

生命科学中的物理问题

罗辽复

(内蒙古大学物理系, 呼和浩特 010021)

将生命科学中的物理问题分为四类: 第一, 关于生命的起源和进化; 第二, 关于密码-序列-构象-动力学(功能)的反向生物学路线; 第三, 关于生物自组织; 第四, 关于脑信息处理和向生命学习的问题. 文中就每一类列举了若干基本的物理问题.

Abstract

The physical problems in life science can be classified into four classes; namely, (1) those on the origin and evolution of life, (2) those on a line of reverse biology: code-sequence-conformation-dynamics (function), (3) those on the self-organization of living body, and (4) those on the information processing in brain and the problems of learning from life. Several basic problems of physical importance are cited for each category.

生命现象一直是物理学家关注的重点. 我们可以列举出一大堆名字, 如玻尔、薛定谔 Delbrück, 伽莫夫, Crick, (分子生物学), Bekesy (听觉), Cooper (视觉), 安德森, Hopfield (神经网络), Cormack (X 射线断层摄影), Nehr (膜离子通道), 等. 在这些物理学家中, 有很多都对生命科学的发展作出了巨大贡献. 尽管如此, 什么是生命物理学? 这个问题的回答仍然是众说纷纭. 一个搞遗传学的人和搞生物膜的人, 一个搞生物流变学的人和搞视觉的人, 所关心的问题可能完全不同. 原因在于生命现象的复杂性和丰富性, 也由于生物物理还不是一个发展得很成熟的科学领域, 因此十个人就有十种回答. 这个问题也可从另一角度来看, 物理学在推动近代生命科学发展中的贡献有哪些? 主要在两个方面. 一是为生命科学提供现代化的实验手段, 如电子显微镜、X 射线衍射、CT 技术、核磁共振、扫描隧道显微镜等; 二是为生命科学提供理论概念和方法. 后一方面对于物理学家和生物学家来说, 贡献是相互的. 物理学家既给生物学以量子力学、原子分子理论、热力学等理论工具, 也不断从生命科学的发展中吸取营养. 化学家 Volkenstein 说得好:

“如果物理学给生物学以显微镜, 那么生物学给物理学的是能量守恒.” 的确, 上一世纪中叶, 迈耶尔和亥姆霍兹等正是从生命现象的研究中领悟到能量守恒的. 迈耶尔注意到热带海员患者的静脉血比他们在欧洲时红, 认为这是由于血中含较多 O_2 之故. 因为在热带有机体只需要吸收食物中较少热量就可维持生命, 食物的燃烧过程减弱就使得静脉里留下较多 O_2 . 他认为食物的化学能可转化为热能, 并从空气的定压、定容比热关系算出热功当量. 所以发现能量守恒的都不是某一领域的专家, 他们的视野是整个自然界. 19 世纪的这段历史是令人鼓舞的, 今天重新来探讨物理学和生物学的关系时, 不要忘记生物学对于物理学发展的作用.

但是当我们稍为深入地探讨生物学问题时, 便会发现生物学与物理学二者之间的巨大差异. 第一, 物理学习惯于研究简单的系统, 而生物学则研究自然界最复杂的生命. 第二, 物理学研究大几率事件, 可以根据抽样来对它进行预测, 而生物学研究的却是那些小几率事件. 生命本身就是同类组织结构中无数可能状态中一种罕见状态. 有人认为它属于完全非均一类中的一员, 根据抽样来对它预测是不可能的^[1]. 第

三,物理学家喜欢问 why?探究原因;生物学家喜欢问 what for? 寻找目的.

尽管二者存在巨大差异,但生命科学中还是有很多有趣的物理问题.什么是“物理”?费因曼说:“物理学家常有这样的习惯,研究现象的最简单例子,并把它叫做‘物理’,把那些更复杂的例子留给其他学科去研究.”对于复杂的生命,我们也将遵循这种方法来挑选物理问题.Crick 曾经指出,分子生物学的突破是在研究生命起源和进化早期现象时作出的.简单的过程往往是自然界复杂性的基础,突破从这里开始是容易理解的.因此,下面列举的问题也较多地和生命起源及早期进化有关.

一、关于生命的起源和进化

生物体内蛋白质的氨基酸的立体结构是左手性的(L型),而核酸中的五碳糖是右手性的(D型).这种立体结构上的左右不对称性和它们的光学性质有关,所以又称为光活性.为什么生物大分子只有一种手性?这是因为不同手性的对映异构体性质差别很大.例如,用部分D氨基酸来取代L氨基酸, α 螺旋就要破坏,成为无规线圈.所以分子手性决定了高级构象;而高级构象是生物功能的基础.因此,一种手性的物质是有毒的,相反手性的物质却是无害的,D型天冬氨酸是甜的,而L型天冬氨酸却是无味的.细菌只吃一种对映异构体的食物.一种L型生物生活在D型生物的世界中是必定要饿死的.地球大约在46亿年前诞生,从那时开始到31—32亿年前的这段时期里,地球上发生了通过非生物途径合成有机物的过程,这阶段称为化学进化.在有机物产生、复杂化、组织化的同时,光活性也就起源和发展了.光活性物质的产生加速了物质组织化的过程,随之而来的是原始生命的产生和生物进化的阶段.所以手性分子系统的演化对于生命起源是极端重要的.那么手性分子是如何产生的呢?有两种可能解释:一是不对称力的影响;二是对称性自发破缺.看来两方面都有作用^[7].由于光活性在生命进化

物理

过程中的特殊地位,Crick 把生物分子只有一种对映异构体称为生物化学的第一原理.

和生物分子手性起源相关的是密码起源问题^[3].生命是复杂的,地球上有150万种生命.但是存在内在的单纯性——普适的大分子组成、普适的手性和普适的密码.这说明生物有共同的祖先.密码的普适性说明它产生于生命的极早阶段,产生于生物学进化得到那么多的物种分岔之前.值得一提的是遗传密码最早由伽莫夫提出.“伽莫夫工作的重要性在于,这是一种真正抽象的密码理论,没有那些冗赘而不必要的化学细节,尽管他的基本观念——认为双链DNA是蛋白质合成的模板——显然是完全错误的.”关于遗传密码的逻辑,作者已有另文评述^[4],兹不赘述.

生命信息储存于核酸序列中.这个概念最早是由薛定谔提出的^[5].薛定谔从Delbrück的推测——基因包含于 10 \AA 大小的微观体积中——出发,建议大分子作为一种非周期固体,可作为遗传信息的负荷者,他举了一个例子,用五个符号编写长度25的文字,可能排列数为 4×10^{17} 个.所以局限于一个很小空间的微型密码就可以包含足够多的信息,描写足够复杂的内容.自从1955年Sanger发表氨基酸序列分析技术和1977年发表DNA快速测序技术以来,国际基因库中的序列资料迅猛增长,截至1990年底,已达5000万核苷酸.这是一部5000万言的巨著,比24史加在一起还厚!氨基酸序列决定了蛋白质的一级结构、高级结构和功能,核苷酸的序列又指导着氨基酸序列的形成,决定着基因的表达和调控.它们决定了包括生命形成和演化以及个体发育、新陈代谢等在内的一切生命活动的基本方面.因此,这些丰富的序列资料向研究者提供了生命系统的最重要信息,展示在我们面前的是一本厚厚的生命之书.然而直到今天,人们还只是找到了它的一小部分,离开读懂它就更远了.试读这部“天书”的第一步工作为Dayhoff所做,她通过氨基酸序列比较,寻找同源关系,从分子水平上导出了进化树.由此可导出基本普适的进化时标—— 10^{-9} /氨基酸·年.

这意味着高等动物基因组突变时标为1.75年。在群体遗传学中 Haldane 证明了一个定理,据此一个典型的基因取代要300代才能实现,这和基因组突变时标1.75年矛盾。为了解决这个矛盾, Kimura, King, Jukes (1968, 1969年)提出了中性突变假设。认为在分子水平上,大多数突变对于自然选择来讲是中性或近中性的随机变化。例如不改变氨基酸的同义突变,非功能性DNA序列中的碱基突变,或不影响功能的氨基酸突变,都属于这个类型^[6]。中性突变是对达尔文自然选择的一个重要补充。

以上是对氨基酸序列的分析。如果对核酸序列进行统计研究,可以得到更多有趣的结论。这主要是遗传语言的进化特征。例如碱基关联随进化逐渐增加,而且这种关联多半是短程的。这些统计规律提供了构造核酸和蛋白质序列进化模型的依据,也给出了一个实例,说明在核酸这个开放系中熵是如何变化的。这对于理解热力学的箭头是有启发性的^[7]。

上面谈到生命进化有一个时标,大约10亿年。经过10亿年,每个氨基酸更新换代一次,就出现生命。生命的出现在空间范围也有一定大小。人们估计,为了成功地进行细胞分裂,为了信息不致丢失,细胞不能太小。当细胞半径 $R=100 \text{ \AA}$ 时,成功几率仅 6×10^{-16} ,而当 $R=500 \text{ \AA}$ 时,成功几率就非常接近于1。更细致的分析表明,细胞要具有比较完整的功能,其直径至少要 1000 \AA ,这是下限。细胞大小有没有上限呢?细胞要通过其表面不断与环境进行物质交换。没有足够表面积,代谢作用不能正常进行。而表面积与体积之比反比于直径,因此细胞直径一般在 $10^{-4} - 10^{-5} \text{ cm}$ 。 10^{-5} cm 是生命发生的空间尺度。因此可以认为生命发生于微观与宏观之间。人的基因组包含 $3 \times 10^9 \text{ bp}$ 碱基对,大脑有 10^{11} 个神经元。若把微观尺度 ~ 1 ,宏观 $\sim 10^{23} (N)$,上述数字约为 \sqrt{N} ,也说明生命发生于微观宏观之间。

生命进化中的一个有趣问题是有那么多物种形成。这和耗散结构理论里的分岔颇为相象。但要认真构成一个物种进化理论并不简单,看

来还需要一些物理原理。其中之一就是竞争排斥原理。这是一个生态学原理。生态学中把物种栖息的空间称为生态位。由于竞争,不能有两个以上种群占据同一实际的生态位,这个生物学中的泡利原理是可以实验证明的。竞争排斥在自然界中是常见的。有时看到一些不同物种能利用相同资源而共存,但仔细观察,它们在利用资源上总有差别。例如危害作物的一些害虫,在取食部位、时间等方面总有微小而重要的差别。这是一种生态分离现象。由于生态分离,两部分生物受到的选择力将出现明显差异,导致它们的基因序列向不同的稳定态漂变,最后形成不同的物种。

二、关于密码—序列—构象—动力学 (功能)的反向生物学路线

传统的生物学的研究是由形态到细胞,到分子水平的结构,由生理到微观调控,由表型到基因型。现在看来有可能建立反向生物学,由基因入手,由序列而构象,由结构而功能,从而阐明千变万化的生命现象。密码—序列—构象—动力学(功能),就是一条可能的反向生物学路线。这种反向生物学的建立,将全面地革新生命科学的面貌,使生命科学的体系与数学的理性精神相协调,使它更富有预见性^{[8], [11]}。伽里略在论述“自然”这部宏大的书时说,它是用数学的语言写成的。没有数学,一个人只能在黑暗的迷宫里徘徊。今天对生命科学,也应作如是观。

首先是密码。通常遗传密码是指核苷酸顺序和氨基酸之间指定的关系,是指信息从核酸到蛋白质的流动中的密码识别。但是这还不够,基因表达是在严格调控下进行的。一个受精卵或一个体细胞都含有发育成为完整个体的全部信息。为什么不同类型的体细胞含有相同的基因,而形状、功能、代谢特征却各不相同?这是由于基因在何时何条件下表达,受到严格的调节控制,这样才能保证生命有条不紊地进行。关于

1) 罗订夏,大自然探索,7-4(1988)。

原核基因的调节, Monod 提出了一个乳糖操作子模型. 这是一个调控线路, 分成两部分: 一是结构基因, 乳糖分解需要三种酶, 细菌只有在需要这三种酶时才合成它, 结构基因就是编码这三种酶的 DNA 片段; 二是调节基因, 调节基因编码阻遏物蛋白质. 阻遏物能识别 DNA 的特定片段(叫操作子), 它们靠氢键结合, 结合后就使启动子关闭, 结构基因停止表达. 当阻遏物与诱导物结合时, 它就失活, 因而恢复酶的合成. 阻遏物与诱导物间的键是微弱的, 能很快形成和破裂, 所以调节是很灵敏的. 真核基因的调节复杂得多, 但基本机理相同. 这个自动控制线路包含两个过程: 一是 DNA 转录和转译, 生成 RNA 和蛋白质(酶), 二是蛋白质反过来和 DNA 作用, 如上例中的阻遏物蛋白和操作子之间的分子识别. 因此, 除了信息从 DNA 流向蛋白质过程中的遗传密码外, 还有基因表达调控中蛋白质对 DNA 特定片段的识别. 我们说的密码就是指核酸和蛋白质两类序列相互识别的密码关系.

DNA 指导生成蛋白质, 蛋白质又反作用于核酸, 调节 DNA 信息的表达. 这是生命活动的“怪圈”. 巴赫为了奉献弗里德利希国王, 创作了“音乐的奉献”. 在这个乐曲中, 他用了一种特殊技巧, 使听众有一种不断增调的感觉. 转了几圈之后, 听众感到离原调已很远了. 可是奇妙的是, 通过一种变调, 又使结尾平滑地过渡到开头, 回到原调. 巴赫原意是采用这种方法使升调过程无限地进行下去. 因此他在乐谱上注了“陛下的荣耀也随着变调而增高”. 巴赫创作的是音乐中的“怪圈”, 现在, 通过 DNA 与蛋白质的“怪圈”, 生命活动逐步向高级发展, 不是和巴赫创作的怪圈颇为相象吗?

再来看序列和构象的关系. 生物学家认为, 在一定条件下, 氨基酸序列决定了蛋白质的折叠, 从而决定了它的功能. 目前已有 700 余个不同蛋白质的空间结构被测定, 而测定顺序的氨基酸序列数约四万余条. 如果由序列能预测构象, 则大大增加了人们对各种蛋白质的功能的了解. 基因工程已经进入到生产自然界不存在物理

蛋白质的蛋白质工程阶段. 为了获得新型的具有一定功能的蛋白质, 需要选择某些氨基酸位点进行定向突变. 当然, 这需要序列和构象关系的理论指导. 所以蛋白质折叠已经成为一个热门话题. 60 年代以来, 人们把蛋白质结构分为四个层次. 序列是一级结构, 多肽链的规则构型如 α 螺旋、 β 折叠等是二级结构, 三维构型是三级结构, 若干亚单位构成聚集体是四级结构. 一级结构是基础, 它决定了高级结构. 近年来弄清了由序列决定高级结构不是单一的, 实际情况是由序列决定有若干微小差异的空间结构, 它们称为构象子态. 它们一般有相同功能, 只是实现功能的速率有所不同罢了^[9]. 体内新生肽链的折叠更为复杂, 需要有一定的促进或催化折叠的物质(酶和一种称为分子伴侣的蛋白质)的帮助. 折叠经过一定的启动和一定的中间体, 落入众多构象子态中的某一个, 然后在这些子态间弛豫.

分子生物学的诞生以双螺旋模型的提出为标志. 这是对 DNA 二级结构的认识. 但是这个模型并没有完全解决核酸构象的问题. 由于嘌呤体积大, 挤来挤去使实际双螺旋中碱基对排列不如预想那样平坦而规则, 这种局部偏差可能蕴含蛋白质凭以识别的信息^[10]. 正象蛋白质功能须在一定折叠状态下才能表现出来一样, 核酸功能——转录复制及其动力学, 也和它的堆集状态有关.

确定线型大分子折叠的力是什么? 是弱力. 生物体中原子、分子间作用力根据结合能的大小区分为两类: 一类是强力, 如共价键、离子键, 结合能约为 50—100kcal/mol; 一类是弱力, 如涉及到键长、键角变化的能量为 1—7kcal/mol, 涉及到扭角变化的能量为 0.1—1kcal/mol. 这些都是弱力. 弱力来源于范德瓦耳斯作用、氢键和疏水键作用. 为什么弱力对生命运动特别重要? 生理温度下平均热能为 0.6kcal/mol. 弱键容易受热运动影响, 它是敏感的、易变化的因素. 因此弱键储存的能量是活性能, 弱力是导演生命运动的主要力. 当然, 这并不是说电子运动不重要, 电子在生命活动中也扮演着重要角色.

例如嘌呤和嘧啶都是共轭分子,具有非局域的电子轨道,它们称为 π 电子。 π 电子在确定核酸性质方面起着基本的作用。又如电子从施主元素(如H)到受主元素(如O)的转移,有一条电子传递链,在生化过程(如呼吸作用)中特别重要,这也是电子在生命活动中扮演的主角。既考虑构象(分子形状),又考虑电子(特别是前沿电子),建立一个构象——电子系统的量子理论,对于解释生物活性和生物功能具有特别重要的意义。由于有构象坐标的参与,这已不是通常的量子化学或量子生物学(即量子生物化学),需要有新的物理思想和方法^[8,11]。

生命是高度有序的系统。然而具体计算表明,一个有机体的序和一块同样质量石头的序没有多大差别,生物大分子的结构特点是它象一种非周期固体,具有丰富的信息内容。如何研究这类既不具有周期性又不是无序的系统?如何从噪声背景中提取有用的信息?这是一个新的物理问题。

三、关于生物自组织

合作现象出现于生命活动的分子的、细胞的和整体的各个水平上。例如血红蛋白和氧的结合呈S型曲线,保证在氧气稠密处吸收氧,到空气稀薄处释放氧,变化灵敏。这样血红蛋白才成为有效的负载氧的工具。合作现象可用伊辛模型去讨论。但这仅是生物自组织的最简单例子。实际情况要复杂困难得多。特别是和物质自组织相联系,还有一个能量自组织的问题。例如视觉中网膜视紫质具有单光子敏感性,吸收一个光子后产生脉冲的能量却是被吸收光子的 10^8 倍,这些能量是如何在 10^{-2} s这样短的时间内集中起来的?又如肌肉收缩是由于肌球蛋白头部和肌动蛋白纤维的交替作用,为了举起1g重物移动1cm,需要 10^{14} 个ATP分子能量。这些原来存储于不同分子的能量如何同时释放出来用于肌肉收缩而不致于中途耗散?生物体中能量和信号的传递不象一块石头投于水中激起的波纹,而象蛛丝,一端扰动可沿丝传递很远而

邻近的却纹丝不动。这些关于能量的相干产生和长距离输运,机理尚不清楚,有待进一步研究。

生命是一台化学机器。利用自己制造的有机催化剂,促进一系列循序衔接的反应,进行能量的转换和它本身组分的合成。这里能量转换的效率是很高的,机器是按照最经济的原则运转的。例如一个细菌细胞不但能迅速制造单个蛋白质分子,而且按精确克分子比值同时制造3000种以上的蛋白质。这台化学机器中包含的物理原理是值得人们去探索的。

生命从种系角度讲是遗传和进化,从个体角度讲是发育。个体发育是涉及遗传信息表达、细胞分化等的复杂问题,但仅就形态发育而言,还是较易着手、可用自组织理论去处理的^[12]。早在1952年,Turing就讨论过这类问题。这里的一个基本观点是位置信息。1969年,Wolpert提出细胞分化决定于某种位置信息,后者来源于和细胞相互作用的化学物质(貌基)浓度,如果不问来源,把位置信息抽象化,就得到发育场的概念。发育场是发育有机体的特定区域,其中图样形成显示出定位的、整合的机制,它的活性在某些由动力学决定的点上最强,随着距离增大而衰减。发育场的概念是发育生物学家从实验中提出的,和物理学中的场似乎不尽相同,Trainor等用物理学中场论方法去研究它,解释了蝾螈断肢再植实验中产生若干赘余肢的有趣实验结果^[13]。

细胞膜是细胞和外界物质、能量、信息的转换站。膜的自组织是生物自组织的一个重要方面。磷脂双分子层作为一个物理系统来研究已经有很多工作。但和膜功能有关更重要的是镶嵌于其中的蛋白质,例如离子通道和受体蛋白。这又回到了前面讲的密码—序列—构象—动力学的路线。

生物自组织的热力学原理是又一个值得注意的问题。在《生命是什么?》这部名著中,薛定谔提出“生命以负熵为生”。人们认为这是生命的热力学基础。但是为了解释生物的高度自组织性能,似乎这还不够。生物体在正常状态下可

能遵循“熵产生最小”的规律. 如果这被证明是正确的, 那么在生命的热力学基础中还应加上一句话: “生命以熵产生最小而活”.

四、关于脑信息处理和向生命学习的问题

在理论天体物理中有两大类问题: 一是宇宙演化的早期; 一是恒星演化的晚期. 同样, 在讨论生命科学中的物理问题时, 也应注意两头. 一头与生命起源和早期进化有关, 另一头则与生命进化最高阶段的产物——大脑有关. 对于后一头, 主要的物理问题可能集中于脑信息处理. 这方面研究最深入的是视觉. 根据视觉系统信息加工的研究, 人们知道视觉不是外界图形光强分布的简单模写, 而是一个信息加工过程, 其中有高级智能活动的成分^[14]. 这似乎说明认识是一种主观和客观相互作用的过程. 对于大脑信息加工, 现在还了解得很少和很不完全. 近年最重要的进展是神经网络. 这方面已有很多介绍, 本文不谈. 我们要指出的是, 现在流行的 Hopfield 神经网络实际上只是利用了大脑信息处理方式众多特点中的一个——分布表示, 即高级神经活动是大量神经元集体活动的结果, 信息不是贮存在单个神经元中, 而是贮存于整个网络中. 由此派生出一系列有趣的和有用的特点, 如并行方式、图象识别能力、误差自动校正等^[15]. 在 Hopfield 神经网络理论上发展起来的神经计算机, 在和传统的数字计算机相比中已经显示出巨大的威力. 如果我们更加深入而全面地向大脑学习信息加工方式, 定会得到更多的酬报.

1987年, 日本通产省和英国皇家学会搞了一个人类前沿科学计划, 喊出了“向生命学习”的口号^[16]. 生命具有极为精巧的信息处理和极为有效的物质、能量转换功能, 比现有科学技术创造出来的人工功能要精细和有效得多. 探索

生命的奥秘, 向生命学习, 有可能创造科学技术的新模式. 那么这个学习有哪些内容呢? 主要有两个方面: 一是物质和能量的转换; 二是脑信息处理加工. 所以向大脑学习是向生命学习的重要一环.

上面列举了四类物理问题. 还有未谈到的, 如生物学研究中物理仪器的应用, 物理因素对生物结构和功能的影响等, 都是大有可为的.

尽管物理学和生物学两个学科传统差异很大, 但他们都要发展, 都不会停留在原先的模式上. 远缘杂交提供了绝好的创造机会, 作者愿引用数学家哈达玛的一句话作为本文的结束: “很明显, 发明或发现, 无论是数学的还是其他方面的, 都发生于观念的组合.”

- [1] 西伯斯马著, 沈淑敏、王大成等译, 生物物理学引论, 科学出版社, (1979).
- [2] Avetisov VA et al., *Biosystems*, **25** (1991), 141; *Physics Today*, **44**(1991), 33.
- [3] Lacey JC et al., *Origins of Life*, **22**(1992), 243.
- [4] 罗辽复, 科学, **42-3**(1990), 187; 百科知识, No. 2, (1992)62.
- [5] E. Schrödinger, *What is Life?* Cambridge Univ. Press, (1945).
- [6] M. Kimura, *The Neutral Theory of Molecular Evolution*, Cambridge Univ. Press, (1983).
- [7] 罗辽复, 科学**42-4**, (1990), 271.
- [8] 罗辽复, 物理, **20**(1991), 85.
- [9] H. Frauenfelder et al., *Ann. Rev. Biophys. Biophys. Chem.*, **17**(1988), 451.
- [10] G. Heijne, *Sequence Analysis in Molecular Biology*, Acad Press, Inc., (1987).
- [11] M.V. Volkenstein, *Molecular Biophysics*, Acad Press, London, (1977).
- [12] Murray JD, *Bull. Math. Biol.*, **52**(1990), 119.
- [13] Trainor LEH, *Theoretical Biology* (Ed: Goodwin B, Saunders P) Edinburgh Univ. Press, (1989).
- [14] 汪云九, 科学, **44-4**(1992), 9.
- [15] 王宝翰, 生物物理学报, **5**(1989), 97.
- [16] 中国生物物理学会编译, 人类前沿科学计划, (1987).