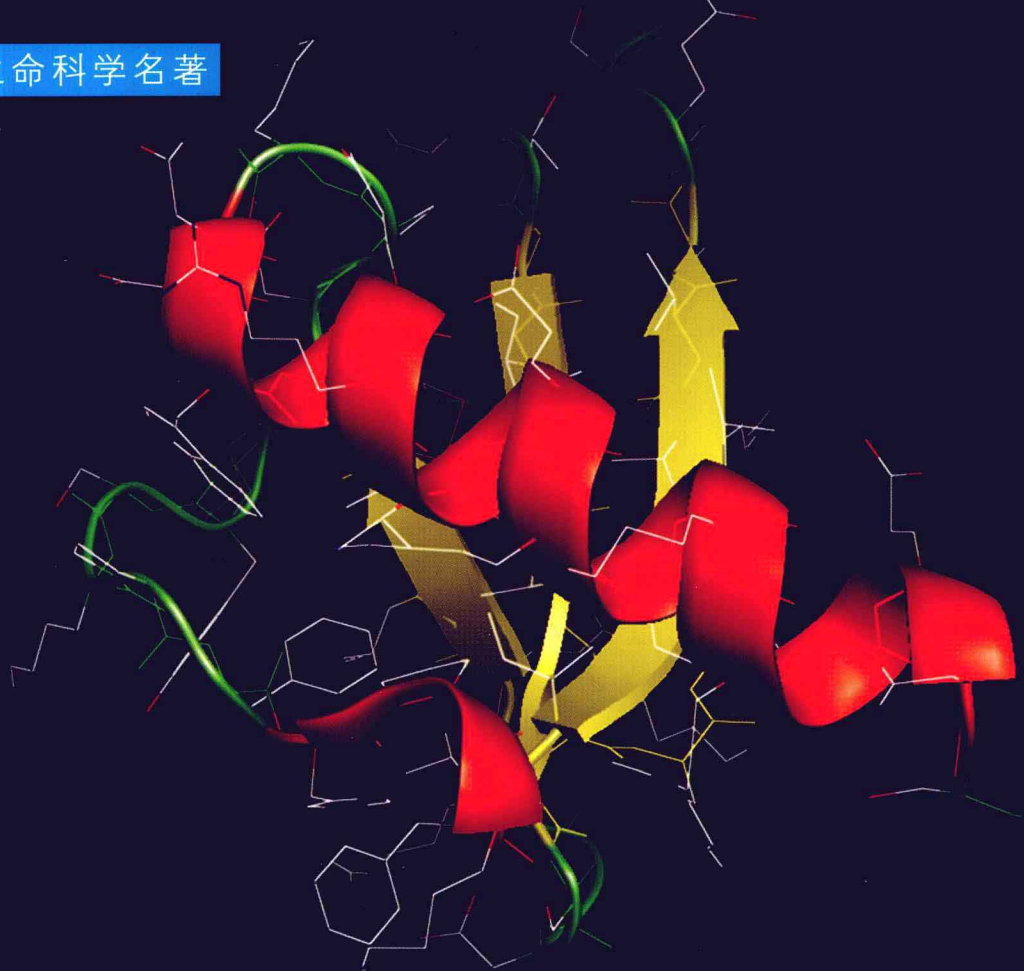


生命科学名著



蛋白质物理 (原书第四版)

Protein Physics (Fourth Edition)

[俄] A. V. 芬克尔施泰因 (Alexei V. Finkelstein)

O. B. 普季岑 (Oleg B. Ptitsyn)

著

李安邦 译



科学出版社

生命科学名著

Protein Physics 蛋白质物理

(原书第四版)

[俄] A.V. 芬克尔施泰因 (Alexei V. Finkelstein)

著

O.B. 普季岑 (Oleg B. Ptitsyn)

李安邦 译

本书的出版得到了国家自然科学基金(项目号:11175068)的支持

科学出版社

北京

图字:01-2013-2616 号

内 容 简 介

本书是一部蛋白质物理学研究领域的系统性著作,由享誉世界的俄罗斯科学家 A. V. 芬克尔施泰因和 O. B. 普季岑根据其于俄罗斯科学院和莫斯科大学的学生讲授蛋白质物理课程时的讲义撰写而成,具有很高的理论和学术价值。本书内容广泛,包括了蛋白质的结构分类、变性与折叠、结构预测与设计等几个部分,其主旨是以物理学(特别是统计物理和热力学)的观点和方法来研究与蛋白质相关的问题。本书循序渐进,深入浅出,不仅以简明易懂的语言和简繁适当的数学估算介绍了与蛋白质相关的基础问题,而且对本领域中的重点研究问题都高屋建瓴地予以严谨的分析和讨论,在研究思想上有很强的指导性和启发性。本书英文版已在美、英、德等国多所大学采用为相关课程的教材。

本书适合于生物物理学、分子生物学、结构生物学等专业的高年级本科生、研究生,以及相关专业的研究人员阅读。

图书在版编目(CIP)数据

蛋白质物理 / (俄) 芬克尔施泰因 (Finkelstein, A. V.), 普季岑 (Ptitsyn, O. B.) 著; 李安邦译. —4 版. —北京: 科学出版社, 2013. 6

(生命科学名著)

书名原文: Protein Physics

ISBN 978-7-03-037698-5

I. ①蛋… II. ①芬… ②李… III. ①蛋白质-物理学-研究生-教材
IV. ①Q510.2

中国版本图书馆 CIP 数据核字(2013)第 118261 号

The original Russian work Protein Physics (Физика белка. Курс лекций) has been published by KDU publishing house.

©2012 by Finkelstein A. V. and Ptitsyn O. B. (Финкельштейн А. В. и Птицын О. Б.)
All right reserved.

责任编辑:李悦 张琬 / 责任校对:包志虹

责任印制:钱玉芬 / 封面设计:陈敬

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

源海印刷有限责任公司 印刷

科学出版社发行 各地新华书店经销

*

2013 年 6 月第 一 版 开本:B5(720×1000)

2013 年 6 月第一次印刷 印张:21 3/4 插页:6

字数:441 000

定价:98.00 元

(如有印装质量问题,我社负责调换)

中文版序

我非常高兴,在我的朋友李安邦——他在莫斯科大学听过我的蛋白质物理课——的巨大努力之下,蛋白质物理课可以进入中国的科学讲堂,世界上最大的讲堂。在这些讲义翻译为中文之前的全新编辑过程中,我大大地增加了供读者练习的习题的数量,当然,还尽力把在前一俄语版出版以来最近7年蛋白质物理学中所出现的所有最突出的成果包括在本书中。

A. V. 芬克尔施泰因 (Alexei V. Finkelstein)

2012年12月

前言(第三版俄文版)

新版《蛋白质物理》在 Oleg B. Ptitsyn(奥列格·普季岑)诞辰 75 周年和他过早地逝世 5 周年之际付印了。

在筹划这个版本时,我首先改正了暴露出的文字错误、某些表格和图片中的不准确或不明确之处。然后,我在第 11 讲和第 18~22 讲中引入了新内容。之后,在 A. K. Цатурян 和 С. Ю. Бершицкий 的帮助下(我对他们深表感谢),我重新编写了第 25 讲中关于肌肉工作的部分。最后,我在新版中增加了一定数量供读者自主练习的习题。

我感谢《蛋白质物理》俄文版和英文版的读者们以及听我的讲座的听众们,他们向我说出了或寄给我他们的意见,使得我对这些讲义进行修正。我特别要感谢以下向我提出意见和提供附加材料的以下人士:Д. Н. Иванков, Е. С. Надеждина, В. Е. Бычкова, О. В. Галзитска, С. О. Гарбузинский, Н. С. Богатырева, Г. В. Семисотнов, Н. Ю. Марченко, Е. Н. Барышникова, М. Г. Шарапов, Л. Б. Переяславц, Ф. К. Гюева, А. Я. Бадретдиний, А. А. Климов, Д. А. Климов, В. Е. Финкульштейн, А. Александреск, А. Кистер, Д. Бейкур 和 К. Плакско。

我感谢 А. А. Веденов, А. М. Гутин, Е. И. Шахнович, Б. А. Реве, М. Д. Франк-Каменецкий, А. М. Дыхне, С. Чотиа, М. С. Гельфанд, И. М. Гельфанд, К. Добсон, А. Фершт, Г. Фринд, А. Трамонтано 和 А. Леск 对讲义中涉及的问题所进行的讨论。

我感谢 М. Г. Никитина 在本书新版本出版工作过程中给予的帮助,感谢我的妻子 М. С. Вильнер-Мармер,没有她长期的支持、关注和协助,本书的准备工作不可能最终完成。

最后,我感谢莫斯科市津贴、ISSEP 基金、俄罗斯科学院蛋白质研究所学术中心和 Algodayn 公司对我的教学活动的支持,感谢俄罗斯科学院、俄罗斯科学院分子和细胞生物学计划、“领先科学流派”计划,俄罗斯基础研究基金^①和霍华德·休斯医学研究所对我们的(由 Oleg B. Ptitsyn 创建的)实验室的长期支持。

A. V. 芬克尔施泰因(Alexei V. Finkelstein)

2004

^①相当于中国的国家自然科学基金

序(第二版英文版)

1967年6月,Oleg Ptitsyn(奥列格·普季岑)成为普希洛新成立的蛋白质研究所的蛋白质物理实验室的负责人。三个月之后,Alexei V. Finkelstein(阿列克谢·V. 芬克尔施泰因)参与到他的工作中来——最初是作为研究生,后来是作为同事。由于受到高分子聚合物物理学的俄罗斯学派的强烈影响,他们研究蛋白质的方法与西方的常用方法并不相同。该流派中最著名的成员之一,Michael Volkenstein,曾经是 Ptitsyn 的博士生导师。Oleg Ptitsyn 和 Alexei V. Finkelstein 一起在普希洛创建了一座世界上杰出的蛋白质物理和化学研究中心。

他们在某些领域特别是关于蛋白质折叠的工作,或是直接通过他们的科学论文,或是间接地通过他们的两位学生——在美国的 Eugene I. Shakhnovich^①,在英国的 Alexey G. Murzin^②——对其工作的辛勤耕耘,已经在欧洲、印度和美国广为人知。但是,很显然,当 Finkelstein 或 Ptitsyn 与同事们交谈时,他们的工作范围远远超过了西方学者所熟知的。我们会发现,我们当前所关注的一些基础问题已经早已被他们思考过,并且已经给出了一些能得出答案的优美的计算。以前,他们没有抽出时间来出版他们的著作。现在,为了让他们的全部成就不只是让他们的朋友和莫斯科大学的学生们可以接触到,Finkelstein 在他和 Ptitsyn 对他们的学生所做的讲座的基础上写成了此书。

无论是从论述范围的宽广性还是从分析和思维一致的严谨性来看,此书都是绝世之作(*tour de force*)。考虑到蛋白质物理和化学所涉及的内容,我毫不怀疑,任何人,无论是他们的学生或高级工作者,都会从中找到新的、有用的且重要的思想、解释和信息。

Cyrus Chothia^③

英国剑桥大学 MRC 分子生物学实验室^④

①Eugene I. Shakhnovich,美国哈佛大学化学与化学生物学系教授。他通过使用统计物理方法等多种方法,对蛋白质折叠问题作出了基础性的重要贡献。例如,他发现了多肽序列成为类蛋白质的判据,在动力学上他发现了成核机制和折叠核。参见本书第16、18和21讲。

②Alexey G. Murzin,英国剑桥 MRC 分子生物学实验室教授,蛋白质结构分类法 SCOP 的发明者。SCOP 提供了对已知蛋白质的结构之间的相互关系的详尽说明,已成为结构基因组学中对新蛋白质进行功能性注释的基础资源。由于 Alexey Murzin 在 SCOP 分类法和数据库中的核心作用,他可能是世界上唯一对所有蛋白质结构都了然于胸的人。参见本书第14、15和23讲。

③Cyrus Chothia,英国皇家科学院院士,英国剑桥 MRC 分子生物学实验室教授。他和 Michael Levitt 共同提出了对蛋白质的“ α 型、 β 型、 $\alpha+\beta$ 型和 α/β 型”结构分类,与 Alexey Murzin 合作建设 SCOP 数据库,与 Julian Gough 共同建立了 SUPERFAMILY 数据库。

④英国剑桥大学 MRC 分子生物学实验室是世界著名的分子生物学研究机构,是现代分子生物学的发祥地之一。

前言(第一版俄文版)

呈现于你们面前的这个讲义是关于蛋白质物理学的,即关于蛋白质分子的最基本的结构、自组织和功能问题。

本讲义是基于我们(开始时是 Oleg B. Ptitsyn, 后来是 Alexei V. Finkelstein) 最初在莫斯科物理技术学院、后来在普希洛^①大学和莫斯科大学^②普希洛分校所做的讲座。讲座最初是针对物理学家的,然后主要是针对生物学家的(部分地对于化学家)。因此在最近这个讲义不仅进行了很大的更新(科学不是停滞不动的),而且进行了彻底修订,以适应生物学课堂。

因此,在大家面前看到的,是课程讲义手稿,而不是专著或教科书。这里难免有重复之处(主要是图片的重复)。因为在实际讲课中不可能说“请看前一讲中的图2和公式3”;不过,我已经尽量把这些重复之处减至最少。

我们深深地感谢在本书中用到的我们的工作的所有共同作者,以及阅读本书手稿、讨论并给出了大量有益评论的 А. И. Четверин, Д. И. Харакоз, В. У. Бычкова, Ю. В. Митин, А. С. Спирин, В. А. Колб, М. А. Ройтберг, В. В. Великов, А. В. Ефимов, И. Г. Птицина, Г. И. Гительзов, Д. С. Рыкунов, Г. В. Семисотнов, А. В. Скугарев и Д. Н. Иванков。

我们感谢在我们俄罗斯科学院蛋白质研究所的所有同事,特别是在我们的物理实验室里工作过和正在工作的。没有与他们的持续不断的交流(科学方面和简单的私人方面的),本书就不可能写成。

我们感谢本讲义的所有听众:我们的学生,以及国际学习班和俄罗斯学习班的参与者,他们的问题和意见让作者们更明确地,首先对于我们自己,处理本讲义中涉及的许多问题。

我们要特别感谢 Г. А. Морозов, Д. С. Рыкунов, А. В. Скугарев, Д. Н. Иванков, М. Ю. Лобанов, Н. Ю. Марченко, И. В. Соколовской, А. А. Шумилин,

^①普希洛科学中心(Pushchino Scientific Centre, <http://www.psn.ru>)是位于莫斯科南部110km处的一个科学城,建立于1963年,现有十多个研究所,主要从事物理化学生物技术领域的研究。目前有3200多人在那里工作,其中1300多人为科研人员,其中850人有副博士或科学博士学位(在苏联/俄罗斯,“副博士”相当于国际上通行的“博士”,而“科学博士”是科学工作者通过多年的科研工作之后进行论文答辩所能获得的最高学位)。

^②罗蒙诺索夫国立莫斯科大学是苏联/俄罗斯最好的综合性大学,列居世界排名前十名内。1755年由教育家М. В. 罗蒙诺索夫倡议并创办,在20世纪50年代扩建。该校目前有超过8600位教师,其中一半有副博士或科学博士学位,125位是俄罗斯科学院院士,在校学生约3万人。

М. Г. Дашкувич, Т. Ю. Сальникова, Н. С. Богатырева, А. Э. Жумаев 和 А. А. Финкульштейн 等在讲义的图片设计和文字排版方面给予的帮助,也要特别感谢俄罗斯基础研究基金对本书出版给予的资助。

事先的半哲学性的注解

关于“讲课人”。所有讲义是按照它们在教室里讲课时那样写作的:即以“讲课人”这个人物的第一人称方式进行。

关于“课堂里的声音(inner voice)”。“讲课人”身份是两个共同作者(Oleg B. Ptitsyn 和 Alexei V. Finkelstein)合二为一而成的。但这并不意味着,共同作者们对所讲述的内容都相互完全赞同。而且,有时候某个作者觉得讲义涉及有争议的、远未最终解释清楚的问题。我们不想掩饰这些反对意见和争论,因此“讲课人”从容不迫的讲述有时会被“课堂里的声音”打断,或提出异议。而有时候“课堂里的声音”只是说出经常向我们提到的问题,使“讲课人”的叙述更明确,或加以深化。

为什么书名是《蛋白质物理》?因为我们惊奇地看到一幅图景,生物进化是如何强烈地增强、巩固并向我们证明作为蛋白质中分子相互作用的基础的物理定律的结果。我们还惊奇地看到,我们对基本生物学(尤其是蛋白质系统)的理解是如何通过物理学方法的运用而得以推进。我们可以从蛋白质的质谱法到电子显微镜、到 X 射线和核磁共振研究中观察到这一切。也许,现代科学中没有哪一个其他领域,科学与其哲学的传统边界被如此巨大的益处所冲破。

关于本书中的物理学和生物学。在讲课过程中我们没有放弃讲述物理概念(尤其是统计力学和量子力学的基本原理)的机会。一方面,在我们看来,它们是绝对必不可少的,这不仅是对理解蛋白质的结构和功能而言,而且是对基本科学素养也是如此,“普通的”生物学学生要不就在专业开始时已经完全忘掉了这些物理概念,要不就根本不知道。

另一方面,在有关蛋白质功能的信息海洋中,我们只舀取了对于展示蛋白质空间结构在它们的生物学——更准确些,生物化学——活动中所发挥的作用的绝对必需的内容。

关于实验、物理理论和计算。实验为我们对现象的认识提供了基本事实,以及大量准确的细节。理论让我们理解现象的本质和联系,有助于设计内容丰富的实验。计算把理论和实验联系起来,在理论的关键点进行检验。不过,需要注意的是,并非所有能计算的东西都需要计算出来。例如,蛋白质或水的密度更容易被测量出来。不只是更容易,而且更可信:详细的计算取决于许多很难被精确估算的参数。

在我们的讲义中包括了对一些基本物理理论的讲述,自然,是以最简单的形式。这不仅是因为这些理论让我们能以普适的观点来包容、简化和理解数量众多的实验资料,同样重要的是,这些理论很优美。除此之外,我们认为,对基本的物理

学理论和模型的理解即使对于通识教育也是必不可少的。

关于体内(*in vivo*)和体外(*in vitro*)的实验。谈到“体内”和“体外”实验,物理学家和生物学家经常有不同的理解。总的来说,并非没有原因:在纯“体内”和纯“体外”之间有许多允许双重解释的层次。例如,在非细胞系统(包含所有核糖体、启动子、终止子、分子伴侣等)中的蛋白质折叠,从物理学家的观点来看很显然是“体内”实验(对他们而说,“体外”应该是溶液中的单个蛋白质;非细胞系统……已经有很多生命实在了!),而在生物学家观点来看它显然是“体外”实验(对他们来说,“体内”是活的,最好是完整的生物体中)。然而结构学研究,如单个蛋白质,在生物体中实际上是不可能的。因此在实践上总是在合理的妥协中进行的:感兴趣的体内生物学现象用可达到的体外实验观测现象来近似。

关于物理模型,粗略估算和计算机模拟(*in silicio*)^①实验。在课程进行中我们会经常讨论简单的即与实际相比大大简化的模型,并进行粗略估算。而且我们希望你们在研读完这些讲义之后,也学会做粗略估算,并在你们感兴趣的事件中使用简单模型。初看上去,为什么所有这些是必需的呢?许多人以为,拥有了强大的计算机,可以把“所有的东西”:水分子、盐、蛋白质和 DNA 的原子坐标,以及其他等等,都输入到里面,给定温度,“所有东西都会准确计算出来”。但实际上这是一个乌托邦的图景。计算,我们指的是详细计算(通常是使用所谓“分子动力学”^②来进行;对整个蛋白质及其周围环境的量子力学计算基本上不可能),需要花费数天,且只覆盖蛋白质生命中的几个纳秒:因为你必须观察数千个原子的热运动和相互作用。顺便要说到,这种计算不可能达到完全精确:对所有的基本相互作用,我们都只是近似地知道。对系统的描述越详细,考虑到的基本相互作用越多,则在我们的计算中产生的细小误差就越多(更别说所耗时间的增长)。因此所有同样得到的不是“终极真理”,而是多少有点近似的某种估算——超级计算机工作多天的估算。而大家更感兴趣的是简单但快速型的估算:在蛋白质某个位置是否可以引入电荷,或蛋白质在这种情况下是否会爆裂?因此我们的目标之一就是:教会你们做这种粗略估算。

但这并不意味着,我们将会完全忽略计算机实验。这种实验也是实验(通过计算机模拟,而不是体外或体内),而且是用非常复杂的系统:它建立起需要将来解释的事实,从而产生简化的、但清晰的理论。

关于如何处理公式。我们知道,数学公式对生物学家来说是很复杂的,因此尽力避免它们,留下的只是确实必需的。我们想给一个建议:在阅读时请“用公式来

^①*in silicio*(以前曾写为“*in silico*”)表示“在计算机上进行的”,或“通过计算机模拟的”之意,有时候也翻译为“硅片实验”。它只应用于模拟自然科学中的自然过程或实验室过程,而不表示由计算机执行的通用计算。

^②在分子动力学中,使用经典力学(而非量子力学)对原子进行描述。

检验文字”！当然，只阅读文字更简单，但是它们常常会产生歧义；因此，用公式检验文字并反向进行，你们可以检查自己的理解程度。

关于小字体文字。它们给出了有益的，但并非必需的题外话，补充和注释。

关于口味。这些讲稿，毫无疑问，表现了我们的个人口味和嗜好。它们更专注于理解现象的本质，而非细致描述它们的细节；在讲义中有很多理论和物理问题，仅有最少的必需的实验事实（几乎没有实验技术；尤其是我们实际上根本没有讲述X射线晶体学和NMR谱学，这些方法是我们为关于蛋白质结构的知识而最应该感谢的）。因此我们的讲义决不能代替正规的富含事实的生物学和生物化学“蛋白质”课程，也不应该作为手册来使用。涉及具体的蛋白质，我们只给出了最重要的（从我们的观点出发！）例子；表格中只包含绝对必需的数据；图片特意做成概要性的；所有数字舍去了尾数；如此等等。

关于这些讲义中的个人成分。我们没有放过机会提到我们自己和我们俄罗斯科学院蛋白质研究所的同事在蛋白质物理方面所做的工作。这当然加大了讲义的“个人成分”，而我们希望能使它们更有生气……

目 录

中文版序

前言(第三版俄文版)

序(第二版英文版)

前言(第一版俄文版)

绪论

第1讲 (1)

蛋白质的主要功能。氨基酸序列决定空间结构,空间结构决定功能。反之不然。球状蛋白,纤维蛋白和膜蛋白。蛋白质的一级、二级、三级和四级结构。蛋白质的生物合成。蛋白质在体内(*in vivo*)和在体外(*in vitro*)时的折叠。翻译后修饰。

蛋白质内部和周围环境的基本相互作用

第2讲 (10)

L型氨基酸残基的立体化学。共价键连接和它们之间的角度。共价键的振动。围绕共价键的转动。肽基团。顺式(*trans*)和反式(*cis*)脯氨酸(Pro)。

第3讲 (16)

范德华相互作用;远距时吸引,近距时排斥。氨基酸残基的允许构象(甘氨酸、丙氨酸、缬氨酸和脯氨酸的拉氏图)。

第4讲 (23)

水环境的影响。氢键。氢键的电本性。氢键的能量。氢键在晶体中的几何形状。氢键在水中的松弛。熵和自由能的概念。肽链中的氢键取代了该链与水的氢键结合;结果使得肽链中的氢键获得了(在水环境中)熵性质。

第5讲(两倍) (32)

热力学基础。自由能与化学势。疏水作用。它们与氢键在水中的饱和必然性的联系。氨基酸的水可及非极性表面及其疏水性。

第6讲(两倍) (42)

水环境对静电相互作用的影响。蛋白质表面和内部的静电场。介电系数。盐溶液中的电荷屏蔽。利用蛋白质工程测量蛋白质中的电场。二硫键。配位键。

多肽链的二级结构

第7讲 (56)

多肽的二级结构。螺旋: $2_7, 3_{10}, \alpha, \pi, \text{poly(Pro) II}$ 。反平行和平行 β 结构, β 转折。测定二级结构的实验方法。

第8讲 (64)

统计物理原理。温度与能量和熵变化的关系。不同能量状态的概率(玻尔兹

曼-吉布斯分布)。配分函数及其与自由能的关系。构象转变。一级相变(“全或无”转变)和非相变的概念。构象转变过程克服自由能垒的内能。绝对反应速度理论的概念。并行过程和串行过程。扩散过程的特征速度。

第9讲 (79)

α 螺旋初始化和延长的自由能。朗道理论与螺旋-线团转变的非相变性。螺旋-线团转变中的协同区段的尺寸。 α 螺旋在水中的稳定性。 β 结构在水中的稳定性。 α 螺旋和 β 结构的形成速度。“线团”是什么?

第10讲 (90)

氨基酸残基侧链基团的性质。氨基酸残基参与到二级结构中。丙氨酸(Ala)、甘氨酸(Gly)、脯氨酸(Pro)、缬氨酸(Val)。非极性的、短极性和长极性的侧链基团。带电侧链基团。蛋白质二级结构的疏水性表面。

蛋白质的空间结构

第11讲 (96)

纤维蛋白,其功能和周期性一级与二级结构; α 角蛋白, β 丝蛋白,胶原。长 α 螺旋和宽 β 片的拼装。形成基质的蛋白质;弹性蛋白。蛋白质遗传缺陷与疾病。淀粉样纤维。

第12讲 (104)

膜蛋白,它们的结构与功能的特性。细菌视紫红质,受体和 G-蛋白,孔蛋白。光合作用中心。膜通道的选择通透性。光合作用中心的工作方式。隧道效应的概念。光子-构象相互作用的概念。

第13讲 (113)

球蛋白。蛋白质球结构的简化表示;结构分类。 β 型蛋白质的结构; β 折叠片,它们的纵向和垂直正交拼装。 β 蛋白质中反平行结构占优势。折叠片右手扭转。 β 蛋白质的拓扑结构。

第14讲 (125)

α 型蛋白质的构造。螺旋束和螺旋层。 α 螺旋构成的准球模型。 α 螺旋接触下的紧凑拼装。 α/β 型蛋白质构造; α 螺旋覆盖的平行 β 折叠片,以及 α/β 圆筒。 β - α - β 亚基的拓扑结构。 $\alpha+\beta$ 型蛋白质的构造。蛋白质的架构与其功能之间并无直接联系。

第15讲 (136)

蛋白质结构的分类。未观察到蛋白质链折叠的“宏观进化”,虽然观察到了它们的“微观进化”。基因复制与专一化。域混合产生的进化。“标准的”三级结构。氨基酸在球状蛋白一级结构中的“准随机”交替的典型性,与纤维蛋白的周期性结构和膜蛋白的块状结构的对比。蛋白质球体构造的物理原理。在蛋白质球体结构中观察到的基本规律:存在单独的 α 和 β 层;罕见有环线的交叠;罕见有沿链相邻结构片段的平行;罕见有左 β - α - β 超螺旋。结构中很少存在“能量缺陷”和“熵缺陷”,这些“缺陷”与能使“缺陷”稳定的蛋白质序列

的相对罕见性之间的关系。多数原则。

第 16 讲 (149)

对随机和准随机氨基酸序列,可以期望有哪种二级结构? 准随机长序列的稳定空间结构的域构造最有可能。蛋白质结构细节的准玻尔兹曼统计。准玻尔兹曼统计来自于对稳定的蛋白质结构的物理性选择。结构元件的稳定性对球状蛋白未破坏的空间结构的一级结构的选择严格性的影响,或者说:为什么一种蛋白质结构经常出现,而其他的很罕见,在大蛋白质球中心哪一种结构, α 或 β ,更值得期待?“熵性缺陷”与“能量缺陷”之间的联系。球状蛋白是作为“中选”的随机肽链而出现的吗? 蛋白质工程中“蛋白质有利的”随机序列的选择。

蛋白质分子中的协同转变

第 17 讲(两倍) (163)

蛋白质的变性。天然解折叠的蛋白质。协同转变。蛋白质变性的可逆性。球状蛋白的变性是“全或无”型转变。“全或无”型转变的范特霍夫(Van't Hoff)判据。热变性和冷变性。蛋白质分子的相状态图。变性蛋白质看上去怎么样? 线团与熔球。熔球的不均匀性。“普通”高分子聚合物球体膨胀时不存在“全或无”型的相变。

第 18 讲 (179)

为什么球状蛋白的变性是“全或无”型转变? 蛋白质核心的紧凑拼装包的解体和侧链基团的解放。溶剂渗入到变性蛋白质中,熔球的破坏,变性蛋白质链随着溶剂力的增强而解折叠。蛋白质链的自然折叠与它的其他球状折叠之间的能隙:蛋白质链与随机高分子同聚物的主要物理差别。“中选的”高分子杂聚物(有能隙)与随机高分子共聚物在熔化时的差别。

第 19 讲 (189)

蛋白质结构在体内(*in vivo*)和在体外(*in vitro*)的形成。在体内自动折叠中的辅助机制:共翻译折叠,分子伴侣及其他。自发的自动折叠在体外有可能。“Levinthal 佯谬”。非细胞系统中的蛋白质折叠实验,以及关于单词“在体外(*in vitro*)”的不同理解。蛋白质自动折叠的分级机制。许多蛋白质中发现了亚稳的(积聚的)折叠中间体。熔球——通常(但并非必须)在天然条件下观察到的蛋白质折叠中间体。某些蛋白质的最简单(单步骤)折叠,没有任何积聚的亚稳中间体。膜蛋白的自动折叠。

第 20 讲 (201)

小蛋白质的单步骤折叠。过渡态理论。在实验上寻找和研究蛋白质折叠中的不稳定过渡态。天然蛋白质结构的折叠核。蛋白质工程上用体外方法对它们的实验观察。蛋白质折叠的成核机制。

第 21 讲 (211)

“Levinthal 佯谬”的解决:快速折叠路径网络自动引导肽链达到稳定结构。为此必需的只是在肽链的天然折叠和其他球状折叠之间存在明显的能隙。对几种蛋

白质[塞尔平(serpins)和普利昂(prion)]中形成稳定结构异常慢速的讨论。蛋白质“能景”的概念。蛋白质结构:自组织物理学和自组织链的自然选择。

蛋白质结构的预测与设计

第22讲 (226)

根据氨基酸序列预测蛋白质结构的需求。根据序列同源性对蛋白质结构和功能的“识别”。蛋白质家族的一级结构谱(profile)。蛋白质结构的关键区域和功能位点。蛋白质链的稳定空间结构的释出。蛋白质结构的“模板”。我们总是被迫根据作用于链中的部分相互作用判断预测得的蛋白质结构。结果:概率性预测。使多肽二级结构稳定和破坏的相互作用。非球状多肽的二级结构的计算。蛋白质二级结构的预测。

第23讲 (238)

根据氨基酸序列预测蛋白质空间结构的方法概述。蛋白质结构数据库。远同源链的通用折叠预测降低了蛋白质结构识别中的不确定性。结构基因组学和蛋白质组学。生物信息学。蛋白质工程和蛋白质设计。蛋白质结构设计中的初步成功。

蛋白质功能的物理学原理

第24讲(两倍) (252)

蛋白质功能与结构。基本功能。结合蛋白:DNA结合蛋白,免疫球蛋白。酶。活性中心:球形酶功能的“缺陷”。蛋白质的硬度对基本酶功能很重要。催化和底物结合中心。抑制作用。辅因子。多价离子。酶催化的机制。例子:丝氨酸蛋白酶。催化中的过渡态理论及蛋白质工程方法对其的证明。抗体酶。催化作用的特异性。“钥匙-锁”概念。

第25讲 (268)

功能的组合。底物从一个活性中心到另一个活性中心的转移。“双层筛”增加功能特异性。蛋白质结构对它的基本酶活性的相对独立性。蛋白质结构与其环境的可见联系。蛋白质基本功能与其结构的匹配。诱导契合。蛋白质域的活动性。蛋白质域在进化中的移动。域结构:激酶,脱氢酶。变构:阳离子中心的相互作用。蛋白质功能的变构调节。变构与蛋白质四级结构。血红蛋白与肌红蛋白。肌肉收缩机制。肌丝行走。

后记 (282)

译后记 (283)

推荐文献 (284)

练习、解答与评注 (286)

索引 (328)

彩色立体插图

绪 论

第 1 讲

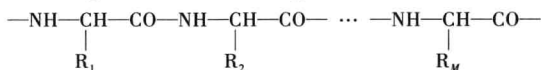
蛋白质的主要功能。氨基酸序列决定空间结构,空间结构决定功能。反之不然。球状蛋白,纤维蛋白和膜蛋白。蛋白质的一级、二级、三级和四级结构。蛋白质的生物合成。蛋白质在体内(*in vivo*)和在体外(*in vitro*)时的折叠。翻译后修饰。

蛋白质是分子机器、建筑模块和活细胞的铠甲。蛋白质最重要的,也几乎是唯一的功能,是对细胞内和细胞周围的化学转换进行酶促催化作用。除此之外,调节蛋白控制基因表达,受体蛋白(它们位于脂膜上)保障细胞间信号的感知,而信号常常是由激素(也是蛋白质)传递的。免疫蛋白和与之类似的组织相容性蛋白质既能识别和结合“异种”分子,也能识别和结合“自身的”细胞,从而让这些细胞能正确地安置在生物体中。结构性蛋白构成微丝和微管,以及纤丝、毛发、丝和其他保护性覆盖物,它们加固细胞膜,支撑细胞和组织的结构。运输蛋白转运(而储存蛋白——储存)其他分子。跨膜传递质子和电子的蛋白质保障所有的生物能学过程:光吸收,呼吸,ATP 生产。其他蛋白质,通过“燃烧”ATP,保障机械化学活动,它们在肌肉中工作或移动细胞成分。

尽管有这些多样性,蛋白质的工作总是以它们和与之发生作用的分子之间的相互作用的高度专一性:就像钥匙和锁(更准确些:像柔软的钥匙和柔软的锁)为基础的。为了保障这种相互作用,蛋白质必须有足够“坚硬”(“工作中的”蛋白质在任何情况下)的空间结构。所以蛋白质(以及对生命非常重要的其他大分子:DNA 和 RNA)的生物学功能与它们特定的三维结构密切相关。这些结构的破坏,甚至只是小小的改变,常常都会导致蛋白质活性的丧失或显著改变。

有关蛋白质分子三维结构的知识对理解蛋白质分子的功能是必不可少的。因此我将在最后才讲述蛋白质功能的物理学,开始时集中注意力于它们的结构,关注它们的稳定性和自组织能力的本质。

■ 蛋白质是高分子聚合物^①；由氨基酸残基组成的多肽链^②；这是由 E. Fischer 于 20 世纪初所发现的。在 20 世纪 50 年代早期, Sanger 证明蛋白质链具有唯一的链节序列, 其链节是氨基酸残基。“残基”是游离氨基酸聚合成为蛋白质链之后留下来的部分。该链具有化学上规则的骨架(“主链”), 从骨架上伸出各种不同的氨基酸侧链基团: R_1 原子团, R_2 原子团, …… , R_M 原子团。



蛋白质链中的链节数目 M 是由基因编码的, 有时是几十个, 有时可达几千个。

存在有 20 种基本氨基酸残基。它们在蛋白质链中的位置是由基因编码的。但是后续的蛋白质修饰有时会增加氨基酸的多样性。此外, 在某些蛋白质中还包含其他分子(辅因子)。

在“工作中的”蛋白质中, 肽链以严格确定的方式折叠起来。在 20 世纪 50 年代末期, Perutz 和 Kendrew 解析出了第一批蛋白质空间结构, 并证明了它们的结构具有高度复杂性和唯一性。然而, 蛋白质分子结构的严格确定性实际上是在 19 世纪 60 年代由 Hoppe-Zeiler 首次证明的(这是后来才弄清楚的), 他获得了血红蛋白的晶体: 因为在晶体中每个分子的每个原子都占据唯一的位置。

蛋白质在晶体中和溶液中的结构是否相同? 这一问题讨论了多年: 当时只有间接的数据可以证明, 直到最后被核磁共振(nuclear magnetic resonance, NMR)证实, 实际上通常是一致的(在误差范围内)。

蛋白质“生活”在各种不同的环境中, 其结构带有环境的明显烙印。蛋白质周围的水越少, 越难弥补那些拉紧肽链的强力氢键的断裂, 这些氢键(正是它们稳固了蛋白质分子的骨架结构)对蛋白质就越重要, 蛋白质的稳定结构就必须越规则。

粗略地说, 根据“生活环境”和基本构造类型, 蛋白质可以分为三类。

1. 纤维蛋白: 它们形成巨大的聚集体, 其中极少有水; 它们的结构通常非常规则, 基本上由不同的链之间的相互作用所维系。

2. 膜蛋白: 它们“生活”在没有水的膜中, 但某些部分从膜内伸出到水中。这类蛋白质的膜内部分(纤维蛋白也是如此)非常规则且由氢键结合, 但是这些规则部分的尺寸受限于膜的厚度。

3. 水溶性球状蛋白: 它们生活在水中, 最不规则(尤其是较小的); 其结构由蛋

①高分子聚合物(polymer, 也常简称为“高分子”)指由许多相同的、简单的结构单元(或称为链节、单体)通过聚合反应后由共价键重复连接而成的相对分子质量很高(通常可达 $10 \sim 10^6$)的化合物。如果所有结构单元是完全相同的, 则称为同质聚合物(同聚物), 如果聚合物包含多种不同的结构单元, 则称为异质聚合物(异聚物、杂聚物或共聚物)。橡胶、纤维和塑料等都是常见的高分子聚合物, 而蛋白质、DNA 和 RNA 等生物分子也都是高分子聚合物。

②由两个氨基酸分子脱水聚合而成的化合物称为二肽, 同理类推还有三肽、四肽、五肽等。通常称由 20 个以上的氨基酸残基组成的化合物称为多肽。(也有文献把由 2~10 个氨基酸组成的肽称为寡肽或小分子肽; 10~50 个氨基酸组成的肽称为多肽; 由 50 个以上的氨基酸组成的肽就称为蛋白质。)

白质链自身的相互作用所维系,而且链上相距较远、但空间上接近的烃基^①(“疏水的”,即“怕水的”)基团之间的相互作用尤其重要,有时候蛋白质链与辅因子之间的相互作用也很重要。

最后,较小的,或者缺少烃基基团的多肽可能自身并不具有固定的结构,但是在与其他大分子相互作用时而获得固定结构。

无疑,上述分类是极为粗糙的。有时候蛋白质可以由纤维状的“尾部”和球状的“头部”(如肌球蛋白)构成,如此等等。

迄今(2012年)已经知道几百万条蛋白质氨基酸序列(已经建立了专门的计算机数据库储存它们,如 UniProtKB^②)和约十万个蛋白质空间结构(它们储存在计算机蛋白质结构数据库中:Protein Data Bank,缩写为 PDB^③)。已知三维结构的蛋白质,绝大部分都属于水溶性球状蛋白。对于膜蛋白和纤维蛋白,已解析出的仅仅是相对较少的空间结构或它们的独立片段。原因很简单:水溶性蛋白容易分离成单个分子,在溶液中它们的结构容易用 X 射线研究(在晶体中,或者用核磁共振谱研究)。因此,讲到“蛋白质结构”、“蛋白质结构的形成”等时,实际上指的是仅仅在水溶性球状蛋白中已证实的规律。在阅读书籍或文章时,应该对此牢记在心,阅读本书也是如此。此外,还需要记住,由于实验的原因,当代所有的蛋白质物理学,也主要是关于小蛋白质的。大蛋白质分子及其复合体的物理学,才刚刚开始发展。

■ 维系蛋白质空间构造的非共价键相互作用,比在蛋白质链中固定单体(残基)序列的共价键要弱得多。这种氨基酸残基序列被称为“蛋白质一级结构”(图 1-1)这是在根据写在基因中的“指令”而进行的生物化学基质合成过程中生成的。

蛋白质的架构(architecture)^④,尤其是球状的水溶性蛋白,非常复杂且种类繁多,与 DNA 双链单一性全然不同(顺便说一句,单链 RNA 像是介于两者之间)。但是在蛋白质中发现了一套“标准”的结构,关于它们我们将会非常详细地讲到。

在这里首先要谈到蛋白质的规则二级结构:α 螺旋和 β 结构;α 螺旋通常用螺旋状的丝带(图 1-1)或圆筒表示,而伸直的 β 结构片段(它们黏在一起会形成薄片)用箭头表示(图 1-1)。二级结构的特点是主链具有规则的周期性形状[或者如通常说的“构象(conformation)”],尽管侧链基团的构象各不相同。

①碳原子和氢原子所构成的化合物称为烃类。烃类分子中去掉一个或几个氢原子而成的基团称为烃基。例如,甲基 CH_3 —、乙基 CH_3CH_2 —、苯基 C_6H_5 —等。

②UniProt(<http://www.uniprot.org/>)是信息最丰富、资源最广的蛋白质数据库。它由整合 Swiss-Prot、TrEMBL 和 PIR-PSD 三大数据库的数据而成。UniProtKB(UniProt 知识库)主要由两部分组成:UniProtKB/Swiss-Prot(包含手工注释、已校验的条目)和 UniProtKB/TrEMBL(包含自动注释、未校验的条目)。

③蛋白质结构数据库(PDB)是美国纽约 Brookhaven 国家实验室于 1971 年创建的。1998 年由美国国家自然科学基金委员会、能源部和卫生研究院资助,成立了结构生物学合作研究协会(RCSB)来管理 PDB(<http://www.rcsb.org/pdb>)。

④本书中“architecture”翻译为“架构”,指的是骨架结构。