

EL CODIGO GENETICO COMO PUNTO CRITICO EN LA EVOLUCION DE LOS SISTEMAS BIOLÓGICOS

(Estudio de las posibilidades evolutivas del ARN enzimático)

Alvaro MORENO BERGARECHE y Julio FERNANDEZ

ABSTRACT

Firstly we consider the new results about enzymatic capabilities in the RNA. In this framework we analyse the sequence-folding duality as a precursor of the genotype/phenotype duality. We discuss then which are the evolutive potentialities and limitations for a system with the absence and the presence of a nucleic acid to proteins translator code. We study the arise of the code in the living systems as a form of deep interlocking between the logic of the machinery and its hardware, discussing some problems that abstract modelization of this process implies. Finally we propose to redefine the informational vocabulary usually employed in Molecular Biology.

INTRODUCCION

El problema del origen del código genético es, desde su descubrimiento, una cuestión ampliamente discutida dentro de la Biología Teórica. La cuestión es más complicada de lo que podría parecer en un primer momento por cuanto no solo se trata de proponer un mecanismo para su aparición, sino de resolver también algunas cuestiones que la aparición del código suscita.

Trataremos en este trabajo de abordar tres aspectos estrechamente relacionados con la aparición del código genético:

- (i) La aparición de la dualidad genotipo (cadena de caracteres para ser expresados)/fenotipo (estructura funcional).
- (ii) La validez del modelo darwiniano de evolución en los estados prebióticos.

- (iii) La capacidad de evolución de un sistema en presencia y ausencia de código respectivamente.

En la última década han aparecido teorías y modelos (Eigen & Schuster 1979, Csanvi & Kampis 1985, Pattes 1977 y 1982) que tratan de solucionar, en parte, los problemas planteados, pero que por otro lado hacen aparecer explícitamente algunos puntos oscuros. Lo que estos modelos han puesto de manifiesto es la dificultad de determinar cual es el papel del código en relación con el origen del mecanismo de evolución darwiniana, así como cual es la relación entre la aparición del mecanismo del código genético (información, traducción, etc.) y los requisitos lógicos que Von Neumann planteara a principios de los 50 en su bien conocida teoría de autómatas autorreproducibles (Von Neumann 1966). El problema al que Von Neumann intentará dar respuesta es el mismo que trataremos aquí, que requisitos debe cumplir un sistema para superar el umbral crítico a partir del cual ocurre un aumento de complejidad a través de sucesivos pasos de autorreproducción. En otras palabras: qué requisitos debe cumplir un sistema para poder ser considerado capaz de evolucionar. Si los sistemas vivientes cumplen con toda evidencia estas condiciones, la cuestión a la que intentamos contestar es: ¿cuáles son, en un sentido estrictamente físico, los requisitos fundamentales que permiten la evolución biológica?

DOBLE PAPEL FUNCIONAL Y REPLICATIVO DE LAS MOLECULAS DE ARN

Se considera que la evolución de los sistemas biológicos debió pasar desde sistemas moleculares con síntesis sin instrucciones de macromoléculas biológicas -lo que se ha dado en llamar evolución química- hasta la evolución biológica darwiniana clásica (desarrollo desde sistemas genéticos primitivos hasta organismos multicelulares complejos), a través de una fase intermedia -llamada evolución prebiótica- que consistirá en una síntesis con instrucciones de macromoléculas biológicas y organización de sistemas genéticos autorreproductores (Küppers 1983), donde debió tener lugar el proceso de aparición del código genético.

Durante este período de evolución prebiótica el salto fundamental debería haber ocurrido cuando se pasó de sistemas que contenían, exclusivamente, ácido nucleico replicándose a sistemas donde ya existiera

EL CODIGO GENETICO COMO PUNTO CRITICO

una traducción a proteínas. La necesidad de este paso venía dada por la esencial diferenciación entre las funciones respectivas de ácidos nucleicos y proteínas, funciones que respondían a rasgos básicos de sus estructuras. En efecto, de siempre se había considerado a los ácidos nucleicos como moléculas capaces de transportar y almacenar información dado que su estructura primaria y alguna de sus propiedades (alta capacidad de copia por molde, por ejemplo) son muy adecuadas para esta tarea. Sin embargo, nunca habían sido estudiadas en profundidad, aunque si consideradas (Crick 1976), sus capacidades enzimáticas, puesto que se suponían casi exclusivas de las proteínas. Estas por su parte, derivaban de su versatilidad secuencial en la estructura primaria de aminoácidos, una potencialmente enorme versatilidad estérica en su estructura plegada tridimensional (3D), y por ende funcional. Sin embargo el reciente descubrimiento de una serie de importantes actividades enzimáticas secuencialmente dependientes en el ARN (Cech & Bass 1986, Orgel 1986a) ha supuesto una brecha en el rígido esquema anterior. En otras palabras, tenemos un escenario en el que una misma clase de moléculas -los ARN- son, a la vez, capaces de autorreproducción y de actividades funcionales.

Evidentemente existen notables diferencias entre las capacidades enzimáticas del ARN y las proteínas y la principal proviene del siguiente hecho: mientras que las proteínas están constituidas por veinte aminoácidos diferentes y de diferentes tipos, en los ARN existe una gran homogeneidad proporcionada por las similares características químicas de las cuatro bases (Cech & Bass 1986). Así mientras que teóricamente se podría encontrar casi cualquier tipo de actividad enzimática en una proteína a base de buscar los aminoácidos adecuados en una estructura terciaria adecuada, esta versatilidad funcional sería mucho menor en el caso del ARN sin que, no obstante, podamos establecer límites precisos a su capacidad.

Pero lo importante es que ya no resulta irreal la hipótesis de que en un primer momento fueran los ARN los que soportaran -vía actividad enzimática- un rudimentario aparato de traducción que fuera evolucionando hasta ser sustituido, progresivamente, por proteínas.

De una u otra forma, está actualmente muy extendida la opinión de que los ARN con capacidad enzimática (denominados ribozimas, Cech & Bass 1986) pudieron tener un papel fundamental en el proceso

de origen del código genético, que en buena medida se debió apoyar en la posibilidad de actuación del ribozima a dos niveles: cómo soporte de una información para ser expresada y cómo medio de expresión de esta información.

Esto nos lleva a la cuestión de saber en qué momento y con qué criterios es posible distinguir entre la estructura configuracional de la información secuencial y los mecanismos funcionales que la expresan. Pues si bien en la teoría esta distinción aparece nítida y necesaria, en el estudio de los sistemas prebióticos las cosas se presentan mucho más confusas. Si nos situamos en un momento anterior a la existencia de sistemas ácido nucleico proteína, la cuestión que se plantea entonces es la de determinar si es posible -y en caso afirmativo, en qué condiciones- la existencia en sentido propio de un sistema de símbolos en el marco de un modelo puramente ribozímico. Más en concreto, el problema reside en saber si en un sistema de tipo ribozímico es posible considerar como mutuamente distinguibles los mecanismos de símbolos (secuencias) y de lectura-expresión (plegamientos funcionales) respectivamente.

Trataremos de analizar en las páginas que siguen el proceso de surgimiento, formación y constitución de un sistema de código genético a partir de sistemas moleculares autorreplicativos.

FORMACION DE SISTEMAS AUTORREPLICATIVOS A TRAVES DE COPIA POR MOLDE

Se considera el principio de la evolución prebiótica como el estadio en el cual sistemas químicos macromoleculares se reproducen por copia en un proceso continuo tras el incremento de complejidad que debió ocurrir durante la evolución química. Han sido muchas las vías de aproximación al estudio de este tipo de sistemas. De un lado, basadas en el desarrollo que la escuela de Bruselas hiciera de la termodinámica de procesos irreversibles, (Glansdorf & Prigogine 1977, Nicolis & Prigogine 1977) varios autores (Schuster 1986, Orgel 1986b) han desarrollado sus trabajos evaluando en los sistemas otras características que no fueran el aumento de entropía como, por ejemplo, el grado de complejidad estructural o funcional que el sistema presenta.

Por otro lado, algunos otros trabajos han estado dirigidos a lograr simulaciones de los procesos dinámicos que pudieron llevar a los sistemas primitivos hasta estadios de mayor complejidad que la que presentan

EL CODIGO GENETICO COMO PUNTO CRITICO

los estudiados por la evolución química. Según Langton (1986) estas situaciones de mayor complejidad se presentan en sistemas asimilables a los autómatas celulares cuyo comportamiento es intermedio entre los que vienen representados por puntos fijos, en el espacio fase, y los caóticos.

Huberman y Hogg (1986), por su parte, señalan que los valores de complejidad varían con el valor de orden del sistema, alcanzándose máximos para situaciones intermedias entre las muy ordenadas y las muy desordenadas. Así un aumento de orden sería una necesidad para que el sistema aumentara su grado de complejidad pero nunca constituiría una explicación de dicho proceso. En el fondo de la cuestión está el requisito que ya señalaran Ashby (1958) y Atlan (1979), de la necesidad de un equilibrio entre orden y variedad para que un sistema pueda evolucionar.

En el curso de la evolución química nos interesa el punto en que se llegó a la formación de polímeros, ya que algunas clases de estas moléculas (especialmente los ácidos nucleicos) presentan dos importantes propiedades: estructuración en cadenas de sub-unidades tipo, y capacidad de autorreproducción por copia por molde. Estas propiedades hacen, que los mencionados polímeros sean los sustratos más adecuados para una potencial evolución prebiótica de un sistema que los contenga.

La primera propiedad mencionada es la base de la constitución de un orden configuracional o 'conservatorio' (Eigen & Winkler 1982), y que nosotros denominaremos complejidad formal de la secuencia. Trataremos de explicar a continuación en que sentido es importante la progresiva separación entre la disponibilidad de complejidad formal en las secuencias, y su requerimiento funcional vía plegamiento para la reproducción del polímero.

Debemos distinguir claramente en este proceso lo que es la complejidad topológica (3D) de la secuencia, de lo que es su complejidad formal aunque ambas estén soportadas por un mismo elemento. Pondremos un ejemplo para distinguirlas; supongamos una secuencia de n elementos que pertenecen al conjunto A y otra secuencia de n elementos que pertenecen al conjunto B, donde A y B tienen el mismo tamaño pero donde los elementos de A tienen unas características físicas más homogéneas que los de B, entonces algunas secuencias A y B ($1a, 2a, \dots, na,$ y $1b, 2b, \dots, nb$; por ejemplo) tendrán la misma complejidad formal de

secuencia pero muy distinta complejidad topológica. Y esta distinción nos lleva a otra nueva que nos permitirá discriminar en lo sucesivo entre información en el sentido shannoniano (formal) -que está asociada a la complejidad formal de la secuencia- e información funcional (semántica) -que va a tener su base en la creación de variedad topológica por plegamiento-.

La creación de variedad de formas topológicas, aunque por sí sola no constituye el establecimiento de funciones nuevas, sí es la base para la aparición de esas funciones a través de la potencialidad de las formas para ser procesadas en un estadio ulterior. Pero tampoco basta con la aparición de funciones nuevas a partir de la creación de variedad basada en la complejidad topológica, pues hay que contar con la evaluación para el sistema de estas nuevas funciones.

Veamos ahora que importancia tiene la capacidad de copia por molde en este tipo de sistemas. Según nuestro punto de vista, la capacidad de copia por molde es el mecanismo clave que hace posible el desarrollo del sistema por vías que le lleven a situaciones en las que va a poder superar el umbral crítico de complejidad definido por Von Neumann. El hecho mismo de la copia por molde permite al sistema la disponibilidad de secuencias para posteriores usos funcionales no relacionados directamente con la supervivencia de la unidad que la contiene en el momento de emergencia de la función. Y esto es así porque sí, en un primer momento, tuvieramos un sistema autorreproductivo en cuyas copias la secuencia, por el uso de su expresión funcional, estuviera totalmente requerida para ('ocupada en') realizar la función autorreproductiva, tal sistema no tendría posibilidad de evolucionar. Esta forma de autorreproducción que no se basa en el empleo de instrucciones, es considerada trivial (Langton 1984, Kampis & Csanyi 1987).

En resumen, la capacidad de copia por molde va a permitir la separación en el sistema de dos formas de uso de la secuencia: la primera, como la parte no expresada funcionalmente, y la segunda como la parte expresada. Las sumas de ambas constituye lo que hemos llamado complejidad formal de la secuencia, mientras que la segunda, a la que llamamos 'no disponible', es la expresión tridimensional (3D) de la naturaleza específica de los componentes materiales de la secuencia, y

EL CODIGO GENETICO COMO PUNTO CRITICO

juega un rol funcional fundamental en el ciclo autorreproductivo.

RELACION SECUENCIA-PLEGAMIENTO Y APARICION DE INFORMACION FUNCIONAL

Tras el aumento de complejidad que se seguiría de los procesos expuestos anteriormente vamos a ver cómo surge ahora un nuevo sentido para el término información, que no coincide con el shannoniano (información formal) porque tiene un contenido semántico (Sibatani 1986), aunque éste sea estrictamente autofuncional.

Numerosos autores han señalado la importancia fundamental del papel que desempeñaría la información funcional, en contraposición con la idea simple de complejidad, en la capacidad evolutiva de los sistemas vivos (Wicken 1979, Atlan 1979, Csanyi & Kampis 1985, Kampis & Csanyi 1986). Pero cuando se trata de abordar el problema del origen de esta información funcional en los sistemas vivos las opiniones se muestran divergentes. Para Wicken es la acción de la selección natural la causa que originaría esta información funcional, mientras que para otros autores, como Atlan, ésta hay que situarla en función del cambio de nivel descriptivo efectuado por el observador al estudiar un sistema altamente complejo como son los seres vivos. El problema es que entonces, la creación de información dependería de un observador que a su vez la presupone.

Pero si es verdad como dice Wicken, que la selección natural es la que crea información funcional, la inversa -la existencia de información funcional como condición del mecanismo de selección darwiniano- no es menos cierta. Porque la 'evaluativa' acción del medio ambiente requiere de diversos prerequisites organizativos: competencia, capacidad de creación (incremento) de variedad, autorreproducción fiable, etc. Y en última instancia la 'evaluabilidad' es más consecuencia de la organización interna de los sistemas que de las propiedades predeterminadas del exterior. Así pues, no parece que recurrir a la selección natural constituya una explicación adecuada del origen de la información en los seres vivos.

¿A través de que procesos puede entonces originarse esta información funcional en la evolución prebiótica de los sistemas moleculares? Nuestra opinión es que en la dualidad secuencia-plegamiento del modelo ribozímico existe ya alguna forma de creación de información funcional

y por tanto de incremento de complejidad.

Supongamos un sistema únicamente constituido por ARN. Vamos a distinguir en él dos formas de estas moléculas para el análisis que nos interesa: ARN como secuencias unidimensionales y ARN como estructura terciaria. Si hacemos abstracción de la consideración de las posibles variaciones físicas del medio, generalmente se admite que en el proceso del plegamiento las secuencias unidimensionales poseen un 'significado' intrínseco. En realidad lo que se quiere decir únicamente es que no existe un sistema enzimático de reconocimiento arbitrario (emparejamiento) susceptible de producir formas diferentes de expresión de una misma secuencia por variación de dicha secuencia.

En un sistema dinámico circular, como modelo ribozímico, la relación secuencia-plegamiento supone la congelación ('rate-independent') configuracional de un conjunto de procesos temporales, continuos y dinámicos ('rate-dependent'), y el establecimiento de un homomorfismo. Cuando nos planteamos la cuestión de determinar si es necesaria la misma complejidad descriptiva para la secuencia ribozímica (1D) que para su plegamiento (3D), vemos que el problema consiste en determinar si la función de transición secuencia-plegamiento debe ser implícita o explícitamente considerada. Si la función de transición es implícitamente considerada, la relación secuencia-plegamiento es, a efectos prácticos, de uno-a-uno. Pero si la función de transición es explícitamente requerida, entonces en nuestra descripción del proceso de plegamiento hay un déficit de información. Y dicho de otra manera, la proyección es de uno-a-varios. Pero en el sistema no existe función de transición localizable, ni el proceso del plegamiento se desarrolla ambiguamente. Entonces esto pone de manifiesto algún defecto de nuestro método de descripción del sistema.

El problema consiste en que, a diferencia de lo que ocurre en los modelos formales (en autómatas celulares por ejemplo), la función de transición no es otra cosa que la especificidad material (bases concretas que forman la cadena) del ARN, mientras que en la forma de la secuencia (la cantidad de información contenida) no existe ninguna información interpretativa de la descripción. Por eso, si el observador exterior considera que la secuencia es la autodescripción o registro del plegamiento tendrá que reconocer un incremento de complejidad (o del contenido de información) en la descripción del plegamiento

EL CODIGO GENETICO COMO PUNTO CRITICO

con respecto al contenido de la secuencia (1D), y por tanto, una ambigüedad (para él) en el proceso de transformación. Pero realmente, como hemos dicho, en el sistema en su conjunto no hay ambigüedad.

Pero justamente el problema epistemológico es saber si a todos los niveles del proceso de autorreproducción ribozímico, el tipo de descripción requerido ha de ser el mismo, es decir, si los elementos implícitos o explícitos de la relación secuencia-plegamiento son igualmente requeridos para una adecuada descripción. Y esto es también importante a la hora de elaborar modelos y/o realizar simulaciones por cuanto, en estos casos, no se puede utilizar la materialidad de los componentes, y las relaciones entre los elementos no pueden ser implícitas sino que deben explicitarse en todos los casos.

De una manera más general, puede decirse que lo que un modelo formal gana en universalidad conlleva, inevitablemente, la necesidad de explicitación descriptiva de la función de transición. Este es el precio a pagar por la abstracción de la materialidad que implica la formalización.

Ahora bien, en los sistemas naturales que estudiamos, la función de transición llega a perderse confundiendo con la materialidad de los componentes debido a la poca universalidad de estos (nótese que un sistema ribozímico sólo sería posible con ARN), de tal modo que se llega a la situación de que en la secuencia se contiene a la función de transición de forma implícita. Un sistema ribozímico solo puede construirse con elementos ribozímicos porque son éstos los que proporcionan las relaciones adecuadas (función de transición implícita). A su vez, lo que hace posible esta construcción, además de la información secuencial, es la capacidad de copia por molde de los ácidos nucleicos en su duplicación. Este proceso (copia por molde) posibilita la inclusión de la función de transición en la materialidad de la secuencia. Si no existiera copia por molde, al intentar construir una copia de un elemento del sistema (ARN) en su forma plegada (funcional) el 'aparato' de construcción debería 'leer' (intentar copiar) el plegamiento, y en ese caso la nueva construcción contendría en todos los casos el resultado de aplicar la función de transición a los elementos del sistema, perdiéndose así la dualidad secuencia-plegamiento (o, registro-hecho registrado). Con la copia por molde, sin embargo, ocurre un traspaso de información sin interpretación de esa información, esto es, sin utilización de la

función de transición, cuando los nucleótidos polimerizan, uno a uno, siguiendo la guía de la cadena molde, sin explicitar que forma tendrá el conjunto final, pero mandando en cada caso una parte de la función de transición a través de la materialidad de un nucleótido concreto, de tal modo que, al final, la función de transición será una composición de las contribuciones particulares de los nucleótidos.

Otra cosa es lo que ocurre en los sistemas formales donde la función de transición debe ser incluida como 'software' (vía explicitación formal) por lo que el efecto 'abreviatura' (la parte que contiene el todo) no se puede lograr. En definitiva, no es por casualidad que el punto donde debe surgir la relación símbolo-materia no se pueda modelizar. En los sistemas ribozímicos, al ser la configuración secuencial una abreviatura, la acción de procesamiento del plegamiento sobre ellas hace que, una vez expresada la nueva secuencia, aparezcan nuevas funcionalidades, lo que da como resultado una imagen de incremento global de la complejidad a través de la evolución temporal del sistema. Y es en el proceso del auto-procesamiento donde se pone de manifiesto, a través de la continua interacción mutua entre secuencia y plegamiento, esta paradoja para el observador. Porque la posibilidad de la actuación dinámica sobre autorregistros (= procesamiento) que engendra nuevas funcionalidades, es una consecuencia de que una misma entidad, el ARN, aparezca en las dos formas mencionadas. Y en el fondo, cuando consideramos a las secuencias como registros, no hacemos más que reconocer esta situación de radical e irresoluble ambigüedad de nuestra forma de descripción del sistema.

Así pues las secuencias unidimensionales en el ribozima, debido a su implícita expresión o traducción dinámico-funcional a través del plegamiento, prefigurarían un sistema de información genética. Ciertamente, este sistema no sería plenamente 'lingüístico', porque en un primer momento la presencia de secuencias semánticas referidas a las propiedades dinámicas de la propia molécula (esto es, el conjunto tridimensional) no presentaría todos los requisitos teóricamente necesarios para ser definida como sustrato de símbolos. Faltaría principalmente la arbitrariedad en la relación establecida entre las secuencias y sus 'representaciones' dinámicas. No obstante, existiría ya un homomorfismo entre las estructuras discretas de las secuencias y los correspondientes procesos (continuos, dinámicos) que determina. En otras palabras la

EL CODIGO GENETICO COMO PUNTO CRITICO

especificación de la secuencia (la estructura primaria) constituiría una abreviatura de la molécula tridimensional plegada con sus potencialidades enzimáticas específicas.

Queremos hacer notar aquí que, a diferencia de lo habitual, la abreviatura o simplificación no es un efecto (un sistema dinámico complicado que es sencillo en cuanto a las salidas que proporciona, pudiendo así el observador desentenderse de los detalles (Varela 1979, 1981)), sino que es una causa de la organización interna misma. Las abreviaturas son causa y condición de la posibilidad de la complejidad pormenorizada de la dinámica del conjunto del sistema.

De otra parte, la relativa estabilidad del orden configuracional de la secuencia es, a su vez, fundamento (junto con las condiciones anteriormente mencionadas) de que podamos considerarlo como un requisito o anotación ('rate-independent') de sus propias funcionalidades dinámicas ('rate-dependent').

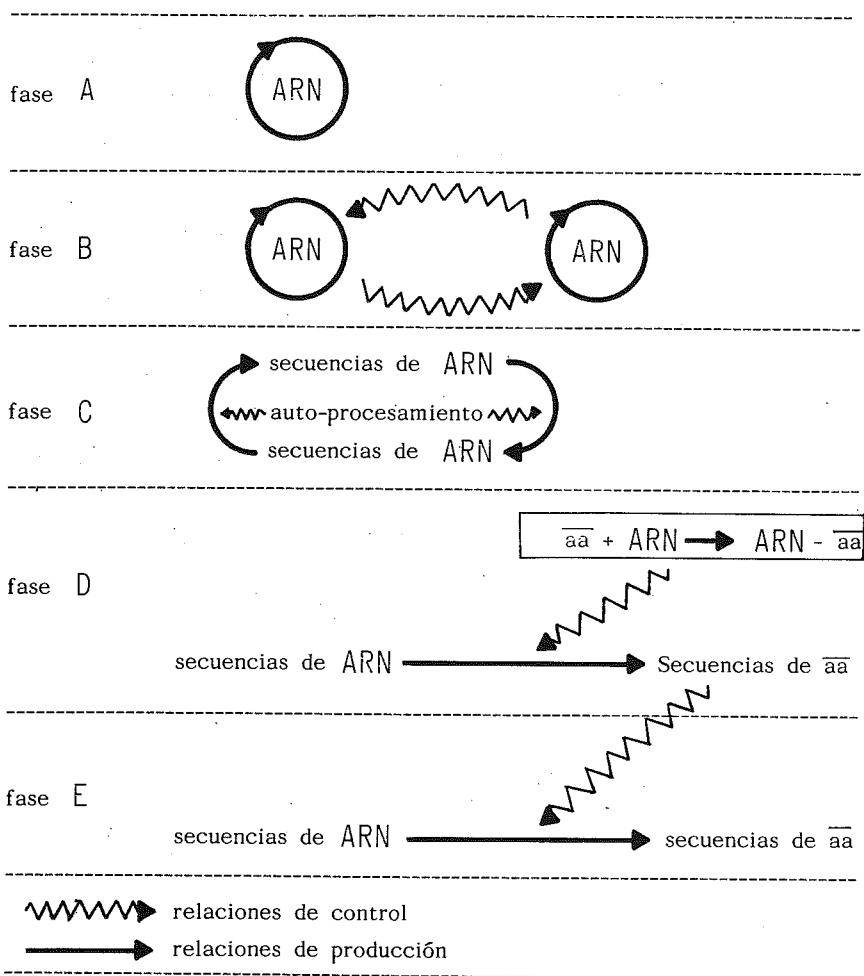
Además, el establecimiento de la dualidad secuencia-plegamiento permite un funcionamiento en dos tiempos o velocidades en el sistema; uno lento relacionado con las transformaciones configuracionales, discretas y secuenciales, que es susceptible de producir consecuencias dinámicas nuevas a otros niveles del sistema, y otro basado en las actividades funcionales (enzimáticas) de las estructuras tridimensionales estereoespecíficas. El interés de esta doble dinámica reside en que al ser unas estructuras 'registros' de otras se da origen a dos niveles mutuamente interrelacionados en el sistema.

Sin embargo, lo más importante para nuestro propósito es que una vez que el sistema se ha dotado de este mecanismo de dualidad configuración (1D)/estructura (3D) funcional, se abren nuevas posibilidades de incremento de complejidad. en efecto, un sistema puede aumentar su variedad interna sin disminuir su redundancia a condición de actuar sobre registros de sí mismo.

Pero como hemos dicho la relación secuencia-plegamiento no puede constituir, en sí misma, una verdadera forma de representación, sino únicamente una abreviatura en el estado configuracional de la forma compleja 3D dinámica. Por eso, el código genético tuvo que aparecer en un momento dado para aumentar la variabilidad del sistema sin disminuir su tamaño, mediante los mecanismos de emparejamiento arbitrario en la traducción de ácidos nucleicos a proteínas.

TRADUCCION COMO CONSTITUCION DE EMPAREJAMIENTOS ARBITRARIOS

Expondremos a continuación un esquema hipotético que no pretende ser un modelo general para el origen del código genético ni mucho menos incluye todos los problemas relativos al origen de un sistema viviente. Únicamente intentamos remarcar aquellas fases que consideramos puntos importantes.



EL CODIGO GENETICO COMO PUNTO CRITICO

Dentro de este esquema general (figura) vamos a examinar las posibilidades de evolución de un sistema viviente sin código. En un sistema de este tipo la capacidad evolutiva se vería claramente limitada, y ello fundamentalmente por la imposibilidad de auto-creación de variaciones que permite el establecimiento de un código arbitrario. En efecto, sin la constitución de un código el sistema no puede desarrollar una dualidad entre genotipo y fenotipo, lo que a su vez implica la imposibilidad de una selección darwiniana en un sentido estricto, ya que la evaluación selectiva de variantes fenotípicas no podría ser autónoma de los constreñimientos selectivos genotípicos.

Esta situación viene representada por las fases:

- A) Estadio en el cual los ácidos nucleicos se duplican por copia. Este estadio evolutivo del sistema sería el resultado último de la evolución química, y la consecuencia de las actividades polimérica y de copia por molde que presentan los ácidos nucleicos.
- B) Una organización de tipo hipercíclico (Schuster 1986) aseguraría la continua conservación de las secuencias durante la replicación de los ácidos nucleicos -mantendría el grado de redundancia en el contenido informativo de las secuencias-.

Como resultado de las fases anteriores el sistema, después de haber superado la primera crisis informativa, aparece situado en un punto fijo (proporcionado por un atractor dinámico como consecuencia de la organización hipercíclica) del que sólo saldrá si ocurre un incremento de variedad dentro del sistema. El incremento de variedad interna que viene posibilitado por la creación de un código arbitrario es en realidad el resultado implícito de las diferentes formas de descripción o registro que pueden referirse a un mismo objeto. Ahora bien, para que este fenómeno -distintas formas de representación de un mismo objeto- sea posible, se requiere además de las configuraciones 'rate-independent' de registro, la acción complementaria de un sistema de fijación de los emparejamientos. Puede decirse, por tanto, que la arbitrariedad del código es una consecuencia derivada de las potencialidades universales (= gratuitas) de emparejamiento que pueden realizar los enzimas. El problema es que la construcción (síntesis) de estos enzimas depende a su vez de la información contenida en la secuencia de nucleótidos. Parece, por tanto, bastante lógico pensar que, de ser posible alguna

función de emparejamiento (similar a las de las aminoacil tRNA sintetasas) entre secuencias de ácidos nucleicos y secuencias polipeptídicas realizable por ribozimas, ésta sería la vía de transición hacia la forma desarrollada del código.

En las fases que siguen se propone una acción enzimática del ribozima que conduciría a la aparición de la primera diferencia entre secuencia y función;

- C) Representa una situación en la que los ARN interaccionarían en una modalidad de auto-procesamiento que aunque no está definitivamente caracterizado, si está claramente definido para buen número de casos.
- D) Esta fase que significaría la acción del ARN como adaptador tiene el problema de la no probada -hasta ahora- capacidad enzimática de los ARN sobre sustratos distintos de sí mismos. Pero no parecen existir límites, en principio, absolutos para la capacidad enzimática de los ARN. Existen, desde luego, razones estructurales que limitan seriamente su versatilidad enzimática, en comparación con las proteínas (Cech & Bass 1986). Ahora bien, parece verosímil que puedan encontrarse ARN con sitios de unión para sustratos distintos de los nucleótidos (Orgel 1986a).

El siguiente paso en el orden de evolución debió venir representado por la que aparece como última fase en nuestro esquema;

- E) Durante esta fase debió ocurrir la sustitución de los ácidos nucleicos (concretamente al ARN) en las tareas enzimáticas por proteínas, conformándose definitivamente un sistema circularmente cerrado de dos niveles: dinámico y lingüístico, mutuamente complementarios (y que Pattee define como 'semantic closure' (1982, 1986). En este tipo de sistemas ya podemos hablar con propiedad de información funcional universal.

Durante esta fase consideramos que una condición necesaria (aunque quizás no suficiente) es la que Küppers (1983) menciona en la siguiente forma:

"Si la aparición de un código en un sistema va a ser estabilizada, entonces los elementos transportadores de la información deberán incluir algo que después de la traducción promueva el uso del

EL CODIGO GENETICO COMO PUNTO CRITICO

mismo código. (...) Una traducción reproducible solo es posible cuando se prefiere la aplicación del mismo código utilizado en la vuelta anterior de traducción".

Efectivamente, supongamos un modelo como el presentado en el cual por el momento, solo poseemos ácidos nucleicos (ARN por ejemplo) e imaginemos ahora que los ARN empiezan a funcionar como adaptadores entre el ARN y los aminoácidos. Bien, en principio cualquier proteína se puede formar con este rudimentario aparato de traducción, pero ¿cuando una proteína sustituirá al ARN en sus labores de adaptación? Sólo cuando la proteína sea capaz de cumplir la función que hasta el momento desarrollaba el ARN. De otra manera, si esto no ocurriera así, el proceso de traducción destruiría la información acumulada en el sistema (en cuanto a la traducción destruiría al sistema, aunque éste pudiera perpetuarse en su forma replicativa) y sólo cuando dos vueltas consecutivas (la primera con ARN y la segunda con proteínas) el proceso fuera el mismo, la sustitución quedaría fijada. En otras palabras, buscamos sistemas donde existan secuencias originadas al azar cuya funcionalidad de emparejamiento proporcione los mismos emparejamientos que le han dado origen al azar.

Una vez alcanzada esta condición -autorreproducibilidad- podemos considerar que la naturaleza ha constituido un sistema dinámico-lingüístico en sentido estricto, (por muy primitivo que pueda parecernos), ya que el sistema de traducción en proteínas permite una universalización funcional codificable en la secuencia de bases. A partir de aquí, la maquinaria enzimática bajo control de la información contenida en la secuencia de bases (de los ácidos nucleicos) contribuye a reforzar el aumento de la complejidad formal de las secuencias y la fidelidad de su reproducción, a la vez que este proceso refuerza la creación de nuevas y mejores funcionalidades enzimáticas.

De esta forma, el organismo viviente primitivo está en condiciones de alcanzar ulteriores procesos de complejización. Sin embargo, los rasgos esenciales de la autoorganización biológica están ya en juego.

CONCLUSIONES

A la luz de lo expuesto anteriormente nos parece oportuno tratar de redefinir el significado de varios conceptos importantes, por cuanto

no solo se trata de nociones fundamentales para la Biología Teórica, sino que lo son también por su alcance epistemológico.

El punto central de nuestras conclusiones es el siguiente: en el ámbito biológico, el código genético establece la frontera entre un sistema capaz de evolucionar indefinidamente, y los sistemas ribozímicos que le precedieron. Definimos código como el conjunto de transformaciones que pueden sufrir distintos sistemas de símbolos sin dejar de corresponder a una misma representación.

Las transformaciones que solo afectan a las relaciones de los símbolos entre sí únicamente existen como abstracción, pero hacen posible la aparición de lenguajes formales. Sin embargo, la relación representacional, que constituye la esencia del problema símbolo/materia, y que está en la raíz del origen de los sistemas vivientes, no puede ser simulada adecuadamente en un sistema formal puramente abstracto. La vida se basa en la máxima utilización de las condiciones específicas de su 'hardware' en funciones implícitas de 'software', a costa de perder universalidad.

El último y quizás el más difícil de definir de los conceptos empleados aquí es el de representación. Hemos tratado de mostrar cómo su génesis se confunde prácticamente con la de los sistemas evolutivos, es decir, con la vida. Cómo surge embrionariamente -en forma exclusivamente autorreferencial- en los sistemas ribozímicos, y cómo, a través de la constitución del código, alcanza su forma plena o de universalidad semántica. En realidad, la mejor definición que proponemos es genética, mostrando las tres fases sucesivas por las que pasa para llegar a su plena constitución: a) como registros autorreferenciales; b) como configuración de formas de emparejamiento físicamente arbitrarias; y c) como traducción a través de un código. Este proceso comporta dos clases de condiciones:

- 1) El establecimiento de procesos dinámicos circulares (ciclos autocatalíticos, hiperciclos, ...).
- 2) El establecimiento de correspondencias (emparejamientos homomórficos): secuencia de bases → secuencia de aminoácidos → proteínas tridimensionales → funcionalidades dinámicas.

La fijación de las correspondencias se debe al carácter y la traduc-

EL CODIGO GENETICO COMO PUNTO CRITICO

ción son solo posibles mediante la acción de enzimas, cuya especificidad funcional es a su vez secuencialmente dependiente de las propias cadenas que reconocen y traducen. En su origen, sin embargo, este proceso no fue más que un accidente congelado, o mejor, la selección entre múltiples formas de accidente congelado, de las que se han eliminado aquellas menos aptas (Hofstadter 1985), puesto que físico-químicamente hablando, las relaciones de emparejamiento que se establecieron son en lo esencial arbitrarias. Ciertamente se puede decir que una estructura puede representar o codificar a otra si las anteriores circunstancias están presentes. En este sentido, la relación secuencia/plegamiento de un sistema puramente ribozímico constituiría, de alguna manera, una forma de autorrepresentación. Pero también hemos intentado exponer a lo largo de este trabajo cuales son las limitaciones del plegamiento como interpretación-expresión de las secuencias en un sistema sin procesos de traducción de ácidos nucleicos en proteínas. Por eso, en un sentido propio, la representación solo cobra su verdadera dimensión en la medida en que se articulan complementariamente las secuencias informacionales con los mecanismos enzimáticos de reconocimiento y traducción en un sistema de organización circular o cerrada.

Pero en última instancia la posibilidad del desdoblamiento de un sistema dinámico, a través de la creación de un homomorfismo que permita formas de autodescripción, solo se consigue utilizando la propia especificidad material en el doble cometido de soporte configuracional (homomorfismo de los procesos dinámicos vía plegamiento) y como expresión de su naturaleza físico-química concreta, construyendo las leyes físicas en un proceso que interpreta una estructura unidimensional, en otra tridimensional. Este proceso puede parecer trivial en la relación secuencia/plegamiento de una misma molécula, pero no lo es cuando se convierte, por la intervención de un código, en la traducción en principio arbitraria de la secuencia de una molécula en el plegamiento de otra diferente.

REFERENCIAS

- ASHBY, W.R., (1958): "The requisite variety and its implications for the control of complex systems", en *Cybernetica* 1, pp. 83-99.
- ATLAN, H. (1979): *Entre le cristal et la fumée*, Paris, Seuil.
- CECH, T. & BASS, B. (1986): "Biological catalysis by RNA" en *Ann. Rev. Biochem.*, 55, pp. 599-629.
- CRICK, F., [et alia] (1976): "A Speculation on the Origin on the Protein Synthesis", en *Origins of life*, 7, pp. 389.
- CSANYI, V. & KAMPIS, G. (1985): "Autogenesis: The Evolution of Replicative Systems", en *J. Theor. Biol.*, 114, pp. 303-321.
- EIGEN, M. & SCHUSTER, P. (1979): *The Hipercycle: A Principle of Natural Self-Organization*, Berlin, Springer Verlag.
- EIGEN, M. & WINKLER, R. (1982): *Laws of the Game*, London, Penguin Books.
- GLANSDORF, P. & PRIGOGINE, I. (1977): *Thermodynamic Theory of Structure Stability and Fluctuations*, New York, Wiley-Interscience.
- HOFSTADTER, D. (1985): *Metamaical Themes*, New York, Basic Books.
- HUBERMAN, B. & HOGG, T. (1986): "Complexity and Adaptation", en *Physica*, 22D, pp. 376-384.
- KAMPIS, G. & CSANYI, V. (1986): "Replication is informational", versión preliminar multicopiada
- KAMPIS, G. & CSANYI, V. (1987): "Replication in abstract and natural systems", en *Biosystems*, 20, pp. 143-152.
- KÜPPERS, B. (1983): *Molecular Theory of Evolution*, Berlín, Springer Verlag.

EL CODIGO GENETICO COMO PUNTO CRITICO

- LANGTON, C. (1984): "Self-reproduction in Cellular Automata", en *Physica* 10D, pp. 135-144.
- LANGTON, C. (1986): "Studying Artificial Life with Cellular Automata", en *Physica* 22D, pp. 120-149.
- VON NEUMANN, J. (1966): *Theory of Self-Reproducing Automata*, A. Burks (ed.) University of Illinois Press.
- NICOLIS, G. & PRIGOGINE, I. (1977): *Self-organization in Non equilibrium Systems*, New York, Wiley-Interscience.
- ORGEL, L. (1986a): "RNA Catalysis and the Origins of Life", en *J. Theor. Biol.* 123, pp. 127-149.
- ORGEL, L. (1986b): *Molecular Replication and the Origins of Life*. J. de Boer et alia (eds.), *The lesson of Quantum Theory*, New York, Elsevier Science Publishers.
- PATTEE, H. (1977): "Dynamic and Linguistic modes of complex Systems", en *Int. J. General Systems*, 3, pp. 259-266.
- PATTEE, H. (1982): "Cell Psychology: an evolutionary approach to the symbol-matter problem", en *Cognition and Brain Theory*, 5, pp. 325-341.
- PATTEE, H. (1986): *Universal Principles of Measurement and Language Functions in the evolving Systems*. In L. Casti (ed.), *Complexity - language and life*, Berlin, Springer Verlag.
- SCHUSTER, P. (1986): *The interface between Chemistry and Biology. Laws determining Regularities in the Early Evolution*, en G. Pifat-Mizljak (ed.), *Supramolecular Structure and Function*, Berlin, Springer Verlag.
- SIBATANI, A. (1986): *Information as a misnomer in Biology and a warning to contemporary*. Comunicación presentada a "Today symposium"

Alvaro MORENO y Julio FERNANDEZ

sium Information and its functions", 6-9 de Octubre de 1986,
Tokyo University.

VARELA, F. (1979): *Principles of Biological Autonomy*, New York, Elsevier, North Holland.

VARELA, F. (1981): *Describing the Logic of the Living*, en M. Zeleny (ed.), *Autopoiesis*, New York, Elsevier North-Holland.

WICKEN, J. (1979): "The Generation of Complexity in Evolution: A Thermodynamic and Information-Theoretical Discussion", en *J. Theor. Biol.* 77, pp. 349-365.

Depto. de Lógica y Filosofía de la Ciencia
Universidad del País Vasco/E.H.U.
San Sebastián

Becario del Plan de Formación del Personal Investigador
del Ministerio de Educación y Ciencia
Universidad del País Vasco/E.H.U.
San Sebastián