

# **UN SIGLO DE MACROMOLÉCULAS BIOLÓGICAS**

## **De la estructura y función de los conceptos**

### **al concepto de estructura y función**<sup>\*</sup>

**Ricardo Ehrlich**

Sección Bioquímica

Departamento de Biología Celular y Molecular

Instituto de Biología, Facultad de Ciencias

### **Introducción**

Mario Calcagno, personaje central del libro de Eduardo Mizraji: “El segundo secreto de la vida”, fue uno de los primeros uruguayos purificadores de moléculas. Él purificaba una enzima en una vieja torre, donde funcionaba el Laboratorio de Bioquímica de la Facultad de Medicina. A mediados de los años 60, Mario Calcagno le da un frasco, que tenía en su laboratorio, a un estudiante, etiquetado NADP (el NADP es un compuesto que interviene en una serie de reacciones biológicas) y le pide que prepare una solución, para lo cual debe pesar este producto. El estudiante se acerca al frasco con la etiqueta NADP y queda como petrificado, porque no puede creer que el frasco tenga NADP realmente, acostumbrado a que el NADP era un concepto.

En la historia de las macromoléculas biológicas, se nos mezclan los conceptos y las realidades, y nos hemos encontrado en esta historia con todo tipo de situaciones. Se imaginan conceptos, moléculas, reacciones químicas, que tienen que existir y se pueden deducir de una serie de estudios con un alto componente fenomenológico, o con un alto componente intuitivo, como son las experiencias y reflexiones de la genética formal, de una disciplina esencialmente intuitiva como la matemática. Esto permite deducir conceptos muy fuertes: la existencia de algunas reacciones químicas particulares de las que se deducen funciones alteradas, o la existencia de alguna molécula con propiedades muy especiales, y así se puede llegar a definir cuáles son sus propiedades, cuál puede ser su estructura. Por experiencias de análisis genético se puede saber si una molécula es un monómero, es un dímero o es un tetrámero.

Muchas veces los conceptos preceden a las moléculas, se ubica a la molécula porque se dice que debe existir una molécula de ese tipo. En los últimos años nos pasa exactamente lo contrario, se encuentran una gran cantidad de moléculas y no se sabe para qué sirven, o donde están. Los investigadores entonces hacen el recorrido inverso, desde las moléculas hacia los conceptos, que es mucho más complejo. Esto está muy bien ilustrado en una frase de Francois Jacob, en su libro: “El juego de lo posible”, que dice así: “Los comienzos de la ciencia moderna se remontan al momento en que las preguntas generales fueron sustituidas por preguntas limitadas, es decir, cuando en vez de preguntar: ¿Cómo se creó el universo?, ¿en qué consiste la materia?, ¿cuál es la esencia de la vida?, aparecen preguntas del tipo: ¿Cómo cae una piedra?, ¿cómo avanza el agua por un tubo?, ¿cuál es el recorrido de la sangre por el interior de un cuerpo? Este cambio produjo

---

\* Versión aceptada por el autor.

un resultado sorprendente. Mientras que para las preguntas generales sólo se encontraban respuestas limitadas, para las preguntas limitadas empezaron a surgir respuestas cada vez más generales”. Esto ha sido válido durante muchísimo tiempo, y en este momento se nos empieza a invertir la situación.

En esta presentación, voy a resumir primero una visión de conjunto de algunos grandes momentos en la historia de los conceptos y de las macromoléculas biológicas, para después detenerme en algunos aspectos particulares.

### **Acumulación de información y rupturas conceptuales**

Cuando uno revisa la trayectoria de un área en particular y trata de hacer una historia de ideas y conceptos, hay dos grandes orientaciones posibles. Una es la galería de retratos, donde uno empieza a hablar de los aportes que fueron haciendo tales y cuales científicos, y es una manera grata de seguir la historia. Muchas veces nos acompañan anécdotas que cuentan circunstancias en que se produjeron determinados descubrimientos. La otra es analizar todo ese período a través de lo que fue la historia de las ideas, los movimientos sociales desde el punto de vista científico, que fueron haciendo avanzar la ciencia. De las dos maneras, galería de retratos o movimientos sociales, rápidamente podemos ver que hay una acumulación de información en todas las áreas, a un cierto ritmo.

Esta acumulación de información se procesa de manera distinta según las épocas. Para dar una idea, en las áreas básicas de Biología y Biomedicina se están publicando cerca de 400 a 500 mil artículos por año, y hay que suponer que se envían a las revistas científicas por lo menos tres veces más, considerando un promedio de aceptación de uno cada tres, a veces de uno cada cinco o cada diez artículos. Eso da una idea de la velocidad a la cual se acumula la información. Se ha dicho y se repite muchas veces que la información científica se duplica cada tanto tiempo. Lo que es cierto es que si ustedes van a los bancos de datos, donde figuran las publicaciones de las revistas en el área Biomédica y de Biología básica, lo que se encuentra es que hace unos 25 años se publicaban unos 200 mil artículos por año. La duplicación, las progresiones geométricas de que se hablaba no son tales, sin embargo el aumento de información es significativo. Se va acumulando información pero de vez en cuando se producen cambios cualitativos importantes, se producen rupturas.

Estas rupturas son los grandes saltos conceptuales, cambios contextuales, que transforman la trayectoria de una disciplina y abren en general nuevas perspectivas y nuevas disciplinas. Cambios de contexto y cambios de concepto que conducen entonces a nuevas síntesis. Muy frecuentemente, sobre todo en las áreas de las ciencias de la vida, se producen otros tipos de cambios, que no producen nuevas síntesis ni nuevos conceptos rápidamente, pero sí nuevas perspectivas. Me refiero a cambios metodológicos, en el desarrollo de nuevas herramientas y de nuevas disciplinas, que permiten pasar a nuevas dimensiones en cuanto al acceso de la información y al desciframiento de la misma. Existen numerosos ejemplos de estos cambios; aquí se van a dar algunos ejemplos a lo largo de toda la historia de la Biología, en particular en lo que tiene que ver con las macromoléculas biológicas.

De manera muy esquemática, podemos decir que hasta el siglo XIX en el área de ciencias de

la vida estamos en la época de los grandes clasificadores y ordenadores de la información. En el siglo XVIII aparecen los grandes sistemas de clasificación, con Linneo, Cuvier, Dufont. Con el siglo XIX empieza a haber un desarrollo tecnológico importante, comienzan a ser más fáciles los viajes, empiezan a desarrollarse los grandes imperios coloniales, aparecen los grandes viajeros y la preocupación del censo de todo lo conocido, del censo de la naturaleza. Se desarrollan las grandes disciplinas analíticas: la Química y la Fisiología, y lo que se llamó Química-Fisiológica, que precede al nacimiento de la Bioquímica.

Todo esto va a concluir en tres fuertes teorías con las cuales comienza el siglo XX, que son: la Teoría celular, que identifica a la estructura de la célula como el lugar donde ocurren todas las reacciones biológicas, como la unidad de los seres vivos; la conocida Teoría de la evolución y la Teoría de la unidad de los procesos bioquímicos, que varias décadas más tarde lleva a decir a Jacques Monod que lo que es válido para la bacteria (*Escherichia coli*) es válido para los elefantes. Las reacciones químicas que ocurren en los procesos biológicos en los distintos organismos, son esencialmente similares, así como la lógica y el tipo de moléculas implicadas. La teoría celular es de fines del siglo XIX: la vida ocurre en sistemas cerrados, delimitados con un funcionamiento propio, independiente. Un científico alemán que hizo grandes contribuciones a la patología humana dijo que toda célula debe provenir de otra célula. Por primera vez establece lo que es un organismo vivo independiente, que tiene que provenir de otro organismo vivo. Esto representó un gran avance conceptual. Sobre estos tres pilares se van a constituir las modernas ciencias de la vida.

Luego se van a desarrollar avances conceptuales importantes. En la segunda década del siglo XIX se sintetiza la urea, la primer molécula biológica en ser sintetizada. La síntesis química de la urea, que hoy nos parece algo completamente banal, provocó un cambio conceptual mayor. Se lo relaciona al nacimiento de la Química Orgánica; la urea, que es una molécula sencilla (la diamida), era producida por los riñones –se la había identificado como producida por los riñones del perro–. En los seres vivos los fenómenos biológicos se explican por fenómenos químicos. El hecho de que la urea fuera sintetizada en un laboratorio provocó una ruptura en el pensamiento científico moderno, y mostró que es posible estudiar los fenómenos químicos que ocurren en los seres vivos.

Aproximadamente al mismo tiempo se describe la fotosíntesis y se empieza a ver que es posible que una planta asimile sustancias inorgánicas del suelo y las transforme. Otro avance notable es el de la asimetría molecular, conocido por las vacunas, por el desciframiento de la presencia de microorganismos, es decir los procesos y la demolición de lo que era la Teoría de la generación espontánea. Luis Pasteur empieza su formación como químico, trabajando en un laboratorio que estaba equipado con instrumentos de óptica. Analizaba los elementos químicos a través de la propiedad de los compuestos para desviar la luz polarizada. En ese momento él estudiaba las sales que se acumulaban en los toneles de vino, las sales de ácido tartárico, y encuentra que un compuesto químico idéntico, al analizar sus propiedades ópticas tenía dos componentes. Sorprendentemente llega a la conclusión, en su trabajo de tesis, de que es el mismo compuesto químico que existe en dos formas que son isómeras. Lo que va a llevar a que más adelante Kelvin defina a la asimetría molecular como quiralidad, en referencia a que los isómeros ópticos son imágenes especulares. Este es el nacimiento del estudio de la estructura de las moléculas biológicas, la primera vez que se tiene información sobre la estructura de estas

moléculas y va a ser la semillita de lo que décadas más tarde será el estudio cristalográfico de las moléculas biológicas.

En el siglo XX se puede distinguir un primer período de clasificación de los conocimientos, marcado por un fuerte racionalismo. Un punto de ruptura en este período lo constituye el desarrollo de la  $\alpha$  hélice por Linus Pauling en el año 1948. Este período se cierra en el momento en que se desarrollan poderosísimas herramientas, una gracias a la contribución de Pauling y otra por la incipiente Biología Molecular, donde se puede incluir la contribución de Jacob y Monod sobre la primera descripción del funcionamiento del control de la expresión génica.

Hay un segundo período que podemos llamar romántico. En este período, los investigadores en el tema se apasionan por la belleza del objeto de estudio y la escritura de los artículos científicos traduce esa actitud del investigador. La finalización de este segundo período se establece en el momento en que las bases de la Biología Molecular están claramente constituidas, se termina de descifrar lo que se llamó el código genético, que en realidad es una parte de todo el código genético (la relación entre el mensaje escrito en los ácidos nucleicos y las proteínas). Esta ruptura se puede fijar en una fecha precisa: el año 1966. Un investigador que acompañó todo el proceso de desarrollo de la Biología Molecular fue Sidney Brenner, un biólogo inglés asociado al descubrimiento de los ARN mensajeros (las estructuras donde se copia el mensaje genético y a partir de las cuales se construye una proteína). Este cambio va acompañado de un cambio en la literatura científica que se destacaba por ser muy elaborada y con expresiones hasta poéticas en el área, con un gran énfasis en el manejo de las ideas y una gran osadía en avanzar hipótesis. Recordemos el aforismo de Jacques Monod que decía que una teoría formulada de manera bonita no necesariamente es verdadera, pero si su formulación es fea seguramente es falsa. Luego aparece un pragmatismo muy importante donde se enfatiza en la comunicación científica basada en el hecho y en la descripción de elementos, y se va disminuyendo cada vez más lo que es la elaboración intelectual, con un componente ingenieril importante en el desarrollo del trabajo.

Rápidamente enumeramos otros elementos de este segundo período: la postulación del concepto de gen por Morgan y la definición de la posible función de los cromosomas por Boveri. Hoy es muy fácil adquirir un gen, pero cuando Morgan lo hizo a partir de su genética en *Drosophila* no era así. Un gen es una unidad funcional responsable de caracteres precisos en un organismo, que tiene un locus físico preciso. Hasta ese momento se conocían las leyes de Mendel pero no los genes. Además, De Vries propone el concepto de mutación.

Volviendo un poco atrás, podemos decir que el trabajo con las fermentaciones es otro de los grandes avances de Pasteur, que las describe como fundamentalmente asociadas a los seres vivos, a la levadura. Demuestra que no existe generación espontánea. En Alemania se dice que la fermentación está asociada a organismos en descomposición. A partir de esa polémica se describen las enzimas. Se describen transformaciones químicas tanto en organismos vivos como en sus productos. A principios del siglo XX se producen una serie de contribuciones notables que tienen que ver con las propiedades químicas de macromoléculas biológicas. Algunas macromoléculas biológicas puestas en un tubo de ensayo, cumplían funciones particulares, en las que catalizaban procesos químicos. Se habla así por primera vez de las enzimas, y muy tempranamente se describe su comportamiento. Las primeras enzimas que se estudian son las que degradan azúcares, por estar asociadas a las fermentaciones. Fischer propone el modelo de la

llave y la cerradura, cuando dos moléculas tienen que encontrarse de manera específica para constituir una estructura o para asegurar una función. En los años 20 del siglo XX se realizan las primeras purificaciones enzimáticas (purificar es separar una molécula del resto de la célula). La primera enzima que se intentó purificar fue la xantina oxidasa y lo hizo Malcom Dixon, que casi lo logra. Poco tiempo después se logra cristalizar una enzima. Más precisamente en 1924 se cristaliza la ureasa (enzima que degrada la urea).

Primero se veía a las enzimas como una especie de caja negra, después se dice que son proteínas, porque tienen un comportamiento termolábil. Hay que esperar casi hasta la tercera década del siglo XX para decir que una enzima es una proteína. Hoy sabemos que hay enzimas que son proteínas, pero también existen otro tipo de moléculas que tienen actividad enzimática. A partir del estudio del comportamiento de las enzimas se desarrollan modelos, que permiten caracterizar fenomenológicamente cómo funcionan esos mecanismos enzimáticos. Brown, un investigador inglés, observa que existe una relación entre la velocidad de degradación de una enzima y la cantidad del producto obtenido, y esta relación es distinta a la que se conocía en los problemas de catálisis químicas. Hoy esos modelos fenomenológicos son los estudios de cinética enzimática de los alemanes Michaelis y Menten, herramientas que siguen siendo tremendamente poderosas y por las cuales continuamos estudiando las enzimas. Se empiezan a purificar algunas moléculas, con lo cual ahora es posible estudiar su función en un tubo de ensayo, se pueden comenzar a aislar.

Siguen dos corrientes culturales: una de la física que empieza con las leyes de Bragg y Bragg en las primeras décadas del siglo XX; y la otra es un acercamiento de las leyes de la difracción a las moléculas biológicas por Bernal. El encuentro entre Bernal y Pauling acelera el avance en este campo. Bragg padre y Bragg hijo en Inglaterra, describen las primeras leyes que permiten recomponer la estructura de un objeto a partir de la difracción de rayos X. La forma en como difracta un haz de rayos X en un objeto dado, sigue ciertas reglas. Es posible recomponer el objeto a partir de las manchitas que corresponden a los haces difractados. ¿Cómo se hace esto? Las leyes de Bragg permiten construir leyes matemáticas, para poder reconstruir el objeto. Cuando se descubren esas relaciones, los Bragg obtienen el premio Nobel, pero se considera absolutamente imposible que eso pueda ser aplicado a una molécula biológica, porque son moléculas de un cierto tamaño. Esta corriente cultural se cruza con otra proveniente de los genetistas y los bioquímicos. Las primeras estimaciones de tamaños de moléculas biológicas vienen de las experiencias de Svedberg, que desarrolló la centrifugación. Las proteínas se empiezan a caracterizar por una serie de propiedades. Con respecto a la estructura química se sabe que son polímeros de aminoácidos. A principios del siglo XX, Fischer describe el enlace peptídico, define a las proteínas como cadenas donde se repiten una serie de elementos y se van identificando todos los aminoácidos. Cuando Svedberg pone a punto la centrifugación, analiza distintas proteínas y empieza a obtener una serie de coeficientes de sedimentación que ahora llamamos "s". Por técnicas de ultracentrifugación se puede calcular el peso molecular de una partícula, así Svedberg obtuvo los siguientes pesos moleculares: 8, 16, 24, 32 KDa. Por lo cual llega a la conclusión de que las proteínas son polímeros de monómeros de 8 KDa. Esto significaba ya tamaños extraordinarios. Con esos tamaños, que hoy nos parecen proteínas pequeñas, aplicar las leyes de Bragg parecía absolutamente imposible. Sin embargo aparece Bernal, que juega un rol muy importante en la ciencia, por sus contribuciones personales y por su visión de la misma. En los años 30, Bernal dice que es posible cristalizar estas moléculas y debe

ser posible también resolver la estructura de las mismas por difracción de rayos X.

Poco tiempo después, Linus Pauling hace un par de contribuciones mayores. Primero describe la importancia de los enlaces débiles en las moléculas biológicas. Estudia las moléculas que están enmarcadas por la química, por los enlaces covalentes, y pone el acento en la importancia de los enlaces débiles, de cuyos estudios surge el libro llamado: “El puente de hidrógeno”. Señala la importancia de las uniones débiles, en particular de los puentes de hidrógeno, en la construcción de las macromoléculas biológicas. Por otro lado es, como ya mencionamos, quien describe por primera vez, en los años 40, que la estructura de una proteína puede enrollarse en una  $\alpha$  hélice.

### **Grandes encuentros**

A partir de este momento empiezan a confluir dos áreas de trabajo completamente distintas. Por un lado, mencionamos a Morgan con la definición de los cromosomas y la definición de cuál era la posible función de los cromosomas por Boveri; por otro lado, además del componente genético nace un componente estructural, que permite ir progresivamente estudiando y caracterizando macromoléculas biológicas. Bernal dice que pueden cristalizarse las proteínas, Svedberg que se pueden medir pesos moleculares y Pauling habla de la importancia de los enlaces débiles.

Además tenemos los estudios de Morgan sobre *Drosophila*, y estudios sobre levaduras, sobre microorganismos y demás, que van a desarrollar la genética moderna, y llegan a un máximo cuando Jacob y Monod postulan, analizando bacterias, el primer modelo de regulación de la expresión génica. A partir de cruzamientos entre bacterias que tenían distintos tipos de deficiencias y estudiando la función que resultaba de ese cruzamiento, empiezan a deducir cómo se regula el sistema que tiene la información para construir ciertas moléculas biológicas. Monod estudia cómo crece un cultivo bacteriano en una mezcla de glucosa y lactosa. Este experimento lo lleva a dos interpretaciones distintas. Por un lado describe el comportamiento nutricional de las bacterias como un fenómeno de adaptación enzimática, la enzima necesaria para degradar cierto tipo de compuesto no está preparada por lo que se requiere de un ajuste. Esto correspondía a un problema en la regulación de los genes que sintetizaban las enzimas necesarias para degradar glucosa o lactosa y observó que mientras había glucosa era imposible degradar lactosa, porque las enzimas necesarias para su degradación no se fabricaban. Así postula el primer modelo de regulación de la expresión génica junto a Jacob y define conceptos que son válidos hasta hoy, el concepto de gen regulador (un gen que puede encender o reprimir la expresión de otros genes). El gen represor fue aislado por Gilbert a mediados de los años 60.

Por otro lado, Monod aporta una reflexión derivada del fenómeno de adaptación enzimática y por primera vez cuestiona seriamente la imagen que había propuesto Fischer sobre cómo dos moléculas pueden reconocerse, el modelo de la llave y la cerradura. Monod describe la existencia de una serie de etapas de adaptación en el reconocimiento de dos moléculas biológicas. Esta adaptación mutua entre moléculas biológicas cuando se reconocen va a ser parte de la herencia que el siglo XX le deje al XXI. El modelo del operón de Jacob y Monod es de principios de los años 60. Entre este componente genético y el componente estructural, la distancia era enorme y hasta ese momento seguían vías casi paralelas. Mientras tanto, los bioquímicos se afanaban en

tratar de elaborar conceptos como los de Jacob y Monod, que definen qué tipo de moléculas se tienen que encender o apagar en un momento determinado, y buscan trabajar en la misma área. Por estudios fisiológicos o genéticos empiezan a describir determinadas funciones o procesos biológicos. Tratan de encontrar qué moléculas producen un funcionamiento como el que se observa en los tubos de ensayo. Los estudios de Jacob y Monod, entonces, postulan un modelo de regulación primero, y rápidamente empiezan los bioquímicos a poner tubos de ensayo a las etiquetas que Jacob y Monod habían inscripto como una molécula que tiene que ser represora, una molécula que tiene que ser capaz de copiar los mensajes, etc. Estos son los grandes encuentros donde confluyen las distintas vías.

A través de distintos avances mayores, lo que fue el estudio de desciframiento de los procesos metabólicos, época de oro de la Bioquímica, se empiezan a ver los mapas metabólicos, complejos circuitos donde se detallan todos los procesos químicos que tienen lugar en la célula. El descubrimiento y la postulación del modelo de doble hélice del ADN en el año 1953, sin ninguna duda constituye un primer punto de confluencia. El impacto social de la doble hélice empieza en esa época, pero va a tener su peso mayor durante los años 60. El descubrimiento del modelo de doble hélice, aparece menos de diez años después de la demostración hecha por Avery, de que el ADN es el portador de la información genética, a partir de experiencias con neumococos y con virus donde se demuestra que transfiriendo ADN se transfiere información. Esto es a fines de los años 40, principios de los 50, y en ese descubrimiento se sintetizan tres observaciones y tres ideas mayores del siglo XIX, que aparecen casi en el mismo momento: teoría de la evolución de 1859, descubrimiento de los ácidos nucleicos pocos años después, y las observaciones de Mendel de las relaciones cuantitativas en los caracteres transmitidos hereditariamente. Esas tres observaciones, esas tres grandes contribuciones, si se quiere tres grandes pilares fundadores de una disciplina, confluyen en el descubrimiento de que el ADN es el depositario de la información genética.

El tema de la doble hélice representa más de lo que ahora es la Biología Molecular, la Biotecnología. El modelo de doble hélice es un producto de la clarividencia de Bernal y de la intuición de ciertos ámbitos académicos, en particular en Inglaterra –y en Estados Unidos en el Laboratorio de Linus Pauling– de que era posible estudiar la estructura de macromoléculas biológicas. Comienzan en Inglaterra Kendrew y Perutz, dos jóvenes científicos, a intentar resolver la estructura de dos moléculas proteicas como eran la mioglobina y la hemoglobina, porque eran las únicas que podían purificar en grandísimas cantidades. Sabían que para cristalizar una molécula tenía que estar absolutamente pura. Las únicas moléculas que podían purificar en grandes cantidades eran esas dos, si hubieran tomado a la ovoalbúmina o alguna otra, hubieran tenido algún problema. Las cristalizaron y les llevó muchísimo tiempo resolver la estructura de estas dos moléculas. Al mismo tiempo, un par de laboratorios ingleses y también Linus Pauling, inician el estudio de la estructura de los ácidos nucleicos. Más que la estructura de la doble hélice y de la mioglobina y la hemoglobina, la intención es mostrar que es posible resolver la estructura de las macromoléculas biológicas, y eso abre un campo nuevo de pensamiento. Progresivamente, distintas líneas de pensamiento (la Genética, la Bioquímica, la Físicoquímica biológica o la Bioquímica estructural) confluyen. Con esa historia quería contarles algunos problemas en la construcción de algunos conceptos, en torno a ciertas macromoléculas biológicas.

### **Acerca de la estructura de las proteínas**

Voy a saltarme el Código Genético que es muy conocido, y voy a hablar de Anfisen, Gajdusek y Prusiner. Como dije anteriormente, se postula cómo está estructurada químicamente una proteína a principios del siglo XX, con la teoría de Fischer en 1903, donde describe el enlace peptídico como un enlace covalente entre aminoácidos. Estas moléculas rápidamente se asocian a ciertos procesos biológicos. Una buena parte de ellas intervienen en reacciones químicas y van a formar parte de las enzimas. No se sabe lo que son, pero se empieza a poder estudiar su función a través de las reacciones químicas en las cuales pueden intervenir. Se comunican los resultados de Svedberg, relacionados con la centrifugación de moléculas biológicas, por los años 20, y aparece ahí un personaje, que es un científico chino llamado Wsieh Wu. Este es el responsable de una gran cantidad de problemas, que los estudiantes de Biología y Bioquímica tenían con las proteínas, las cuales eran unas moléculas bastante misteriosas con unas propiedades extrañas que hace algo más de treinta años. Si se calentaban, de pronto, podían pasar varias reacciones raras: coagulaban, se unían formando una masa insoluble como le pasa a la clara de huevo; o de pronto si a una solución conteniendo proteínas se le agregaba sal entonces floculaban y precipitaban. Se veían una serie de propiedades extrañas, que en realidad eran cambios de estado de la proteína. Wsieh Wu en realidad estaba tratando de deducir qué era una proteína y cómo estaba estructurada a partir de esos cambios, cuando se alteraban las condiciones del medio. Wsieh Wu demostró que no hay cambio químico de la proteína sino físico en la desnaturalización. Entonces si calentamos una proteína se insolubiliza, si agregamos sal, si cambiamos de pH, también cambia su estado. ¿Qué quiere decir todo esto? Si las proteínas tienen que estar en estado lábil –que se llamaría metaestable luego– de alguna manera todo esto prepara a Linus Pauling para decir poco después que tiene que haber enlaces débiles, y también de alguna manera está definiendo que en ellos están implicados puentes de hidrógeno. Es la contribución sin ninguna duda fundamental para empezar a resolver la estructura de las proteínas.

Después las cosas suceden rápido, ya viene Pauling con la descripción de los enlaces débiles, el puente de hidrógeno y la  $\alpha$  hélice. Pauling describe a la  $\alpha$  hélice de la siguiente manera: distintos residuos aminoácidos unidos por enlaces peptídicos, que forman una estructura helicoidal estabilizada por puentes de hidrógeno. Cuenta la leyenda que encontramos en la galería de retratos, que Pauling dibujó una cadena polipeptídica en un papel y empezó a enrollarlo, para ver como podía estructurar eso en el espacio. Así se dio cuenta de que formaba una estructura regular, que podía intercambiar puentes de hidrógeno en ciertos grupos particulares y que además ese modelo podía explicar muchas de las propiedades que tenían las moléculas biológicas.

El siguiente paso sin ninguna duda lo dan Kendrew y Perutz. Estamos en los años 20 a principios de siglo, años 40, principios de los 50, allí resuelven la estructura de la mioglobina y la hemoglobina. Ven también otro tipo de estructura, en la que las cadenas polipeptídicas pueden disponerse de tal manera que forman puentes entre ellas y describen las hojas plegadas  $\beta$ , como parte de las estructuras posibles dentro de las proteínas. A partir de ahí se desarrolla una gran cantidad de esfuerzo para empezar a relacionar y estudiar, lo que llamaríamos la relación estructura-función de macromoléculas. Si tenemos la estructura cristalográfica, conocemos las propiedades, y podemos empezar a ver que pasa con cada uno de los componentes y como esos componentes pueden estar implicados en una función biológica precisa.

La comparación de la hemoglobina de Petrucci y la mioglobina de Kendrew, ambas fijadoras de oxígeno, pero con propiedades diferentes, permite avanzar mucho en la comprensión del funcionamiento de estas moléculas, y motivan por un lado a que haya muchas inversiones en los grandes sistemas científicos para estudiar la estructura de las macromoléculas biológicas y que muchos científicos se acerquen a este tema.

Estamos a fines de los años 60, principios de los 70, con las contribuciones de Anfinsen, que es otro de los grandes hitos. Aparece entonces la siguiente pregunta: ¿Cómo está determinada, cómo está codificada, la estructura tridimensional de una molécula? Rápidamente se ve que las propiedades de una proteína están determinadas por la estructura tridimensional. Lo que le pasaba a las moléculas de Wu era que al agregarle sales, al calentar, de alguna manera interfería con todos los enlaces débiles en particular con los puentes de hidrógeno, y hacía que la molécula, según el término introducido por él, se desnaturalizara. Ese estado desnaturalizado, se pensaba que correspondía a una estructura que se desenrollaba y quedaba bajo la forma de un polímero completamente desenrollado. Se empiezan a hacer estudios fisicoquímicos que permiten caracterizar ese estado desnaturalizado. Si se coloca un agente desnaturalizante, que puede ser un agente químico, el calor, o un agente físico, y se mide algún tipo de parámetro que permita caracterizar la estructura desenrollada, como la absorción de luz ultravioleta, la fluorescencia de los residuos aromáticos que depende del entorno, algo más sencillo como es la medición en los cambios en las propiedades de absorción de la luz polarizada que varía según la estructura que tenga una macromolécula biológica; entonces según el tipo de parámetro que se mida se van a obtener resultados diversos. Pero se ve en general el pasaje desde un estado nativo donde la proteína o enzima tiene una función, a un estado desnaturalizado, por transiciones que son cambios de fase, en el cual la misma proteína o enzima no tiene dicha función. Esto nos parece ahora muy banal pero entonces era una herramienta muy potente.

Todo está muy bien, hasta que a alguien se le ocurre tratar de pasar desde el estado desnaturalizado al nativo y lo logra. Es decir se parte de un tubo de ensayo con una proteína completamente desplegada y se vuelve a las condiciones nativas, logrando recuperar la función que tenía esa molécula. Esta experiencia la produce Anfinsen, con lo cual gana el premio Nobel en 1972, con una proteína que es una ribonucleasa. Toda la información necesaria para construir el edificio tridimensional de una proteína, está en la sucesión de los elementos que la componen. Las proteínas difieren únicamente, aparte de su tamaño, por la sucesión de los monómeros, de los componentes primarios. Así la sucesión de aminoácidos determina la estructura tridimensional. Eso provocó un impacto notable, y fue sancionado con premio Nobel.

A partir de ese momento se inician varias líneas de trabajo. Si toda la información está en los aminoácidos, es decir que podemos descifrar un código, que nos permita predecir a partir de una sucesión de aminoácidos cómo va a ser la estructura tridimensional de la proteína. Después, con el advenimiento de la Biología Molecular, podemos pasar de ADN a proteínas, conociendo la estructura del gen podemos tener la estructura tridimensional de la proteína. Si toda la información está contenida en la sucesión de aminoácidos, entonces dos estructuras con una sucesión similar tienen que dar lugar a un edificio tridimensional similar. Pasamos de la información lineal a construir un edificio en tres dimensiones.

Entonces se trata de descifrar el código, y hay dos aproximaciones que siguen siendo válidas

hoy. Estamos en la época de los años 70, todavía las computadoras son estructuras gigantescas. Se hacen dos tipos de estudios: estudios de propiedades físicas de las moléculas, para ver cuáles son las propiedades físicas de los aminoácidos, que van a determinar una estructura dada, y estudios probabilísticos. Para eso se sabe que distintos residuos aminoacídicos tienen propiedades y tamaños diferentes. La cadena lateral de un aminoácido, tiene propiedades físicas muy diversas: puede ser hidrofóbica, polar, cargada positiva o negativamente, aromática, etc. Se empieza a tratar de ver cómo dos aminoácidos unidos por un enlace peptídico pueden girar en torno a ciertos ejes y las disposiciones estéricas que determinan los distintos constituyentes, tratando entonces de descifrar la información bajo su forma física. Apenas se dispone de métodos computacionales de más comodidad, empiezan los análisis estadísticos. Se estudia en una estructura conocida, cuales son las series de aminoácidos que van a determinar las propiedades. Se dice que determinada secuencia aminoacídica tiene determinada probabilidad de estar en hoja  $\beta$  o en una hélice  $\alpha$ . Estos métodos se siguen usando, a tal punto que hoy están disponibles en la mayor parte de los paquetes computacionales, de todas las computadoras de todos los laboratorios. Con una secuencia conocida se puede saber la estructura de la proteína, según distintos algoritmos estadísticos o basados en métodos físico-químicos. Se aceptan esas propuestas como una aproximación bastante cercana a la realidad.

### **¿La caída de un dogma?**

Lo que dijo Anfisen entonces de que toda la información necesaria para determinar la estructura tridimensional de una proteína está en su estructura primaria, pasa a ser un dogma, hasta hace muy poco tiempo, que aparecieron una serie de problemas. Pero lo más notable se produce con la aparición de una molécula, con la que el dogma empezó a caerse solo.

La Resonancia Magnética Nuclear permite resolver la estructura de proteínas pequeñas de alrededor de 120 a 150 aminoácidos en solución, por difracción de rayos X se resuelven cristales. Por este método se postula la existencia de una molécula, en 1982 por Prusiner, y se la identifica como la responsable de una enfermedad en las ovejas llamada Scrapie, el Prurigo Lumbar en español. Es una enfermedad conocida del siglo XVIII, donde las ovejas tienen una patología nerviosa cuticular. Esta aparecía en forma esporádica en las manadas y no planteaba demasiados problemas. Los veterinarios estaban preocupados por controlarla, se pensaba que correspondía a una enfermedad viral. Pero no era una enfermedad común. El problema aparece cuando Gajdusek, un virólogo norteamericano, va a estudiar una epidemia que se desarrolló en una pequeña isla de Nueva Guinea, una enfermedad neurológica muy severa llamada Kuru, contagiosa, que estaba diezmando los nativos. Llega y ve que esa enfermedad se parece muchísimo a otra enfermedad conocida desde principios de siglo, que es la enfermedad de Creutzfeld-Jacob. Otra enfermedad neurológica que daba patologías esponjiiformes, en las autopsias. Estas patologías esponjiiformes son depósitos de proteínas, precipitaciones o agregados de proteínas, similares a la coagulación de la clara de huevo. Esta patología rara aparecía en un caso cada un millón de habitantes. Lo que observa Gajdusek, en la isla de Papúa en Nueva Guinea, son síndromes neurológicos muy similares a la enfermedad de Creutzfeld-Jacob, pero son contagiados. Trata de aislar un virus y enseguida se da cuenta de que el contagio se produce por los rituales funerarios de los nativos, que eran caníbales. Se logran implementar medidas sanitarias, que cortan y terminan con el canibalismo y en poco tiempo la enfermedad desaparece.

A partir de ahí, se busca identificar un virus, y es Prusiner quien encuentra a esta molécula que precipita, y forma esos agregados. Al precipitar la proteína se agregan sus superficies hidrofílicas, y la proteína debe presentar alteraciones o plegamientos que expongan sus superficies hidrofóbicas. Si se calienta la ovoalbúmina, se rompen los puentes de hidrógeno, y en algún momento la ovoalbúmina comienza a exponer sus superficies hidrofóbicas. Esto es lo que observaba Wsieh Wu, que ya en los años 20 hablaba sobre la posibilidad de la intervención de superficies hidrofóbicas. Lo que es más interesante es lo que encuentra Prusiner, es que esa molécula que estaba causando la patología y que precipitaba, tenía una conformación tridimensional cuando era normal y al precipitar tenía una conformación diferente. Por lo que la misma proteína estaba en dos estados diferentes.

Dos proteínas con la misma secuencia, con la misma sucesión de aminoácidos, una toma una estructura tridimensional y la otra toma otra. Por lo cual no toda la información para la estructura tridimensional está en la sucesión de aminoácidos, una misma sucesión puede conducir a dos estructuras diferentes.

Estas precipitaciones se ven también en la enfermedad de Alzheimer, en el mal de Parkinson, enfermedades a las que los médicos patólogos llamaban enfermedades amiloides, porque proteínas que hoy sabemos que son proteínas normales cambian su estructura tridimensional, exponen sus superficies hidrofóbicas y precipitan. Los depósitos, que se llamaban depósitos amiloides, tienen proteínas que son solubles y de pronto forman agregados y fibras. Entonces pasamos del enlace peptídico a la predicción de la estructura tridimensional y terminamos diciendo que no está todo en la estructura lineal de la proteína. Lo que es más complicado es que esa estructura anormal, la estructura que forma esos agregados, es capaz de contagiar a una proteína con estructura normal. O sea que esta partícula que Prusiner llamó Prión, puede transformar a la proteína normal en anormal.

Hemos recorrido casi un siglo de estructura de proteínas, pasando por un dogma tremendamente fuerte como fue el de Anfinsen, para llegar hoy, a principios del siglo XXI, a no saber exactamente cómo se determina la estructura tridimensional de las proteínas. ¿Es totalmente cierta esta afirmación? No lo es. De alguna manera lo que observó Anfinsen es cierto, y de alguna manera es muy claro que la sucesión de aminoácidos, que la sucesión de unidades de una proteína, van autorizando ciertas posiciones y desautorizando otras, pero eso no alcanza. ¿Qué es lo que falta? Eso está llevando muchísimo trabajo, es uno de los grandes desafíos de hoy. Para el estudio de este tipo de problemas se han destinado medios de cómputo gigantescos, se han desarrollado computadoras superpotentes, seguramente sea uno de los grandes temas para la próxima década. ¿Cómo se determina la estructura de las proteínas? Era ya un tema mayor, hasta que se lo asocia a ciertas patologías, que de pronto tienen muy poca incidencia hasta hoy en salud humana.

La proteína que encontró Gajdusek y que Prusiner identifica con un prión es el agente infeccioso de la enfermedad de la Vaca Loca. En este momento hay cien casos descritos en humanos, de contagio de la enfermedad de la Vaca Loca. Cien casos no son muchos, pero esta patología crea un gran sentimiento de inseguridad y de desconfianza. No es posible confiar más en los alimentos, problema de consumo. Parte de esta inseguridad es porque es una enfermedad que se incuba durante muchísimo tiempo y tiene un impacto importante desde todo punto de vista

en las sociedades. El problema de que la estructura tridimensional de una proteína está dada no sólo por la sucesión de aminoácidos, es importante ahora no sólo por los priones, sino porque patologías similares sin el componente transmisible que es un prión (esa molécula anormal que contagia el cambio de una estructura a la otra que sería la base de la enfermedad priónica), el mismo tipo de mecanismo sin el componente priónico, se da en una gran cantidad de patologías del sistema nervioso. Hay una gran cantidad de patologías donde precipitan las proteínas más diversas, proteínas que están tranquilas dentro de una célula y de pronto precipitan. Los patólogos llamaban a esas enfermedades idiopáticas, algunas de ellas las clasificaban en el grupo de enfermedades autoinmunes. En todos estos casos hay una proteína que pasa de tener una estructura tridimensional a otra. Hay proteínas que siguen comportándose como la ribonucleasa de Anfinsen, son proteínas tipo Anfinsen, y otras que no.

En este recorrido centrado en la estructura de las proteínas, estamos recorriendo prácticamente un siglo, hemos pasado por dogmas, pasamos por trabajos que resultan impresionantes, para hoy ver conceptos muy sencillos, ideas muy sencillas que cambian toda una corriente de pensamiento. Los diagramas energéticos, son gráficos tridimensionales donde en una superficie se representan los estados desnaturalizado y nativo de una proteína. El estado nativo se observa como un pozo de potencial. La construcción del diagrama energético se realiza a partir de experiencias de desnaturalización, se estudia el comportamiento de la proteína, cómo se va desplegando y replegando, y se va descifrando entonces la presencia de ciertos pozos de potencial. En el caso de algunas proteínas, cuando se parte de la misma completamente desplegada y se vuelve a las condiciones iniciales nativas, se encuentra sólo un pozo de potencial. Esta es una proteína tipo Anfinsen (aprox. el 20 o 25 % de los casos).

Pero otras proteínas, si están completamente desplegadas y se las vuelve a renaturalizar, pueden tener caminos alternativos, caer en pozos de potencial diferentes. Es posible definir un número pequeño de pozos mayores, que se puede determinar experimentalmente. Lo que es interesante es que si tenemos este tipo de situaciones, tal vez se pueda desnaturalizar una proteína y hacerla explorar otros caminos de renaturalización.

Hace exactamente veinte años, a principios de los años 80, se describe una molécula, una proteína que al purificarla no era muy activa, por lo que debería ser del tipo de proteínas que tienen más de un pozo de potencial en el diagrama energético. Se descubre que esta proteína asociada a otra recupera la actividad, hay una molécula que acompaña a la primera y se la bautiza chaperona. Estas chaperonas son un tipo de moléculas que ayudan a que una proteína adquiera su estructura tridimensional adecuada. Esto causa un gran revuelo y a partir de la primera que se encuentra, se empiezan a encontrar otras y se descubre toda una familia de proteínas que ayudan a otras a adquirir su estructura tridimensional. En particular aparecen algunos elementos interesantes. El primero de ellos es que esta familia de proteínas, bautizadas chaperonas, aumenta en la célula en condiciones de estrés, cuando la célula se tiene que adaptar a una nueva situación, y ayudan a que otras proteínas encuentren su conformación nativa. Este tipo de proteínas chaperonas aumentan cada vez que una célula cambia su estado de base, cuando una célula se va a diferenciar, a cambiar su programa de expresión génica para transformarse en otro tipo de célula, cuando se va a dividir, o cuando sufre agresiones de diferente tipo. Allí aumentan las chaperonas que ayudan a que las proteínas adquieran su conformación adecuada.

Lo más interesante es que, en el año 1995, se describe por microscopía electrónica una familia de estas chaperonas, y se descubre que son proteínas con un agujero. Se piensa que puede ser un artefacto, pero la proteína se cristaliza y se ve que en realidad está compuesta por una gran cantidad de subunidades, según el organismo siete u ocho, y que en realidad no es un agujero sino dos. La superficie interna del agujero está tapizada por residuos fuertemente hidrofóbicos. Dentro del agujero puede entrar una proteína, y quedar allí si está exponiendo superficies hidrofóbicas. Si hay una proteína que expone residuos hidrofóbicos dentro de la célula, ésta va a asociarse rápidamente con otras superficies similares. Esto puede ser un comportamiento fisiológico normal, cuando la proteína va a fabricar un edificio más complejo con otras proteínas, entonces esa superficie hidrofóbica es presentada durante un tiempo corto, hasta que se encuentra con la otra proteína. Pero si su destino no es formar un edificio macromolecular, esa proteína que expone una superficie hidrofóbica es un peligro, porque inmediatamente va a encontrarse con otras moléculas similares y entonces precipita. ¿Qué es lo que hace el complejo de la chaperona? Cuando entra una proteína que fija residuos hidrofóbicos por un proceso enteramente estocástico, se cierra la tapa del agujero y entonces se desencadena un proceso muy interesante en el que se consume energía. Empiezan a hidrolizarse moléculas de ATP (la hidrólisis del ATP libera energía), que lo que hacen es encender un reloj, se va contando el tiempo, lleva bastantes minutos este proceso, y lo que va pasando adentro es lo que ocurre en los libros de niños con las imágenes que van cambiando continuamente y no quedan fijas, una imagen sustituye a la otra. Los residuos hidrofóbicos que tapizaban esta cavidad son sustituidos por residuos hidrofílicos y ahí se produce una catástrofe que hace funcionar a esto como una especie de olla a presión. Cuando se describe este funcionamiento se dice: y entonces se renaturaliza la proteína que está adentro. Esto llevó años de trabajo, hoy lo que decimos es: este proceso, y otros en los cuales están implicadas estas moléculas, lo que hacen es ayudar a que una proteína salga de uno de sus pozos de potencial y pueda explorar otras alternativas de plegamiento.

Termino esta parte con lo siguiente, una proteína con error de plegamiento dentro de la célula puede tener destinos múltiples; esto es importante por las patologías señaladas, también puede estar vinculada a la senescencia. Las proteínas anormales que presentan superficies hidrofóbicas pueden formar complejos específicos. Hay complejos macromoleculares que están estabilizados por interacciones hidrofóbicas, pueden pasar por las chaperonas que le van a permitir explorar otras vías de plegamiento, hay un sistema celular que permite degradar las proteínas anormales y compite con el sistema de las chaperonas, o si no sigue ninguna de estas vías se van a agregar y precipitar.

### **El modelo de la doble hélice**

En los años 30 empieza la época de los polímeros, los polímeros naturales, los polímeros de síntesis. Nacen la física de los polímeros y la química de los polímeros, son físico-químicos de polímeros que luego se transforman en físico-químicos de biopolímeros. Muchas de las leyes descubiertas y de la metodología del estudio de polímeros sintéticos terminan siendo aplicadas en moléculas biológicas. El poliestireno fue una de las grandes vedette, del cual se extrajeron una cantidad de reglas que luego se aplicaron a la resolución de la estructura del ADN y de la

cromatina.

A principios de los años cincuenta, el modelo de Watson y Crick del ADN en doble hélice (dos cadenas polinucleotídicas enlazadas, donde se suceden una serie de grupos químicos, pares de bases apareados por puentes de hidrógeno) es un modelo a partir de estudios cristalográficos a muy baja resolución. Van a pasar casi veinte años para que haya estudios cristalográficos finos y que muestren que hay una gran variedad de estructuras de ADN en torno a este motivo de doble hélice dextrógira, con variantes levógiras, y llegar a que esta molécula es altamente flexible.

El ADN es una molécula kilométrica, por ejemplo nuestro genoma tiene más de  $3 \times 10^9$  pares de bases, los cromosomas a escala celular son moléculas muy largas. Siempre apareció como una molécula rígida, hoy es considerada una molécula esencialmente flexible, capaz de modificar su conformación en forma muy variada. Esa flexibilidad conformacional, lleva a que si las posibilidades de codificar información están dadas por los grupos químicos, las variaciones conformacionales de torsión, flexión, etc., van a cambiar en función de la conformación, de las propiedades físicas de esta molécula. De una manera general, hoy se habla de que la información codificada en esta molécula es de tipo digital, en cuanto a que hay ciertos grupos químicos particulares que están o no están. Pero por otra parte hay una información de tipo analógico, que son las variaciones de la posición de todos esos grupos químicos dependiendo de la estructura de la hélice.

A mediados de los años 60 salen los primeros métodos de secuenciación, de elucidación de la estructura primaria de estas moléculas, donde se puede leer la sucesión de pares de bases, se lee la información genética. A partir de ahí se dan un par de premios Nobel a Gilbert y Ptashne por métodos que hasta hoy se utilizan, métodos manuales. En diez años aparecen los primeros métodos automáticos, se reproducen los métodos manuales en sistemas enteramente automatizados. A principios de los años 90, en un laboratorio medianamente equipado se podían leer unas 50 mil bases por año. Hoy, el promedio en un laboratorio medianamente equipado es de 50 mil bases por hora.

Estamos en un período muy especial de desciframiento de la información genética global de distintos organismos. Se está leyendo y descifrando la estructura de una gran cantidad de genomas. A nivel humano se tienen 3 mil millones de letras. El gran desafío es leer. La secuencia de bases del ADN es posible pasarla a aminoácidos, por el código genético, donde tres bases corresponden a un aminoácido dado, se puede traducir a estructura de proteínas lineal. Pero el genoma también tiene información de dónde, cuándo y cuánto se va a tener que leer el mensaje. Por ejemplo, la insulina es una proteína que tiene que aparecer en un momento dado, en las células que se van a diferenciar en el páncreas, un grupo de células muy particular en los islotes de Langerhans van a comenzar a producir insulina en un momento dado de la vida y del desarrollo del individuo.

¿Cuántas hojas ocupa el genoma humano? Un libro de bolsillo común y corriente puede tener 700 hojas y 500 mil caracteres, si se divide  $3,5 \times 10^9$  por  $5 \times 10^5$ , eso da 7.000. O sea 7.000 libros de bolsillo. El desafío es leer. Entonces, por un lado se está desarrollando una disciplina que se llama Bioinformática, y, por otro lado, hay una gran cantidad de trabajos y aproximaciones que permiten estudiar mensajes precisos, porque se conoce mucho en cuanto a propiedades físicas y

químicas de determinados alineamientos de letras que se pueden empezar a leer. Pero hay que leer mucho.

Para poder leer 50 mil bases por hora se utiliza una máquina en la que se introduce la molécula con cortes en distintos lugares correspondientes a una u otra base y se va separando por distintos principios, en particular por electroforesis. Hay un detector que va leyendo, se utilizan sondas coloreadas y lo que sale se lee en una computadora que reconoce a las distintas bases según el color. El aparato de la Facultad de Ciencias tiene 32 pistas y logra leer en cada pista 500 bases en 2 horas, son 16.000 bases en 2 horas, en una jornada de trabajo liviano de aproximadamente 8 horas se leen 64.000 bases por día, trabajando en un aparato mediano para chico. En principio ahí tenemos toda la información. En el genoma humano hay de 30 a 40 mil genes distintos, que pueden dar lugar a una variedad mayor de proteínas por distintas combinaciones que tienen lugar en la célula y esa información se va a procesar y leer de distinta manera en las distintas células. Se puede tomar un tipo celular y estudiar todas las proteínas que tiene una célula dada, en un momento dado. Estas proteínas son marcadas con un precursor radiactivo, que permite la detección de cantidades incluso muy pequeñas de proteína y se separan según procedimientos que permiten desplegarlas en dos dimensiones; se observan como un conjunto de manchitas en un gel bidimensional. Esto es muy sencillo de realizar, se hace en varios laboratorios en Uruguay, y se pueden ver varios miles de proteínas. Hoy es posible recortar una de las manchitas e identificar qué proteína es. Se la puede poner en un espectrómetro de masa y calcular su masa con una mínima fracción de error. Se puede fragmentar a la proteína por métodos físicos, en particular por espectrofotometría de masa. El espectrofotómetro de masas que tenemos en el Uruguay mide cuánto se desplaza una molécula en un campo. Se mide el tiempo que demora una proteína estabilizada en llegar a un campo, se mide el tiempo de vuelo, cuánto demora la proteína en alcanzar ese campo, lo que va a depender de su tamaño y de su carga, así se puede calcular su peso molecular con gran precisión. Tomando una manchita de una célula humana, como tenemos todo el genoma, podemos traducirlo a proteínas y podemos identificar qué proteína es con ese peso molecular. Entonces, en un momento dado, haciendo este tipo de análisis, se puede ver qué está pasando en una célula, qué cambios en la expresión de los genes se producen, frente a determinadas alteraciones patológicas o fisiológicas. Pasamos de purificar una única proteína a poder estudiar el conjunto de las proteínas de la célula.

¿Cómo se lee toda la información que está contenida en una secuencia, en esa sucesión de cuatro letras? Se construyen microarreglos de ADN (que son microchips de ADN) con métodos similares a los utilizados para los chips o procesadores clásicos. Se hacen arreglos que contienen distintos mensajes, distintas partes de secuencias de ADN de los mensajes posibles que puede tener una célula. En un momento dado se puede extraer, de una célula dada, los mensajes que esa célula está leyendo. De toda la información genética que tiene, la que está leyendo. Cuando se lee el ADN, la lectura es la copia de una de las hebras, el ARN mensajero. Este ARN mensajero tiene una estructura complementaria a una de las hebras. Entonces, basándonos en este fenómeno, lo que se puede hacer es partir de un arreglo donde hay una gran cantidad de secuencias de mensajes posibles, y obtener los mensajeros que una célula tiene en un momento dado; supongamos que a la secuencia del microchip le puedo poner algo que tenga una reacción coloreada, entonces allí se van a pegar los distintos mensajes complementarios a la información que se puso. En un momento dado, entonces, se tiene una fotografía de todos los genes o de una gran cantidad de genes que se están expresando en una célula. De todos esos 3.000 millones de

pares de bases que están en la célula humana, organizados en 30 a 40 mil genes, se puede saber cuáles se están expresando más o cuáles se están expresando menos, cuáles se están leyendo y cuáles no.

A partir del año 1976 se empiezan a desarrollar las herramientas que permiten trabajar con el ADN, con la información genética. Empiezan las técnicas de ADN recombinante a principios de los años 70. Así se comienzan a estudiar nuevos territorios y cada vez más pequeños. Este período es sucedido por un período muy pragmático, con grandes posibilidades de aplicaciones inmediatas a partir de los resultados que se obtienen con las nuevas técnicas. A partir de entonces se recorren un par de décadas de fuerte componente descriptivo. El siglo XX termina con un conocimiento, un manejo e intervención importantes del investigador sobre el conjunto de moléculas que constituyen los seres vivos. Uno de los problemas centrales, disponiendo de la metodología impresionante en ese momento, es resolver las bases físicas y químicas de la función de las macromoléculas biológicas, que va desde la estructura y función de las moléculas depositarias de la información genética, la estructura y función de las moléculas que constituyen los edificios celulares, hasta la de las moléculas que aseguran las funciones y los procesos químicos.

### **El panorama a comienzos del siglo XXI**

Con esto llegamos entonces a principios del siglo XXI, donde estamos exactamente en la situación contraria a la de la frase que diría Jacob. Se están manejando volúmenes enormes de información, se están estudiando genomas completos, y hay métodos que permiten seguir el funcionamiento de todo el aparato de regulación de la expresión génica, del procesamiento de información y de la biosíntesis con métodos tremendamente potentes. Se están revelando grandes cantidades de información y es difícil lograr síntesis. Estamos en un proceso de desarrollo gigantesco de la información.

Por otro lado entramos en una etapa que llamaríamos de promiscuidad molecular. Todas esas moléculas que se habían estudiado durante años, por parte de los bioquímicos, dentro de tubos de ensayo, hoy se pueden estudiar dentro de la célula. Lo que aparece entonces es que ninguna molécula se encuentra sola, sino acompañada de otras moléculas, existe una gran promiscuidad molecular. Dada cualquier molécula en particular, va a interactuar con una gran variedad de otros socios. En algunos casos la interacción va a ser más específica que en otros. La especificidad la podemos concebir de la manera siguiente: mide el tiempo de residencia de una molécula sobre otra. Cada una de las moléculas interactúa con una gran variedad de otras, lo cual lleva a construir complejos circuitos de interacción entre distintas proteínas.

Si se toma cualquier libro de Bioquímica, Biología Celular o Biología Molecular, va a estar lleno en este momento de circuitos con bolitas de colores que representan a proteínas de distinta naturaleza, y donde cada una de las proteínas va a tener múltiples socios. Hoy uno de los grandes esfuerzos está en, por un lado, descifrar toda esa información, establecer el código de los códigos en última instancia, y, por otro lado, en entender como interactúan las moléculas entre sí. De todos los modelos existieron una o dos estructuras moderadas de macromoléculas. Las moléculas se han transformado en bolitas de colores. Nos olvidamos que las letras A, T, G, C, representan estructuras químicas que van a tener propiedades físicas particulares y que ahí está la información, y que las bolitas de colores son una forma muy esquemática de representar las

moléculas. Tenemos que lograr ver las limitaciones que nos impone ese lenguaje.

Sin ninguna duda, ordenar, descifrar lo que son las reglas que determinan estructuras, que determinan interacciones específicas, y descifrar al mismo tiempo esas relaciones entre estructura y función, cómo una estructura en particular determina propiedades específicas que van a explicar el funcionamiento de una molécula, son los temas que están abiertos aún hoy.

Quiero terminar con lo siguiente, tratando de ver algunas estructuras conceptuales importantes.

Mencioné el tema del operón; el modelo del operón de Jacob y Monod marcó un momento particular en la historia de la Biología, la Genética, las macromoléculas biológicas.

En el último cuarto del siglo XIX, Claude Bernard de alguna manera funda la Medicina experimental. Él descubre, entre otras cosas, el poder de ciertas drogas. Estudia mucho tiempo el curare sacado de una calabacita sudamericana, y eso introduce el método experimental en Medicina. A partir de ese momento se empiezan a utilizar drogas diversas, que existían en la naturaleza y que tenían efectos extraños, para ver cuáles eran los efectos a nivel de un organismo. Y a principios de 1900, una persona crea el concepto de receptor farmacológico. La idea de receptor farmacológico quiere decir que cualquier droga, cualquier sustancia que tenga efecto sobre un organismo, sobre un individuo, tiene efecto porque tiene del lado del tejido, de las células del organismo, un receptor, otra molécula que lo recibe. Esta idea llevó a avances notables, porque a partir de cualquier droga, desde el mate hasta el café, la aspirina o cualquiera de los alucinógenos, se sabe que tienen efecto porque actúan sobre el receptor de la célula. Esto llevó a resolver el modo de acción de una gran cantidad de sustancias y finalmente a identificar cómo es que una droga dada, puede tener efecto en cambiar el perfil de expresión de genes de una célula o producir determinado tipo de respuesta en la secreción de ciertas hormonas.

Un premio Nobel descubrió una serie de sustancias que se llaman endorfinas. René Dubos, un gran microbiólogo, fue un ecólogo microbiano antes de que existiera la Ecología Microbiana, en los años 20. En esos años se encontró con Avery, que llamó principio transformante al ADN. Porque el ADN era capaz de pasar la información de un neumococo virulento a uno no virulento y transformarlo en virulento. Eso llevó décadas en el laboratorio de Avery hasta el final de los años 40. Pero en los años 20 Avery estaba con René Dubos, y lo que le preocupaba era tratar de demostrar que el neumococo virulento tenía alguna sustancia transformante, y esto lo va a llevar al descubrimiento de que el ADN es el portador de la información genética. Además plantea que en la naturaleza hay equilibrios dados por los propios microorganismos, anticipándose al descubrimiento de los antibióticos. En 1929 Fleming descubre la penicilina.

Se pueden unir dos pares de bases cualquiera, una adenina con una adenina y la sucesión adenina, adenina siempre va a ser igual, sea cual sea la primera y la segunda, lo mismo puede pasar con los aminoácidos. Se tienen 4 bases que dan una serie de posibilidades de establecer mensajes dados, con 20 aminoácidos se tiene una gran variedad de proteínas. Pero si pasamos a los azúcares, los glúcidos tienen distintas posibilidades de unirse y dos azúcares pueden dar lugar a varios apareamientos distintos. Los azúcares intervienen en la comunicación de señales entre células. Aquí hay un territorio a explorar del cual apenas comenzamos a conocer algunos signos.

Agosto de 2001.