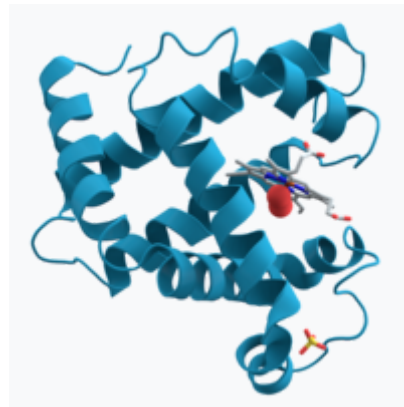


蛋白质

维基百科，自由的百科全书

蛋白质（英語：protein，旧称“**朊**”^{[1][2]}）是大型生物分子，或高分子，它由一个或多个由α-氨基酸残基组成的长链条组成。α-氨基酸分子呈线性排列，相邻α-氨基酸残基的羧基和氨基通过肽键连接在一起，最後經過折疊形成有功能的立體結構。蛋白质的α-氨基酸序列是由对应基因所编码。除了遗传密码所编码的20种“标准”氨基酸，在蛋白质中，某些α-氨基酸残基还可以被改變原子的排序而发生化学结构的变化，从而对蛋白质进行激活或调控。多个蛋白质和礦物質可以一起，往往是通过结合在一起形成稳定的蛋白质复合物，這樣的大分子結構就像機械一樣，來发挥某一特定功能。

与其他生物大分子（如多糖和核酸）一样，蛋白质是地球上生物体中的必要组成成分，参与了细胞生命活动的每一个进程。蛋白酶是最常见的一类蛋白质，它们催化生物化学反应，尤其对于生物体的代谢至关重要。除了酶之外，还有许多结构性或机械性蛋白质，如肌肉中的肌动蛋白和肌球蛋白，以及细胞骨架中的微管蛋白（参与形成细胞内的支撑网络以维持细胞外形）。另外一些蛋白质则参与细胞信号传导、免疫反应、细胞黏附和细胞周期调控等。同时，蛋白质也是动物饮食中必需的营养物质，这是因为动物自身无法合成所有蛋白氨基酸，动物需要和必须从食物中获取必需氨基酸。通过消化过程将蛋白质降解为自由氨基酸，动物就可以将它们用于自身的代谢。



肌球蛋白三维结构的飘带图，其主要由α螺旋所构成。这种蛋白质是第一个通过X射线晶体学解析其结构。

目录

生物化学性质

合成

[蛋白质生物合成](#)

[化学合成](#)

降解

结构

细胞功能

[催化作用](#)

[訊號傳遞和物質运输](#)

[變性作用](#)

[變性蛋白質的特性](#)

[结构蛋白](#)

研究方法

[蛋白质纯化](#)

[细胞内定位](#)

[蛋白质组学](#)

[生物信息学](#)

[结构预测与模拟](#)

蛋白質含量檢測

营养作用

历史和词源

参见

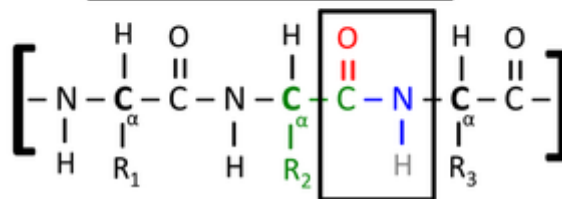
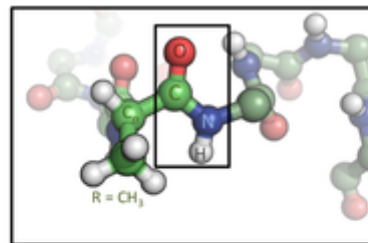
参考文献

外部链接

生物化学性质

蛋白质是由不同的L型α氨基酸所形成的线性聚合物。目前在绝大多数已鉴定的天然蛋白质中发现的氨基酸有20种（参见標準蛋白氨基酸列表）。不过在自然界中还存在着一些特殊的氨基酸，例如在一种海洋寡毛纲小蠕虫*Olavius algarvensis*以及与之存在共生关系的细菌δ1（该细菌属于δ變形菌）中存在着高含量的硒代半胱氨酸，由原本为终止密码子的UGA编码，和吡咯赖胺酸，由终止密码子UAG编码^[3]。

所有氨基酸都有共同的结构特征，包括与氨基连接的 α 碳原子，一个羧基和连接在 α 碳原子上的不同的侧链。但脯氨酸有着与这种基本结构不同之处：它含有一个侧链与氨基连接在一起所形成的特殊的环状结构，使得其氨基在肽键中的构象相对固定。^[4] 标准氨基酸的侧链是构成蛋白质结构的重要元素，它们具有不同的化学性质，因此对于蛋白质的功能至关重要。多肽链中的氨基酸之间是通过脱水反应所形成的肽键来互相连接；一旦形成肽键成为蛋白质的一部分，氨基酸就被称为“残基”，而连接在链的碳、氮、氧原子被称为“主链”或“蛋白质骨架”。由于肽键有两种共振态，具有一定的双键特性，使得相邻 α 碳之间形成肽平面；而肽键两侧的二面角确定了蛋白质骨架的局部形态。



蛋白质结构中氨基酸和丙氨酸通过肽键（用方塊標出）連接在一起。

由于氨基酸的非对称性（两端分别具有氨基和羧基），蛋白质链具有方向性。蛋白质链的起始端有自由的氨基，被称为N端或氨基端；尾端则有自由的羧基，被称为C端或羧基端。



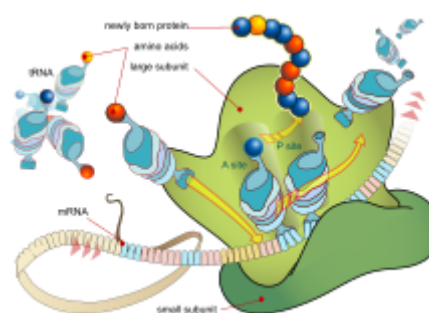
肽键的共振结构。氨基酸通过肽键连接在一起形成蛋白质聚合物。

“蛋白质”、“多肽”和“肽”这些名词的含义在一定程度上有重叠，经常容易混淆。“蛋白质”通常指具有完整生物学功能并有稳定结构的分子；而“肽”则通常指一段较短的氨基酸寡聚体，常常没有稳定的三维结构。然而，“蛋白质”和“肽”之间的界限很模糊，通常以20-30个残基为界。^[5]“多肽”可以指任何长度的氨基酸线性单链分子，但常常表示缺少稳定的三级结构。

合成

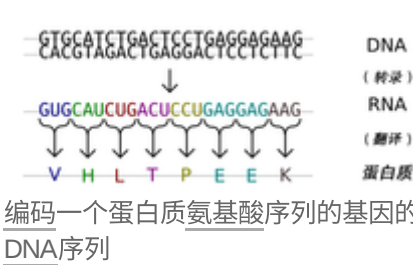
蛋白质生物合成

每一种蛋白质都有自己独特的氨基酸序列，而氨基酸序列的组成信息则由编码对应蛋白质的基因的核苷酸序列所决定。遗传密码是一套由三个核苷酸组成的密码子，每一种三个核苷酸的组合可以编码一种特定氨基酸，如mRNA上的AUG（在DNA中为ATG）编码甲硫氨酸。由于DNA含有四种核苷酸（A、T、C、G），所以对应的可能的密码子有 $4 \times 4 \times 4 = 64$ 种；而标准氨基酸只有20种，因此有部分密码子是冗余的，即部分氨基酸可以由多个不同的密码子所编码。DNA中的基因首先在RNA聚合酶等蛋白质的作用下被转录为前mRNA。在大多数生物体中，前mRNA（或初始转录产物）要经过转录后修饰以形成成熟的mRNA，随后mRNA就可以经由核糖体被用作蛋白质合成的模板。在原核生物中，mRNA可能可以在生成后被直接用于蛋白质合成，或者在离开细胞核后就结合核糖体。而在真核生物中，mRNA在细胞核中被合成，然后通过核膜被转运到细胞质中；在细胞质中，mRNA才可以被用于蛋白质合成。原核生物的蛋白质合成速率可以达到每秒20个氨基酸，要高于真核生物。^[6]



核糖體產生一種蛋白質，使用信使RNA模板。

从一个mRNA模板合成一个蛋白质的过程被称为轉譯。在轉譯过程中，mRNA被一些蛋白质携带到核糖体上；然后核糖体在mRNA上从5'端到3'端寻找起始密码子（大多数情况下为AUG）；找到起始密码子后，即核糖体上起始tRNA的反密码子与起始密码子配对后，翻译就可以开始进行；在起始密码子后，核糖体每一次阅读三个核苷酸（或一个密码子），同样是通过携带对应氨基酸的tRNA上反密码子与密码子配对。其中，氨酰tRNA合成酶可以将tRNA分子与正确的氨基酸连接到一起。不断延长的多肽链通常被称为“新生链”。生物体中的蛋白质合成总是从N-端到C-端。



合成的蛋白质的大小可以通过其含有的氨基酸数目或者其分子量（以道尔顿或千道尔顿，即kDa为单位）来衡量。酵母蛋白的平均长度为466个氨基酸或平均分子量为53kDa。^[5]目前已知的最大蛋白质是肌联蛋白，它是肌肉中肌节的组分之一，其分子量为近3,000 kDa，含有近27,000个氨基酸。^[7]

化学合成

除了生物合成外，一些小的蛋白质可以通过多种化学途径来合成。这些合成方法又被称为肽合成，其依赖于有机合成技术，如化学连接来高通量生产肽。^[8]化学合成允许在合成的肽链中引入非天然氨基酸，如加入荧光标记的氨基酸。^[9]这些合成方法所合成的产物被大量应用于生物化学和细胞生物学实验。但是，化学合成无法有效合成残基数多于300的蛋白质，而且合成的蛋白质可能不具有天然的三级结构。大多数化学合成方法都是从C-端到N-端进行合成，刚好和生物合成反应的方向相反。

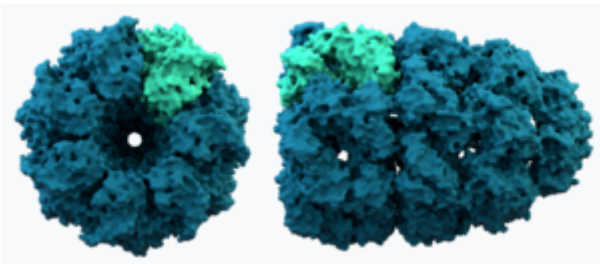
降解

对于细胞来说，蛋白质降解有多种用途，包括去除分泌蛋白的N末端信号肽，对前体蛋白进行剪切以产生“成熟”蛋白等。细胞不需要的或受到损伤的非跨膜蛋白质一般由蛋白酶体来进行降解，而真核生物的跨膜蛋白则通过内体运送到溶酶体（动物细胞）或液泡（酵母）中进行降解。^[10]降解所生成的氨基酸分子可以被用于合成新的蛋白质。一些蛋白质可以发生自降解。此外，细胞中存在的大量蛋白酶（特别是溶酶体中），可以对外来的蛋白质进行降解，这也是一种细胞自我保护的机制。

生物学实验中，也经常对蛋白质进行降解分析；例如在蛋白质组学中，利用蛋白酶对特定蛋白质进行降解，并对降解产物进行质谱分析而获得对应蛋白质的序列信息和修饰情况；此外，生物化学实验中，埃德曼降解法常被用于对蛋白质进行氨基酸序列分析。

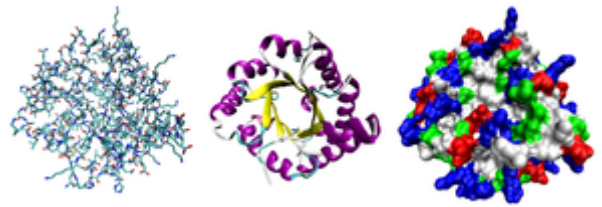
结构

大多数的蛋白质都自然折叠为一个特定的三维结构，这一特定结构被称为天然状态。虽然多数蛋白可以通过本身氨基酸序列的性质进行自我折叠，但还是有许多蛋白质需要分子伴侣的帮助来进行正确的折叠。在高温或极端pH等条件下，蛋白质会失去其天然结构和活性，这一现象就称为变性。生物化学家常常用以下四个方面来表示蛋白质的结构：



分子伴侣的晶体结构。分子伴侣帮助蛋白质折叠。

- 蛋白质一级结构：组成蛋白质多肽链的线性氨基酸序列。一个蛋白质是一个聚酰胺。
- 蛋白质二级结构：依靠不同氨基酸之间的C=O和N-H基团间的氢键形成的稳定结构，主要为α螺旋和β折叠。^[11]因为二级结构是局部的，不同的二级结构的许多区域可存在于相同的蛋白质分子。
- 蛋白质三级结构：通过多个二级结构元素在三维空间的排列所形成的一个蛋白质分子的三维结构，是单个蛋白质分子的整体形状。蛋白质的三级结构大都有一个疏水核心来稳定结构，同时具有稳定作用的还有盐桥 (蛋白质)、氢键和二硫键，甚至翻译后修饰。“三级结构”常常可以用“折叠”一词来表示。三级结构控制蛋白质的基本功能。
- 蛋白质四级结构：由几个蛋白质分子（多肽链），通常称为蛋白质亚基所形成的结构，在功能上作为一个蛋白质复合体。



三种显示蛋白质三维结构的方式。图中蛋白质为磷酸丙糖异构酶。左：显示全部原子，并以原子类型标色（碳原子为蓝绿色，氧原子为红色，氮原子为蓝色）；中：只显示主链构象，以二级结构类型标色（α螺旋为紫色，β折叠为黄色）；右：显示“溶剂可及表面”，以残基类型标色（酸性氨基酸为红色，碱性氨基酸为蓝色，极性氨基酸为绿色，非极性氨基酸为白色）。



细胞色素c的NMR溶液结构，显示了蛋白质的动态结构。

蛋白质并不完全是刚性分子。许多蛋白质在执行生物学功能时可以在多个相关结构中相互转换。在进行功能型结构重排时，这些相关的三级或四级结构通常被定义为不同“构象”，而这些结构之间的转换就被称为“构象变换”。例如，酶的构象变换常常是由底物结合到活性位点所导致。在溶液中，所有的蛋白质都会发生结构上的动态变化，主要表现为热振动和与其他分子之间碰撞所导致的运动。

蛋白质可以由三级结构的不同大致分为三个主要类别：球蛋白、纤维蛋白和膜蛋白。几乎所有

的球蛋白都是水溶性的，许多球蛋白是酶。纤维蛋白多为结构性的，例如，结缔组织的主要成分胶原蛋白，或头发和指甲的蛋白质成分角蛋白。膜蛋白常常作为受体，或提供通道极性的或带电的分子通过细胞膜。



不同大小的蛋白质的分子表面。从左到右依次为：抗体 (IgG)、血红蛋白、胰岛素、腺苷酸激酶和谷氨酰氨合成酶。

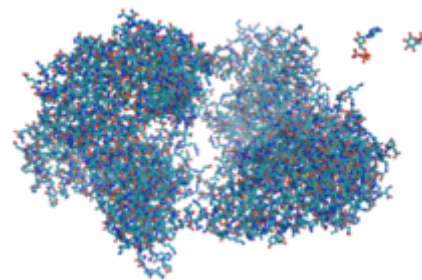
要了解特定蛋白质的功能，获得其三级结构或四级结构可以提供重要的结构信息。目前用于蛋白质的原子分辨率结构测定的方法主要是X射线晶体学和NMR光谱学。冷冻电子显微学也可以提供超大蛋白质复合物（如病毒、核糖体等）的低分辨率结构信息。^[11]而电子晶体学在一些情况下也可以提供较高分辨率的结构信息，特别是对于膜蛋白的二维晶体。^[12]解析的结构（包括原子坐标和结构解析的相关信息）通常存放到蛋白质数据库（PDB），供全世界研究者免费下载。蛋白质结构预测也可以为未知结构（实验结构）的蛋白质提供结构信息。

细胞功能

蛋白质是细胞中的主要功能分子。^[5]除了特定类别的RNA，大多数的其他生物分子都需要蛋白质来调控。蛋白质也是细胞中含量最为丰富的分子之一；例如，蛋白质占大肠杆菌细胞干重的一半，而其他大分子如DNA和RNA则只分别占3%和20%。^[13]在一个特定细胞或细胞类型中表达的

所有蛋白被称为对应细胞的蛋白质组。

蛋白质能够在细胞中发挥多种多样的功能，涵盖了细胞生命活动的各个方面：发挥催化作用的酶；参与生物体内的新陈代谢的调节作用，如胰岛素；一些蛋白质具有运输代谢物质的作用，如离子泵和血红蛋白；可以消耗能量來轉動並運動的分子馬達；控制神經和反應的神經傳遞物；发挥储存作用，如植物种子中的大量蛋白质，就是用来萌发时的储备；许多结构蛋白被用于细胞骨架等的形成，如肌球蛋白；还有免疫、细胞分化、细胞凋亡等过程中都有大量蛋白质参与。



酶己糖激酶被显示为常规的球-棒分子模型。为了显示比例，在右上角标出它的底物，ATP和葡萄糖。

蛋白质功能发挥的关键在于能够特异性地并且以不同的亲和力与其他各类分子，包括蛋白质分子结合。蛋白质结合其他分子的区域被称为结合位点，而结合位点常常是从蛋白质分子表面下陷的一个“口袋”；而结合能力与蛋白质的三级结构密切相关，因为结构决定了结合位点的形状和化学性质（即结合位点周围的氨基酸残基的侧链的化学性质）。蛋白质结合的紧密性和特异性可以非常高；例如，核糖核酸酶抑制蛋白可以与人的血管促生蛋白以亚飞摩尔（sub-femtomolar，即 $<10^{-15}$ M）量级的解离常数进行结合，^[14]但却完全不结合（解离常数 >1 M）angiogenin在两栖动物中的同源蛋白抗肿瘤核糖核酸酶。）非常微小的化学结构变化，如在结合位点的某一残基侧链上添加一个甲基基团，有时就可以几乎完全破坏结合；例如，氨酰tRNA合成酶可以分辨侧链结构非常类似的纈氨酸和异亮氨酸，而这两种氨基酸的差别就在于异亮氨酸的侧链多出一个甲基。相同的蛋白质分子结合在一起就可形成同源寡聚体或多聚体，有些多聚体可以形成纤维；而这些形成纤维的蛋白质往往是结构蛋白，它们在单体状态下是球蛋白，通过自结合来形成刚性的纤维。蛋白-蛋白相互作用可以调控酶的活性和细胞周期中的各种进程，并可以使大型的蛋白质复合物得以形成，这样可以将参与同一生物学功能的分子结合到一起，从而提高其工作效率；而结合所诱导的蛋白构象变化对于复杂的信号传导网络的构建也是必不可少的。还有一些蛋白质（如膜蛋白）可以结合或者插入到细胞膜中。

催化作用

细胞中，酶是最被广泛了解和研究最多的蛋白质，它的特点是催化细胞中的各类化学反应。酶的催化反应具有高度的专一性和极高的催化效率。酶在大多数与代谢和异化作用以及DNA的复制、修复和RNA合成等相关的反应中发挥作用。在翻译后修饰作用中，一些酶（如激酶和磷酸酶）可以在其底物蛋白质上增加或去除特定化学基团（如磷酸基团）。目前已知的酶催化的反应有约4000种。^[15]酶可以极大地加速其所催化的反应；例如，与没有酶催化的情况相比，乳清酸核苷-5'-单磷酸脱羧酶的加速作用最高可达 10^{17} 倍（形象地说，在没有酶的情况下完成反应需要七千八百万年，而存在酶的情况下反应只需18毫秒）。^[16]

结合于酶上，并在酶的作用下发生反应的分子被称为底物。虽然酶分子通常含有数百个氨基酸残基，但参与与底物结合的残基只占其中的一小部分，而直接参与底物催化反应的残基则更少（平均为3-4个残基）。^[17]这部分参与底物结合和催化的区域被称为活性位点。有一些酶需要结合一些小分子（辅酶或辅因子）才能够有效发挥催化作用。酶的活性还可以被酶抑制剂所抑制，或被酶激活剂所提高。

訊號傳遞和物質运输

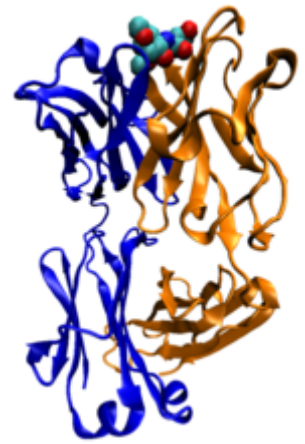
许多蛋白质都参与了细胞中和细胞间的訊號傳遞。一些蛋白质，如胰岛素，作为细胞外蛋白质，可以将訊號从一个细胞（合成这些蛋白质的细胞）传送到身体其他组织中的细胞。还有一些蛋白质，如属于膜蛋白的受体，可以结合细胞外的訊號分子来引发细胞内的生物化学反应；多数受体

都有一个位于细胞外表面的结合域〔结合訊號分子〕和一个位于细胞内的效应结构域（可能具有酶活性或可以发生构象变化以诱发与细胞内其他蛋白质的结合），两者之间通过跨膜域连接。

抗体是适应性免疫系统中重要的组成蛋白质，其作用为结合抗原或机体中的其他外来物质，通过识别来引发免疫系统消除这些物质。抗体可以被分泌到细胞外环境中或锚合到特异性B细胞（即浆细胞）的细胞膜中。抗体和抗原之间存在很高的亲和力，使得抗体可以很快地识别抗原。

在多细胞生物体中，載體运输蛋白能够结合特定的小分子并将它们运送到机体中的特定位置。这些蛋白质在运输的起点（往往具有较高的浓度）必须以高的亲和力结合它们的運輸物，而在目的组织中（運輸物質浓度较低）则必须释放所结合的物質。这就需要运输蛋白和所结合的物質之间有特定的亲和力。一个典型的例子是血红蛋白，它作用是将氧气从肺中运输到其他组织和器官中。[18]

通道蛋白也是重要的物质运输蛋白，它们能够改变细胞膜的通透性，使得一些小分子和离子可以进出细胞。膜本身是疏水性的，极性或带电分子无法通过扩散作用穿过。作为跨膜蛋白的通道蛋白，含有可控制的内部通道，在一定条件下允许这些分子进出细胞。通道蛋白也有专一性，许多离子通道蛋白只选择性地对特定离子起作用；例如，钾离子和钠离子通道分别只识别钾离子或钠离子。[19]



小鼠的抗霍乱抗体与一个糖分子抗原结合的复合物结构图。

變性作用

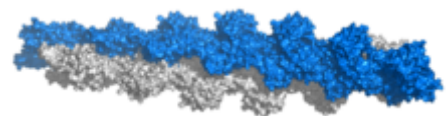
蛋白質受到酸、鹼、尿素、有機溶酶、重金屬、熱、紫外光及X-射線等物理或化學的破壞，引起蛋白質自然之分子結構的改變，並引起生理活性的消失。變性作用破壞了蛋白質的二級、三級、四級結構，一般不會影響其初級結構。

變性蛋白質的特性

1. 溶解度降低。
2. 生物活性消失。
3. 不能結晶。
4. 易受蛋白酶的水解。
5. 滴定曲线改變，因可滴定的官能基增加。
6. -SH等基團的反應活性增加。

结构蛋白

结构蛋白能够形成相对更为刚性的生物组分。多数结构蛋白为纤维蛋白；胶原蛋白和弹性蛋白是结缔组织（如软骨）中关键的组分，而角蛋白则存在于头发、指甲、羽毛、蹄和一些贝壳中。此外，肌动蛋白和微管蛋白作为单体是球状可溶蛋白，但一旦多聚化便形成刚性的纤维用于组成细胞骨架，以保持细胞的大小和形态。



聚合成链状的F型肌动蛋白

其他结构蛋白还包括马达蛋白，如肌球蛋白、运动蛋白和动力蛋白，它们能够产生动力。这些蛋白质对于细胞能动性（特别是精子的运动）、细胞内物质运输和细胞分裂都具有重要作用；^{[20][21]}它们也为肌肉收缩提供动力。^[22]

研究方法

蛋白质是被研究得最多的一类生物分子，对它们的研究包括“体内”(in vivo)、“体外”(in vitro)、和“在硅之中”(in silico)。体外研究多应用于纯化后的蛋白质，将它们置于可控制的环境中，以期获得它们的功能信息；例如，酶动力学相关的研究可以揭示酶催化反应的化学机制和与不同底物分子之间的相对亲和力。相反的，**体内**研究实验着重于蛋白质在细胞或者整个生物中的活性作用，从而可以了解蛋白质发挥功能的场所和相应的调节机制。**在硅之中**研究使用的计算方法来研究蛋白质。

蛋白质纯化

为了进行体外(in vitro)研究，必须先将目的蛋白质从其他细胞组分中分离提纯出来。这一过程通常从细胞裂解开始（对于分泌性蛋白质的提纯则不需要裂解细胞），通过破坏细胞膜将细胞内含物释放到溶液中，从而获得含有目的蛋白质的细胞裂解液。然后通过超速离心将细胞裂解液中膜脂和膜蛋白、细胞器、核酸以及含有可溶蛋白质的混合物。盐析法是一种较为常用的通过沉淀从裂解液中分离浓缩蛋白质的方法。基于目的蛋白质的化学性质（如分子量、带电情况和结合活性），可以利用不同的色谱法来进一步分离提纯蛋白质。纯化的程度可以用电泳（已知目的蛋白质的分子量）、光谱学（目的蛋白质具有独特的光谱学特征）或者酶活分析反应（目的蛋白质具有特定的酶活性）来衡量。另外，蛋白质可以使用电聚焦根据其电荷被分离^[23]。

对于天然蛋白质，可能需要一系列的纯化步骤才能获得纯度足以用于实验室应用的蛋白质。为了简化这一过程，通常采用基因工程的手段在目的蛋白质上添加一些化学特性，在不改变其结构和生物学活性的情况下使纯化过程更为简单。通常是将含有特定氨基酸序列的“标签”连接在目的蛋白质的N-端或C-端。例如，含有连续多个组氨酸的序列，称为组氨酸标签；将含有带组氨酸标签蛋白质的裂解液流过含有镍的亲合层析柱，组氨酸就可以与镍螯合从而结合在柱子上，而裂解液中其他蛋白质由于没有组氨酸标签而直接流出柱子，从而达到分离目的。^[24]通过基因工程（即DNA重组）改造而获得的蛋白质被称为重组蛋白质。

细胞内定位

in vivo的蛋白质研究常常专注于蛋白质在细胞中的合成和定位。虽然已经知道许多细胞内蛋白质是在细胞质中合成，而膜结合蛋白质或分泌性蛋白质是在内质网中合成，但蛋白质定位到特定细胞器或细胞结构的特异性是如何达到的，目前还不清楚。一些有助于获得特定蛋白质在细胞中定位的方法得到了发展，特别是用基因工程将目的蛋白质上连接上“报告者”（如绿色荧光蛋白），将这样的融合蛋白在细胞中表达后，就可以通过显微镜观察荧光来了解融合蛋白在细胞中的分布。^[25]

另一种常用的同样是基因工程的方法为定点突变。利用这一方法，研究者可以改变蛋白质序列，从而改变其结构、细胞内定位以及调控机制；而这些改变可以在in vivo的情况下通过连接绿色荧光蛋白，或者在in vitro的情况下通过酶动力学的方法以及结合实验进行观察。

蛋白质组学

在一定时间内一个细胞或一类细胞中存在的所有蛋白质被称为蛋白质组，研究如此大规模的数据的领域就被称为蛋白质组学，与基因组学的命名方式相似。蛋白质组学中关键的实验技术包括用于检测细胞中大量种类蛋白质相对水平的蛋白质微阵列技术，和用于系统性研究蛋白-蛋白相互作用的双杂交筛选技术。此外，还有探究所有组分之间的可能的生物学相互作用的相互作用组学，以及系统性地解析蛋白质结构，并揭示其中的可能的折叠类型的结构基因组学。

分析基因组和蛋白质组数据，这就需要应用计算技术于生物学问题，如基因识别和支序分类。生物信息学领域现在是基因和蛋白质的分析必不可少的。

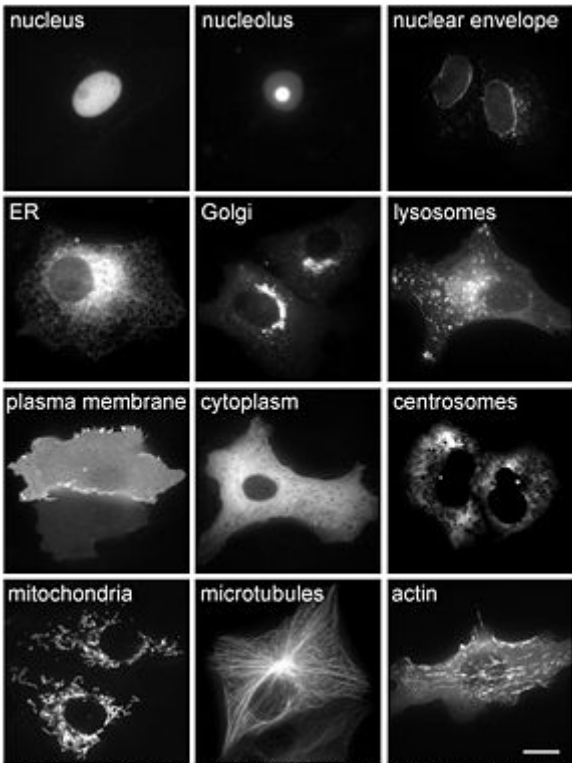
生物信息学

大量的计算方法已经被开发来分析蛋白质的结构，功能和进化。

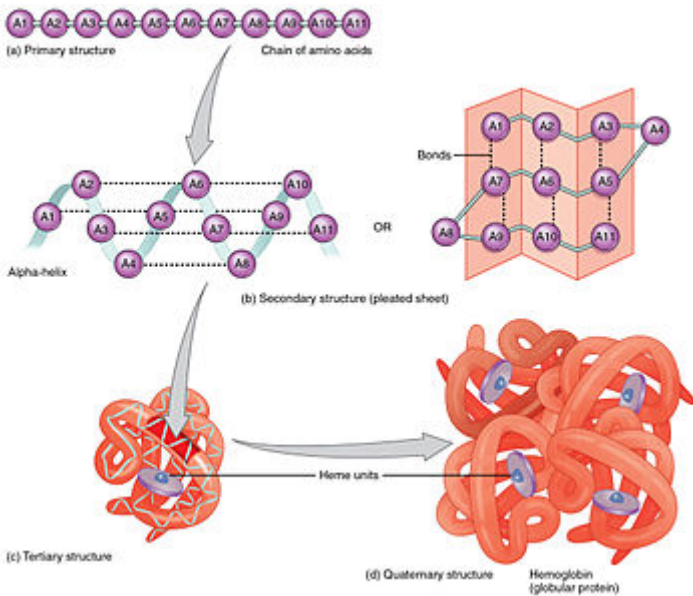
这样的工具的开发是由可用的多种生物体（包括人类基因组）的大量基因组和蛋白质组数据驱动的。根本不可能通过实验来研究所有的蛋白质，因此只有少数蛋白质被得到实验室实验，而计算工具用于外推到相似的蛋白质。

结构预测与模拟

作为结构基因组研究的互补，蛋白质结构预测的目标是发展出有效的能够提供未知结构（未通过实验方法得到）蛋白质的可信的结构模型。目前最为成功的结构预测方法是同源建模；这一方法是利用序列相似的蛋白质（已知结构）的结构作为“模板”。而结构基因组的目标正是通过解析大量蛋白质的结构来为同源建模提供足够的模板以获得剩余的未解析的同源蛋白结构。从序列相似性较差的模板计算出精确的结构模型对于同源建模法还是一个挑战，问题在于序列比准确性的影响，如果能够获得“完美”的比对结果，则能够获得精确的结构模型。^[26]许多结构预测方法已经被用于在蛋白质工程领域，在这些方法的帮助下，研究者们设计出一些新型的蛋白质折叠类型。^[27]更为复杂的结构计算是预测蛋白质分子之间的相互作用，需要应用分子对接法和蛋白-蛋白相互作用预测。



with friendly permission of Jeremy Simson and Rainer Pepperkok
带有绿色荧光蛋白标签的蛋白质在不同的细胞区室和细胞结构中的分布图。荧光以白色来显示。左边从上到下依次为，细胞核、内质网、质膜和线粒体；中间从上到下依次为，核小体、高尔基体、细胞质和微管；右边从上到下依次为，核膜、溶酶体、中心体和微丝。



组成氨基酸可以被分析以预测二级的，三级的和四级的蛋白质结构，在这个例子情况下是含有血红素的血红素单元。

利用分子动力学的方法可以模拟蛋白质的折叠和结合过程。通过分布式计算，如Folding@Home计划，为分子动力学模拟注入了活力。小的α螺旋蛋白结构域，如绒毛蛋白的头部^[28]和HIV辅助蛋白^[29]已经成功地在计算机中（*in silico*）被模拟。将标准的分子动力学和量子力学计算相结合的混合方法已经被用于探索视紫红质分子的电子态。^[30]

蛋白質含量檢測

蛋白質含量檢測，通常是檢測樣本裡的含氮量比例。但是，凱氏分析與杜马斯方法，分辨不出蛋白質與三聚氰胺的差異。一份純瘦肉的雞胸肉以凱氏分析，可以測得數值約31%左右。

- 凱氏分析（凱氏定氮法）
- 杜马斯方法（杜马斯燃烧定氮法）
- 近红外光谱法

营养作用

大多数微生物和植物能够合成所有20种标准氨基酸；动物则由于缺乏某些氨基酸合成途径中特定氨基酸合成反应所需的关键酶，如从天冬氨酸生成赖氨酸、甲硫氨酸和苏氨酸的合成反应第一步中发挥催化作用的天冬氨酸激酶，而只能合成部分氨基酸。因此，动物必须从食物中获取这些自身无法合成的氨基酸。^[13]一个生物体所无法合成而需从食物中获取的氨基酸被称为必需氨基酸；而食物中缺少必需氨基酸的蛋白质，被定位为不完全蛋白质。如果环境中存在所需氨基酸，微生物能够直接摄取这些氨基酸，而下调其自身的合成水平，从而节省了原来需要用于合成反应的能量。

动物所摄取的氨基酸来源于食物中所含的蛋白质，每公克蛋白質可供給4大卡熱量。摄入的蛋白质通过消化作用而被降解，这一过程通常包括蛋白质在消化系统的酸性环境下发生变性，变性后的蛋白质被蛋白酶水解成氨基酸或小段的肽。随后这些降解片段就可以被吸收。部分吸收后的氨基酸被用于蛋白质的合成，其余的则通过糖异生作用被转化为葡萄糖或进入三羧酸循环进行代谢。蛋白质的营养作用在饥饿环境下显得特别重要，此时机体可以利用自身的蛋白质，特别是肌肉中的蛋白质，来产生能量以维持生命活动。^[31]蛋白质/氨基酸也是食物中重要的氮源。

蛋白質進入口腔時無法被分解，進而到胃。胃蛋白酶可斷裂芳香族氨基酸或亮氨酸氨基端肽鍵，凝乳酶可將牛奶中的酪蛋白（Casein）催化成含鈣的副乾酪素（Ca paracaseinate）。到了小腸，胰蛋白酶、糜蛋白酶、羧基肽酶開始作用；胰蛋白酶可斷裂赖氨酸或精氨酸的羧基端肽鍵，糜蛋白酶可斷裂色氨酸、酪氨酸、苯丙氨酸的羧基端肽鍵，羧基肽酶可斷裂肽鏈羧基端的最後一個肽鍵。人體所需蛋白質在許多食物中都含量豐富，如動物肌肉、乳製品、蛋、豆類、穀類和蕈類等。人體中蛋白質缺乏可以導致全身浮腫、皮膚乾燥病變、頭髮稀疏脫色、肌肉重量減輕、免疫力下降等。^[32]

食物中的蛋白质有时会引起过敏反应。

高蛋白饮食会加重心脏病的状况，胆固醇问题会使其上升到不健康的水平。^{[33][34]}

历史和词源

在18世纪，安东尼奥·弗朗索瓦和其他一些研究者发现蛋白质是一类独特的生物分子，他们发现用酸处理一些分子能够使其凝结或絮凝。当时他们注意到的例子有来自蛋清、血液、血清白蛋白、纤维素和小麦面筋裡的蛋白质。荷兰化学家格哈杜斯·约翰内斯·穆德对一般的蛋白质进行元素分

析发现几乎所有的蛋白质都有相同的实验式。用“蛋白质”这一名词来描述这类分子是由穆德的合作者永斯·贝采利乌斯于1838年提出。穆德随后鉴定出蛋白质的降解产物，并发现其中含有为氨基酸的亮氨酸，并且得到它（非常接近正确值）的分子量为131原子质量单位。

对于早期的生物化学家来说，研究蛋白质的困难在于难以纯化大量的蛋白质以用于研究。因此，早期的研究工作集中于能够容易地纯化的蛋白质，如血液、蛋清、各种毒素中的蛋白质以及消化性和代谢酶（获取自屠宰场）。1950年代后期，Armour Hot Dog公司纯化了一公斤纯的牛胰腺中的核糖核酸酶A，并免费提供给全世界科学家使用。目前，科学家可以从生物公司购买越来越多的各类纯蛋白质。

著名化学家莱纳斯·鲍林成功地预测了基于氢键的规则蛋白质二级结构，而这一构想最早是由威廉·阿斯特伯里于1933年提出。随后，沃尔特·考兹曼在总结自己对变性的研究成果和之前凯伊·林诺斯特伦·郎的研究工作的基础上，提出了蛋白质折叠是由疏水相互作用所介导的。1949年，弗雷德里克·桑格首次正确地测定了胰岛素的氨基酸序列，并验证了蛋白质是由氨基酸所形成的线性（不具有分叉或其他形式）多聚体。原子分辨率的蛋白质结构首先在1960年代通过X射线晶体学获得解析；到了1980年代，核磁共振也被应用于蛋白质结构的解析；近年来，冷冻电子显微学被广泛用于对于超大分子复合体的结构进行解析。



约翰·肯德鲁与在进展中的肌红蛋白模型。

蛋白质这一概念最早是由瑞典化学家永斯·贝采利乌斯于1838年提出，但当时人们对于蛋白质在机体中的核心作用并不了解。1926年，詹姆斯·B·萨姆纳揭示尿素酶是蛋白质，首次证明了酶是蛋白质。^[35]

第一个被测序的蛋白质是胰岛素，由弗雷德里克·桑格完成，他也因此获得1958年度的诺贝尔化学奖。^[36]首先被解析的蛋白质结构包括血红蛋白和肌红蛋白的结构，所用方法为X射线晶体学；^{[37][38]}该工作由马克斯·佩鲁茨和约翰·肯德鲁于1958年分别完成，他们也因此获得1962年度的诺贝尔化学奖^[39]。截至到2017年，蛋白质数据库(PDB)中已存有接近126,000个原子级分辨率的蛋白质及其相关复合物的三维结构的坐标^[40]。

参见

- 氨基酸和标准氨基酸列表
- 必需氨基酸
- 蛋白质组
- 蛋白质工程
- 蛋白质设计
- 蛋白质内含子
- 重组蛋白质列表
- 蛋白质列表
- 朊病毒
- PDSS2
- 每单位面积土地可食用蛋白质
- 營養學中的營養素蛋白質

参考文献

1. 阮. 汉典. [2020-06-03]. (原始内容存档于2021-05-02) (中文 (中国大陆)) .
2. 现代汉语词典 (第7版) . 商务印书馆. ISBN 978-7-100-12450-8.
3. (英文) Yan Zhang and Vadim N. Gladyshev. High content of proteins containing 21st and 22nd amino acids, selenocysteine and pyrrolysine, in a symbiotic *deltaproteobacterium* of gutless worm *Olavius algarvensis*. *Nucleic Acids Res.* 2008, **35** (15): 4952–4963 [04.03]. PMID 17626042.
4. (英文) Nelson, D. L. and Cox, M. M. (2005) *Lehninger's Principles of Biochemistry*, 4th Edition, W. H. Freeman and Company, New York.
5. (英文) Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipurksy SL, Darnell J. (2004). *Molecular Cell Biology* 5th ed. WH Freeman and Company: New York, NY.
6. (英文) Dobson CM. (2000). The nature and significance of protein folding. In *Mechanisms of Protein Folding* 2nd ed. Ed. RH Pain. *Frontiers in Molecular Biology* series. Oxford University Press: New York, NY.
7. (英文) Fulton A, Isaacs W. Titin, a huge, elastic sarcomeric protein with a probable role in morphogenesis. *Bioessays.* 1991, **13** (4): 157–61. PMID 1859393.
8. (英文) Bruckdorfer T, Marder O, Albericio F. From production of peptides in milligram amounts for research to multi-tons quantities for drugs of the future. *Curr Pharm Biotechnol.* 2004, **5** (1): 29–43. PMID 14965208.
9. (英文) Schwarzer D, Cole P. Protein semisynthesis and expressed protein ligation: chasing a protein's tail. *Curr Opin Chem Biol.* 2005, **9** (6): 561–9. PMID 16226484.
10. (英文) Nickerson DP, Russell MR, Odorizzi G. A concentric circle model of multivesicular body cargo sorting. *EMBO Rep.* 2007, **8**: 644–50. PMID 17603537.
11. (英文) Branden C, Tooze J. (1999). *Introduction to Protein Structure* 2nd ed. Garland Publishing: New York, NY
12. (英文) Gonen T, Cheng Y, Sliz P, Hiroaki Y, Fujiyoshi Y, Harrison SC, Walz T. (2005). Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438(7068):633-8.
13. (英文) Voet D, Voet JG. (2004). *Biochemistry* Vol 1 3rd ed. Wiley: Hoboken, NJ.
14. (英文) Lee FS, Shapiro R, Vallee BL. Tight-binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. *Biochemistry.* 1989, **28** (1): 225–230. PMID 2706246.
15. (英文) Bairoch A. The ENZYME database in 2000 (PDF). *Nucleic Acids Res.* 2000, **28**: 304–305 [2008-03-13]. PMID 10592255. (原始内容 (PDF)存档于2011-06-01) .
16. (英文) Radzicka A, Wolfenden R. A proficient enzyme.. *Science.* 1995, **6** (267): 90–931. PMID 7809611.
17. (英文) The Catalytic Site Atlas at The European Bioinformatics Institute (<http://www.ebi.ac.uk/thornton-srv/databases/CSA/>) (页面存档备份 (<https://web.archive.org/web/20180927214709/http://www.ebi.ac.uk/thornton-srv/databases/CSA/>), 存于互联网档案馆)
18. (英文) Marden MC, Griffon N, Poyart C. Oxygen delivery and autoxidation of hemoglobin. *Transfus Clin Biol.* 1995, **2** (6): 473–480. PMID 8646344.
19. (英文) Shi N, Ye S, Alam A, Chen L, Jiang Y. Atomic structure of a Na⁺- and K⁺-conducting channel. *Nature.* 2006, **440** (7083): 570–574. PMID 16467789.
20. (英文) Hirokawa N, Noda Y, Okada Y. Kinesin and dynein superfamily proteins in organelle transport and cell division. *Curr Opin Cell Biol.* 1998, **10** (1): 60–73. PMID 9484596.
21. (英文) Gibbons IR. The role of dynein in microtubule-based motility. *Cell Struct Funct.* 1996, **21** (5): 331–342. PMID 9118238.
22. (英文) Curtin NA, Woledge RC. Energy changes and muscular contraction. *Physiol Rev.* 1978, **58** (3): 690–761. PMID 28541.

23. Hey J, Posch A, Cohen A, Liu N, Harbers A. Fractionation of complex protein mixtures by liquid-phase isoelectric focusing. *Methods in Molecular Biology*. *Methods in Molecular Biology*TM. 2008, **424**: 225–39. ISBN 978-1-58829-722-8. PMID 18369866. doi:10.1007/978-1-60327-064-9_19.
24. (英文) Hochuli E, Döbeli H, Schacher A. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J Chromatogr*. 1987, **411**: 177–184. PMID 3443622.
25. (英文) Yuste R. Fluorescence microscopy today. *Nat Methods*. 2005, **2** (12): 902–4. PMID 16299474.
26. (英文) Zhang Y, Skolnick J. (2005). The protein structure prediction problem could be solved using the current PDB library. *Proc Natl Acad Sci USA* 102(4):1029-34.
27. (英文) Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science* 302(5649):1364-8.
28. (英文) Zagrovic B, Snow CD, Shirts MR, Pande VS. (2002). Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing. *J Mol Biol* 323(5):927-37.
29. (英文) Herges T, Wenzel W. (2005). In silico folding of a three helix protein and characterization of its free-energy landscape in an all-atom force field. *Phys Rev Let* 94(1):018101.
30. (英文) Hoffmann M, Wanko M, Strodel P, König PH, Frauenheim T, Schulten K, Thiel W, Tajkhorshid E, Elstner M. (2006). Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II. *J Am Chem Soc* 128(33):10808-18.
31. (英文) Brosnan J. Interorgan amino acid transport and its regulation. *J Nutr*. 2003, **133** (6 Suppl 1): 2068S–72S [2008-03-12]. PMID 12771367. (原始内容存档于2009-04-14) .
32. (中文) 默克诊疗手册, 第一章, 第二节: 蛋白质-能量营养不良. [2008年3月11日]. (原始内容存档于2008年3月13日) .
33. 补充蛋白质有什么好处, 蛋白质的作用及功能, 蛋白质对人体的作用 (一) . WebMD. [2018-12-07] (中文 (中国大陆)) .
34. 补充蛋白质有什么好处, 蛋白质的作用及功能, 蛋白质对人体的作用 (二) . WebMD. [2018-12-07] (中文 (中国大陆)) .
35. (英文) Sumner, JB. The Isolation and Crystallization of the Enzyme Urease. Preliminary Paper (PDF). *J Biol Chem*. 1926, **69**: 435–441 [2008-03-01]. (原始内容存档 (PDF)于2007-09-29) .
36. (英文) The Nobel Prize in Chemistry 1958. Nobelprize.org. [2008-10-06]. (原始内容存档于2018-12-26) .
37. (英文) Muirhead H, Perutz M. Structure of hemoglobin. A three-dimensional fourier synthesis of reduced human hemoglobin at 5.5 Å resolution. *Nature*. 1963, **199** (4894): 633–638. PMID 14074546.
38. (英文) Kendrew J, Bodo G, Dintzis H, Parrish R, Wyckoff H, Phillips D. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*. 1958, **181** (4610): 662–666. PMID 13517261.
39. (英文) The Nobel Prize in Chemistry 1962. Nobelprize.org. [2015-02-20]. (原始内容存档于2019-01-06) .
40. (英文) RCSB Protein Data Bank. [2017-01-19]. (原始内容存档于2015-04-18) .

外部链接

- (英文) 蛋白质数据库 (<http://www.rcsb.org>) (页面存档备份 (<https://web.archive.org/web/20160303165729/http://www.rcsb.org/>), 存于互联网档案馆)
- (英文) UniProt上的蛋白质资源 (<https://web.archive.org/web/20080608183902/http://www.expasy.uniprot.org/>)
- (英文) 人类蛋白质图集 (<http://www.proteinatlas.org>) (页面存档备份 (<https://web.archive.org/web/20170501045727/http://www.proteinatlas.org/>), 存于互联网档案馆)
- (英文) 有关于蛋白质的超链接信息 (<http://www.ihop-net.org/UniPub/iHOP/>) (页面存档备份 (<https://web.archive.org/web/20051017154959/http://www.ihop-net.org/UniPub/iHOP/>), 存于互联网档案馆)
- (英文) 美国国家生物技术信息中心中Entrez上的蛋白质数据库 (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein>) (页面存档备份 (<https://web.archive.org/web/20041117054909/http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein>), 存于互联网档案馆)
- (英文) 美国国家生物技术信息中心的蛋白质结构数据库 (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Structure>) (页面存档备份 (<https://web.archive.org/web/20041117055014/http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Structure>), 存于互联网档案馆)
- (英文) 人类蛋白质参考数据库 (<https://web.archive.org/web/20060424071622/http://www.hprd.org/>)
- (英文) 人类蛋白质百科 (<https://web.archive.org/web/20070314135408/http://www.humanproteinpedia.org/>)
- (英文) 斯坦福大学的Folding@Home (<http://folding.stanford.edu/>) Archived (<https://www.webcitation.org/6Aqq8Z3E7?url=http://folding.stanford.edu/English/HomePage>) 2012-09-21 at WebCite
- (英文) 蛋白质：生物合成降解-生物化学和细胞生物学的虚拟图书馆 (<https://web.archive.org/web/20050219090405/http://www.biochemweb.org/proteins.shtml>)
- (英文) 氨基酸的代谢 (<http://web.indstate.edu/thcme/mwking/amino-acid-metabolism.html>) (页面存档备份 (<https://web.archive.org/web/20190217040602/http://web.indstate.edu/thcme/mwking/amino-acid-metabolism.html>), 存于互联网档案馆)

取自“<https://zh.wikipedia.org/w/index.php?title=蛋白质&oldid=70818749>”

本页面最后修订于2022年3月26日 (星期六) 04:19。

本站的全部文字在知识共享 署名-相同方式共享 3.0协议之条款下提供，附加条款亦可能应用。（请参阅使用条款）

Wikipedia®和维基百科标志是维基媒体基金会的注册商标；维基™是维基媒体基金会的商标。

维基媒体基金会是按美国国内税法501(c)(3)登记的非营利慈善机构。