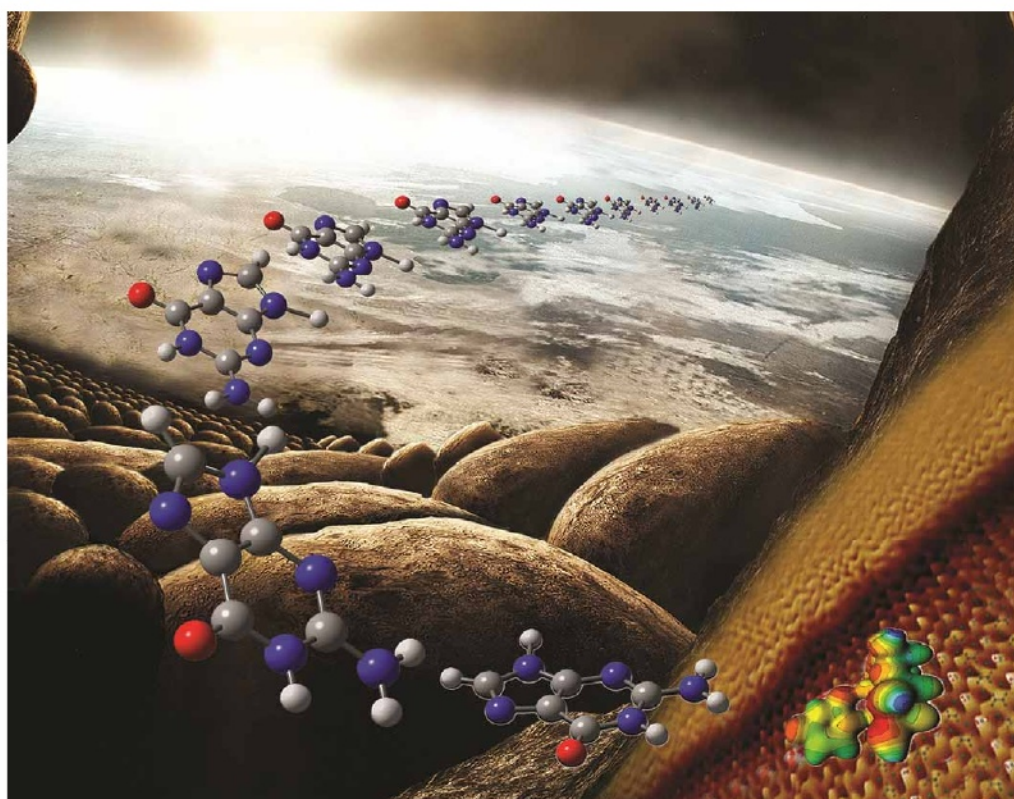


Chem Soc Rev

这篇文章是作为
益生元化学主题专刊

特邀编辑 jean - francois Lambert、Mariona Sodupe 和 Piero
Ugliengo

请查看2012年第16期[目录](#)，以访问本主题期的[其他评论](#)



Cite this: Chem. Soc. Rev., 2012, 41, 5430–5446

www.rsc.org/csr

TUTORIAL REVIEW

生命起源研究中的理论、建模与仿真

Peter V. Coveney,^a Jacob B. Swadling,^a Jonathan A. D. Wattis^b 和 H. Christopher Greenwell^c

收稿日期: 2012.01.19 DOI:

10.1039/c2cs35018a

生命起源研究代表了一个令人兴奋的、高度跨学科的研究领域。在这篇综述中,我们将重点关注理论、建模和仿真对解决该领域基本问题的贡献,以及这些方法在该领域所取得的进展。基于对非线性催化化学反应网络显著特性的分析,理论方法将继续在“系统化学”层面产生重大影响,非线性催化化学反应网络的显著特性是由于许多与自我复制和自我复制相关的假定过程的自催化和交叉催化性质而产生的。通过这种方式,我们描述了原始达尔文汤中RNA复制的非线性动力学模型,以及同手性和同手性聚合的起源。然后,我们讨论了目前正在部署的最先进的基于计算的分子建模技术,以研究与生命起源相关的各种场景。

1. 介绍

在大众的想象中,很少有以化学为基础的话题能与地球上(当然还有宇宙中其他地方)的生命起源相媲美。考虑到我们与这些关键事件之间相隔了很长一段时间,今天,这个主题在很大程度上仍然是一个理论和猜想。我们首先简要讨论当前生命起源研究理论的各个方面,特别强调生命前化学,其中化学物种和定义生命的基本现象(包括代谢,信息复制,区系化)出现。在确立了当前的概念立场后,我们对各种关键生命起源过程的理论建模进行了回顾,然后介绍了该领域内计算化学模拟的方法和应用。我们表明,多年来,理论和建模已经为这些复杂且通常是非线性过程的物理化学现象提供了相当大的见解,并阐明了生物分子或原生物分子与矿物表面之间的化学相互作用。

“地球上的生命是如何以及在哪里出现的?”这个问题一直以来都引起了广泛的关注,近年来更是如此。生命起源研究代表了一个令人兴奋的、高度跨学科的研究领域,它融合了地质学家、物理学家、生物学家、数学家、化学家和计算机科学家等人的贡献。在这篇综述中,我们重点关注建模和模拟对回答生命起源问题的贡献,以及这些技术在这一领域取得的更广泛的进展。

所有生物体生物化学的统一性表明,生命起源于一个最后的普遍共同祖先(LUCA)。如果LUCA从地球早期的非生物成分进化的阶段是已知的,原则上我们应该能够编制一个完整生命起源的描述。在生命起源研究中进行的大部分研究旨在解决这样一个问题:“构成第一个自我复制系统的有机分子的来源是什么,在哪里?”以及“生物组织是如何从一个非生物的有机分子来源进化而来的?”^{1,2}

最早被科学界采纳的生命起源假说之一,是查尔斯·达尔文提出的“温暖的小池塘”理论。很难想象传统的“温暖的小池塘”概念,也被称为“益生元汤”,是如何产生生命的,这不仅是因为缺乏地质证据,还因为在这样一个系统中没有明显的自由能来源。形成RNA(或蛋白质)所需的聚合发生有两点是必要的:能量和合理浓度的核糖核苷酸(或氨基酸)。此外,通常设想的益生元汤处于热力学平衡状态,而所有生命过程本质上都远非如此

^a Centre for Computational Science, Department of Chemistry, UCL, 20 Gordon Street, London, WC1H 0AJ, UK.

E-mail: p.v.coveney@ucl.ac.uk, jacob.swadling.09@ucl.ac.uk; Tel: 020 7679 4802

^b School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK.

E-mail: jonathan.wattis@nottingham.ac.uk

^c Department of Earth Sciences, Durham University, Durham, DH1 3LE, UK. E-mail: chris.greenwell@durham.ac.uk

w Part of the prebiotic chemistry themed issue.

z Electronic supplementary information (ESI) available: See DOI: 10.1039/c2cs35018a

平衡;作为任何这样的平衡的结果,核苷酸将不愿进一步反应。为了解决这个问题,已经假设和测试了许多物理和/或化学梯度施加远离平衡条件的方法,包括使用矿物质浓缩单体,³盐诱导的模仿潮汐池的肽形成反应,⁴和冷冻过程中单体的浓度。⁵

重要的是,我们要从驳斥一个毫无意义的论点开始,这个论点经常被生命起源界的成员令人惊讶地阐述出来。这种观点认为,通过随机组装过程合成一克“一个”(或几个)特定的自我复制序列的概率,将需要比地球上现有的全部物质质量更大的物质,因此不可能发生。³⁴这一论点基于一种低级的观念,即RNA序列是通过随机合成(即好像处于平衡状态)形成汤状的,完全忽略了它们动态自组装的非线性性质。它所确定的是,并不是所有的序列都被制造出来了:大自然并没有对所有可能的化学物种进行全局的最优性搜索。在决定产生了哪些序列时,有很大的偶然性。

生命的确是由一系列发生在远离平衡状态的化学过程所驱动的。⁶为了维持这些过程,所有的生物体都是开放的系统;它们的复杂性建立在反馈的基础上,这些反馈涉及自催化和交叉催化分子,这些分子辅助反应而不会在过程中被破坏。一种代谢或调节途径可能产生一种分子,这种分子可以加速其他途径(交叉催化),这些途径通过大量相互关联的化学反应,最终可能间接催化原始途径。

米勒和尤里对非生物合成的开创性贡献解决了与益生元汤相关的许多缺点。他们通过甲烷、氨和氢的混合物进行放电,产生氨基酸和其他产物。这证明了能量输入可以产生生物化学物质。米勒继续这项研究,表明氨基酸的主要合成途径是通过斯特雷克反应。⁷米勒的反应是霍尔丹在他1929年关于生命起源的文章中提出的。霍尔丹提出,紫外线(UV)辐射提供了将甲烷、氨和水转化为早期海洋中第一批有机化合物的能量。⁸早期地球大气层的历史还没有完全解决,尽管有证据表明早期地球没有强烈的还原大气层,这使得米勒-尤里实验不像最初认为的那样相关⁹

20世纪70年代海底**深海热液喷口**的发现最终导致了一个进一步的、似是而非的替代假设,为早期地球的大气条件提供了一个可能的解决方案。这些热液喷口与周围的海水不平衡。¹⁰“黑烟”是第一类被发现的热液喷口;它们在一定程度上是由于岩浆在海洋扩张带与海水相互作用而产生的。然而,作为生命起源的候选者,黑烟带提出了一些挑战,包括它们的极高温、低pH值、短寿命和缺乏三维分隔,¹¹可以想象,这将导

致早期海洋中原始生物分子的不可挽回的稀释。

第二种类型的热液喷口是在千禧年之交发现的。¹²与黑烟不同,这些“**碱性喷口**”不是火山喷发形成的,而是通过蛇形化形成的。蛇形化是橄榄石的氧化,随后水被还原成氢。¹³这个过程会产生大约150-200摄氏度的温度和富含氢的pH值为9-11的强碱性液体。Russell等人在1993年提出了这些碱性喷口系统在益生元情景中的作用,之后它们最终在大西洋中部的“失落之城”被发现。¹⁴Russell及其同事假设碱性喷口是产生生命的潜在候选者,因为它们相对温和的温度和压力,在新兴的碱性流体中通过质子和氧化还原电位梯度孵化原始代谢的明显能力,以及它们微妙的内部多孔结构作为原始细胞。

“**RNA世界**”假说是关于地球上生命起源的一个成熟的理论,该假说认为,在目前脱氧核糖核酸(DNA)/蛋白质世界出现之前的某个时刻,核糖核酸(RNA)为催化和复制提供了分子基础。¹⁵这个理论很有吸引力,因为它解决了Miller-Urey实验提出的复杂性问题:RNA世界只需要一种生物分子的合成。支持RNA世界假说的证据来自于切赫和奥特曼的工作,他们因独立发现一组被称为核酶的催化RNA分子而获得1989年诺贝尔化学奖;¹⁶进一步来自于斯蒂茨和摩尔发现的核糖体的结构,核糖体是一种制造蛋白质的细胞成分。¹⁷这些后一项研究有助于证明核糖体实际上是一种核酶。最近,萨瑟兰和他的同事提供了一种在接近益生元条件下合成核苷酸的非生物途径(见图1)。¹⁸然而,尽管RNA世界假说在益生元化学研究中普遍存在并迄今占据主导地位,但另一种假说,即蛋白质最初是由现成的氨基酸产生的,并产生了一个蛋白质世界,仍然受到关注,并且也有很强的理由。¹⁹

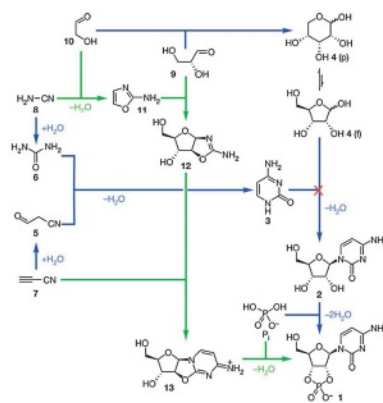


图1嘧啶核糖核苷酸的组装选项。先前假设合成b-核糖苷-29,39-环磷酸1(蓝色;注意胞嘧啶3和核糖4被提议浓缩在一起的步骤的失败)和萨瑟兰等人的工作中描述的成功的新合成(绿色)。p,吡喃糖;f,呋喃糖。经参考文献18许可转载。

在这篇综述中,我们将重点介绍在益生元背景下使用理论建模和模拟方法取得的进展。从本质上讲,我们可以将讨论的建模方法分为两组:

(i)那些描述“更高层次”现象的方法,如手性扩增,或囊泡形成动力学,使用质量作用定律,没有直接的化学结构信息;

(ii)使用量子数和/或分子力学等计算化学方法在电子结构、原子或分子水平上详细描述化学结构和过程的。

这两种方法的一个共同点是,它们往往被实验益生元化学界的许多部门所忽视。尽管事实上,生命起源前的化学必然要严重依赖抽象和理论,像20年代的“硫化铁世界”这样的概念,其根源就在于对化学反应能量学的计算。

2. 理论非线性化学动力学

理论能说些什么来指导我们设定参数,限制对生命起源的探索呢?对于一个化学反应环环相扣的复杂系统,热力学第二定律是起点。它定义了一组热力学势,这些势的极值指定了物质的平衡状态。特别是,在化学反应的通常条件下(固定温度和压力,或固定温度和体积),吉布斯和亥姆霍兹自由能作为势,其全局极小值完全定义了平衡状态。

21

如果系统由于施加外部约束(例如温度或浓度梯度等)而无法达到热力学平衡,但能够充分接近平衡(从技术上讲,热力学力和通量在其消失的平衡值的位移中是线性的),那么适当的热力学势是系统的内部熵产。²²也就是说,它的极值(通常是最小值,但在某些情况下是最大值)定义了系统的定态。

然而,对于远离平衡态的系统(即热力学力和通量不再线性依赖于与平衡态的距离),不存在指导系统演化的热力学势。虽然这一结果是在35年前建立的,但很少有人注意到它。^{20,23}除了最重要的不可逆性(即熵总是在增加)之外,传统意义上的化学热力学对远离平衡的系统的全局行为几乎没有什么可说的,比如物质的生命状态和生命起源所涉及的那些。相反,为了获得更多的定量见解,我们必须转向非线性动力学的研究,这是复杂反应动力学的理论基础所在的形式数学领域,所有这些都与第二定律相容。

在过去的四十年左右,我们对非线性动力系统行为的理解取得了巨大的进步,其中大部分可以追溯到图灵关于形态

发生的化学基础的研讨会论文。²⁴远离平衡的化学反应提供了一个绝佳的例子,因为基于质量作用定律的许多速率过程(包括自催化和交叉催化过程)都是非线性的。定性地说,非线性系统之所以如此迷人,是因为从传统的线性、平衡或稳态理论的角度来看,它们表现出了“意想不到”的特性。非线性系统更容易让人联想到我们所居住的世界:它们可以为相同的方程和参数表现出多种状态(如多稳定性和滞后),各种形式的空间、时间和时空组织(通常被称为耗散结构或“自组织”),混沌(在技术意义上)及其推论,对初始条件的敏感依赖。²⁰正是通过应用非线性动力系统的数学,我们才有望更好地理解生命的起源²⁰

2.1 超循环

艾根和舒斯特在题为《超循环:自然自组织原理》的三篇论文中,以²⁵的形式提出了与生命起源有关的最早的过程模型。在他们的论文中,作者讨论了反应循环的联锁,以此来解释益生元系统的自组织。作者描述了这种特殊的所谓超循环型功能组织,并证明了其与生命起源和进化的可能相关性。艾根和舒斯特用“准物种”的概念来表示自我复制的大分子,如RNA和DNA。他们将“准种”定义为大分子物种的一种分布,其密切相关的序列由一个或几个主拷贝主导。外部约束强制选择最适合的分布,通常被称为“野生型”。对于达尔文行为来说,最重要的是准物种内部稳定性的某些特定标准。结果表明,RNA或DNA分子的选择和进化是有限的,相对于可以存储在单个复制单元中的信息量而言。

利用动力系统的数学理论,艾根和舒斯特表明,只有一种机制可以满足存储在每个复制单元(或繁殖周期)中的信息必须保持的要求,即各自的主副本必须与其误差分布有利地竞争。尽管存在竞争行为,但这些单位必须建立一种包括所有功能整合物种在内的合作关系。另一方面,这个循环作为一个整体必须继续与任何其他单一实体或相互联系的整体进行激烈的竞争,这不利于其整合功能。艾根和舒斯特提出的要求对于选择最适合的、功能联系的整体及其进化优化至关重要。只有超循环组织能够满足这些要求(自主繁殖周期之间的非循环连接,如链或分支,树状网络缺乏此类性质)。

艾根和舒斯特的工作提出了一个与遗传密码起源相关的超循环的现实模型

平移机械。这些模型包括以下特征:超旋回具有足够简单的结构,可以在预言条件下以有限概率承认起源;它允许从密切相关的(转移RNA样)前体持续出现,最初是稳定的RNA准物种的成员,被放大到更高的丰度水平。这种超循环的单个功能单元的组织结构和性质反映在原核细胞翻译装置内的遗传密码中,也反映在某些细菌病毒中。

这种超循环很快就遭到了批评:进化生物学家梅纳德·史密斯(Maynard Smith)指出,在寄生虫分子的存在下,这种超循环是不稳定的,因为寄生虫分子会为了自己的私利而破坏这种机制。²⁶寄生虫利用超循环的交叉催化特性来强化自己的数量,却没有给予任何回报。因此,梅纳德·史密斯认为,这些超级周期会在任何一段相当长的时间内消亡。他的分析是基于一个理想化的场景,在这个场景中,超级周期的成分在空间中均匀分布。Eigen和Schuster在八年前撰写的原始论文中得出结论,要保持超循环的正常运行,需要某种形式的分隔(即空间非均匀性)。事实上,在“区隔论者”和生命共同体起源的其他人之间,关于区隔先于RNA复制出现,还是相反,存在着长期的争论。²⁷有人提出,由于隔室通透性对物质输入和生长至关重要,因此寄生不仅是可能的,而且是不可避免的。然后,由寄生虫引起的传染性的程度将取决于区隔化的具体情况。

2.2 最近对RNA世界假说的评估

尽管RNA的世界在概念上有各种各样的吸引力,但还没有人能从实验上证明它的可行性。因此,与该场景相关的各种理论问题往往成为分析的对象。

绍斯塔克等人²⁸强调了合成生命领域的潜力和目标,特别是RNA领域。一个核心挑战是RNA复制酶的发现。这个问题也被马和吴注意到了²⁹,他们认为自我复制的RNA分子一定是RNA世界的先驱,信息内容随后进入。即使有充足的底物供应,RNA链的组装也是低效的;因此绍斯塔克等人认为可能需要活化的核苷酸。没有天然存在的核酶可以催化必要的化学反应。复制酶必须既是用于复制的核酶模板(这表明它具有天然开放的结构),又是具有活性的聚合酶(这表明它具有稳定的折叠结构)。

这些明显矛盾的说法可以通过假设环境经历某种形式的规则振荡行为来调和;例如潮汐循环导致温度循环或单体浓度的周期性波动。早期的撞击创造了月球,也导致地球经历了快速的自转,周期在2-6 h。³⁰这导致了沿海地区显著的潮汐活动,

洪水和干燥产生了盐度和温度的振荡,使得DNA链之间的链间相互作用交替促进结合和分离。这将允许复制链中模板化的信息。由于在晚期重轰炸结束和生命起源之间只有相对较短的几亿年,因此该理论带来了在大约30万年的相当短的时间尺度上解释生命起源的需要。

Fernando等人³¹支持这样的观点,即有电镀但没有酶催化的体系会发生链延长,从而导致链的总浓度降低,整体反应减慢。他们构建了一个基于计算机的这一过程模型,并发现潮汐循环显著提高了延伸率。

Fernando和迪保罗研究了另一个复制模型³²,即化学子。甘蒂(Ganti)提出了生命的定义³³,认为生命涉及新陈代谢、调节、控制、稳定和具有信息子系统。化学子具有自催化、膜和信息子系统。这提倡了模板复制可以在不需要酶催化的情况下发生的观点。他们的模型结合了细胞样结构的复制、双链聚合物和代谢。他们发现,在整个细胞周期中,聚合物只有罕见的伸长和单体浓度的振荡。

除了携带信息的RNA的复制外,绍斯塔克等人²⁸也注意到,为了成功复制,细胞需要能够生长和分裂其细胞膜。这些细胞膜还需要满足关键特性,即小分子应该能够在合理的时间尺度上扩散,而大的大分子通常应该被限制在它们的内部。此外,膜内的化学环境应该同时有利于膜的生长和RNA的复制。

乔伊斯还讨论了rna世界假说的各种潜在机制。³⁴在他看来,RNA世界假说的主要挑战是克服与益生元化学有关的混乱。他指出,尽管目前的复制是以残基对残基的方式进行的,但在益生元环境中却不一定如此。只要生产速率超过衰变速率,聚合物或细胞就会增殖。此外,即使它存在,RNA世界也不一定是生命的起点,因为它之前可能有更简单的遗传系统,例如TNA或PNA(三糖核酸或肽核酸;(见图2),然后转化为RNA世界。或者,可能存在其他允许模板化的复制机制,例如涉及无机粘土。³⁵乔伊斯考虑了RNA和蛋白质合成的潜在途径的更多细节-由RNA催化裂解。再次提出了区隔化的问题;此类反应可能发生的细胞样结构的优点之一是维持足够浓度的产物种类和反应环境的适宜性。乔伊斯指出,这些隔室不一定是现代细胞常见的磷脂双层;它们可以是空气中的气溶胶、岩石孔隙、其他多肽的b层或萜类化合物。另一个悬而未决的问题是

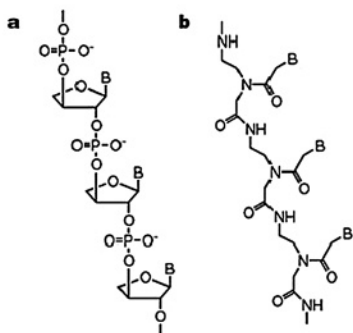


图2 (a)三磷酸核苷酸的化学结构;(b)核苷酸, 其中b为核苷酸碱基。

RNA世界是如何过渡到dna -RNA-制造-蛋白质生物化学的。

作为酶和底物的化学物质的自我复制的错误阈值由 Obermeyer和Frey³⁵通过以下形式的方程模型系统进行了分析

$$\frac{dX_i}{dt} = \sum_j M_{ij} R_j X_j - X_i \sum_j R_j X_j, \quad (1)$$

其中 M_{ij} 是突变概率, $X_j(t)$ 是化学物种 j 的浓度, $R_j = A_j + \text{PiBj}$, iX_i 是物种 j 的复制率。他们注意到非酶复制的潜力是值得怀疑的。他们发现, 为了在主序列上定位物种的分布, 只需要弱特异性。更强的特异性约束提供了更强的定位, 因此更长的序列可以被复制, 这也提供了更强的抗突变耐受性。

2.3 RNA世界的起源

艾根和舒斯特的工作在瓦蒂斯和科文尼所做的理论工作中有了新的方向。³⁶他们的研究重点是构建和分析在益生元环境下单体b- d核糖核苷酸前体产生长链RNA的微观动力学模型。瓦蒂斯和科文尼的理论是从与艾根和舒斯特相似但更普遍的化学假设出发的, 即催化复制可以产生大量的长链。这些模型包含了(i)直接链生长(ii)模板辅助合成和(iii) RNA复制酶核酶催化的可能性, 所有这些都具有不同程度的效率。然而, 所有的化学过程都是“开放的”;他们不假设存在维持长链种群的封闭超循环, 而是担心单体核苷酸最初出现一组自我维持的RNA链的可行性。他们的工作直接面对了问题的非线性特征, 这在很大程度上被艾根和舒斯特所忽视。详细的微观动力学模型导致动力学方程是贝克-杜林系统的推广, 用于团簇或聚合物链的逐步生长;它们位于一个一般的理论框架内, 科文尼和瓦蒂斯已经成功地应用于广泛的复杂化学动力学问题。³⁷

科文尼和瓦蒂斯考虑的最详细的模型包含Becker-Doering聚集项, 以及一般的Smoluchowski断裂项, 以模拟RNA聚合物的竞争性水解。³⁸最后这个过程是一个关键的过程, 它使所有的化学成分都能被循环利用, 并导致一些自我复制的序列在所有其他序列上大量扩增。这种生长链经历的主要反应是(i)控制链生长的基本贝克尔-杜林速率过程

$$C_r^\gamma + N_i \xrightleftharpoons{\text{slow}} C_{r+1}^{\gamma+N_i}. \quad (2)$$

在这里, 四个核苷酸碱基用 N_i 表示(其中 $i = 1, 2, 3$ 或 4), 寡聚核糖核苷酸序列用 C_r^γ 表示, 其中 r 表示序列中碱基的数量, γ 表示它们发生的特定顺序。(ii)基于模板的链合成(互补链上核糖核苷酸的Watson - Crick碱基配对介导的一种催化形式)

$$C_r^\gamma + N_i + C_s^\theta \xrightleftharpoons{\text{fast}} C_{r+1}^{\gamma+N_i} + C_s^\theta. \quad (3)$$

(iii)水解, 即长链被分成两条较短的链。从化学上讲, 这与该过程相对应

$$C_{r+s}^{\gamma+\theta} \rightarrow C_r^\gamma + C_s^\theta. \quad (4)$$

这是一种由斯摩鲁霍夫斯基方程模拟的一般断裂过程的形式, 这种机制增加了链的数量, 但减少了平均链的长度。(iv)酶复制(复制酶核酶活性), 其中第三条链有助于序列的生长, 该序列已经与作为模板的另一条链密切接触, 由

$$C_r^\gamma + N_i + C_{r+k}^{\gamma*+\theta*} + C_s^\xi \rightleftharpoons C_{r+1}^{\gamma+N_i} + C_{r+k}^{\gamma*+\theta*} + C_s^\xi. \quad (5)$$

这里, γ 与 N_i 的组合是链 $\gamma + y$ 的子序列, 而 C_s^ξ 则扮演复制酶核酶的角色。不用说, 其中一些复制酶的效率将比(大多数)其他复制酶高得多。

瓦蒂斯和科文尼证明, 在合理的时间内, 从对化学和初始条件的合理假设开始, 有可能实现某些自我复制的RNA聚合物链的选择, 这些假设可能在由b-D-ribonucleotide单体组成的假定的益生元汤中普遍存在。^{36,37}

2.4 rna世界元素的建模

科文尼和瓦蒂斯建立的所有模型都是基于确定性非线性动力学, 源自质量-作用的宏观定律。对于可能经常涉及细胞水平的情况, 这一“定律”是否总是有效, 绝不是明确的。在涉及少量分子的情况下, 更固有的概率模型可以发挥重要作用。因此, Hanel等人³⁹考虑了随机最小非线性模型的动力学, 其形式为: $\dot{x}_i = \sum_j P_{ij} x_j + \sum_j v_{ij} x_j$, 其中连性矩阵 A 具有Erdos-Renyi网络的形式(节点之间的边根据某些随机变量独立形成的网络)。在选择 A 为随机矩阵的同时, 他们施加了约束 x_{40} , 这使得

系统非线性, 因此Lyapunov指数可能是正的、负的或零。如果所有指数都为负, 则系统收敛到稳态或平衡解; 如果最大指数严格为正, 则系统是混沌的, 并表现出强烈的不稳定性、对初始条件的敏感依赖和其他不利于生命系统持久性的“随机”行为。然而, 如果最大指数为零, 则系统处于临界状态, 此时可以探索更大范围的状态, 非线性效应占主导地位, 有助于控制行为; 这种状态在生命系统的建模中非常有意义, 其中表现出的行为类型包括振荡。网络的平均度 $h_{ki} = L/N$, 其中 L 为链路数, N 为节点数。 h_{ki} 有相当大的范围, 其中最大特征值接近于零, 表明其适用于生命系统。

Lehman⁴⁰提出, 即使在早期的RNA世界中, 自我复制也是通过RNA链断裂成更短的链来复制子链, 重组形成原始RNA分子的多个副本。造成这种情况的原因有很多: 重组是一个能量中性的过程, 因此在益生元条件下会很常见; RNA复制酶必须很长, 通常超过100个核苷酸。它太长了, 无法通过催化准确地复制, 而较短的RNA链可以很容易地复制。这一假设的一个后果是, 随着时间的推移, 人群中的RNA会延长; 这与从简单的益生元生命到更高级功能的发展完全一致。

如前所述, 分析RNA世界模型的一个问题是可能的物种数量巨大(组合数量大)。一般来说, 没有足够的质量让所有的物种都被探索, 也就是说, 甚至每个物种的一个分子都不能被创造出来, 所以这不是一个理论上哪个具有最稳定结构的问题, 而是那些被创造出来的物种中哪个最擅长复制的问题。相反, 许多作者提倡的一个更简单的问题是考虑一个简化的物种系统, 在这个系统中, 只有两个物种在争夺一个底物。两个物种的一个很好的例子是在许多生命系统中发现的左手和右手手性结构。在化学上、物理上和热力学上, 它们具有相同的稳定性, 但竞争和其他非线性效应导致其中一个被淘汰, 另一个成为主导。因此, 在第3节中, 我们将详细关注可能导致这种对称性破坏的机制, 在两个物种(左手和右手)竞争的系统中形成手性结构, 可能也具有非手性状态。

2.5 区隔化: 自我复制的胶束和囊泡

路易斯的团队已经成功地开发出了导致自我复制囊泡的实验场景。⁴²Walde等人描述了辛酸和油酸在水中形成的囊泡进行自我繁殖的条件。它们种群数量的增加是由于在囊泡本身的球形边界内发生的反应。这是通过让辛酸或油酸酐在

碱性pH下脱氢来实现的。酸/羧酸盐浓度的初始增加是缓慢的, 但超过临界浓度的囊泡的存在带来了快速的第二相, 导致在整个自催化过程中形成更多的囊泡。Walde等人还研究了辛酸和油酸囊泡的自复制过程与温度的关系。他们表明, 通过提高温度(高达70°C), 快速囊泡形成的区域变得更加陡峭, 而漫长的初始慢相则显著缩短。

两位现任作者已经提出并分析了自我复制的囊泡模型。⁴¹Coveney和Wattis基于Becker-Doering聚类方程的新推广构建了一个动力学模型, 该模型描述了囊泡结构的逐步生长和破碎。他们的非线性动力学模型⁴¹很复杂, 涉及许多微观过程; 然而, 通过将完整的动力学方程集系统地收缩到宏观极限, 他们证明了该模型正确地捕获了实验观察到的长缓慢诱导阶段, 快速自催化阶段随后缓慢收敛到平衡的行为。Bolton和Wattis进一步推广了该模型⁴¹, 以解释Lonchin等人⁴²和Berclaz等人观察到的尺寸模板效应⁴³当研究的系统在没有囊泡的情况下开始时, 大范围的囊泡大小发展缓慢; 而在初始分布一种特定大小的囊泡(本质上是单分散的)的系统中, 相同大小的囊泡产生得更快。

3 同手性的起源 不用看这部分

发生手性对称性破缺效应的一个领域是亚原子物理的弱电相互作用中的宇称破坏。许多作者已经研究了这种能量上的微小差异(10^{15} kcal/mol)在更大的聚合或结晶水平上导致对称性破坏的机制。例如, 参见Faglioni等人的工作,⁴⁴和其中的参考文献。然而, 对于本文的其余部分, 我们不考虑这种可能性, 而是倾向于非线性效应导致外消旋态不稳定的理论, 随机波动选择一个手性而不是另一个手性, 非线性相互作用然后驱动系统达到手性纯度。

从涉及大约相等数量的左手和右手手性物质的复杂化学转变为只涉及手性分子的单手性的系统, 是生命起源的一个重要阶段。在本节中, 我们将重点讨论表现出这种对称性破坏转变的过程的数学模型。我们考虑在聚合和固体晶体形成过程中发生这种情况的系统。我们的目标是描述导致对称性破坏的机制。

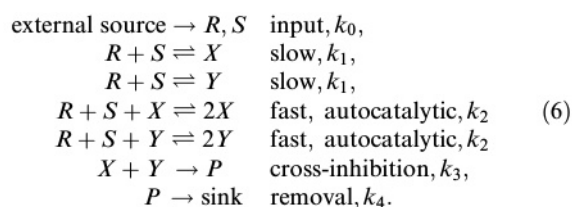
2001年, Sharpless、Noyori和Knowles因为他们在手性合成方面的工作而共同获得了诺贝尔化学奖。这项开创性的研究起源于20世纪70年代, 其中包括包括卡根在内的更大团队的工作。⁴⁵后来的重要实验证明了通过化学反应放大了手性的初始小不平衡, 包括

Soai et al.⁴⁶:在该反应中, 在实验开始时催化剂中的少量对映体过量导致反应结束时产物手性的更大不平衡。在手性噻啉烷醇的不对称自催化中, 对映体的过量达到了85%以上。然而, 手性对称性断裂过程的理论和建模历史要长得多。

3.1 历史背景

多年来, 已经提出了许多模型来解释从最初的外消旋混合物中出现的同手性。第一个是Frank⁽⁴⁷⁾的模型, 他提出两个前体R和S通过自催化反应结合形成左手或右手的物质X、Y。这样做的效果是放大任何超过Y的X或反之亦然($ee_{\text{abs}} = XY$), 两个手性产物然后结合形成产物P, P被移除;因此, 只有对映体过量的X超过Y(或反之亦然)留在系统中。这一阶段放大了相对手性, $ee_{\text{rel}} = (XY)/(X + Y)$ 。

这些过程可以总结为



我们用小写的变量来表示每种物质的浓度, 这些变量是时间的函数。为简单起见, 我们忽略可逆反应, 用质数表示时间导数。这个系统用微分方程来描述

$$\begin{aligned}
 r' = s' = k_0 - 2k_1rs - k_2rs(x + y) + k_{-1}(x + y) \\
 + k_{-2}(x^2 + y^2), \quad p' = k_3xy - k_4p,
 \end{aligned} \quad (7)$$

$$x' = k_1rs + k_2rsx - k_3xy - k_{-1}x - k_{-2}x^2, \quad (8)$$

$$y' = k_1rs + k_2rsy - k_3xy - k_{-1}y - k_{-2}y^2, \quad (9)$$

从中我们注意到, 在稳态下, 我们有

$$rs = \frac{k_0 + k_{-1}(x + y) + k_{-1}(x^2 + y^2)}{2k_1 + k_2(x + y)}. \quad (10)$$

我们把绝对对映体过量写成 $ee = x - y$, 总浓度写成 $s = x + y$; 加减 dx/dt 和 dy/dt 的方程, 我们得到 $s^2 = 2k_0/k_3 + ee^2$, 和

$$ee \left[\frac{k_2(k_{-2}ee^2 + k_{-2}s^2 + 2k_{-1}s + 2k_0)}{2(2k_1 + k_2s)} - k_{-1} - k_{-2}s \right] = 0. \quad (11)$$

注意, $ee = 0$ 总是一个解;但是, 也有其他的解, 在速率常数满足一定条件, 包括 $k_3 > k_2$, k_0 足够大的情况下, ee 不为0。

在这里, 产生不对称(即手性)解的重要特征是

(i)系统的开放和非平衡性质, 即连续输入新鲜的R, S是必要的;此外, 为

了在系统中保持有限质量的物质, 我们还要求去除产物P;

(ii)系统具有交叉抑制, 去除等量的X和Y, 放大由初始数据或输入速率的随机波动引起的任何差异;

(iii)手性物质X, Y的自催化生成。

产生不对称状态需要如此多的特殊条件, 这可以看作是Frank模型的一个弱点;因此, 其他作者试图在不做那么多假设的情况下, 产生出表现出对称性破缺的系统。

齐藤和Hyuga提出并分析了一系列“玩具”模型。⁴⁸这些模型包括非线性自催化和回收的不同强度, 总结为

$$r' = k_r(1 - r - s) - \lambda r, \quad s' = k_s(1 - r - s) - \lambda s, \quad (12)$$

其中 $r(t)$ 和 $s(t)$, 表示两种对映体的浓度。最初, 他们考虑了 $l = 0$ 时 $k_r = k_s = k$ 的情况, 发现对映体过量 rs 没有增加。在 $k_r = k_r$, $k_s = k_s$, $l = 0$ 的情况下, 相对对映体过量 $rs/(r + s)$ 是常数;虽然 $r + s$ 的绝对过量 rs 增加, 但手性产物 $r + s$ 的总量以相同的速率增加。在 $k_r = k_r^2$, $k_s = k_s^2$, $l = 0$ 的更复杂情况下, 相对对映体过剩被放大, 而这种放大的过剩在 $l = 0$ 的更一般情况下保持不变。这再次证明, 强自催化可以引起同手性化, 但尚不清楚在任何特定应用中应使用哪种形式的速率系数(k_r, k_s, l)。与Frank的模型相比, 这些模型是封闭的, 没有交叉抑制, 但自催化在对称性破坏发生之前必须是强的。

齐藤和Hyuga⁴⁸分析了描述聚合的各种方程式系统, 其中包括各种大小的簇。他们观察到, 截断四聚体的模型与截断六聚体的模型表现出不同的行为。特别是, 虽然对称破缺在六聚体模型中表现出来, 但在四聚体模型中是不可能的。这是对理论家的一个警告, 他们希望通过分析一个显著简化的版本来推断复杂模型的性质。

Goldanskii和Kuzmin⁴⁹考虑了一个“生命的冷史前”, 他们的意思是生命可能起源的各种非传统机制。这些机制包括发生在固体阶段的过程, 在寒冷的环境中, 例如在“星际尘埃颗粒的脏冰幔”中, 或者在经历一次或一系列温度升高的“热爆炸”的环境中。反应通过涉及分子隧道和“缩聚”的机制发生。因此, 有机分子是在太空或其他行星上制造出来的, 并通过陨石带到地球, 这是一种可信的可能性。在如此低的温度下, 熵不起作用, 这样的反应可能涉及隧道机制;他们声称时间尺度是数百万年的数量级。左旋和右旋结构相互转换以及与声子相互作用的量子力学模型是

由Berlin等人提出⁵⁰他们表明, 这些效应的耦合可以抑制左手和右手结构的相互转换, 从而稳定手性状态。其他具有手性产物的自催化的简单化学反应方案会破坏对称状态的稳定性, 从而导致干草叉分叉, Goldanskii和合著者后来考虑了这一点。⁵¹

鉴于生命表现出几个关键的物理性质, 即同手性、模板式自我复制和进化, Avetisov等人⁵²考虑了这些性质发生的顺序。他们认为, 由于非催化的同手性聚合几乎是不可能的, 酶的复制必须先于向手性世界的转变。然而, 他们继续承认, 结构和功能的相互作用并不是一个简单的鸡生蛋还是蛋生鸡的问题; 从简单的自催化到信息存储聚合物(如RNA)的模板化自我复制, 存在手性优势和复制水平的差异。

3.2 同手性结晶的实验观察

Kondepudi等人⁵³后来Kondepudi等人⁵³的实验表明, 在没有搅拌的情况下, 形成了大约等量的左旋和右旋晶体。然而, 当搅拌速率超过临界(阈值)值时, 每种手性晶体数量的不平衡增加; 而在足够大的搅拌速率下, 则实现了总手性纯度。因此, 从过饱和氯酸钠溶液中产生的晶体的最终手性分布可以通过将搅拌速率作为分岔参数来描述。

在较高的搅拌速率下出现同手性的原因是, 所有的晶体都来自同一个“母”晶体, 这是第一个生长到超临界尺寸的晶体, 这样当它碰到叶轮或壁面时, 碎片从它上面断裂下来, 所有其他的晶体都是从这些碎片中生长出来的(与LUCA相比)。早些时候, Kondepudi和Nelson⁵³研究了一个有点不太可能的理论, 即违反宇称的微扰可能被放大, 并导致生命前化学中的手性对称性破坏。他们的研究结果表明, 这样一个过程大约需要 10^{-4} 年。Avetisov等人对固体的手性结构进行了计算机模拟,⁵⁴得出了对同手性结构域形成的现象学描述。

维德玛首先注意到, 研磨手性晶体的混合物最终会导致晶体的同手性分布(也就是说, 所有的晶体都具有相同的手性)。⁵⁵他跟随Kondepudi等人⁵³, 使用保持过饱和状态的氯酸钠晶体, 用搅拌器和小玻璃球不断粉碎晶体。随着时间的推移, 左手型和右手型晶体的比例稳定地从大约外消旋型转变为完全的同手性。

Noorduyn等人⁵⁶对氨基酸也观察到类似的效果; 这些分子与生命起源研究的相关性要大得多。McBride和Tully对Noorduyn等人和维德玛的研究成果进行了综述。⁵⁷研磨是一个动态过程, 通过溶解和再结晶, 将一种手性的固体转化为另一种手性。一旦系统处于手性纯

净状态, 系统就被“锁定”在那个状态: 极不可能形成任何另一种手性的核。手性纯态并不是平衡态, 相反, 研磨的施加意味着能量不断被输入系统, 这破坏了平衡构型的稳定, 反而稳定了不对称的稳态。

McBride和Tully⁵⁷将一种对映异构体的生长和另一种对映异构体的溶解描述为一种奥斯特瓦尔德成熟过程; 其含义是, 不对称稳态的稳定是由于表面积: 体积或尺寸: 曲率比的变化。虽然小晶体由于其较大的表面积与体积比而具有快速的溶解速率, 但较大的晶体具有较低的表面积与体积比, 这意味着相对而言, 它们溶解得较慢。奥斯特瓦尔德成熟的发生是由于溶解与生长的比率, 这取决于晶体的曲率, 并且由于晶体通常不是球形的, 因此它以复杂的方式随晶体尺寸的变化而变化。然而, 下面分析的模型表明, 没有必要调用这种尺寸依赖效应来解释研磨如何导致同手性化。

3.3 通过研磨对同手性进行建模

同样, 在任何实验证据发表之前, 乌瓦哈写下了一个描述上述详细过程的粗略模型。⁵⁸我们用 $c(t)$ 表示无定形(或非手性)单体的随时间变化的浓度, 用 $x_i(t)$ 、 $y_i(t)$ 分别表示小的左(右)旋簇, 用 $x_2(t)$ 、 $y_2(t)$ 表示较大的左(右)旋晶体。

这些速率方程模拟了以下过程:

- (i) 小手性结晶团(x_1, y_1)从非手性单体池中成核(c)速率 k_0 ;
 - (ii) 相同手性的小手性核(x_1, y_1)以 k_c 的速率结合成更大的手性团簇(x_2, y_2)
 - (iii) 相同手性的小簇和大簇结合成更大的簇, 速率为 k_u ;
 - (iv) 通过与非手性单体结合生长的大晶体(x_2, y_2)(c)以 k_1 的速率生长
 - (v) 较大的晶体(x_2, y_2)以 λ_u 的速率溶解成较小的簇(x_1, y_1)
 - (vi) 较小的核(x_1, y_1)以 λ_0 的速率溶解成非手性单体
 - (vii) 较大的晶体(x_2, y_2)以 λ_1 的速率溶解为非手性单体
- 根据这些量, 乌瓦哈的化学反应模型有这样的形式

$$c' = -2k_0c^2 - k_1c(x_1 + y_1) + \lambda_1(x_2 + y_2) + \lambda_0(x_1 + y_1), \quad (13)$$

$$x_1' = k_0c^2 - k_u x_1 x_2 - k_c x_1^2 + \lambda_u x_2 + \lambda_0 x_1, \quad (14)$$

$$x_2' = k_1 x_2 c + k_u x_1 x_2 + k_c x_1^2 - \lambda_1 x_2 - \lambda_u x_2, \quad (15)$$

$$y_1' = k_0c^2 - k_u y_1 y_2 - k_c y_1^2 + \lambda_u y_2 + \lambda_0 y_1, \quad (16)$$

$$y_2' = k_1 y_2 c + k_u y_1 y_2 + k_c y_1^2 - \lambda_1 y_2 - \lambda_u y_2, \quad (17)$$

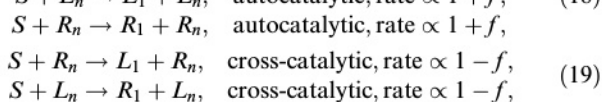
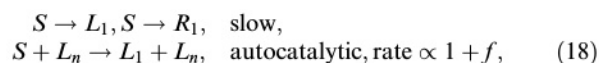
由于所有变量描述的都是每种状态下物质的总质量, 所以这个模型中没有纳入簇的大小。

因此,分析中不涉及表面积体积比。然而,在这个模型中,对称破缺到 x_1a x_2 和 x_2a y_2 的解仍然是可能的。乌瓦哈指出,涉及 l 的术语描述了物质的再循环,对形成完全的同手性状态至关重要。当系统处于手性不对称状态时,左手晶体和右手晶体的典型尺寸可能会有很大的不同。虽然不期望该模型在定量上准确,但它仍然应该能很好地定性描述行为。

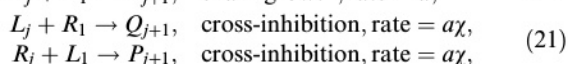
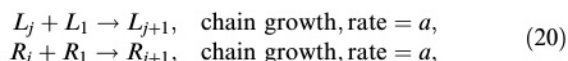
Wattis⁵⁹引用了一系列从详细的微观模型开始的模型,其中包括标准晶体生长和溶解。这个模型被推广到包括研磨的影响,研磨会去除较大的晶体碎片,并且有三种晶体同时生长,即无定形、左旋和右旋晶体。左手和右手晶体没有直接的相互作用(没有交叉抑制),尽管它们竞争非手性单体和小晶核,这些小晶核在聚并时具有较大晶体的手性。然后对模型进行简化,以显示研磨、逐步生长和非晶态材料竞争的基本过程如何导致对称性破坏。除了每个物种的尺寸分布的微观模型外,还使用各种粗粒化程序来推导出晶体总浓度、每种手性晶体材料的总质量和每种类型的小核数量的常微分方程的小系统。

3.4 通过催化反馈进行的同手性聚合

Sandars在2003年提出了一种有趣而新颖的手性聚合物形成对称破缺模型。⁶⁰这个玩具模型与Frank 1953年关于自发不对称合成的论文有许多共同的特点⁴⁷(参见3.1节),但它适用于早期的肽世界。在这个模型中,一种非手性底物(S)会自发地、缓慢地分裂成手性单体 L_1, R_1 ,并以更快的速度被聚合物催化。正如最初提出的那样,该模型具有最大聚合物长度(n),并且只有最长的同手性链能起到催化剂的作用。其催化作用既有自催化作用,也有交叉催化作用,具体如下:长左手链 L_n 催化 S 生成左手单体 L_1 ,同时也生成右手单体 R_1



聚合被认为是未催化的,因此是线性的;交叉抑制是对称断裂所需的另一种机制,通过允许将相反手性的单体添加到生长链中来结合。这被认为是为了防止链的进一步生长。因此



当 $j = 1, 2, \dots, n$ 时,分析模型的目的是找出参数 k, a, w, f 破坏对称稳态的条件。

随着自催化与交叉催化比例的增加,即通过增加保真度参数 f ,得到干草叉分叉(在 $f=f_c$ 处)。干草叉的位置可以出奇地低,例如,在 w 大的情况下, $f_c = 0.2$ 。更大的交叉抑制(更大的 w)减少 f_c 的事实,乍一看,是相当违反直觉的。将这种行为合理化的一种方法是注意到,如果要使一种手性支配另一种手性,那么更重要的是阻止另一种手性的链的生长,而不是生长更长的占主导地位的手性链。

其他相关但独立的工作来自Brandenburg等人关于早期肽世界的同手性⁶¹和Wattis和Coveney。⁶²在后一项工作中,假设催化步骤对所有聚合物长度(n, A, n, j, A, n)都有效,从而可以分析确定对称溶液的稳定性标准。Wattis和Coveney还分析了聚合物形成的动力学,在这种情况下,系统从零聚合物浓度开始,然后在短时间内形成聚合物。在这段时间内,随机外部扰动引入系统的不对称性可能会被大量放大。然后,系统接近其中一种稳态解,其中的产物绝大多数是同手性的。

Brandenburg等⁶³考虑了Sandars模型的不同变体,其中一些反应是可逆的。他们加入了 $L_n \cdot L_m + L_{n-m}$ 和 $L_n R_1 - L_m R_1 + L_{n-m}$ 。这种体系的问题在于,它们主要由短聚合物组成,因此平均聚合物长度短得不切实际。因此,这些作者考虑了其他可能的修饰,其中片段立即被回收成单体。最终,他们发现了与Sandars原始模型相似的结果,并且随着 f 的增加(约为0.5),干草叉分叉发生,导致不对称稳态。

Blanco和Hochberg⁶⁴将可逆性添加到Sandars模型中的所有反应中,并分析了反应每个阶段的熵产。他们发现,在接近稳态或平衡构型之前,系统会经历手性的振荡。当系统打破对称性时(从最初几乎外消旋的状态),或者在其手性中经历反转时,熵的产生率达到最大值。

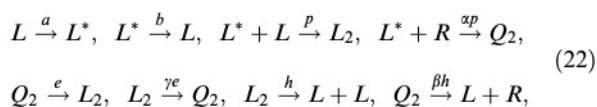
Gleiser et al.⁶⁵将Sandars的模型推广到一个空间域(二维和三维),方法是用扩散算子($d/dt - q/(\eta t - k, \eta)$)替换控制方程中的时间导数,并加入高斯白噪声来模拟随机波动对系统的影响。他们发现,随机初始条件导致粗化行为,其同手性区域通过表面张力驱动的动力学生长和进化。他们还发现了间断平衡,其中噪声可以抹去系统对其手性历史的记忆,并建立新的手性,从而扩展了Eldredge和Gould关于进化中的间断平衡的概念。⁶⁶

3.5 通过外映化实现的同手性聚合

Plasson等人也试图放宽Frank的假设,特别是他们想要一个封闭和

没有自催化作用。他们的激活聚合表观聚合解聚合(ape)方案⁶⁷具有每种手性(L, R)的单体, 可以通过向系统输入能量(到L*, R*)来激活。激活、聚合、外聚和解聚合这四个过程的速率取决于它们是作用于均聚物还是异聚物。活化的单体可以与未活化的单体结合形成二聚体(polymerisation, L₂, R₂, Q₂, 前两者为均聚物, 后者为异聚物)。在解聚过程中, 聚合物被分解成它们的组成单体。所有这些过程都使每种手性的分子总数保持不变。因此, 出现了一种新的过程, 外映化, 在这种过程中, 聚合物中的最终单体可以改变其手性; 这将均聚物转化为异聚物, 反之亦然, 并且可以被视为先前模型中自催化的替代品。根据它们是作用于均聚物还是异聚物, 过程具有不同的速率。

Plasson等人的方案中的过程可以总结为



类似的反应用R代替L, 反之亦然。虽然这个方案很容易扩展到任意长度的聚合物, 但Plasson的分析仅限于单体和二聚体的情况, 并得出了一个有趣的结果, 即在低总质量下, 只有“死”状态存在, 其中大多数物质以单体形式存在, 所有单体都被激活。在更大的质量下还存在另外三种状态: 总的来说, 系统是不稳定的, 在两种手性之间振荡; 在a的中间值处, 对称态是稳定的; 在a值较低时, 对称状态是不稳定的, 有两个稳定的非对称状态, 系统接近其中一个。因此, Plasson等人成功地创造了一个封闭的体系, 没有催化, 也没有同手性产物的交叉抑制。同手性化所需的正反馈是由外聚化而不是自催化提供的。在较高的总质量下, 外聚化的总体平均效应有利于物质从少数手性转化为优势对映体, 破坏外消旋(对称)稳态。

Gleiser和Walker对Plasson等人的分析进行了扩展, 以研究模型中空间变化形式的扩散效应。⁶⁸他们的动机是研究这种模型是否可以描述原始细胞的形成, 即包围特定手性区域的细胞膜, 从而形成与周围环境在热力学上分离的结构。Gleiser和Walker考虑了系统处处过饱和的齐次初始条件和只有局部空间域过饱和的系统。他们发现环状结构形成时, 内部是同手性的, 壁是外消旋的, 具有高浓度的异二聚体。虽然原始细胞具有许多生命系统共同的特征(生长、代谢、繁殖、对环境变化做出反应的能力), 但它

们只是活细胞的一个组成部分。Gleiser和Walker承认, 他们的系统动力学不能描述新陈代谢或繁殖。尽管如此, 他们的论文证实, 细胞样结构可能先于自我复制的遗传分子, 这对更广泛的关于谁先出现的争论做出了重要贡献。

3.6 热液孔隙系统中核苷酸的积累

RNA世界需要高浓度的小的生命前分子, 而地球化学推断表明存在一个浓度与现代值相当的稀释的生命前海洋。⁶⁹Braun等人通过联合实验和模拟提出了一种潜在的解决方案来解决这种浓度差异。⁷⁰利用有限元求解器对Navier-Stokes方程、分子扩散和传热的组合解进行二维模拟, 以模拟受热梯度影响的细长热液孔隙系统中的分子运输。他们发现, 在各种各样的堵塞孔隙中存在分子的极端积累。作者指出, 该机制能够提供高度浓缩的单核苷酸, 适合在生命起源时RNA世界的运作。⁷⁰积累是由孔内的热梯度驱动的。流体通过热对流沿孔隙穿梭, 而分子则通过热扩散驱动在孔隙中漂移。Baaske等人的研究表明, 与没有热扩散的情况相比, 毫米大小的孔隙甚至比10⁻⁸-fold更多地将单个核苷酸积聚到微米大小的区域。在靠近孔隙封闭末端的散装水区域中发现了分子浓度的增强。由于积累指数依赖于孔隙长度和温差, 因此相对于裂隙几何形状和分子尺寸的变化, 它是稳健的。Baaske等人的研究结果表明, 生命的进化最初并不需要复杂的主动膜运输。相反, 热梯度中相互连接的矿物孔隙可以为生命的分子进化提供一个高浓度的起点。

4 分子模型

在前面的章节中, 我们一直关注研究一系列化学动力学模型, 这些模型主要基于质量作用的宏观定律。现在我们转而考虑更微观的分子理论对生命起源的看法。迄今为止, 生命起源研究很少使用基于计算机模拟的分子建模技术来理解第一个生物分子形成的可能化学途径。模拟方法提供了关于系统的几乎无限的分子信息, 这些系统由于缺乏测量分辨率和/或与这些生命起源系统相关的极端温度和压力条件而难以用实验来表征。然而, 使用第一性原理技术计算这些生命起源系统的反应机制和动力学仍然非常困难。在本节中, 我们将从生命起源研究的角度考虑所使用的计算方法和从这些分子模拟技术中获得的科学见解。

4.1 分子模拟方法

这里我们对所使用的计算技术做一个简短的总结, 以便让普通读者熟悉这些方法。分子模拟方法通常根据它们计算原子间相互作用的准确性以及它们产生的结构数据的类型进行分类。

4.1.1 量子化学方法。量子化学模拟方法通过尝试求解量子力学方程(基于薛定谔方程)来模拟电子与原子系统的原子核之间的相互作用。量子化学方法使我们能够在诸如成键和断键等过程中对电动力学进行建模, 因为它们明确地表示电子。量子方法的优点之一是, 除了电子的质量和电荷等基本常数外, 它们只需要输入原子序数、原子核的初始构型和电子数。量子方法的主要缺点是它们在计算上非常昂贵, 这限制了目前研究几百个原子组装的方法。关于量子化学方法的很好的介绍可以在参考文献71中找到。

量子化学计算, 有时被称为从头计算或“第一性原理”计算, 在凝聚态中经常使用密度泛函理论(DFT)。DFT使用Hohenberg - Kohn定理, 该定理指出分子或固体的电子基态是电子密度的唯一泛函。该泛函由动能、库仑能、电子交换能和相关能等项组成。

有机分子和无机矿物之间相互作用的量子化学计算, 就像生命起源研究中考虑的那样, 提供了精确的能量、电荷分布和几何形状, 但即使使用大型超级计算机, 这些模型也只能在几皮秒内模拟数百个原子。由于模型大小和模拟长度的限制, 经典的分子模拟技术方法(见第4.1.2节)更适合研究大型生物分子的动态相互作用。经典分子模拟技术的准确性在很大程度上取决于所使用的力场的有效性。从头算可以用来确定这种经典分子动力学模拟的精确力场。

4.1.2 分子动力学。经典分子动力学技术是以经典力学为基础的。每个原子或原子群被视为一个单一的实体, 它们之间的力用势能梯度来表示。分子动力学既需要输入初始原子位置, 也需要输入力场参数。这些参数可以从实验中得到, 也可以从4.1.1节讨论的从头计算中得到。分子动力学模拟非常适合于模拟由键和非键相互作用控制但没有任何化学转化的系统。原子间电势的使用意味着目前可以模拟数百万个原子, 通常是几十纳秒, 但在某些情况

下长达微秒。在ref. 72中可以找到对分子动力学领域的很好的介绍。

分子动力学(Molecular dynamics, MD)详细描述了分子系统内原子的个体运动和集体运动。MD尝试在已知系统起始结构的情况下, 对所有相关参数进行从头预测。有了足够的计算能力, MD能够计算经典多体系统的平衡描述, 但它也明显地提供了动态信息。假设系统中的原子在原子间势能的影响下运动。模拟中的每个原子都有质量和大小。将牛顿的运动方程整合起来, 计算组成系统的N个原子的动力学行为。因此

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i, \quad (23)$$

其中 m_i 是原子i的质量, \mathbf{r}_i 是原子i的笛卡尔3向量, \mathbf{r}_i 的二阶导数是原子i的加速度, \mathbf{F}_i 是作用在原子i上的力向量。力 \mathbf{F}_i 通过方程 $\mathbf{F}_i = -\nabla V$ 与多体势 V 相关, 其中 ∇ 表示梯度算子。我们只需要根据一组初始原子坐标来指定系统的初始构型, 同时根据麦克斯韦-玻尔兹曼速度分布随机分配初始原子速度。分子动力学的主要限制之一是它所能达到的时间尺度。由于数值稳定性的原因, MD模拟中典型的时间步长不能超过几飞秒, 因此必须集成数十亿个时间步长才能达到纳秒域。

在MD的背景下, 力场提供了用于描述粒子系统势能的函数形式和参数。力场函数和参数集通常来自实验和计算昂贵的量子力学计算。“全原子”力场, 如用于描述核酸和蛋白质相互作用的AMBER力场⁽⁷³⁾和用于描述粘土和无机矿物相互作用的ClayFF力场⁽⁷⁴⁾, 为系统中所有原子提供参数。

4.1.3 计算需求

为了进行与生命起源相关的分子模拟, 我们需要模拟现实尺寸模型;虽然今天经常可能只有数万个原子大小, 但它们可能达到数十万到数百万个原子。我们的目标是在一个可接受的时钟时间内模拟这些系统。因此, 为了执行逼真的模拟, 我们需要在强大的超级计算机上运行具有出色加速和缩放特性的代码来运行模拟。经典的分子动力学代码, 如LAMMPS, ⁷⁵NAMD⁷⁶和GROMACS⁷⁷, 随着用于执行计算的内核数量和模拟尺寸不断增加的系统的功能, 显示出近似线性的加速。

4.1.4 粘土和生物聚合物。大多数粘土由连续的二维薄片以共面方式堆叠而成, 具有随机的面内相。⁷⁸各种类型的粘土是

其特征是薄片中的原子种类不同, 层间区域中的客体种类不同。蒙脱石类阳离子粘土由四面体和八面体铝硅酸盐片的组合组成。当二价离子存在于八面体薄片时, 所有的八面体位点都被填满, 形成三八面体粘土。如果三价离子存在, 则三个位点中只有两个被占据, 从而得到二八面体粘土。同形取代进一步使粘土多样化。最常见的是四面体层中的硅和八面体层中的铝被不同价的原子部分取代。这些取代在粘土层上产生净负电荷, 通过将电荷平衡阳离子吸附到层间来平衡。阴离子粘土(更准确地说是层状双氢氧化物)在结构上与水镁石有关。水镁石由羟基配位的八面体 Mg^{2+} 离子组成。在大多数阴离子粘土中, Mg^{2+} 被 Al^{3+} 等 M^{3+} 离子部分取代, 使粘土层带正电。阴离子粘土的电荷由层与层之间的阴离子平衡。

文献中粘土矿物的建模主要由基于力场的经典分子动力学模拟组成。利用ClayFF力场,⁷⁴ Kirkpatrick等人通过分子动力学模拟解释了层状双氢氧化物(LDHs)的NMR和FTIR数据。⁷⁹其他人, 包括Coveney和Skipper, 已经研究了蒙脱石粘土中水和阳离子的层间结构。⁸⁰Greenwell等人发表了一篇关于模拟技术在粘土矿物中的应用的综述。⁸¹

4.2 量子力学模拟在生命起源研究中的应用

正如前几节所讨论的, 量子力学(QM)处理提供了以高精度模拟少量原子的方法。QM研究也被称为从头算或第一性原理方法, 因为与分子力学不同, 它们从解决感兴趣的系统的基态电子结构开始, 除了存在的原子的电子构型和原子质量之外, 基本上完全没有参数。这些方法提供了对键形成机制的独特见解, 或者是依赖于局部电子结构微小差异的详细结构和光谱方面。因此, QM在起源研究中的主要应用是探测矿物的催化作用(主要是矿物表面介导的)和手性选择性。为了使QM计算易于处理并适用于扩展结构(如矿物), 通常在所有三个空间方向上施加周期性边界条件, 以便在所有空间中数学上重复一个小的模拟单元, 以近似一个无限系统。到目前为止, 最常用的QM技术是密度泛函理论(DFT)方法。

4.2.1 矿物表面的反应性。与生命起源研究中的其他研究领域相比, QM模拟的大部分重点都放在理解矿物质催化肽键形成的过程上, 许多研究小组选择模拟这条通往蛋白质世界的道路, 而不是那些倾向于RNA世界的方法。造成这种偏差的一个原因是,

肽形成反应涉及相对简单的机制, 反应物较少, 因此与核酸合成中涉及的许多反应相比, 在计算上是可处理的。Lambert⁽⁸²⁾对矿物质上的肽形成进行了全面的回顾, 包括实验和分子模型研究, 他确定了1988年至2007年间在硅酸盐、二氧化钛、粘土和黄铁矿表面进行的一系列模拟研究。

多肽形成反应需要脱水 and 缩合机制。可以很容易地理解, 这对稀溶液中的氨基酸是非常不利的。因此, 矿物表面³和浓缩盐溶液(盐诱导的肽形成, SIPF⁽⁸³⁾)已经被假设和实验证明可以促进肽键的形成, 特别是当这些与湿/干循环耦合时, 如在假定的冥古宙地球上的潮汐泻湖中似乎是合理的。值得注意的是, SIPF的整个研究领域源于对氯化钠溶液脱水的蒙特卡罗计算机模拟, 结果显示在3M浓度以上的钠离子具有不饱和的内水化壳。⁸³

矿物诱导的多肽形成也有着悠久的历史。从20世纪70年代中期开始, Lahav和他的同事们就对各种激活氨基酸的方法进行了实验。⁸⁴然后由low等人进行了计算研究, 他们使用QM(非dft)方法来观察这些激活分子的影响(H_3PO_4 , H_2SO_4), 并包括原子簇来表示粘土边缘($\text{Al}(\text{OH})_4$ 和 $\text{Si}(\text{OH})_4$)。⁸⁵

Aquino等人观察到在许多关于粘土矿物和氧化铝催化肽键形成的实验报告中缺乏机制细节, 他们使用DFT研究酰胺键的形成, 选择包括显式和隐式水分子, 尽管以较简单的氨基酸醋酸($\text{CH}_3\text{CO}_2\text{H}$)和甲胺(CH_3NH_2)作为反应物来近似氨基酸。⁸⁶作者扩展了low等人的早期研究, 利用不断提高的计算能力来模拟各种刘易斯和Brønsted酸位点, 包括 $\text{Al}(\text{OH})_3$, $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$, H^+ , H_3O^+ 和 H_2O , 从而计算具有相应热力学输出的活化能。一个关键的结论是刘易斯酸位点的重要性和pH值的影响, 作者认为这是该领域未来发展的一个挑战。

在Rode等人对SIPF的开创性实验工作⁸⁷的基础上, Rimola等人⁸⁸使用DFT模拟方法确定, 具有明确水分子的溶剂化 Cu^{2+} 阳离子将二甘氨酸形成肽键的激活屏障从55 kJmol^{-1} 降低到20 kJmol^{-1} , 因此有利于肽形成的正向脱水反应。

4.2.2 水溶液中氨基酸的反应性。除了在各种矿物表面上关于益生元肽键形成的新兴文献外, Nair等人利用第一线分子动力学原理表明, 活化剂, 如羰基硫化物(COS), 在非平衡高压和温度条件下, 就在水的临界点以下, 也可以介导肽键的形成, 而无需额外调用

矿物表面, 虽然这并不能完全解决所讨论的浓度问题。⁸⁹这些过程的实验证据已经由Leman等人在COS活化的情况下提供,⁹⁰和Huber和Wa¨chtershau¨ser在热液条件下提供。²⁰作者确定了该反应的关键限制条件, 包括:

(i) 极端条件稳定所需的反应物的中性形式, 从而提高在中性pH值条件下的潜在反应性;

(ii) 高温确保了热反应性的增加;

(iii) 高压水将选择性改变为共合反应而不是分步反应途径。

4.2.3 矿物中的手性选择性。DFT可以增加洞察力的一个关键领域是理解手性原生物分子和生物分子在手性表面上的特定结合相互作用, 以阐明初始不对称不平衡可能导致动力学分解过程的地方。Yu等人利用DFT进行的早期研究之一⁹¹探索了一种模型二肽的对映体的结构, 该对映体与一种富铁粘土——非铁土的中间层相互作用。发现丙氨酸(Ala)的二肽根据对映体表现出不同的结构和不同的结合能。1-ala-D-矿物体系比相应的D-Ala体系稳定6 kcalmol⁻¹, 两种对映体-矿物体系之间的构象结构差异很容易识别。作者报告了粘土结构的重复基序与二肽的结构特征的相容性。这样的基序也被Greenwell和Coveney发现在其他层状矿物-生物分子系统中很重要。⁹²

4.3 矿物表面的催化性质

如第4.1.2节所述, Mathew和luthy - schulten利用分子动力学研究了一种涉及粘土蒙脱土及其在活化核苷酸形成寡核苷酸中的催化作用的生命起源设想。⁹³他们的模拟提供了多核苷酸形成之前的反应物构象的原子细节, 为Huang等人报道的实验观察提供了见解⁹⁴模拟阐明了金属离子的催化作用, 证明了导致正确连接的反应主要发生在中间层中, 并解释了观察到的链伸长中的序列选择性。Mathew等人进一步比较了涉及L手性和D手性反应物的反应概率, 发现蒙脱土催化时, 同手性产物比异手性产物更强。模拟结果证实, 插入时寡核苷酸的合成应沿着3'-0-5'-0方向进行, 就像RNA聚合酶中模板定向合成一样。有趣的是, 就生命起源研究而言, Mathew和luthy - schulten的模拟表明, 3'-0-50的区域选择性比竞争性的2'-0-50键形成的区域选择性更高, 而且当反应发生在中间层时, 催化效应总体上有所增强。

4.4 粘土矿物对RNA折叠的影响

Swadling等人利用分子动力学对散装水中和含水蒙脱土表面的各种RN

A进行了大规模模拟。⁹⁵Swadling等人报道的模拟阐明了核酸-粘土相互作用和反应性的许多方面, 这些方面在以前的实验研究中已经观察到, 迄今为止在分子水平上仍未得到解释。在生命起源研究中, 一个流行的理论是, 在益生元条件下的RNA可能被限制, 就像今天活细胞的核糖体中的功能性RNA一样。由于自然界中无处不在的粘土矿物及其相关阳离子的存在, 这些水环境可能为新生的核酸聚合物提供了益生元栖息地。事实上, 粘土可能因此在这些聚合物和/或其前体的形成和保存中发挥了关键作用。Swadling等人的模拟表明, 通过对RNA动力学的详细分析, RNA与这些表面相互作用的行为与散装水中的行为显着不同。更具体地说, 他们的模拟表明粘土表面、RNA和阳离子之间存在强大的静电力。模拟结果表明, 二价阳离子比一价阳离子更有效地介导表面与RNA之间的相互作用。RNA链的静电吸引和拴在表面上导致它被钉住, 因此与散装水中自由的RNA相比, 它的运动受到限制。此外, 作者发现, 当蒙脱土表面存在时, RNA折叠的速度大大提高, 与在大量水中相比。这种折叠很重要, 因为早期的RNA聚合物需要产生明确的折叠来支持酶活性, 以催化RNA链的自我复制, 这是分子进化的基本步骤。

4.5 游离和矿物保护核酸的稳定性

尽管深海热液喷口作为最早生命形式的可能来源引起了人们的特别兴趣, 但仍然存在一个问题, 即RNA等生物聚合物如何在这些喷口周围的高温高压下保持完整。一种可能的解释是, 粘土状颗粒可能在核酸形成后起到支撑和保护核酸的结构作用。实验表明, 在类似海底热液喷口的条件下, 甲醇与蒙脱石(如蒙脱石)发生反应时, 会形成烷烃。虽然层状双氢氧化物(LDHs)在当今时代并不像现在这样自然地广泛存在, 但有证据表明, 在地球早期, 即太古宙时代, 由于大气中缺乏氧气, 绿锈[Fe₂Fe^{III}(OH)₆]ClH₂O等矿物质可能更为普遍⁹⁶

由于层状双氢氧化物(LDHs)层间区域的纳米级尺寸, 嵌入的DNA的确切构象很难通过实验来阐明。Thyveetil等人使用在超级计算网格上执行的分子动力学技术, 对插入镁铝LDH内的双链、线性和质粒DNA进行了大规模模拟(最长可达480碱基对)(见图3)⁹⁶

Thyveetil等人的模型被发现与实验观察一致, 根据实验观察, 水合作用是决定DNA结构稳定性的关键因素。

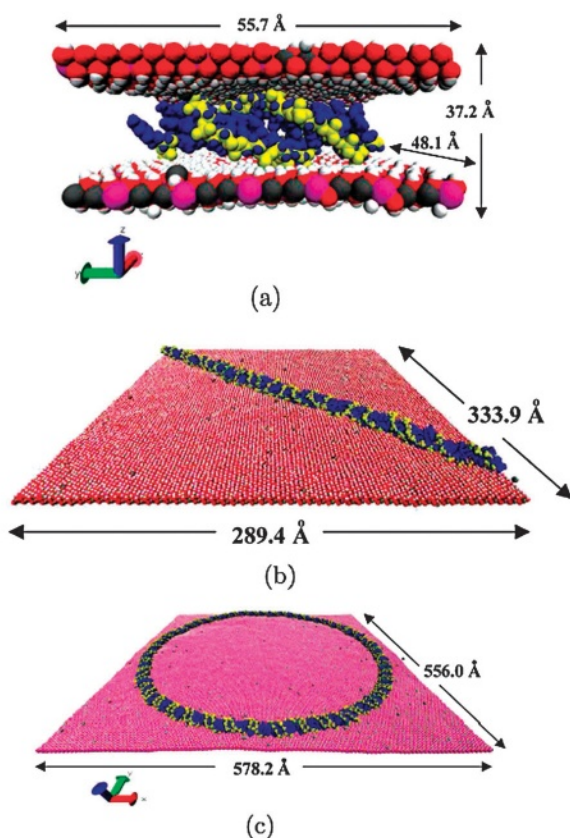


图3 (a)插入到LDH中的12碱基对DNA的分子动力学模拟快照可视化;(b)插入到LDH中的108bp DNA;(c)在环境条件下(300 K和1 atm),将480bp的质粒插入到LDH中。LDH片中的镁、铝、氧原子和氢原子分别用灰色、粉红色、红色和白色球体表示。DNA链用黄色表示磷酸主链,用蓝色表示糖基和碱基对。水分子没有显示,(b)和(c)只显示了一张LDH表。图摘自参考文献96,已获得许可。

磷酸盐主基显示与铝晶格位置对齐。在与生命起源研究相关的高温高压环境下,一些研究认为最早的生命形式起源于深海热液喷口附近(见第1节),与大量水中的DNA相比,嵌入ldh的DNA结构稳定性大大增强。Thyveetil等人的模拟使用高温和高压来观察在这些条件下插入时DNA的结构和稳定性是如何改变的。模拟中温度和压力的增加证实了DNA一旦插入是稳定的,因为在类似的条件下,DNA在大量水中的沃森-克里克氢键的数量会迅速降低。

在第二组模拟中,Swadling等人研究了三种不同的核酸嵌入镁铝层状双氢氧化物矿物的结构稳定性,在不同程度的水化作用下,在水溶液中游离。⁹⁷Swadling等研究的核苷酸为核糖核酸(RNA)、脱氧核糖核酸(DNA)和肽核酸(PNA);见图4。模拟表明,DNA增强了沃森-克里克氢

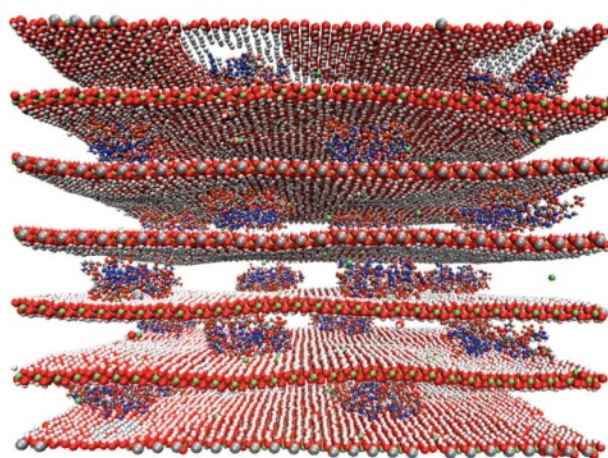


图4全原子MD模拟20ns后, Mg_2Al LDH插入24个25mer RNA核酸。C、N、O、H、Mg和Al原子分别以灰色、蓝色、红色、白色、绿色和银色球体表示。转载自ref. 97, 已获得许可。

与嵌入RNA和PNA相比,当嵌入LDH粘土层时,它们之间的键合效果更好。模拟还证实,水合作用在确定三种嵌入核苷酸的结构稳定性方面起着重要作用。关于沃森-克里克氢键,与插入RNA和PNA双链相比,DNA双链保留了更大的结构完整性,后者在碱基配对中表现出明显的降解。当不带电的PNA蛋白主链作为双相插入时,会对聚合物的整体稳定性产生不利影响,因为它与带电的层状双氢氧化物片的静电相互作用大大减少。Swadling等人报告的结果表明,以矿物为基础的生命起源很可能与在生命起源场景中更常见的以水为基础的水起源大不相同,在这种情况下DNA是最稳定的遗传物质。因此,人们可能会得出这样的结论:矿物导的生命起源可能更倾向于DNA作为信息存储的生物分子,而不是竞争的RNA和PNA,为生物从RNA世界进化提供了动力。

4.6 宿主层柔韧性在DNA客体嵌入中的作用

Thyveetil等人进行的另一组模拟使用了相同的分子动力学技术来研究层状双氢氧化物如何形成分阶段的中间结构。⁹⁸这些分阶段的中间体已经在实验中观察到,尽管它们的形成机制尚未确定。Thyveetil等人表明,ldh具有足够的灵活性,可以在DNA等大体积插入物周围形成波纹。模拟探索了DNA嵌入 Mg_2Al -LDH过程中可能形成的三种可能的中间结构,以及这些模型在能量上的差异。结果表明,当DNA链彼此直接堆叠时,LDH系统具有比交错堆叠或层间堆叠时更高的势能。模拟结果表明,平均而言,dauma - he罗尔德构型的DNA链的扩散系数比Ru道夫构型的要大

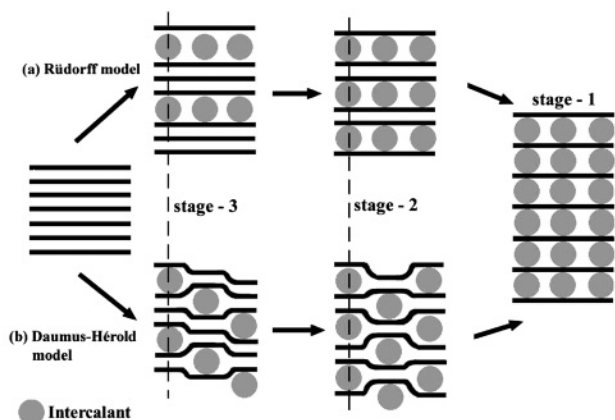


图5层状材料嵌入离子种的可能途径示意图。ldh被认为通过Rüdorff模型形成第2阶段中间体(a)。更灵活的材料,如石墨,被认为遵循Daumas-Hérold途径(b)。Thyvetil等人的模拟表明, $\text{Mg}_2\text{Al}(\text{OH})_6$ 能够由于插入DNA和氯离子之间的尺寸差异而扭曲,支持Daumas-Hérold途径插入大分子。虚线表示两种途径中相似的分阶段区域。图摘自参考文献98, 经许可。

模型和阶段1结构-图5显示了所描述的各种阶段结构的示意图。

与Rüdorff和阶段1结构相比,蠕虫模式在daumas-Hérold结构中更为突出,并支持一种机制,通过这种机制,大块的插入分子(如DNA)在中间层中迅速扩散。

5 结论

在这篇综述中,我们讨论了化学动力学和分子建模方法,这些方法现在对与地球上(以及宇宙中其他地方)生命起源相关的许多具有挑战性的问题提供了相当大的帮助。可用的方法功能强大且范围广泛。它们跨越了大量的长度和时间尺度,从电子动力学的量子力学描述,到经典(牛顿)力学最常描述的原子和分子水平,再到更大的介观和宏观水平,这些水平代表了更大的反应和自我复制分子的集体动力学行为。随着该领域继续从科学机构中获得尊重,我们预计这些理论方法和相关实验研究之间会有更强的相互作用,这只能巩固我们对地球上生命诞生之初发生的重大事件的理解的进一步进展。

建模已经成为一种被接受的益生元化学方法的一个领域,并获得了一些最伟大的喝彩,这是在对称破坏、囊泡形成和RNA世界起源等过程的动力学建模中。这些过程的动力学本质上是复杂的,不容易接受基于还原论的简单解释。动力学模型已经说

明,例如,一个对映体的小过量与手性放大反馈回路耦合可能导致在比先前假设的时间尺度短得多的时间尺度上产生同手性,从而避免了在大规模扩展的时间尺度上操作不利的手性选择性过程的需要。

建模方法的内在价值怎么估计也不为过。众多实验组的研究表明,益生元化学家感兴趣的许多反应涉及有序矿物催化表面上反应物分子的无序对或基团之间的反应。这样的反应不适合直接的化学或物理分析,尤其是当一些反应发生在催化矿物宿主深处的纳米孔或微孔环境中时。在这里,使用计算化学方法的模拟增加了对粘土表面肽键形成机制和能量学的见解;硫化铁系统中的一氧化碳聚合;粘土夹层内RNA的性质;以及核酸在粘土表面的稳定性和折叠动力学。目前,矿物界面上氨基酸活化和肽形成的详细量子化学研究与矿物-核酸系统的大规模分子模拟之间存在着某种不连续性。

最近计算能力的进步使得对更复杂的化学系统的大规模模拟成为可能。这样的模拟不仅为实验工作增加了洞察力,而且还为新的研究方向提供了信息。考文尼及其同事最近的工作表明,带电矿物表面有可能从根本上改变核酸折叠等物理化学过程的动力学,这可能在未来几年被证明是重要的,可以支撑研究益生元化学的一个巨大挑战,即确定遗传信息在何时何地首次成为重要的鉴别器,以及达尔文进化的征程开始了它的无情推进。⁹⁵随着计算资源的日益强大,以及大规模电子基础设施环境的日益复杂,模拟在解决益生元化学领域的基础科学问题方面的能力将不断增强。

致谢。

作者感谢国家自然科学基金会通过国家计算机科学研究所在NRAC拨款MCA04N014和MRAC拨款DMR070013N下提供的XSEDE资源给予的支持,以及阿贡国家实验室阿贡领导计算设施(Intrepid)的资源,由美国能源部科学办公室根据合同DE-AC02-06CH11357提供支持。并在PRACE拨款PRA044下使用Jülich超级计算中心的JUGENE机器。我们还要感谢J. R. G. Evans教授对本文的有益评论。JBS由EPSRC博士奖学金资助。

注释和参考文献

- 1 W. Martin and M. J. Russell, *Philos. Trans. R. Soc. London, Ser. B*, 2007, 362 (1486), 1887–1925.
- 2 L. Orgel, *Trends Biochem. Sci.*, 1998, 23, 491–495.

- 3 R. M. Hazen and D. A. Sverjensky, *Cold Spring Harbor Perspect. Biol.*, 2010, 2(5), a002162.
- 4 B. M. Rode, A. H. Ederand and Y. Yongyai, *Inorg. Chim. Acta*, 1997, 254(2), 309–314.
- 5 H. Trinks, W. Schroder and C. K. Biebricher, *Origins Life Evol. Biosphere*, 2005, 35(5), 429–445.
- 6 P. V. Coveney and P. W. Fowler, *J. R. Soc. Interface*, 2005, 2(4), 267; P. V. Coveney and R. Highfield, *The Arrow of Time*, WH Allen, London, 1991; P. V. Coveney and R. Highfield, *Frontiers of Complexity: The Search for Order in a Chaotic World*, Faber and Faber, London, 1995.
- 7 S. L. Miller, *Science*, 1953, 117, 528–529.
- 8 J. B. S. Haldane, *Rationalist Annual*, 1929, 3, 3–10.
- 9 J. F. Kasting, *Science*, 1993, 259, 920–926.
- 10 J. A. Baross and S. E. Hoffman, *Origins Life Evol. Biosphere*, 1985, 15, 327–345.
- 11 P. Walde, R. Wick, M. Fresta, A. Mangone and P. L. Luisi, *J. Am. Chem. Soc.*, 1994, 116, 11649–11654.
- 12 D. S. Kelley, J. A. Karson and D. K. Blackman, *Nature*, 2001, 412, 145–149.
- 13 J. I. Drever, *The Chemistry of Weathering*, Springer, 1985.
- 14 M. J. Russell, R. M. Daniel and A. J. Hall, *J. Mol. Evol.*, 1994, 39, 231–243.
- 15 *The RNA World*, ed. R. F. Gesteland and J. F. Atkins, Cold Spring Harbour Press, New York, 1993.
- 16 T. R. Cech, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, 83, 4360–4363.
- 17 N. Ban, P. Nissen, J. Hansen, P. B. Moore and T. A. Steitz, *Science*, 2000, 289, 905–920.
- 18 M. W. Powner, B. Gerland and J. D. Sutherland, *Nature*, 2009, 459, 239–242.
- 19 K. Plankensteiner, H. Reiner and B. M. Rode, *Curr. Org. Chem.*, 2005, 9(12), 1107–1114.
- 20 C. Huber and G. Wächtershäuser, *Science*, 1998, 5377(281), 670.
- 21 P. W. Atkins and J. De Paula, *Physical Chemistry*, Oxford University Press, 1998.
- 22 L. Onsager, *Phys. Rev.*, 1931, 37(4), 405–426.
- 23 G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, Wiley, 1977.
- 24 A. M. Turing, *Philos. Trans. R. Soc. London, Ser. B*, 1952, 237(641), 37–72.
- 25 M. Eigen and P. Schuster, *Naturwissenschaften*, 1977, 64, 541–565; M. Eigen and P. Schuster, *Naturwissenschaften*, 1978, 65, 7–41; M. Eigen and P. Schuster, *Naturwissenschaften*, 1978, 65, 341–369. 26 J. Maynard Smith, *Nature*, 1979, 280, 445–446.
- 27 *Self-production of supramolecular structures from synthetic structures: to models of minimal living systems*, ed. G. R. Fleischaker and S. Colonna, Kluwer, 1994.
- 28 J. W. Szostak, D. P. Bartel and P. L. Luisi, *Nature*, 2001, 409, 387–390.
- 29 W. Ma and C. Yu, *Origins Life Evol. Biosphere*, 2005, 36, 413–420.
- 30 R. Lathe, *Icarus*, 2004, 168, 18–22; U. Wiechert, A. N. Halliday, D. C. Lee, G. A. Snyder, L. A. Taylor and D. Rumble, *Science*, 2001, 294, 345–348.
- 31 C. Fernando, G. von Kiedrowski and E. Szathmáry, *J. Mol. Evol.*, 2007, 64, 572–585.
- 32 C. Fernando and E. Di Paolo, *The Chemoton: a model for the origin of long RNA templates*, *Proc. of the 9th Int. Conf. on Simulations and Synthesis of Artificial Life (ALIFE)*, Boston, MA, 2004.
- 33 T. Ganti, *The Principle of Life (In Hungarian)*, Budapest, Gondolat, 1971; T. Ganti, *The Principles of Life*, Oxford University Press, 2003.
- 34 G. F. Joyce, *Nature*, 2002, 418, 214–221.
- 35 B. Obermayer and E. Frey, *J. Theor. Biol.*, 2010, 267, 653–662.
- 36 J. A. D. Wattis and P. V. Coveney, *J. Phys. Chem. B*, 1999, 103, 4231–4250.
- 37 P. V. Coveney and J. A. D. Wattis, *Mol. Phys.*, 2006, 104, 177–185. 38 J. A. D. Wattis, *Phys. D*, 2006, 222, 1–20.
- 39 R. Hanel, M. Pochacker and N. S. Thurner, *Philos. Trans. R. Soc. London, Ser. A*, 2010, 368, 5583–5596.
- 40 N. Lehman, *Chem. Biodiversity*, 2008, 5, 1707–1717.
- 41 P. V. Coveney and J. A. D. Wattis, *J. Chem. Soc., Faraday Trans.*, 1998, 102, 233–246; C. D. Bolton and J. A. D. Wattis, *J. Phys. Chem. B*, 2003, 107, 7126–7134.
- 42 S. Lonchin, P. L. Luisi, P. Walde and B. H. Robinson, *J. Phys. Chem. B*, 1999, 103, 10910–10916.
- 43 N. Berclaz, E. Blochliger, M. Müller and P. L. Luisi, *J. Phys. Chem. B*, 2001, 105, 1065–1071; N. Berclaz, M. Müller, P. Walde and P. L. Luisi, *J. Phys. Chem. B*, 2001, 105, 1056–1064.
- 44 F. Faglion, A. Passalacqua and P. Lazzeretti, *Origins Life Evol. Biosphere*, 2005, 35, 461–475.
- 45 H. B. Kagan and Dang-Tuan-Phat, *J. Am. Chem. Soc.*, 1972, 94, 6429–6433.
- 46 K. Soai, T. Shibata, H. Morioka and K. Choji, *Nature*, 1995, 378, 767–768.
- 47 F. C. Frank, *Biochim. Biophys. Acta*, 1953, 11, 459–463.
- 48 Y. Saito and H. Hyuga, *J. Phys. Soc. Jpn.*, 2004, 73, 33–35.
- 49 V. I. Goldanskii and V. V. Kuzmin, *Nature*, 1991, 352, 114–114; V. I. Goldanskii, *J. Phys. Chem. A*, 1997, 101, 3424–3432.
- 50 Y. A. Berlin, A. L. Burin and V. I. Goldanskii, *Z. Phys. D, At., Mol. Clusters*, 1996, 37, 333–339.
- 51 V. I. Goldanskii and V. V. Kuzmin, *Usp. Fiz. Nauk*, 1989, 157, 3–50; V. A. Avetisov and V. I. Goldanskii, *Usp. Fiz. Nauk*, 1996, 166, 873–891; V. A. Avetisov and V. I. Goldanskii, *Phys. Lett. A*, 1993, 172, 407–410.
- 52 V. A. Avetisov, V. I. Goldanskii and V. V. Kuzmin, *Phys. Today*, 1991, 44, 33–41.
- 53 D. K. Kondepudi, R. J. Kaufman and N. Singh, *Science*, 1990, 250, 975–976; D. K. Kondepudi, K. L. Bullock, J. A. Digits and P. D. Yarborough, *J. Am. Chem. Soc.*, 1995, 117, 401–404; D. K. Kondepudi and G. W. Nelson, *Phys. Lett. A*, 1984, 106, 203–206; D. K. Kondepudi and G. W. Nelson, *Nature*, 1985, 314, 438–441.
- 54 V. A. Avetisov, V. I. Goldanskii, S. N. Grechukha and V. V. Kuzmin, *Chem. Phys. Lett.*, 1991, 184, 526–530.
- 55 C. Viedma, *Phys. Rev. Lett.*, 2005, 94(065504), 1–4.
- 56 W. L. Noorduin, T. Izumi, A. Millemaggi, M. Leeman, H. Meekes, W. J. P. van Enkevort, R. M. Kellogg, B. Kaptein, E. Vlieg and D. G. Blackmond, *J. Am. Chem. Soc.*, 2008, 130, 1158–1159.
- 57 J. M. McBride and J. C. Tully, *Nature*, 2008, 452, 161–162.
- 58 M. Uwaha, *J. Phys. Soc. Jpn.*, 2004, 73, 2601–2603.
- 59 J. A. D. Wattis, *Origins Life Evol. Biosphere*, 2011, 41, 133–173.
- 60 P. G. H. Sandars, *Origins Life Evol. Biospheres*, 2003, 33, 575–587.
- 61 A. Brandenburg, A. C. Andersen, S. Höfner and M. Nilsson, *Origins Life Evol. Biosphere*, 2005, 35(3), 225–241.
- 62 J. A. D. Wattis and P. V. Coveney, *Origins Life Evol. Biosphere*, 2005, 35, 243–273.
- 63 A. Brandenburg, A. C. Andersen and M. Nilsson, *Origins Life Evol. Biosphere*, 2005, 35, 507–521.
- 64 C. Blanco and D. Hochberg, *Phys. Chem. Chem. Phys.*, 2011, 13, 839–849.
- 65 M. Gleiser, J. Thorarinson and S. I. Walker, *Origins Life Evol. Biosphere*, 2008, 38, 499–508.
- 66 N. Eldridge and S. J. Gould, *Models in Paleobiology*, Freeman, San Francisco, 1972, pp. 82–115.
- 67 R. Plasson, H. Bersini and A. Commeyras, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 16733–16738.
- 68 M. Gleiser and S. I. Walker, *Origins Life Evol. Biosphere*, 2009, 39, 479–493.
- 69 S. J. Mojzsis, T. M. Harrison and R. T. Pidgeon, *Nature*, 2001, 409, 178–181.
- 70 P. Baaske, F. M. Weinert, S. Duhr, K. H. Lemke, M. J. Russell and D. Braun, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, 104, 9346–9351; D. Braun and A. Libchaber, *Thermal force approach to molecular evolution*, *Physical Biology*, 2004, 1, P1.
- 71 C. J. Cramer, *Essentials of Computational Chemistry: Theory and Models*, John Wiley & Sons Inc, 2004; J. Kohanoff, *Electronic Structure Calculations for Solids and Molecules: Theory and Computational Methods*, Cambridge University Press, 2006.
- 72 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, 1999; D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Elsevier (formerly published by Academic Press), 1996.
- 73 J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, 21(12), 1049–1074.
- 74 R. T. Cygan, J. J. Liang and A. G. Kalinichev, *J. Phys. Chem. B*, 2004, 108(4), 1255–1266.
- 75 S. Plimpton, *J. Comput. Phys.*, 1995, 117(1), 1–19.
- 76 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, 26(16), 1781–1802.

- 77 H. J. C. Berendsen, D. van der Spoel and R. Van Drunen, *Comput. Phys. Commun.*, 1995, 91(1–3), 43–56.
- 78 G. Sposito, *The surface chemistry of soils*, Oxford University Press, New York, 1984; R. E. Grim, *Clay Mineralogy*, McGraw Hill, New York, Second edn, 1968.
- 79 J. Wang, A. G. Kalinichev, J. E. Amonette and R. J. Kirkpatrick, *American Mineralogist*, 2003, 88(2), 398; X. Hou, A. G. Kalinichev and R. J. Kirkpatrick, *Chem. Mater.*, 2002, 14(5), 2078–2085.
- 80 J. O. Titiloye and N. T. Skipper, *Chem. Phys. Lett.*, 2000, 329(1–2), 23–28; J. L. Suter, P. V. Coveney, H. C. Greenwell and M.-A. Thyveetil, *J. Phys. Chem. C*, 2007, 111(23), 8248–8259; E. S. Boek, P. V. Coveney and N. T. Skipper, *J. Am. Chem. Soc.*, 1995, 117(50), 12608–12617.
- 81 H. C. Greenwell, W. Jones, P. V. Coveney and S. Stackhouse, *J. Mater. Chem.*, 2006, 16(8), 708–723.
- 82 J. F. Lambert, *Origins Life Evol. Biosphere*, 2008, 38(3), 211–242.
- 83 M. G. Schwendinger and B. Rode, *Phys. Lett.*, 1989, 155(6), 527–532.
- 84 N. Lahav, D. White and S. Chang, *Science*, 1978, 201(4350), 67; N. Lahav and D. White, *J. Mol. Evol.*, 1980, 16(1), 11–21.
- 85 J. R. Collins, G. H. Loew, B. T. Luke and D. H. White, *Origins Life Evol. Biosphere*, 1988, 18(1–2), 107–119.
- 86 A. J. A. Aquino, D. Tunega, M. H. Gerzabek and H. Lischka, *J. Phys. Chem. B*, 2004, 108(28), 10120–10130.
- 87 B. M. Rode, *Peptides*, 1999, 20(6), 773–786; B. M. Rode and Y. Suwannachot, *Coord. Chem. Rev.*, 1999, 190(1), 1085–1099; K. R. Liedl and B. M. Rode, *Chem. Phys. Lett.*, 1992, 197(1–2), 181–186; M. Remko and B. M. Rode, *Chem. Phys. Lett.*, 2000, 316(5–6), 489–494.
- 88 A. Rimola, L. Rodriguez-Santiago, P. Ugliengo and M. Sodupe, *J. Phys. Chem. B*, 2007, 111(20), 5740–5747.
- 89 N. N. Nair, E. Schreiner and D. Marx, *J. Am. Chem. Soc.*, 2006, 128(42), 13815–13826.
- 90 L. Leman, L. Orgel and M. R. Ghadiri, *Science*, 2004, 306(5694), 283–286.
- 91 C. H. Yu, S. Q. Newton, D. M. Miller, B. J. Teppen and L. Schafer, *Struct. Chem.*, 2001, 12(5), 393–398.
- 92 H. C. Greenwell and P. V. Coveney, *Origins Life Evol. Biosphere*, 2006, 36(1), 13–37.
- 93 D. C. Mathew and Z. Luthey-Schulten, *Origins Life Evol. Biosphere*, 2010, 40(3), 303–317.
- 94 W. Huang and J. P. Ferris, *J. Am. Chem. Soc.*, 2006, 128(27), 8914–8919.
- 95 J. B. Swadling, P. V. Coveney and H. C. Greenwell, *J. Am. Chem. Soc.*, 2010, 132(39), 13750–13764.
- 96 M.-A. Thyveetil, P. V. Coveney, H. C. Greenwell and J. L. Suter, *J. Am. Chem. Soc.*, 2008, 130, 12485–12495; G. Arrhenius, *Helv. Chim. Acta*, 2003, 86, 1569–1586.
- 97 J. B. Swadling, P. V. Coveney and H. C. Greenwell, *Geochim. Cosmochim. Acta*, 2012, 83, 360–378.
- 98 M.-A. Thyveetil, P. V. Coveney, H. C. Greenwell and J. L. Suter, *J. Am. Chem. Soc.*, 2008, 130, 4742–4756.
- 99 A. G. Cairns-Smith, *Genetic Takeover and the Mineral Origins of Life*, Cambridge University Press, New York, 1982.