

科学 专著丛书

生命进化 的物理观

PHYSICAL ASPECTS
ON LIFE EVOLUTION

罗辽复 著

上海科学技术出版社

科学专著丛书

生命进化的物理观

罗辽复 著

上海科学技术出版社

内 容 提 要

本书是理论生物学领域的一部学术著作,作者用物理学的观点和方法,提出和讨论了生命进化的一系列重大问题。本书也是一名理论物理学工作者潜心转向生命科学研究十几年的学术总结。全书共四章。第1章介绍生物分子手性的起源和进化;第2章探讨了遗传密码的起源和进化;第3、第4章则重点论述核酸、蛋白质序列的进化和有关的信息学研究。

本书可供从事生命科学与理论物理交叉领域研究的学者、教师及研究生阅读,也可作“理论生物物理学引论”等研究生课程的参考读物。

图书在版编目(CIP)数据

生命进化的物理观/罗辽复著. —上海:上海科学技术出版社,2000.10

(科学专著丛书)

ISBN 7-5323-5562-4

I. 生... II. 罗... III. 生物-进化-研究

IV. Q11

中国版本图书馆 CIP 数据核字(2000)第 40627 号

上海科学技术出版社出版发行

(上海瑞金二路 450 号 邮政编码 200020)

商务印书馆上海印刷股份有限公司印刷

新华书店上海发行所经销

2000 年 10 月第 1 版 2000 年 10 月第 1 次印刷

开本 787×1092 小 1/16 印张 25 插页 4 字数 377 000

印数 1—1 500 定价: 52.00 元

本书如有缺页、错装或坏损等严重质量问题,
请向本社出版科联系调换

本书由华夏英才基金资助出版

《科学专著丛书》序

如果说科学研究论文是创造性科学工作的发表性记录,那么科学技术学术专著则是创造性科学工作的总结性记录。前者注重的是优先权,后者注重的是系统化。

在大量科学研究的基础上,对一个专题或一个领域的研究成果,作系统的整理总结,著书立说,乃是科学研究工作不可少的一个组成部分。著书立说,既是丰富人类知识宝库的需要,也是探索未知领域、开拓人类知识新疆界的需要。特别是在科学各门类的那些基本问题上,一部优秀的学术专著常常成为本学科或相关学科取得突破性进展的基石。所以,科学技术学术专著的著述和出版是一项十分重要的工作。

近20年来,中国的科学事业有了迅速的发展,涌现了许多优秀的科学研究成果,为出版学术专著提供了坚实的基础。值此20世纪90年代,在出版学术专著方面,中国的科学界和出版界都在抓紧为本世纪再加些积累,为迎接新世纪多作些开拓。我高兴地看到,作为这种努力的一个部分,《科学》杂志的出版者——上海科学技术出版社推出了这套《科学专著丛书》。

上海科学技术出版社是科学技术界熟悉和信赖的一家出版社,历来注重科学技术学术专著的出版。《科学》杂志的编者组织编辑学术系列丛书,也不是第一次。在本世纪三四十年代,就曾推出过《科学丛书》,其中不乏佳作,对当时的学术研究起了很好的作用。

《科学》在中国是一份历史最长的综合性科学刊物,80年来与科学技术界建立了广泛的密切联系。现在推出的这套《科学专著丛书》正是这种联系的产物。我相信,加强这种联系,著者与编者、出版者,科技界与出版界共同努力,精心选题,精心编辑,精心出版,一定能使这套专著丛书反映

出中国科学技术研究的最新水平,为本世纪多留下几本中国学者的优秀专著,为迈向新世纪多铺下几块引路的基石!

周光召

(《科学》杂志编委会主编)

1994年8月

序 言

生命是自然界最伟大的创造。生命的起源和进化一直激荡着每一个求知者的心。这是一项综合性极强的课题,其中每一条款的真正解决都要依靠多门学科的通力合作。物理学以研究自然界基本规律为己任,自然对此问题有着浓厚兴趣。本书将着重于生命进化早期的某些事件,并特别强调从微观的分子角度来进行研究。当然,题材的选取主要反映了作者个人的观点和研究经验,不敢要求它具有全面性。事实上,在一个课题如此宽广、发展如此迅速的领域,全面性是很难达到的。因此,从某种意义上说,把这本书看成一个理论物理学工作者转向生命科学的读书笔记和研究报告,也许更为合适。

1950年代分子生物学的诞生,把生命科学推进到分子原子水平上去;即将到来的一场新的革命,将把生命科学提高到理性的综合的水平。理论生物学将成为生命科学的一个重要组成部分,正如理论物理学之于物理科学一样。本书曾作为研究生课程“理论生物物理学引论”的一部分试用。按照作者的观点,这门课应该包括两部分,一是生命进化的物理观;二是生物功能的物理观。因此这本书可看成理论生物物理学引论的第一卷。有些既与进化有关又与功能有关的问题则要在两卷中适当安排。本书分三部分,一是生物分子手性的起源和进化;二是遗传密码的起源和进化;三是核酸和蛋白质序列的进化和有关的信息学研究。将第三部分分成两章纯系为了行文的方便,没有更重要的理由。另外,与进化有关的物种竞争拟放到下一卷中去。

科学研究既要有交流,又要有个人的静心钻研。感谢命运把我安排在一个沙漠的边缘,有着创造的自由,能够在这里连续若干年地进行多方面独立思考。本书中大部分内容都是作者和他的学生在这个独特环境中的工作成果,其中不少论文发表在《内蒙古大学学报》上。在将学术刊物按“影响力”打分的时尚中,一个大学学报往往是不被人注意的。尽管在真理

面前,一切刊物皆平等,有些热心的朋友还是劝我把过去在《内蒙古大学学报》上发表的论文系统地整理成书,让更多的读者了解和参与讨论。这本书也算是对这个建议的一个回答。近年还发现我研究组过去的工作常常以某种形式被人重复,如果总结在本书中的研究成果能替后来者起一点铺路石的作用,这是作者最大的欣慰。

生命是作为一种基因的装置而存在的。对人来说,除了生物基因外,还有另一种酷似基因可复制可传播并进化着的东西,这就是文化基因。我们可以遗留给后代的东西只有两种:基因和文化基因。柏拉图在《饮宴篇》中论证了这样一条哲理:一切有死之物都希望通过生育繁衍这个唯一可能的方式达到永生。“性爱是希望肉体长存的欲望的表现,而各种创造性的艺术作品都是精神上的产儿。不论在哪一方面,这种以繁衍的方式达到永生的欲望是动人的。”当然,科学作品也在这个范围。唯有创造才是欢乐,唯有创造的生灵才是生灵。在本书的创作中我感到了这种欢乐。

感谢与本书有关课题的探索中共同工作过的多位学生,感谢国家基金“blue sky”计划和内蒙古大学物理系的支持,感谢华夏英才基金对本书出版的资助。感谢内蒙古大学生物系方天祺教授、中国科学院理论物理研究所郝柏林教授和刘寄星教授在百忙中审阅了本书初稿,并提出了很多宝贵意见。在本书即将脱稿之时,作者内心不由地产生了对父母和先人的深切怀念之情,因为没有他们赐予的良好基因和身体,是无法在这短暂的一年中完成这项写作的。

罗辽复

于内蒙古大学

1998. 12. 28

目 录

《科学专著丛书》序

序 言

第 1 章 生物手性的起源·····	1
§ 1.1 引言——生命进化时标和地外生命问题 ·····	1
§ 1.2 分子手性和光活性 ·····	5
§ 1.3 光活性的微观理论·····	11
§ 1.4 分子手性的起源和发展·····	14
§ 1.5 β 衰变电子和手性分子的相互作用 ·····	25
§ 1.6 化学反应动力学的对称性破缺·····	31
§ 1.7 手性聚合链的合作产生·····	36
§ 1.8 手性系统的演化问题·····	38
§ 1.9 前 RNA 世界和光活性 展望 ·····	50
参考文献 ·····	53
第 2 章 遗传密码的逻辑和进化 ·····	56
§ 2.1 遗传密码的发现·····	56
§ 2.2 遗传密码的进化和古密码的可能形式·····	61
§ 2.3 遗传密码的对称性·····	67
§ 2.4 遗传密码的阴阳对偶性·····	72
§ 2.5 遗传密码的简并规则·····	78
§ 2.6 终止密码子在密码表上的分布·····	87
§ 2.7 密码表中氨基酸的亲水-疏水畴 ·····	93
§ 2.8 普适密码的反常,氨基酸的简并度·····	102
§ 2.9 遗传密码突变危险性的全局极小化 ·····	107

§ 2.10 密码表突变危险度的进一步讨论	114
§ 2.11 密码子使用频率	118
§ 2.12 氨基酸突变的统计理论	127
参考文献	133
第 3 章 序列的进化和基因信息学	135
§ 3.1 群体的基因动力学	135
§ 3.2 分子进化速率	149
§ 3.3 基因组中信息的积累	157
§ 3.4 基因组的尺度和结构	161
§ 3.5 核苷酸的统计关联及其短程为主性	168
§ 3.6 基因序列信息参数的进化相关性	183
§ 3.7 基因序列中的偏好模	189
§ 3.8 序列的谱分析与碱基长程关联	201
§ 3.9 序列的阅读框架和不均匀性	210
§ 3.10 DNA 行走与分维	216
§ 3.11 核苷酸频数的图示	223
§ 3.12 马尔可夫链与核酸概率矩阵	227
参考文献	235
第 4 章 序列的进化和基因信息学(续)	239
§ 4.1 序列进化机制——最大信息原理(MIP)	239
§ 4.2 分子进化树	247
§ 4.3 基因的复制、移位和横向进化	258
§ 4.4 DNA 保守位点和双螺旋结构局部偏差	265
§ 4.5 序列进化方程	277
§ 4.6 序列的复杂性	302
§ 4.7 基因表达、密码子使用与碱基关联	310
§ 4.8 构象动力学和蛋白质折叠速率	323
§ 4.9 从序列到结构的反向生物学路线	331
§ 4.10 蛋白质的框架结构 结构进化的保守性	346

参考文献·····	365
结束语:分子进化的几个普适性质 ·····	369
附录·····	371
主题索引·····	376

第 1 章 生物手性的起源

§ 1.1 引言——生命进化时标和地外生命问题

生命的发生是宇宙演化中的一个特异事件。这一事件产生在地球上,它是在什么年代产生的?除了地球以外,其他天体上有没有生命?让我们从这两个问题谈起。

据古生物学的研究,从不同地质年代的化石中可以找到生命进化的链条。从化石(包括微观化石)中显示的生命最早痕迹大约在 35 亿年前^[1]。从沉积岩石中碳同位素 ^{12}C 和 ^{13}C 含量比的测定,还可以把生命发生的年代提早到 38 亿年前^[2]。碳有两种稳定同位素—— ^{12}C 和 ^{13}C 。由于运动学效应,光合作用的固碳过程更有利于 ^{12}C 。因此植物中含较多的 ^{12}C ,而地球表面的其他碳氧化物中含较多的 ^{13}C 。 $^{12}\text{C}/^{13}\text{C}$ 比值就成为光合作用的标志。1987—1988 年 Schidlowski 等发现,35 亿年前到 38 亿年前沉积岩中有机碳的 $^{12}\text{C}/^{13}\text{C}$ 比明显高于一般的碳酸盐,从而证明了 38 亿年前的地球上已经存在着光合作用的生态系统。据化石的记录,人们把 35 亿年前到 13 亿年前称为细菌和蓝藻时代。从 13 亿年前到 4.4 亿年前(包括了地质年代古生代的寒武纪和奥陶纪),是真核藻类和无脊椎动物时代。从 4.4 亿年前到 3.5 亿年前(志留纪和泥盆纪),是裸蕨植物和鱼类时代。从 3.5 亿年前到 2.7 亿年前(石炭纪),是蕨类和两栖动物时代。从 2.7 亿年前到 7000 万年前(古生代的二迭纪和全部中生代,即三迭纪、侏罗纪和白垩纪),是裸子植物和爬行动物(恐龙)时代。从 7000 万年前到 300 万年前(新生代的第三纪),是被子植物和哺乳动物时代。在 300 万年前(新生代的第四纪)地球上开始有了人类,进入了人类时代。生物进化和地质纪年的平行性,说明了物种的形成及延续和它的生存环境的统一性。

地球的形成可以追溯到 46 亿年前。从 46 亿年前到 39 亿年前,这是一个太阳系初始形成和星胚生长的阶段,地球受到外界微小星体(宇宙尘、彗星、小行星、陨星等)的连续轰击^[3]。从这个阶段的结束到生命的开始,时间间隔是很短的,Miller 认为前生物阶段的化学反应(例如合成氨基酸等有机物)进行很快,大约只需要不到 1 000 万年的时间,甚至可能短至 1 万年^[4, 5]。当然,第一代的生物“爆炸”式地诞生了,那还不是微化石中发现的 35 亿年前的类蓝藻细菌那样的比较完整生命。著名粒子物理学家 Gell-Mann 在《夸克与美洲豹》一书中把生命看作一种进化的复杂适应系统。它不断获取环境及自身与环境间相互作用的信息,以此为基础在实际中采取相应的行动。通过不同模式的竞争和自然选择,使系统逐步向高级进化。在进化中模式要不断更新换代。对于生命来说,如果每代模式替换需要 1 000 万年,那么 3~4 亿年对应于几十代模式替换。看来,这么多的变革对于形成一个基本完整的单细胞生物来说是必要的。

从天文学的观点看,太阳系的内层空间缺少生命所必需的化学组分:碳、氮、水等。例如,地球上的碳就比天王星、海王星低几个数量级。外层空间富含这些成分,但是温度又太低,不适于水以液态存在,也不利于形成生命。这是一个矛盾。为了在地球上产生生命,必须依靠彗星的撞击或其他星际物质的积聚,把大量生命组分带到地球上来。实际上,原始地球上的有机分子来源有三个途径:一是靠地球外天体的直接输送;二是靠地球受巨大撞击后引发的有机合成;三是靠其他能源(紫外辐射、电离放电等)作用下的有机合成。可以估计这三者的贡献,证明它们都是重要的^[6]。由此看来,地球上的生命演化和太阳系内的物质运动密切相关。

有生必有死。如何估计地球上生命的前途命运? 假定一个过程发生于两个时刻即 t_a (开始)和 t_b (终结)之间,观察者的现时刻为 t_0 。假定 t_0 没有任何特殊,它随机地分布于 t_a 和 t_b 间(这个假定很关键)。记

$$r_1 = \frac{t_0 - t_a}{t_b - t_a},$$

则 r_1 是一个 0~1 间的随机数。据数理统计, r_1 处于 $0.025 < r_1 < 0.975$ 的概率为 $p = 0.95$ 。对于观察者来说,未来的时间记作 $t_f = t_b - t_0$, 过去的时间记作 $t_p = t_0 - t_a$ 。上述结果意味着

$$\frac{1}{39}t_p < t_f < 39t_p, \quad (95\% \text{ 置信度})$$

同样地可有

$$\frac{1}{3}t_p < t_f < 3t_p. \quad (50\% \text{ 置信度})$$

上面的论证称为 Delta t 证法^[7]。例如,内蒙古大学的历史为 41 年, $t_p = 41$ 年,可用 Delta t 证法预测它还能存在的时间 t_f 在 1.05 年与 1599 年之间。人类的 t_p 大约 20 万年,预测其未来寿命为 5 100 年与 780 万年之间(95%置信度)。地球上生命的 $t_p = 38$ 亿年,预测还能存活的时间为 0.97 亿年到 1 482 亿年之间(95%置信度)。后者已远远超过宇宙的年龄,我们可以高枕无忧。

生命是高度有序的物质体系,它要求其组成分子也是高度有序的。1894 年 Fisher 提出了锁钥假设,他认为有效率的生物化学必须是单一手性的,即只用一种手性的对映异构体。手性分子指结构上具有空间不对称性(镜像不对称性或左右不对称性)的分子。例如,氨基酸和核酸中的核糖都是手性分子,有 L 型(左手性)和 D 型(右手性)之分。手性分子正好适应了体系具有高度有序性的要求。事实上,尽管地球上有多至数百万种的生物,但它们都是由 L 型氨基酸和 D 型核糖组成的。生物分子手性的起源一直是生命起源问题的核心内容之一。人们普遍认为,单一手性的实现必须在前生物进化阶段或生物进化的最初阶段完成^[8]。所以,生命进化的时标也给出了从混合手性进化到单一手性的时间尺度。

再来看看地外生命问题。完全像地球上的那种建立在碳和水的生物化学基础上的生物,只适合于在中等温度下生存,例如 -80°C 到 100°C 。某些地球生物能在液氮温度(-195°C)甚至液氦温度(-269°C)下被保存,但没有生物能在这种温度下生长。因为在太低的温度下,重要的生化过程进行太慢,并且只有少数几种溶剂能保持液体状态。在我们的太阳系中,水星很热而且没有大气,不可能有生命。天王星、海王星和冥王星太冷,也不可能有生命。木星和土星主要由氢和氦两种气体组成,若有生命,只能停留在具有适当温度的某一局部大气层中,这种可能性是不大的。金星的大气主要是 CO_2 , 表面温度过高(近 500°C),也不适于生命存在。因此太阳系中除地球外,可能存在生命的行星是火星。火星赤道表面温度在

-100°C 到 30°C 间,大气中主要是 CO_2 和水,但明显缺氮,这对生命是不利的。若大气中含 0.1% 的 N_2 ,就可能维持生命,目前资料还不排斥此数据。不过从近年火星航天器的探寻实验来看,还没有找到生命的迹象。有人认为,早期火星的环境与原始地球相似,可能那时有生命(如在一个标号为 ALH84001 的火星陨石缝隙中发现有丰富的多环芳香烃化合物,有人认为可能与过去的生命活动有关^[9])。另外一个引人注意的天体是土卫六 Titan。Titan 大气中富含氮和甲烷,表面上有乙烷、甲烷等液体组成的海洋,是一个研究有机物前生物进化的理想实验室^[10]。彗星中含有由 C、H、O、N 等元素组成的有机物,并且不同彗星的化学组成不同,对于太阳系内生命的诞生可能有影响和贡献。

太阳系内除地球外迄今还未发现生命存在的证据。那么太阳系外的情况怎样呢? Miller 和 Orgel 在《地球上生命的起源》(1974)中曾这样论述^[11]:我们的银河系大约有 10^{11} 个恒星,考虑到恒星形成速率及其有行星的平均数,行星中有适于生命环境和已发生了生命的分数,有生命行星中已有智慧生命的行星分数,在有智慧生命的行星中文明程度已经达到无线电通信的行星分数,以及每种文明的寿命,便可求出有通信文明行星的数目。但由于每个分数都有很大的不确定性,据此求得我们银河系中具有通信文明的行星可能数目在 10^{-7} (最悲观的估计)到 10^9 (最乐观的估计)之间! 没能得到任何有实际意义的结论!

费米认为:如果银河系中具有智慧生物的星体足够多,至少其中之一已进行星际殖民并遍布全银河。然而地球上没有他们的足迹,故地外智慧生物在银河系中是不存在的。

哥白尼认为:如果没有特殊的理由,作这样的假定——我们处于宇宙中的特殊地位——是错误的。根据哥白尼原理,地球不是宇宙的中心,人类不是宇宙中唯一的智慧生物。

两位科学大师的看法恰好相反!地外智慧生命的问题一直在讨论着,我们见到最近在 *Nature* 上发表的一种看法是:银河系中具有通信文明的星体数 ≤ 120 ^[12]。

由于射电天文学的发展,近年已发现了很多星际空间的有机分子,包括一些前生物物质的前体。例如,氨基酸和嘌呤的前生物合成中的氰化氢、甘氨酸和各种糖类的前生物合成的前体甲醛、胞嘧啶和尿嘧啶合成的

前体丙炔腈等。不仅如此,人们还在陨石上发现了多种氨基酸。特别有名的是1969年9月28日的默奇孙陨石上的几十种氨基酸,并且有充分证据说明它们不是地球的污染物^[1]。这说明宇宙空间、太阳系内到处有生命的组成成分。但是对默奇孙陨石上的74种氨基酸进行构型分析,发现带有不对称碳的氨基酸都是消旋混合物,即L型(左手型)和D型(右手型)等量出现,和地球上生物体内的氨基酸具有一定的手性(L型)相违背。分子手性可以作为生命微观基础的指示灯,因此这个实验资料并未支持地外生命的存在,而是地外生命的否定。

探寻地外生命的否定结果给我们的启示是:生命起源是发生在地球上的过程。外部空间提供了部分物质和能源,但没有向地球输送生命,甚至它的基础——手性物质。

惠更斯在1670年写过一本书,书名为《关于行星世界的新猜测:居民和物产》,书中论证了木星上产麻。论证如下:木星有四个月亮。为何有四个月亮?必然有其目的。地球的月亮可作海员的导航之用,木星的四个卫星必然意味着木星上有众多的航海者。航海者需要船,船需要帆,帆需要绳,绳需要麻,因此木星上有麻。惠更斯是伟大的物理学家,光的波动理论的建立者,土卫六Titan的发现者。但是他的上述论证和逻辑,今天我们读起来不禁要失声摇头。当然,这些错误并不妨碍他是一位真正的科学巨人。这只是说明科学和文化在近几百年中以怎样惊人的速度和规模发展着和传播着。明日视今犹如今之视昔。希望今天关于地球外生命的种种推测,日后不要像木星上的大麻的论证那样令人发笑。

爱因斯坦说过,让我们每个人从莱辛的名言中得到鼓舞——“对真理的追求比对真理的占有更可贵”。我们将继续探索地球上生命演化的关键步骤,并且首先把焦点集中在手性的起源和演化上。

§ 1.2 分子手性和光活性

分子手性是指分子结构的左右不对称性(镜像不对称性)。最早研究这种不对称性的是巴斯德。19世纪中叶,年轻的巴斯德发现,酒石酸盐在偏振光中的不同反应是由于分子结构上的立体异构现象(左右不对称

性)。他认为“一个不对称现象必须有一个不对称原因”。他的这项研究对立体化学的发展起了决定性作用。巴斯德还十分注意这种立体异构和生命过程的关系。他说:“如果没有生命,就不能产生不对称分子。”其涵义是:在无生命的自然界中,左右不对称的两种对映异构体是等量存在的。而在生物体中,左手性的和右手性的分子是不等量的,总有一种占绝对优势。因此,靠生命为融媒,才能产生不对称分子。

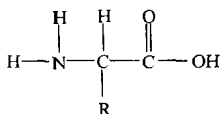
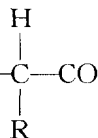


图 1.1 氨基酸的分子结构

蛋白质和核酸是最重要的生物大分子,蛋白质中的氨基酸和核酸中的核糖都是具有手性的。以蛋白质为例,它是由氨基酸组成的长链。氨基酸的分子量约 100,分子结构如图 1.1 所示。

氨基酸之间的连接称为肽键,它是靠两个氨基酸相连时脱去 1 分子水(H_2O)而实现的(图 1.2)。图中 $\uparrow$ 所指位置就是两个氨基酸相连的肽键。蛋白质分子中每一氨基酸单位



称为残基,R 称为侧链。R 有 20 种。也就是说,构成蛋白质的氨基酸有 20 种,它的符号、命名见表 1.1。除了这 20 种标准氨基酸外,还有一些它们的衍生物,如羟基脯氨酸,存在于骨有机质中;羟基赖氨酸,存在于某些酶中。

表 1.1 氨基酸的符号和性质

符号	缩写	氨基酸	疏水性	符号	缩写	氨基酸	疏水性
A	Ala	丙氨酸	疏水	M	Met	甲硫氨酸	疏水
C	Cys	半胱氨酸	疏水*	N	Asn	天冬酰胺	强极性
D	Asp	天冬氨酸	强极性,负电	P	Pro	脯氨酸	弱疏水*
E	Glu	谷氨酸	强极性,负电	Q	Gln	谷氨酰胺	强极性
F	Phe	苯丙氨酸	强疏水	R	Arg	精氨酸	强极性,正电
G	Gly	甘氨酸	弱极性*	S	Ser	丝氨酸	极性
H	His	组氨酸	强极性,正电	T	Thr	苏氨酸	极性*
I	Ile	异亮氨酸	强疏水	V	Val	缬氨酸	强疏水
K	Lys	赖氨酸	强极性,正电	W	Trp	色氨酸	疏水*
L	Leu	亮氨酸	强疏水	Y	Tyr	酪氨酸	极性*

疏水性有多种标度,*号表示可能有不同归类。