

分子模拟的

理论与实践

陈正隆 徐为人 汤立达 编著

分子模拟的概况

力场

能量最小化

分子力学的应用

分子动力学模拟计算的原理

特殊形态的分子动力学计算

分子动力学计算的应用及实例

长程作用力的计算

蒙地卡罗计算方法

蒙地卡罗计算的应用及实例

布朗动力学计算方法

耗散粒子动力学计算方法

溶解自由能

生化分子系统的模拟

分子模拟与药物设计



化学工业出版社

生物·医药出版分社



分子模拟的

理论与实践

陈正隆 徐为人 汤立达 编著



化学工业出版社
生物·医药出版分社

·北京·

分子模拟技术是一项广泛应用于材料、化学、生化与药物研究的通用技术。通过计算模拟，可以得出实验无法得到的资料，因此该技术应用领域十分广泛。但由于其公式较多，一直以来都令人感觉深奥难懂。本书在写作过程中略去公式的繁琐推导，力图结合丰富的应用实例使读者尽快掌握各种模拟方法的精髓。书中具体介绍了分子力学方法、分子动力模拟、蒙地卡罗计算及布朗动力模拟等方法。图文并茂，实例生动。

本书适用于从事材料、生物化学以及药学相关专业的研究人员，以及相关专业的学生使用。

图书在版编目 (CIP) 数据

分子模拟的理论与实践/陈正隆，徐为人，汤立达编
著. —北京：化学工业出版社，2007.7
ISBN 978-7-122-00439-0

I. 分… II. ①陈…②徐…③汤… III. 计算机模拟-应用-分子物理学-研究 IV. 0561

中国版本图书馆 CIP 数据核字 (2007) 第 070525 号

责任编辑：陈燕杰 孙小芳 文字编辑：颜克俭
责任校对：吴 静 装帧设计：关 飞

出版发行：化学工业出版社 生物·医药出版分社
(北京市东城区青年湖南街 13 号 邮政编码 100011)
印 刷：北京云浩印刷有限责任公司
装 订：三河市万龙印装有限公司
720mm×1000mm 1/16 印张 22½ 字数 384 千字 2007 年 8 月北京第 1 版第 1 次印刷

购书咨询：010-64518888 (传真：010-64519686) 售后服务：010-64518899
网 址：<http://www.cip.com.cn>
凡购买本书，如有缺损质量问题，本社销售中心负责调换。

定 价：49.00 元

版权所有 违者必究

前 言

本书旨在介绍分子模拟技巧的基本原理与其在实际体系中的应用。分子模拟的方法通常包括以量子力学与经典力学为基础的两部分。为了适应各种领域的读者，本书只介绍以经典力学为基础的模拟方法，并强调这些方法的应用。分子模拟方法已发展数十年，适用的范围愈来愈广、精确度愈来愈高，已成为材料、溶液、表面、生化、药物等方面研究不可缺少的工具。

目前已有一些关于这些方法的外文著作，但大部分艰深难懂，或是局限于某些特殊的领域而不适合分子模拟技术的推广和普及。本书省略各种繁琐公式的推导，尽量以简单易懂的方式说明各种模拟所需的原理。书中以实际的计算例子显示这些原理的应用，以加强读者对应用的认识。本书介绍的模拟原理包括各种力场 (force field)、分子力学方法 (molecular mechanics, MM)、分子动力模拟 (molecular dynamics simulation, MD)、蒙地卡罗模拟 (Monte Carlo simulation, MC)、布朗动力模拟 (Brownian dynamics simulation, BD)、耗散粒子动力模拟 (dissipative particle dynamics simulation, DPD) 等。这些是目前应用最广、最普遍的模拟方法。书中介绍的应用领域包括：溶液、自组装纳米级材料、光电材料、非晶相高分子材料、离子型表面活性剂、生化体系、药物系统等。书中引用分子模拟方法在这些领域的实际工作，其中许多应用例子为编著者的研究成果且为研究中作者实际遇到的各种问题。各种模拟实例均附有清楚的图形或表格，便于读者了解。

分子模拟可应用的领域十分宽广，发展十分迅速，各文献中有丰富的应用实例。由于篇幅的限制，书中不能一一介绍，但读者可由书中的相关章节了解模拟方法的基本原理与类似的应用实例。希望经过本书的介绍，读者能很快掌握各种模拟方法的精髓，并应用于自身的研究中，做出更多、更深入的研究成果。

李正隆

2007年4月

目 录

第 1 章 分子模拟的概况

参考文献	5
------------	---

第 2 章 力场

2.1 力场简述	6
2.2 力场作用项的一般式	9
2.3 常见的力场	13
2.3.1 MM 形态力场	13
2.3.2 AMBER 力场	14
2.3.3 CHARMM 力场	14
2.3.4 CVFF 力场	15
2.3.5 第二代力场	16
2.4 涵盖周期表元素的力场	17
2.5 特殊的力场	18
参考文献	21

第 3 章 能量最小化

3.1 势能图与势能面	24
3.2 函数的极小值	26
3.3 利用导数求极小值的方法	28
3.4 一次导数求极值法	29
3.4.1 最速下降法	30
3.4.2 共轭梯度法	32
3.5 二次导数求极值法	33

3.5.1	牛顿-拉森法	33
3.5.2	近似牛顿-拉森法	34
3.5.3	对角线块状牛顿-拉森法	35
3.6	求分子最低能量构象方法的选择	35
3.7	能量最小化的应用	36
3.8	多原子分子振动频率的计算	37
3.9	以能量最小化方法预测团簇体的大小	42
	参考文献	48

第4章 分子力学的应用

4.1	分子结构与势能面	49
4.2	势能面的最小值点、最大值点、鞍点	51
4.3	势能面与化学反应	52
4.3.1	过渡态结构与反应路径	52
4.3.2	鞍点与二次区域	54
4.3.3	决定鞍点的方法	56
4.4	分子力学计算的热力学性质	58
4.4.1	构象能	58
4.4.2	分子的生成焓	58
4.4.3	反应热与生成焓	61
4.5	晶体结构的计算	62
4.6	立体效应	64
	参考文献	65

第5章 分子动力模拟计算的原理

5.1	分子动力计算的基本原理	67
5.2	牛顿运动方程式的数值解法	69
5.3	周期性边界条件与最近镜像	72
5.4	积分步程的选取	75
5.5	简化单位	76
5.6	分子动力计算流程	77
5.7	分子动力计算的初始设定	79
5.8	定温算法	81
5.9	系统形状的改变计算	83
5.10	各种系综的分子动力计算方法	84

参考文献	88
------------	----

第 6 章 特殊形态的分子动力计算

6.1 刚性线形分子的计算	90
6.2 刚性非线性分子的计算	92
6.3 键长限制法	95
6.4 限制键长与键角的计算	97
6.5 聚合物的限制动力计算	99
参考文献	103

第 7 章 分子动力计算的应用

7.1 运动轨迹分析	104
7.2 热力学特性的计算	107
7.3 径向分布函数	110
7.4 与时间相关物理量的计算	112
7.4.1 均方位移	112
7.4.2 相关函数	113
7.5 自由能的计算	121
7.5.1 自由能的变化	121
7.5.2 微扰计算	123
7.5.3 耦合系数	124
7.5.4 热力学积分方法	125
7.5.5 缓慢成长法	127
7.6 相对自由能的变化	128
7.7 构型自由能	129
参考文献	132

第 8 章 分子动力计算的应用实例

8.1 水分子系统的计算	133
8.1.1 简单的水分子的力场	133
8.1.2 ab initio 水分子力场	136
8.1.3 水分子自由能的计算	137
8.2 正戊烷构象的研究	139
8.3 链状分子的计算	142
8.4 链状双性体的分子动力计算	144

8.5 聚合物的特殊运动	147
参考文献	151

第9章 长程作用力的计算

9.1 中央多极展开法	152
9.2 Ewald 求和法	156
9.3 作用场与镜像电荷算法	159
9.4 非键结作用项的格胞多极计算方法	161
9.5 溶剂介电质模型	164
参考文献	166

第10章 蒙地卡罗计算方法

10.1 基本原理	167
10.2 重要取样法	170
10.3 Metropolis 算法	174
10.4 各类系综的蒙地卡罗计算	176
10.4.1 定温定压系综	176
10.4.2 巨正则系综	177
10.5 进阶蒙地卡罗算法	179
10.5.1 优先取样蒙地卡罗算法	180
10.5.2 力场导向蒙地卡罗算法	181
10.5.3 明智的蒙地卡罗算法	182
参考文献	183

第11章 蒙地卡罗计算的应用

11.1 压力与自由能的变化	185
11.2 分子液体	187
11.3 聚合物晶格模式计算	188
11.3.1 晶格模式	188
11.3.2 晶格模式计算的执行步骤	188
11.3.3 研究聚合物的物理量	190
11.3.4 MC 计算的结果	192
11.4 改良型晶格计算	193
11.5 蛇形滑动模式	194
11.6 聚合物的改进计算模式	197

11.6.1 球与棒模式	197
11.6.2 球与弹簧模式	199
11.7 随机搜寻最佳结构方法	200
参考文献	202

第 12 章 蒙地卡罗计算的实例

12.1 液晶的蒙地卡罗计算	203
12.2 两相并存系统的蒙地卡罗计算	208
12.3 沸石等温吸附线的计算	212
12.4 蛋白质的折叠	217
参考文献	220

第 13 章 布朗动力计算方法

13.1 布朗粒子运动力学	221
13.2 布朗动力计算	223
13.3 布朗动力计算的技巧	226
13.4 流动效应	227
13.5 布朗动力计算在生化系统的应用	228
13.6 布朗动力计算在分子溶液的应用	231
13.7 分子动力计算与布朗动力计算的比较	233
13.8 朗之万方程式的一般解	237
参考文献	238

第 14 章 耗散粒子动力计算法

14.1 耗散粒子动力计算	240
14.2 离子型表面活性剂胶束的计算	242
14.3 非离子型表面活性剂的相变化研究	246
14.4 脂质分子的相行为计算	249
14.5 纳米尺寸金粒子的 DPD 模拟	252
参考文献	255

第 15 章 溶解自由能

15.1 Born 与 Onsager 模型	256
15.2 连续模型	258

15.3	朗之万偶极矩模型	261
15.4	泊松-玻耳兹曼方程式	262
	参考文献	266

第 16 章 生化分子系统的模拟

16.1	生化分子的结构	267
16.2	生化分子体系间的亲水性与疏水性	275
16.3	蛋白质分子折叠的模拟	276
16.4	细胞膜的融合模拟	279
16.5	疏水基团的堆积	284
	参考文献	287

第 17 章 分子模拟与药物设计

17.1	小分子的定量构效关系	288
17.1.1	传统的构效关系	288
17.1.2	三维定量构效关系	289
17.1.3	构效关系的数学基础	291
17.2	药效团模型	297
17.2.1	药效团模型的建立	298
17.2.2	药效团识别的软件	299
17.3	分子对接计算	300
17.3.1	FlexX	300
17.3.2	LigandFit	301
17.3.3	Autodock 3.05	302
17.3.4	DOCK 5.2	303
17.3.5	分子对接计算的注意点	303
17.4	数据库搜索	305
17.5	从头设计	306
17.6	计算实例	307
17.6.1	3D-QSAR 研究实例	307
17.6.2	对接计算和分子设计实例	309
	参考文献	316

附录 A 经典运动力学

A-1	牛顿运动力学	317
-----	--------------	-----

A-2	拉格朗日运动力学	319
A-3	哈密顿运动力学	322

附录 B 统计力学

B-1	基本统计原理	324
B-2	系综	326
B-3	配分函数	326
B-4	统计涨落	328
B-5	统计力学与模拟计算的关系	329
B-6	统计力学与量子力学的关系	330
B-7	经典力学与统计力学的关系	331
	参考文献	332

附录 C 天然存在的氨基酸的结构

索引

第1章 分子模拟的概况

自20世纪量子力学 (quantum mechanics) 的快速发展后, 几乎有关分子的一切性质, 如结构 (structure)、构象 (conformation)、偶极矩 (dipole moment)、电离能 (ionization potential)、电子亲和力 (electron affinity)、电子密度 (electron density) 等, 皆可由量子力学计算获得。计算与实验的结果往往相当吻合, 并且可由分析计算的结果得到一些实验无法获得的资料, 有助于对实际问题的了解。

与实验相比较, 利用计算机计算研究化学有下列几项优点: ①成本降低; ②增加安全性; ③可研究极快速的反应或变化; ④得到较佳的准确度; ⑤增进对问题的了解。基于这些原因, 分子的量子力学计算自1970年后逐渐受到重视。利用计算先行了解分子的特性, 已成为合成化学家和药物设计学家所依赖的重要方法。化学家们借此可设计出最佳的反应途径, 预测合成的可能性, 并评估所欲合成分子的适用性, 节省许多时间和避免材料的浪费。以欧美的许多大型药厂为例, 在采用计算以前, 合成新药的成功率约为17%~20%, 但自1980年后, 由于在合成前先利用计算预测, 其成功率已提高至50%~60%。图1-1为1995年美国化学会数据库 (ACS database) 所做的统计图。图中的纵坐标为引用计算机计算程序报告所占的比例, 横坐标为年份。

由图中可清楚地看出计算受重视的程度逐年增加。除了药物的应用外, 分子计算已广泛地被应用于研究如高分子聚合物等材料、生物等复杂庞大的体系。如今化学的各个领域都将计算视为不可缺少的工具, 而且, 计算效果正随着方法的改良与计算机的发展而快速地提高。

量子力学是以分子中电子的非定域化 (delocalization) 为基础, 一切电子的行为以其波函数 (wave function) 表示。根据海森伯 (Heisenberg) 的测不准原理 (uncertainty principle), 量子力学仅能计算区间内电

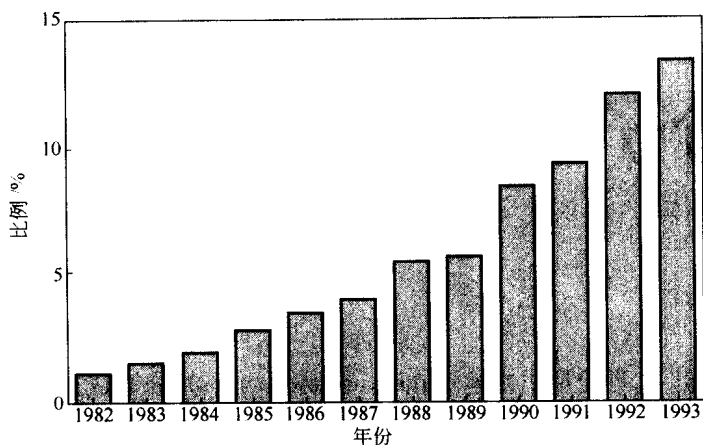


图 1-1 美国化学会所发表的计算机计算在化学报告的比例

子出现的概率，其概率正比于波函数绝对值的平方。欲得到电子的波函数，则需解薛定谔方程式 (Schrödinger equation)，即：

$$\hat{H}\Psi = E\Psi \quad (1-1)$$

式中， \hat{H} 为薛定谔算子，是一些数学的指令； Ψ 为电子的波函数； E 为能量。由于原子与分子含有许多电子（原子序愈大，电子数愈多），解此方程式并非易事。最为普遍的量子力学计算方法为从头算计算方法 (ab initio method)。这种分子轨道 (molecular orbital) 算法，利用变分原理 (variation principle)，将系统电子的波函数展开为原子轨道波函数的组合，而原子轨道的波函数又为一些特定数学函数（如高斯函数）的组合。这种计算方法虽然精确，却甚为缓慢，所能计算的系统亦极为有限，通常不超过 100 个原子。为了增进量子力学计算法的效益，自 1960 年起，人们陆续发展出一些较为简便的量子计算方法。这些方法多引用一些实验值为参数，以取代计算真正的积分项。这样的计算称为半经验分子轨道算法 (semi-empirical molecular orbital method)。半经验分子轨道计算方法有许多类型，各有其优点及缺点，使用时需视实际情况而定。目前常用的半经验分子轨道计算方法有 CNDO、INDO、MINDO、MNDO、ZINDO、AM1 等。利用半经验分子轨道方法可计算较大的分子，其计算结果往往与精确的量子力学方法一致。但即使利用半经验方法和最先进的计算机技术（如平行化处理、计算机团簇），目前所能解的量子系统实际上不

超过 1000 个电子。今日，科学家们一方面致力于改进量子计算的方法及增进其精准度，如 1999 年诺贝尔物理学奖所颁给的密度泛函数理论 (density functional theory, DFT)，即为非常精确的量子计算方法；另一方面，科学家们积极研究提升计算机的计算速度，以期计算较多电子的分子系统。目前已发展的商业软件如 GAUSSIAN 2000 包含各种量子力学计算方法，使用非常方便，可用以计算许多分子的特性。

由上可知，量子力学的方法适用于简单的分子或电子数量较少的体系。但自然界与工业上的许多系统，譬如生化分子（蛋白质、酶、核酸、多糖类……），聚合物（橡胶、安全玻璃、脂肪、油类分子……）等均含有大量的原子及电子。此外，如金属材料、聚合物材料、浓稠溶液、固态混合物、纳米材料等系统，不但需要了解单一分子的性质及了解分子间的交互作用，最重要的是了解整个系统的各种集合性质、动态行为与热力学的性质。像这样复杂的体系，因其电子数目过多，迄今仍不可能完全仰赖量子力学计算。为了解决这样庞大体系的问题，科学家们从 1960 年左右开始着手研究各种可行的非量子计算方法，迄今成果丰富。利用这些非量子力学的计算方法可以解决许多复杂系统的问题，并往往得到相当精准的结果。本书的主旨在于介绍这些以经典力学为主的非量子计算方法的原理及应用，希望能给初学者建立良好的基础。以下先简单介绍这些方法的一般基本原理。

分子力学方法 (molecular mechanics, MM) 起源于 1970 年左右，是依据经典力学 (classical mechanics) 的计算方法。此种方法主要依据分子的力场 (force field) 计算分子的各种特性。依照玻恩-奥本海默近似 (Born-Oppenheimer approximation) 原理，计算中将电子的运动忽略，而将系统的能量视为原子核位置的函数。分子的力场含有许多参数，这些参数可由量子力学计算或实验方法得到。利用分子力学方法可计算庞大与复杂分子的稳定构象、热力学特性及振动光谱等资料。与量子力学相较，此方法简便得多；而且，往往可快速得到分子的各种性质。某些情形下，由分子力学方法所得的结果几乎与高阶量子力学方法所得的结果一致，但其所需要的计算时间却远小于量子力学的计算。分子力学方法常被用于药物、团簇体、生化大分子的研究。

最早对庞大系统采用的非量子计算方法为蒙地卡罗算法 (Monte Carlo method)，简称为 MC 算法。蒙地卡罗算法借由系统中质点 (原子或分子) 的随机运动，结合统计力学的概率分配原理，以得到体系的统计及热力学资料。此计算方法多用以研究复杂体系及金属的结构及其

相变化性质。蒙地卡罗计算的弱点在于只能计算统计的平均值，无法得到系统的动态信息。此计算所依据的随机运动并不适于物理学的运动原理，并且，与其他的非量子计算方法相较亦非特别经济快速，因此，自分子动力学计算逐渐盛行后，蒙地卡罗计算方法已较少为人所采用。

分子动力模拟 (molecular dynamics simulation, MD)，是时下最广泛为人采用的计算庞大复杂系统的方法。自 1970 年起，由于分子力学的发展迅速，人们又系统地建立了许多适用于生化分子体系、聚合物、金属与非金属材料的力场，使得计算复杂体系的结构与一些热力学与光谱性质的能力及精准性大为提升。分子动力模拟是应用这些力场及根据牛顿运动力学原理所发展的计算方法。与蒙地卡罗计算方法比较，分子动力模拟中系统中粒子的运动有正确的物理依据。此方法的优点为精准性高，可同时获得系统的动态与热力学统计资料，并可广泛地适用于各种系统及各类特性的探讨。分子动力模拟的计算技巧经过许多改进，现已日趋成熟，由于其计算能力强，能满足各类问题的需求，许多适用方便的分子动力模拟商业化计算软件也已问世。在先进国家的学校、工厂、医院等的实验室里，这些商业化的计算软件已成为不可或缺的重要研究工具。但是，分子动力模拟方法本身亦有一定的限制。由于此计算需要引用数理积分方法，因此仅能研究系统短时间范围内的运动，而无法模拟一些时间较长（如蛋白质的折叠）的运动问题。现今科学家正努力提升计算机的计算能力及改进计算方法，以期能利用这种计算方法研究更长时间范围内的运动。

另有一种与分子动力模拟类似的计算方法为布朗动力模拟 (Brownian dynamics simulation, BD)。布朗动力模拟适用于大分子的溶液系统，计算中，将大分子的运动分为依力场作用的运动与来自溶剂分子的随机力作用。利用解布朗运动方程式可得到大分子运动的轨迹及一些统计与热力学的性质。布朗动力模拟通常适用于计算生化分子（如多肽、蛋白质、DNA…）的水溶液。此计算的优点在于能计算大分子于较长时间范围内（纳秒级）的运动，其缺点则为将溶剂分子的运动视为布朗运动粒子的假设未必正确。

如今从事计算化学或物理的科学家们，正致力于将量子计算与各种经典力学的模拟计算方法结合的研究，以期能将其精准性提高、应用范围增广。另外，发展高速计算能力的工作站或应用个人计算机集群 (PC cluster) 提升计算速度与资料储存量，也是增进计算能力的方法。可预期的是，随着计算能力的不断提升、计算体系的不断扩大，计算定将成为主要的研究方法之一。

本书将介绍分子力学方法、分子动力模拟、蒙地卡罗计算及布朗动力模拟等方法。书中介绍这些方法的原理与各种运用实例，希望帮助对化学计算有兴趣的读者了解各种计算方法与应用范畴。

参 考 文 献

- 1 Lawley K P. "AB Initio Methods in Quantum Chemistry", Part I, II. John Wiley and Sons, New York, USA, 1987
- 2 Chong D P. "Recent Advances in Density Functional Methods", Part I, II. World Scientific, New Jersey, USA, 1995
- 3 Segal G A. "Electronic Structure Calculation", Part A, B. Plenum Press, New York, USA, 1977
- 4 Burkert U, Allinger N L. "Molecular Mechanics", ACS Monograph 177, USA, 1989
- 5 Leach A R. "Molecular Modeling", Addison Wesley Longman Limited, England, 1996
- 6 Allen M P, Tildesley D J. "Computer Simulation of Liquids", Clarendon Press, Oxford, England, 1987
- 7 Frenkel D, Smit B. "Understanding Molecular Simulation", Academic Press, San Diego, USA, 1996

第2章 力场

力场可以看作是势能面的经验表达式，是后面介绍的分子力学、分子动力学、蒙地卡罗模拟等分子模拟内容的基础。针对特定的目的，力场分为许多不同的形式，具有不同的适用范围与局限性，计算的结果可靠性与选用的力场密切相关。本章分两部分内容，2.1~2.2介绍了力场的一般形式及其各种作用项的定义；2.3~2.5介绍了适用有机分子、生化分子、无机配合物、溶液等分子模拟中多种常用的力场。

2.1 力场简述

分子的总能量为动能与势能的和，分子的势能通常可表示为简单的几何坐标的函数。例如，可将双原子分子 AB 的振动势能表示为 A 与 B 间键长的函数，即：

$$U(r) = \frac{1}{2}k(r-r^0)^2 \quad (2-1)$$

式中， k 为弹力常数； r 为键长； r^0 为 AB 的平衡键长。这样以简单数学形式表示的势能函数称为力场 (force field)。经典力学的计算以力场为依据，力场的完备与否决定计算的正确程度。复杂分子的总势能一般可分为各类型势能的和，这些类型包括：

总势能 = 非键结势能 + 键伸缩势能 + 键角弯曲势能 +
二面角扭曲势能 + 离平面振动势能 + 库仑静电势能
势能项习惯上以符号表示为：

$$U = U_{nb} + U_b + U_\theta + U_\phi + U_\chi + U_{el} \quad (2-2)$$

各种类型的势能分别叙述如下。