



新生物学丛书

蛋白质模拟

——原理、发展和应用

王存新 等 编著



科学出版社

新生物学丛书

蛋白质模拟 ——原理、发展和应用

王存新 等 编著



科学出版社

北京

内 容 简 介

本书对蛋白质分子模拟领域的原理、发展和应用,特别是对该领域的热点和难点问题,结合实际进行了深入的讨论。本书力求做到理论联系实际,学术思想新颖,内容主要包括分子动力学模拟方法、蛋白质复合物结构预测、用分子模拟方法研究蛋白质折叠、粗粒化模型、长程静电相互作用,以及药物分子设计方法与应用。

本书不仅适合于从事计算生物学、蛋白质分子模拟和分子设计的专业技术人员,刚开始接触生物分子模拟的人员学习参考,而且可供高等学校及科研院所的教师、研究人员和研究生参考,也可选为分子模拟和生物信息学、系统生物学等课程的指定教材和参考书。

图书在版编目(CIP)数据

蛋白质模拟:原理、发展和应用/王存新等编著. —北京:科学出版社, 2016.3

(新生物学丛书)

ISBN 978-7-03-047327-1

I. ①蛋… II. ①王… III. ①蛋白质—模拟方法 IV. ①Q51

中国版本图书馆 CIP 数据核字(2016)第 026738 号

责任编辑:罗 静 刘 晶 / 责任校对:郑金红

责任印制:徐晓晨 / 封面设计:刘新新

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>

北京教图印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2016 年 3 月第 一 版 开本:787×1092 1/16

2016 年 3 月第一次印刷 印张:27 3/4

字数:650 000

定价:138.00 元

(如有印装质量问题,我社负责调换)

《新生物学丛书》专家委员会成员名单

主 任：蒲慕明

副主任：吴家睿

专家委员会成员（按姓氏汉语拼音排序）：

昌增益	陈洛南	陈晔光	邓兴旺	高 福
韩忠朝	贺福初	黄大昉	蒋华良	金 力
康 乐	李家洋	林其谁	马克平	孟安明
裴 钢	饶 毅	饶子和	施一公	舒红兵
王 琛	王梅祥	王小宁	吴仲义	徐安龙
许智宏	薛红卫	詹启敏	张先恩	赵国屏
赵立平	钟 扬	周 琪	周忠和	朱 祯

《蛋白质模拟——原理、发展和应用》编委会名单

主任：王存新

编委（按姓氏汉语拼音排序）：

常 珊 陈慰祖 丛肖静 龚新奇 何红秋

胡建平 焦 雄 孔 韧 卢本卓 李春华

刘 斌 刘 明 彭 波 齐立省 苏计国

孙庭广 谭建军 王存新 许先进 张小轶

编著者简介



王存新 1943 年 11 月生，教授，博士生导师。1968 年毕业于中国科学技术大学，1978~1998 年在中国科大任教，1998 年作为学术带头人调入北京工业大学工作。曾赴美国、法国、意大利从事合作研究多年。主要从事蛋白质模拟和药物设计研究。主持完成国家及省部级科研项目 20 余项，培养硕士、博士研究生 50 余人，在国内外学术期刊上发表论文 200 余篇，出版专著 2 部、译著 1 部，获国家发明专利 10 余项。曾获国家自然科学奖三等奖、中国科学院自然科学奖二等奖、国务院颁发的政府特殊津贴、北京市突出贡献专家、北京市优秀教师、北京市教学名师等多项奖励和荣誉称号。

(按姓氏汉语拼音排序)



常珊，副教授。2004 年于江苏工业学院信息科学系获学士学位，2009 年获北京工业大学生物医学工程博士学位，2012~2013 年在美国密苏里大学哥伦比亚分校做博士后研究。现任江苏理工学院生物信息与医药工程研究所所长，从事生物大分子复合物结构预测和药物设计研究。



陈慰祖，女，教授。1968 年毕业于中国科学技术大学，1981~1998 年在中国科学技术大学任教，1998 年调入北京工业大学工作。曾赴美国、法国、意大利进行学术访问和研修。主要从事蛋白质结构和药物设计研究。主持并参与完成国家及省部级科研项目 10 余项，在国内外学术期刊上发表论文 100 余篇，参与完成国家发明专利多项。



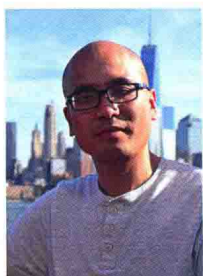
丛肖静，女，2005 年于南开大学化学学院获学士学位，2009 年获北京工业大学生物物理学硕士学位，2013 年获意大利国际高等研究院功能与结构基因组学博士学位。目前从事医药领域中生物大分子体系的分子模拟和药物设计研究。



龚新奇，2001 年于北京航空航天大学获学士学位，2010 年于北京工业大学获生物医学工程博士学位，曾在清华大学做博士后、助理研究员。目前在中国人民大学数学科学研究院任副教授、Tenure-track 研究员、博士生导师、课题组组长，从事生物信息学与数学的交叉理论、方法与应用研究。



何红秋，副研究员。2004 年获昆明理工大学生物工程学士学位，2010 年获北京工业大学生物医学工程博士学位，2010 年至今在重庆市科学技术研究院生物医药与器械研究中心工作。



胡建平，1999 年于吉林大学生命科学学院获学士学位，2003 年于北京工业大学获硕士学位，2008 年于北京工业大学获生物医学工程博士学位；2015~2016 年在美国匹兹堡大学做访问学者。目前从事分子模拟、药物设计和生物活性评价等方面研究。



焦雄，2007 年于北京工业大学获生物医学工程博士学位，现为太原理工大学应用力学与生物医学工程研究所副教授。主要从事系统生物学与生物信息学等研究。



孔韧，女，2002 年获天津科技大学精细化工学士学位，2008 年于北京工业大学获生物医学工程博士学位，2008~2011 年在苏州爱斯鹏药物研发有限公司从事抗肿瘤新药研发，2011~2015 年在美国康奈尔大学威尔医学院休斯顿卫理医院研究所做博士后研究，目前在江苏理工学院生物信息与医药工程研究所从事分子模拟与药物设计等研究。



卢本卓，1993 年于华东师范大学物理系获学士学位，2002 年获中国科学技术大学生物化学与分子生物学博士学位。2003~2008 年在美国加州大学圣地亚哥分校及霍华德休斯医学研究所做博士后及研究人员。现为中国科学院计算数学与科学工程计算研究所研究员，从事计算化学/生物/数学交叉研究。



李春华，女，北京工业大学生命科学与生物工程学院教授，博士生导师。分别于 1997 年、2000 年获内蒙古大学物理学士学位和生物物理硕士学位。2003 年获北京工业大学流体力学博士学位，并获全国百篇优秀博士学位论文提名奖。目前从事蛋白质-蛋白质/RNA 识别、蛋白质折叠和药物设计研究。



刘斌，助理研究员。2005 年获扬州大学理学学士学位，2011 年获北京工业大学生物医学工程博士学位并在北京工业大学从事博士后研究，致力于高通量药物筛选与剂型设计工作。2013 年博士后出站后在中国科学院老专家技术中心工作。



刘明，副研究员。2005 年获北京工业大学生物医学工程学士学位，2010 年获北京工业大学生物医学工程博士学位。后在华东理工大学做博士后，致力于治疗性抗体的优化。2013~2015 年在中国医学科学院药物研究所继续从事治疗性抗体研发工作。2015 年至今，于舒泰神（北京）生物制药股份有限公司任技术总监，负责治疗性抗体和蛋白质的发现与优化。



彭波，2010 年于湘潭大学数学与计算科学学院获得学士学位，2010 年至今于中国科学院科学与工程计算国家重点实验室读博士学位，研究方向为生物分子模拟计算。



齐立省，女，副教授。2009 年于北京工业大学获生物医学工程博士学位。现在山东省生物物理重点实验室工作，主要从事基于分子动力学模拟的蛋白质结构-功能关系研究。



苏计国，副教授。2000 年于兰州大学物理科学与技术学院获学士学位，2011 年获北京工业大学生物医学工程博士学位。目前在燕山大学理学院应用物理系工作，主要从事蛋白质和 RNA 的折叠机理及结构-功能关系研究。



孙庭广，副教授。2002 年于西北大学生命科学学院获学士学位，2009 年获北京工业大学生物医学工程博士学位，2009 年至今在广西科技大学工作，主要从事生物大分子结构与功能关系方面的研究。



谭建军，北京工业大学生命科学与生物工程学院副研究员。2007 年于北京工业大学获流体力学博士学位。现任 *Pharmaceutical Chemistry Review* 与《计算生物学》期刊编委。主要从事药物设计研究。主持参与多项国家科研项目，发表学术论文 60 余篇，获国家发明专利多项。



许先进，2008 年于湖南大学获应用物理学学士学位，2014 年于北京工业大学获生物医学工程博士学位。2011~2012 年于意大利图西雅大学生物物理与纳米科学中心交流学习。目前在密苏里大学道尔顿心血管研究中心做博士后，从事蛋白质-配体相互作用及识别机制研究。



张小轶，女，北京工业大学生命科学与生物工程学院副教授。2000 年于华东师范大学获生物化学与分子生物学硕士学位，2009 年于北京工业大学获生物医学工程博士学位。主要从事计算机辅助药物设计研究。

《新生物学丛书》丛书序

当前，一场新的生物学革命正在展开（为此，美国国家科学院研究理事会于 2009 年发布了一份战略研究报告，提出一个“新生物学”（New Biology）时代即将来临。这个“新生物学”，一方面是生物学内部各种分支学科的重组与融合；另一方面是化学、物理、信息科学、材料科学等众多非生命学科与生物学的紧密交叉与整合。

在这样一个全球生命科学发展变革的时代，我国的生命科学研究也正在高速发展，并进入了一个充满机遇和挑战的黄金期。在这个时期，将会产生许多具有影响力、推动力的科研成果。因此，有必要通过系统性集成和出版相关主题的国内外优秀图书，为后人留下一笔宝贵的“新生物学”时代精神财富。

科学出版社联合国内一批有志于推进生命科学发展的专家与学者，联合打造了一个 21 世纪中国生命科学的传播平台——《新生物学丛书》。希望通过这套丛书的出版，记录生命科学的进步，传递对生物技术发展的梦想。

《新生物学丛书》下设三个子系列：科学风向标，着重收集科学发展战略和态势分析报告，为科学管理者和科研人员展示科学的最新动向；科学百家园，重点收录国内外专家与学者的科研专著，为专业工作者提供新思想和新方法；科学新视窗，主要发表高级科普著作，为不同领域的研究人员和科学爱好者普及生命科学的前沿知识。

如果说科学出版社是一个“支点”，这套丛书就像一根“杠杆”，那么读者就能够借助这根“杠杆”成为撬动“地球”的人。编委会相信，不同类型的读者都能够从这套丛书中得到新的知识信息，获得思考与启迪。

《新生物学丛书》专家委员会

主 任：蒲慕明

副主任：吴家睿

2012 年 3 月

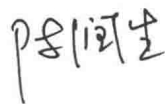
序

生物大分子的计算机模拟是伴随着生命科学和信息科学技术的发展而产生的一门新学科，它是由生物学、物理学、化学、计算机科学等多学科组成的交叉学科。其基本思想是从体系内相互作用出发，借助于计算机模拟，用来研究生物大分子体系的结构和动力学性质。该学科在后基因组时代蛋白质科学的理论和应用研究中将发挥越来越重要的作用。

该书的主编王存新教授是国内最早从事蛋白质计算模拟的研究成员之一。20 世纪 80 年代初，他在中国科技大学与施蕴渝院士一起开始涉足该领域，1998 年作为学术带头人调入北京工业大学，参与筹建该校生命科学与生物工程学院，继续从事分子设计方面的研究，结合所承担的国家科研课题，采用理论与应用相结合的研究方法，将基础研究和应用研究紧密结合，获得了一系列研究成果。

该书全面、系统地阐述了蛋白质模拟的基本原理、最新方法和应用，内容翔实、资料丰富、特点鲜明。全书包括 6 部分共 20 章。第一部分涉及分子动力学模拟的原理、方法和应用，包括分子动力学模拟方法在膜蛋白体系、蛋白质与 DNA 相互作用体系中的应用，以及拉伸分子动力学模拟方法及其应用等。第二部分主要是结合作者参加国际上用分子对接方法预测蛋白质复合物结构竞赛（CAPRI）的实例，阐明蛋白质复合物结构预测的热点、难点问题及其解决途径。第三部分是用分子模拟方法研究蛋白质折叠，主要涉及蛋白质折叠的最新研究进展，如何用复杂网络方法和相对熵方法研究蛋白质折叠问题，并结合实例分析了蛋白质折叠路径和折叠核的预测问题。第四部分是关于粗粒化模型，主要探讨 Gō 模型和弹性网络模型在蛋白质折叠及结构-功能关系研究中的应用。第五部分是如何处理分子模拟中长程静电相互作用的计算瓶颈问题，包括长程静电相互作用的研究进展，以及处理长程静电相互作用的主要理论模型、方法及其在研究蛋白质相互作用中的应用。第六部分主要涉及分子模拟方法在药物设计中的应用，包括正向和反向虚拟筛选方法、抗体药物设计、耐药性机理及高通量药物筛选技术等。

该书的作者及其研究团队总结了 30 多年来从事蛋白质模拟的丰富经验，该书是他们多年研究成果的总结，学术思想扎实，紧密联系当前该领域发展的前沿科学问题，创新性强，具有较高的可读性和参考价值。望该书早日出版，以飨读者。



中国科学院院士

2015 年 9 月于北京

前 言

生物大分子的计算机模拟是伴随着生命科学和信息科学技术的发展而产生的一门新学科，它是一门由生物学、物理学、化学、计算机科学等多学科组成的交叉学科。其基本思想是从体系内原子间的相互作用出发，借助于计算机模拟，用来研究生物分子的结构和动力学性质。2013 年诺贝尔化学奖授予马丁·卡普拉斯（Martin Karplus）、迈克尔·莱维特（Michael Levitt）和亚利耶·瓦谢尔（Arieh Warshel），以奖励他们在发展复杂化学及生物体系多尺度模型方面所做的贡献。可见生物大分子的计算机模拟已成为一个重要的研究领域，发展越来越快，对后基因组时代蛋白质科学的理论和应用研究具有深远的科学意义及重要的应用价值。在蛋白质结构预测、蛋白质折叠机理、蛋白质与配体的相互作用与识别、药物设计与筛选等方面的研究中将发挥越来越重要的作用。

作者在 20 世纪 80 年代初期开始涉足这一领域，是国内较早从事蛋白质计算模拟的研究人员之一。作者及其课题组针对国家重大需求与国际学术发展趋势，结合承担的国家科研项目，获得了一系列研究成果。经与科学出版社协商，计划结合本领域国内外发展现状，全面系统地总结 30 多年的工作经验，出版一本关于蛋白质模拟的专著，从交叉学科的角度，把握国内外最新研究动向，探讨该领域发展的热点和难点问题，并结合课题组工作实际，给出应用实例，力求为从事该领域的研究人员提供新思想和新方法，以促进该学科的发展。

本书的内容涉及分子模拟领域十分广阔的范畴，各章节的作者都在该研究领域进行过多年的实践，并做出过许多优秀的研究工作，具有丰富的分子模拟研究经验。本书在出版过程中，课题组陈慰祖教授参与了出版方案的讨论和制定、书稿撰写的组织及内容审阅、修改和校对，做了大量认真细致的工作。本书将对蛋白质分子模拟领域的原理、发展和应用进行介绍，重点针对该领域的热点和难点问题，结合实际进行深入的讨论，力求做到理论联系实际、学术思想新颖。本书的内容主要包括以下六部分。

第一部分，分子动力学模拟方法。除了通常的分子动力学模拟原理和方法外，重点讨论分子动力学模拟方法在研究膜蛋白体系、蛋白质与 DNA 相互作用体系中的技术处理方法，以及如何用拉伸分子动力学模拟方法研究蛋白质与配体结合或解离过程的动力学性质。该部分的负责人是胡建平，撰写人有丛肖静、许先进、刘明、孙庭广、胡建平。

第二部分，蛋白质复合物结构预测。主要结合课题组多次参与国际上用分子对接方法预测蛋白质复合物结构竞赛（CAPRI）的实践经历，讲述蛋白质-蛋白质对接方法的关键难点问题，诸如如何充分搜索复合物构象、如何确定结合位点、如何设计正确有效的打分函数等问题。该部分的负责人是李春华，撰写人有李春华、龚新奇、常珊。

第三部分，用分子模拟方法研究蛋白质折叠。蛋白质折叠的机理是分子生物学领域目前尚未解决的、具有挑战性的问题。这部分内容主要涉及蛋白质折叠的最新研究进展、

如何用复杂网络方法和相对熵方法研究蛋白质折叠问题,并结合实际分析了大家十分关心的蛋白质折叠路径和折叠核的预测问题。该部分的负责人是苏计国,撰写人有苏计国、焦雄、齐立省。

第四部分,关于粗粒化模型。该模型是目前国际上处理复杂生物体系常用的理论方法,本书主要结合作者工作实际,探讨弹性网络模型和 Gō 模型在蛋白质结构-功能关系研究中的应用。该部分的负责人和撰写人是苏计国。

第五部分,关于长程静电相互作用。这是分子模拟技术能否获得准确结果的关键问题,也是一个尚未完全解决的难点。作者结合自身工作实际,讲述长程静电相互作用的研究进展、处理长程静电相互作用的主要理论模型和方法,以及该方法在研究蛋白质相互作用中的应用。该部分的负责人是卢本卓,撰写人有彭波、卢本卓。

第六部分,药物分子设计方法与应用。这部分内容虽然已有许多专著详细论述过,但本书主要结合作者的工作实际,突出分子模拟方法在药物设计中的应用,力求具有新意,并重点讲述正向虚拟筛选和反向虚拟筛选方法、抗体药物设计、耐药性机理,以及高通量药物筛选技术和应用。该部分的负责人是谭建军,撰写人有谭建军、孔韧、许先进、刘明、张小轶、何红秋、刘斌。

本书不仅适合于从事计算生物学、蛋白质分子模拟和分子设计的专业技术人员,而且可供刚开始接触生物分子模拟的人员学习参考,可提供给高等学校及科研院所的教师、研究人员和研究生参考,也可选为分子模拟和生物信息学、系统生物学等课程的指定教材和参考书。

本书在出版过程中得到科学出版社罗静编辑的大力支持,并得到北京工业大学研究生院和北京工业大学生命科学与生物工程学院的资助,课题组在完成本书涉及的有关研究工作中,得到国家自然科学基金项目、科技部国际合作项目及北京市基金项目的支持,在此一并表示衷心感谢。

王存新

2015年8月于北京工业大学

目 录

第一部分 分子动力学模拟方法

第 1 章 分子动力学模拟的原理、方法与进展	3
1.1 引言	3
1.2 生物大分子的经典力学模型与常见力场	4
1.2.1 经典力学模型	4
1.2.2 常见分子力场	5
1.3 积分算法	7
1.3.1 体系的动力学方程	7
1.3.2 动力学方程的数值解法	7
1.4 周期性边界条件	9
1.5 约束条件动力学模拟	10
1.5.1 SHAKE 算法	10
1.5.2 LINCS 算法	12
1.6 非键相互作用	12
1.6.1 短程相互作用	13
1.6.2 MD 模拟中长程静电相互作用的常用算法	13
1.7 恒温恒压分子动力学模拟	17
1.7.1 温度控制方法	17
1.7.2 压力控制方法	21
1.8 溶剂模型	23
1.8.1 隐含溶剂模型	23
1.8.2 显含溶剂模型	24
1.9 分子动力学模拟的主要步骤	24
1.10 蛋白质分子动力学模拟的进展与前景	31
参考文献	32
第 2 章 拉伸分子动力学模拟	37
2.1 引言	37
2.2 拉伸分子动力学模拟方法	38

2.3 拉伸分子动力学模拟实例	39
2.3.1 周质结合蛋白与配体相互识别研究	39
2.3.2 抗癌多肽 p28 与肿瘤抑制蛋白 p53 的拉伸分子动力学研究	46
参考文献	52
第 3 章 膜蛋白体系的分子动力学模拟	56
3.1 引言	56
3.2 膜蛋白分子动力学模拟研究进展	57
3.2.1 膜的性质和脂质分子的类型	57
3.2.2 膜蛋白的性质和类型	57
3.2.3 分子动力学模拟在膜研究方面的应用	58
3.2.4 分子动力学模拟在膜蛋白体系的应用	59
3.3 膜蛋白体系分子动力学模拟的基本方法和步骤	60
3.4 BtuC-POPC 膜蛋白体系的分子动力学模拟	62
3.4.1 研究背景	62
3.4.2 材料和方法	62
3.4.3 结果和讨论	64
3.4.4 总结和展望	71
参考文献	71
第 4 章 蛋白质与 DNA 相互作用的分子动力学模拟	76
4.1 引言	76
4.2 蛋白质与 DNA 识别的结构特征	76
4.2.1 DNA 结合蛋白的结构特征	76
4.2.2 蛋白质-DNA 复合物的作用位点特征	78
4.3 蛋白质-DNA 识别的研究方法	79
4.3.1 蛋白质与 DNA 的相互作用模式	79
4.3.2 蛋白质-DNA 相互作用的实验方法	81
4.3.3 蛋白质-DNA 识别研究的分子模拟方法	82
4.4 HIV-1 整合酶与病毒 DNA 识别的分子动力学模拟	84
4.4.1 研究背景及意义	84
4.4.2 体系和方法	85
4.4.3 结果和讨论	87
4.5 小结	103
参考文献	104

第二部分 蛋白质复合物结构预测

第 5 章 用分子对接方法预测蛋白质复合物结构	111
5.1 引言	111
5.2 蛋白质-蛋白质分子对接方法	112
5.2.1 分子对接的基本原理	112
5.2.2 分子对接的关键步骤	114
5.3 分子对接方法的研究现状	116
5.3.1 分子对接方法的分类	116
5.3.2 几种重要的分子对接方法	117
5.3.3 国际 CAPRI 蛋白质复合物结构预测简介	124
5.4 难点和亟待解决的问题	126
参考文献	127
第 6 章 蛋白质结合位点预测	131
6.1 引言	131
6.2 蛋白质结合位点的分类	132
6.2.1 结合位点上的热点残基	132
6.2.2 锚残基结合位点	132
6.2.3 模块结合位点	133
6.3 常见的蛋白质结合位点预测方法	133
6.3.1 基于序列的预测方法	133
6.3.2 基于结构的预测方法	133
6.3.3 基于理化性质的预测方法	134
6.4 蛋白质结合位点预测实例	134
6.4.1 基于主链氢键包埋的预测方法	134
6.4.2 基于蛋白质表面氨基酸模块内部接触和外部暴露的预测方法 (PAMA)	139
6.5 展望	145
参考文献	145
第 7 章 蛋白质分子对接打分函数设计	150
7.1 引言	150
7.2 经典打分参量	150
7.2.1 几何互补项	150
7.2.2 界面接触面积	151