en el complejo de preiniciación. Se observa claramente que es imposible formar este complejo cuando hay un nucleosoma sobre la caja TATA (véase la figura 15).

lente: acetilación, fosforilación y metilación. La acetilación puede afectar a 26 residuos de lisina por nucleosoma y tiene unas consecuencias químicas importantes: se pierde la carga positiva que posee la cadena lateral del aminoácido. La fosforilación y la metilación son más limitadas en cuanto al número de aminoácidos potencialmente modificables.

De esas modificaciones, la acetilación es la mejor estudiada. Especialmente, los últimos 7 años han sido testigos de increíbles avances, como consecuencia de la clonación de los genes que codifican los dos tipos de enzimas implicadas en el proceso: las histona acetiltransferasas (HAT) y las histona desacetilasas (HDAC). Las primeras catalizan la entrada de grupos acetilo en las histonas a partir de acetil-coenzimaA, el donador universal de acetato. Las segundas catalizan la eliminación hidrolítica del resto acetilo, lo que permite que la acetilación de histonas sea un proceso reversible. Las consecuencias de esta modificación son de dos tipos. Por un lado, se admite actualmente que la reducción de carga positiva inherente a la acetilación, altera las interacciones de las colas de algunas histonas de un nucleosoma no con el DNA, como se pensó al principio, sino con el nucleosoma vecino. Así, la acetilación de esas histonas impide la adquisición de estructuras de orden superior. Una función de la acetilación de histonas es, pues, controlar la compactación de la cromatina. En realidad, se venía sospechando ese papel desde el descubrimiento de la acetilación de histonas en 1964, pero ha sido ahora cuando se han comprendido las causas moleculares de ese proceso.

Pero hay un segundo modo de acción de la acetilación de histonas. La introducción de grupos acetilo en determinados residuos de lisina constituye una señal que permite o impide la interacción de determinados factores con las histonas a través de sus colas N-terminales, accesibles en los nucleosomas. Esta posibilidad había sido propuesta en 1988 por Peter Loidl, el primero que habló del posible papel señalizador de la acetilación de histonas. Recientemente, la hipótesis de Loidl ha sido ampliada por Strahl y Allis para incluir las otras modificaciones comentadas: fosforilación y metilación. Según esa idea, cada combinación de acetilaciones, metilaciones y fosforilaciones en los diversos residuos de una histona, constituye un código, que significa una precisa función de la cromatina en esa región. Evidentemente, dado el elevado

número de residuos modificables, el número de combinaciones es virtualmente ilimitado y es evidente que no todas ellas son significativas funcionalmente. Pero se van obteniendo datos que indican claramente el mensaje que encierran algunos de estos códigos. Por ejemplo, la fosforilación del residuo 10 de serina de la histona H3, seguida de la acetilación de las lisinas 9 y 14 constituye una señal para la transcripción. Por el contrario, cuando la fosforilación de la serina 10 va seguida de una nueva fosforilación en la serina 28, se interpreta esa modificación como una señal para la condensación mitótica, en la que la cromatina permanece totalmente inactiva.

Pero aún se plantean dos problemas. Por un lado hay que saber cómo se lee el código de las modificaciones de las histonas. Por otra parte, es preciso determinar de qué manera se dirige la precisa combinación de modificaciones al lugar adecuado. En el caso de que el código signifique "transcripción", estos dos problemas se concretan en las preguntas siguientes:

Teniendo en cuenta que una determinada señalización sobre una histona depende de las actividades relativas de las enzimas capaces de introducir o eliminar la correspondiente modificación, ¿cómo se reclutan las enzimas adecuadas sobre el gen diana en el momento correcto?

¿Qué factores proteicos reconocen el código? ¿Cómo actúan para que los nucleosomas que existen sobre el gen dejen de impedir el ensamblamiento del complejo de preiniciación y el posterior avance de la RNA polimerasa?

Honradamente, hay que advertir que, aunque se han dado pasos de gigante para resolver estos problemas en los últimos 10 años, son aún muy numerosos los interrogantes planteados. Por sólo citar un ejemplo, se puede mencionar el caso de la activación del gen *MAT2A* en hepatocitos de rata. Este gen codifica la subunidad catalítica de una de las isoenzimas de la metionina adenosiltransferasa. La enzima es clave en el metabolismo de la metionina y, teniendo en cuenta que el producto de la reacción que cataliza, la S-adenosilmetionina es el donador universal de grupos metilo, la reacción resulta clave en el contexto de las metilaciones biológicas.

Además del gen mencionado, también codifica una subunidad catalítica de la metionina adenosiltransferasa el gen *MAT1A*. Ambos productos génicos son homólogos, pero dan lugar a isoenzimas con distintas propiedades cinéticas. *MAT1A* se expresa en hígado, y las propiedades de las isoenzimas resultantes se adaptan perfectamente a las necesidades metabólicas de este órgano. Por el contrario, *MAT2A* se expresa en tejidos extrahepáticos, pero también en el hígado cuando éste se encuentra en rápido crecimiento, como ocurre en el hígado fetal, en el hígado en regeneración después de una hepatectomía parcial, o en un hepatocarcinoma.

Así pues, en un cultivo primario de hepatocitos¹⁴ el gen *MAT1A* se expresa pero *MAT2A*, no. De acuerdo con las ideas expuestas antes, las histonas de los nucleosomas del promotor de este último gen se encuentran hipacetiladas. La situación es paralela a la que ocurre en el hígado intacto. Pues bien, es posible inducir la división celular en el cultivo primario añadiendo al medio el factor de crecimiento hepático (HGF, del inglés hepatic growth factor). HGF es una proteína que no penetra en la célula, pero que interacciona con un receptor de su membrana e inicia una ruta de señalización intracelular que da lugar, entre otras cosas, al inicio de la división celular. Los mecanismos generales por los que opera esa transducción de señales se han descrito ya en otro volumen de este Programa de Promoción de la Cultura Científica. Baste ahora recordar que implican cascadas de fosforilación en las que una proteína quinasa —una enzima que puede catalizar la fosforilación de una proteína a expensas del ATP— pasa a su estado activo al ser fosforilada por otra quinasa y, de ese modo se hace capaz de fosforilar a la siguiente enzima de la cascada.

En el caso presente, el receptor de HGF adquiere actividad de proteína quinasa al interactuar con su ligando. La transducción de esa señal a través de una cascada de fosforilación, llega hasta el núcleo y se activa la expresión de *MAT2A*. De hecho, a los 90 min de adición de HGF ya se observa la presencia del mRNA de *MAT2A*. Pero antes, ya a los 30 min, se puede detectar un incremento en la acetilación de las histonas en los nucleosomas del promotor del gen. Ni

la acetilación de histonas ni la transcripción se dan cuando se inhibe la actividad de proteína quinasa del receptor de HGF o cuando se inhiben algunas otras quinasas de la cascada. Por otro lado, en presencia de metil-tioadenosina —un producto secundario del metabolismo de S-adenosilmetionina, que actúa como inhibidor de la metilación de proteínas— no se produce la transcripción de *MAT2A* estimulada por HGF, pero sigue teniendo lugar la acetilación de histonas. En su conjunto, estos resultados recientes muestran que el reclutamiento de histona acetiltransferasas sobre el promotor de *MAT2A*, requisito previo a la expresión del gen, está regulado por una ruta de transducción de señales, concretamente por una cascada de fosforilación. Además, ponen de manifiesto que la acetilación de histonas es necesaria, pero no suficiente para la transcripción. Dicho de otra manera, los resultados que se acaban de mencionar —los primeros que muestran la relación de la acetilación de histonas con una cascada de fosforilación— responden parcialmente a la primera de las preguntas planteadas más arriba.

En cuanto a la segunda cuestión, hace 10 años que se detectó la existencia de los denominados complejos de remodelación de la cromatina. Estos complejos, presentes en todos los eucariotas, están formados por la unión de un número variable de proteínas y son capaces de alterar la estructura de la cromatina utilizando la energía procedente de la hidrólisis del ATP. Esa alteración puede ser simplemente un deslizamiento de los nucleosomas, que deja así libre el DNA sobre el que ha de ensamblarse el complejo de preiniciación. En otros casos, la remodelación puede implicar un cambio de la estructura nucleosomal que permita la interacción del DNA con factores proteicos. En cualquier caso, la estabilidad del nucleosoma, tanto en lo que se refiere a las interacciones histona-DNA como histona-histona, implica que el proceso de remodelación consuma energía.

Finalmente, queda latente la pregunta que se ha apuntado a lo largo de las líneas precedentes: ¿cómo se lee el código que las modificaciones covalentes introducen en las histonas? Aunque una respuesta completa requerirá aún más esfuerzo por parte de los investiga-

¹⁴ Es posible aislar hepatocitos a partir de hígado de rata. Estas células, pueden mantenerse vivas en cultivo durante más de 24 h, pero no tienen capacidad de reproducirse. Un cultivo con estas características recibe el nombre de “cultivo primario”.

dores, se sabe que hay diferentes dominios en los factores proteicos capaces de reconocer los distintos motivos, acetil-lisina, grupos metilo, etc., que se introducen en las histonas. La presencia de esos dominios en enzimas de modificación, componentes de los complejos de remodelación, etc. hace posible que las diferentes modificaciones de las histonas y, en su caso, la remodelación de la cromatina, procedan de un modo secuencial.

CONSIDERACIÓN FINAL

Desde el descubrimiento del gen hasta nuestros días han transcurrido casi 140 años. Los hallazgos sobre su naturaleza, organización y funcionamiento han aumentado de forma exponencial hasta constituir, en los últimos años, uno de los temas favoritos de la investigación en Biología Molecular. Algunos de esos avances se han resumido en las líneas precedentes que, inevitablemente, constituyen una exposición excesivamente resumida de los logros obtenidos. Pero, como ocurre siempre en la investigación científica, si los logros han crecido de modo exponencial, también lo han hecho los nuevos interrogantes que han planteado. Lejos de resultar frustrante, esta situación constituye un acicate para los investigadores: la contemplación de un amplio panorama abierto frente a nosotros siempre es un ilusionante motivo para seguir adelante.

BIBLIOGRAFÍA

1. ARENTS, G., BURLINGAME, R. W., WANG, B. -C., LOVE, W. E., Y MOUDRIANAKIS, E. N. (1991) «The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix» *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10148-10152.
2. DABAN, J. R. (2000) «Physical constraints in the condensation of eukaryotic chromosomes. Local concentration of DNA versus linear packing ratio in higher order chromatin structures» *Biochemistry* **39**, 3861-3866.
3. DABAN, J. R. Y BERMÚDEZ, A. (1998) «Interdigitated solenoid model for compact chromatin fibers» *Biochemistry* **37**, 4299-4304.
4. LUGER, K., MÄDER, A. W., RICHMOND, R. K., SARGENT, D. F. Y RICHMOND, T. J. (1997) «Crystal structure of the nucleosome core particle at 2.8 Å resolution» *Nature* **389**, 251-260.
5. MUNICIO A. M (2000) «Medicamentos viejos para enfermedades nuevas», en Horizontes Culturales (Real Academia de Ciencias), pp. 53-79. Espasa, Madrid.
6. STRAHL, B. D. Y ALLIS, C. D. (2000) «The language of covalent histone modification» *Nature* **403**, 41-45.
7. VAN HOLDE, K. E. (1989) *Chromatin*. Springer-Verlag, New York.
8. WATSON, J. D. Y CRICK, F. H. C. (1953) *Nature* **171**, 737-738.
9. WATSON, J. D. Y CRICK, F. H. C. (1953) «Genetic implications of the structure of deoxyribonucleic acid» *Nature* **171**, 964-967.
10. WOODCOCK, C. L. Y DIMITROV, S. (2001) «Higher order structure of chromatin and chromosomes» *Curr. Opin. Genet. Develop.* **11**, 130-135.
11. ZHANG, Y Y REINBERG, D. (2001) «Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails» *Genes & Develop.* **15**, 2343-2360.