

从分子生物学的历程看学科交叉*

——纪念金螺旋论文发表 50 周年

赵凯华[†]

(北京大学物理学院 北京 100871)

摘要 文章从遗传学到细胞学、从生物化学到分子生物学,一步步回顾了基因概念及其载体的具体化,直到 DNA 双螺旋模型建立和遗传密码破译的整个历程,从中分析了生物学家、化学家和物理学家各自发挥的作用。

关键词 基因,分子结构,DNA 双螺旋,遗传密码

Crossfertilization of disciplines as seen from the development of molecular biology ——50th anniversary of the golden helix

ZHAO Kai-Hua[†]

(School of Physics, Peking University, Beijing 100871, China)

Abstract The development of the concept of genes and their carriers, up to the discovery of the DNA double helix and genetic codes, is reviewed from genetics to cytology, from biochemistry to molecular biology, while the role played by biologists, chemists and physicists is examined.

Key words gene, molecular structure, DNA double helix, genetic codes

很少有一项科学成果像 DNA 分子双螺旋结构的发现那样轰动和广为人知。由 DNA 分子双螺旋结构的发现引出了遗传密码的破译并开创了遗传工程,对人类和社会产生了深远的影响。可以说,这项成果与相对论、量子力学一样,同是 20 世纪最重要的科学发现之一。有关 DNA 分子双螺旋结构的论文共四篇,1953 年发表在 Nature 杂志上,前三篇发表于 4 月份,后一篇发表于 5 月份,至今整整 50 周年。这里涉及四位作者:Wilkins M H F 和 Crick F H C 是物理学家,Franklin R E 是化学家,Watson J D 是生物学家。除 Franklin 女士患子宫癌于 1958 年英年早逝外,其他三位学者荣获 1962 年诺贝尔生理学或医学奖。这项成果是学科交叉与合作的典范。

1 从遗传学到细胞学

遗传学(genetics)是从一位奥地利的修道院修士孟德尔(Mende Gregor J)开始的,在他之前只有些不系统的植物杂交试验。植物有多种的相对性状,

如豌豆有红花/白花、高植株/矮植株、花腋生/花顶生等。1854 年夏天,孟德尔从严格控制的豌豆杂交实验中首先提出了遗传因子(hereditary factor,现称基因 gene)的概念,并建立了数量化的定律——孟德尔定律(1866 年发表)。若在多种性状中选择一种,譬如花的颜色,开红花的和开白花的亲本豌豆杂交出子一代(F_1),结果全部开红花。再以 F_1 自花授粉方式培育出子二代(F_2)结果开红花的概率为 $3/4$,开白花的概率为 $1/4$ 。若用 F_1 代中开红花的与开白花的杂交,下一代开红花和开白花的概率分别为 $2/3$ 和 $1/3$ 。孟德尔提出:植物的每一性状(如开花的颜色)由一对遗传因子决定。令 C 和 c 分别代表开红花和开白花的遗传因子,开红花和开白花的亲本携带的遗传因子对分别为 CC 和 cc。杂交时亲本的遗传因子对拆开,与对方的随机组合(分离定律),于是 F_1 代的遗传因子对为 Cc 和 cC 的概率各半, F_2

* 2003-03-24 收到初稿,2003-05-12 修回

† E-mail khzhao@pku.edu.cn

代的遗传因子对为 CC、Cc、cC 和 cc 的概率各 1/4。孟德尔假定：性状有显性(dominant)和隐性(recessive)之分，红色是显性，白色是隐性。遗传因子对为 CC、Cc、cC 的都开红花，只有 cc 开白花。所以 F₁ 代全部开红花，F₂ 代中 3/4 开红花，1/4 开白花(见图 1)。若以 CC、Cc、cC 各 1/3 与 cc 杂交，则形成 Cc 和 cC 组合(红花)的概率为 $1/3 + (1/3) \times (1/2) + (1/3) \times (1/2) = 2/3$ ，形成 cc 组合(白花)的概率为 1/3(见图 2)。这样就解释了上述豌豆杂交的实验结果。

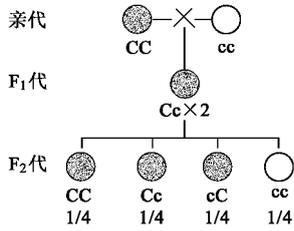


图 1 孟德尔定律

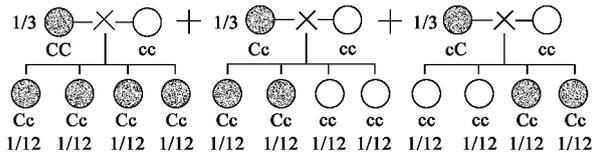


图 2 F₂ 代杂交

孟德尔假定与每一对相对性状对应的有一遗传因子对，它们在后代中传播的概率分布是相互独立的(自由组合定律)。这一假定为他本人的豌豆杂交实验所证实，但 20 世纪的许多实验表明，这一规律不是普遍的。

孟德尔的论文被冷落了 30 多年，直到 20 世纪初，他的遗传学说才被三位植物学家(de Vries H, von Tschermak E, Weismann A)各自独立地重新发现。

随着显微镜的进步，生物学家对细胞的观察愈来愈细致。科学家开始关心孟德尔遗传因子的载体是什么？19 世纪后半叶细胞分裂和受精时两个配子(精子和卵子)的融合已被发现。使用染色剂提高细微结构的可见性，生物学家在细胞核内发现了一种线状体，它们被命名为染色体(chromosome)。在细胞分裂前每根染色体变成双股，分裂后的子细胞各持其中的一股，于是染色体便成为遗传因子载体的候选者。但是，1887 年 Weismann 提出受精卵中的染色体数目应该不变，所以组成它的配子中染色体数目应该减半。他把细胞分为体细胞(somatic cell)

和生殖细胞(germ cell)两类，后者实行减数分裂(meiosis)，配子融合后形成正常数目染色体的细胞。几年后 Weismann 的想法被显微镜的观察所证实，如图 3 所示。为了画面简洁，图中只画了一对染色体，实际上减数分裂分两个阶段。细胞中的染色体一半来自父本，一半来自母本。第一阶段每根染色体变成双股，来自父本和来自母本的染色体单股(称为染色单体 chromatid)在随机地点交叉，并置换被交叉点分割的片段，然后分裂成两个细胞。第二阶段每个细胞再减数分裂，最后形成四个染色体数目减半的细胞。这四个细胞中的单倍染色体是经过杂交的，随机地继承了父本或母本的染色体片段。



图 3 染色体的杂交

不同物种的染色体数目从几对到几百对，差别悬殊，而且染色体数目的多寡与物种的进化程度没有必然联系。显然不可能一条染色体对应一个孟德尔遗传因子。1926 年，美国生物学家摩根(Morgan T H)提出，每条染色体包含大量的遗传因子。不久在某些蝇类的唾腺细胞中发现染色体有横向明暗带，它们的变化可以与遗传突变一样被 X 射线的照射所加速。这就证实了摩根的猜想。摩根选取果蝇做实验材料，发现一些遗传因子，如孟德尔所说，是彼此独立的，但另一些遗传因子则有不同程度的连锁概率。1928 年，摩根提出了基因论，认为基因组成一些连锁群分布在各条染色体上，在同一条染色体上基因呈线性排列。不在同根染色体上的基因之间没有连锁效应，同一根染色体上的基因之间连锁概率的大小与它们在染色体上的距离有关。摩根据此测定了一些基因在染色体上排列顺序的图谱。1933 年，摩根获诺贝尔生理学或医学奖。

2 染色体的化学成分

染色体的主要成分是一种特殊的蛋白质——组蛋白(histone)和 DNA，此外还有痕量的 RNA。染色体内基因的载体是蛋白质还是 DNA 或 RNA，成为遗传生物化学长期关注的问题。下面我们先看看蛋白质、DNA 和 RNA 本身的化学成分。

蛋白质的英文名字 protein 来源于希腊文 proteios，意思是“一流的”。1904 年 Emile Fischer 证明，蛋白质是氨基酸(amino acid)经脱水肽键连接起来

的多肽链 (polypeptide chain) ,其分子式和链接过程如下 :



其中 R 是某种氨基酸的残基. 在生物体内的蛋白质中有 20 种“标准的”氨基酸,它们的名称和代号见表 1.

表 1 蛋白质中的 20 种氨基酸

名称	代号	名称	代号
丙氨酸 alanine	Ala	亮氨酸 leucine	Leu
精氨酸 arginine	Arg	赖氨酸 lysine	Lys
天冬氨酸 asparagine	Asn	蛋氨酸 methionine	Met
天冬酰胺 aspartate	Asp	苯丙氨酸 phenylalanine	Phe
半胱氨酸 cysteine	Cys	脯氨酸 proline	Pro
谷氨酰胺 glutamate	Gln	丝氨酸 serine	Ser
谷氨酸 glutamine	Glu	苏氨酸 threonine	Thr
甘氨酸 glycine	Gly	色氨酸 tryptophan	Trp
组氨酸 histidine	His	酪氨酸 tyrosine	Tyr
异亮氨酸 isoleucine	Ilu	缬氨酸 valine	Val

虽然每种氨基酸化学成分各异,但它们的分子有很强的相似性. 它们都有一个羧基 COOH 酸性端和一个氨基 NH₂ 碱性端. 两氨基酸的酸性端与碱性端结合,从前者去掉一个 H,从后者去掉一个 OH,即脱出一个水分子 H₂O,连接起来就是一个肽键. 这就是前面说的脱水缩合的含义. 蛋白质分子是几十个、上百个氨基酸连接起来的长链. 20 世纪 50 年代初,发明了测定蛋白质分子中氨基酸排列顺序的色层分析法,1954 年, Frederick Sanger 测定了第一个蛋白质——胰岛素分子中 51 个氨基酸的顺序.

DNA 全名是脱氧核糖核酸(deoxyribonucleic acid), RNA 全名是核糖核酸(ribonucleic acid),二者都是核酸(nucleic acid). 核酸是由核苷酸(nucleotide)链接起来的多聚体,核苷酸则由核苷(nucleoside)和磷酸组成,核苷由碱基和脱氧核糖(deoxyribose)或核糖(ribose)组成. DNA 和 RNA 中的碱基都有四种,其中有三种是相同的,有一种不同(详见表 2). 这些成分分子式如图 4 所示.

1898 年, Miescher F 发现细胞内含磷物质,取名 nuclein,后改名为核酸. 在小牛的胸腺和鲑鱼的

表 2 DNA 和 RNA 的化学成分

	核苷酸 (nucleotide)	
	糖	碱基
DNA	脱氧核糖 deoxyribose	腺嘌呤 guanine, G 鸟嘌呤 cytosine, C 胸腺嘧啶 thymine, T
RNA	核糖 ribose	腺嘌呤 adenine, A 鸟嘌呤 guanine, G 胞嘧啶 cytosine, C 尿嘧啶 uracil, U

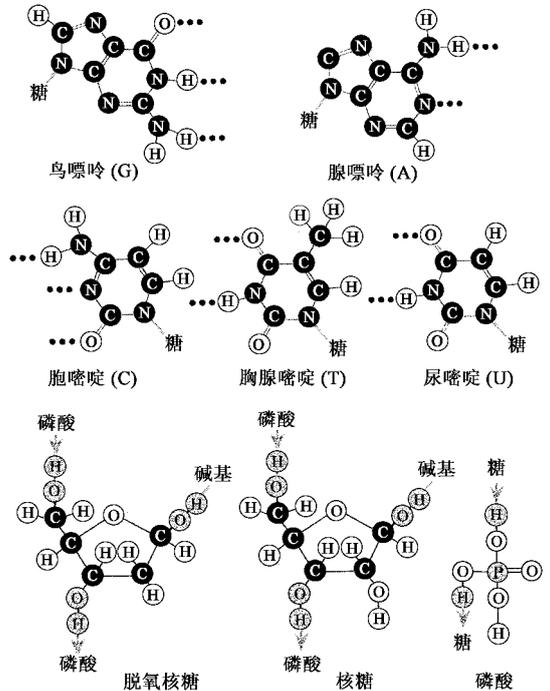


图 4 核苷酸组分的分子结构

精子中这种物质特别多. Kossel A 发现动植物细胞内四个碱基之一不同, Levene P 发现糖也不同,即动物细胞内的核酸是 DNA,植物细胞内的核酸是 RNA. 1914 年, Feulgen R 用分别只给 DNA 和 RNA 着色的染料进行研究,结果表明,二者在动植物细胞内都有, DNA 在细胞核内, RNA 主要在细胞质内. 1933 年, Brachet J 用灵敏的染色剂将 DNA 精确定位,确定它在染色体内.

3 从遗传学到生物化学

回顾遗传学向生物化学迈出的步伐,我们要追溯到 19—20 世纪之交. 1898 年,英国医生 Garrod A 接收了一个患尿黑酸病(alkaptonuria)的男孩, 1901 年第 5 个患尿黑酸病的婴儿在同一个家庭出生了,孩子的父母是嫡表兄妹. 这种病人的尿一接触空气

就变成黑色. Garrod 认为这种病是由一种稀有的遗传因子引起的,患者体内缺乏一种新陈代谢过程中使苯环断裂的酶.

1941 年,美国 George Beadle 和 Edward Tatum 提出“一个基因对应一个酶”的假说,这是遗传学朝着生物化学迈进的重大的一步.他们的实验材料是链孢霉(*neurospora*)真菌,用辐射处理其孢子以产生大量的变异,孢子杂交后把可能发生突变的孢子分别加入到正常培养基和添加了维生素 B₁ 和 B₆ 的培养基中.实验中他们得到了几十个需要某种维生素才能存活营养缺陷型菌株.通过这些菌株的杂交,他们发现每一营养缺陷型的菌株都是由单个基因突变引起的.接着他们又得到一种体内不能合成色氨酸(tryptophan,一种氨基酸)的变异菌株,进一步的研究表明,色氨酸生物合成的每一步都是由一个不同的基因控制的.这就是说“一个基因一个酶”.Beadle, Tatum 和另一位生物学家 Lederberg J 荣获 1958 年诺贝尔生理学或医学奖.

肺炎球菌(*pneumococcus*)有光滑型(S)和粗糙型(R)的不同菌株,只有 S 型菌株能引起疾病,R 型菌株无害.1928 年,Griffith F 用热杀死的 S 型细菌和活的 R 型细菌注入小鼠体内,小鼠不但死亡,且在它们的血液中发现活的 S 型细菌.若只注入死的 S 型细菌而不含活的 R 型细菌,则小鼠无恙.据信有一种转化因子从死的 S 型细菌转入 R 型细菌,使后者转化为 S 型.从 1935 年起,Avery O T 和他的同事们在离体的条件下完成了转化过程.他们用一系列物理和化学的方法把转化因子中的 DNA、蛋白质和其他物质分开,实验证明,只有 DNA 有转化功能.经过 10 年工夫,Avery 的论文发表于 1944 年.

下面一项重要工作是具有很高声望的“噬菌体研究组”的科学家做出的.噬菌体(phage)是一种感染细菌的病毒.现在我们知道,噬菌体的构造是包藏在蛋白质外壳内的 DNA,它们感染细菌时尾部吸附在细菌上,像注射器那样将自己的 DNA 注射到细菌内,将蛋白质外壳留在细菌外边(见图 5).细菌被感染后本身停止繁殖,最后被溶解,释放出大量噬菌体的后代.1952 年,Alfred Hershey 和 Martha Chase 作噬菌体 T₂ 感染大肠杆菌的实验.噬菌体 T₂ 约含 60% 的蛋白质和 40% 的 DNA,蛋白质中含硫(S)而 DNA 中不含硫,但 99% 以上的磷(P)在 DNA 中.Hershey 和 Chase 分别用放射性同位素³⁵S 和³²P 来标记蛋白质和 DNA.实验第一步分别用含有³⁵S 和³²P 的培养基培养宿主细菌,然后用噬菌体 T₂ 感染它

们,得到分别具有³⁵S 和³²P 原子标记的噬菌体后代.实验第二步用不同标记的噬菌体去感染没有标记的细菌,结果大多数³⁵S 标记留在宿主细胞的外边,而³²P 标记大多数进入宿主细胞内.

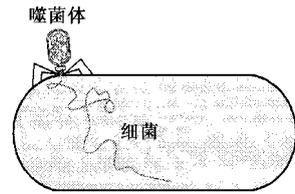


图 5 噬菌体模型

病毒(virus)是介于生物和化学之间的物质,它能够繁殖和突变,也能够结晶.1935 年,Stanley W M 用纯化蛋白质的方法完成了烟草花叶病毒(tobacco mosaic virus, TMV)的提纯和结晶,获得 1946 年诺贝尔化学奖.这项成功是用物理化学方法研究生物现象的代表.1940 年,有人试图重复 Stanley 的实验时发现, TMV 并非纯蛋白质,其中包含 6% 的 RNA.那时人们对 DNA 和 RNA 在遗传中作用的认识是模糊不清的,认为这无关紧要.1956 年,当人们不再怀疑核酸是遗传信息的载体时, Schramm G 和 Fraenkel - Conrat H 为此提供了一份新的佐证.他们使 TMV 在水和苯酚中震荡,将蛋白质和 RNA 分离,用它们分别去感染烟草,结果发现是 RNA 使烟草感染,并产生了病毒的后裔.

上述几个实验直接或间接地证明了,遗传信息的物质载体是核酸(DNA 或 RNA)而不是蛋白质,这无疑是分子生物学最基本的实验基础.载入这一贡献史册的主要有 Avery 和 Hershey, Chase 两组科学家,他们先后相差 8 年,受到了一些不太公平的待遇.在 Avery 时代,人们还没太意识到什么是遗传物质载体,他给了这个不存在的问题一个答案,他的先驱性工作受到相对的冷遇.在 Hershey, Chase 时代,人们已不太怀疑遗传的物质载体是核酸,他们给了这个已有答案的问题一个意料中的答案(他们的实验也确实更具有说服力),受到了热烈的欢迎.这就是历史!

Avery 的工作还是有影响的.在他的成果的影响下, Erwin Chargaff 和他的学生和同事们从事 1947 年 DNA 中碱基含量比例的分析.1949 年, Chargaff 从大量实验数据中忽然悟出,尽管有的生物(如结核杆菌)腺嘌呤(A)和胸腺嘧啶(T)含量很高,有的生物(如某些真菌)鸟嘌呤(G)和胞嘧啶(C)含量很高,即(G+C)/(A+T)的比例(摩尔比)随物种的

变化很大,但 $(G/C) = (A/T) = 1$ 却是一条与物种无关的普遍法则. 这条现称为“Chargaff 定则”的规律在遗传学上的意义是很大的.

4 物理学家的看法

分子生物学的诞生和发展激发了许多物理学家的兴趣,其中不少人成为积极的参与者. 下面谈谈在 20 世纪 30—40 年代两位影响最大的人物——Delbrück 和 Schrödinger.

1932 年夏,年轻的德国理论物理学家 Max Delbrück 为 Niels Bohr 的一次题为“光与生命”的报告所鼓舞,转向了生物学. 在柏林他开始与 Timoféeff N W (遗传学家)和 Zimmer K G (物理学家)合作,他们两位做果蝇的 X 射线诱变实验,Delbrück 作理论分析. 1935 年三人合写一篇论文发表在德国 Göttingen 的一本冷僻的杂志上,此文被谑称为“三人篇(three-man-work)”. 使 Zimmer 惊讶的是, X 射线形式的能量无论剂量怎样小都会诱发基因突变,而一些其他形式能量(譬如加热)则不能.“三人篇”的结论认为,诱发基因突变的元过程是遗传物质的最小感受单元(基因的体积)被“个别击中”而电离或激发的事件,具有量子力学的性质. Delbrück 估计基因的体积内约含 10^3 个原子,比细胞学原先的估计(300Å)³小多了. 一些科学家欣赏这个估计,因为它与蛋白质分子的大小相当. 看来基因的稳定性是由原子间键合力的强度来保证的,基因的突变是分子不同构型之间越过势垒的跃迁.

量子力学创始人之一 Erwin Schrödinger 于 1943 年在英国的 Dublin 作了一次报告,次年出版一本小册子,名为《生命是什么(What is Life)?》. Schrödinger 赞同“三人篇”中 Delbrück 的量子力学观点,认为基因既稳定又能发生突变,这只能用量子论中能量的不连续性和量子跃迁的突发性来说明,经典物理学是不能解释的. Schrödinger 认为,基因的变化是生物大分子同分异构体之间的量子跃迁,其间的能量阈值保证了基因在室温下的稳定性. Schrödinger 提出,基因是一种“非周期性的晶体”,其中包含了压缩的“微型密码”. 遗传密码的概念是 Schrödinger 首先提出来的.

虽然 Delbrück 和 Schrödinger 从物理学得到有关基因的看法是接近的,但他们二人的哲学理念却很不一样. Bohr 将波粒二象性的互补原理运用到生物学中,认为生物学规律和物理、化学规律在生物学中是互补的,两方面都不能单独说明全部生命现象.

Delbrück 的想法是从 Bohr 那里来的,他指望在生物学的研究中能发现新的物理定律. Schrödinger 则坚信生命现象最终是能用物理学和化学来解释的.

虽然 Delbrück 和 Schrödinger 二人的深刻思想在当时并未得到充分的理解,但都对后来产生了深刻的影响,尤其是把许多有才华的年轻物理学家吸引到生命科学的研究中来了.

Delbrück 还建树了一项不朽的功勋,1940 年他和微生物学家 Salvador Luria 发起了上文谈到的那个噬菌体研究组,1943 年 Hershey 加入该组,吸引了许多不同行的科学家加入,在噬菌体方面的研究成绩卓著,三人荣获 1969 年诺贝尔生理学或医学奖.

5 蛋白质的分子结构

化学是 Delbrück 和 Schrödinger 两人共同的盲点,而分子生物学的建立和发展要求对生物大分子的结构有具体而深入的了解,这只得仰仗化学家了.

Linus Pauling 被公认是 20 世纪最伟大的化学家,两次诺贝尔奖金的获得者(1954 化学奖,1962 和平奖). 1925 年他在美国加州理工学院获得博士学位时,已补上了数学和物理训练的不足. 1926 年他访问欧洲一年,正值量子力学开创并蓬勃发展的时期. 1927 年 Heitler 和 London 用量子力学计算了氢分子的问题,首次揭示了共价键的本质——反向自旋电子对在两个原子轨道状态之间的交换. 共价键的理论(或者说计算方法)有二:分子轨函(molecular orbital, MO)法和价键(valence bond, VB)法. 前者忽略了电子间的相互作用,后者将此因素考虑过份. VB 法是 Pauling 和 Slater 分别创立的. Pauling 熟知 MO 法,但偏爱 VB 法. 研究分子结构的结构化学是以量子化学为基础的,可以说,Pauling 是结构化学的奠基人. 翻开一本结构化学的书,你可以看到,其中大部分基本概念和原理,如离子半径、各种化学键的键长、键角、轨道波函数的杂化和分子的立体构型、共振(电子的交换与配对)、电负性,等等,都是 Pauling 提出来的. Pauling 惯用的方法是已有的经验事实与理论计算相结合,他用的量子理论计算往往不很严格,但提出的模型图像和概念鲜明,且很实用. 他写的书《化学键的本质(The Nature of the Chemical Bond)》(1939 年初版,1940、1960 年两次再版)是该领域的经典著作,几十年不失其光辉.

起初,Pauling 主要研究的是无机分子和晶体,20 世纪 30 年代,当他获得一项只能用于生命科学研究的 Rockfeller 基金后,他的兴趣转向了生物大分

子,特别是蛋白质。Pauling 在这个领域内提出了几个重要的开创性概念,一是氢键的生物大分子中的特殊作用,二是生物大分子空间构型的互补性。酶的催化和生理功能以及抗体的免疫功能都具有极强的专一性,它们都来源于分子构型的互补性,而互补分子之间的啮合主要靠氢键。下面我们专门谈一下这两个问题。

氢键的概念是 Latimer 和 Rodelbush 于 1920 年首先提出来的。氢键($X-H\cdots Y$)是处于两个电负性较强的原子 X 和 Y(如 F, O, N 等)之间的 H 原子与 Y 原子的结合力(用虚线 \cdots 表示)。这里原子 X 和 H 之间是共价键(用实线 $-$ 表示),氢键本质上是一种范德瓦耳斯力。分子 $X-H$ 具有较强的电偶极性,它与电负性强的 Y 原子之间的静电吸引力构成了氢键。氢键比共价键差不多要弱一个数量级,键长也大得多,所以室温下容易开合和重组,Pauling 首先指出,氢键在生物化学和生理过程中起重要作用。

1938 年, Jordan P 提出,相同分子之间的量子力学稳定化相互作用可能促进生物分子的合成。1940 年, Pauling 指出,首选的因素应该是分子的互补性而不是全同性。这时他正在考虑免疫学问题,奥地利免疫学家 Landsteiner K 问他抗体和抗原之间的专一性如何解释,经过思考,他得到了他的互补性原理,即抗体与抗原的分子表面的形状有一部分是互补的,它们能够紧密地啮合在一起,沉降下来被排除体外。Pauling 还假定,这种互补分子之间的啮合主要是氢键或其他弱键。生物体内有各种酶(enzyme),它们各自专门对生理过程中某一特定的化学反应进行高效率的催化。Pauling 认为,这也是由于酶与目标物在分子构型上互补所致。所以在 Pauling 看来,生物大分子与无机分子的化学反应不同。在无机分子中原子的个性影响着化学键的性质,而生物化学反应中多半只涉及氢键(通常是氧和氮原子之间的氢键),原子的个性不突出了,重要的是分子的空间构型。

抗体和酶都是特殊的蛋白质,蛋白质分子是氨基酸连接起来的长链。现在我们知道,蛋白质分子的结构有四个层次:在多肽链中氨基酸的排序属一级结构;二级结构是多肽链绕成螺旋(α 螺旋)或反复摺成片层(β 片);三级结构就更复杂,螺旋和螺旋可以再缠绕;四级结构是一个以上多肽链的组合结构。对蛋白质分子的折叠起重要作用的除氢键外,还有半胱氨酸(cys)之间的双硫桥。在生物体内正常的生理环境中,每种蛋白质分子折叠成一定的空间

构型,这是它的自然构型。加热或改变酸碱度,蛋白质可以变性(denaturation)。蛋白质变性时一级结构不变,改变的是高级结构,即氢键等弱键的断裂和重组。变性的蛋白质失去了其“正确”的构型和生理功能。有关蛋白质变性的理论也是 Pauling 首先提出的。Pauling 在蛋白质结构方面最重要的贡献是他的 α 螺旋理论。

1948 年, Pauling 在牛津做访问教授期间,一次因受寒而卧床不起。为了消磨时间,除了看侦探小说外,他拣起十多年前研究过的 α 角蛋白(keratin)结构,企图用纸、笔和直尺来解决问题。多肽链骨架中与氨基酸残基 R 连接的碳叫 α 碳,记作 C_α 。相邻两个 C_α 之间是酰胺,其结构如图 6 所示, C, O 之间的双键是 π 键,不能转动; C, N 之间是单键,但根据 Pauling 多年的经验,这个键不是单纯的 σ 键,上述 π 键有一部分转移到这里来,能量更划算。于是 C, N 间的键也是不能转动的。这样一来,图 6 中的 6 个原子总保持共面,只有 C_α 原子处是可以转动的。

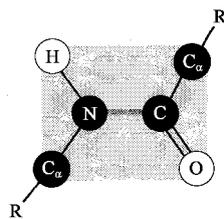


图 6 酰胺键

Pauling 按已知的键长和键角把这一共面结构画在纸上,在 C_α 处一个接一个连接起来,盘旋成螺旋状(图 7)。螺旋是靠氢键来固定的,当某节上的 $N-H$ 与盘旋上去另一节的 O 之间的距离刚好是氢键的长度时,它们之间就建立了氢键连接。从这个模型可以定出,每一圈中含 3.7 个氨基酸残基,螺距 5.4\AA 。由于这数据与前人(Astbury W T)的 X 射线衍射实验数据 5.1\AA 有点差距,Pauling 就把这个问题搁下了。他回到加州后,与合作者一起核对了模型的细节,并试探着一些别的方案,没得到什么新结果。1950 年,英国 Cavendish 实验室 Bragg, Kendrew, Perutz 的一篇文章发表了,Pauling 认为他们允许酰胺中 $C-N$ 键转动是不可接受的。加之此时一些新的实验提供的旁证表明,螺距 5.1\AA 这个数据不那么神圣,Pauling 与合作者们于 1950—1951 年一连发表了几篇文章,详细报道了他们的 α 螺旋模型。人工合成多肽链的 X 射线照片表明, 5.4\AA 的螺距是对的。

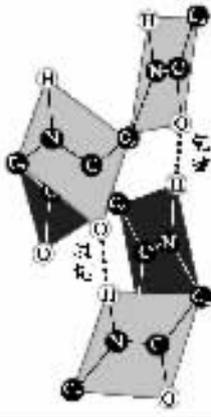


图 7 多肽链骨架排成螺旋状

6 DNA 双螺旋结构

我们将进入本篇叙述的最重要阶段——DNA 双螺旋结构的发现,在此之前先介绍一下 DNA 双螺旋是怎么一回事.

图 8 所示为 Watson 和 Crick 两人 1953 年最后提出的模型,其中外边两股长链相互缠绕,构成骨架.如图 9 所示,长链是磷酸和脱氧核糖连接起来的,左链由上到下,右链由下到上,走向彼此相反.碱基在两股长链之间,构成像梯子一样的横档,每档由一对被氢键(见图中虚线)连接起来的一对碱基构成,它们的分子都是平面环状的,平面与螺旋的轴线垂直(这一点未在图 9 中正确地反映出来,实际情况应该是螺旋轴线在图面内,碱基对的平面与图面垂直).四种碱基的配对方式是严格的:腺嘌呤(A)与胸腺嘧啶(T)结合,鸟嘌呤(G)与胞嘧啶(C)结合,两链上碱基是互补的,一链上碱基的顺序完全决定了另一链上碱基的顺序,组成一套遗传密码.如图 8 所示,梯子的横档(即碱基对)的间隔为 3.4\AA ,螺旋距 34\AA ,即螺旋每盘旋一圈,包含 10 个横档,每档转过 36° 角.

以上也是我们今天对 DNA 双螺旋的认识.当年得到这个认识,知道了 DNA 的化学组成后还要在观念上跨过一系列障碍.首先得坚信 DNA 是遗传物质的载体,其次要相信 DNA 具有螺旋结构,再次要搞清楚螺旋有几股,然后是磷酸和脱氧核糖连接起来的骨架和碱基何者在内何者在外,最后的问题是碱基怎样组合和排列.跨出每一步都是很艰辛的.

获得 DNA 双螺旋结构模型的整个过程大约是两年半(1951 年到 1953 年上半年),主角是两个单位的四个人,即伦敦 King's 学院的 Wilkins 和 Frank-

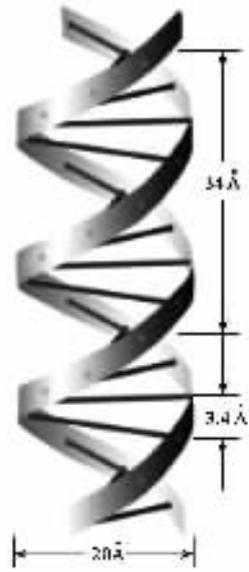


图 8 DNA 双螺旋模型

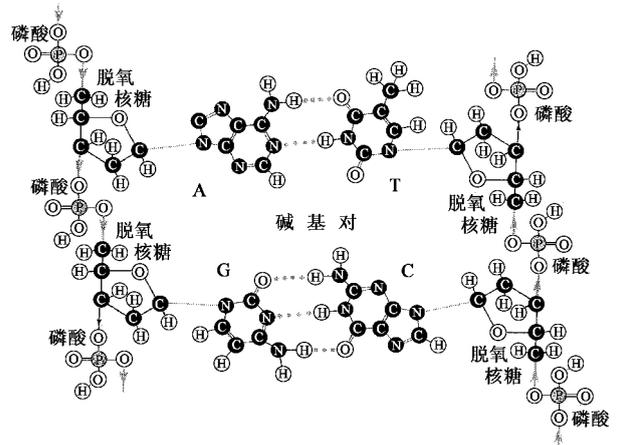


图 9 DNA 的分子结构

lin, 剑桥 Cavendish 实验室的 Click 和 Watson.

Wilkins M H F, 物理学家,二战时参加过制作原子弹的曼哈顿计划,1946 年来到 King's 学院医学科学院生物物理研究室.1950 年 5 月,他参加一次学术会议时得到报告人 Signer R 分发的一小瓶高质量的 DNA 钠盐.回来后他对这材料做了一些拉丝和光学观察,同时大学刚毕业的 Gosling R G 做些 X 射线衍射的研究,那时 Wilkins 尚未搞过 X 射线衍射.是 Wilkins 和他的同事 Stokes A R 首先直接从核酸纤维的研究中提出 DNA 的螺旋结构的,尽管早在 1949 年从挪威来伦敦的研究生 Furberg V S 已从核苷与核苷酸的 X 射线衍射研究中推论出一个 DNA 的单螺旋结构.1951 年夏,Wilkins 请 Stokes 做出螺旋衍射的理论.螺旋的傅里叶变换式含各阶 Bessel

函数,图解显示,衍射极强由原点向外移动,移动方向与子午面的夹角等于螺旋的斜率,在子午面上留下空白.1951年秋,当 Wilkins 看到 Franklin 和 Gosling 拍摄的 DNA 晶体 X 射线衍射照片清晰显示出螺旋的特征时,大为兴奋.

Franklin R E, 物理化学家,做过多年的 X 射线晶体学工作.1951年1月,她来到 King's 学院后即被 Randall J T 爵士委派做 DNA 的 X 射线衍射实验.当时正值 Wilkins 在外,当他回来时,Franklin 误以为委派给她的工作是由她一人主持的,Wilkins 只好让位,把 Signer 给的 DNA 钠盐留给她单独使用. Franklin 和 Gosling 在实验中发现,在不同湿度下 DNA 钠盐有两种形态:晶体(crystalline, A 型)和类晶(paracrystalline, B 型),两者之间的转换是可逆的.1951年9月,她拍摄到那张令 Wilkins 惊叹不已的照片,但她本人却处之淡然. Franklin 是位专业的结构晶体学家,她希望由实验直接定出晶体的结构而不外加任何前提和假设,她不信任直观和猜测.她并不反对螺旋结构,但认为证据尚不充分.1951年11月,在 King's 学院举办的学术讨论会上,Franklin 汇报了她的工作,根据她的 X 射线衍射照片提出了螺旋结构的设想,得到 27\AA 的螺距数据.1952年5月,她拍摄到 DNA 单纤维更清晰的 X 射线衍射照片,她对数据作了澄清:A 型向 B 型转化时纤维伸长 25%,空间周期从 27\AA 增加到 34\AA ,链上碱基层的间隔 3.4\AA ,每个周期内 10 层.她根据 DNA 与水的关系提出磷原子暴露在外面的观点.从她的晶体单胞数据可以看出,DNA 的晶型属单斜晶系, C₂ 对称性.1953年3月中旬,Franklin 离开了 King's 学院,到 Birkbeck 学院去做烟草花叶病毒的研究.

Crick F 于 1937 年伦敦大学学院(University College London)物理系毕业,二战期间他在英国海军部的一个实验室工作,战后受 Schrödinger 的《生命是什么?》一书的影响,转向了生物物理.1949年7月,他到 Cavendish 实验室后加入 Max Perutz 新组建的一个课题组,搞血红蛋白(haemoglobin)结构的研究.

Watson J D, 年在美国 Indiana 大学获动物学博士学位,导师是噬菌体研究组创建人之 Luria. Luria 和 Delbrück 都认为欧洲的科研人员比他们美国同行更富于想像力,他们把 Watson 送到哥本哈根作博士后.在欧期间,1951年5月,他参加了在意大利 Naples 举办的一次学术会议,Wilkins 有关 DNA 晶体的报告引发了 Watson 极大的兴趣,活物质能够

结晶增加了在分子水平上解释生物现象的信心.于是 Watson 想转到英国搞晶体学工作,1951年10月初转到了 Cavendish 实验室,在那里他结识了 Crick,虽然他们在年龄上差 12 岁,却兴趣相投,一见如故,立即开始了他们 DNA 结构的研究.

Crick 从 Pauling 蛋白质 α 螺旋模型的成功得到的启示是,仔细制作的精确模型能够体现正确答案必须满足的限制条件,有时用这种办法可以借助最少量实验证据得出正确的结构.这正是他们(他和 Watson)与 Franklin, Wilkins 科研路线不同的地方.1951年12月,即 Franklin 在 King's 学院讨论会上汇报工作之后一周, Watson 和 Crick 匆匆搭起一个 DNA 模型:三股磷酸钠与糖接起来的长链在中央拧在一起,作为骨架,碱基分布在周围. King's 学院的人应邀来看说,这模型破绽百出,特别是 Watson 把水的含量搞错了,少了一半.这是他们第一次惨败.其后果是双方领导达成协议,把 DNA 的工作留给 King's 学院, Cavendish 实验室还是搞他们的蛋白质研究.如前所述,1952年, King's 学院的人继续作了一些 DNA 的实验,积累了一些新的数据,拍了一些更好的 X 射线衍射照片,详情 Watson 和 Crick 不太知晓.1952年7月,Chargaff 访问剑桥时与 Crick 和 Watson 见过面,由于当时他们对碱基的结构与可能的结合方式还很生疏,Chargaff 定则在他们建模的初期并未起到约束作用.1952年,他们考虑碱基的排列问题是不得要领的.

继 α 螺旋模型的成功之后,1952年, Pauling 也把注意力集中到 DNA 上.他的模型也是三股螺旋在中央,碱基在周围. Pauling 于 1952年12月将论文投出,该文将于 1953年2月发表.他同时写信把这个消息告诉了他在英国的儿子 Peter Pauling, Peter 把这信交给 Watson 和 Crick 传阅.二人急于知道详情,请 Peter 写信给他父亲要手稿,1953年1月底他们看到了手稿. Watson 仔细钻研了手稿,怀疑它是错的.竞争意识很强的 Watson 和 Crick 非常着急,担心老 Pauling 很快发现并改正错误,到那时他们就没有指望了.

Crick 同意 Watson 立即去伦敦把这手稿拿给 Wilkins 看. Wilkins 与 Crick 的关系很好,但与 Franklin 不融洽,虽然他们的实验室在同一层楼内,却很少互通消息.这次 Watson 来找他时,他把私自复制的 Franklin 上一年 5 月的报告拿给 Watson 看, Franklin 的那张 B 型 DNA 的 X 射线衍射照片显示的螺旋结构如此清晰,使 Watson 看得目瞪口呆.这

时他才明白 :DNA 由 A 型转 B 型时纤维伸长 20% ,螺距增加到 34Å ,每股 10 个碱基 ,纤维的直径 20Å .在回剑桥的火车上他在报纸的空边上作了些估算 ,认为螺旋的股数应为 2. 当 Crick 听了 Watson 的汇报后同意着手双螺旋模型的试制 ,不过他还有些犹豫 ,除非得到 DNA 纤维具有 C2 对称性的信息. 幸运的是 ,上一年 12 月 King's 学院给医学研究委员会 (MRC) 写了一份报告 ,其中包括 Franklin 的工作总结. Randall 将它分发给所有生物物理研究委员会的委员 ,其中有 Perutz. 1953 年 2 月的第二周 ,Perutz 将这份秘密报告给了 Crick. Crick 看到其中确有 DNA 晶体属单斜晶系 C2 对称性的分析(这一点是 Watson 所不懂的) ,由空间群理论得知 ,螺旋不仅是双股的 ,且走向相反. Crick 很兴奋 ,Watson 可以放心地搭他的双股螺旋模型了 ,但 Watson 还要为碱基如何配置发愁.

这已是 2 月的第三周了 ,Watson 从一本核酸生物化学的书上抄下碱基的结构式 ,查阅了一些文献 ,认识到一直被他和 Crick 排除在外的氢键在碱基配对中的重要性. 2 月 20 日 ,他设计出一种同类(腺嘌呤配腺嘌呤 ,鸟嘌呤配鸟嘌呤)通过氢键结合的方案. 与他同办公室的 Jerry Donohue(原在 Pauling 处专搞氢键化合物的专家)看了他的“杰作”之后指出 ,因氢原子的位置不同 ,碱基有许多互变异构体 (tautomer) ,他抄的碱基结构式是错误的一种. Crick 也认为这模型不符合 C2 对称性. 2 月 27 日 ,Watson 摆弄他的硬纸壳碱基模型时突然意识到 ,若使腺嘌呤 (A) 与胸腺嘧啶 (T) 配对 ,鸟嘌呤 (G) 与胞嘧啶 (C) 配对 ,两者形状一样 ,就不必担心两股螺旋之间距离忽长忽短不均匀了. 再问 Donohue 对这个新的模型有什么意见 ,回答说没有. Crick 认为 ,这新模型不仅符合 C2 对称性 ,也解释了 Chargaff 定则. 成功了 ! 剩下的事情就是构建一个符合立体化学要求的模型 ,用铅垂和皮尺测量模型中每个原子的坐标了.

1953 年 3 月上旬 ,Crick 及时地把情况通报给 Wilkins , Watson 也把详情写信告诉 Delbrück ,并转告了 Pauling. Wilkins 肯定了他们的工作 ,并建议自己和 Franklin 都应各写一篇实验性的文章一起发表 ,作为佐证. 4 月初三篇文章到达 Nature 编辑部 ,4 月 25 日同期发表了. Crick 感到 4 月份的文章对他们的模型在遗传学上的意义阐述不够 ,他们又在 Nature 上补了一篇文章 ,5 月 30 日发表. 这就是 50 年前四篇划时代的历史性科学文献.

DNA 双螺旋结构模型被充分接受的过程是需

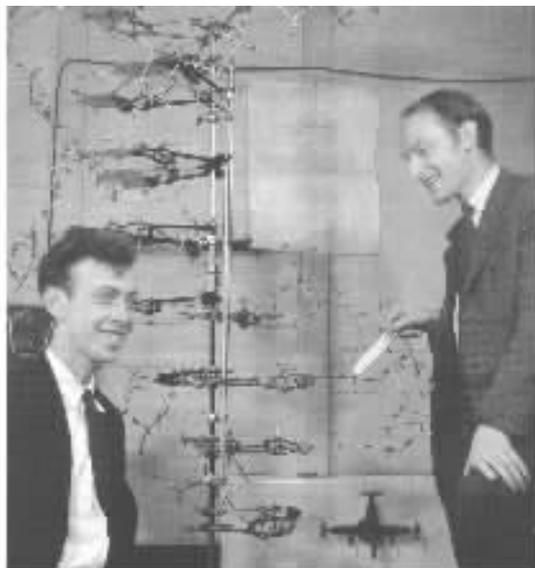


图 10 Watson(左)、Crick 和他们的模型

要时日的. 直到 1962 年诺贝尔生理或医学奖才授予 Wilkins、Watson 和 Crick. 人们公认 ,Franklin 的功绩决不亚于获奖者 ,遗憾的是那时她已不在人世了. 顺便说起 ,她对 Watson 和 Crick 暗中使用了她的重要数据始终不知情 ,Watson 于 20 世纪 70 年代才公开透露这一点.

7 遗传密码

Watson 和 Crick 在 1953 年的第二篇论文中写道 : 碱基的精确顺序就是携带有遗传信息的密码. ”这种阐述仅表示“密码”的概念已被人们接受 ,然而那时生物学界对蛋白质结构的了解尚很模糊 ,对密码怎样起作用并无具体的构想. 1953 年夏 ,他们意外地收到一封核物理和宇宙学家 George Gamow 的来信. Gamow 说 ,在看到他们在 Nature 上的文章后很快断定 ,沿 DNA 螺旋链碱基序列的排列构成 20 种不同形状的孔洞 ,与蛋白质中 20 种氨基酸分子的形状像锁和钥匙那样匹配. 所以 DNA 本身就是合成蛋白质的模板. Gamow 还提出“三联体密码子”的想法 ,因碱基有 4 种 ,至少相连 3 个碱基才能为 20 种氨基酸编码. Crick 看了 Gamow 的新奇想法后第一个反应是想证明他是错的. 为什么是 20 种氨基酸 ? Gamow 显然漏掉了一些. Crick 仔细把蛋白质中存在的氨基酸分一下类 :在多数蛋白质中都有的算作标准的 ,只在少数几种不常见的蛋白质中存在的归入异常类. 他惊讶地发现 ,标准类的氨基酸刚好 20 种 ! DNA 中碱基对构成密码子与 20 种氨基酸对

应,控制着蛋白质的合成,Gamow 这一思想还是被接受了。

此后多年里,Crick 和一些其他人对遗传密码作了许多猜测,例如三联体可以是不重叠的(即 $c_1c_2c_3, c_4c_5c_6, \dots$),也可以是重叠的(即 $c_1c_2c_3, c_2c_3c_4, \dots$)。4 个碱基构成的三联体有 $4^3 = 64$ 种,远大于 20。不重叠密码还存在一个“阅读框架”问题,起点错了位,将读出完全不同的“含义”。所以有人设想一种无标点密码,其中 64 种编码不都是“有意义”的,只有句子断对了,每个三联字符串才有意义。这就给三联体一定的限制,譬如由三个相同字母组成的密码子一定是无意义的。经过仔细推算,“有意义”的三联体刚好 20 个,这想法简直太美妙了!似乎它不能是不对的。遗传密码最终是 1961—1964 年间由生物化学家“试管”蛋白质合成实验破译的。在那以后,有关遗传密码的各种理论构想都变得毫无意义了。

在叙述生物化学家破译遗传密码之前,我们先介绍一下生物体内蛋白质合成的过程。前已述及,DNA 在细胞核内的染色体中,RNA 则在细胞核内和细胞核外的细胞质内都有。RNA 有多种形式。一种是核糖体(ribosome),大约由 RNA 和蛋白质各半组成,存在于核里和核外;蛋白质的合成是在核外进行的,核糖体是“加工厂”。另一种叫做信使 RNA(messenger-RNA, mRNA)的小分子可以自由穿越细胞核壁,它们是在细胞核内经 DNA 的转录而成的一段互补的遗传密码副本(转录时 DNA 中的 T 换成 RNA 中的 U),为核糖体合成蛋白质提供“蓝图”,这蓝图用过后即被销毁,故 mRNA 存在时间短暂,较晚才被人们发现。还有一种转移 RNA(transfer-RNA, tRNA),为蛋白质合成输送原料——氨基酸。tRNA 有 20 种“牌号”,每种专门识别并俘获一种氨基酸,提供给核糖体。蛋白质合成需要能量,则由专门携带能量的分子——三磷酸腺苷(adenosine triphosphate, ATP)负责输送。

1954 年,哈佛大学的 Zamecnik P 小组建立了一个有蛋白质合成活性的体外系统,其中是超速离心提取的清液,内含氨基酸、ATP、核糖体,及细胞粗提物,而无活细胞。1955 年,Ochoa S 和 Grunberg-Mango M 发现一种在试管中合成 RNA 的酶。与此同时,Kornberg A 还发现一种在试管中合成 DNA 的酶。这些都为破译遗传密码准备了条件。

1961 年 5 月,生物学家 Matthaei J H 在一根试管里放上离心分离的细菌提取物、tRNA、20 种氨基

酸(其中 16 种作了放射性标记)、ATP、盐和保持 pH 值恒定的液体。接着他加入几微克的人工合成的多聚尿嘧啶(U)核苷酸,这是单一尿嘧啶连接起来的核苷酸链。结果由单一氨基酸——苯丙氨酸(Phe)连接起来的蛋白质被合成了。这就是说,Matthaei 鉴定了第一个遗传密码:UUU 对应 Phe。Matthaei 所在工作小组的领导人 Nirenberg M 很快将这一结果公布了,使圈内人(包括 Crick)大为震惊,无标点密码的设想显然错了。用同样的方法可以鉴定出密码 AAA、GGG、CCC 的含义。用多种比例不同的碱基合成的随机混杂核苷酸链可以鉴定一些组合密码,如 UUG、UGU 等,但无法确定其字母的顺序。1964 年,Nirenberg 和 Leder P 发明了一种运用三联核苷酸鉴定密码的技术,而另一单位的 Khorana H G 则完善了一种精确技术,可以合成顺序完全确定的 RNA 长链。到 1966 年,两家都破译了所有 64 个三联体密码子(见表 3)。密码是不重叠的,其中有句点(stop)——UAA、UAG 和 UGA,有起始符——AUG [同时又是蛋氨酸(Met)唯一的编码]或 GUG,有冗余度(许多氨基酸有重复编码)。1968 年,Khorana, Nirenberg 和另外一个人——Holley R W 分享了诺贝尔生理或医学奖。

表 3 RNA 的遗传密码

第一碱基 (5'端)	第二碱基				第三碱基 (3'端)
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Tyr	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

8 后记

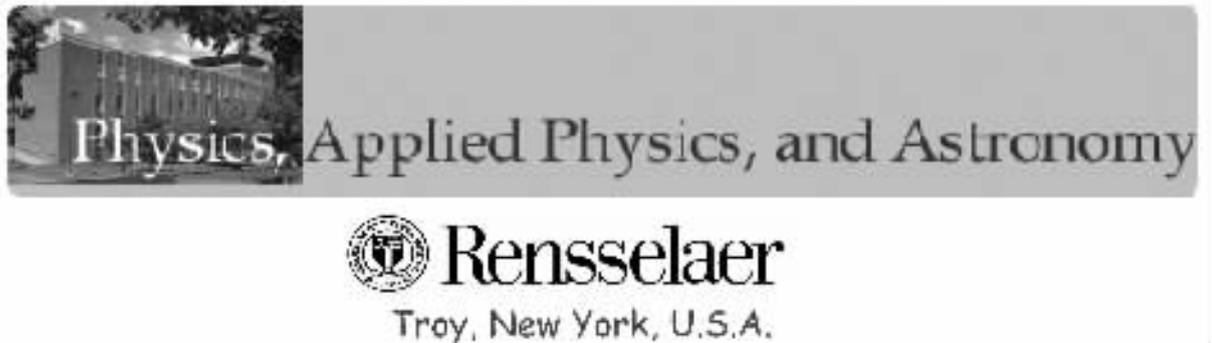
在双螺旋论文发表 50 周年之际,我们回顾了分子生物学创始和发展的主要脉络。我们看到,这一对

本世纪影响深远的学科,是物理、化学和生物学几个学科的科学家协作的结果。他们各有所长,知识结构和思想方法互补,才结出丰硕的科学果实。综合学科需要的是各方面专家的合作,而不是一批什么都懂一点的人的集合。物理学家进入生命科学领域已成为当前的潮流,看来不懂一些化学是不行的。物理学家追求理论上的普遍性和简洁美,这对研究生命现象有时是有启发意义的,但要小心,因为现存的物种不是优化设计出来的,而是进化而来的。物种进化靠适者生存的竞争机制,但在漫长的演化道路上有很多分岔点,在这些地方当初走上哪个分支,全属偶然(对称破缺),靠纯逻辑推理往往会失败。

参 考 文 献

- [1] Olby R. The Path to the Double Helix. Seattle :Univ. Washington Press ,1974
- [2] 米歇尔·莫朗热著. 昌增益译. 二十世纪生物学的分子革命. 北京:科学出版社,2002[Morange M. A History of Molecular Biology. Chinese Trans. Chang Z Y. Beijing :Science Press ,2002]
- [3] Watson J D. The Double Helix. New York :Atheneum ,1968
- [4] 弗朗西斯·克里克著. 吕向东,唐孝威译. 狂热的追求. 合肥:中国科学技术大学出版社,1994[Crick F. What Mad Pursuit—A Personal View of Scientific Discovery. Chinese Trans. Lu X D ,Tang X W. Hefei :The University of Science and Technology of China Press ,1994]
- [5] Cotterill R. The Cambridge Guide to the Material World. London :Cambridge University Press ,1985

· 信息服务 ·



美国伦斯勒理工学院招生信息

JOIN OUR GRADUATE SCHOOL IN PHYSICS

Ph. D. in Department of Physics ,Applied Physics ,and Astronomy Areas of Research :Astronomy ,Elementary Particles Physics ,Nano-Structure Physics ,Origins of Life ,THz Imaging ,THz Electronics.

Teaching ,research assistantships ,and fellowships are available.

Application <http://www.rpi.edu/dept/grad-services/>

Information <http://www.rpi.edu/dept/phys/>

E-mail :gradphysics@rpi.edu

更 正

本刊 2003 年第 10 期有几处错误更正如下:

(1)第 641 页右栏,倒数第四行起;“成为中国工程院院士的有赵伊君、任阵海…….”更正为“成为中国工程院院士的有何德全、赵伊君、任阵海…….”;

(2)第 642 页右栏,倒数第 11 行;“投湖自尽”更正为“投环自尽”;

(3)第 644 页的封面说明有两处更正,第 1 行“1958 年”更正为“1957 年”,第 2 行“1.4 万平方米”更正为“1.95 万平方米”.特此说明并向读者和作者致歉.

《物理》编辑部