

计算机在生物医学研究中应用简介

王 曼 霖

(中国科学院生物物理研究所)

电子计算机自问世以来,在短短的三十余年中,得到了飞速的发展,运算速度愈来愈快、贮存容量愈来愈大、体积愈来愈小,而每分钟的上机费用却愈来愈低,它们差不多都有三个数量级的变化。由于计算机的通用性,所以已成为几乎一切领域中的一个强有力的工具。

我们来回顾一下它的发展,可以看到它对生物学和医学方面的影响越来越大。

本文一般性地介绍计算机在生物医学方面的应用。

一、数字计算机的功能

数字计算机是一种能自动地、高速地进行大量计算工作的电子设备。它能通过对输入数据进行数值运算和逻辑运算,来求解各种问题。

1. 信息加工

最容易想到的信息加工是对数值进行计算,这也是计算机发展初期的专长,它因此而获得了“计算”机这个名称。确实,计算机能相当迅速地进行各种庞大复杂的数值运算,例如在生物大分子结构分析中根据X线衍射数据求出各个原子的相对座标;在统计学问题中配回归曲线;在放射性样品测量中用最小二乘法分解复杂 γ 能谱等等。在这些问题中,输入到计算机的信息都是“数值”,进行的加工方式都是算术运算,即计算机根据预先存入的程序将输入的数据进行计算后再输出结果。

除了数值外,只要添上相应的数据获取系统,计算机还能接受各种各样的信息,如图像、文字、各种语言程序以及各种实时送入的随时变化的测量信号(如心电、脑电或各种神经冲

动脉冲信号)等等,对应于每种不同的输入信息,计算机可作完全不同于算术运算的加工处理。例如医疗部门广泛应用计算机化的病历查询及病员护理,查询人员送到计算机的是特定的语言信息,“告诉”机器待查的内容、类别等,而机器将存放在外存贮器(磁盘或磁带)内信息库中的有关部分调入计算机,并进行“查询”处理,然后通过显示设备或打印设备给出回答。这里,计算机的加工处理不是数值的计算,而是进行“检索”,“查询”。

又如在图象辨认的工作中,一般经过三个必要的处理过程来实现“辨认”,即所谓收集数据的过程、图象分析过程和图象分类过程。收集数据过程的任务是将图象信息(不同的灰度或彩色)转换成计算机能接受的二进制编码信息,该任务可由传感器、扫描器和数字化装置来完成;此后便开始分析过程,如进行图象的特征提取和选择;根据分析得到的资料进行分类。分析和分类两个过程可能要在人的干预下反复进行几次才能完成。

总之计算机能与任何形式的信息打交道,并能根据程序的要求,对该种信息进行特定的处理加工,最后按要求输出结果。

2. 过程控制

计算机具有控制各类装置设备的能力。即根据特定的要求,能利用机器本身输出的信息去控制任何工具、仪器、机器的操作。这方面的例子很多,军事部门的指挥仪就是各种武器的火力控制计算机;各种工业控制机,除了能巡回检测、自动记录、统计制表、监视报警和自动启动和停机外,还可以直接调节和控制生产过程,使其保持于“最佳”的工作状态,以及实现经济

核算、组织管理生产等。

控制机的使用方式分计算机的开环控制和计算机的闭环控制两种。

在计算机开环控制系统中计算机不接受被控过程所输出的信息，而仅仅按照操作人员的要求确定控制方案(见图 1a)；或者，计算机不直接控制受控过程，而仅将其处理结果提供操作者参考，由人去决定控制方案并执行控制作用(见图 1 b)。

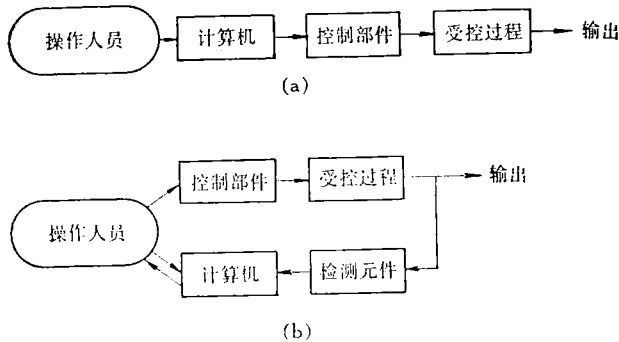


图 1 计算机开环控制的两种方式

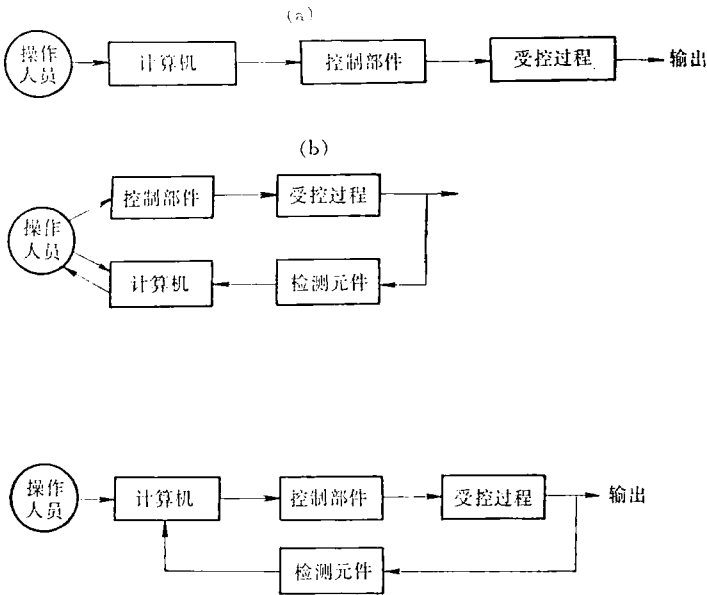


图 2 计算机的闭环控制

在计算机的闭环控制系统中，计算机能自动地接受检测到的数据资料、自动计算控制方案并直接指挥控制部件的动作以控制整个过程。而操作人员的作用，仅在于初始时送入必要的程序和在进行必要的监督、干

预(图 2)。

3. 仿真模拟

计算机对生物医学科学研究产生了巨大影响，主要由于它能对自然界的客观事物或过程进行仿真模拟。

所谓用计算机对某客观对象进行模拟，就是在计算机内部建立一个反映该对象的仿真模型，并且建立一个反映它变化规律的程序，因此计算机内部的代码的变化，反映了外界真实

的变化。这样，代表实际状况的代码和代表变化规律的程序形成了一个数学系统。在生物学研究中，一般地说来，这种数学系统就是一组方程式。模拟技术，就是对这组方程式，简单地计算一个又一个特殊情况，而不是进行全面的数学分析，这是因为实际问题中的系统往往比较复杂，不易作全面分析，另外，即使对特殊情况进行计算时，往往计算量和数据保存量都是大量的，所以这种模拟只能由计算机来实现。

建立数学模型，在模拟技术中具有很重要的地位。所谓建立数学模型就是把研究对象或过程所遵循的数学方程和特性曲线，全部在计算机上实现。建成数学模型后，各种研究和试验就能在该模型上进行。在科学研究的实验活动中，利用上述手段益处甚多。首先，可用该技术代替一些实验，减少实验费用和缩短试验时间；有时即使不能代替实验，至少可用来设计实验，并从中预计观察的内容、观察时间的长短和所期待的实验结果等等；数学模型还可用来模拟一些实验室中

办不到的实验，例如在细胞膜渗透性的研究中，对水的渗透细孔存在与否(直径为 7 埃左右)的假定进行检验^[7]；数学模型还常常是一种提供待研究系统特征参量的方法；特别值得一提的优点是，此方法能定量地发现一些新的“实验”

现象。墨鱼突轴中非传导性膜的动作电位,不是以“全或无”方式作用,就是一个用计算机模拟方法发现的极好例证^[8]。特别是,计算机模拟所揭示的这个现象,仅仅在一定的温度范围内出现,这样就满意地解释了分级响应现象的实验。

总之,计算机的模拟技术使实验工作发展到了一个新的阶段。

二、计算机技术的发展对 生物医学领域的影响

在最初的十年中(50年代到60年代),计算机主要用在军事部门、物理学研究、化学研究和工程技术中。因为在这些部门和领域中,有大量的计算工作,而这恰好是计算机发展的最初阶段的特长。到了60年代由于计算机对科学和社会的影响已为人们所公认,它在生物医学上的应用价值也引起了广泛的重视。计算技术在医学生物学中的应用,有了很大的发展。例如美国洛克菲勒大学的生物实验室^[1]的计算机数目,从1963年的第一台发展到1976年的三十七台,按指数往上增加。其应用范围,从最早对整个器官行为的研究,各种感受器及中枢神经系统的研究方面,发展到目前能对各种生物化学实验进行监测、控制和数据处理;能模拟多种电生理和心理学实验。虽然这仅仅是一个生物实验室的例子,但表明了计算机在生物医学方面的应用其发展之快和范围之广。

计算技术在近十多年来的迅速发展,是促进生物医学领域广泛应用的重要原因。譬如在生物化学领域中计算机的应用,在六十年代发展得很慢。这是由于当时在生化技术中作用的仪器,如气相、液相层析分析仪、质谱仪、电子顺磁共振及核磁共振谱仪等仪器,所要配的计算机既贵而又复杂,计算机和上述各实验仪器间的连接装置——接口*,也较复杂。后来由于集成电路的发展出现了小型计算机,大大降低了使用计算机的费用,又因为小型机采用了标准化的控制信号和“总线”结构等新技术,大大

简化了接口的设计和制作。这样一来,在生化实验中可广泛应用计算机。目前在分析化学方面,运用了大型、小型、微型各种形式的计算机系统^[2,3]。用微型计算机来控制仪器的操作,例如逐个处理样品、调整电场和磁场的参数或调节样品瓶的机械运动等等;用小型计算机来收集实验数据、编排一系列试验的顺序或管理大量的数据资料;用大型的通用计算机进行大量复杂的数据处理,如在电子顺磁共振谱分析技术和核磁共振谱分析技术中,用以作快速付里叶变换的计算,或者在某些实验中,用大计算机作最小二乘法配曲线等运算。

1. 显示设备的发展

计算机的显示设备的飞速发展,是计算机化的断层照相技术以及白血球分类技术获得成功的条件之一。这是计算机技术的发展,促进了它在生物医学领域中的应用的又一例证。白血球分类的计算机技术和计算机化的断层照相技术,都早在五十年代就开始研究了^[4,5],但是迟至七十年代才获得成功,1974年有白血球分类的商品仪器,1972年英国的EMI公司出售第一台商品的断层扫描仪。这和计算机图形显示设备的发展有密切的关系。在1965年以前,计算机的图形显示设备发展很慢,它的硬设备和软件均价格昂贵,产品的数量和品种也少。从七十年代开始,计算机图形显示设备方面取得了迅速的发展。据1971年统计,世界上各公司产品品种已多达130余种。特别是当时出现了低价格的图形显示设备及智能终端显示设备,一些上万美元的设备价格下降到4,000美元以下,使得原来主要用于军事部门的信息显示技术能够广泛地应用在国民经济的各个领域内^[6]。另外智能显示终端的出现,使得显示设备不仅作为计算机的简单的输入输出手段,而且它可以根据主计算机的命令或操作人员的命令,完成字母、数字或图形信息的输入、显示、缓

* 接口(interface)在计算机中央处理和外部设备之间、计算机和计算机之间以及计算机和各种其他仪器、装置、设备间的连接装置。它包括许多信息传输线及其逻辑控制电路。

存及检验等任务。它通常带有一台小型的或微型的计算机。这种图形显示设备,是实时的“人-机”通讯的有力工具,也是利用计算机分析生物图象时必不可少的工具。例如在用于临床白血球分类的商品仪器中,有的仍必须通过“人-机”对话系统,来辨认一些机器所“不认识”的细胞^[4]。

2. 集成电路的发展

近年来随着大规模集成电路的发展出现了具有许多独特优点的微型处理机,它推动了计算机在生物医学领域中更为有效和广泛的应用。

由于生物医学研究和临床应用的多样性和复杂性,不同的实验(或临床应用)所用的仪器和要求往往是完全不相同的,过去传统的做法是根据实验的专门需要,设计和制作一些专用的逻辑电路或模拟电路,甚至制作专用的计算机及接口设备,这不仅在设计 and 制作过程中要耗费大量的时间和人力,使实验周期历时很长,而且当一种实验结束后,开始另一种实验时,上述专用接口和专用计算机有时就无能为力了。

自从大规模集成电路发展后,情况就大不相同了。下述各功能均可分别由一块大规模集成电路组件来实现:(1)中央处理机(因做在一块组件上,又称微处理机);(2)系统时钟;(3)随机读写存贮器;(4)可编程序的只读存贮器;(5)实验室控制或收集数据用的并行输入输出接口;(6)显示终端用的串行输入输出接口装置;(7)模-数转换装置和数-模转换装置。利用这些组件,在一块印刷电路板上就可以构成专门的实验设备,也可以构成计算机的各种接口,以及构成一台微型计算机。此时由于用的“元件”少,连线少,可靠性大大提高。又由于微型计算机可以组成满足各种需要的实验装置和接口,仪器设备的灵活性也大大提高,成本比小型计算机低,使用起来又省电又灵巧(微型计算机像袖珍字典一样大小)。特别是在同一个实验中,可以用多台微处理机构成一套装置,使得多种软件功能可以同时并行执行,而不像小型计算机控制器那样,系统的全部软件必须按时间顺

序先后执行。以微处理机为基础的这种软件的积木化功能,是一般小型计算机或其他通用机所不易做到的。例如在一般电生理实验中,可采用以微处理机为基础构成的刺激器,产生各种感官(如听觉、视觉等等)的电刺激信号,同时又可用另外一些微处理机为基础的装置,控制微电极在相应器官中(如脑、心等等)的位置和提取有关的响应信号,并分析这些信号以产生新的指挥刺激器的命令。这种产生各类刺激的程序、控制微电极插入位置的程序及提取响应并进行分析的程序构成了一个程序系统,其中的各个部分可以同时并行执行,从而大大简化了一般的串行结构的程序系统,并提高了程序系统(即软件)的灵活性。

计算机发展至今已有三十多年的历史了,它已成为人智力的一个很大的扩展,但是这个工具的发展将是无止境的。因为科学和技术的日新月异发展,不断对计算机提出了新的任务,特别在现代生物学的研究工作中,需要更大规模、更高速度的计算技术。因为在生物系统中分子水平、亚分子水平的反应可快达 10^{-10} 秒量级,而生物图象分析工作中的信息量又是非常大的,如X线断层造影的三维图象重建的研究工作,需要信息容量达十万亿比特以上和运算速度高达每秒百亿次的高速计算机。目前正处于发展阶段的光计算机和激光技术,提供了这种可能性。利用微微秒或亚微微秒脉冲激光器,有可能研究寿命为 10^{-9} — 10^{-12} 秒的快速反应过程;用光计算机处理图象信息,可能比现有数据处理技术快几百万倍,若把这些新技术结合起来运用到生物医学领域中,定将出现新的飞跃。

三、应用计算机的几个例子

1. 计算机化的断层照相^[9]

计算机化的断层照相(简称“CT”)技术是一种把小型计算机和X射线扫描设备结合起来的新技术。在X射线领域中,这一技术的成功,被认为是自从1895年发现X射线以来的最重大的发展。它把从X射线扫描得到的数据,实

时地传送到计算机处理并在计算机的图形显示屏上，显示出人体各个不同部位的横断面的象。

CT 技术的基本原理，是利用一组投影图叠加后得到物体之原像。例如用一个含有圆洞的正方形代表待测物的横断面，利用经过准直后的 X 射线，从三个不同的角度进行平移扫描照射。在待测物的另一侧有与 X 光源作同步扫描的探测器，它所测到的透射 X 线的分布图，即所谓“投影图”。如图 3 所示，在待测物体另一侧的三个相应的角度处，可得到三个一维分布的“投影图”，用它们在对应的三个方向上进行叠加，便可得到图 3b，显然如投影的方向取得愈多，则重建象的图上的交叉点也愈接近原来的圆洞，并且它相对周围的“辐射”状条的光密度值也愈大。

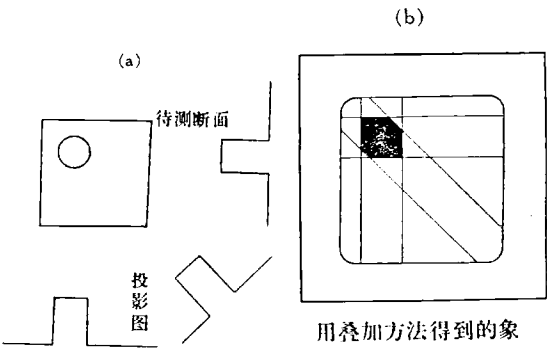


图 3 断层照相的投影叠加原理示意图

可由下式表示这种最简单的算法：

$$R(\gamma, \phi) = \sum_i \sum_j P_{\theta_i}(x_j)$$

式中 $R(\gamma, \phi)$ 代表在横断面上坐标为 (γ, ϕ) 点的重建后的象； $P_{\theta_i}(x_j)$ 是在 θ_i 方向上的投影， x_j 是 θ_i 方向上平移扫描的坐标。事实上在扫描仪中采用的是较精确的算法，如付里叶变换或卷积的算法。在 CT 扫描仪中，与传统的 X 射线透视不同，人是横卧在床上，X 射线在待测部位处绕病人转（见图 4），在 X 光管对面有一个闪烁计数器，它与 X 光管的扫描同步，通常只需二分到五分钟，就可以在计算机显示屏上观察到有关脏器的断面图了。

图 5 介绍了美国 Georgetown 大学医学中心

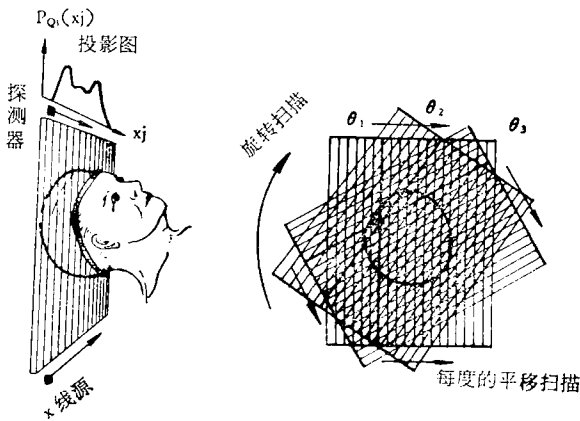


图 4 CT 扫描仪的运动方式示意图

附属的国立生物医学研究基金会研制的 ACTA 扫描仪的方块图。在这台仪器中可进行平移与旋转相结合的扫描，即机械扫描臂每转 1 度后，可以在一直线范围上作 320 步的平移扫描，与此同时闪烁计数器每隔 4 毫秒测量一次透过病人身体的 X 射线强度，一个完整的扫描需旋转 180 度，即共测 57,600 点，历时约 4.25 分钟；若机械臂每次转 2 度，则测量点数减少一半，机械运动和测量共历时 2.25 分钟。机械运动一结束，彩色显示屏与黑白显示屏上马上产生一幅清晰的脏器横断面图象，分辨率可达 1.5 毫米，图象的层次很清楚，因为无论是彩色或黑白显示，均可划分为十六种不同的灰度水平（或颜色），为了便于参考，还在显示屏上端标出十六种灰度（或颜色）的色标及对应的数据。根据需要，利用“人-机”对话，显示屏的画面可以进行调整或舍取，以突出有兴趣的部分。据报道一种新型的扫描仪，作一次完整的扫描只需 2 秒钟^[10]。

CT 技术的最主要的优点是，能产生人体内各种软组织的图象，分辨能力较高，因而可用来扫描脑、心、肝、肺和胃；可检查各种肿瘤、血管的畸形，使诊断工作大为简化，减少危险性；也可用来研究心脏的形状、大小及其动态的变化。此外作一断层片子，病人所接受的剂量也较低。因而这门技术发展之快和应用之广都是罕见的。自 1972 年第一台商用 CT 扫描仪问世以来，现在每年美国有二千万张计算机化的断层

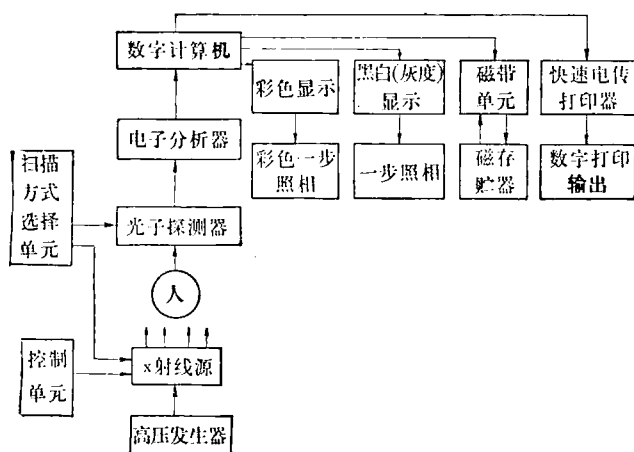


图5 ACTA 扫描仪方块图

照片,据估计,今后的五年到十年中,CT 仪的数目至少将增加二个数量级^[10]。

1975 年和 1976 年已召开了二次国际“CT”讨论会,并于 1976 年秋又召开了头部“CT”的国际讨论会^[11],后者共报道了六十三篇论文,可见在这短短不到十年中,从设想到实现,从实现到发展,进度之迅速十分惊人,这的确是有计算机后对医学方面的一个重大突破。

2. 白血球自动分类^[10]

七十年代白血球自动分类技术和计算机化的断层技术的成功,被认为是计算机应用在生物医学图象辨认上的两大突破。

白血球分类(下简称“WBCD”)技术,即用计算机和电视摄像机代替人的脑子和眼睛,对显微镜下的人体的白血球细胞进行分析并分类。

自动化的“WBCD”技术,不像“CT”技术那样采用解析的算法,而是根据所提取的图像特征来辨认细胞的。

计算机是怎样一步步地提取形状信息的呢?譬如说有一幅被红血球围着的白血球的电视图象,把光密度值的域值(即灰度的“门坎”水平)先后调三次,以得到三张黑白图(1比特的图): 1. 细胞核的稠密部分的结构,即染色质图;2. 细胞核部分;3. 细胞核加上周围原生质中稠密的颗粒状的部分和红血球中的稠密部分。然后由计算机把这三张图组成一张具有 4 个灰

度水平的图象,称这张图为“卡通”(Cartoon)。有时一张待辨认的图象,可以分解成若干张子图,也称这些子图为“卡通”。计算机从这些“卡通”图中又可以求得若干张形态学上的特征图,从而实现辨认。上述分析属图象的“语言结构”分析方法,除此之外还有一种利用特征抽取来分析白血球的方法^[12]。也还有人在使用特殊的染色剂后,用比色的方法来提供白血球的分类信息^[13]。后两种方法也都采用计算机,实现了自动分析和分类。

目前的商业的“WBCD”仪,每分钟可处理五十到一百张图象。在美国每年需要进行观察和分类的白血球涂片有 3×10^{10} 张,在 1976 年约有 2×10^8 张是由计算机进行自动地辨认分类,占总数的 0.6%。

3. 膜渗透性的计算机模拟^[7]

膜的生理学研究之一是关于其输运性质与膜结构的关系。细胞膜中存在有水的通道或细孔的假定,成功地解释了许多观察到的现象。然而目前企图用直接的实验来研究它是不可能的,因为用来测量细孔半径的所有方法,都在某种程度上依赖于连续流体动力学的结论,而后者能否应用到半径小到 4 埃左右的小孔上,还是有问题的。目前还没有一个方法可测定这种细孔的半径。

Levitt 运用了分子动力学的方法,以计算机模拟的方式着手研究了上述问题。他用计算机模拟了一个实验,在该实验中,膜上面有一个半径为 3.2 埃的细孔,膜的两侧是“刚性球”模型的液体,然后“测量”了“刚性球液体”的扩散系数和水压渗透性,从这些数据再定出膜细孔半径的“实验值”,并且将“实验值”和模拟中细孔的“真实”半径比较。

Levitt 的计算机实验的具体内容如下,假定细胞膜的细孔半径为 3.2 埃、长 11 埃,水分子模型为 1 埃半径的光滑刚性球,对每一个分子分配给一个随机的初始位置和速度(平均温度为 313°K),然后计算每一个分子的实际运动

过程,也就是在分子之间的相互作用及分子与膜孔壁的相互作用下,每个分子的真实的路径。选定水分子的总数目为 20 个和 11 个两种。计算机每算一次即所谓做一次“实验”,因此每算一次的时间愈长以及算的次数愈多,“实验”结果也愈精确。人为地规定在低密度时,每次“实验”经历 1,200 次碰撞,在高密度时为 2,400 次碰撞。实验总次数的选择,由平均标准偏差为 5% 来定出。在 CDC 6,600 数字计算机上,每一个“实验”所需的时间,对低密度和高密度条件,分别为 80 秒和 280 秒。

在上述模拟实验中可直接测量出下述数值:净通量和扩散通量、压力、温度、在细孔内的扩散系数、扩散的有效面积、细孔内刚性球位置的几率分布以及扩散和对流间的相互影响等等。液体的密度、内摩擦系数及扩散系数则由刚性球的致密气体理论可确定。由这些数据又可算出相当的细孔半径。

上述分子动力学实验提供了到目前为止还是唯一的,液体在膜细孔内行为的“最直接的”知识;得到的细孔半径为 3.7 埃(低密度)和 4.0 埃(高密度时),比模型中的“实际的”孔半径(3.2 埃)稍大一点;特别值得指出的是,连续的流体动力学理论能够定性地外推到细孔半径为 3.2 埃的情形。

四、结 束 语

计算技术在生物学医学方面的应用,尚需开展大量的工作,尤其在下述两方面应予以注意。

1. 开展面向生物医学领域的应用软件研制。生物学科人员没能广泛地采用先进的计算技术,往往与缺乏合适的灵活而又简便的软件有关。在国外有一种可应用于生物实验的“模拟语言”对一个毫无使用计算机经验的人只需要经过一个小时的练习便可熟练的掌握它。

2. 接口的研究,要把计算机变成实验仪器的一个组成部分,必须解决它和相应的实验仪器间的数据或控制信息的交换问题。也就是通常所谓“接口”问题。欧洲核电子学标准委员会

于 1968 年发表了标准接口系统——CAMAC 标准(或称 CAMAC 系统),1970 年美国原子能委员会正式承认该标准为美国仪器标准,到 1975 年美国电气及电子工程师学会和国际电工委员会也相继宣布接受 CAMAC 系统,现在世界上已有 20 多个国家采用它,CAMAC 系统已完全进入了应用阶段。

CAMAC 是一种具有数据通道特性的仪器系统,是用于交换计算机和联机(on-line)设备之间的数据和控制信息的硬件和软件系统的总称。它规定了信息交换逻辑、信号电平、连接器和电源,它的数据传输速度为 1 微秒,测量点和控制点很多,单机箱可测 368 个点,单分支系统(每个分支 7 个机箱)可测 2,576 点,如果组成多分支系统,就可以测量或控制更多的参数了。整个系统具有高度的灵活性和通用性,提高了计算机的使用能力。如 DJS-130 机最多可带 62 种外围设备,而单分支的 CAMAC 接口则可扩充到 161 个外围设备,并且对于计算机而言只占用一个外部设备的位置。与该标准接口相配合的还有 CAMAC 软件系统。正因系统的高度灵活性和通用性,许多国家都积极采用和推广它。从 1973 年和 1975 年召开的两次国际性 CAMAC 会议来看,它已完全进入了应用阶段。美国和加拿大,都已应用于医学界,用来收集神经生理实验、药理实验和生化实验等信息。

随着我国电子工业和计算机科学的快速发展,计算机在生物医学研究中一定会得到更广泛的应用,为我国早日实现四个现代化做出贡献。

本文承郑竺英同志热情指导,在此致谢。

参 考 文 献

- [1] Schoenfeld, Robert R. L. et al.: *Computer* 10, (5), 56, 1977.
- [2] Bowen, B. E. et al.: *Anal. Chem.*, 45, (13), 2185, 1973.
- [3] Brouer, G. and Jansen J. A. J.: *Anal. Chem.*, 45, (13) 2239, 1973.
- [4] Tolles, W. E.: *Trans. N. Y. Acad. Sci. Ser II*, 17 (3), 250, 1955.

- [5] Howry, D. H.: *Ultrasonnd in Biology and Medicine* (E. Kelly, ed.), Waverly Press, Baltimore, p. 49, 1957.
- [6] 长沙工学院情报资料室:《计算机图形显示综述》, 1975,10。
- [7] Levitt, D. G.: *Biophys. J.* **13**, 186, 1973.
- [8] Macey, R. I. and Wadzinski, L. T. *Computers in Life Science Research*. p. 29, (Edited by. William Siler), Plenum Press, 1975.
- [9] P. 亨德森主编, 上海生物化学所等译《生物学与人类未来》, 第 321 页, 科学出版社, 1977 年。
- [10] Ledley, R. S. *Comput. Biol. Med.*, **6**, 239, 1976.
- [11] Preston, K. Jr: *Computer*, **9**, 54, 1976.
- [12] Janksch, W. and Kazner, E (ed): *Cranial. Computerized Tomography*, Springer-Verlag, 1976.
- [13] 谷口庆治、川端信賢、尾崎弘: 电子通讯学会论文志, Vol. J. 59—C. No. 1, p15 1976.
- [14] Prewitt, J. M. S: *Adv. Computers*, **12**, 285, 1972.

【本文于 1978 年 6 月 10 日收到】

信使核糖核酸的核苷酸序列分析

魏 西 平

(中国科学院生物物理研究所)

核酸是重要的遗传物质。信使核糖核酸 (mRNA)* 从 DNA 转录生物遗传信息, 并将此信息翻译成蛋白质。测定 mRNA 的核苷酸排列顺序是研究结构功能关系的基础。近年, 越来越多的学者致力于测定 mRNA 的一级结构, 目前已有了一定进展。现仅就 mRNA 的一般特性、纯化方法和核苷酸序列分析技术作一简单介绍。

一、mRNA 的一般特性

细胞中信使核糖核酸占核糖核酸总量的百分之三到五, 分子量一般为 $0.5-2 \times 10^6$ 道尔顿。其寿命较短, 可能是多核苷酸磷酸化酶“终止”它的寿命。在细胞核中, 依赖于 DNA 的 RNA 聚合酶以 DNA 为模板转录出 mRNA 的前体, 又在酶促作用下, 于 mRNA 前体的 3' 端加上多聚腺苷酸 (poly A), 经过选择性降解, 最后形成成熟 mRNA, 进入细胞质。

mRNA 的主要生物功能是与核糖体结合合成蛋白质。任何蛋白质的氨基酸排列顺序都是

由其基因——DNA 碱基排列顺序决定的, 这信息是通过 mRNA 从 DNA 传递到蛋白质,

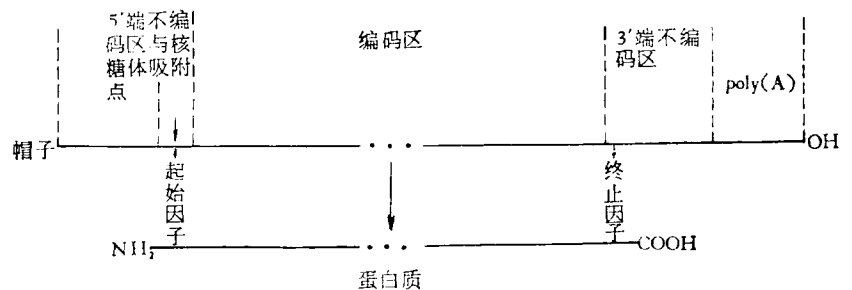


图 1 mRNA 分子的结构示意图

mRNA 分子上密码区的碱基排列决定了蛋白质的氨基酸排列。

可将 mRNA 分子分为三个部分, 中间部分是蛋白质编码区, 链的两端分别为 3' 端不编

* 缩写注释

mRNA	信使核糖核酸
DNA	脱氧核糖核酸
dNTP	某一种脱氧核糖核苷三磷酸
rNTP	某一种核糖核苷三磷酸
A	腺苷
G	鸟苷
C	胞苷
T	胸腺嘧啶核苷
poly(A)	多聚腺苷酸
oligo(dT)	寡聚胸腺嘧啶核苷酸
oligo(U)	寡聚尿苷酸