

De las redes de reacciones al surgimiento del metabolismo. Modelamiento del origen de la vida mediante la Teoría de la Organización Química

From reaction networks to the emergence of metabolism. Chemical Organization Theory applied to modeling the origin of life

Diego Becerra¹

¹ *Facultad de Ciencias, Universidad de Valparaíso, Valparaíso, Chile*

Fecha de recepción del manuscrito: 02/12/2022

Fecha de aceptación del manuscrito: 12/12/2022

Fecha de publicación: 30/12/2022

Resumen— Los sistemas vivos son casos de estructuras disipativas lejos del equilibrio termodinámico. Diversas teorías acerca de su origen han sido propuestas, pero son difíciles de evaluar empíricamente. Una estrategia prometedora consiste en buscar el núcleo biosintético mínimo común a los organismos unicelulares más ancestrales (e.g. bacterias acetógenas y arqueas metanógenas), y examinar si las especies químicas de dicha red están presentes en escenarios candidatos a origen de la vida; y si las reacciones de la red son termodinámicamente favorables, autocatalíticas, dinámicamente estables, etc. La Teoría de la Organización Química [COT] es un marco matemático que nos permite representar y computar el conjunto de todas las configuraciones observables posibles dada una red de reacciones; y puede utilizarse para testear hipótesis *in silico* acerca de auto-organización molecular previa a la primera red de reacciones metabólica dentro de un sistema vivo. En este artículo se realiza una breve introducción a la teoría, luego, se analiza una red biosintética termodinámicamente favorable mediante herramientas de COT. Las organizaciones encontradas contienen las especies afluentes del sistema, además de formiato, clústeres FeS, bicarbonato, CO, y H^+ . Los Respiraderos Hidrotermales Alcalinos, candidatos al origen de la vida, contienen el conjunto de especies encontradas en las organizaciones químicas obtenidas en este trabajo. Las propiedades estequiométricas y dinámicas de estas sub-redes ameritan ulterior exploración.

Palabras clave— Origen del metabolismo, Teoría de la Organización Química, Origen de la vida, Respiraderos Hidrotermales Alcalinos, Redes bioquímicas

Abstract— Living systems are instances of far-from-equilibrium dissipative structures. Several theories about their origin have been proposed, but they are hard to test empirically. One promising strategy is to search the minimal biosynthetic core common to the most ancestral unicellular organisms (e.g. acetogenic bacteria and methanogenic archaea), and to test whether the chemical species from the core are present in candidate scenarios for the origin of life, and whether the reactions are thermodynamically favorable, autocatalytic, dynamically stable, etc. Chemical Organization Theory, a mathematical framework which allow us to represent and compute the set of all possible observable configurations from a given reaction network, can be used to test hypotheses *in silico* about molecular self-organization prior to the first metabolic network within a living system. In order to do this, the theory will be introduced, and a thermodynamically favorable biosynthetic network will be analyzed with Chemical Organization Theory' tools. Organizations containing the inflow species from the biosynthetic core (H_2 , CO_2 , NH_3 , P_i , H_2O , and H_2S) plus formate, FeS clusters, bicarbonate, CO, and H^+ were found. Hydrothermal Alkaline Vents, a strong candidate for the origin of life, contains the set of species found in the chemical organizations obtained from this work. The stoichiometric and dynamical properties of those sub-networks merit further exploration.

Keywords— Origin of metabolism, Chemical Organization Theory, Origin of life, Hydrothermal Alkaline Vents, Biochemical networks

INTRODUCCIÓN

Los sistemas vivientes son casos de estructuras disipativas lejos del equilibrio termodinámico, es decir, una colección de componentes en interacción continua cuyo mantenimiento requiere un intercambio constante de energía y materia con el medio circundante (Prigogine y Nicolis, 1971). Los sistemas disipativos no requieren preservar un conjunto de componentes para mantener su identidad, sino, generar una dinámica cohesiva que se mantenga en el tiempo, de manera que los componentes microscópicos son constantemente atraídos, sintetizados, reemplazados, desechados, y/o reparados. En otras palabras, dichos componentes se auto-organizan para formar temporalmente un sistema fluctuante cuya duración los excede. Ejemplos de estructuras disipativas son los huracanes, el fuego, reacciones químicas oscilatorias como la reacción Belousov-Zhabotinskii, y los seres vivos (Barandarian y Moreno 2008; Hudson y Mankin 1981).

Dado un flujo adecuado de energía y materia –y ciertas condiciones de frontera específicas– los componentes microscópicos de una estructura disipativa configuran un orden macroscópico relativamente estable, que produce (y depende de) un borde o interfaz; que tiende al aumento de complejidad y a la diversificación de sus componentes, a la vez que produce una transición espontánea desde macro-estados ricos en microestados hacia macro-estados pobres en microestados, indicando una reducción en su entropía estadística (Pulselli et al., 2009). No obstante, los sistemas vivientes –y en particular, las células, en tanto sistemas autopoieticos– presentan características distintivas respecto a otras estructuras disipativas: sea porque, a diferencia de los tornados o el fuego, son (i) redes de procesos físico-químicos, (ii) capaces de producir un subconjunto de la red de componentes moleculares que los constituyen, y (iii) este subconjunto de componentes, mediante relaciones con otros componentes de la red y componentes del entorno, genera las condiciones necesarias que mantienen la proximidad física de los componentes, formando una unidad espacial discreta en un intervalo de tiempo (Maturana y Varela 1973; Razeto-Barry 2012).

De ese modo, las células son autocatalíticas (i.e. se requieren a sí mismas para persistir y reproducirse) (Steel et al. 2019; Xavier et al. 2020), lo cual nos permite hipotetizar un continuo entre la geoquímica (o química prebiótica) y la bioquímica de los primeros organismos vivientes que surgieron hace 4.28–3.77 mil millones de años atrás (Dodd et al., 2017). Las redes de reacciones autocatalíticas espontáneas podrían haber dado paso, paulatinamente, a los primeros ensayos de organismos vivientes. Por ello, resulta tremendamente relevante estudiar su viabilidad en tres dimensiones: (i) relacional (i.e. las moléculas que componen las redes de reacciones, y sus interacciones relevantes), (ii) estequiométrica (propiedades cuantitativas de las reacciones), y (iii) dinámica (evolución temporal de la red de reacciones) (Veloz y Razeto-Barry, 2017). Para ello, se explorará la viabilidad de modelar las posibles redes de reacciones termodinámicamente favorables al inicio de la vida mediante la Teoría de Organización Química –en adelante COT, por sus siglas en inglés– (Dittrich y Di Fenizio, 2007).

REDES AUTOCATALÍTICAS

Es posible representar un sistema de reacción catalítico (CRS) como una tripleta $\langle X, R, C \rangle$ donde, X es el conjunto de todas las especies moleculares [*molecule types*] que participan en el CRS, tal que el conjunto alimento F –aquél que contiene aquellas especies de moléculas ampliamente disponibles, provenientes del ambiente circundante– es un subconjunto de X . Luego, $R : \{r_1, r_2, \dots, r_n\}$ es el conjunto de todas las reacciones químicas del CRS; y C es un operador de asignación que designa qué especies moleculares catalizan cuáles reacciones. Dada una fuente de entrada de materia (el conjunto alimento F) proporcionada por el medio, una red de reacciones $R : \{r_1, r_2, \dots, r_n\}$ es colectivamente autocatalítica y auto-mantenida [*self-sustaining*] si:

- (i) Es reflexivamente autocatalítica (RA): Cada reacción $r_i \in R$ es catalizada por al menos una especie molecular que es o bien un producto de R o está presente en F .
- (ii) Es F-generada: Cada reactivo involucrado en una reacción $r_i \in R$, puede ser producido por el conjunto alimento F usando una serie de reacciones del conjunto R .

Si una red de reacciones R cumple con (i) y (ii), se le denomina RAF (mantendremos la abreviación en inglés de ‘*reflexively autocatalytic and food-generated*’). Un RAF puede ser representado matemáticamente como un grafo bipartito dirigido (ver Fig. 1, izquierda), o bien, como una lista de reacciones de la forma $r : A \xrightarrow{c_1, c_2, \dots} B$. Dado cualquier subconjunto R' de R , definimos el conjunto cerradura [*closure*] de R' en nuestro CRS, denotado \bar{R} . Aquél es el único subconjunto mínimo de R que contiene a R' y satisface la siguiente propiedad: si una reacción r de R posee todos sus reactivos y al menos un catalizador presente en F , o como producto de una reacción de \bar{R} ; entonces r está en \bar{R} . Finalmente, un RAF es un RAF cerrado, si es igual a su conjunto cerradura (Steel et al., 2019).

Se ha propuesto utilizar RAFs como modelo del origen del metabolismo, tanto mediante sistemas químicos construidos en el laboratorio (Semenov et al. 2016; Miras et al. 2020), como en análisis computacional de redes metabólicas presentes en arqueas metanogénicas (Steel et al., 2020) o en arqueas y bacterias autótrofas que reduzcan CO_2 geoquímico mediante H_2 , ambas especies químicas disponibles abióticamente en respiraderos hidrotermales (McDermott et al. 2015; Preiner et al. 2020) apuntando a encontrar la red metabólica mínima presente en el último ancestro común universal (en adelante LUCA, por sus siglas en inglés: ‘*Last Universal Common Ancestor*’) (Xavier et al., 2020).

Mientras que el metabolismo celular moderno está basado en enzimas, $> 60\%$ de los mecanismos enzimáticos conocidos involucran al menos un cofactor (Ribeiro et al., 2018) y 40% de las proteínas cristalizadas tienen un metal asociado involucrado en su función (Guengerich, 2016). Esto puede indicar que, antes de que hubiese mecanismos de traducción

proteica, los cofactores (metálicos o no) habrían catalizado reacciones en ausencia de enzimas, de ser el caso, habría algunas redes químicas dependientes de cofactores en el metabolismo moderno que (i) constituyan RAFs, y (ii) provengan de una época previa a la codificación y traducción genética de proteínas (Xavier et al., 2020).

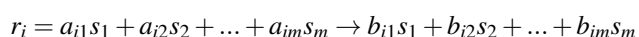
Utilizando una bacteria acetogénica (*Moorella thermoacetica*) y una arquea metanogénica (*Methanococcus maripaludis*), Xavier et al. (2020) encontraron una red primordial de RAFs que genera aminoácidos, nucleósidos y acetil-CoA desde un conjunto de moléculas ‘alimento’ sencillas. Para ello, se eliminó toda reacción química específica de eucariotas en la base de datos KEGG (Enciclopedia de Kyoto de Genes y Genomas, por sus siglas en inglés), y luego se examinaron los metabolismos anaeróbicos, debido a que el O_2 atmosférico es producto de la fotosíntesis de cianobacterias, cuya existencia se remonta a no más de 2400 millones de años atrás (Fischer et al., 2016). De ese modo, quedaron con una red compuesta por 5723 metabolitos y 5994 reacciones. Se definió un conjunto alimento inicial conteniendo $H_2O, H_2, H^+, CO_2, CO, PO_4^{3-}, SO_4^{2-}, HCO_3^-, P_2O_7^{4-}, S, H_2S, NH_3, N_2$, todos los metales, clusters FeS, un aceptor y un donante generalistas. Dicha red generó un maxRAF (el RAF compuesto por la unión de todos los RAFs en un CRS) de 8 reacciones. La adición de formato, metanol, acetato y piruvato duplicó el número de reacciones del maxRAF (16), y la adición secuencial de los 8 cofactores orgánicos más frecuentes identificados en las proteínas de LUCA (ATP, NAD^+ , FMN, SAM, CoA, GTP, molibdopterina, ferredoxina reducida) expandió el maxRAF a 914 reacciones. En vez de ello, agregar sólo las 5 especies moleculares con mayor impacto sobre el tamaño del maxRAF de la red anaeróbica completa (NAD^+ , piridoxal – 5 – fosfato, FMN, H_4F , tiamina difosfato), permite generar un maxRAF de 1248 reacciones (cerca de las 1335 reacciones del maxRAF con todos los cofactores de la red anaeróbica). Lo anterior permite dar cuenta del 25% de la red metabólica anaeróbica ancestral, mientras que agregar catalizadores peptídicos genéricos al conjunto F aumenta el tamaño del maxRAF hasta el 93% de la red anaeróbica completa (Xavier et al., 2020).

TEORÍA DE ORGANIZACIÓN QUÍMICA

Alternativamente a las redes autocatalíticas, es posible representar un CRS –o cualquier sistema dinámico complejo– mediante la Teoría de Organización Química (en adelante COT, por sus siglas en inglés), la cual permite computar el conjunto de todos los subsistemas observables posibles de una red de reacciones (Dittrich y Di Fenizio 2007; Hordijk et al. 2018; Veloz et al. 2022).

Descripción estequiométrica

Una red de reacciones $\langle M, R \rangle$ consiste en un conjunto finito $M = \{s_1, s_2, \dots, s_m\}$ de m especies químicas reaccionando unas con otras de acuerdo a un conjunto finito $R = \{r_1, r_2, \dots, r_n\}$ de n reacciones. Una reacción r_i es representada mediante



donde s_j es la j -ésima especie química (e.g. tipo de molécula) que participa de la reacción y $a_{ij} > 0, b_{ij} > 0$ especifican la estequiometría de la reacción para reactivos y productos, respectivamente. De ese modo, mediante la reacción r_i , la colección de las especies químicas a la izquierda del operador flecha \rightarrow , es transformada en la colección de las especies químicas de la derecha.

COT se enfoca en el estudio de las propiedades de los subsistemas X_k de M ($X_k \subseteq M$): para cada conjunto X , hay un conjunto maximal único de reacciones $R_X \subseteq R$ definido como el conjunto de todas las reacciones cuyos reactivos están en X , por lo tanto, cada conjunto X induce una sub-red $\langle X, R_X \rangle$.

Se le llama conjunto cerrado de especies a cada sub-red $\langle X, R_X \rangle$ que no produce especies fuera de X . En otras palabras, dada las interacciones R_X dentro del sistema, no puede aparecer un nuevo objeto $s_{m+1} \notin X$ (es relevante notar que el concepto de ‘cerradura’ o ‘clausura’ tiene un significado distinto en teoría de redes autocatalíticas y en sistemas (M, R) de Rosen; pero las diferencias no serán exploradas en este ensayo, para una exploración de dicho concepto en distintos marcos teóricos, revisar (Contreras et al. 2011; Hordijk et al. 2018; Reynaert 2013). Un conjunto cerrado puede estar compuesto por varios conjuntos cerrados desconectados más pequeños, por ejemplo, en $X = \{a, b, c, d\}, R = \{a \rightarrow b, b \rightarrow a, c \rightarrow d, d \rightarrow c\}$, encontramos que $\{a, b\}$ y $\{c, d\}$ están conectados y son cerrados, pero no hay reacción alguna entre esos dos sub-conjuntos. Otra propiedad relevante es la semi auto-mantenición: X es un conjunto semi auto-mantenido si y sólo si para cada reactivo $s \in X$ de una reacción $r \in R_X$, existe otra reacción $\bar{r} \in R_X$ tal que s es un producto de \bar{r} . En otras palabras, cada objeto consumido dentro del conjunto tiene al menos una forma de ser generado dentro del conjunto. Pero dicha relación puede no estar balanceada cuantitativamente, como en el caso de $R = \{a \rightarrow b, 2b \rightarrow a\}$, que llevaría a la extinción del proceso al agotarse los reactivos. Si un conjunto de especies es cerrado y semi auto-mantenido, se le denomina semi-organización (Dittrich y Di Fenizio, 2007).

Descripción dinámica

La red de reacciones puede modelarse en tiempo discreto (medido en pasos), o en tiempo continuo (medido en alguna unidad de tiempo pertinente). Para capturar aquello, el vector que especifica la dinámica de la red de reacciones se denomina ‘proceso’ o ‘vector de flujo’ v ; el cual está determinado por el estado ‘inicial’ x del sistema –que corresponde a la disponibilidad inicial de reactivos y catalizadores–; y por su dinámica, la cual consiste en la tasa de ocurrencia de las distintas reacciones. Para construir el vector de flujo $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n \geq 0$, tenemos que v_i especifica la tasa de reacción r_i en el intervalo temporal a considerar, donde r_i es la i -ésima reacción del conjunto de reacciones R , compuesto por n elementos.

Decimos que un vector de flujo v puede aplicarse a un conjunto de m especies químicas X si y sólo si $v_i > 0$ implica que para cada una de las n reacciones r del conjunto R ,

tenemos que $r_i \in R_X$. Luego, en la red de reacciones, los números a_{ij} pueden representarse como una matriz $L_{m \times n}$ de los coeficientes estequiométricos de las especies consumidas (a la izquierda de cada reacción), y los números b_{ij} pueden representarse como una matriz $R_{m \times n}$ de los coeficientes estequiométricos de las especies producidas (a la derecha de cada reacción). De este modo, podemos construir la matriz estequiométrica $S \in \mathbb{N}^{m \times n}$, tal que $S_{ij} = b_{ij} - a_{ij}$ (o equivalentemente, $S = R_{m \times n} - L_{m \times n}$).

Una vez definido el vector de flujo v y la matriz estequiométrica S , podemos calcular la tasa de producción $f = S \cdot v$, donde $f = (f_1, f_2, \dots, f_m)^T$ es un vector que contiene la tasa de producción f_j para cada especie química s_j , como consecuencia de la aplicación del proceso. Sea x un vector de coordenadas no-negativas tal que x corresponde al número (o concentración) inicial de especies químicas s_1, \dots, s_m , entonces x_v es el nuevo estado de la red de reacciones $\langle X, R_X \rangle$ asociada a un estado x mediante la siguiente ecuación: $x_v = x + Sv = x + f$. En dinámica discreta, f representa la variación en número de especies químicas durante el intervalo temporal en el que ocurrió el proceso. En dinámica continua, f es la derivada de x respecto al tiempo: $x_v = x + \frac{dx}{dt}$.

Ahora al fin podemos definir un proceso auto-mantenido. Notemos que no todos los vectores de flujo engendran dinámicas de red viables: por ejemplo, el estado actual del sistema x podría no tener suficientes reactivos como para gatillar una reacción el número de veces que es requerida por el vector de flujo. Por ello, resulta útil formalizar el subconjunto de procesos viables [feasible processes] $\Pi(R_X)$ que puede ser aplicado al conjunto de moléculas X , al cual le llamaremos espacio de procesos. En la mayoría de sistemas químicos, el espacio de procesos está determinado por la ley cinética de acción de masas, donde la tasa $\lambda_i(r)$ de la reacción r_i está dada por la siguiente ecuación

$$\lambda(r_i, x) = k_i \prod_{j=1}^n x_j(t)^{a_{ij}}$$

Donde k_i es la tasa de reacción intrínseca de r_i , $x_j(t)$ es la concentración de especies s_j en el tiempo t , y a_{ij} es la tasa de reactivos de la especie s_j requeridos para gatillar r_i . Finalmente, X es un conjunto auto-mantenido si y sólo si existe un vector de flujo $v \in \Pi(R_X)$, tal que $r_i \in R_X$ implique que $v_i > 0$ y que $x_v[j] \geq x[j]$, para $j = 1, 2, \dots, m$. En otras palabras, un conjunto de moléculas es auto-mantenido si es posible encontrar vectores de flujo en su espacio de procesos tales que todas las reacciones de la red a estudiar tengan tasa positiva, y que todas las especies de moléculas consumidas son producidas en igual o mayor cantidad. De ese modo, si el conjunto de moléculas X es auto-mantenido, puede operar activando todas las reacciones en R_X sin disminuir su concentración (y en algunos casos, aumentándola).

Dicho esto, denominaremos Organización Química a un subsistema $\langle O, R_X \rangle$ de la red $\langle M, R \rangle$ (donde $O \subseteq M$), definido como un conjunto de objetos (moléculas) estructuralmente cerrado y auto-mantenido que consume y produce especies químicas (Dittrich y Di Fenizio 2007; Hordijk et al. 2018; Veloz et al. 2022).

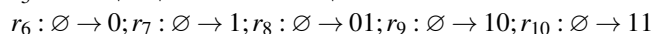
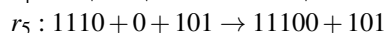
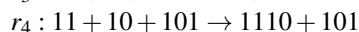
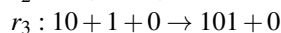
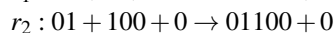
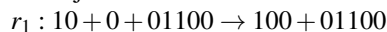
Pero ¿cuál es la relevancia de que una sub-red sea una organización? Al no generar nuevas moléculas y operar

establemente en el tiempo sin perder concentración de reactivos, es posible concebir a las organizaciones químicas como una abstracción de los atractores periódicos en Teoría de Sistemas Dinámicos: puntos fijos, órbitas periódicas, y ciclos límite de la red de reacciones (Veloz et al., 2022). La persistencia en el tiempo de una organización química permite tender puentes entre las redes de reacciones geoquímicas que ocurren espontáneamente y el inicio del metabolismo. Una red que cambia constantemente sus componentes moleculares y sub-procesos, pero que persiste en el tiempo con un conjunto definido de reacciones características (debido a y mientras perdure su clausura y auto-mantenimiento), es un sistema con identidad cualitativa. De este modo, $\langle M, R \rangle$ constituye un universo de reacciones donde no es necesario que haya sistemas definidos de antemano, y todas las organizaciones $\langle O, R_X \rangle$ encontradas son sistemas en dicho universo de reacciones. Eso permite tanto modelar la durabilidad y respuesta frente a ‘estresores’ ambientales de sistemas de interés ya conocidos (como arqueas y bacterias primitivas), como modelar la emergencia espontánea de organizaciones desde redes geoquímicas que ocurren actualmente, o que podrían haber ocurrido hace 4000 millones de años atrás.

Ejemplo de juguete: modelo de polímeros binarios

A falta de un ejemplo sencillo en la literatura que utilice especies químicas, ilustraré las similitudes y diferencias entre los RAFs de las redes autocatalíticas y las Organizaciones de COT mediante el ejemplo de polímeros binarios que Steel et al. (2019) mencionan brevemente. Sea $X = \{0, 1, 01, 10, 11, 100, 101, 1110, 01100, 11100\}$ un conjunto con 10 moléculas abstractas, representadas por cadenas de números binarios, y sea $F = \{0, 1, 01, 10, 11\}$ el conjunto alimento incluido en X , compuesto por todas las cadenas de largo ≤ 2 . Sea $R = \{r_1, r_2, r_3, r_4, r_5\}$ el conjunto de las 5 reacciones del CRS $\langle X, R, C \rangle$, descritas en la Fig. 1 (derecha), obtenidas mediante concatenación de cadenas; luego, los catalizadores son asignados al azar. Tenemos que para este sistema $\{r_1, r_2, r_3, r_4, r_5\}$ es maxRAF, y $\{r_3, r_4, r_5\}$ es el otro RAF encontrado, el cual además es el cierre del RAF $\{r_3\}$. El sistema cuenta en total con 6 RAFs (Fig. 1, derecha).

Aplicando la notación de COT al ejemplo de polímeros binarios, tenemos una red de reacciones $\langle M, R \rangle$, donde $M = \{0, 1, 01, 10, 11, 100, 101, 1110, 01100, 11100\}$, y la sub-red $M' = \{0, 1, 01, 10, 11\}$ presenta una regla de reacción $\emptyset \rightarrow s_j$, tal que hay un influjo constante de $\{0, 1, 01, 10, 11\}$ en la red. Sea entonces $R = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9, r_{10}\}$ el conjunto con las cinco reacciones del ejemplo:



Entonces, podemos obtener la matriz estequiométrica S de la red de 10 tipos de moléculas, y las 5 reacciones iniciales, donde cada fila es una especie química, y cada columna una

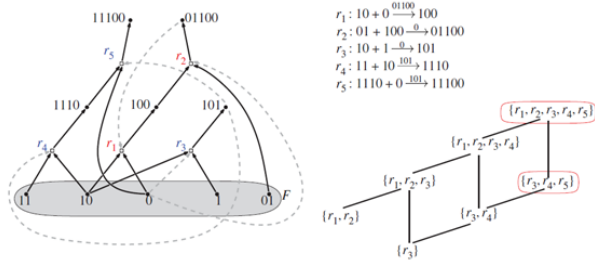


Figura 1: Un RAF sencillo que surge de un caso del modelo de polímeros binarios de Kaufmann. El modelo contiene 5 reacciones (denotadas como r_1, r_2, r_3, r_4, r_5), 10 tipos de moléculas, y un conjunto alimento F de 5 tipos de moléculas. **Izquierda:** red de reacciones del sistema. El conjunto alimento está encerrado en gris, las flechas indican dirección de las reacciones (representadas por cuadros blancos pequeños), las líneas punteadas indican catálisis. **Derecha, arriba:** reacciones del CRS. **Derecha, abajo:** diagrama de Hasse con todos los RAFs del CRS; encerrados en un óvalo rojo están los RAF con cerradura [closure]: un conjunto de reacciones R' está cerrado en un CRS si y sólo si existe un sub-conjunto R'' de R que contenga a R' y donde cada reacción $r \in R$ tenga a todos sus reactivos y al menos un catalizador, o bien, presente en el conjunto F , o bien, como producto de una reacción de R'' . Nótese que la definición de cerradura en redes autocatalíticas no es la misma que en COT. (Extraída de Steel et al., 2019.)

reacción:

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \end{bmatrix} \quad (1)$$

Al modelar la red sin el “conjunto alimento” (i.e. excluyendo $\{r_6, r_7, r_8, r_9, r_{10}\}$) encontramos cero organizaciones y 20 conjuntos reactivamente cerrados (ver Fig.2). Cinco de ellos son básicos (aquellos compuestos por únicamente una reacción) y tienen 4 especies cada uno; un conjunto $\{r_1, r_2\}$ tiene 5 especies, 3 tienen 6 especies, 4 tienen 7 especies, 3 conjuntos tienen 8 y 9 especies; y sólo el conjunto de todas las reacciones tiene todas las especies.

Sin embargo, al agregar las reacciones del conjunto alimento, encontramos dos organizaciones: aquella generada por r_2 y aquella generada por $\{r_1, r_2\}$ (ver Fig.3). La primera organización cuenta con 8 especies (0,1,01,10,11,101,1110,11100) y la segunda organización involucra a todas las especies. Esto permite explorar el ‘lema 1’ del artículo de Hordijk et al. (2018): Si un maxRAF no contiene algún CAF (conjunto constructivamente autocatalítico y F-generado: CAF es una noción más fuerte que RAF, la cual le agrega orden de ocurrencia a las reacciones del RAF), habrá entonces una organización química que consiste sólo en F , y que no corresponderá a ningún RAF cerrado. En el ejemplo $\{r_3, r_4\}$ es un CAF; y consistentemente, no hay ninguna Organización que contenga exclusivamente las

especies $\{0, 1, 01, 10, 11\}$.

COT Y UN ESCENARIO POSIBLE DEL ORIGEN DE LA VIDA: RESPIRADEROS HIDROTERMALES ALCALINOS

Dentro de los escenarios propuestos como sistemas capaces de haber originado la vida en la tierra, uno de los que acumula mayor evidencia a su favor es el de los respiraderos hidrotermales alcalinos (Martin et al. 2008; Lane y Martin 2012; Sojo et al. 2016; Kitadai y Maruyama 2018; Cartwright y Russell 2019; Brunk y Marshall 2021). Estudios recientes han encontrado vías geoquímicas que catalizan la fijación de carbono, muy similares a vías bioquímicas, como la vía del Acetil-CoA. Dichas vías utilizan catalizadores inorgánicos –típicamente, compuestos que incluyen metales de transición, como greigita (Fe_3S_4), magnetita (Fe_3O_4) y awaruita (Ni_3Fe), a temperaturas y pH hallables en respiraderos hidrotermales alcalinos (Preiner et al., 2020). Un requerimiento energético para las reacciones geoquímicas que dieron paso a la bioquímica temprana, es que estas sean exergónicas (i.e. que liberen energía), o bien, que se acoplen con reacciones exergónicas que permitan su ocurrencia espontánea. Por lo tanto, antes del advenimiento de moléculas capaces de almacenar energía (como el ATP), las reacciones biosintéticas independientes del ATP jugaron un rol central en el inicio del metabolismo, contribuyendo a la formación de redes autocatalíticas y organizaciones químicas.

Para ello, Wimmer et al. (2021) polarizaron la red hipotética central de 402 reacciones metabólicas presente en LUCA, de manera que las reacciones se encuentren en la dirección de síntesis celular. Las reacciones fueron obtenidas de KEGG. Partieron suponiendo disponibilidad inicial de $H_2, CO_2, NH_3, H_2S, H_2O$ y P_i , sin introducir alguna otra fuente de energía a la red. De los 18 cofactores presentes en la red biosintética (CoB, CoM, Metanofurán, PLP, ThPP, CoA, NADP, NAD, Biotina, SAM, F_{430} , Cobamida, THF, THMPT, FMN, FAD, Molibdopterina, y F_{420}), la vía del acetil-CoA requiere 10 de ellos para sintetizar piruvato desde H_2 y CO_2 . Las/os investigadores procedieron a calcular energía libre de Gibbs (ΔG_{mol}^{kJ}) para cada reacción individual, a diversas temperaturas entre 25 y 100°C, y variando el pH en el rango 1-14. Se encontró que, de las 351 reacciones calculables para las distintas combinaciones de parámetros, el mayor porcentaje de reacciones exergónicas (98%) se obtenía a 80°C, pH 7, concentraciones de no-equilibrio y retención de agentes reductores orgánicos (NADH, NADPH, flavodoxina reducida, ferredoxina reducida); sin embargo, el reemplazo de dichos agentes por H_2 (tal como ocurre en sistemas geoquímicos con serpentinización) a un pH de 9 resultó en una tasa de 95% de reacciones exergónicas. El único compuesto que fue reducido en las reacciones de la red central fue CO_2 . Se encontró que las reacciones de la red biosintética central de LUCA son energéticamente favorecidas en respiraderos hidrotermales alcalinos mientras haya presencia de fosfato.

Dado el resultado de Wimmer et al. (2021), cabe pre-

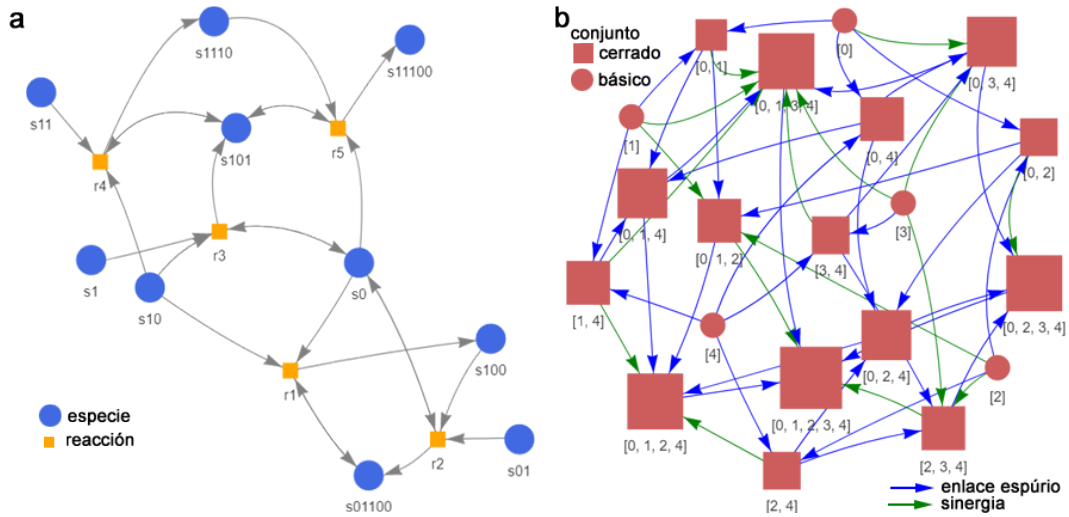


Figura 2: Grafos obtenidos mediante un algoritmo de COT desde el mismo caso del modelo de polímeros binarios de Kaufmann utilizado por Steel et al., 2019; pero sin incluir las reacciones $\emptyset \rightarrow s_j$ equivalentes al conjunto F. **(a)** visualización de la red de reacciones, los círculos azules corresponden a especies de la red, y los cuadrados amarillos a reacciones (donde r_N denota r_N). Las flechas bidireccionales permiten identificar a los catalizadores de una reacción. La red es topológicamente equivalente a la de la figura 1 (izquierda). **(b)** descomposición de las sinergias y visualización de las estructuras de la red. Los conjuntos son representados como círculos si son conjuntos de reacciones básicos (aquellos conjuntos cerrados generados por un generador), y cuadrados si no lo son. El color rojo denota que el conjunto es sólo reactivamente cerrado. En este caso ningún conjunto es semi-auto-mantenido, y por lo tanto, ninguno es una organización. Un par de conjuntos es sinérgico si su dinámica no puede computarse desde las dinámicas de sus constituyentes cerrados (Veloz et al., 2019). Las flechas verdes denotan sinergias y las flechas azules, uniones espurias. Los números entre corchetes denotan los generadores del conjunto. El tamaño de los círculos o cuadrados refleja el número de especies que el conjunto contiene.

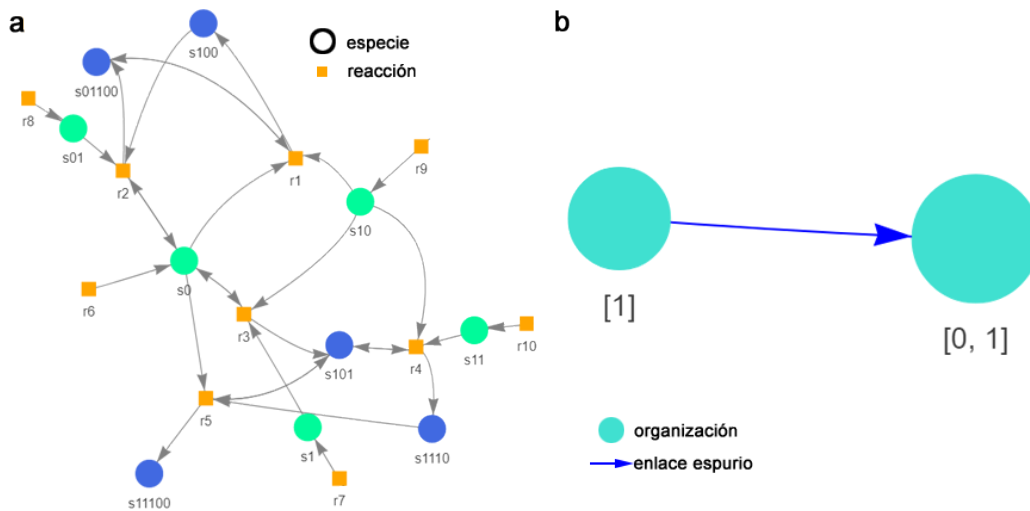


Figura 3: Grafos obtenidos mediante un algoritmo de COT desde el mismo caso del modelo de polímeros binarios de Kaufmann utilizado por Steel et al., 2019; incluyendo el equivalente al conjunto F. **(a)** visualización de la red de reacciones, los círculos corresponden a especies, el color verde indica especies afluentes (el conjunto alimento), y los cuadrados amarillos a reacciones (donde r_N denota r_N). **(b)** descomposición de las sinergias en la red de reacciones y visualización de las estructuras de la red. Los conjuntos son representados como círculos si son el cierre de una o más reacciones, y el color verde denota que el conjunto de reacciones constituye una Organización Química. Tenemos entonces dos organizaciones: aquella generada por r_2 y aquella generada por $\{r_1, r_2\}$.

guntarse cuantas organizaciones emergen espontáneamente en dicha red: la formación e identificación de RAFs no permite conocer la estabilidad en el tiempo de dichas redes autocatalíticas, en cambio, las organizaciones químicas son por definición temporalmente estables. Utilizar COT para encontrar conjuntos cerrados, semi-auto-mantenido, sinergias y organizaciones químicas puede permitirnos detallar el espacio de procesos del origen del metabolismo en un contexto hidrotermal alcalino, además de hacer conjeturas respecto a los posibles regímenes dinámicos más plausibles

previos al surgimiento de la vida: Conocer el ΔG de los procesos relevantes (calculados por (Wimmer et al., 2021)) permite complementar la descripción estequiométrica y pasar a modelar dinámicamente las organizaciones más prometedoras, enfocando de manera óptima los recursos computacionales disponibles.

Utilicé el repositorio pyRN (<https://github.com/pmalдона/pyRN>) en Python 3.8 para generar una descomposición estructural de la red de 402

reacciones, y así poder caracterizar sus organizaciones y las propiedades de su espacio de procesos. Agregué 6 reacciones de flujo ($r_1 : \emptyset \rightarrow H_2; r_2 : \emptyset \rightarrow CO_2; r_3 : \emptyset \rightarrow NH_3; r_4 : \emptyset \rightarrow P_i; r_5 : \emptyset \rightarrow H_2O; r_6 : \emptyset \rightarrow H_2S$), que corresponden a las especies libremente disponibles en el sistema: hidrógeno molecular, dióxido de carbono, amoníaco, ortofosfato, oxidano, y ácido sulfhídrico (Fig.4).

El siguiente paso, es obtener las organizaciones: debido a la complejidad computacional de calcular todos los conjuntos cerrados, semi-automantenidos, y las organizaciones de la red; se realizó inicialmente un muestreo aleatorio que incluyese las 6 reacciones de flujo más un número de reacciones entre 55 y 70; luego, a medida que iban apareciendo organizaciones, se seleccionaron conjuntos de tamaños variables de manera heurística, para aumentar la probabilidad de encontrar organizaciones. Mediante esta aproximación se encontró (i) la organización que contiene exclusivamente el conjunto de especies afluentes ($H_2, CO_2, NH_3, P_i, H_2O$, y H_2S), (ii) la organización que contiene formiato (CH_2O_2), bicarbonato (HCO_3^-) y protones (H^+), además de $H_2, CO_2, NH_3, H_2O, H_2S$, y P_i ; (iii) la organización que además de las especies afluentes contiene ferredoxina reducida y oxidada (las cuales constituyen un clúster hierro-sulfuro [FeS]), y monóxido de carbono (CO). También fueron encontradas varias organizaciones intermedias que consisten en iteraciones de los elementos de las organizaciones previamente mencionadas (ver Fig. 5).

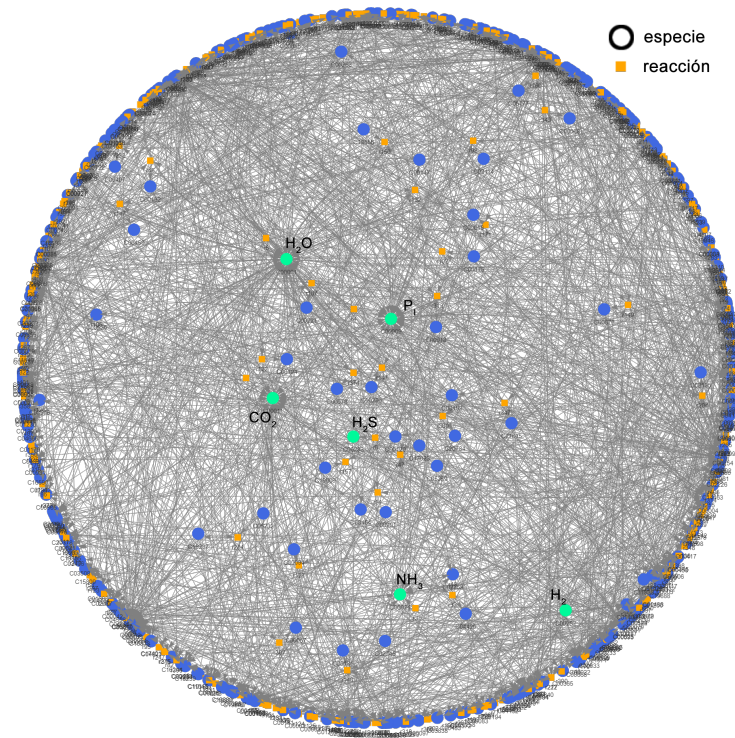


Figura 4: Red biosintética primigenia. Los círculos corresponden a especies, el color verde indica especies afluentes (el conjunto alimento).

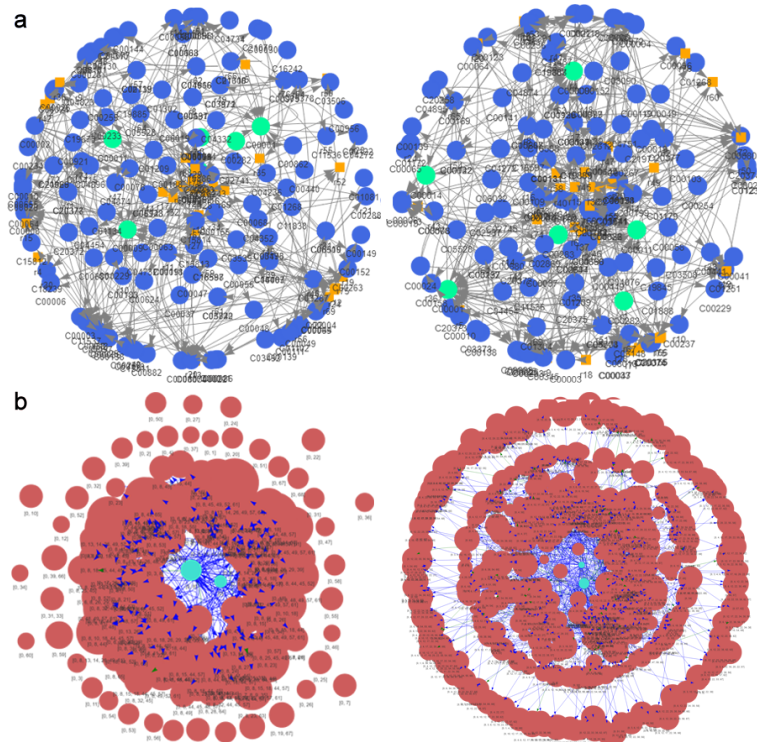


Figura 5: Ejemplos de redes aleatorias de 76 reacciones con dos organizaciones cada una. **(a)** gráficos de las redes de reacciones (cuadrados amarillos), Los círculos corresponden a especies, el color verde indica especies afluentes. La red de la izquierda fue generada completamente al azar, y la red de la derecha fue generada con base en la información previa acerca de las organizaciones del sistema. **(b)** descomposición de las sinergias en la red de reacciones y visualización de las estructuras de la red. Los conjuntos son representados como círculos si son reactivamente cerrados, el color verde denota que el conjunto de reacciones constituye una Organización Química.

CONCLUSIONES

El paso de la geoquímica a la bioquímica requiere una colección de moléculas en interacción periódica. Dicha interacción requiere a su vez un influjo constante de moléculas desde el ambiente, tales como H_2 , CO_2 , fosfato inorgánico o H_2O (representado en redes autocatalíticas como el conjunto F , y en COT como las reacciones de influjo $\emptyset \rightarrow s_j$); además, para que la velocidad de las reacciones sea compatible con la vida, son necesarios catalizadores previos al surgimiento de la traducción de enzimas (e incluso, previos a las ribozimas u otras estructuras de RNA con capacidad catalítica), en otras palabras, hemos de contar con catalizadores inorgánicos como los clústeres FeS o NiS. Dichas reacciones de catálisis son representadas en redes autocatalíticas de la siguiente manera: –tomando un ejemplo de Zhao et al. (2021) $-r : CO + H_2O \xrightarrow{FeS} CO_2 + 2H$, y en COT como $r : CO + H_2O + FeS \rightarrow CO_2 + 2H + FeS$. Pudimos corroborar el ‘Lema 1’ de Hordijk et al. (2018); quienes además muestran que sólo los subconjuntos autocatalíticos cerrados son organizaciones químicas; y que, además, no todas las organizaciones químicas son necesariamente conjuntos autocatalíticos ‘cerrados’ (en la acepción propia de dicha teoría). Dado que el concepto de RAF no es equivalente al concepto de Organización Química (que además de ser auto-mantenida; debe ser cerrada, en la acepción de COT), ambos enfoques resultan complementarios para descubrir propiedades estructurales de los CRS geoquímicos de interés para modelar el surgimiento de los primeros metabolismos. Utilizando la red biosintética primigenia, encontramos que además de la organización de las especies afluentes, tienden a aparecer espontáneamente organizaciones que incluyen especies disponibles en escenarios de respiraderos hidrotermales alcalinos que a su vez se han propuesto como centrales para el origen del metabolismo con independencia de la teoría que se baraje: los clusters de FeS antecesores a la ferredoxina, tales como la Greigita, han sido propuestos como catalizadores inorgánicos primarios presentes en las superficies de los respiraderos (Sojo et al. 2016; Russell y Hall 2006). Se ha visto que el formiato se produce de manera abiótica en condiciones con alto hidrógeno molecular, el cual es un producto de las reacciones de serpentinización que ocurren en los respiraderos hidrotermales alcalinos (Lang et al., 2018), siendo además el primer intermediario en la vía acetil-CoA (o Wood-Ljungdahl). En la interfaz entre el océano (ácido) y los respiraderos alcalinos se producen gradientes de protones (H^+), que se postulan centrales para la generación de energía en los organismos tempranos mediante fuerza protón-motriz (Sojo et al., 2016); previamente, el hidrógeno molecular y el carbono podrían haber funcionado como un par redox (donante-aceptor) para proveer energía a las reacciones en los organismos anteriores a la aparición de la quimiosmosis (Brunk y Marshall, 2021). Finalmente, el bicarbonato es la forma del carbono con mayor chance de interactuar con las superficies catalíticas de FeS (Roldan et al., 2015). El hecho de que se conformen distintas organizaciones que incluyen dichas especies permite conjeturar un mapa dinámico de las reacciones que pueden haber resultado en sistemas autoorganizados estables en los respiraderos hidrotermales, miles de millones de años atrás. Por ejemplo, podría indicar cierta independencia entre la red

de formación de formiato, y la red catalítica prebiótica de clústeres FeS.

La red de reacciones que da paso a los primeros sistemas vivientes puede construirse de dos maneras: reduciendo el metabolismo de bacterias o arqueas consideradas primitivas a sus mínimos componentes comunes; o bien, sumando especies químicas relevantes, capaces de participar en reacciones termodinámicamente favorables unas con otras, y que estén presentes en escenarios plausibles para el surgimiento de la vida, como son los respiraderos hidrotermales alcalinos aquí explorados, pero como también son otros sistemas submarinos como fumarolas negras, océanos fríos abundantes en cianuro de hidrógeno (HCN), superficies cristalinas de minerales, o sistemas hidrotermales sub-aéreos como aguas superficiales cerca de geiseres. Reconstruir las propiedades estructurales de las distintas redes de reacciones resultantes puede favorecer la comparación entre escenarios plausibles del origen del metabolismo.

AGRADECIMIENTOS

Éste trabajo fue financiado mediante la beca doctoral ANID n°21210914. Agradezco a Pedro Maldonado por su ayuda con la librería pyRN, y a Tomás Veloz por comentarios y sugerencias sobre el manuscrito.

DISPONIBILIDAD DE DATOS

Tanto el archivo con las redes de reacciones utilizadas como el código utilizado (archivo .py) pueden ser compartidos previa solicitud al autor de este trabajo. La librería utilizada para implementar COT está disponible en <https://github.com/pmaldona/pyRN>. Los números de las especies químicas utilizadas corresponden a números K de la base de datos KEGG Orthology, disponible en <https://www.genome.jp/kegg/ko.html>.

BIBLIOGRAFÍA

- [1] Barandarian, X. y Moreno, A. (2008). “Adaptivity: From metabolism to behavior”. *Adaptive Behavior*, 16(5):325–44.
- [2] Brunk, C. F. y Marshall, C. R. (2021). “Whole organism, systems biology, and top-down criteria for evaluating scenarios for the origin of life”. *Life*, 11(7):1–26.
- [3] Cartwright, J. H. y Russell, M. J. (2019). “The origin of life: the submarine alkaline vent theory at 30”. *Interface: Focus*, 9(6):20190104.
- [4] Contreras, D., Pereira, U., Hernández, V., Reynaert, B., y Letelier, J. (2011). “A loop conjecture for metabolic closure”. *Advances in Artificial Life*, pp. 176–83.
- [5] Dittrich, P. y Di Fenizio, P. S. (2007). “Chemical organisation theory”. *Bulletin of Mathematical Biology*, 69, 4:1199–231.
- [6] Dodd, M., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., O’Neil, J., y Little, C. (2017). “Evidence for early life in earth’s oldest hydrothermal vent precipitates”. *Nature*, 543(7643):107–48.
- [7] Fischer, W., Hemp, J., y Johnson, J. E. (2016). “Evolution of oxygenic photosynthesis”. *Annual Review of Earth and Planetary Sciences*, 44:60–4.
- [8] Guengerich, F. (2016). “Metals in biology 2016: Molecular basis of selection of metals by enzymes”. *Journal of Biological Chemistry*, 291(40):20838–39.
- [9] Hordijk, W., Steel, M., y Dittrich, P. (2018). “Autocatalytic sets and chemical organizations: Modeling self-sustaining reaction networks at the origin of life”. *New Journal of Physics*, 20(1):647–83.

- [10] Hudson, J. y Mankin, J. (1981). "Chaos in the belousov-zhabotinskii reaction". *The Journal of Chemical Physics*, 74:6171–77.
- [11] Kitadai, N. y Maruyama, S. (2018). "Origins of building blocks of life: A review". *Geoscience Frontiers*, 9(4):1117–1153.
- [12] Lane, N. y Martin, W. F. (2012). "The origin of membrane bioenergetics". *Cell*, 151(7):1406–16.
- [13] Lang, S. Q., Früh-Green, G., Bernasconi, S., Brazelton, W., Schrenk, M., y McGonigle, J. (2018). "Deeply-sourced formate fuels sulfate reducers but not methanogens at lost city hydrothermal field". *Scientific reports*, 8(1):755.
- [14] Martin, W., Baross, J., Kelley, D., y Russell, M. J. (2008). "Hydrothermal vents and the origin of life". *Nature Reviews Microbiology*, 6(11):805–14.
- [15] Maturana, H. y Varela, F. (1973). *De máquinas y seres vivos: Autopoiesis: La organización de lo vivo*. Editorial Universitaria.
- [16] McDermott, J. M., Seewald, J. S., German, C. R., y Sylva, S. P. (2015). "Pathways for abiotic organic synthesis at submarine hydrothermal fields". *Proceedings of the National Academy of Sciences*, 112(25):7668–72.
- [17] Miras, H. N., Mathis, C., Xuan, W., Long, D.-L., Pow, R., y Cronin, L. (2020). "Spontaneous formation of autocatalytic sets with self-replicating inorganic metal oxide clusters." *Proceedings of the National Academy of Sciences*, 117(20):10699–705.
- [18] Preiner, M., Igarashi, K., Muchowska, K. B., Yu, M., Varma, S. J., Kleinermanns, K., Nobu, M. K., Kamagata, Y., Tüysüz, H., Moran, J., y Martin, W. F. (2020). "A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism". *Nature ecology & evolution*, 4(4):534–42.
- [19] Prigogine, I. y Nicolis, G. (1971). "Biological order, structure and instabilities". *Quarterly Reviews of Biophysics*, 4(2–3):107–48.
- [20] Pulselli, R., Simoncini, E., y Tiezzi, E. (2009). "Self-organization in dissipative structures: A thermodynamic theory for the emergence of prebiotic cells and their epigenetic evolution". *BioSystems*, 96(3):237–41.
- [21] Razeto-Barry, P. (2012). "Autopoiesis 40 years later. a review and a reformulation". *Origins of Life and Evolution of Biospheres*, 42(6):543–67.
- [22] Reynaert, B. (2013). *Estudio sobre clausura metabólica*, Tesis de magister. Facultad de Ciencias; Universidad de Chile, Santiago, Chile.
- [23] Ribeiro, A. J., Holliday, G., Furnham, N., Tyzack, J. D., Ferris, K., y Thornton, J. M. (2018). "Mechanism and catalytic site atlas (m-csa): a database of enzyme reaction mechanisms and active sites". *Nucleic Acids Research*, 46(D1):D618–23.
- [24] Roldan, A., Hollingsworth, N., Roffey, A., Islam, H. U., Goodall, J. B., Catlow, C., Darr, J. A., Bras, W., Sankar, G., Holt, K. B., Hogarth, G., y de Leeuw, N. H. (2015). "Bio-inspired CO₂ conversion by iron sulfide catalysts under sustainable conditions". *Chemical communications*, 51(35):7501–504.
- [25] Russell, M. y Hall, A. (2006). "The onset and early evolution of life". *Geological Society of America*, Memoir 198:1–32.
- [26] Semenov, S. N., Kraft, L. J., Ainla, A., Zhao, M., Baghbanzadeh, M., Campbell, V. E., Kang, K., Fox, J. M., y Whitesides, G. (2016). "Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions". *Nature*, 537(7622):656–60.
- [27] Sojo, V., Herschy, B., Whicher, A., Camprubí, E., y Lane, N. (2016). "The origin of life in alkaline hydrothermal vents". *Astrobiology*, 16(2):181–97.
- [28] Steel, M., Hordijk, W., y Xavier, J. C. (2019). "Autocatalytic networks in biology: Structural theory and algorithms". *Journal of the Royal Society: Interface*, 16(151).
- [29] Steel, M., Xavier, J. C., y Huson, D. H. (2020). "The structure of autocatalytic networks, with application to early biochemistry". *Journal of the Royal Society: Interface*, 17(171).
- [30] Veloz, T., Bassi, A., Maldonado, P., y Razeto-Barry, P. (2019). "Towards an analytic framework for system resilience based on reaction networks". *Complexity*.
- [31] Veloz, T., Maldonado, P., Busseniers, E., Bassi, A., Beigi, S., Lenartowicz, M., y Heylighen, F. (2022). "Towards an analytic framework for system resilience based on reaction networks". *Complexity*.
- [32] Veloz, T. y Razeto-Barry, P. (2017). "Reaction networks as a language for systemic modeling: Fundamentals and examples". *Systems*, 5(1):1–16.
- [33] Wimmer, J. L. E., Xavier, J. C., Vieira, A., Pereira, D. P., Leidner, J., Sousa, F. L., Kleinermanns, K., Preiner, M., y Martin, W. F. (2021). "Energy at origins: Favorable thermodynamics of biosynthetic reactions in the last universal common ancestor (LUCA)". *Frontiers in Microbiology*, 12.
- [34] Xavier, J. C., Hordijk, W., Kauffman, S., Steel, M., y Martin, W. F. (2020). "Autocatalytic chemical networks at the origin of metabolism". *Proceedings of the Royal Society B: Biological Sciences*, 287(1922).