# 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。

陈正隆

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** 定温计算法 …… 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** 刚性非线形分子的计算 …… 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** 作用场与镜像电荷计算法 …… 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 计算法 …… 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** 进阶蒙地卡罗计算法 …… 179
10.5.1 优先取样蒙地卡罗计算法 …… 180
10.5.2 力场导向蒙地卡罗计算法 …… 181
10.5.3 明智的蒙地卡罗计算法 …… 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** 布朗动力计算在高分子溶液的应用 …… **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❶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

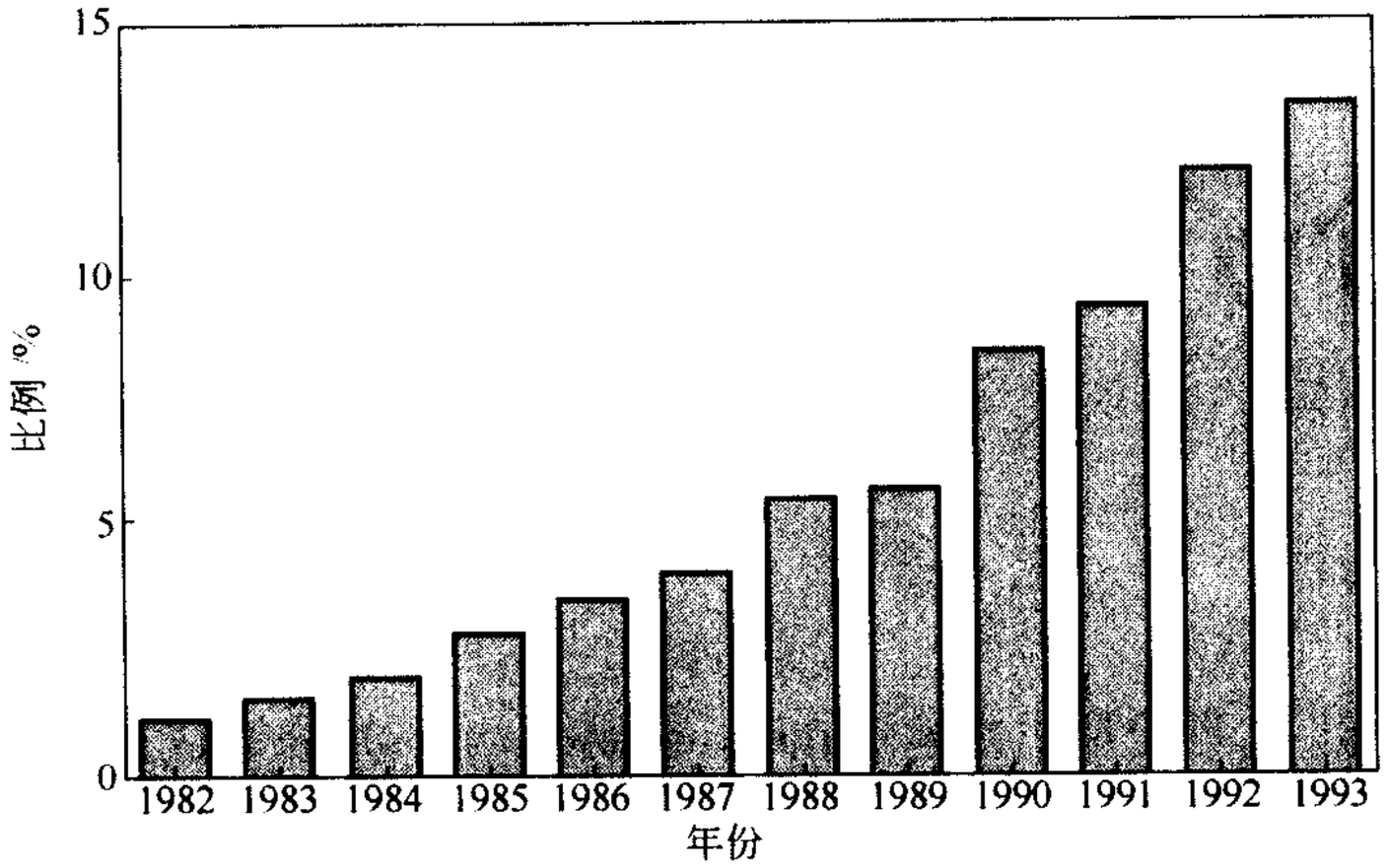



**图 1-1　美国化学会所发表的计算机计算在化学报告的比例**

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：

$$\hat{H}\Psi = E\Psi \tag{1-1}$$

式中，$\hat{H}$ 为薛定谔算子，是一些数学的指令；$\Psi$ 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法（ab initio method）。这种分子轨道（molecular orbital）计算法，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计算法（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，这些参数可经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计算法（Monte Carlo method)，简称为 MC 计算法。蒙地卡罗计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 Limitted, 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❷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、无机配化合物、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：

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

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：

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

各种类型的势能分别叙述如下。

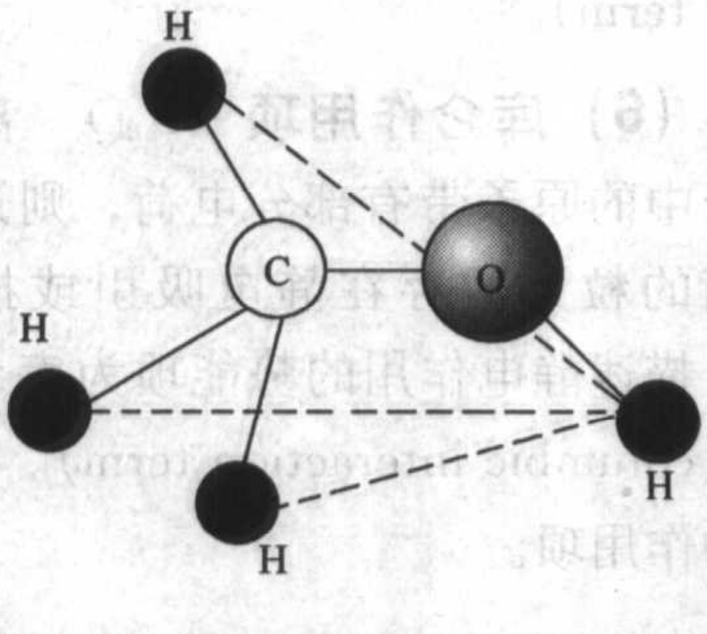



**图 2-1　乙醇分子（$CH_3CH_2OH$）中的非键结作用原子对**

**(1) 非键结势能**（$U_{nb}$）　一般在分子力场中，若 A、B 两原子属于同一分子但其间隔多于两个连接的化学键(如 A—C—C—B)，或两原子分属于两个不同的分子，则原子对间有非键结范德瓦耳斯作用力（van der Waals force)。以乙醇分子 $CH_3CH_2OH$ 为例，图 2-1 的虚线显示乙醇分子中所有非键结（non-bonding）作用原子对。此种作用亦存在于不同分子的原子对之间。

**(2) 键的伸缩项**（$U_b$）　分子中互相键结的原子形成的化学键，如图 2-1 中分子的 C—H 键、C—C 键等。化学键的键长并非维持恒定，而是于其平衡值附近呈小幅度的振荡。描述此种作用的势能项称为键伸缩项（bond stretching term)。

**(3) 键角弯曲项**（$U_\theta$）　分子中连续键结的三原子（如上例乙烷分子的 C—C—C、C—C—H、H—C—H、C—O—H…）形成键角。与键的伸缩一样，这些键角并非维持恒定不变，而是于其平衡值附近呈小幅度的振荡。描述此种作用的势能项为键角弯曲项(angle bending term)。

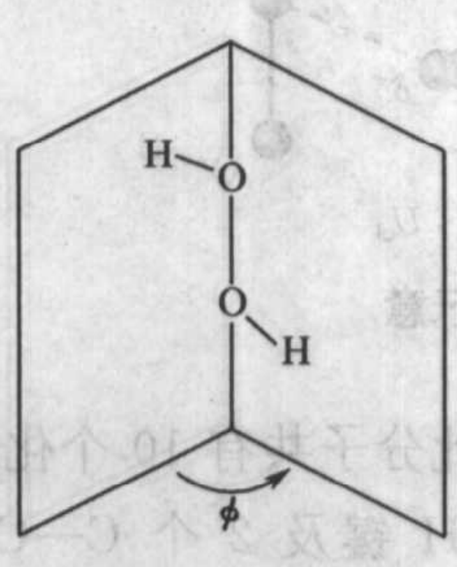



**图 2-2　$H_2O_2$ 分子的二面角扭转项**

**(4) 二面角扭曲项**（$U_\phi$）　分子中连续键结的四个原子，形成二面角（dihedral angle)，如图 2-2 所示 $H_2O_2$ 分子中的二面角。又如丙烷 $CH_3CH_2CH_3$ 中的 H—C—C—C、H—C—C—H 等。一般分子中的二面角较为松软，易于扭动。描述二面角扭转的势能项称为二面角扭曲项（torsion angle term)。

**(5) 离平面振动项**（$U_\chi$）　分子中有些部分的原子有共平面的倾向，如丙酮 $CH_3C(O)CH_3$ 中，碳原子与氧原子的平衡位置位于共同平面，其他如苯环上的碳氢原子、烯类中的碳氢原子等。这些原子有形成共平面的倾向，通常共平面的四个原子的中心原子（如图 2-3 中的中心碳原子）离平面小幅度振动，如图 2-3 中的位移 $\chi$。描述此种振动的势能项

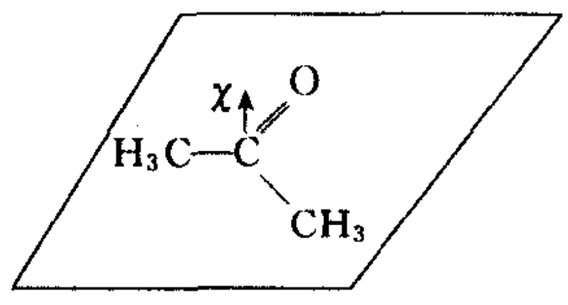



**图 2-3　丙酮分子的离平面振动示意**

称为离平面振动项（out-of-plane bending term）。

**（6）库仑作用项**（$U_{el}$）　离子或分子中的原子带有部分电荷，则这些带电荷的粒子间存在静电吸引或排斥作用。描述静电作用的势能项为库仑作用项（columbic interaction term）。

图 2-4 综合显示分子中势能的各种作用项。

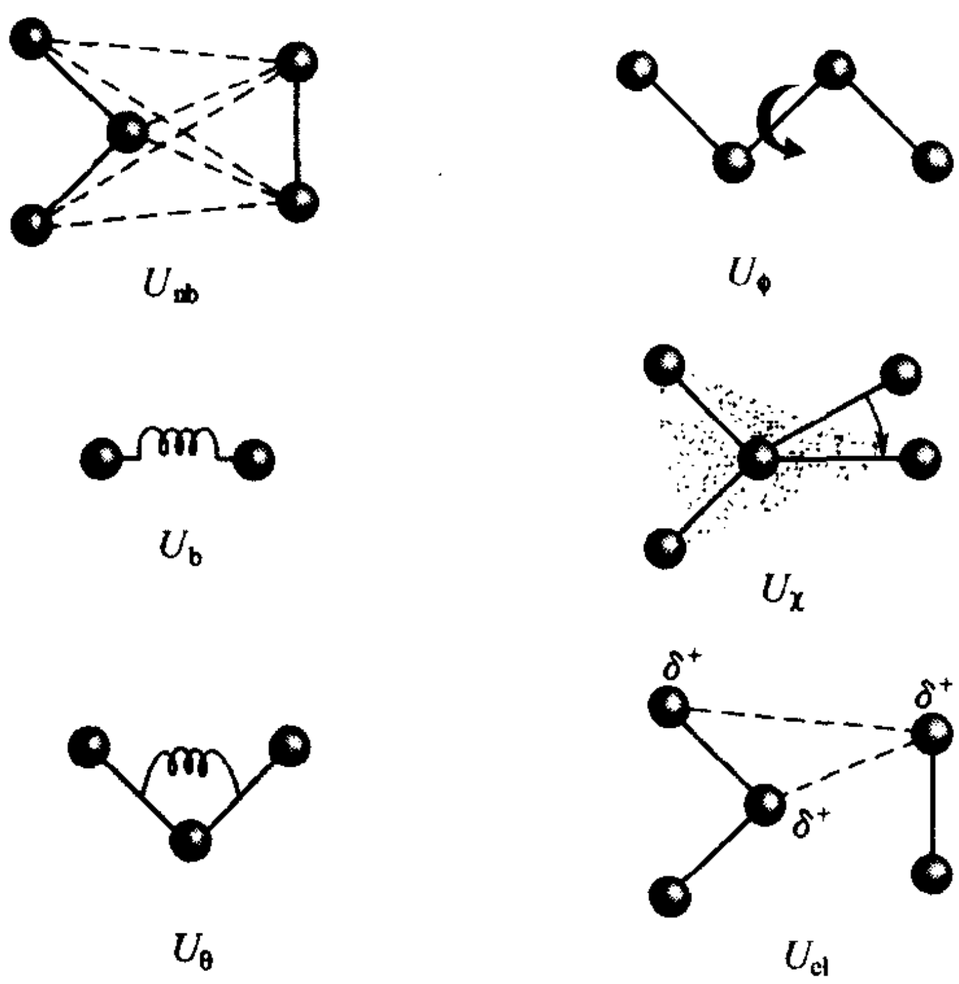



**图 2-4　分子中各种运动的示意**

以丙烷 $CH_3CH_2CH_3$ 为例（图 2-5），此分子共有 10 个化学键、包括 8 个 C—H 键及 2 个 C—C 键。其中，C—C 键为等价键，但 C—H 键则有两类：分子甲基（$CH_3$）的 6 个 C—H 键为一类，而亚甲基（$—CH_2—$）上的 2 个 C—H 键为另外一类。此分子共有 18 个键角，分别为 1 个 C—C—C 键角，10 个 C—C—H 键角、7 个 H—C—H 键角。该分子含有 18 个二面角扭曲项，其中包含 12 个 H—C—C—H 二面角扭曲项与 6 个 H—C—C—C 二面

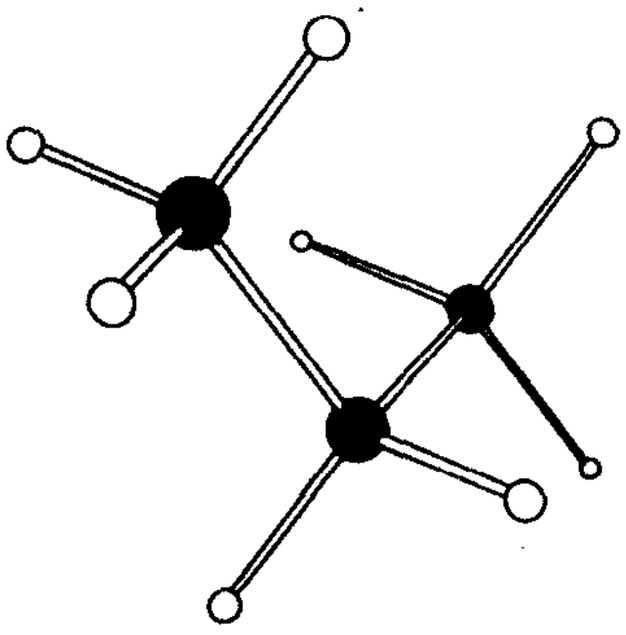


**图 2-5　丙烷分子的三维示意**

角扭曲项。此分子无共平面原子，故无离平面振动项。其非键结作用项共有 27 项，分别为 21 组 H—H 原子对间的作用项与 6 组 H—C 间的作用项。丙烷为中性分子，原子所带的部分电荷很小，可忽略其库仑作用。

## 2.2 力场作用项的一般式

计算非键结作用，通常将原子视为位于其原子核坐标的一点。一般力场中最常见的非键结势能形式为 Lennard-Jones（LJ）势能。此种势能又称为 12-6 势能，其数学式为：

$$U(r)=4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right] \quad (2\text{-}3)$$

式中，$r$ 为原子对间的距离；$\epsilon$ 与 $\sigma$ 为势能参数，因原子的种类而异。图 2-6 为 LJ 势能曲线，其势能的最低点位于 $r=2^{1/6}\sigma$ 处，$\epsilon$ 为由势能最低点至势能为 0 的差。故 $\sigma$ 的大小反映原子间的平衡距离，而 $\epsilon$ 的大小则反映出势能曲线的深度。LJ 势能中，$r^{-12}$ 项为排斥项，$r^{-6}$ 项为吸引项，当 $r$ 很大时 LJ 势能趋近于零，这表示当原子对距离很远时，彼此间已无非键结作用。

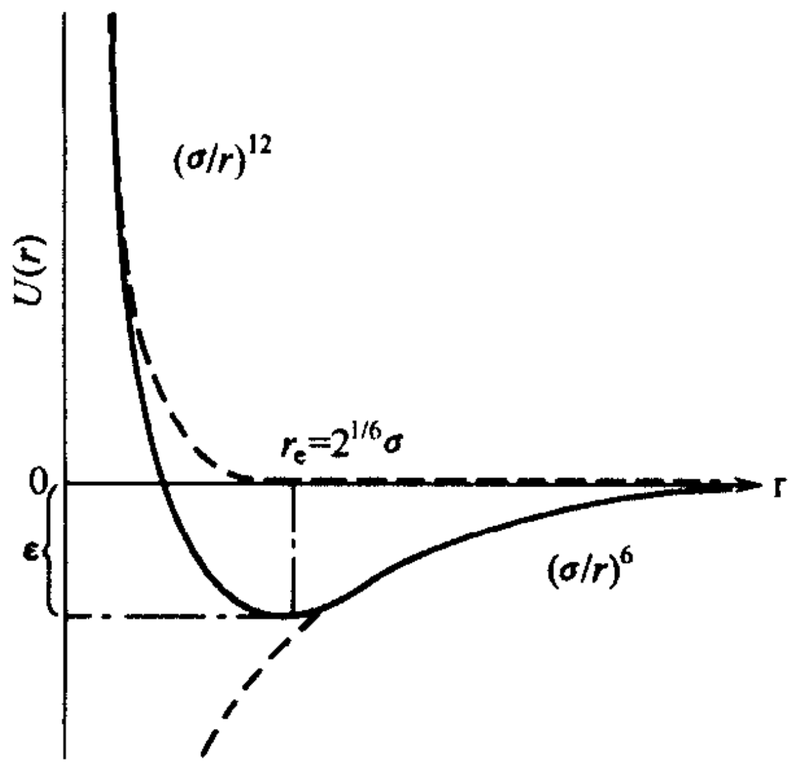



**图 2-6 LJ 势能曲线**

图 2-7 为一些常见分子的 LJ 势能曲线。

表 2-1 列出一些常见的分子的 LJ 势能参数。由此表可估计分子的大小及非键结作用势能的最低值。两种不同原子间的 LJ 作用常数通常以式(2-4) 估计。

$$\sigma_{AB}=\frac{1}{2}(\sigma_A+\sigma_B)$$

$$\epsilon_{AB}=\sqrt{\epsilon_A\epsilon_B} \quad (2\text{-}4)$$

式中，A，B 表示两种不同的原子。

键伸缩势能项之一般形式为简谐振动，即：

$$U_b=\frac{1}{2}\sum_i k_b(r_i-r_i^0)^2 \quad (2\text{-}5)$$

式中，$k_b$ 为键伸缩的弹力常数；$r_i$ 及 $r_i^0$ 分别表示第 $i$ 个键的键长及其平衡键长。弹力常数愈大，振动愈快，振动频率愈大。为了提高计算的

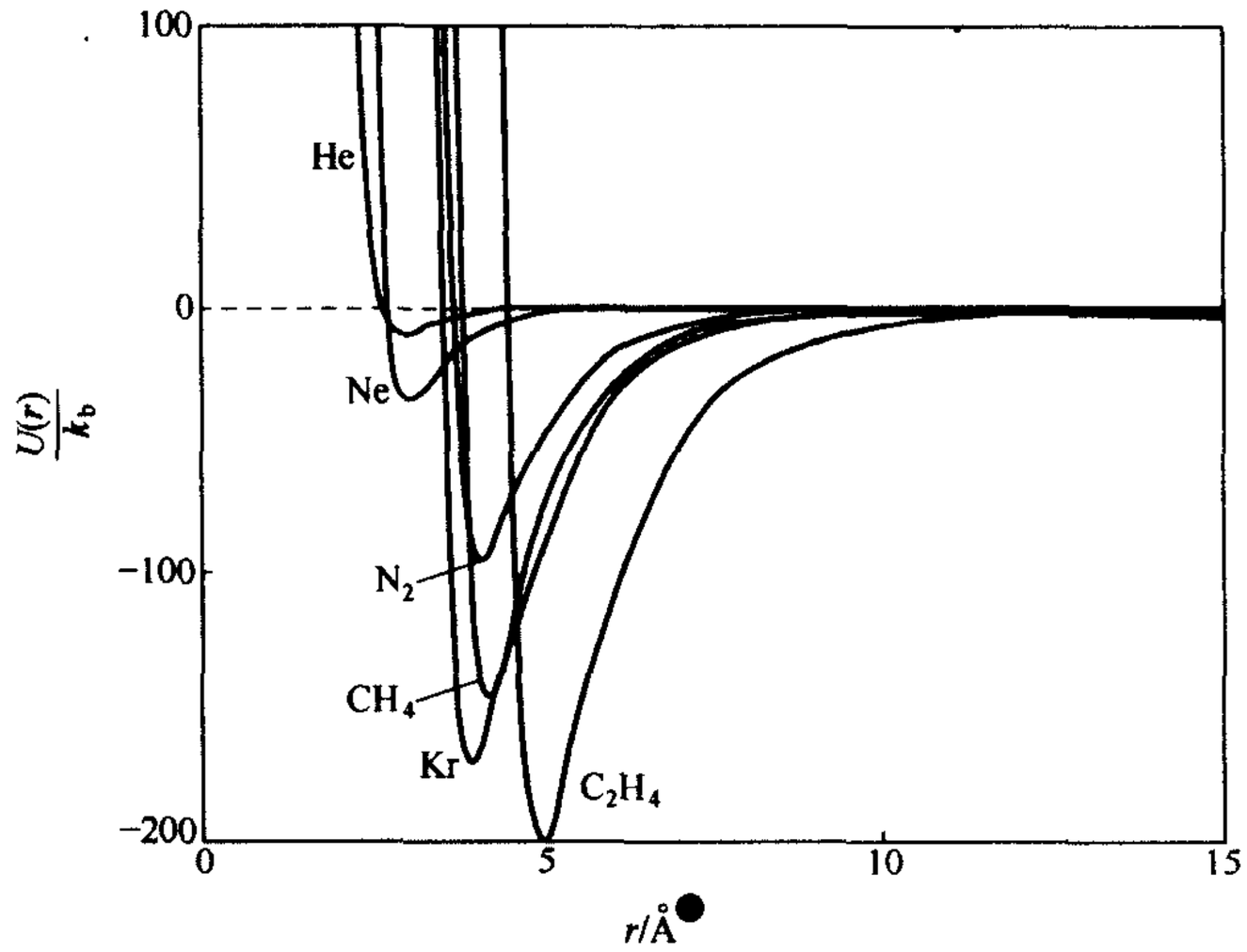



图 2-7　一些常见分子的 LJ 势能曲线

表 2-1　常见分子的 LJ 势能参数

| 原子或分子 | $(\varepsilon/k_b)$/K | r/Å | 原子或分子 | $(\varepsilon/k_b)$/K | r/Å |
|---|---|---|---|---|---|
| He | 10.41 | 2.602 | $SF_6$ | 207.7 | 5.252 |
| Ne | 42.0 | 2.755 | $C_2H_6$ | 241.9 | 4.371 |
| Ar | 141.6 | 3.350 | $C_3H_8$ | 268.5 | 4.992 |
| Kr | 199.8 | 3.581 | $NC_4H_{10}$ | 285.6 | 5.526 |
| Xe | 281.0 | 3.790 | $i$-$C_4H_{10}$ | 260.9 | 5.629 |
| $N_2$ | 104.2 | 3.632 | $C_2H_4$ | 244.3 | 4.070 |
| $O_2$ | 126.3 | 3.382 | $N_2O$ | 266.8 | 3.703 |
| $CO_2$ | 245.3 | 3.762 | $CCl_3F$ | 267.4 | 5.757 |
| $CH_4$ | 161.3 | 3.721 | $CHClF_2$ | 288.3 | 4.647 |
| $CF_4$ | 156.5 | 4.478 | | | |

注：1Å=0.1nm。

精准性，有的力场除了简谐振动项外，亦增加了非简谐振动项。如 MM3 形式的力场，其键伸缩势能的形式为：

$$U_b=\frac{1}{2}k_b(r-r^0)^2+k_3(r-r^0)^3+k_4(r-r^0)^4 \tag{2-6}$$

式中，除了二次的简谐振动项外还包括了三次与四次的非简谐振动项。

键角弯曲项的一般形式为键角的简谐振荡：

$$U_\theta=\frac{1}{2}\sum_i k_\theta(\theta_i-\theta_i^0)^2 \tag{2-7}$$

● 1Å=0.1nm。

式中，$k_\theta$ 为键角弯曲的弹力常数；$\theta_i$ 及 $\theta_i^0$ 分别表示第 $i$ 个键角及其平衡键角的角度。亦可如键伸缩势能一样增加高次非简谐项以提高计算的准度。

二面角扭曲项的一般形式为：

$$U_\phi = \frac{1}{2}\sum_i [V_1(1+\cos\phi) + V_2(1-\cos 2\phi) + V_3(1+\cos 3\phi)] \qquad (2\text{-}8)$$

式中，$V_1$、$V_2$ 及 $V_3$ 为二面角扭曲项的弹力常数，$\phi$ 为二面角的角度。图 2-8 为利用 AMBER 力场计算所得 $OCH_2—CH_2O$ 分子中 O—C—C—O 扭转角的势能图。由图可知最低势能对应于扭转角 $\phi=60°$与 $\phi=300°$。

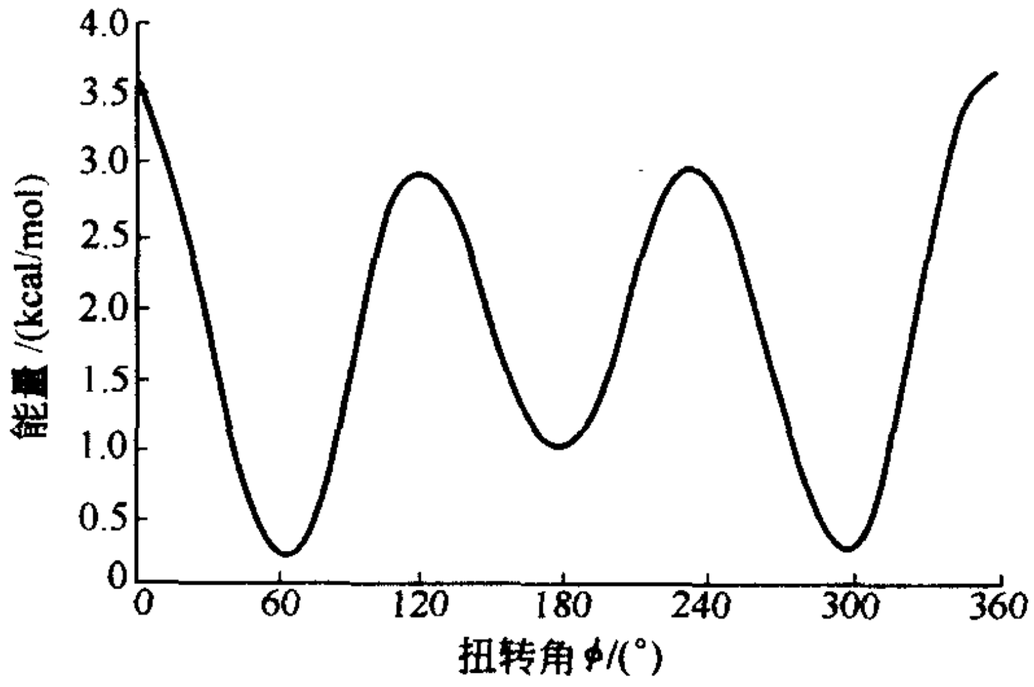



**图 2-8　AMBER 力场计算 $OCH_2—CH_2O$ 分子的扭转势能**

1kcal=4.184kJ

离平面振动项的一般形式为：

$$U_\chi = \frac{1}{2}\sum_i k_\chi \chi^2 \qquad (2\text{-}9)$$

式中，$k_\chi$ 为离平面振动项的弹力常数；$\chi$ 定义为离平面振动的角度。




**图 2-9　分子中偶极作用的坐标**

库仑作用项的形式为

$$U_{el} = \sum_{i,j} \frac{q_i q_j}{D r_{ij}} \qquad (2\text{-}10)$$

式中，$q_i$ 及 $q_j$ 为分子中第 $i$ 个离子与第 $j$ 个离子所带的电荷；$r_{ij}$ 为距离；$D$ 为有效介电常数（effective dielectric constant）。不含离子的分子中，库仑作用项主要为偶极（dipole）间的作用，即：

$$U_{\text{dipole}}=\frac{\mu_i\mu_j}{Dr_{ij}^3}(\cos\chi-3\cos\alpha_i\cos\alpha_j) \tag{2-11}$$

式中，$\mu_i$ 及 $\mu_j$ 为分子中的偶极矩；$\chi$ 及 $\alpha$ 的定义如图 2-9。

另一种在分子力学计算中常见的力场为 Urey-Bradly（UB）所推导的力场，此种力场将未键结作用项写成距离的二次函数。

$$U_{\text{UB}}=\frac{1}{2}\sum_i K_i(\Delta r_i)^2+\frac{1}{2}\sum_i L_i(\Delta\alpha_i)^2+\frac{1}{2}\sum_i M_i(\Delta\rho_i)^2 \tag{2-12}$$

$$+\sum_i k_i\Delta r_i+\sum_i l_i\Delta\alpha_i+\sum_i m_i\Delta\rho_i$$

式中，$\Delta\rho_i$ 为两个未键结原子间的距离；$K$，$L$，$M$，$k$，$l$，$m$ 为力场参数。此力场亦可加入二面角的扭转项及离平面振动项以使其更为精准。

不论力场的形式为何、力场中含有多少参数，最重要的是力场的参数与形式是可以互相转换的。例如，不同的分子若含有同样形式的键结，则这些键的伸缩势能有同样的形式与弹力常数。表 2-2 为 MM2 力场的键伸缩参数，根据此表，凡是分子中含有 $Csp^3$-$Csp^3$ 键（如乙烷与丙烷），其平衡键长均为 1.523Å，其伸缩振动的简谐弹力常数均为 317kcal/(mol·Å²)[1326.3kJ/(mol·Å²)]。

**表 2-2　MM2 力场一些键的伸缩势能参数**［势能形式为 $U_b=\frac{1}{2}k_b(r-r_0)^2$］

| 键　结 | 平衡键长 $r_0$/Å | 弹力常数 $k_b$/[kcal/(mol·Å²)] |
|---|---|---|
| $Csp^3$-$Csp^3$ | 1.523 | 317 |
| $Csp^3$-$Csp^2$ | 1.497 | 317 |
| $Csp^2$—$Csp^2$ | 1.337 | 690 |
| $Csp^3$-H | 1.113 | 662 |
| $Csp^2$—O | 1.208 | 777 |
| $Csp^3$-$Nsp^3$ | 1.438 | 367 |
| C-N(amide) | 1.345 | 719 |

注：1kcal=4.184kJ；1Å=0.1nm。

表 2-3 为 MM2 力场中一些键角弯曲的振动势能参数。同样的，依据此表，凡是分子中含有如 $Csp^3$-$Csp^3$-$Csp^3$ 键角的，其平衡键角值于 MM2 力场中均应为 109.47°。

根据表 2-1、表 2-2 和表 2-3 所列的势能参数可计算简单分子（如水、甲烷……）的任何构型的能量。

表 2-3　一些键的弯曲势能参数［势能形式为 $U_\theta=\frac{1}{2}k_\theta(\theta-\theta_0)^2$］

| 键结 | 平衡键角 $\theta_0$/(°) | 弹力常数 $k_\theta$/[kcal/mol·(°)²] |
|---|---|---|
| $Csp^3$-$Csp^3$-$Csp^3$ | 109.47 | 0.0099 |
| $Csp^3$-$Csp^3$-H | 109.47 | 0.0079 |
| H-$Csp^3$-H | 109.47 | 0.0070 |
| $Csp^3$-$Csp^2$-$Csp^3$ | 117.2 | 0.0099 |
| $Csp^3$-$Csp^2$═$Csp^2$ | 121.4 | 0.0121 |
| $Csp^3$-$Csp^2$═O | 122.5 | 0.0101 |

## 2.3　常见的力场

分子力学与分子动力计算由 20 世纪 70 年代发展起，至今已 30 年。其计算的系统由最初的单原子分子系统延伸至多原子分子、聚合物分子、生化分子系统。计算所使用的力场，也由最简单的范德瓦耳斯作用，随着系统复杂度的增加而增加其复杂性。这些力场多是针对特殊目的所发展的，各有其优点与适用范围。在执行分子力学或分子动力计算时，选择适当的力场极为重要，往往决定计算成果的优劣。本部分内容整理过去与目前较常用的力场，并指出其优点与适用的范围。

### 2.3.1　MM 形态力场

此力场为 Allinger 等所发展的，依其发展的先后顺序分别称为 MM2、MM3、MM4 等。MM 力场将一些常见的原子细分，如将碳原子分为 $sp^3$、$sp^2$、sp、酮基碳、环丙烷碳、碳自由基、碳阳离子等。这些不同形态的碳原子具有不同的力场参数。此力场适用于各种有机化合物、自由基、离子。应用此力场可得到十分精准的构型、构型能、各种热力学性质、振动光谱、晶体能量等。其力场的一般形式为：

$$U=U_{nb}+U_b+U_\theta+U_\phi+U_\chi+U_{el}+U_{cross} \tag{2-13}$$

其各项的形式为

$$U_{nb}(r)=a\varepsilon\cdot e^{-c\sigma/r}-b\varepsilon(\sigma/r)^6$$

$$U_b(r)=\frac{k_b}{2}(r-r_0)^2[1-k'_b(r-r_0)-k''_b(r-r_0)^2-k'''_b(r-r_0)^3]$$

$$U_\theta(\theta)=\frac{k_\theta}{2}(\theta-\theta_0)^2[1-k'_\theta(\theta-\theta_0)-k''_\theta(\theta-\theta_0)^2-k'''_\theta(\theta-\theta_0)^3]$$

$$U_\phi(\phi) = \sum_{n=1}^{3} \frac{V_n}{2}(1+\cos n\phi)$$

$$U_\chi(\chi) = k(1-\cos 2\chi) \tag{2-14}$$

此力场中，$U_{el}$为一般的库仑作用项；$U_{cross}$为交叉作用项。例如与同一原子键结的两个键，其键长分别为 $r_1$ 及 $r_2$。则键与键的交叉作用项为：

$$U(r_1, r_2) = \frac{k_{12}}{2}(r_1 - r_{1,0})(r_2 - r_{2,0}) \tag{2-15}$$

此三个原子构成一个键角，键角为 $\theta$。其键长与键角的交叉作用项为：

$$U(r_1, r_2, \theta) = \frac{k_{12\theta}}{2}[(r_1 - r_{1,0}) + (r_2 - r_{2,0})](\theta - \theta_0) \tag{2-16}$$

此外，二面角所含的键为 $r$，键伸缩与二面角扭曲的交叉作用项为：

$U(r,\phi) = k(r - r_0)\cos n\phi$，或 $U(r,\phi) = k(r - r_0)(1 + \cos n\phi)$ (2-17)

在 MM 形式的力场中仔细考虑了许多交叉作用项，其结果往往优于其他形式的力场。相对的，其力场形式较为复杂，比较不易程序化，计算亦较为费时。

## 2.3.2 AMBER 力场

AMBER (assisted model building with energy minimization) 力场为美国加州大学的 Peter Kollman 等所发展的。此力场主要适用于较小的蛋白质、核酸、多糖等生化分子。应用此力场通常可得到合理的气态分子几何结构、构型能、振动频率与溶剂化自由能 (solvation free energy)。AMBER 力场的参数全来自计算结果与实验值的比对，此力场的标准形式为：

$$U = \sum_b K_2(b - b_0)^2 + \sum_\theta K_\theta(\theta - \theta_0)^2 + \sum_\phi \frac{1}{2}V_0[1 + \cos(n\phi - \phi_0)] +$$

$$\sum \varepsilon\left[\left(\frac{r^*}{r}\right)^{12} - 2\left(\frac{r^*}{r}\right)^{6}\right] + \sum \frac{q_i q_j}{\varepsilon_{ij} r_{ij}} + \sum\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) \tag{2-18}$$

式中，$b$、$\theta$ 与 $\phi$ 为键长、键角与二面角；第 4 项为范德瓦耳斯作用项；第 5 项为静电作用项；第 6 项为氢键的作用项。

## 2.3.3 CHARMM 力场

CHARMM (chemistry at Harvard macromolecular mechanics) 力场为美国哈佛大学所发展的。此力场参数除来自计算结果与实验值的比对外，并引用了大量的量子计算结果为依据。此力场可应用于研究许多分子

系统，包括小的有机分子、溶液、聚合物、生化分子等。除了有机金属分子外，应用此力场通常皆可得到与实验值相近的结构、作用能、构型能、转动能障、振动频率、自由能与许多与时间相关的物理量。CHARMM 力场的形式为：

$$\begin{aligned} U = & U_{\text{bond}} + U_{\text{angle}} + U_{\text{torsion}} + U_{\text{outp}} && \text{(分子内作用项)} \\ & + U_{\text{elec}} + U_{\text{vdw}} && \text{(分子间作用项)} \\ & + U_{\text{constriant}} && \text{(特殊项)} \end{aligned} \tag{2-19}$$

各项详细的形式为：

$$\begin{aligned} U = & \sum k_{\text{b}}(r - r_0)^2 + \sum k_\theta(\theta - \theta_0)^2 + \sum [\,|\,k_\phi\,| - k_\varphi \cos(n\phi)\,] + \\ & \sum k_\chi(\chi - \chi_0)^2 + \sum_{i,j} \frac{q_i q_j}{4\pi\zeta_0 r_{ij}} + \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) \text{sw}(r_{ij}^2, r_{\text{on}}^2, r_{\text{off}}^2) \end{aligned} \tag{2-20}$$

CHARMM 力场计算氢键的能量式为：

$$\begin{aligned} E = & \left(\frac{A}{r_{\text{AD}}^6} - \frac{B}{r_{\text{AD}}^9}\right) \cos^m(\phi_{\text{A-H-D}}) \cos^n(\phi_{\text{AA-A-H}}) \text{sw}(r_{\text{AD}}^2, r_{\text{on}}^2, r_{\text{off}}^2) \times \\ & \text{sw}[\cos^2(\phi_{\text{A-H-D}}), \cos^2(\phi_{\text{on}}), \cos^2(\phi_{\text{off}})] \end{aligned} \tag{2-21}$$

式中，sw 称为开关函数（switching function），用以控制作用的范围，下标“on”与“off”为开始计算与终止计算此函数的键长或角度值。此函数的意义将在第 5 章中介绍。

### 2.3.4 CVFF 力场

CVFF 力场为 Dauber Osguthope 等所发展的，其全名为一致性价力场（consistent valence force field）。此力场最初以生化分子为主，适用于计算氨基酸、水及含各种官能基的分子体系。其后，经过不断的强化，CVFF 力场可适用于计算各种多肽、蛋白质与大量的有机分子。此力场以计算系统的结构与结合能最为准确，亦可提供合理的构型能与振动频率。CVFF 力场的详细形式为：

$$\begin{aligned} U = & \sum_b D_{\text{b}}[1 - \text{e}^{-a(b-b_0)^2}] + \sum_\theta k_\theta(\theta - \theta_0)^2 + \sum_\phi k_\phi[1 + \cos(n\phi)] + \\ & \sum_\chi k_\chi \chi^2 + \sum_b \sum_{b'} k_{\text{bb}'}(b - b_0)(b' - b'_0) + \sum_\theta \sum_{\theta'} k_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta'_0) + \\ & \sum_b \sum_\theta k_{\text{b}\theta}(b - b_0)(\theta - \theta_0) + \sum_\phi \sum_\theta \sum_{\theta'} k_{\phi\theta\theta'} \cos\phi(\theta - \theta_0)(\theta' - \theta'_0) + \\ & \sum_\chi \sum_{\chi'} k_{\chi\chi'} \chi\chi' + \sum \varepsilon\left[\left(\frac{r^*}{r}\right)^{12} - 2\left(\frac{r^*}{r}\right)^{6}\right] + \sum \frac{q_i q_j}{\varepsilon_0 r_{ij}} \end{aligned} \tag{2-22}$$

### 2.3.5 第二代力场

第二代力场的形式远较上述的经典力场复杂，需要大量的力常数。其设计的目的为能精确地计算分子的各种性质、结构、光谱、热力学特性、晶体特性等资料。其力常数的推导除引用大量的实验数据外，并参照精确的量子计算的结果。尤其适用于有机分子或不含过渡金属元素的分子系统。

第二代力场因其参数的不同包括 CFF91、CFF95、PCFF 与 MMFF93 等。其中，前三种称为一致性力场（consistent force field，CFF），其形式为：

$$
\begin{aligned}
U = & \sum_{b}[K_2(b-b_0)^2 + K_3(b-b_0)^3 + K_4(b-b_0)^4] + \\
& \sum_{\phi}\{V_1[1-\cos(\phi-\phi_1^0)] + V_2[1-\cos(2\phi-\phi_2^0)] + V_3[1-\cos(3\phi-\phi_3^0)]\} + \\
& \sum_{b'}\sum_{\phi}(b'-b'_0)(V_1\cos\phi + V_2\cos 2\phi + V_3\cos 3\phi) + \\
& \sum_{\phi}\sum_{\theta}\sum_{\theta'}K_{b\theta\theta'}\cos\phi(\theta-\theta_0)(\theta'-\theta'_0) + \sum_{i>j}\frac{q_i q_j}{\epsilon r_{ij}} + \sum_{i>j}\left(\frac{A_{ij}}{r_{ij}^9} - \frac{B_{ij}}{r_{ij}^6}\right)
\end{aligned}
\tag{2-23}
$$

CFF91 力场适用于研究碳氢化合物、蛋白质、蛋白质-配位基的交互作用。亦可研究小分子的气态结构、振动频率、构型能、扭转能障、晶体结构。CFF91 的力场含有 H、Na、Ca、C、Si、N、P、O、S、F、Cl、Br、I、Ar 等原子的参数，可计算包含这些原子的分子体系。

PCFF 力场由 CFF91 力场衍生而出，适用于计算聚合物及有机物。主要可应用于聚碳酸酯类（polycarbonates）、三聚氰氨树脂（melamine resins）、多糖类（polysaccharides）、其余的聚合物、糖类、脂肪类、核酸、有机物及约 20 种无机物。除了 CFF91 力场的参数外，PCFF 含有 He、Ne、Kr、Xe 等惰性气体原子及 Li、K、Cr、Mo、W、Fe、Ni、Pd、Pt、Cu、Ag、Au、Al、Sn、Pb 等金属原子的力场参数，可计算包含这些原子的分子体系。

CFF95 力场衍生自 CFF91 力场，特殊针对如多糖类、聚碳酸酯等生化分子与有机聚合物所设计的，较适用于生命科学的应用。此力场含有卤素原子及 Li、Na、K、Rb、Cs、Mg、Ca、Fe、Cu、Zn 等金属原子的参数，可计算包含这些原子的分子体系。

MMFF93 力场为美国的 Merck 公司针对有机药物设计所发展的。此力场引用大量的量子计算结果为依据，采取 MM2、MM3 力场的形式。其主要应用于计算固态或液态的小型有机分子系统。可得到很准确的几何结构、振动频率及各种热力学性质。其力场的形式为：

$$U = \sum_b U_b(b) + \sum_\theta U_\theta(\theta) + \sum_{b,\theta} U_{b,\theta}(b,\theta) + \sum_\phi U_\phi(\phi) + \sum_d U_d(d) + \sum U_{VDW} + \sum_{i,j} U_{elec} \tag{2-24}$$

## 2.4 涵盖周期表元素的力场

前述的力场由于最初的设计都是针对有机分子或生化分子，故仅能涵盖周期表中的某部分元素。其力场参数多由拟合计算与实验值或量子计算值而来，较缺乏理论的根据。为使力场能广泛地适用于整个周期表所涵盖的元素，发展了从原子角度为出发点的力场。这些新一代的力场其原子的参数来自实验或理论的计算，具有真实的物理意义。

以原子为根据的力场包括 ESFF 力场、UFF 力场（universal force field）与 DREDING 力场等。ESFF 力场可用于预测气态与凝态的有机分子、无机分子、有机金属分子系统的结构，但不能用以计算构型能或准确的振动频率。其力场涵盖周期表中由氢至氡的元素。DREDING 力场可计算分子聚集体的结构及各项性质，但其力场参数并未涵盖周期表的全部元素。UFF 力场可适用于周期表所涵盖的所有元素，即适用于任何分子与原子系统。以 UFF 力场所计算的分子结构优于由 DREDING 力场所得的结果，但计算与分子间作用有关的性质，则有较大的偏差。

ESFF 力场考虑了各种不同的环形化合物，并将环状化合物其中心原子的混成模式（sp，$sp^2$，$sp^3$）等纳入键角弯曲项。其力场的形式为：

$$U = \sum_b D_b\{1 - \exp[-\alpha(r_b - r_b^0)^2]\} + \begin{cases} \sum_a \frac{K_a}{\sin^2\theta_\alpha^0}(\cos\theta_a - \cos\theta_a^0)^2 \\ \sum_a 2K_a(\cos\theta_a + 1) \\ \sum_a K_a^{\theta_a}\cos^2\theta_a \\ \sum_a \frac{2K_a}{n^2}[1 - \cos(n\theta_a)] + 2K_a^{-\beta(r_{13}-\rho_d)} \end{cases} +$$

$$\sum_\tau D_\tau\left[\frac{\sin^2\theta_1\sin^2\theta_2}{\sin^2\theta_1^0\sin^2\theta_2^0} + \text{sign}\bullet\ \frac{\sin^n\theta_1\sin^2\theta_2}{\sin^n\theta_1^0\sin^n\theta_2^0}\cos(n\tau)\right] +$$

● 符号函数。

$$\sum_{\chi} D_{\chi}\chi^2 + \sum_{nb}\left(\frac{A_i B_j + A_j B_i}{r_{nb}^9} - 3\frac{B_i B_j}{r_{nb}^6}\right) + \sum_{nb}\frac{q_i q_j}{r_{nb}} \tag{2-25}$$

式中，第一项为键伸缩项，第二项为键角弯曲项，第三项为二面角的扭转项，第四项为非平面振动项，第五项为非键结远程作用项，最后项为静电作用项。

ESFF 力场最重要的是引用了合理的方法，推导分子中原子所带的部分电荷量（partial charge）。各原子电荷量的总和必须与分子所带的电荷量相等。静电作用能为：

$$E = \sum_i (E_i^0 + \chi_i q_i + \frac{1}{2}\eta_i q_i^2) + \sum_{i>j} B\frac{q_i q_j}{R_{ij}} \tag{2-26}$$

式中，第一项为原子能量对电荷量的泰勒展开式；$\chi$ 为原子的阴电性（electronegativity）；$\eta$ 为原子的坚硬度（hardness）。将式中的能量对原子的电荷量最小化，可得每个原子的电荷量：

$$q_i = \frac{\lambda - \chi_i - \Delta\chi_i}{\eta_i} \tag{2-27}$$

式中，$\lambda$ 为拉格朗日乘子（Lagrangian multiplier）。原子的阴电性与其坚硬度随分子所处的环境不同而不同。ESFF 力场应用精确的量子力学密度泛函数的计算方法得到此二参量。

## 2.5 特殊的力场

除了上述的一般性力场外，还有一些力场为针对某些特殊系统所设计的。例如专门计算沸石（zeolite）系统的力场、专门计算金属氧化物固体的 COMPASS 力场、专门计算高分子的力场。若力场考虑系统中所有的原子，称为全原子（all-atom）力场，还有一些简化的力场，将原子基团合并考虑的，称为联合原子（united-atom）力场。例如 Rykaert 与 Bellemans 等针对长链烷类化合物所设计的力场，此力场将甲基—$CH_3$ 与亚甲基—$CH_2$—均视为一原子团，质量分别为 15 原子质量单位（amu）及 14 原子质量单位。原子团的中心位于碳原子上。其力场的形式为：

$$U = U_{nb} + U_b + U_\theta + U_\phi$$

$U_{nb} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$，$r$ 为所有相距大于 4 个基团的距离

$$U_b = k_b(r - r_0)^2, \qquad r \text{ 为相连基团间的距离}$$

$$U_\theta = \frac{1}{2}k(\theta - \theta_0)^2, \qquad \theta \text{ 为 3 个相连基团的键角}$$

$$U_\phi = \sum_{n=0}^{5} a_n \cos^n \phi, \qquad a_n \text{ 为常数，} \phi \text{ 为双面扭转角} \qquad (2\text{-}28)$$

另外一种常见的联合原子简化力场为 Gay-Berne（GB）所设计的。此力场最大的作用在于简化非均向性（anisotropic）分子间的作用（如苯分子）。其特点为将势能参数 $\sigma$ 与 $\epsilon$ 视为互相作用分子的距离及排列位向的函数。

Gay-Berne 简化力场将二非均向分子间的作用势能定为

$$U(r_{ij}) = 4\epsilon(\hat{u}_i, \hat{u}_j, \hat{r})\left\{\left[\frac{\sigma_0}{r_{ij} - \sigma(\hat{u}_i, \hat{u}_j, \hat{r}) + \sigma_0}\right]^{12} - \left[\frac{\sigma_s}{r_{ij} - \sigma(\hat{u}_i, \hat{u}_j, \hat{r}) + \sigma_s}\right]^{6}\right\} \qquad (2\text{-}29)$$

式中，LJ 势能的参数均为分子位向的函数；$\hat{u}_i$ 与 $\hat{u}_j$ 为表示分子位向的单位向量；$\hat{r}$ 为连接二分子中心的单位向量。图 2-10 显示分子间位向的各种关系。

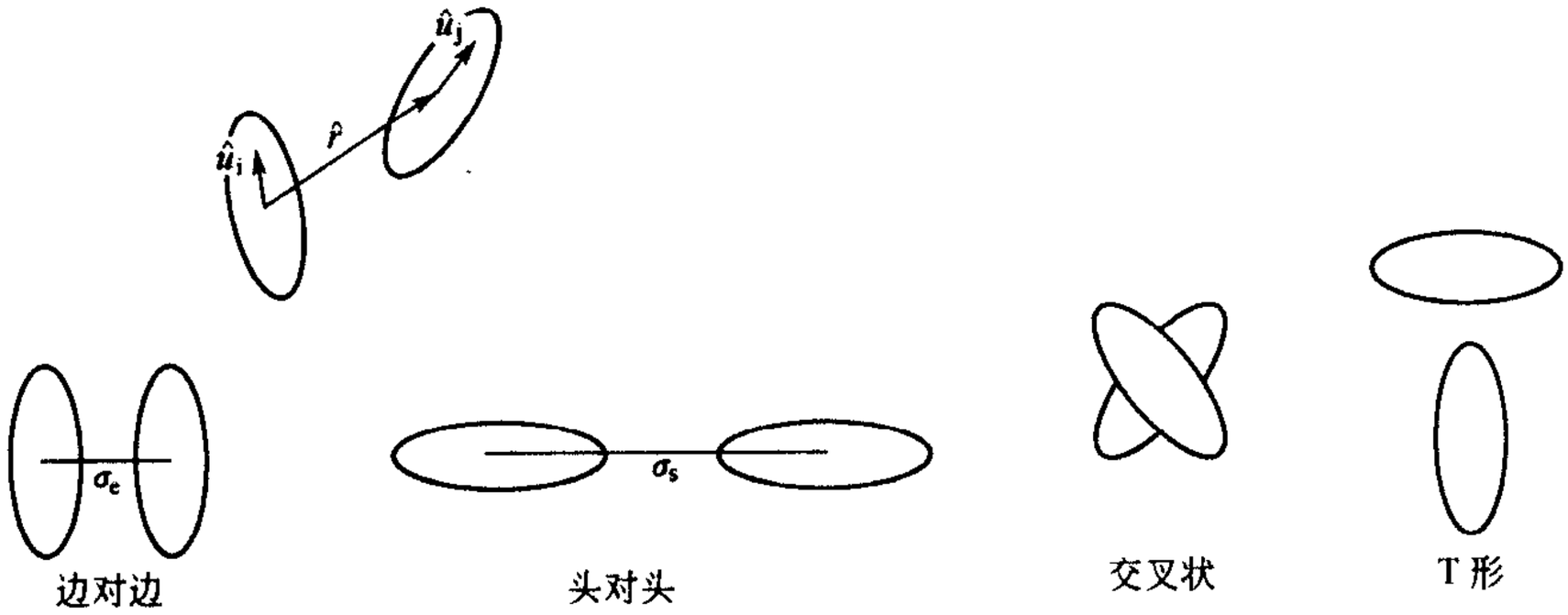



**图 2-10　GB 力场所考虑分子排列的各种可能**

若将分子视为椭圆的饼状，则其形状可由两个参数 $\sigma_s$ 与 $\sigma_e$ 描述。此二参数与势能的关系为：

$$\sigma(\hat{u}_i, \hat{u}_j, \hat{r}) = \sigma_0\left\{1 - \frac{\chi}{2}\left[\frac{(\hat{u}_i \cdot \hat{r} + \hat{u}_j \cdot \hat{r})^2}{1 + \chi(\hat{u}_i \cdot \hat{u}_j)} + \frac{(\hat{u}_i \cdot \hat{r} - \hat{u}_j \cdot \hat{r})^2}{1 - \chi(\hat{u}_i \cdot \hat{u}_j)}\right]\right\}^{-1/2} \qquad (2\text{-}30)$$

其中，

$$\chi = \frac{(\sigma_e/\sigma_s)^2 - 1}{(\sigma_e/\sigma_s)^2 + 1} \qquad (2\text{-}31)$$

$\chi$ 称为非均向性形状参数（anisotropic shape parameter），反映椭圆体的形状。若为圆形，则 $\sigma_e=\sigma_s$，$\chi$ 值为 0；若为长杆状，则 $\chi=1$，若为扁圆面，则 $\chi=-1$。实际计算中，通常令 $\sigma_0=\sigma_s$。

能量参数 $\epsilon$ 与位向的关系为：

$$\epsilon(\hat{u}_i,\hat{u}_j,\hat{r})=\epsilon_0\epsilon'^u(\hat{u}_i,\hat{u}_j,\hat{r})\epsilon^v(\hat{u}_i,\hat{u}_j)$$

$$\epsilon(\hat{u}_i,\hat{u}_j)=[1-\chi'(\hat{u}_i,\hat{u}_j)^2]^{-1/2}$$

$$\epsilon'^u(\hat{u}_i,\hat{u}_j,\hat{r})=\left\{1-\frac{\chi'}{2}\left[\frac{(\hat{u}_i\cdot\hat{r}+\hat{u}_j\cdot\hat{r})^2}{1+\chi'(\hat{u}_i\cdot\hat{u}_j)}+\frac{(\hat{u}_i\cdot\hat{r}-\hat{u}_j\cdot\hat{r})^2}{1-\chi'(\hat{u}_i\cdot\hat{u}_j)}\right]\right\} \quad (2\text{-}32)$$

$\chi'$ 为度量非均向吸引力的参数，即：

$$\chi'=\frac{1-(\epsilon_e/\epsilon_s)^{1/u}}{(\epsilon_e/\epsilon_s)^{1/u}+1} \quad (2\text{-}33)$$

式中，$\epsilon_e$ 为椭圆体头对头排列时势能的深度，此时吸引与排斥的作用互相抵消；$\epsilon_s$ 为二椭圆体边对边排列时势能的深度。Luckhurst 等比较 Gay-Berne 力场与实验的结果，定出 $v=1$，$u=2$ 的关系式。

Gay-Berne 力场的形式虽然复杂，但并不需要计算所有原子间的作用，而是将整个分子视为椭圆体，因此大量地简化了计算的工作，并且可将计算的体系扩大许多。此外，这种力场只需要少数几个参数值，十分简便。Gay-Berne 力场适用于液晶（liquid crystal）分子体系的计算，已获得许多良好的结果。图 2-11 列出一些常见的液晶分子，以供参考。

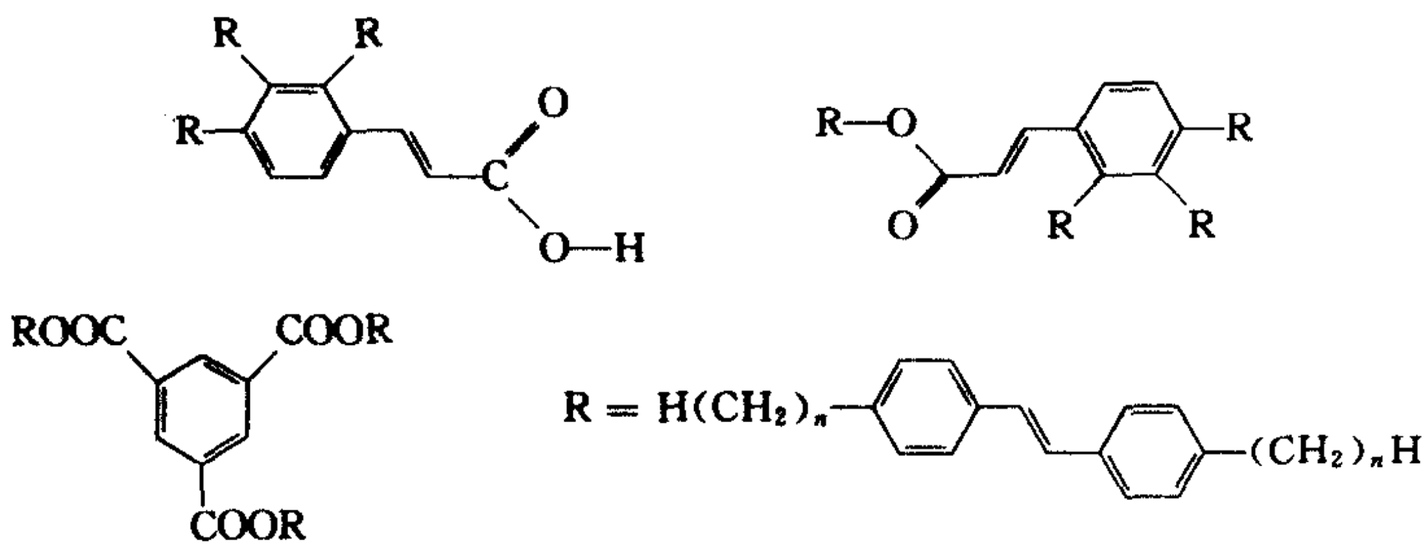



**图 2-11　常见的液晶分子**

这些液晶分子均含有苯环，利用 GB 力场可简化许多计算工作。图 2-12 为 Bates 与 Zannoni 利用 GB 力场计算 12960 个向列型液晶系统的结构。图中清楚地显示系统中液晶分子的排列次序，与实验结果一致。

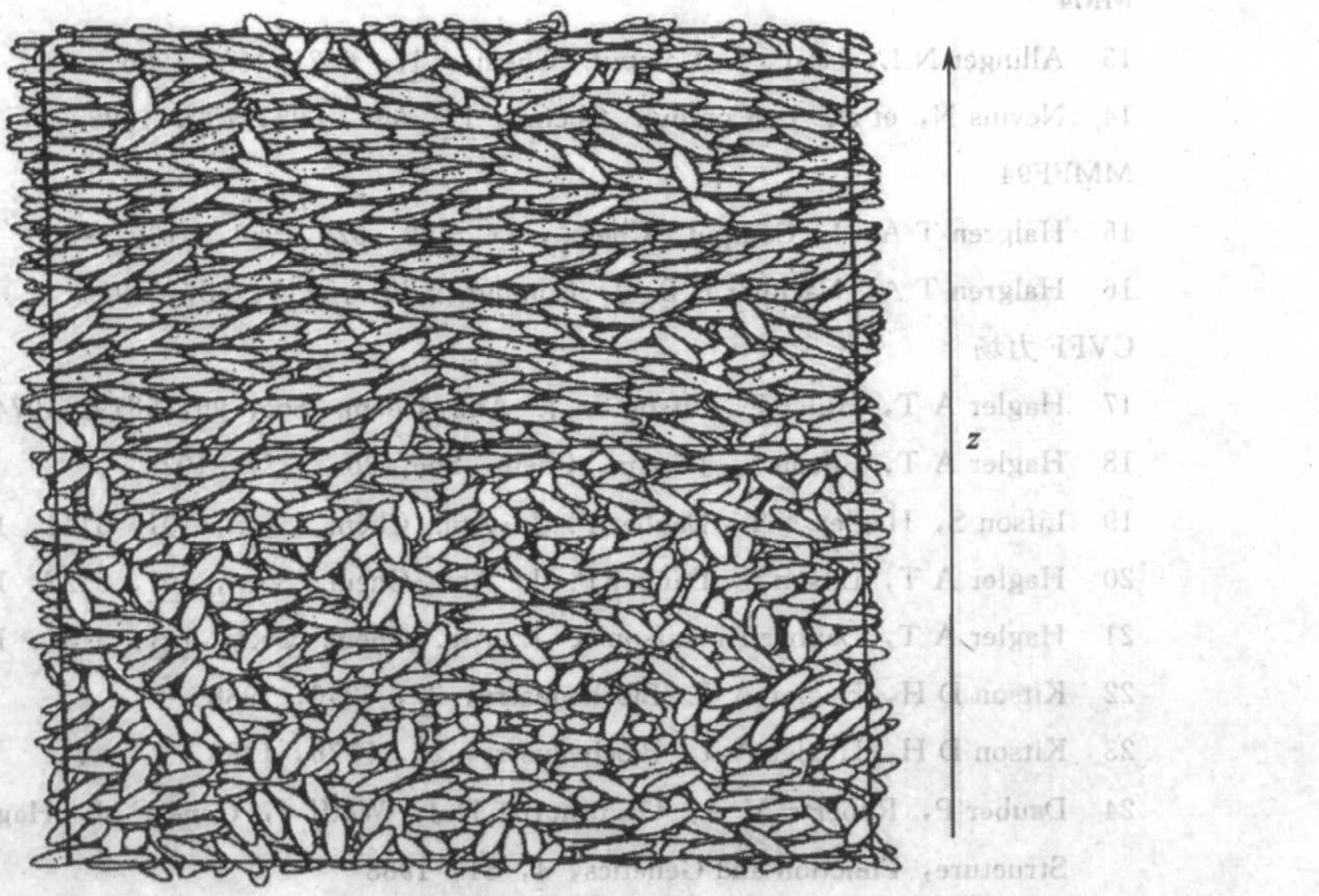



图 2-12　利用 GB 力场计算的向列型液晶系统的结构

## 参　考　文　献


CFF 力场

1　Maple J R, Tacher T S, Dinur U, Hagler T. Chemical Design Automation News, 5 (9), 5, 1990

2　Maple J R, Hwang M J, Stokfisch T P, Hagler A T. Israel J. Chem., 34, 195, 1944b

3　Maple J R, Hwang M J, Stockfisch T P, Dinur U, Waldman M, Ewig C S, Hagler A T. J. Comput. Chem. 15, 162, 1994a

4　Maple J R, Dinur U, Hagler A T. Proc. Nat. Acad. Sci. USA, 85, 5530, 1988

5　Hwang M J, Stockfish T P, Hagler A T. J. Am. Chem. Soc, 116, 2515, 1994

6　Maple J R, et al. J. Comput. Chem., 19, 430, 1998

PCFF 力场

7　Sun H, Mumby J S, Maple J R, Hagler A T. J. Am. Chem. Soc., 116, 2978, 1994

8　Sun H, Mumby J S, Maple J R, Hagler A T. J. Phy. Chem. 99, 5873, 1995

9　Sun H. Macromolecules, 28, 701, 1995

10　Sun H. Macromolecules, 26, 5924, 1994

MM 形态力场

MM2

11　Kao J, Allinger N L. J. Am. Chem. Soc., 99, 975, 1977

MM3

12 Allinger N L, Yan L. J. Am. Chem. Soc., 115, 11918, 1993

MM4

13 Allinger N L, et al. J. Comput. Chem., 17, 642, 747, 1996

14 Nevins N, et al. J. Comput. Chem., 17, 669, 695, 730, 1996

MMFF94

15 Halgren T A. J. Comput. Chem., 17, 490, 520, 553, 616, 1996

16 Halgren T A, Nachbar R B. J. Comput. Chem., 17, 587, 1996

CVFF 力场

17 Hagler A T, Huler E, Lifson S. J. Am. Chem. Soc, 96, 5319, 1974

18 Hagler A T, Lifson S. J. Am. Chem. Soc, 96, 5327, 1974

19 Lifson S, Hagler A T, Dauber P. J. Am. Chem. Soc., 101, 5111, 1979

20 Hagler A T, Lifson S, Dauber P. J. Am. Chem. Soc., 101, 5122, 1979

21 Hagler A T, Dauber P, Lifson S. J. Am. Chem. Soc., 101, 5131, 1979

22 Kitson D H, Hagler A T. Biochemistry, 27, 5246, 1988

23 Kitson D H, Hagler A T. Biochemistry, 27, 7176, 1988

24 Dauber P, Roberts V A, Osguthorpe D J, Wolff J, Genest M, Hagler A T. Proteins: Structure, Function and Genetics, 4, 31, 1988

AMBER 力场

25 Weiner S J, et al. J. Comput. Chem., 7, 320, 1986

26 Weiner S J, et al. J. Am. Chem. Soc., 106, 765, 1984

AMBER94

27 Cornell W D, et al. J. Am. Chem. Soc., 117, 5179, 1995

28 E-mail: www. amber. ucsf. edu

CHARMM 力场

29 Brooks B R, et al. J. Comput. Chem., 4, 187, 1983

30 Momany F A, Rone R. J. Comp. Chem., 13, 888, 1992

31 Mackerell A D, et al. J. Am. Chem. Soc., 117, 11946, 1995

32 Mackerell A D, et al. J. Phys. Chem. B, 102, 3586, 1998

33 E-mail: yuri. Harvard. edu

OPLS 力场

34 Jorgensen W L, Tirado-Rives J. J. Am. Chem. Soc., 110, 1657, 1988

35 Pranata J, et al. J. Am. Chem. Soc., 113, 2810, 1991

OPLS-AA

36 Jorgensen W L, et al. J. Am. Chem. Soc., 118, 11225, 1996

GROMOS 力场

37 Daura X, Mark A E, van Gunsteren W F. J. Comput. Chem., 19, 535, 1998

DREDING 力场

38 Mayo S L, Olafson B D, Goddard W A. J. Phys. Chem., 94, 8897, 1990

UFF 力场

39 Pappé A K, et al. J. Am. Chem. Soc. 114, 10024, 1992

TRIPOS（SYBYL）力场

40 Clark M, et al. J. Comput. Chem., 10, 982, 1989

GB力场

41 Gay J G, Berne B J. J. Chem. Phys., 74, 3316, 1981

42 Luckhurst G R, Stephenes R A, Phippen R W. Liquid Crystals, 8, 451, 1990

43 Bates M A, Zannoni C. Chem. Phys. Lett., 26, 40, 1997

简化聚合物力场

44 Ryckaert J P, Bellemans A. Chem. Phys. Lett., 30, 123, 1975

45 Rigid D, Roe R J. J. Chem. Phys., 87, 7285, 1987

46 Rigid D, Roe R J. Macromolecules, 22, 2259, 1989

COMPASS力场

47 Bunte S W, Sun H. J. Phys. Chem. B, 104, 2477, 2001

# 第❸章 能量最小化

利用第2章介绍的力场，计算分子性能的方法称为分子力学（molecular mechanics，MM），其最重要的内容就是根据适合的力场计算分子各种可能构象的势能，势能最低的构象为最稳定的构象。寻找势能最低点的过程称为能量最小化（energy minimization），所得到的构象称为几何优化构象（geometry optimized configuration）。分子的几何优化构象是计算分子性能的基础。本章共分三部分内容，3.1和3.2节结合势能面图和函数介绍了能量最小点的概念；3.3～3.6节主要介绍了利用一阶导数和二阶导数求极小值的常见方法；3.7节以后结合几个简单例子介绍了能量最小化的应用。

## 3.1 势能图与势能面

除了一些简单的分子以外，大多数分子的势能是分子中一些复杂形式势能的组合。势能为分子中原子坐标的函数，由原子不同的坐标所得到的势能构成所谓的“势能面”（potential energy surface）。在本章中所讨论的“能量”指的是分子的势能。要了解分子的稳定性及一些热力学特性，则必须借助于分子的势能面。若分子含有 $N$ 个原子，其势能为 $3\times N$ 个笛卡儿坐标的函数，即：

$$U(x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_N, y_N, z_N) \tag{3-1}$$

此势能亦可表示为分子 $3N-6$（线性分子为 $3N-5$）个内坐标（internal coordinate）的函数。以水分子为例，其势能为键长及键角内坐标的函数，即：

$$U(\Delta r_1, \Delta r_2, \Delta\theta) \tag{3-2}$$

式中，$\Delta r_i = r_i - r_i^0 (i=1,2)$ 为键长的改变量；$\Delta\theta = \theta - \theta^0$ 为键角的改

变量，上标“0”表示其平衡值；$r_i^0=0.095\text{nm}$，$\theta^0=109.45°$。除了非常简单的分子以外，当然无法将所得到的势能面完全以图形显示（$3N$ 或 $3N-6$ 维）。因此，为了了解势能与标变化的关系，往往只计算部分坐标的势能面。以正戊烷分子为例，图 3-1 中显示分子的两个扭转角 $\tau_1$ 及 $\tau_2$，其扭转的范围均为 0°→360°。

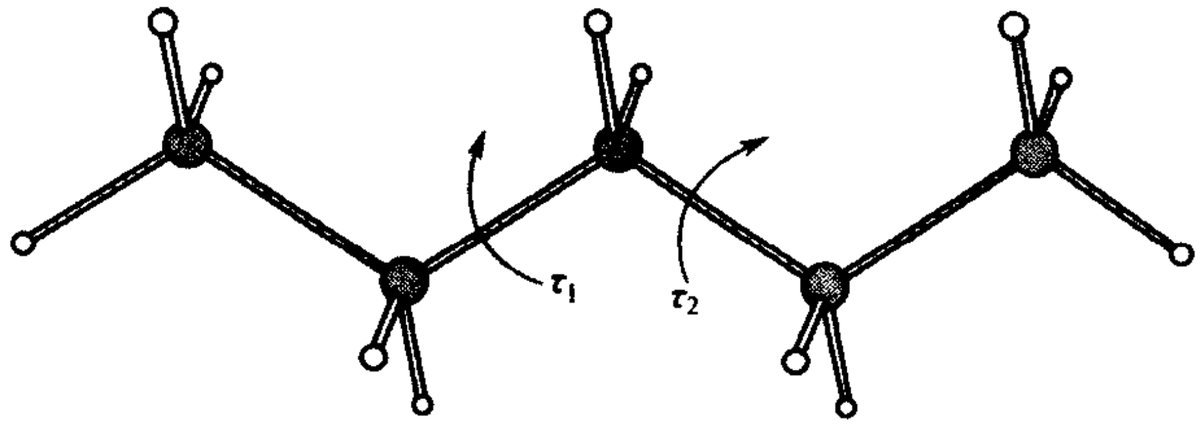



**图 3-1　正戊烷分子示意**

若将分子中其余的键长及键角都固定不变，仅仅将这两个扭转角旋转，则得到的势能面以等高线的方式表示，如图 3-2 所示。

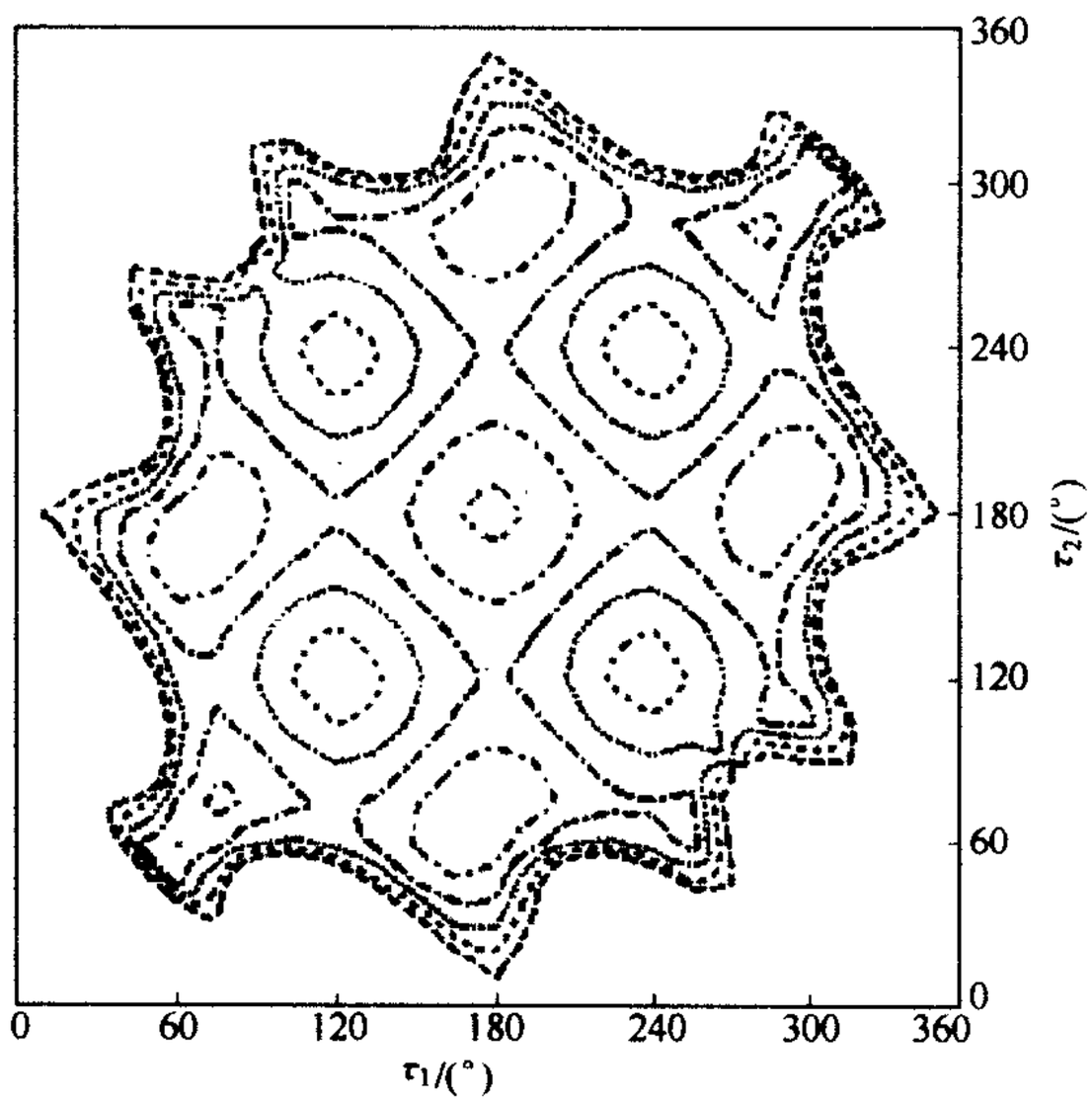



**图 3-2　正戊烷扭转势能的等高线**

若将其以三维的图形显示，则如图 3-3 所示。

图 3-3 中，垂直于角度平面（$\tau_1$，$\tau_2$）为正戊烷分子所对应的势能值。

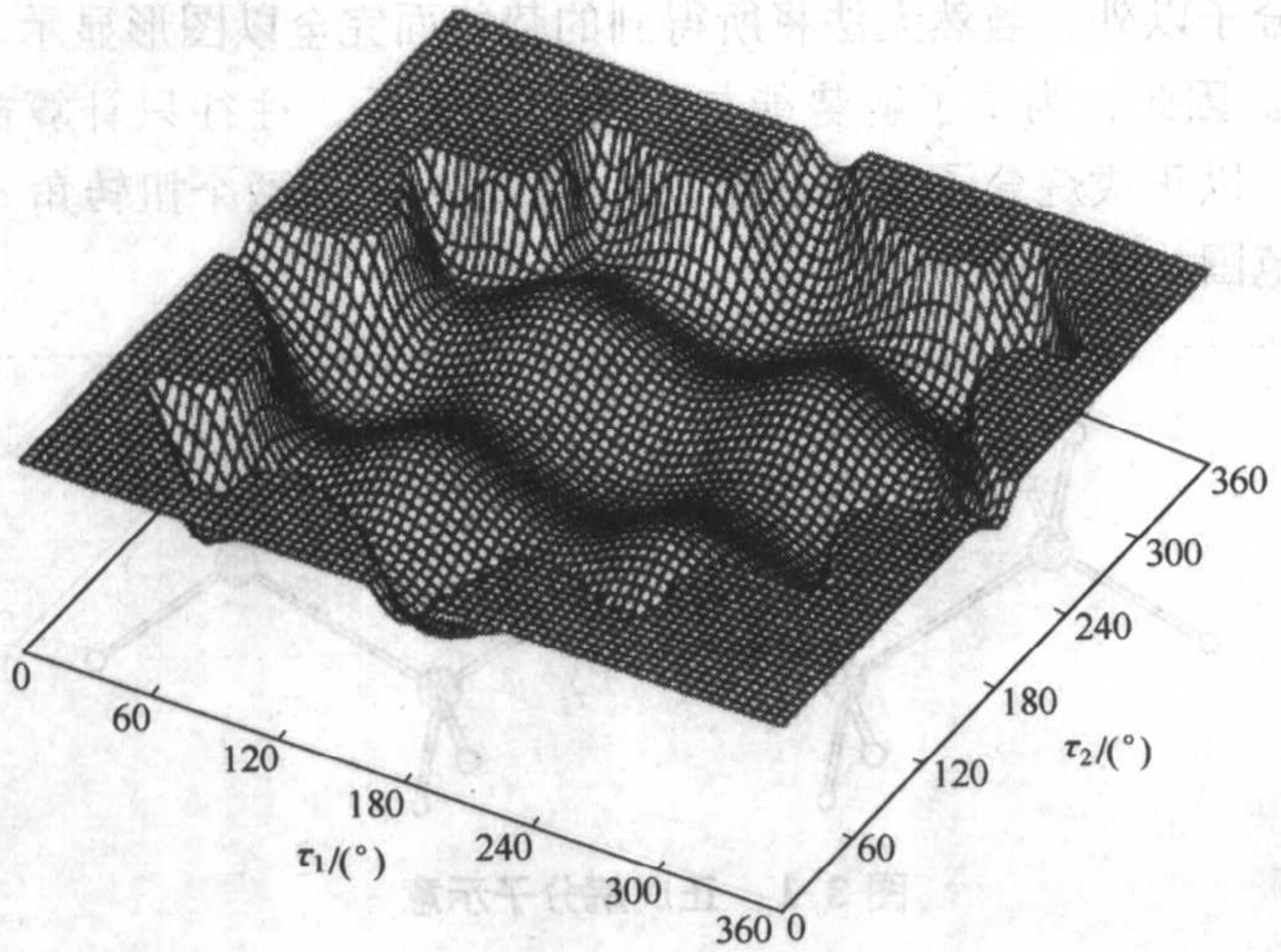



**图 3-3　正戊烷扭转势能面的三维表示**

在讨论分子的稳定性时，最重要的工作是找出势能面极小值的位置。势能的极小值所对应原子的位置，相当于分子稳定状态的构象。若原子的位置偏离分子的稳定构象，则对应于较高能量状态。通常势能面可得到许多极小值的位置，这些极小值中能量最低的称为最低极小值（global energy minimum），相当于分子最稳定的构象。由势能面求最低极小值的过程称为能量最优化（energy optimization），其所对应的结构称为最优化结构（optimized structure）。

## 3.2　函数的极小值

能量最优化过程相当于求势能函数的极小值。单一变量的函数 $f(x)$ 具有极小值的条件为：

$$f'(x)=0,\ f''(x)>0 \tag{3-3}$$

含有 $n$ 个变数的多变量函数 $f(x_1,x_2,\cdots,x_n)$ 具有极小值的条件为：

$$\frac{\partial f}{\partial x_i}=0;\ \frac{\partial^2}{\partial x_i^2}f>0,\ i=1,2,\cdots,n \tag{3-4}$$

以两个变量的函数 $f(x,y)=x^2+2y^2$ 为例，此函数值的等高线表示如图 3-4 所示。

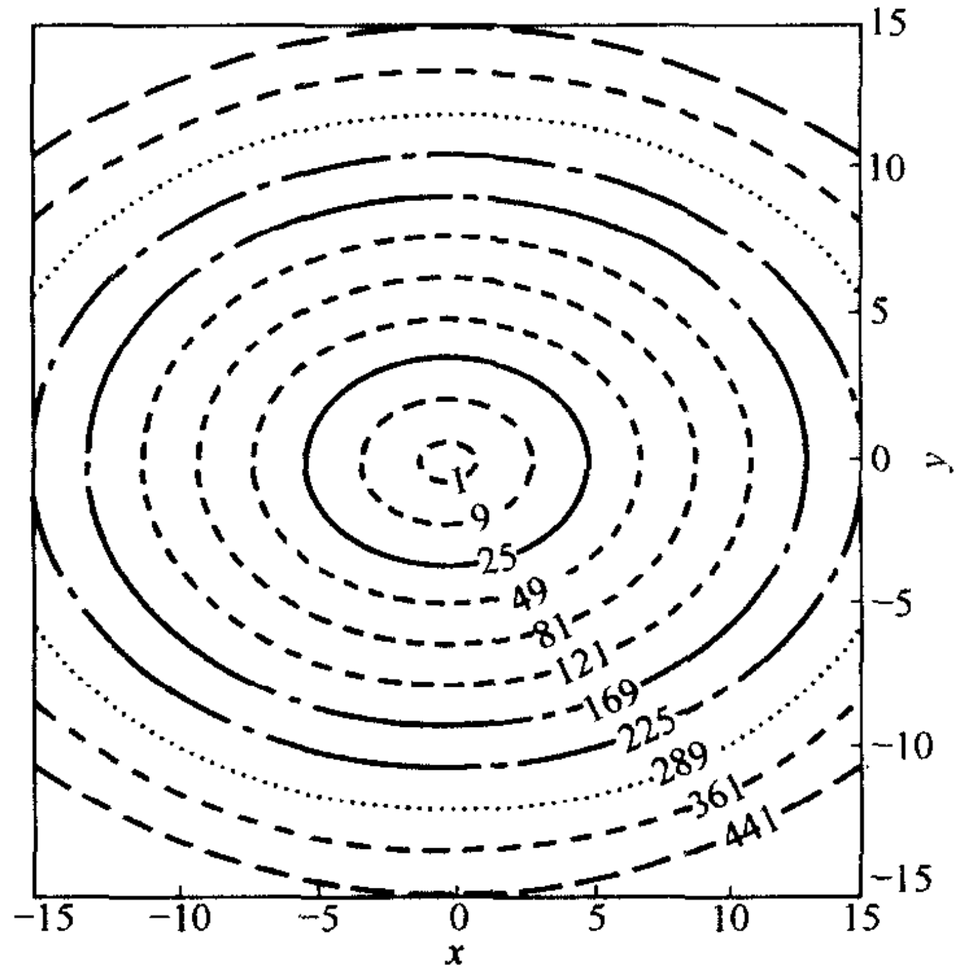



**图 3-4　函数 $f(x,y)=x^2+2y^2$ 的等高线**

图中，等高线所标示的数字为此函数的值。此函数仅有一最小值，位于图 3-4 中的原点。现以此函数为例，举出各种常用的求极小值的方法。

一般求函数的极小值的方法为利用函数的导数，以最快速度与最小储存容量为基本原则。有时，也可能将多种方法并用，视实际需要而定。图 3-5 表示了一般求极小值可能的情形。图中，$A$，$B$，$C$ 为不同的起始点，$G_1$ 与 $G_2$ 则为函数的两个极小值点。

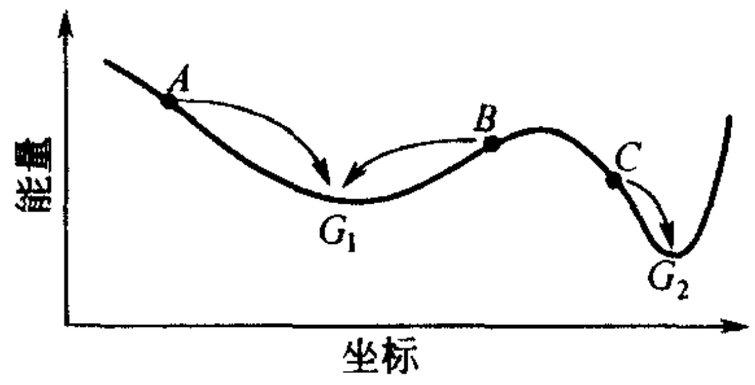



**图 3-5　求函数极小值可能的情形**

求取函数得极小值如同由起始点开始，沿着函数曲线滚球一样。设由 $A$ 或 $B$ 点开始，求函数的极小值，则得到点 $G_1$；但若由 $C$ 点为起点，求函数的极小值，则得到点 $G_2$。此时，欲求函数其他的极小值，则必须改变起始点。

求取函数的导数，可利用数值方法或分析方法。分析方法即直接由函数的数学式求其微分的数学式，由此可得到完全精准的导数，而且速度快。数值方法比较费时，所求得的导数亦可能有误差，但适用于非常复杂的函数。以一元函数 $f(x)$ 为例，其导数为：

$$f'(x)=\lim_{\delta x\to 0}\frac{f(x+\delta x)-f(x)}{\delta x} \tag{3-5}$$

式中的数值方法为计算起始点 $x_0$ 的函数值 $f(x_0)$，再由 $x_0$ 变动一很小值 $\delta x$ 至 $x_0+\delta x$，计算此点的函数值 $f(x_0+\delta x)$。再将两个函数值此代入式(3-5) 可得函数当 $x=x_0$ 的导数。更精确的计算方法为计算 $f(x_0+\delta x)$ 与 $f(x_0-\delta x)$ 的函数值，得导数为：

$$\therefore\ f'(x_0)=\frac{f(x_0+\delta x)-f(x_0-\delta x)}{2\delta x} \tag{3-6}$$

以函数 $f(x)=2x^3-5x+1$ 为例，设起始点为 $x=1$。由分析方法得到此点的导数为：

$$f'(x=1)=6x^2-5=6(1)^2-5=1 \tag{3-7}$$

以数值方法，设 $\delta x=0.01$，由式(3-5)，得：

$$f(x=1)=2-5+1=-2$$

$$f(x=1+0.01)=2(1.01)^3-5(1.01)+1=-1.9894$$

$$f'(x)=\frac{-1.9894-(-2)}{0.01}=1.062 \tag{3-8}$$

误差约为 6%。或以较精确的方法，由式(3-6)，得：

$$f(x=1-0.01)=2(0.99)^3-5(0.99)+1=-2.0094$$

$$f'(x)=\frac{-1.9894-(-2.0094)}{2(0.01)}=1 \tag{3-9}$$

所得的结果与完全精确的分析方法的结果相比较，几乎无误差。

## 3.3 利用导数求极小值的方法

分子力学中，求取势能的极小值多利用导数极小化法。势能的一次导数称为梯度（gradient），为一向量。此向量的方向指向极小值点，其大小代表目前所处位置的倾斜度（steepness）。势能函数的二次导数为势能面在此位置的弯曲度（curvature）。分子中原子所受的力为与所处位置势能函数的梯度方向相反，即：

$$\vec{F}_i=-\nabla_i U=-\left(\vec{i}\,\frac{\partial U}{\partial x_i}+\vec{j}\,\frac{\partial U}{\partial y_i}+\vec{k}\,\frac{\partial U}{\partial z_i}\right),\ i=1,2,\cdots,N \tag{3-10}$$

式中，$\nabla$ 为梯度的符号。各原子可依所受的力方向移动以找寻势能的极小值。

利用导数求函数的极小值需将函数 $f(x)$ 以泰勒级数（Taylor series）对 $x_k$ 点展开，即：

$$f(x)=f(x_k)+(x-x_k)f'(x_k)+(x-x_k)^2 f''(x_k)/2+\cdots \quad (3\text{-}11)$$

若为多元函数，则以向量 $\vec{x}=\begin{Bmatrix}x_1\\x_2\\\cdots\\x_N\end{Bmatrix}$ 取代式(3-11) 中的变量 $x$，得：

$$f(\vec{x})=f(\vec{x}_k)+(\vec{x}-\vec{x}_k)f'(\vec{x}_k)+(\vec{x}-\vec{x}_k)^T f''(\vec{x}_k)(\vec{x}-\vec{x}_k)/2+\cdots \quad (3\text{-}12)$$

式中，$f'(\vec{x}_k)$ 为一向量，为函数对各个坐标的偏微分，即：

$$f'(\vec{x}_k)=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_N}\right)_k \quad (3\text{-}13)$$

括号外下标“$k$”表示各微分项于 $\vec{x}_k$ 的值。式(3-12) 中，$f''(\vec{x}_k)$ 为一矩阵，为函数对坐标的二次微分，即：

$$f''(\vec{x}_k)=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_N}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_N}\\ \cdots & \cdots & \cdots & \cdots\\ \frac{\partial^2 f}{\partial x_N\partial x_1} & \frac{\partial^2 f}{\partial x_N\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_N^2}\end{bmatrix}_k \quad (3\text{-}14)$$

$(\vec{x}-\vec{x}_k)^T$ 表示向量 $(\vec{x}-\vec{x}_k)$ 的转置向量（transpose)。若分子含 $N$ 个原子，其势能 $U$ 为所有原子坐标的函数，含 $3N$ 个变量。故 $U'(\vec{x}_k)$ 为 $3N$ 维向量，$U''(\vec{x}_k)$ 为 $3N\times 3N$ 矩阵，一般称此二次微分矩阵为 Hessian 矩阵（Hessian matrix)，或力常数矩阵（force constant matrix)。利用如式(3-12) 的泰勒展开式，可由 $\vec{x}_k$ 开始，逐步求取使能量降低的结构，直至找到能量极小值为止。利用一次导数求极小值称为一次求极值法，利用一次及二次导数求极小值称为二次求极值法，余类推。

## 3.4 一次导数求极值法

分子力学常采用一次求极值法以定出势能极小值的分子构象。最常用的两种方法为最速下降法（steepest descents）与共轭梯度法（conjugate

gradient)。一次求极值法由 $\vec{x}_k$ 位置开始，逐渐更动原子的坐标，使分子逐渐移向其最低能量的构象。这两种方法各有其优点、缺点及适用的范围，分述如下。

### 3.4.1 最速下降法

设向量 $\vec{g}_k$ 表示势能函数在分子位于 $\vec{x}_k$ 构象的梯度［或一次导数 $f'(\vec{x}_k)$］，则可由移动原子的坐标求出沿此方向的能量最小的构象 $\vec{x}_{k+1}$，即：

$$\vec{x}_{k+1}=\vec{x}_k+\lambda\vec{s}_k$$

$$\vec{s}_k=-\frac{\vec{g}_k}{|\vec{g}_k|} \tag{3-15}$$

式中，$\vec{g}_k$ 为各原子所受的力；$\vec{s}_k$ 表示沿力方向的单位向量；$\lambda$ 为找到极小值所需移动的幅度参数。以函数 $f(x,y)=x^2+2y^2$ 为例，设起始的 $k$ 点坐标为（9，9），此点的函数值为 $f(\vec{x}_k)=(9)^2+2(9)^2=243$，即：

$$\vec{g}_k=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right)_k=(2x,4y)_k=(18,36)$$

$$\vec{s}_k=-\frac{\vec{g}_k}{|\vec{g}_k|}=-\frac{(18,36)}{\sqrt{18^2+36^2}}=(-0.447,-0.894)$$

$$\vec{x}_{k+1}=\vec{x}_k+\lambda\vec{s}_k=\begin{pmatrix}9\\9\end{pmatrix}+\lambda\begin{pmatrix}-0.447\\-0.894\end{pmatrix}=\begin{pmatrix}9-0.447\lambda\\9-0.894\lambda\end{pmatrix} \tag{3-16}$$

表 3-1 列出各 $\lambda$ 值的计算结果。

**表 3-1　一次导数方法求函数 $f(x,y)=x^2+2y^2$ 的极小值的计算结果**

| $\lambda$ | $\vec{x}_{k+1}$ | $f(\vec{x}_{k+1})$ | $\lambda$ | $\vec{x}_{k+1}$ | $f(\vec{x}_{k+1})$ |
|---|---|---|---|---|---|
| 1.0 | (8.53,8.11) | 204.57 | 8.0 | (5.42,1.85) | 36.25 |
| 2.0 | (8.11,7.21) | 169.73 | 9.0 | (4.98,0.95) | 26.59 |
| 3.0 | (7.66,6.32) | 138.49 | 10.0 | (4.53,0.06) | 20.53 |
| 4.0 | (7.21,5.42) | 110.85 | 11.0 | (4.08,−0.83) | 18.06 |
| 5.0 | (6.76,4.53) | 86.754 | 11.2 | (4.00,−1.00) | 18.00 |
| 6.0 | (6.32,3.63) | 66.36 | 12.0 | (3.64,−1.73) | 19.19 |
| 7.0 | (5.87,2.79) | 49.51 | | | |

计算最小值的过程如图 3-6 所示，图中的斜线表示起始点梯度的方向。沿此线所求得的能量示于图的下方，此能量逐渐降低，经过最低点后又升高。能量的最小值点可由线性寻找法（line search）定出。检视各点的能量计算值直至三连续点的能量为 $f(\vec{x}_n)>f(\vec{x}_{n+1})<f(\vec{x}_{n+2})$（如表

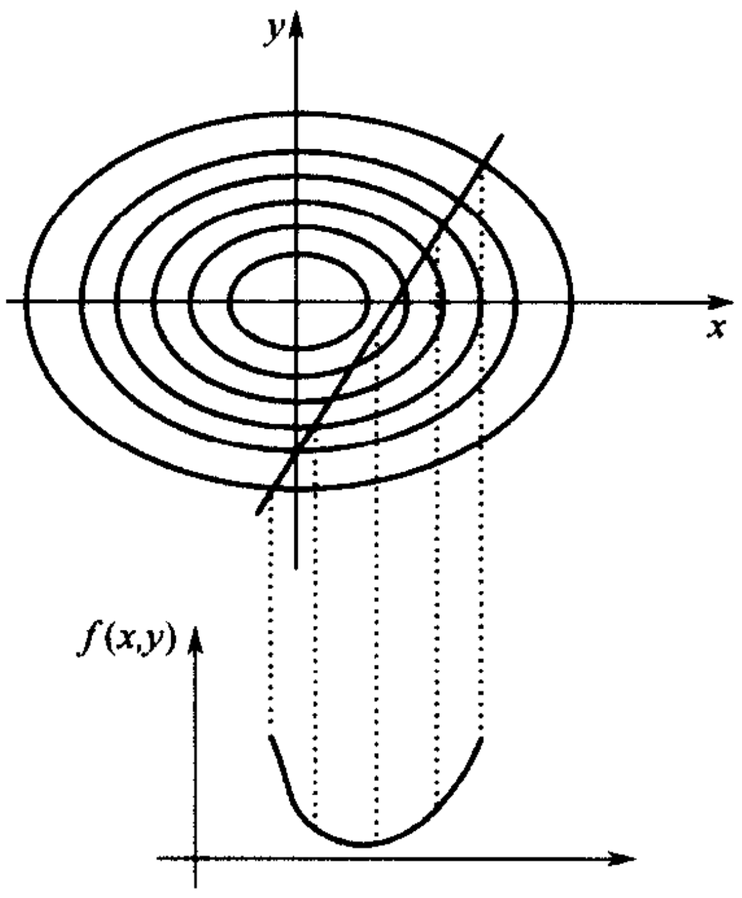



**图 3-6　最速下降法寻找函数极小值**

3-1 中的 $\lambda=11.0$、11.2、12.0 三点)。此时，必有一最小值点处于 $\vec{x}_n$ 与 $\vec{x}_{n+2}$ 之间。若找出这样的三点，则可由 $\vec{x}_n$ 开始，逐步缩小原子移动的幅度，以找出最小值点。但这样的程序较为费时，需要大量的计算，因此常以其他较快的近似法取代。

常用的近似法，为以此三点拟合一二次抛物线函数。由此抛物线函数的数学型式，可以微分方法找出最小值点。再由新得到的三点找出新的抛物线及其最低点。重复这样的步骤，可定出函数的极小值点。图 3-7 例示这样的过程。

图 3-7 中，1，2，3 三点为由最速下降法所找出的点，对应的函数值为 $f(x_1)>f(x_2)<f(x_3)$。由此三点可拟合得抛物线 $f_1$，$f_1(x)$ 的最小值点对应于点 4。再取 1，2，4 三点，拟合得抛物线 $f_2$。$f_2(x)$ 的最小值对应于点 5，此最小值已非常接近函数 $f(x)$ 的最小值。在上例中，沿着起始点 (9.0，9.0) 的梯度所找出的最小值点为 $\vec{x}_{k+1}=(4.0,-1.0)$，其所对应的函数值为 18.0。以此点为另一起始点，由式(3-16) 得出其梯度 $\vec{g}_{k+1}=(8.0,-4.0)$，再重复以上的步骤，沿着梯度的方向寻找另一最小值点。如此，重复此步骤，可找出函数的最小值点 (0.0，0.0)。最速下降法每步所得出的梯度，有如下的正交关系：

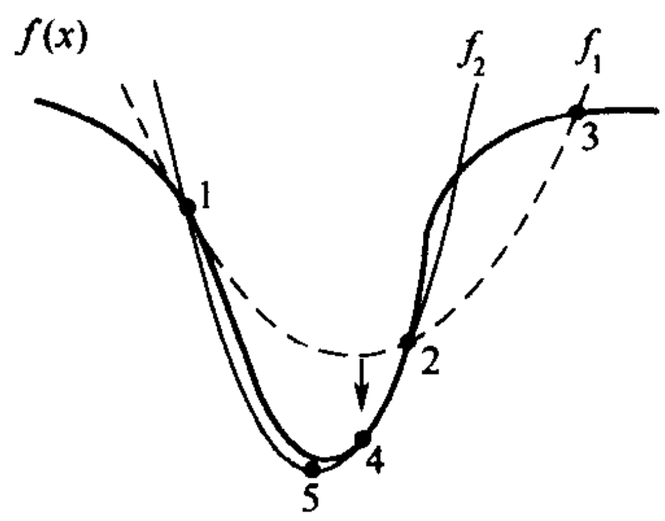



**图 3-7　抛物线拟合求极小值的示意**

$$\vec{g}_{k+1}\cdot\vec{g}_k=0 \tag{3-17}$$

意即 $k$ 点的梯度方向与 $k+1$ 点的梯度方向互相垂直。图 3-8 例示由此方法寻找函数 $x^2+2y^2$ 极小值点的情形。

除了线性寻找最小值的方法外，另一种常用以寻找最小值的方法为可调步程寻找法 (arbitrary step approach)。其基本式为：

$$\vec{x}_{k+1}=\vec{x}_k+\lambda_k\vec{s}_k \tag{3-18}$$

式中，$\lambda_k$ 表示每一步移动幅度的大小。通常先将 $\lambda_k$ 设为 1，由最速

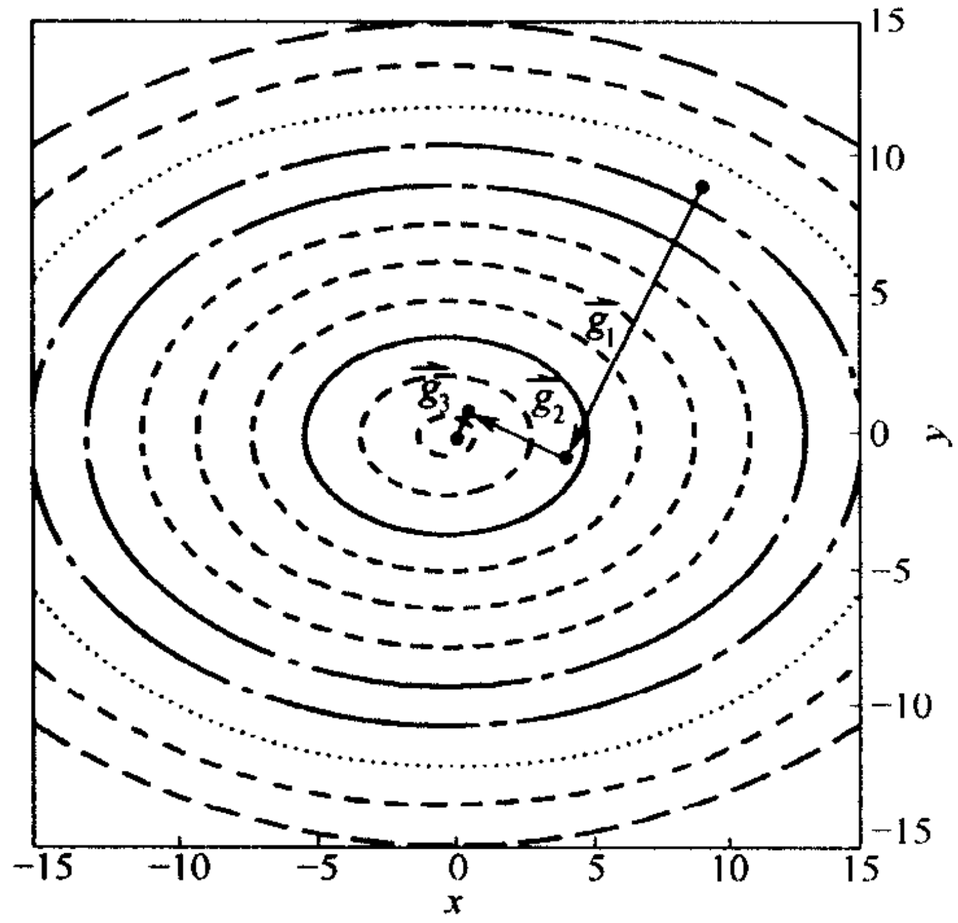



**图 3-8　最速下降法寻找最小值的梯度向量关系**

下降法得到第一步的结构。若能量小于起始结构的能量，则增加 $\lambda_k$ 之值为 1.2，再由式(3-18) 得到第二步的结构与能量。若能量继续降低，则再增加 $\lambda_k$ 值为前一步的 1.2 倍。如此，逐步重复增加 $\lambda_k$ 之值，至能量增加的情形出现。此时，应已超过了最小值点。因此，回到前一点的结构，并将 $\lambda_k$ 值减小为原先的 0.5 倍，再试验是否得到能量降低的结构。以此，逐步改变 $\lambda_k$ 值至找出最小值点为止。

### 3.4.2　共轭梯度法

共轭梯度法是将原子由 $\vec{x}_k$ 处的构象沿着向量 $\vec{v}_k$ 移动；$\vec{v}_k$ 为由 $k$ 点的梯度与前一步的 $\vec{v}_{k-1}$所计算出，即：

$$\vec{v}_k = -\vec{g}_k + \gamma_k \vec{v}_{k-1} \tag{3-19}$$

$\gamma_k$ 为一纯量，其值为：

$$\gamma_k = \frac{\vec{g}_k \cdot \vec{g}_k}{\vec{g}_{k-1} \cdot \vec{g}_{k-1}} \tag{3-20}$$

各步的梯度与方向向量间满足如下的关系式，即：

$$\begin{aligned} &\vec{g}_i \cdot \vec{g}_j = 0 \\ &\vec{v}_i \cdot U''_{ij} \cdot \vec{v}_j = 0 \\ &\vec{g}_i \cdot \vec{v}_j = 0 \end{aligned} \tag{3-21}$$

以函数 $x^2+2y^2$ 示例如下。由 $\vec{x}_1=(9.0,9.0)$ 点及最速下降法可找出第一个沿着负的梯度方向 $-\vec{g}_1=(-18,-36)$ 的极小值点为 $\vec{x}_2=(4.0,$

$-1.0)$。此点的负梯度为$-\vec{g}_2=(-8.0,4.0)$。代入式(3-19) 得：

$$\vec{v}_2=\begin{pmatrix}-8\\4\end{pmatrix}+\frac{(-8)^2+(4)^2}{(-18)^2+(-36)^2}\begin{pmatrix}-18\\-36\end{pmatrix}=\begin{pmatrix}-8.889\\2.222\end{pmatrix} \tag{3-22}$$

由 $\vec{x}_2$ 沿 $\vec{v}_2$ 方向利用线性寻找法所找到的最小值点为 $(0,0)$，即函数的最小值点。

## 3.5 二次导数求极值法

与一次导数求极值的方法相比较，利用函数的二次微分求极小值的方法较为复杂，但通常可得到较为准确的结果。常用的二次导数法为牛顿-拉森法(Newton-Raphson method)。这种方法必须计算所需要矩阵，计算量较大，因此发展出一些简化的方法，如近似牛顿-拉森法（Quasi Newton-Raphson method）及对角线块状牛顿-拉森法（Block-diagonal Newton-Raphson method）等。

### 3.5.1 牛顿-拉森法

利用函数的一次及二次导数求最小值的方法中，最简单的是牛顿-拉森法。

一元多次函数 $U(x)$ 对 $x_k$ 点的二次泰勒展开式为：

$$U(x)=U(x_k)+(x-x_k)U'(x_k)+(x-x_k)^2U''(x_k)/2+\cdots \tag{3-23}$$

将式(3-23) 对 $x$ 微分得 $U(x)$ 的一次导数为：

$$U'(x)=U'(x_k)+(x-x_k)U''(x_k) \tag{3-24}$$

在极小值点，$x=x^*$ 函数的一次微分为零，$U'(x^*)=0$。将此关系代入式(3-24)，得：

$$U'(x^*)=0=U'(x_k)+(x^*-x_k)U''(x_k)$$

$$x^*=x_k-\frac{U'(x_k)}{U''(x_k)} \tag{3-25}$$

若函数为多元函数，则可得类似的关系式，即：

$$\vec{x}^*=\vec{x}_k-\tilde{U}''(\vec{x}_k)^{-1}\tilde{U}'(\vec{x}_k) \tag{3-26}$$

式中，$\tilde{U}''(\vec{x}_k)^{-1}$ 为 Hessian 矩阵的反矩阵。以函数 $f(x,y)=x^2+2y^2$ 为例，其 Hessian 矩阵为：

$$\tilde{f}''=\begin{pmatrix}\dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\partial y}\\[2ex] \dfrac{\partial^2 f}{\partial y\partial x} & \dfrac{\partial^2 f}{\partial y^2}\end{pmatrix}=\begin{pmatrix}2 & 0\\0 & 4\end{pmatrix} \tag{3-27}$$

其反矩阵为：

$$\tilde{f}''^{-1}=\begin{pmatrix}1/2 & 0\\ 0 & 1/4\end{pmatrix} \tag{3-28}$$

故依式(3-26)，若由 (9.0,9.0) 开始寻找函数 $f(x,y)$ 的最小值点，得：

$$\vec{x}^{*}=\begin{pmatrix}9.0\\ 9.0\end{pmatrix}-\begin{pmatrix}1/2 & 0\\ 0 & 1/4\end{pmatrix}\begin{pmatrix}18\\ 36\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix} \tag{3-29}$$

此点即此函数的最小值点。牛顿-拉森法利用一次及二次导数，起始位置离最小值点愈近则效果愈好。此方法因为需要求矩阵的反矩阵，因此需要较多的储存空间。此外，Hessian 矩阵的本征值必须全为正数，否则会导致移到能量较高点的情形发生。牛顿-拉森法因其计算量较大，一般适用于较小的分子，通常应用于求约含 100 个原子的分子势能最低的构象。

### 3.5.2 近似牛顿-拉森法

牛顿-拉森法最困难处在于计算 Hessian 矩阵的反矩阵。此步骤往往非常费时，而且有时计算函数的二次微分亦有相当困难。近似牛顿-拉森法则利用连续的计算法得到 Hessian 矩阵的反矩阵。假设一系列矩阵 $\tilde{H}_k$ 其极限为 Hessian 矩阵的反矩阵，即：

$$\lim_{k\to\infty}\tilde{H}_k=(\tilde{U}'')^{-1} \tag{3-30}$$

设由结构 $\vec{x}_k$ 开始，新的位置可由 $\vec{x}_k$ 的梯度与 $k$ 次的 Hessian 矩阵的近似矩阵 $\tilde{H}_k$ 求得。

$$\vec{x}_{k+1}=\vec{x}_k-\tilde{H}_k\vec{g}_k \tag{3-31}$$

下面分别列出常用的三种求 $\tilde{H}_k$ 的方法，以供参考。

#### (1) Davidon-Fletcher-Powell (DFP) 方法

$$\tilde{H}_{k+1}=\tilde{H}_k+\frac{(\vec{x}_{k+1}-\vec{x}_k)\otimes(\vec{x}_{k-1}-\vec{x}_k)}{(\vec{x}_{k+1}-\vec{x}_k)\cdot(\vec{g}_{k+1}-\vec{g}_k)}-\frac{[\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)]\otimes[\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)]}{(\vec{g}_{k+1}-\vec{g}_k)\cdot\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)} \tag{3-32}$$

式中，“⊗”的符号表示两向量相乘形成矩阵。例如 $(\vec{u}\otimes\vec{v})$ 矩阵的 $i$ 列 $j$ 行元素为 $\vec{u}$ 的第 $i$ 个分量乘以 $\vec{v}$ 的第 $j$ 个分量。

#### (2) Broyden-Fletcher-Goldfarb-Shanno (BFGS) 方法

$$\tilde{H}_{k+1}=\tilde{H}_k+\frac{(\vec{x}_{k+1}-\vec{x}_k)\otimes(\vec{x}_{k+1}-\vec{x}_k)}{(\vec{x}_{k+1}-\vec{x}_k)\cdot(\vec{g}_{k+1}-\vec{g}_k)}-\frac{[\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)]\otimes[\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)]}{(\vec{g}_{k+1}-\vec{g}_k)\cdot\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)}+$$

$$[(\vec{g}_{k+1}-\vec{g}_k)\cdot\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)]\vec{u}\otimes\vec{u}$$

$$\vec{u}=\frac{\vec{x}_{k+1}-\vec{x}_k}{(\vec{x}_{k+1}-\vec{x}_k)\cdot(\vec{g}_{k+1}-\vec{g}_k)}-\frac{\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)}{(\vec{g}_{k+1}-\vec{g}_k)\cdot\tilde{H}_k\cdot(\vec{g}_{k+1}-\vec{g}_k)} \tag{3-33}$$

#### (3) Murtaigh-Sargent (MS) 方法

$$\tilde{H}_{k+1}=\tilde{H}_k+\frac{[(\vec{x}_{k+1}-\vec{x}_k)-\tilde{H}_k(\vec{g}_{k+1}-\vec{g}_k)]\otimes[(\vec{x}_{k+1}-\vec{x}_k)-\tilde{H}_k(\vec{g}_{k+1}-\vec{g}_k)]}{[(\vec{x}_{k+1}-\vec{x}_k)-\tilde{H}_k(\vec{g}_{k+1}-\vec{g}_k)]\cdot(\vec{g}_{k+1}-\vec{g}_k)} \tag{3-34}$$

在应用近似牛顿-拉森法时，将第一个 $\tilde{H}_1$ 设为单位矩阵，然后用上述方法求得下一个 $\tilde{H}_2$，由此继续求出一系列的 $\tilde{H}_k$。

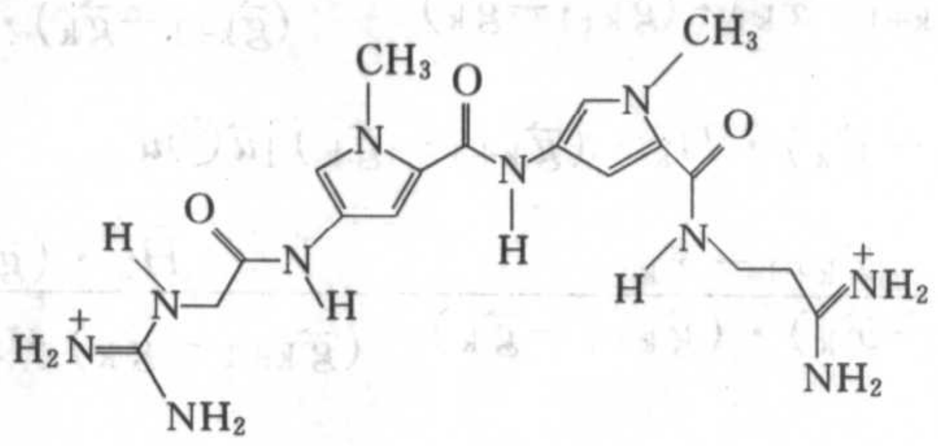
### 3.5.3 对角线块状牛顿-拉森法

除了上述的近似法外，还有一种常被采用的近似法为对角线块状牛顿-拉森法。以此方法求分子的最低能量构象时每次只移动一个原子。设移动的原子为 $k$，除了此原子外，所有其余的二次微分项，如 $\frac{\partial^2 U}{\partial x_i \partial x_j}, \frac{\partial^2 U}{\partial x_i \partial y_j}, \cdots$ $(i,j\neq k)$ 均等于零。此方法将原本为 $3N\times 3N$ 的 Hessian 矩阵简化为 $N$ 个3×3的矩阵。而求 3×3 矩阵的反矩阵远较求 $3N\times 3N$ 矩阵的反矩阵容易得多。有些分子中的原子间耦合性强，如苯环上的原子，则不能单独移动一个原子以求最低能量的构象。此时，对角线块状牛顿-拉森方法即不能适用。

## 3.6 求分子最低能量构象方法的选择

如前诸节所述，求分子的最低能量构象有多种可用的方法。选择适当的方法可以避免时间的浪费与节省计算机的储存量。以常用的抗生素分子为例。设分子的起始结构与其最低能量构象差异很大，求能量最小值的过程分成两阶段进行；第一阶段的能量最小化的目的在于粗略地找出低能量的分子结构，第二阶段的目的在于找出真正最低能量的分子结构。表 3-2

比较利用最速下降法与共轭梯度法求能量最小化结构的差异。



常用的抗生素分子

**表 3-2　利用最速下降法与共轭梯度法求能量最小化结构的差异比较**

| 项　目 | 第　一　阶　段 | | 第　二　阶　段 | |
|---|---|---|---|---|
| 方法 | 平均梯度 CPU 时间/s | ＜1kcal/Å² 计算次数 | 平均梯度 CPU 时间/s | ＜0.1kcal/Å² 计算次数 |
| 最速下降法 | 67 | 98 | 1405 | 1893 |
| 共轭梯度法 | 149 | 213 | 257 | 367 |

注：1kcal＝4.184kJ；1Å＝0.1nm。

表中的结果显示，若起始结构与最低能量结构相差很大，则利用最大倾斜法可以很快地找出粗略的低能量构象，此时的构象应与真正最低能量构象相去不远。但若欲由此精确地得到最低能量结构，则共轭梯度法应为较佳的选择。

实际上，利用分子力学求得分子真正最低能量构象的概率很小，通常只能得到与真正最低能量构象相近的结构。常用以决定计算是否该停止的方法有三种。第一种方法用能量差判别，例如设能量差的判别值为 $\epsilon$，则当 $n$ 次计算与前次 $n-1$ 计算的能量差符合 $\Delta U=U_n-U_{n-1}<\epsilon$ 的条件时计算停止。第二种方法利用原子坐标移动的幅度判别，当原子坐标更动的幅度符合 $\Delta x^2=|\vec{x}_n-\vec{x}_{n-1}|^2<\epsilon$ 的条件时中止计算。第三种方法利用计算的梯度判别，梯度的均方根误差（root-mean-square deviation，RMS）为：

$$\mathrm{RMS}=\sqrt{\frac{\vec{g}^T\vec{g}}{3N}} \tag{3-35}$$

若此偏差符合 $\mathrm{RMS}<\epsilon$ 的条件则计算终止。但通常需检查梯度向量的最大值，以免不准度集中在某个特别的坐标。

## 3.7　能量最小化的应用

分子力学计算的能量，相当于分子在假想的无动能的 0K 状态。但在

某一温度下，分子具有移动、转动与振动等运动。因此，欲正确地比较分子力学计算与实验度量的结果，必须加上一些校正项。通常将分子在热力学温度 $T$ 时的能量 $U(T)$ 可写为：

$$U(T)=U_{\text{trans}}(T)+U_{\text{rot}}(T)+U_{\text{vib}}(T)+U_{\text{vib}}(0) \tag{3-36}$$

式中，$U_{\text{trans}}$为移动能量；$U_{\text{rot}}$为转动能量；$U_{\text{vib}}$为振动能量；$U_{\text{vib}}(0)$为分子处于 0K 时的振动能，称为零点能量（zero point energy）。在一般的温度下（如 300K），分子的移动能与转动能可利用统计力学的均分理论（equipartition theory）估计，此两种能量分别为 $U_{\text{trans}}(T)=\frac{3}{2}k_{\text{B}}T$，$U_{\text{rot}}(T)=\frac{3}{2}k_{\text{B}}T$（非线性分子）或 $U_{\text{rot}}(T)=k_{\text{B}}T$（线性分子）；但欲计算分子的振动能则需确切地知道分子的振动频率。非线性分子和线性分子分别有 $3N-6$ 种及 $3N-5$ 种振动模式（vibration mode）。若将每种振动模式视为简谐振子，其所对应的振动频率为 $\upsilon_{\text{i}}$，则依照统计力学得振动的能量为：

$$U_{\text{vib}}(T)=\sum_{i=1}^{3N-6\text{或}3N-5}\left[\frac{h\upsilon_{\text{i}}}{2}+\frac{h\upsilon_{\text{i}}}{\exp\left(\frac{h\upsilon_{\text{i}}}{k_{\text{B}}T}\right)-1}\right] \tag{3-37}$$

式中，$h$ 为普朗克常数；$k_{\text{B}}$ 为玻耳兹曼常数。分子的零点能量为各振动模式最低能量的和：

$$U_{\text{vib}}(0)=\sum_{i=1}\frac{1}{2}h\upsilon_{\text{i}} \tag{3-38}$$

不同的分子具有不同的振动频率与零点能量。表 3-3 列出甲烷与水分子的振动频率与零点能量。表中，振动频率的括号为该振动模式的简并度（degeneracy）。

**表 3-3　甲烷与水分子的振动频率与零点能量**

| 分　子 | 振动频率/$cm^{-1}$ | 零点能量/$cm^{-1}$ |
|---|---|---|
| 甲烷($CH_4$) | 2895,1513(2),2999(3),1298(3) | 9406 |
| 水($H_2O$) | 3721,3582,1590 | 4447 |

## 3.8　多原子分子振动频率的计算

以简谐振子的方法描述分子振动的坐标称为简正坐标（normal coordinate），此坐标为分子内原子移动的组合。各种振动模式称为简正模式

(normal mode) 或称为正则振动，所对应的振动频率称为简正频率 (normal frequency)。图 3-9 为水分子的简振模式与其对应的简正频率。

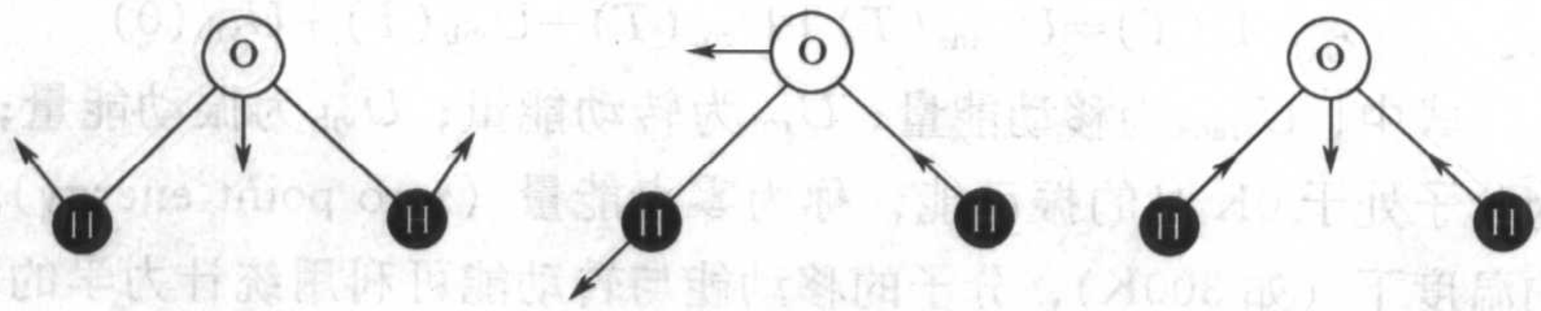



**图 3-9　水分子的简正模式与简正频率**

简正频率可利用分子的 Hessian 矩阵计算。因简正振动与原子的质量相关，故计算振动频率时需先将原子在笛卡儿坐标系的位移转换为质权坐标位移 (mass-weighted displacement)。质权位移的定义为：

$$q_1=\frac{1}{\sqrt{m_1}}\Delta x_1,\ q_2=\frac{1}{\sqrt{m_1}}\Delta y_1,\ q_3=\frac{1}{\sqrt{m_1}}\Delta z_1,\ \cdots,\ q_{3N}=\frac{1}{\sqrt{m_N}}\Delta z_N \tag{3-39}$$

式中，$m_1,\cdots,m_N$ 为分子中第 1 个，…，第 $N$ 个原子的质量；$\Delta x_1$，$\Delta y_1$，$\Delta z_1$ 为第 1 个原子在 $x,y,z$ 方向的位移。将笛卡儿坐标系的 Hessian 矩阵转换为质权坐标系的力矩阵的关系式为：

$$\widetilde{F}_q=\widetilde{M}^{-1/2}\widetilde{U}''\widetilde{M}^{-1/2} \tag{3-40}$$

式中，$\widetilde{M}^{-1/2}$ 的形式为：

$$\widetilde{M}^{-1/2}=\begin{pmatrix} m_1^{-1/2} & & & & & & \\ & m_1^{-1/2} & & & & & \\ & & m_1^{-1/2} & & & & \\ & & & m_2^{-1/2} & & & \\ & & & & m_2^{-1/2} & & \\ & & & & & m_2^{-1/2} & \\ & & & & & & \ddots \\ & & & & & & & m_N^{-1/2} & \\ & & & & & & & & m_N^{-1/2} & \\ & & & & & & & & & m_N^{-1/2} \end{pmatrix} \tag{3-41}$$

矩阵中非对角线元素均为零。求 $\widetilde{F}_q$ 矩阵的本征值即相当于解特征方程

式 $\widetilde{F}_q\widetilde{A}=\widetilde{\lambda}\widetilde{A}$；其本征值可由矩阵的对角线化方法求出，或者可利用解行列式 $|F_q-\lambda I|=0$ 求得；其中 $\lambda$ 为本征值，$I$ 为单位矩阵。以此方法求解，非线性分子有 6 个本征值为零，线性分子则有 5 个本征值为零。这些为零的本征值对应于分子的移动与转动运动。简正频率 $\upsilon_i$ 与本征值 $\lambda_i$ 的关系为：

$$\upsilon_i=\sqrt{\frac{\lambda_i}{2\pi}} \tag{3-42}$$

以图 3-10 所示三原子的线性分子（A-B-A 型）例示简正频率的计算。设只考虑沿分子轴方向的振动。

**图 3-10　A-B-A 形线性分子示意**

图中，原子 1，2 与 2，3 间的平衡键长均为 $l_0$，原子离其平衡位置的位移为 $\xi_i$。设每个原子的位移与键长相比皆很小，故振动符合虎克定律（Hook's law）。此分子的振动势能为：

$$U=\frac{1}{2}k(\xi_1-\xi_2)^2+\frac{1}{2}k(\xi_2-\xi_3)^2 \tag{3-43}$$

其势能函数对坐标的一次微分为：

$$\frac{\partial U}{\partial \xi_1}=k(\xi_1-\xi_2)$$

$$\frac{\partial U}{\partial \xi_2}=-k(\xi_1-\xi_2)+k(\xi_2-\xi_3)$$

$$\frac{\partial U}{\partial \xi_3}=-k(\xi_2-\xi_3) \tag{3-44}$$

由势能函数的二次微分，得 Hessian 矩阵，即：

$$\widetilde{U}''=\begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \tag{3-45}$$

将此矩阵转换为质权坐标系的力常数矩阵，即：

$$\begin{aligned}\widetilde{F}_q&=\begin{pmatrix} m_A^{-1/2} & 0 & 0 \\ 0 & m_B^{-1/2} & 0 \\ 0 & 0 & m_A^{-1/2} \end{pmatrix}\begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}\begin{pmatrix} m_A^{-1/2} & 0 & 0 \\ 0 & m_B^{-1/2} & 0 \\ 0 & 0 & m_A^{-1/2} \end{pmatrix} \\ &=\begin{pmatrix} \dfrac{k}{m_A} & -\dfrac{k}{\sqrt{m_A}\sqrt{m_B}} & 0 \\ -\dfrac{k}{\sqrt{m_A}\sqrt{m_B}} & \dfrac{2k}{m_B} & -\dfrac{k}{\sqrt{m_B}\sqrt{m_A}} \\ 0 & -\dfrac{k}{\sqrt{m_B}\sqrt{m_A}} & \dfrac{k}{m_A} \end{pmatrix}\end{aligned} \tag{3-46}$$

此矩阵的本征值可由解行列式 $|F_q-\lambda I|=0$ 求得，即：

$$\begin{vmatrix} \frac{k}{m_A}-\lambda & -\frac{k}{\sqrt{m_A}\sqrt{m_B}} & 0 \\ -\frac{k}{\sqrt{m_A}\sqrt{m_B}} & \frac{2k}{m_B}-\lambda & -\frac{k}{\sqrt{m_B}\sqrt{m_A}} \\ 0 & -\frac{k}{\sqrt{m_B}\sqrt{m_A}} & \frac{k}{m_A}-\lambda \end{vmatrix}=0 \qquad (3\text{-}47)$$

将此行列式展开得：

$$\left(\frac{k}{m_A}-\lambda\right)\left(\frac{2k}{m_B}-\lambda\right)\left(\frac{k}{m_A}-\lambda\right)-\frac{2k^2}{\sqrt{m_A}\sqrt{m_B}}\left(\frac{k}{m_A}-\lambda\right)=0 \qquad (3\text{-}48)$$

其方程式的三个根分别为：

$$\lambda_1=\frac{k}{m_A},\ \lambda_2=0,\ \lambda_3=\frac{2m_A+m_B}{m_A m_B}k \qquad (3\text{-}49)$$

由此本征值，可依式(3-42) 计算其分别对应的振动频率。由解特征方程式 $F_qA=\lambda A$，可得与各本征值对应的分子振动时原子的位移幅度。设 $A_1$，$A_2$ 与 $A_3$ 分别表示三个原子位移的幅度，则可得：

$$\lambda_1=\frac{k}{m_A}: \qquad A_1=-A_3;\ A_2=0$$

$$\lambda_2=0: \qquad A_1=A_3; \qquad A_2=\sqrt{\frac{m_B}{m_A}}A_1$$

$$\lambda_3=k\frac{m_B+2m_A}{m_A m_B}: \quad A_1=A_3; \qquad A_2=-2\sqrt{\frac{m_B}{m_A}}A_1 \qquad (3\text{-}50)$$

将各原子的振动幅度以图形表示，可得分子的振动模式。图 3-11 显示以上的计算结果。

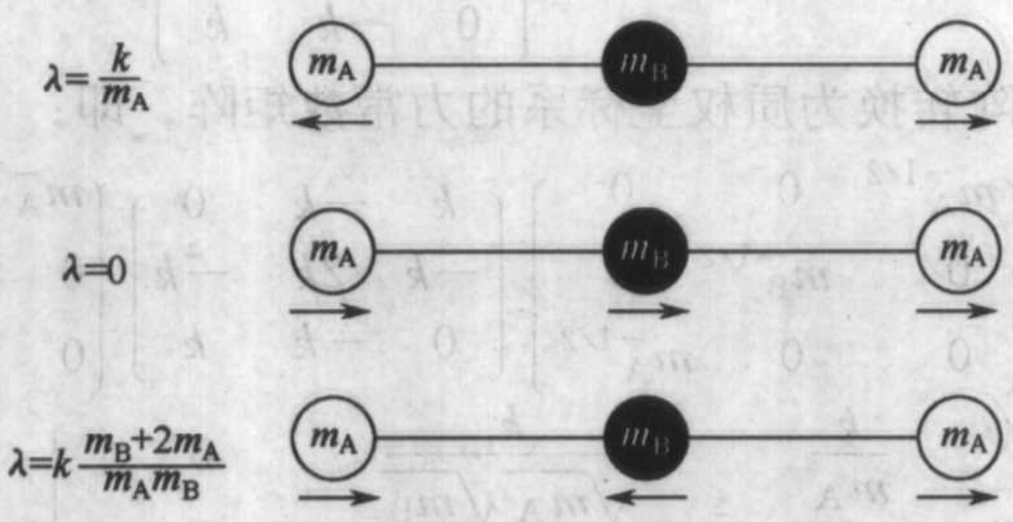



**图 3-11　线性分子的简振模式**

依图 3-11 可知此线性分子与本征值对应的三种简正模式分别为对称伸缩（symmetric stretch）、移动（translation）与非对称伸缩（asymmetric stretch）。

利用多原子分子的振动频率可计算分子的热力学性质，所计算的振动频率可与实验相对比以检验力场的精确性。值得一提的是，利用分子力场所计算的多原子分子的振动频率往往较量子力学方法所计算的振动频率更接近实验值，但分子力场法所需要的计算时间却远远小于任何量子力学方法所需的时间。图 3-12 显示 Brooks 等利用分子力场计算螺旋状多肽分子（由 10 个丙氨酸所组成）所得的简正频率的分析。此分子含有 112 个原子，若执行量子力学的计算需耗费大量的计算量。

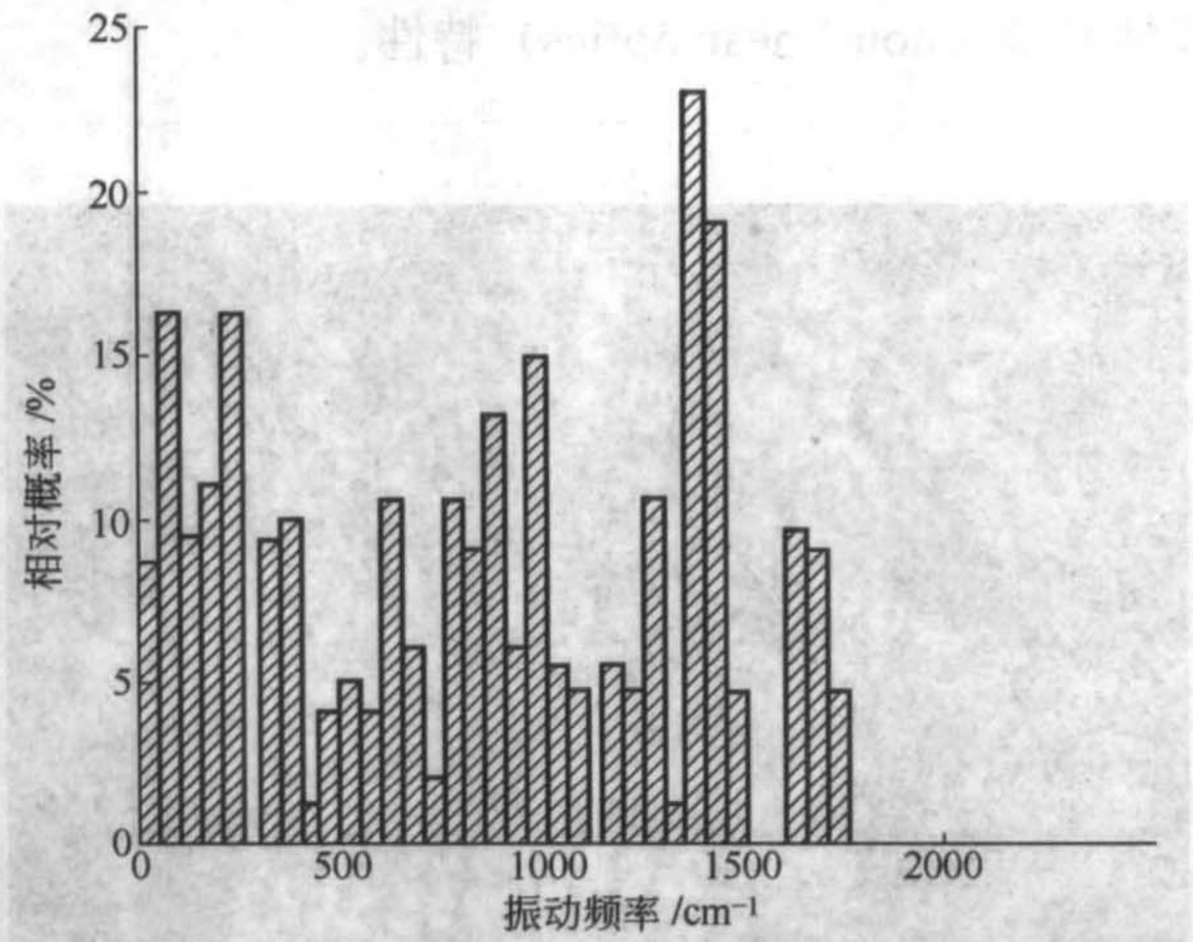



**图 3-12　利用分子力学计算多肽分子的简正（正则）振动分布图**

表 3-4 为 Allinger 等利用 MM3 力场计算所得的甲醇分子的振动频率与实验值的比较，表中显示计算的振动频率非常接近实验值。

**表 3-4　MM3 力场计算方法甲醇分子的振动频率与实验值的比较**

| 实验值/$cm^{-1}$ | MM3 力场计算值/$cm^{-1}$ | 振　动　模　式 |
|---|---|---|
| 3682 | 3679 | O—H 伸缩振荡(str.) |
| 2977 | 2976 | $CH_3$ 非对称伸缩(asym. str.) |
| | 2970 | |
| 2844 | 2873 | $CH_3$ 对称伸缩(sym. str.) |
| 1477 | 1484 | $CH_3$ 扭曲(def.) |
| 1455 | 1446 | |
| 1430 | 1430 | |
| 1340 | 1287 | C—O—H 弯曲(bend) |
| | 1106 | $CH_3$ 摆动(rock.) |
| | 1086 | $CH_3$ 摇摆(wag.) |
| 1034 | 1052 | C—O 伸缩(str.) |
| 270 | 263 | $CH_3$—O(tor.) |

## 3.9 以能量最小化方法预测团簇体的大小

自组装团簇体（self-assembled cluster）多由具有共轭片段的有机分子组装而成，通常为纳米尺寸，常用以制造各种形状的光电原件。这种纳米级的团簇体结构复杂，但往往具备非常特殊的性能，为近年来全世界科学界及工业界所发展的重要研究领域。图 3-13 为 Stuppe 等所提出的自组装团簇体 HEMMH 的假想构造。这种团簇体呈香菇状的结构，具有非常明显的非线性光学（non-linear optics）特性。

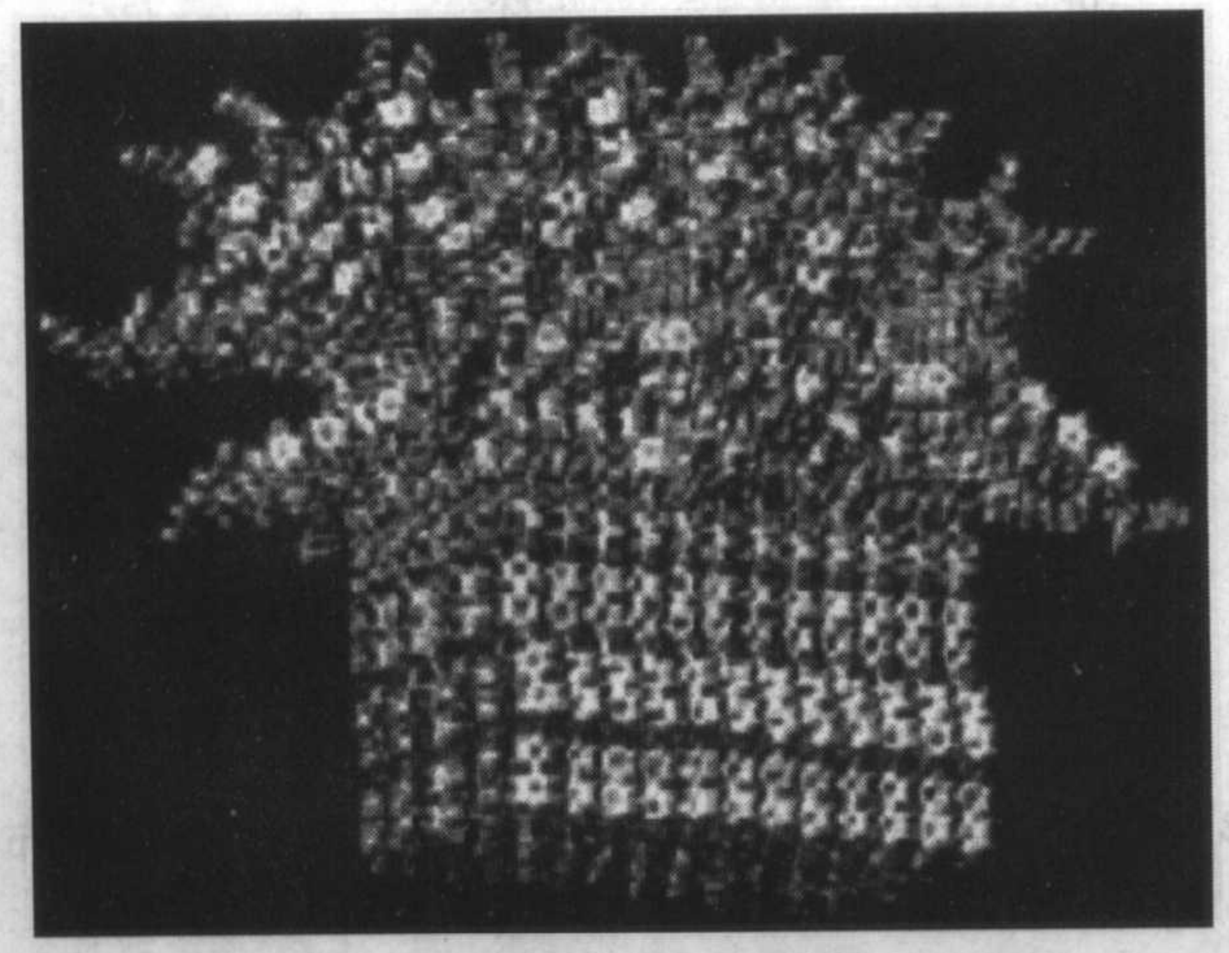


**图 3-13　HEMMH 团簇体的构型示意**

图 3-14 及图 3-15 显示带状及其他一些可能的纳米自组装团簇体的构型。这些纳米自组装团簇体多用以制造光电材料的元件，具有非常特殊的功能。

HEMMH 分子的结构如图 3-16 所示。分子由两部分所构成，一部分为柔软的链状，称为卷线部分（coil）；另一部分由共轭的片段所构成，形状略似硬棒，称为杆状部分（rod）。这种特殊类型的分子通常称为“卷线-杆状”分子（coil-rod molecule）。

Stuppe 等利用分子模拟的能量最小化的方法，计算由几根 HEMMH 组成的小团簇体的构造，其结果如图 3-17 所示。图中显示，卷线-杆状分子 HEMMH 的杆状部分大体呈平行排列，其卷线部分则无明显的排列秩序。各种实验的结果显示，HEMMH 自组装团簇体的直径约为 2～5nm，

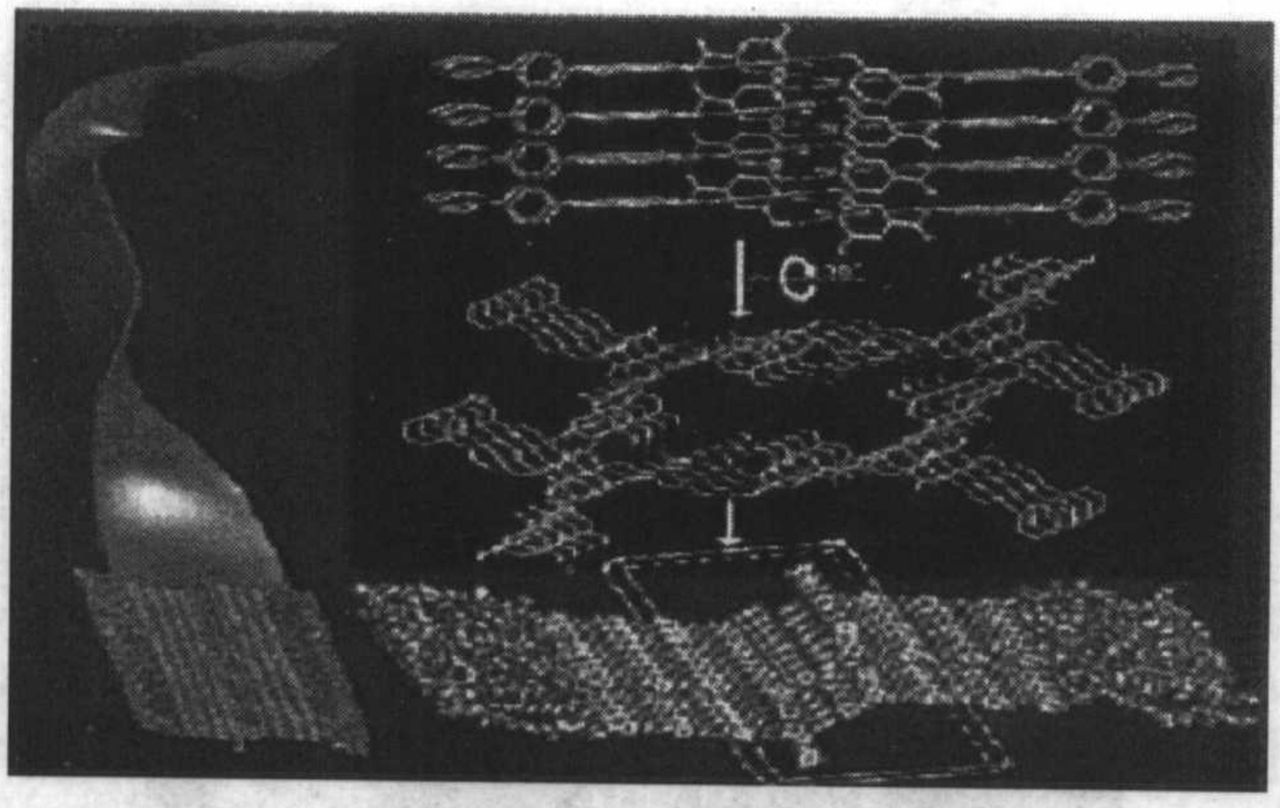


图 3-14　带状自组装团簇体

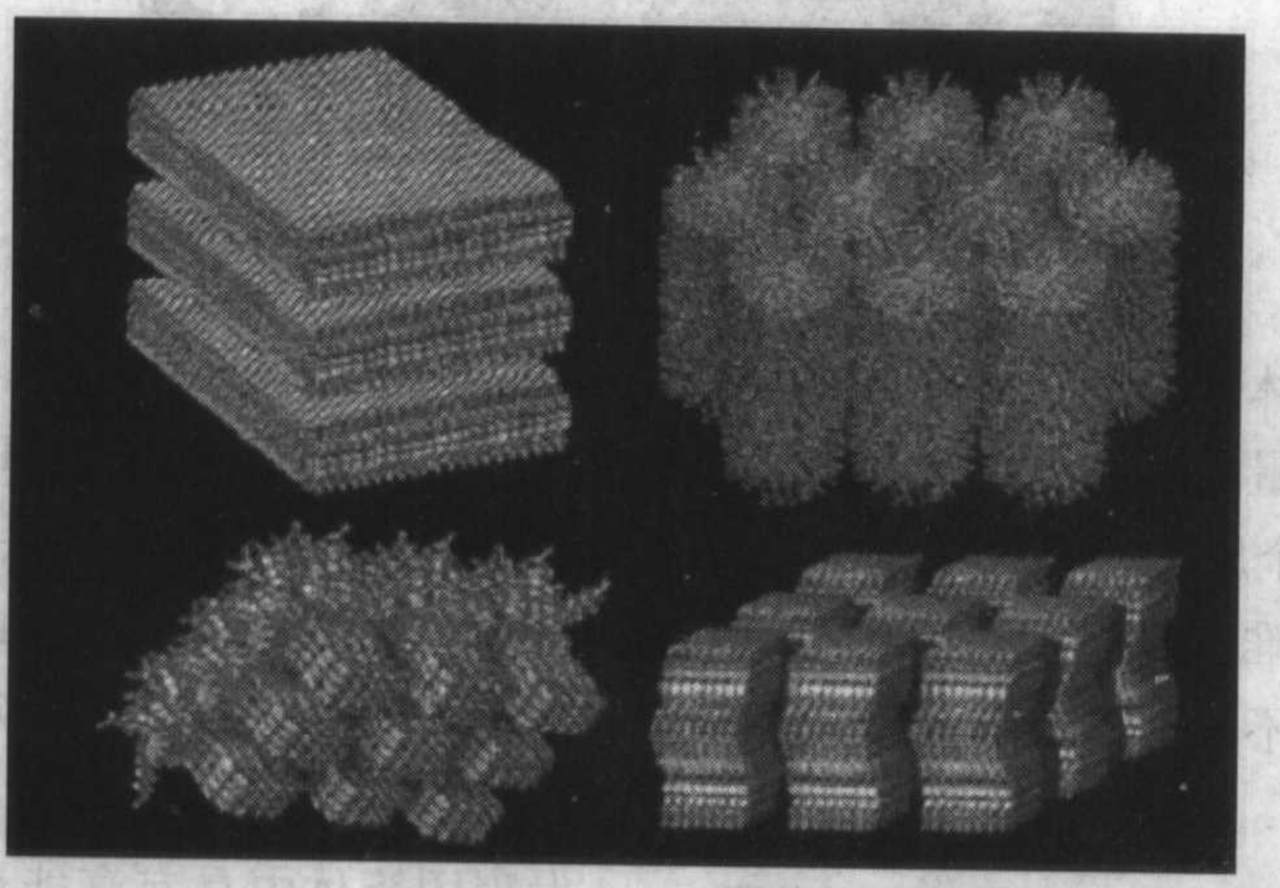


图 3-15　各种形状的自组装团簇体

“卷线-杆状”分子

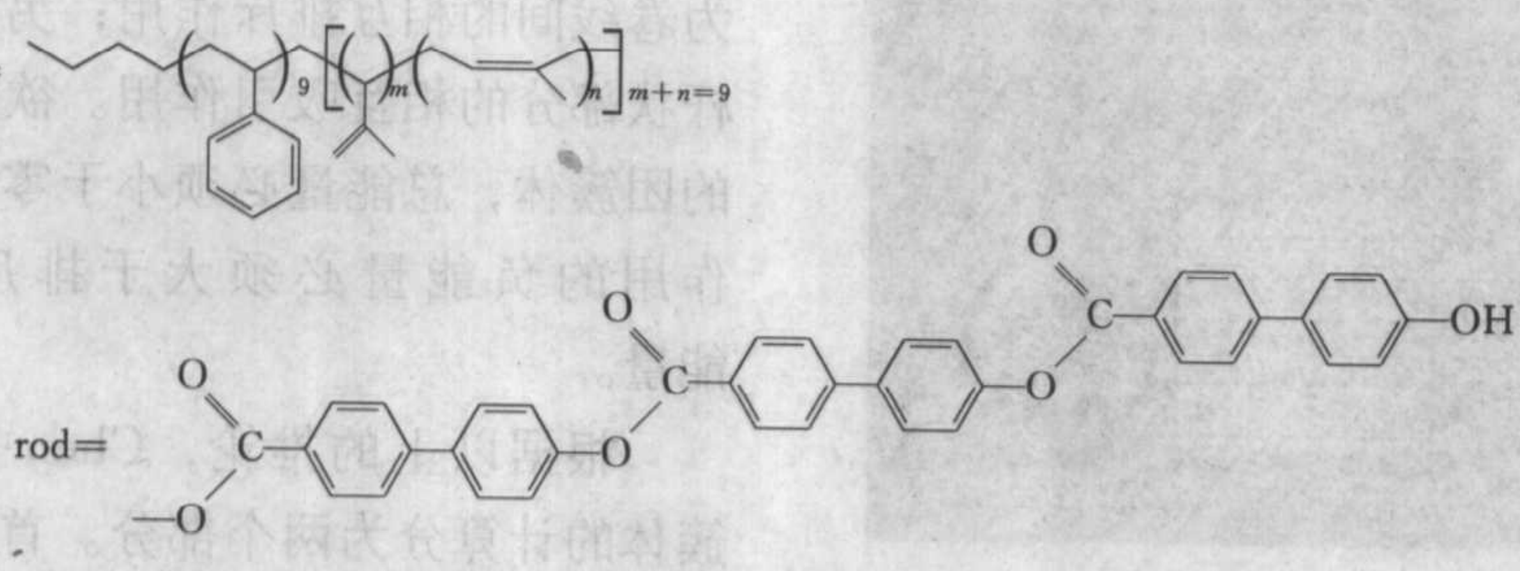



图 3-16　HEMMH 分子的结构式

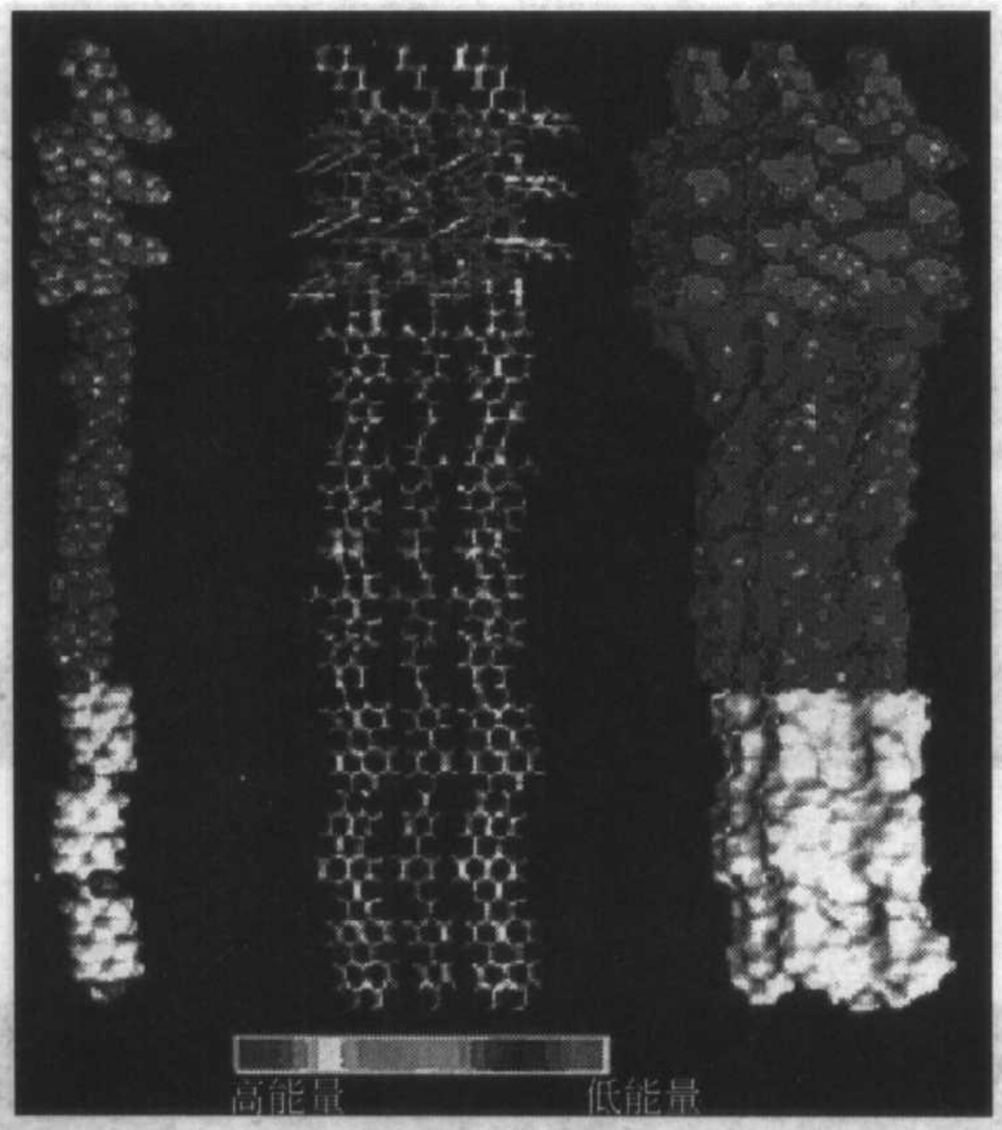



**图 3-17　分子力学计算 HEMMH 小团簇体的构造**

团簇体约由小于 100 根的卷线-杆状分子所组成。理论上，可利用分子模拟能量最小化的方法验证团簇体的大小，但实际上每个 HEMMH 分子含 350 个原子，计算系统非常庞大，使用能量最小化的方法遭遇许多难题。不但难以建立初始的构型，且最小化的构型往往落于区域极小值的范围内，不能得到合理的结果。因此，研究此系统必须采取特殊的方法。

分析 Stuppe 等的分子力学计算，发现团簇体中具有高能量区与低能量区两个部分。因此，可将分子间的作用分成两部分，如图 3-18 所示。其中一部分为卷线间的相互排斥作用；另一部分为杆状部分的相互吸引作用。欲形成稳定的团簇体，总能量必须小于零，即吸引作用的负能量必须大于排斥作用的能量。

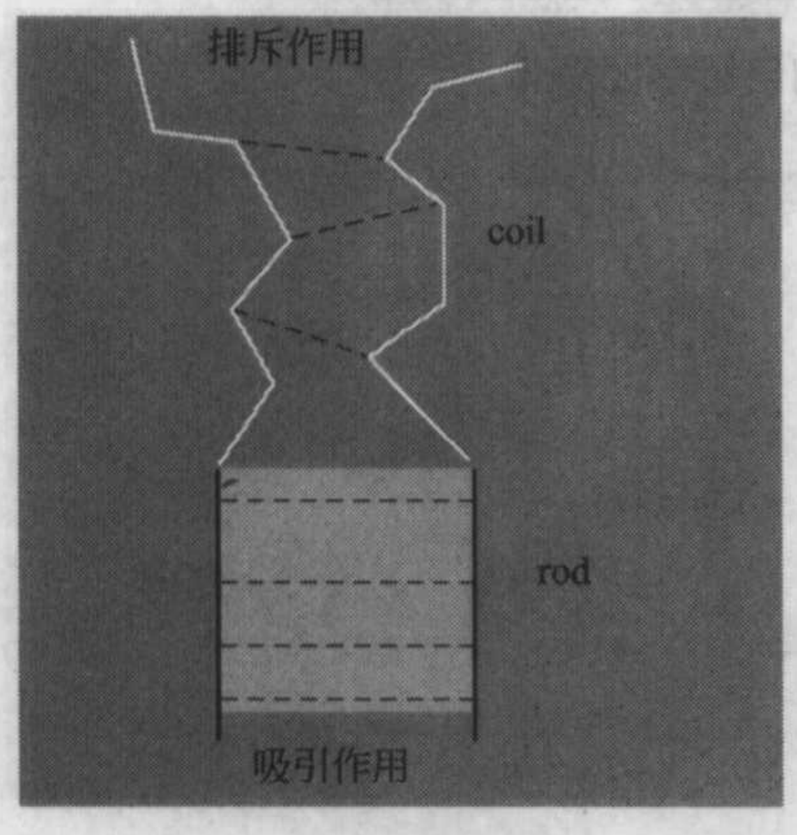



**图 3-18　HEMMH 分子间作用示意**

根据以上的推论，Chen 等将此团簇体的计算分为两个部分。首先，利用 CERIUS2 软件的 UNIVERSAL 力场，

以分子力学的能量最小化方法计算杆状部分的能量。图 3-19 为 Chen 等计算所得的杆状部分聚集体的能量与杆状分子数目的关系。图中显示，这一部分的能量随分子数目的增加而递增。

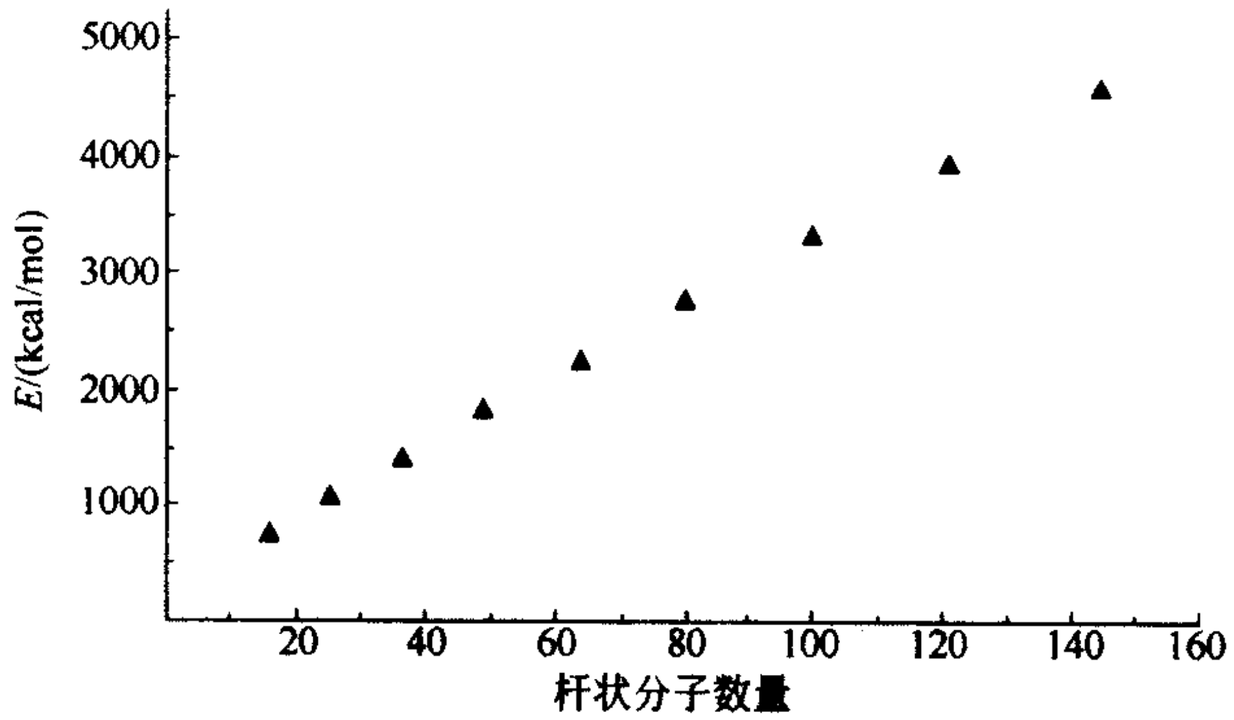



**图 3-19　杆状分子聚集体的能量**

图 3-20 为将杆状部分的总能量减去所有杆状分子单独的能量后的平均值，亦即杆状分子间的平均作用能。

$$\bar{E}_{\text{int}} = \frac{1}{n}\langle E - \sum_{i=1}^{n} E_i^0 \rangle \tag{3-51}$$

式中，$n$ 为杆状分子的数目；$E$ 为总能量；$E^0$ 为杆状分子单独的能量。图 3-20 显示杆状分子间的平均作用能随分子数目的增加而递减。

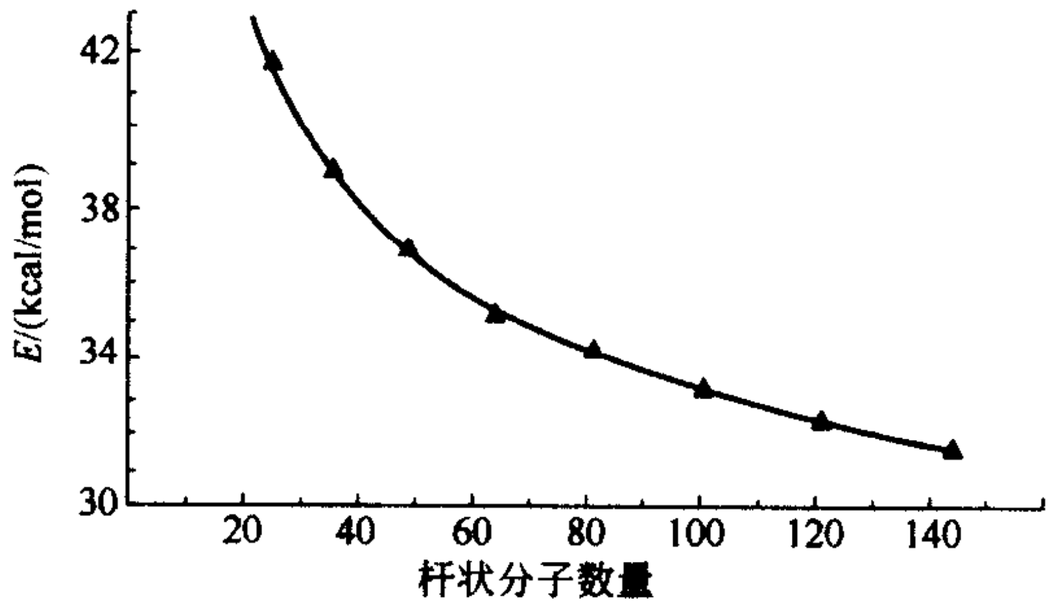



**图 3-20　杆状分子间的平均作用能**

检视计算所得的杆状分子聚集体的能量最小化构型，可知分子间约呈平行排列。将每一分子的端点固定，视为卷线分子部分的起点。第二部分

的计算为由此起点接上卷线部分，执行卷线部分聚集体的分子力学计算。如图 3-21 所示。图中箭头所指之处为杆状与卷线部分的连接处。

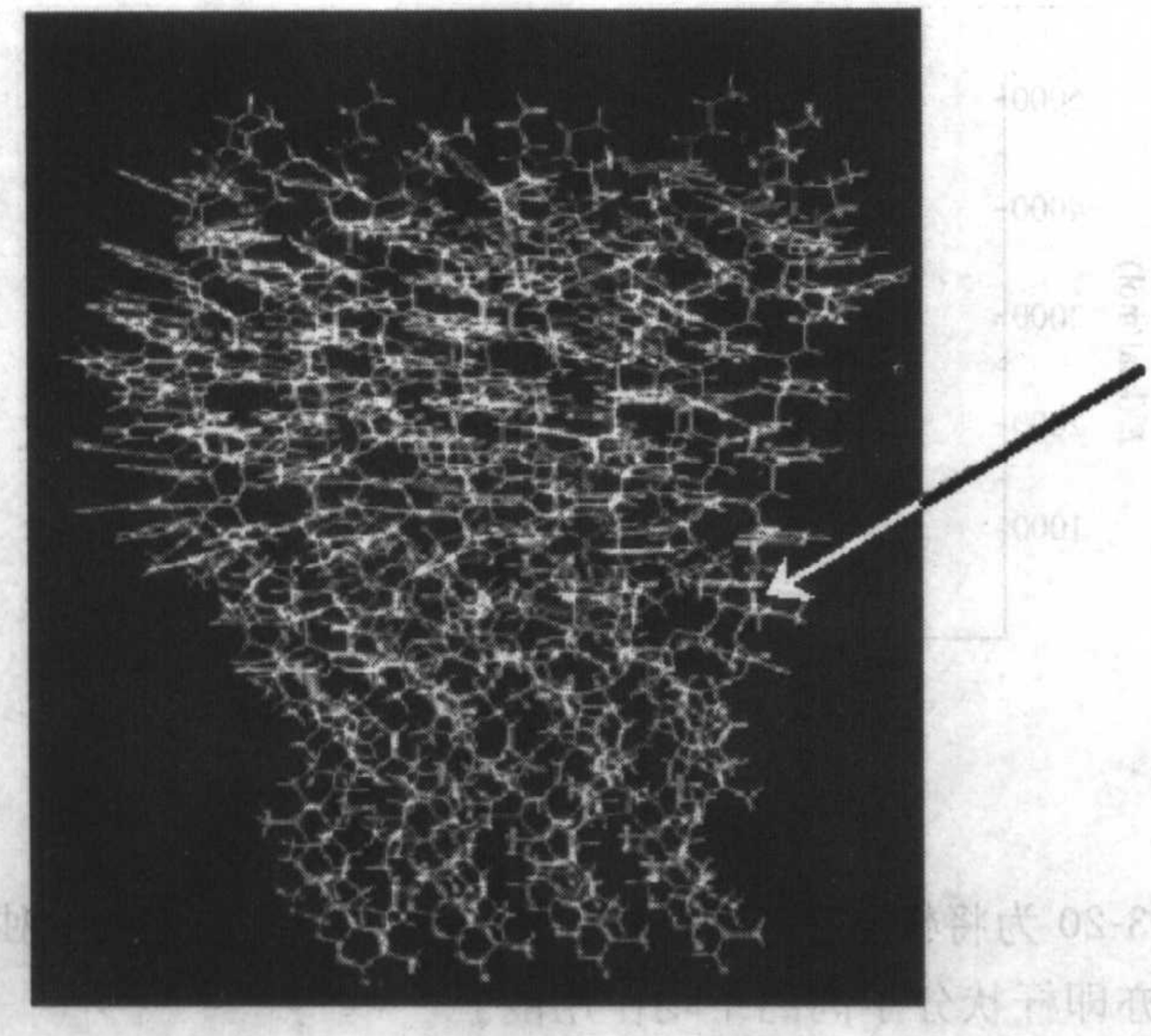


**图 3-21　杆状+ 卷线部分的计算方法**

固定杆状部分的端点，加入卷线部分

图 3-22 为卷线状分子聚集体的能量，由图可知能量随分子数目的增加而递增。

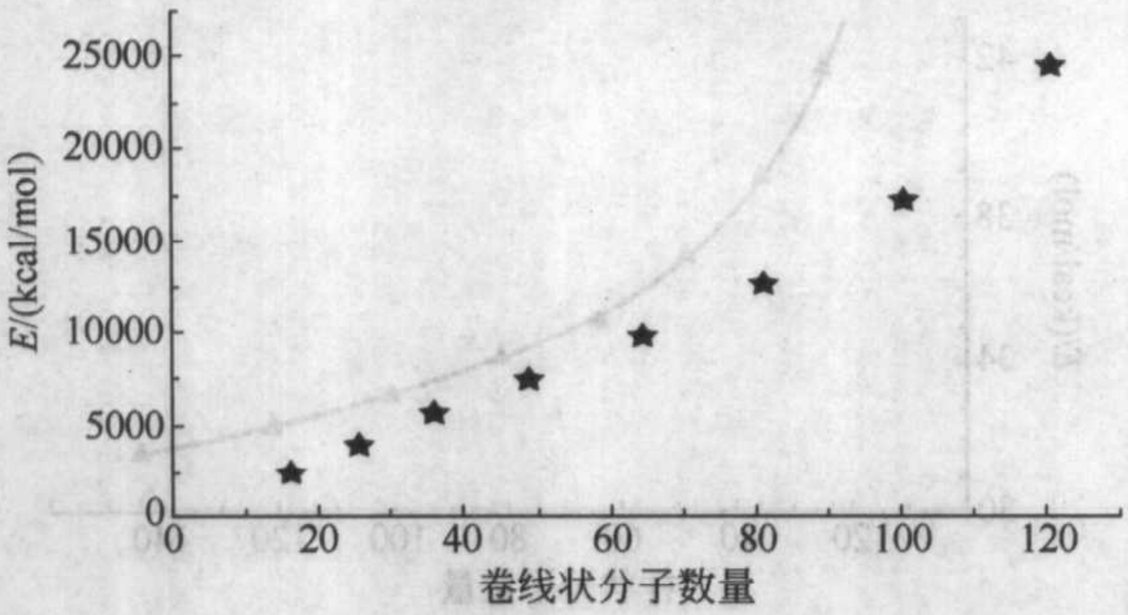



**图 3-22　卷线状分子聚集体的能量**

图 3-23 为卷线部分分子间的平均作用能，其计算方法见式(3-51)。将如上的杆状部分与卷线部分的计算结果合并，所得的分子间的平均

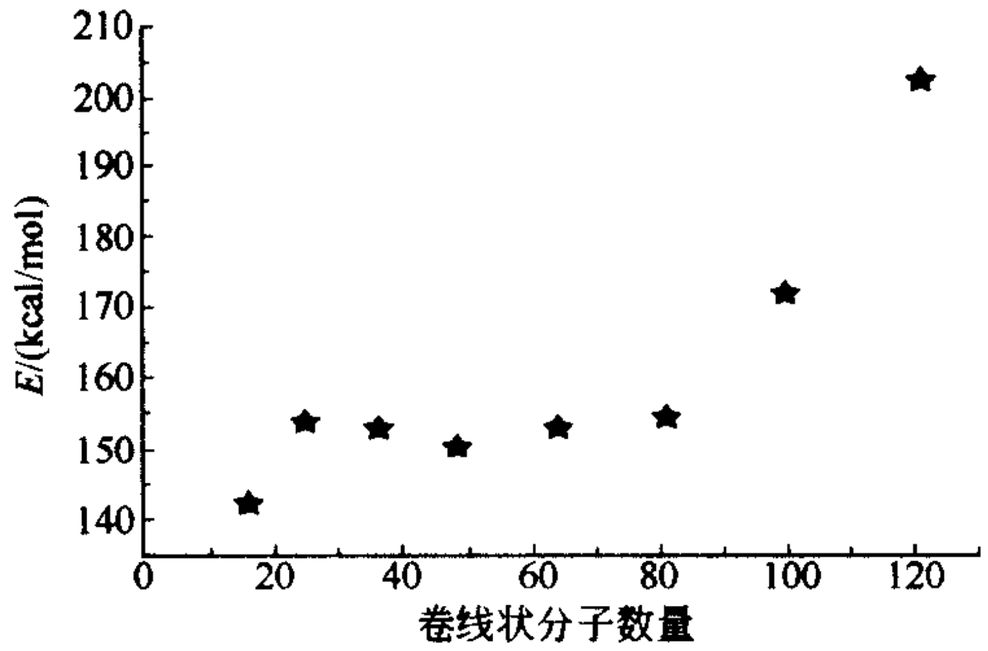



**图 3-23　卷线状分子间的平均作用能**

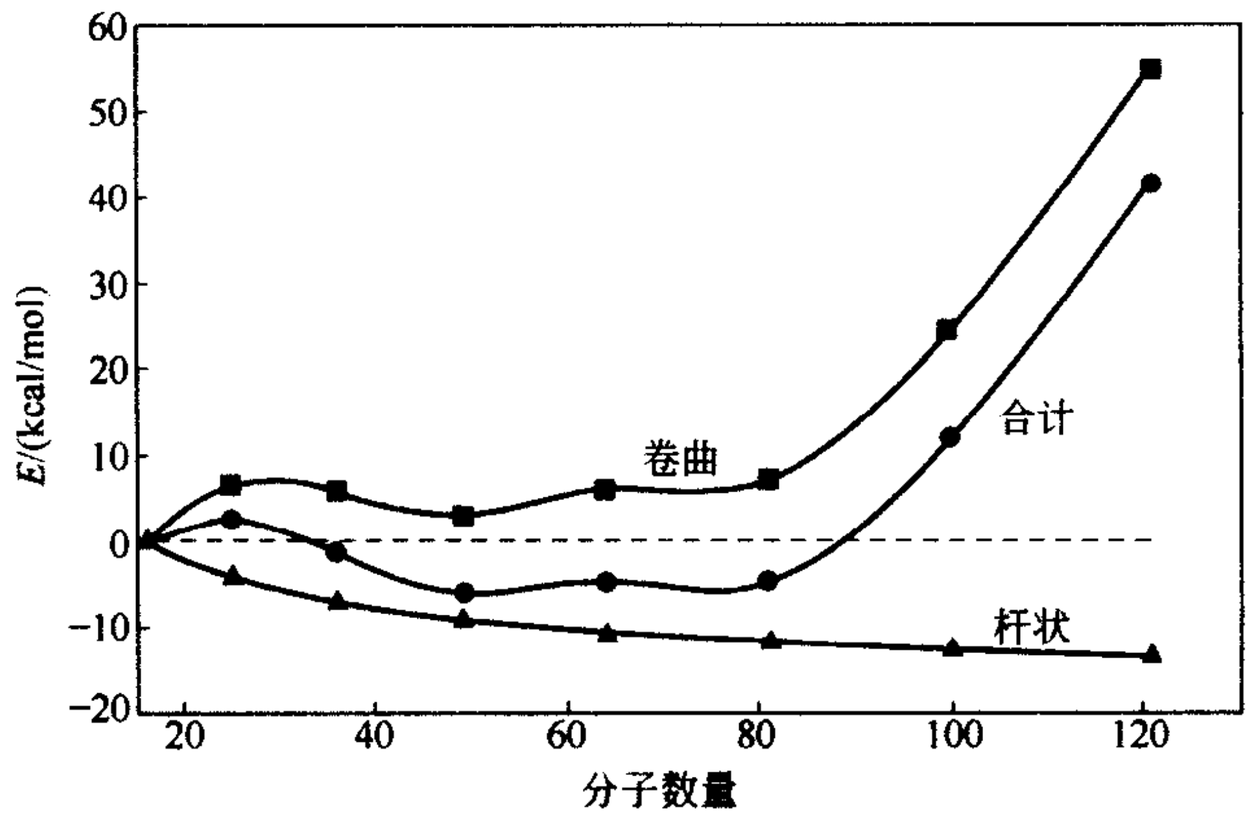



**图 3-24　HEMMH 分子间的平均作用能**

作用能如图 3-24。图中显示分子间的平均作用能于 $n=36\sim90$ 的区间范围内均小于零。由此计算结果可以判定 HEMMH 的团簇体大小，若配制的溶液较为稀薄，则易形成 36 根 HEMMH 链的团簇体；若配制的溶液浓度高，则易形成 90 根链的团簇体。此计算与实验所观测的结果非常吻合，是应用分子力学的成功例子。

图 3-25 为以分子力学能量最小化的构型为起点，执行 36 根链的团簇体的分子动力计算的结果。图中显示团簇体确实呈现香菇状的结构，其直径约为 50Å。计算的团簇体大小亦与实验所观测的结果相符合。

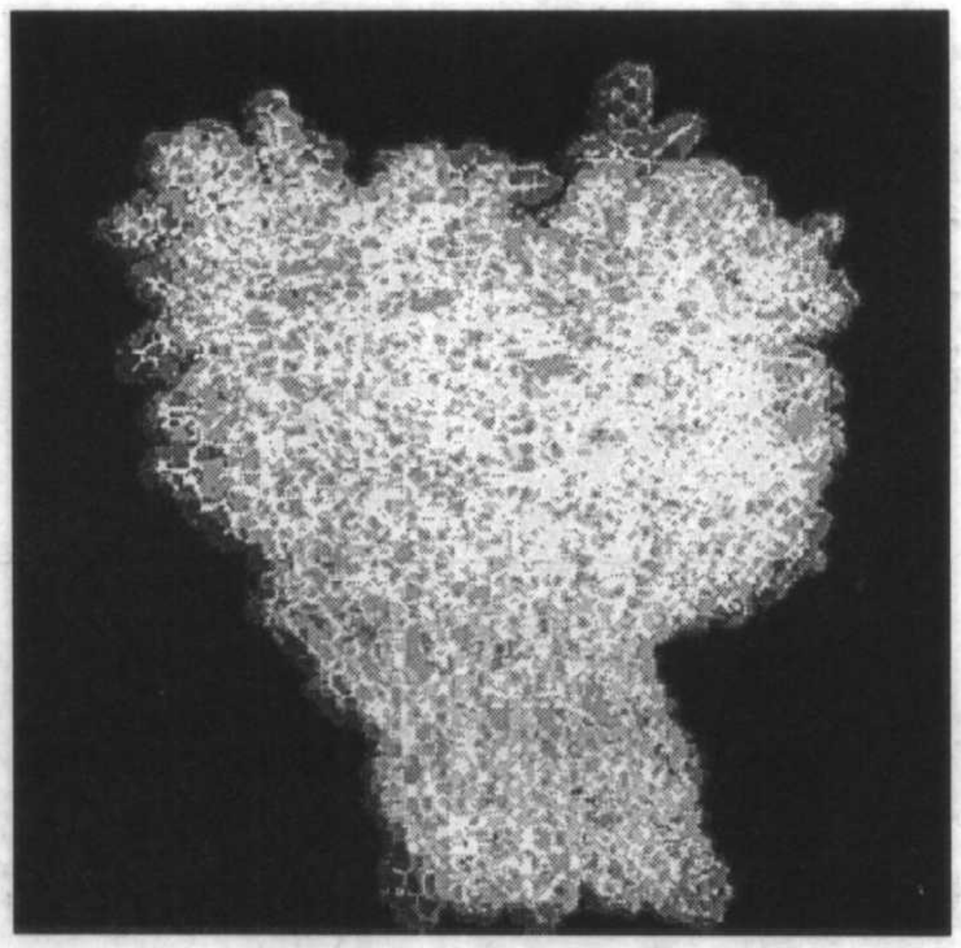


图 3-25　分子动力学计算的 HEMMH 聚集体的结构

## 参 考 文 献


1　Gill P E, Murray W. Practical Optimization, Academic Press, London, England, 1980

2　Allinger N L. Adv. Phys. Org. Chem., 13, 1, 1976

3　Wertz D H. Ph. D Thesis, Univ. of Georgia, 1974

4　White D N. J. Comput. Chem. 1, 225, 1977

5　Allinger N L, Tribble M T, Miller M A, Wertz D H. J. Am. Chem. Soc. 93, 1637, 1971

6　Brooks B, Karplus M. Proc. Nat. Acad. Sci. USA, 80, 6571, 1983

7　Allinger N L, Rahman M, Li J H. J. Am. Chem. Soc. 112, 8293, 1990

8　Stupp S I, LeBonheur V, Walker K, Li L S, Huggins K E, Keser M, Armstutz A. SCIENCE, 276, 384, 1997

9　Chen C L, Yang S C. 8th Nation. Symposium of Quant. Chem. Changchun, 2002

# 第❹章 分子力学的应用

前述两章介绍力场与能量最小化的方法，并举出利用能量最小化方法可获得的各种数据。本章介绍分子力学的模拟方法及分子力学的应用。分子力学的应用面很广，除了必须经由量子力学计算的电子波函数外，分子的各种特性均可经由分子力学计算得到。一般常应用分子力学方法计算分子的最低能量构型、构型转换能障、振动频率和各种热力学性质等。本章分两部分内容，4.1～4.3 节由分子的势能面着手，介绍计算势能面可获得的数据；4.4～4.6 节主要介绍分子力学方法在反应、热力学性质、晶体结构及立体效应各方面的应用。

## 4.1 分子结构与势能面

分子力学利用能量最小化的方法，由起始的结构找出附近区域的极小值结构。因为一般力场的势能函数均为分析加成形式，故可轻易地得到势能函数的一次导数与二次导数，加速最小化的过程。由势能函数的能量最小化可预测分子构象异构体（conformer）的平衡结构。由分子力学计算所得构象异构体平衡结构的能量 $U$ 称为空间位阻能量（steric energy）。$U=0$ 相当于分子间各键长与键角均处于其平衡位置，亦无扭转，范德瓦耳斯作用与静电作用等形式的能量。以不同的力场计算特殊分子的最稳定结构通常得到相近的结果。但是，由不同力场所得到的 $U$ 值却因各力场选定不同的力场参数及形式而彼此不同。事实上，具有明确物理意义的并非由分子力学所得到的 $U$ 值，而是两个不同构象空间位阻能量的差值。一般常利用位阻能量的差值计算比较以下各情形的能差：①同一分子不同构象的能差（如丁烷的反式构象与邻位交叉构象的能差，或环己烷的椅形构象

与船形构象的能差）；②分子空间异构体（stereoisomer）间的能差（如二氯乙烯顺式与反式异构体的能差）；③分子由于绕着某特定键旋转而形成的异构象的能差［如乙烷的重叠构象（eclipse）与对位交叉构象（staggered）的能差］；④相同分子但几何形状不同的能差（如 $NH_3$ 分子三角锥构象与平面构象的能差）；⑤距离远的两个分子与此二分子形成分子间氢键的能差。由分子力学对分子结构预估的准确性，多半获得合理的结果。如 Wertz 与 Allinger 曾计算丁烷的结构，显示反式构象（anti）的能量低于邻位交叉式构象（gauche）的能量。而 Osawa、Collins 与 Schleyer 引用非常不同的力场计算丁烷的各种构象的能量，亦得到相同的结论。

近年来分子力学方法已被广泛地用于探讨生化分子的结构。图 4-1 为双螺旋的核糖核酸（RNA）衍生物的片段。图中的五原子环可能的结构有两种：其中的 4 个原子共平面而第 5 个原子离开此平面的信封状结构（envelope），或是 3 个原子共平面其余 2 个相邻的原子在此平面的相反方向的扭曲结构（twist）。离平面的原子与 $C_{5'}$ 原子在平面的同一边称为 endo

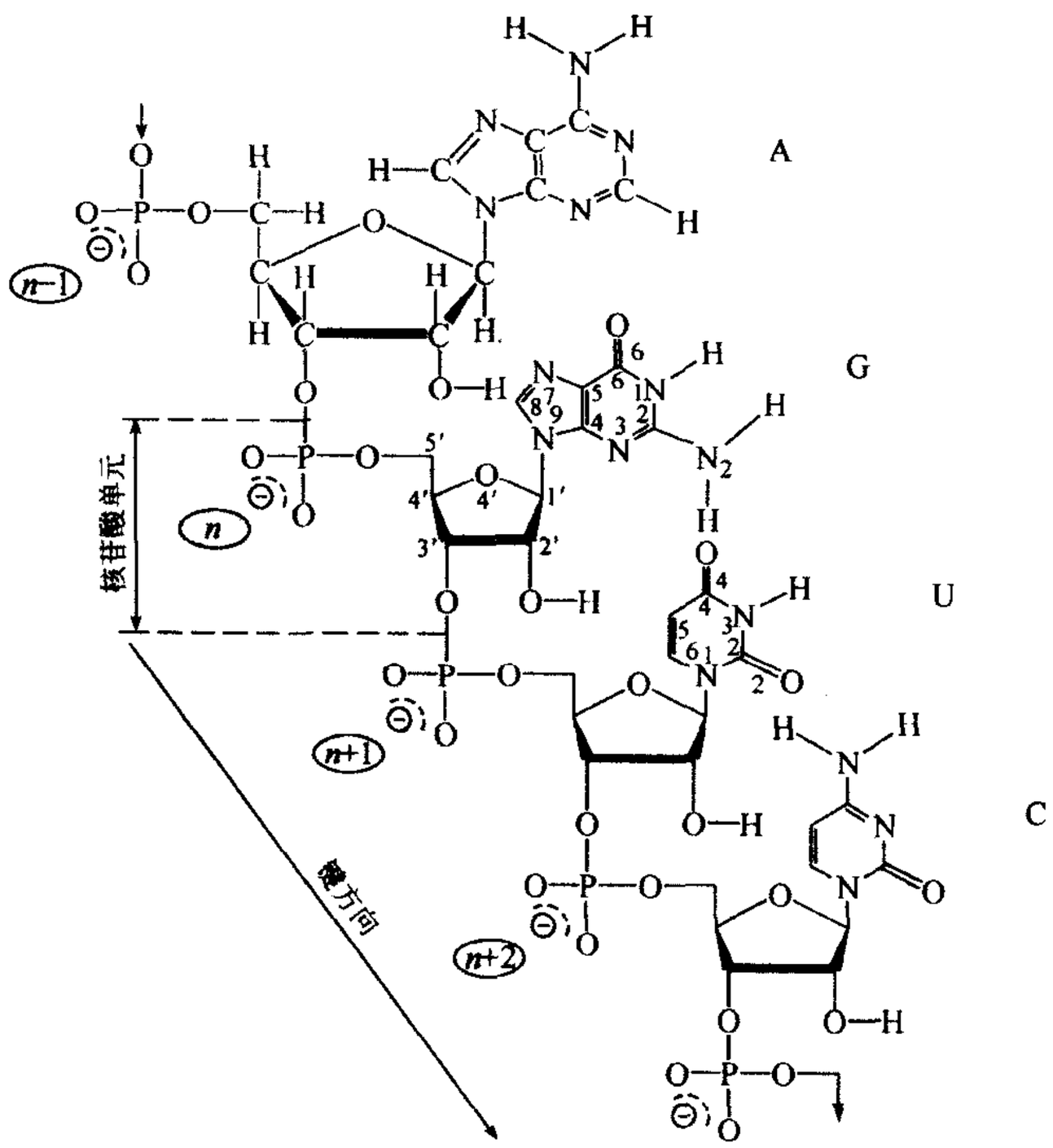



**图 4-1　双螺旋的核糖核酸（RNA）的片段**

结构，若是不在同一边称为 exo 结构。A-型聚核糖螺旋中 $C_{3'}$ 是 endo 结构，B-型聚核糖螺旋中 $C_{2'}$ 为 endo 结构。双股螺旋的 RNA 结构均为右旋的 A 型，但对应的脱氧核糖核酸（DNA）的结构则可能为 A 型或 Z 型。Z 型结构中，连接磷酸酯的线呈拉链状，为非螺旋的左旋结构。Rao 与 Kollman 引用 Weiner 的分子力场，利用分子力学方法计算多核糖由 A 型，B 型与 Z 型互相转换的势能面。其计算结果显示 B 型的 DNA 较 A 型的 DNA 稳定，但 A 型为 RNA 中最稳定的结构。计算所得到各种型的最优化结构及转换的能量与实验所测的一致。此为应用分子力学方法于生化分子研究的成功实例之一。

## 4.2 势能面的最小值点、最大值点、鞍点

势能函数一次微分值为零的结构并不一定表示能量极小值的分子结构，也可能是最大值构象或鞍点（saddle point）。以一个变量的函数 $f(x)$ 为例，当 $f'(x)=0$ 时，函数的极小值点 $f''(x)>0$，而函数的极大值点 $f''(x)<0$。对应于势能函数的一次导数为零时的构象，应由 Hessian 矩阵的本征值判定。设分子由 $N$ 个原子组成，可计算其 $3N\times 3N$ 的 Hessian 矩阵 $\tilde{F}$。此矩阵沿对角线对称，

$$F_{ij}=\frac{\partial^2 U}{\partial q_i \partial q_j}=\frac{\partial^2 U}{\partial q_j \partial q_i}=F_{ji} \tag{4-1}$$

式中，$q$ 为原子坐标。矩阵的本征值可经由矩阵的对角线化（diagonalization）求得，即

$$\tilde{C}^T\tilde{F}\tilde{C}=\tilde{F}_D \tag{4-2}$$

$$\tilde{F}_D=\begin{pmatrix} F_{d1} & 0 & 0 & \cdots & 0 \\ 0 & F_{d2} & 0 & \cdots & 0 \\ 0 & 0 & F_{d3} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & F_{d3N} \end{pmatrix} \tag{4-3}$$

$\tilde{F}_D$ 仅对角线有非零的元素，非对角线的元素皆为零。其对角线的元素称为 $\tilde{F}$ 矩阵的本征值（eigen value）。依此，当势能函数的一次导数为零时，若将分子的 Hessian 矩阵对角线化，则当最小值点时，其本征值中的 6 个元素为零，其余均为正数。此 6 个为零的本征值对应于分子的移动与转动

自由度。若 Hessian 矩阵由分子的内坐标计算求得，因分子有 $3N-6$ 个内坐标，故对应于最小值点的结构所得的本征值均为非零的正数。若为最大值点，则 Hessian 矩阵的本征值除零外均为负值。至于鞍点，则 Hessian 矩阵的本征值将出现正值及负值交互存在的现象。

## 4.3 势能面与化学反应

### 4.3.1 过渡态结构与反应路径

参与化学反应各物种相对能量的高低决定物种的稳定度（stability），反应能否发生取决于此热力学的资料。计算能量面上的极小值点可决定反应物种的相对能量。欲完全了解化学反应的过程，除了知道反应物种的能量外，还需要知道由反应物转换为生成物的速率。因此，除了极小值点的能量外，还需计算远离极小值点的能量面。设考虑能量面上两极小值点间的转换，此两点可能为反应物及生成物，或分子的两种不同的构型。“反应”并非一定指化学键的断裂或重组，亦可能仅为结构的变化，因此称两极小值点间的转换途径为“反应途径”（reaction pathway）。将反应途径中各点能量连成的曲线称为反应能量曲线（energy profile），进行途径的坐标变化称为反应坐标（reaction coordinate）。

由势能面的一极小值点至另一极小值点的反应能量曲线，通常为能量由一极小值点开始攀升，至曲线的鞍点后开始下降，直至另一极小值点。鞍点结构的势能对坐标的一次微分值为零（与极大值点或极小值点类似），但其 Hessian 矩阵产生一个或多个的负本征值。负本征值的数目称为鞍点的次数，通常最重要的为一次鞍点。一次鞍点为反应能量曲线的最大值点，但是此点为与反应途径方向垂直的能量曲线的极小值点，如图 4-2 所示。

图中，箭头标示反应进行的方向。图 4-3 为 $Cl^-$ 与 $CH_3Cl$ 的 $S_N2$ 反应曲线图，图中显示系统于鞍点的过渡状态（transition state）的结构及最小值点的结构。

由图可知，能量最小值点的结构由 $Cl^-$ 离子与被偏极化的 $CH_3Cl$ 偶极组成，称为“离子-偶极复合物”。图 4-4 显示由过渡状态结构及最小值点结构计算的几个较低简正频率与其对应的振动模式。图中，离子-偶极复合物对应于能量曲线的极小值点，其最低频率的振动模式为两个简并的

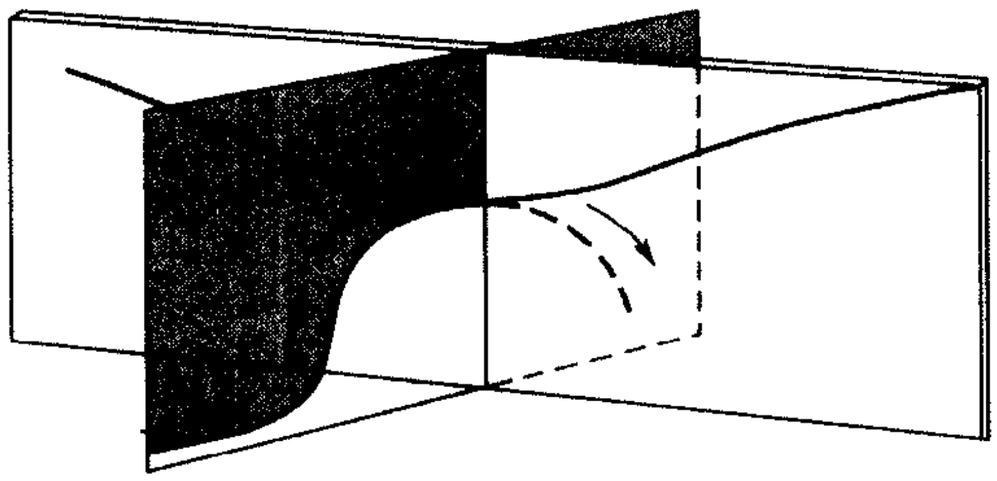


图 4-2　鞍点示意

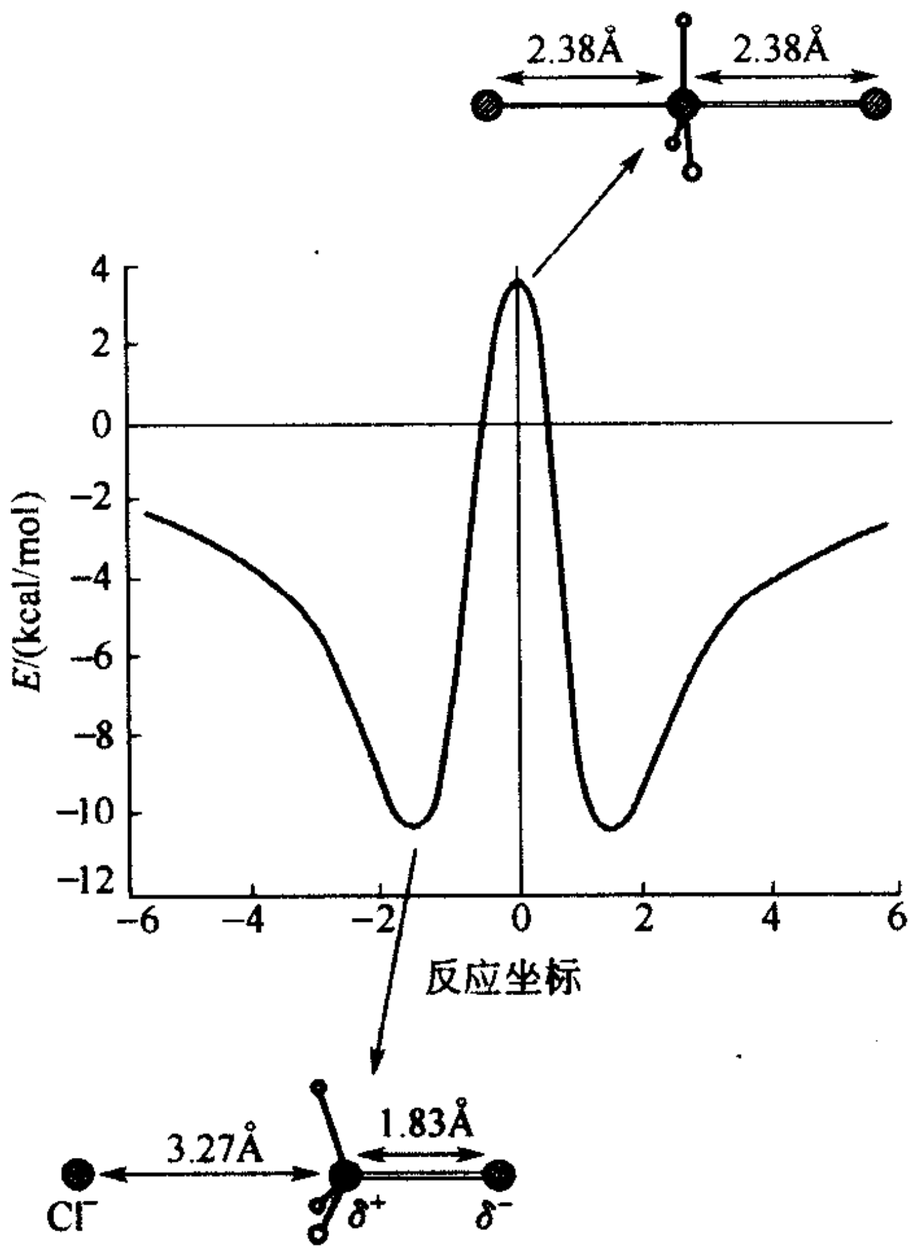



图 4-3　$Cl^-$ 与 $CH_3Cl$ 的 $S_N2$ 反应曲线

摇摆振动（wegging），振动频率均为 71.3cm$^{-1}$；图中的⊕表示振动方向垂直纸面向外伸出。此结构的次低频率（101.0cm$^{-1}$）的振动模式朝着形成过渡态结构的方向进行。由鞍点过渡态结构计算得到一个负本征值，对应于虚频（imaginary frequency）−415.0cm$^{-1}$。图中显示这一模式为沿 Cl—C—C 轴的振动，沿着反应途径向右进行。

图 4-4 的反应能量曲线及系统的结构与振动模式等均来自精确的量子力学计算。

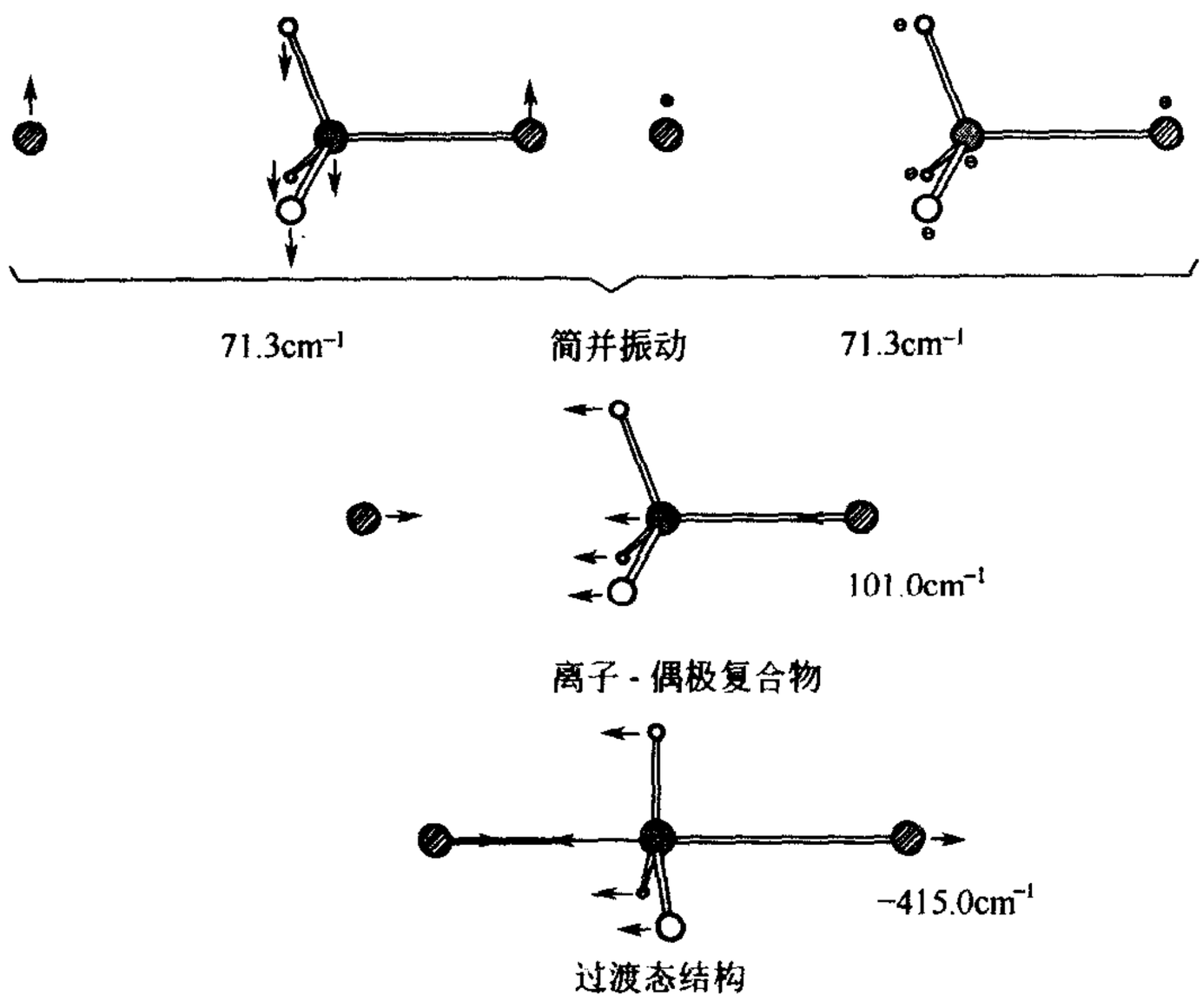



**图 4-4　$Cl^-$ 与 $CH_3Cl$ 的 $S_N2$ 反应过渡态结构与振动模式**

### 4.3.2　鞍点与二次区域

如以上的例子所示，当反应由能量最小值点接近鞍点时，Hessian 矩阵的本征值由全部为正值改变为含有一个负值。鞍点附近所有 Hessian 矩阵含一负本征值的能量面区间称为鞍点的二次区域（quadratic region），而能量最小值点附近所有 Hessian 矩阵无负本征值的能量面区间称为最小值点的二次区域。以下检查一个简单函数 $f(x,y)=x^4+4x^2y^2-2x^2+2y^2$ 的二次区域。此函数的图形如图 4-5 所示，图中的曲面由（$x$，$y$）各点的函数值所构成，另外 $xy$ 平面上显示出对应的等高线图。此函数有极值的条件为：

$$\frac{\partial f(x,y)}{\partial x}=4x^3+8xy^2-4x=0 \text{ 且} \frac{\partial f(x,y)}{\partial y}=8x^2y+4y=0 \quad (4\text{-}4)$$

由此条件，得联立方程组

$$\begin{cases} x(x^2+2y^2-1)=0 \\ y(2x^2+1)=0 \end{cases} \quad (4\text{-}5)$$

解此方程组，若 $x=0$，得 $y=0$；若 $y=0$，得 $x=\pm1$ 或 $x=0$。故此函数的两个最小值点位于（1，0）与（−1，0），一个鞍点位于（0，0）。

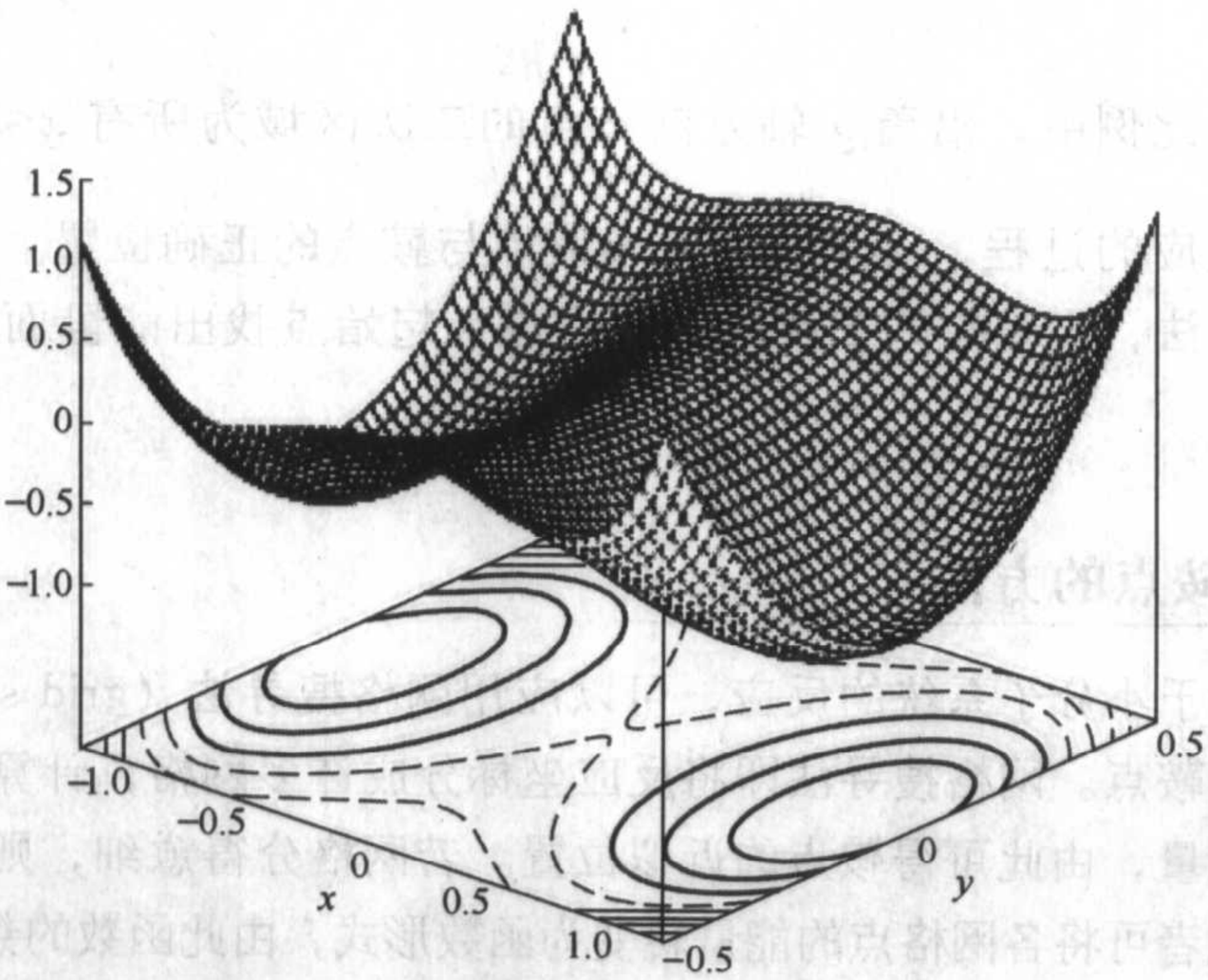



**图 4-5　$f(x,y)=x^4+4x^2y^2-2x^2+2y^2$ 的二次区域**

此函数的 Hessian 矩阵为：

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 12x^2+8y^2-4 & 16xy \\ 16xy & 8x^2+4 \end{pmatrix} \tag{4-6}$$

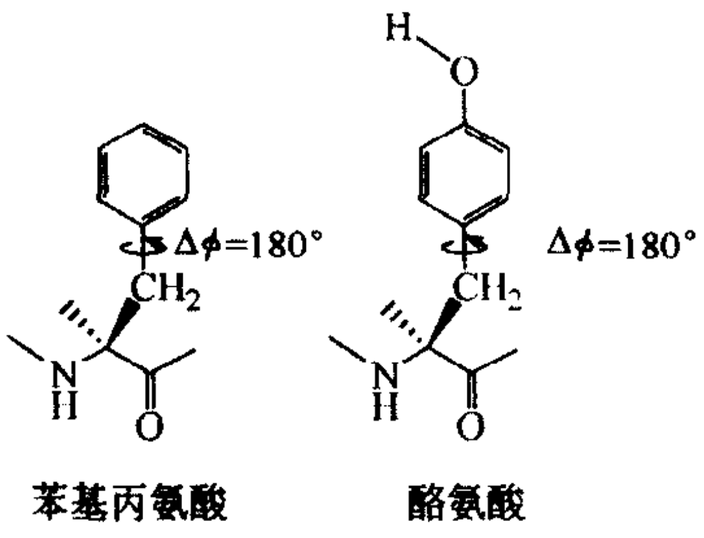
于点（1，0）的 Hessian 矩阵为 $\begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix}$，其本征值 $\lambda=4$，$\lambda=8$ 均为正数，故此点为极小值点。于点（0，0）的 Hessian 矩阵为 $\begin{pmatrix} -4 & 0 \\ 0 & 4 \end{pmatrix}$，其本征值 $\lambda=4$，$\lambda=-4$，一正一负，故此点为鞍点。对应于 $\lambda=4$ 的特征向量为（1，0）；对应于 $\lambda=-4$ 的特征向量为（0，1），此二向量互相垂直。特征向量表示函数梯度改变正负号的方向；故由图中可知沿着 $x=0$ 的方向通过函数的极小值点。若由点（−1，0）通过（0，0）再通过（1，0）将通过函数的极大值点。若将此函数想象为能量面，则在鞍点，对应于过渡状态结构负本征值的特征向量，表示过渡复合物中原子的运动方向。若沿着 $x$ 轴由极小值点（1，0）移动至鞍点，此途径上各点结构的 Hessian 矩阵为 $\begin{pmatrix} 12x^2-4 & 0 \\ 0 & 8x^2+4 \end{pmatrix}$。而 $8x^2+4$ 恒大于零，故当 $12x^2-4\geqslant 0$（$x\geqslant$

$\frac{1}{\sqrt{3}}$）时矩阵的本征值均为正数，否则将出现一个正本征值与一个负本征值。此例中，沿着 $x$ 轴方向鞍点的二次区域为所有 $x<\frac{1}{\sqrt{3}}$ 的点。欲了解反应的过程，必须找出最小值点与鞍点的正确位置。可以运用一些数学方法，由鞍点的二次区域内的点为起始点找出能量面上鞍点的正确位置。

### 4.3.3 决定鞍点的方法

对于小分子系统的反应，可以应用网格搜寻法（grid search）找出势能面的鞍点。网格搜寻法即将反应坐标分成许多网格，计算每一网格点的系统能量，由此可得鞍点的近似位置。若网格分得愈细，则位置的误差愈小。或者可将各网格点的能量转变为函数形式，由此函数的数学形式求出鞍点、最小值点等。网格搜寻法因为需要计算许多网格点的能量，并且有一定的误差存在，只适用于研究如 $H+H_2 \longrightarrow H_2+H$；$Cl^- + CH_3Cl \longrightarrow ClCH_3 + Cl^-$ 等简单的反应系统。此计算法的优点在于能够得到整个能量面而非最小值点与鞍点的附近区域。

由分子的一种最低能量构象转换为另一种最低能量构象的反应，通常仅为少数几个坐标的变化，如图 4-6 中苯基丙氨酸或酪氨酸分子的苯环绕着 $C_1$—$C_4$ 轴旋转 180°的构象转换。




苯基丙氨酸　　酪氨酸

**图 4-6　苯基丙氨酸与酪氨酸分子的扭转示意**

研究这种反应的反应途径，通常可采用如下的近似法。逐渐改变反应的坐标（如图中的 $\Delta\phi$），计算在反应坐标的每一点的能量，除反应坐标外，将系统的其他部分结构进行最优化。以这样的方法所得到的一系列的结构视为反应的途径，而将得到的最高能量点近似为鞍点。以此方法研究主要围绕着键转动的构型变化，称为绝热投像（adiabatic mapping）或扭转角驱动法（torsion angle driving）。分子力学计算广泛地利用绝热投像法以得到大分子的构象变化的能量图。图 4-7 为氘化的麦芽糖 Man（$\alpha$1-3）Man$\beta$ 的结构。麦芽糖由两个单糖基经氧的桥键所组成，其氧桥形成两个扭转角 $\Psi$ 及 $\Phi$。

图 4-8 为利用扭转角驱动法与 MM2 力场计算所得的能量面。计算中，每次改变 $\Phi$ 或 $\Psi$ 为 10°，在固定的（$\Psi$，$\Phi$）求得能量最优化的能量。图中，MA、MB、MC 与 MD 分别为能量极小值点，其鞍点约在（$\Psi$，$\Phi$）=（80°，30°）附近。

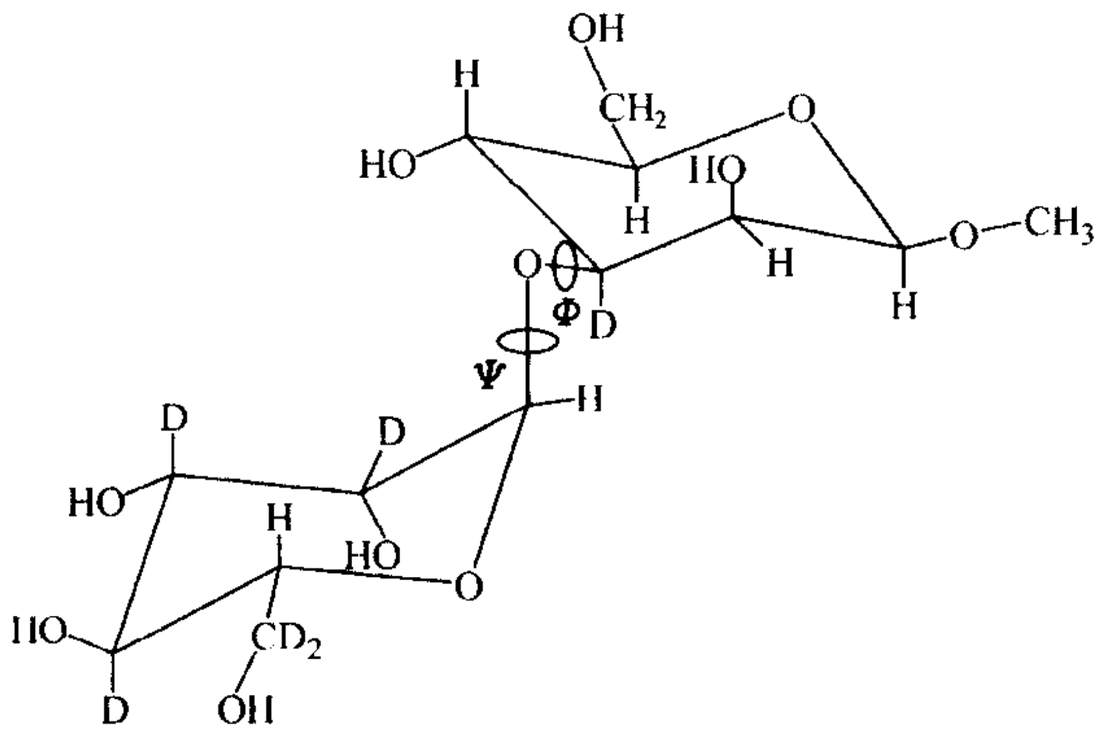



**图 4-7　氘化的麦芽糖 Man（α 1-3）Manβ 的结构**

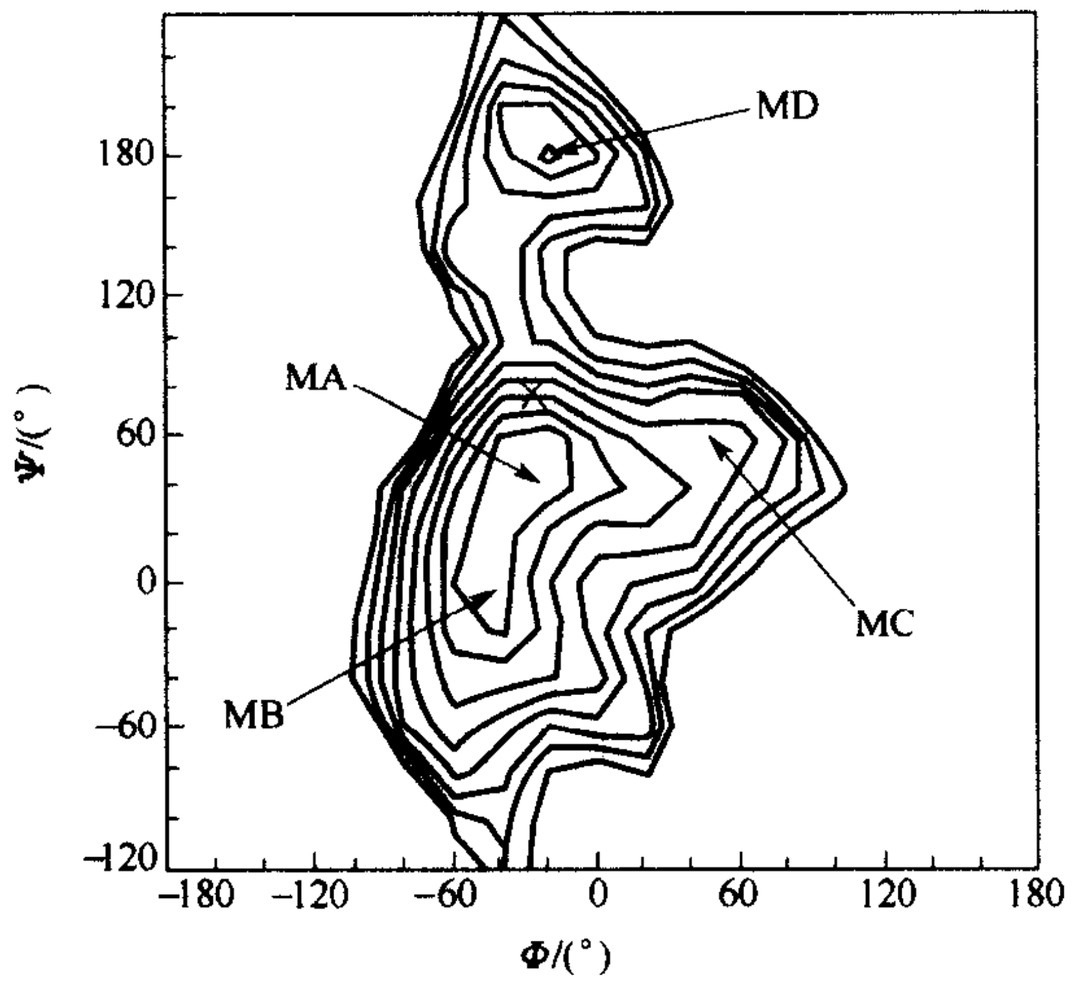



**图 4-8　利用扭转角驱动法与 MM2 力场计算所得氘化的麦芽糖的能量面**

利用此能量面图，可以判定由各构象间互相转换的转换途径，及所需翻越的能量障碍。

# 4.4 分子力学计算的热力学性质

## 4.4.1 构象能

分子力学可计算分子在任一构象的势能，此种势能称为分子的构象能(steric energy，SE)。对多原子分子而言，势能为分子键长、键角、扭转角等内坐标的函数。不同构象所对应的势能则构成分子的多维势能面(multi-dimensional potential surface)。分子的任一几何位置坐标，即任何一种构象（conformation)，对应于势能面的一点，具有一定的势能。分子的势能面通常具有一些极大值与极小值的点。对应于势能面极小值的构象称为构象异构物（conformer)。由分子力学所得到的势能面可计算出各个构象异构物的势能。

由于分子力学所能得到的构象异构体的势能均为分子在 0K 无任何运动状态的假想能量。实际上，即使分子处于 0K 亦有零点振动能（zero-point energy)。在温度不为 0K 的情况下，分子构象异构物的能量除了势能外，还包括移动能、转动能与振动能。气相分子的内能（internal energy）为其势能、平均移动能、转动能和振动能之和。对非线性分子而言，其室温下的移动能与转动能可经由统计力学推得，约为 $3RT$。分子的振动能依照简谐振子的假设为：

$$U_{\text{vib}} = \sum_{i=1}^{3N-6} h\upsilon_i \left( \frac{1}{2} + \frac{1}{e^{h\upsilon_i/kT} - 1} \right) \tag{4-7}$$

理想气体分子的焓（enthalpy）为分子的内能加上压力与体积之功，$RT$。根据以上的结果，可得到分子异构体（isomer）A 与 B 间焓的差值为：

$$\Delta H(\text{A-B}) = \Delta SE(\text{A-B}) + \Delta U_{\text{trans}}(\text{A-B}) + \Delta U_{\text{rot}}(\text{A-B}) + \Delta U_{\text{vib}}(\text{A-B}) \tag{4-8}$$

式中，$\Delta SE$ 为异构物间构象能之差；$\Delta U$ 为各种所标示能量的差。在一般温度状况下，分子异构物的平均移动能与转动能相同。故式(4-8) 又可简化为

$$\Delta H(\text{A-B}) = \Delta SE(\text{A-B}) + \Delta U_{\text{vib}}(\text{A-B}) \tag{4-9}$$

因此，计算由 A→B 焓的变化仅需计算两构象的能差及振动能差。

## 4.4.2 分子的生成焓

分子的生成焓（enthalpy of formation）为分子的焓与其组成原子的

焓的差值。依照热力学，气态分子的标准生成焓（standard enthalpy of formation）$\Delta H^{\ominus}_{f,298}$为

$$\Delta H^{\ominus}_{f,298}=\Delta U^{\ominus}+(\Delta n)RT \tag{4-10}$$

式中，$\Delta U^{\ominus}$为分子内能的改变量；$(\Delta n)RT$为气态物种数目变化时其$P(\Delta V)$的改变量。例如在压力为1bar（1bar$=10^5$Pa），温度为298K条件下的反应：

$$n\mathrm{C}(石墨,\mathrm{s})+m\mathrm{H_2}(\mathrm{g})\longrightarrow \mathrm{C}_n\mathrm{H}_{2m}(\mathrm{g})$$

其物种数目的变化为$\Delta n=1-m$。将此代入式(4-10)得标准生成热，即：

$$\Delta H^{\ominus}_{f,298}=\Delta U^{\ominus}+RT-mRT \tag{4-11}$$

分子内能的变化为移动能、转动能、振动能、电子能与键能等变化量之和，即：

$$\Delta U^{\ominus}=\Delta U^{\ominus}_{trans}+\Delta U^{\ominus}_{rot}+\Delta U^{\ominus}_{vib}+\Delta U^{\ominus}_{elec}+\Delta \mathrm{U}^{\ominus}_{bond} \tag{4-12}$$

在此例中，于温度$T$的分子移动能相对于0K的变化为：

$$\Delta U^{\ominus}_{trans}=(U^{\ominus}_{T,trans}-U^{\ominus}_{0,trans})_{\mathrm{C}_n\mathrm{H}_{2m}(\mathrm{g})}-m(U^{\ominus}_{T,trans}-U^{\ominus}_{0,trans})_{\mathrm{H_2(g)}}-$$
$$n(U^{\ominus}_{T,trans}-U^{\ominus}_{0,trans})_{\mathrm{C}(石墨,\mathrm{s})} \tag{4-13}$$

$(U^{\ominus}_{T,trans}-U^{\ominus}_{0,trans})$项的值，气相分子为$3/2RT$，固相分子为0。将其代入式(4-13)，得：

$$\Delta U^{\ominus}_{trans}=3/2RT-3/2mRT-0 \tag{4-14}$$

非线性分子的转动能为$3/2RT$，物质的量为$m$(mol)的线性$\mathrm{H_2}$分子的转动能为$mRT$，故于温度$T$的分子转动能相对于0K的变化为：

$$\Delta U^{\ominus}_{rot}=3/2RT-mRT-0 \tag{4-15}$$

欲实际计算式(4-12)中的$\Delta U^{\ominus}_{vib}$、$\Delta U^{\ominus}_{elec}$、$\Delta U^{\ominus}_{bond}$各项十分困难。在分子力学中，一般采用所谓的“构群加成法”（group increment method）计算这些能量差。构群加成法中将分子的键或基团视为组成分子的构群。每种不同的构群具有特定的构群能。而分子由于构群数目增加所造成的能量差为固定的构群能的和。分子总的构群加成能（group increment energy，GE）为：

$$\mathrm{GE}=\sum k_i\cdot u_i \tag{4-16}$$

式中，$i$为构群的种类；$k_i$为$i$构群的数目；$u_i$为构群能。例如可将烷类分子视为由一级碳、二级碳、三级碳及四级碳的构群所组成。依此，2，2，3，4，4-五甲基己烷分子包含了7个一级碳构群（$k_p=7$），1个二级碳构群（$k_s=1$），1个三级碳构群（$k_t=1$）及2个四级碳构群（$k_q=$

2)，其总构群加成能为：

$$\mathrm{GE}=7u_{\mathrm{p}}+1u_{\mathrm{s}}+1u_{\mathrm{t}}+2u_{\mathrm{q}} \tag{4-17}$$

利用构群加成能，分子力学将 $\Delta U^{\ominus}_{\mathrm{vib}}$，$\Delta U^{\ominus}_{\mathrm{elec}}$，$\Delta U^{\ominus}_{\mathrm{bond}}$的和近似为：

$$\Delta U^{\ominus}_{\mathrm{vib}}+\Delta U^{\ominus}_{\mathrm{elec}}+\Delta U^{\ominus}_{\mathrm{bond}}\sim \mathrm{GE}+\mathrm{SE} \tag{4-18}$$

将此关系及式(4-14)、式(4-15) 代入式(4-12) 得：

$$\Delta U^{\ominus}=\mathrm{GE}+\mathrm{SE}+3RT-3/2RT-mRT \tag{4-19}$$

故得分子标准生成热为：

$$\Delta H^{\ominus}_{\mathrm{f},298}=\mathrm{GE}+\mathrm{SE}+4RT-7/2mRT \tag{4-20}$$

或将 $7/2mRT$ 合并于 GE 项中得：

$$\Delta H^{\ominus}_{\mathrm{f},298}=\mathrm{GE}+\mathrm{SE}+4RT \tag{4-21}$$

以烷类分子为例，将式(4-21) 重排为：

$$\Delta H^{\ominus}_{\mathrm{f},298}-\mathrm{SE}-4RT=\mathrm{GE}=k_{\mathrm{p}}\cdot u_{\mathrm{p}}+k_{\mathrm{s}}\cdot u_{\mathrm{s}}+k_{\mathrm{t}}\cdot u_{\mathrm{t}}+k_{\mathrm{q}}\cdot u_{\mathrm{q}} \tag{4-22}$$

$\Delta H^{\ominus}_{\mathrm{f},298}$可由一般的物理化学标准参考书查出。若有 $n$ 种不同的烷类分子的资料，则可将式(4-22) 的左边各项集合写为：

$$Y_{\mathrm{n}}=(\Delta H^{\ominus}_{\mathrm{f},298}-\mathrm{SE}-4RT)_{\mathrm{n}}=(k_{\mathrm{p}}\cdot u_{\mathrm{p}}+k_{\mathrm{s}}\cdot u_{\mathrm{s}}+k_{\mathrm{t}}\cdot u_{\mathrm{t}}+k_{\mathrm{q}}\cdot u_{\mathrm{q}})_{\mathrm{n}} \tag{4-23}$$

因此由实验的 $\Delta H^{\ominus}_{\mathrm{f},298}$值及由分子力学计算的 SE 值，可得线性方程组，即：

$$\begin{aligned}
Y_1&=(k_{\mathrm{p}}\cdot u_{\mathrm{p}}+k_{\mathrm{s}}\cdot u_{\mathrm{s}}+k_{\mathrm{t}}\cdot u_{\mathrm{t}}+k_{\mathrm{q}}\cdot u_{\mathrm{q}})_1\\
Y_2&=(k_{\mathrm{p}}\cdot u_{\mathrm{p}}+k_{\mathrm{s}}\cdot u_{\mathrm{s}}+k_{\mathrm{t}}\cdot u_{\mathrm{t}}+k_{\mathrm{q}}\cdot u_{\mathrm{q}})_2\\
Y_3&=(k_{\mathrm{p}}\cdot u_{\mathrm{p}}+k_{\mathrm{s}}\cdot u_{\mathrm{s}}+k_{\mathrm{t}}\cdot u_{\mathrm{t}}+k_{\mathrm{q}}\cdot u_{\mathrm{q}})_3\\
&\vdots\\
Y_{\mathrm{nth}}&=(k_{\mathrm{p}}\cdot u_{\mathrm{p}}+k_{\mathrm{s}}\cdot u_{\mathrm{s}}+k_{\mathrm{t}}\cdot u_{\mathrm{t}}+k_{\mathrm{q}}\cdot u_{\mathrm{q}})_{\mathrm{n}}
\end{aligned} \tag{4-24}$$

由式(4-24)，可应用线性回归法（linear regression）将构群能 $u_{\mathrm{p}}$，$u_{\mathrm{s}}$，$u_{\mathrm{t}}$，$u_{\mathrm{q}}$ 求出。

表 4-1 列出 Kao 与 Allinger 等利用这种方法对数种烷类分子计算的结果。表中，$\Delta H^{\ominus}_{\mathrm{f,exp},298}$为实验的生成热之值，$\Delta H^{\ominus}_{\mathrm{f,cal},298}$为由所计算得到的构群能逆回计算的标准生成热。SE 为利用 MM3 力场所计算的最低构形能。计算误差为：

$$\text{误差}=\left(\frac{\Delta H^{\ominus}_{\mathrm{f,exp}}-\Delta H^{\ominus}_{\mathrm{f,cal}}}{\Delta H^{\ominus}_{\mathrm{f,exp}}}\right)\times 100\% \tag{4-25}$$

表 4-1　一些选择烷类分子的标准生成热

| 烷　类 | $\Delta H_{f,298}^{\ominus}$ 实验值 | $\Delta H_{f,298}^{\ominus}$ 计算值 | $\Delta\Delta H_f^{\ominus}$ | SE | $Y_n$ | $k_p$ | $k_s$ | $k_t$ | $k_q$ | 误差 |
|---|---|---|---|---|---|---|---|---|---|---|
| 2,2,3-三甲基丁烷 | −48.87 | −48.86 | 0.01 | 8.27 | −59.54 | 5 | 0 | 1 | 1 | 0.02 |
| 3,3-二甲基戊烷 | −48.08 | −47.45 | 0.63 | 8.26 | −58.74 | 4 | 2 | 0 | 1 | 1.31 |
| 新戊烷 | −40.14 | −40.78 | −0.64 | 2.27 | −44.81 | 4 | 0 | 0 | 1 | 1.60 |
| 3-甲基戊烷 | −41.33 | −41.34 | −0.01 | 5.53 | −49.26 | 3 | 2 | 1 | 0 | 0.03 |
| 正戊烷 | −35.56 | −36.30 | −0.74 | 2.82 | −40.78 | 2 | 3 | 0 | 0 | 2.07 |
| 正丁烷 | −30.60 | −30.62 | −0.02 | 2.17 | −35.17 | 2 | 2 | 0 | 0 | 0.08 |
| 丙烷 | −24.83 | −24.97 | −0.14 | 1.50 | −28.73 | 2 | 1 | 0 | 0 | 0.57 |
| 乙烷 | −20.24 | −19.33 | 0.91 | 0.82 | −23.46 | 2 | 0 | 0 | 0 | 4.51 |
| 环已烷 | −29.50 | −29.00 | 0.50 | 6.55 | −38.45 | 0 | 6 | 0 | 0 | 1.70 |
| 环丁烷 | 6.78 | 6.32 | −0.46 | 29.22 | −24.84 | 0 | 4 | 0 | 0 | 6.78 |

注：表中 $\Delta H_f^{\ominus}$，SE，$Y_n$ 均以 kcal/mol（1kal=4.184kJ）为单位，误差以%为单位。

### 4.4.3　反应热与生成焓

上节讨论了分子的构象能与标准生成焓 $\Delta H_{f,298}^{\ominus}$ 的关系。构象能亦可用以计算分子构象异构体间的反应热。以正丁烷为例，实验度量气态正丁烷的邻位交叉式（gauche）与反式（trans）间的焓差为 17.15（±0.84）kcal/mol。可将此关系表示为：

$$C_4H_{10}(\text{trans}) \Leftrightarrow C_4H_{10}(\text{gauche})$$

$$\Delta H_{rxn}^{\ominus} = \Delta H_{f,(\text{gauche})}^{\ominus} - \Delta H_{f(\text{trans})}^{\ominus} = 17.15(\pm 0.84)\text{kcal/mol} \quad (4\text{-}26)$$

将式(4-21) 代入得：

$$\Delta H_{rxn}^{\ominus} = (GE+SE+4RT)_{(\text{gauche})} - (GE+SE+4RT)_{(\text{trans})} \quad (4\text{-}27)$$

因为二异构物的构群相同，故式(4-27) 可简化为：

$$\Delta H_{rxn}^{\ominus} = SE_{(\text{gauche})} - SE_{(\text{trans})} \quad (4\text{-}28)$$

这是非常重要的结果，提供了计算异构体间焓差的方法。当然，若分子不同，GE 项不能互相消去，式(4-28) 所得的构形能差并不等于实验的焓差。但是若分子相同，仅是构象的差别，则仅需计算其构象能差。表 4-2列出利用 MM2 力场所计算的一些有机分子异构物对的构象能，及组成构象能各项的能量。由表中的构象能可判定较稳定的异构物。并且，由组成各项的能量可判定主要的差异原因。例如二环 [2,2,1] 庚烷与二环 [3,3,1] 庚烷，其构象能的差异主要在于键角弯曲项的差异。利用这些构象能可以计算异构物间互相转换的焓差，以判定反应的可行性。

表 4-2　一些有机化合物的构象能与其组成项的能量　/(kcal/mol)

| 化　合　物 | 构象能 | 范德瓦耳斯 | 键伸缩 | 键角弯曲 | 二面角扭转 |
|---|---|---|---|---|---|
| 二环[2,2,1]庚烷 | 23.09 | 3.52 | 0.08 | 9.05 | 10.18 |
| 二环[3,3,1]庚烷 | 38.69 | 5.60 | 1.32 | 19.66 | 13.18 |
| 顺式二环[3,2,0]庚烷 | 37.21 | 2.32 | 0.92 | 15.07 | 19.78 |
| 反式二环[3,2,0]庚烷 | 54.63 | 5.84 | 2.15 | 35.06 | 13.31 |
| 邻二异丁基苯 | 40.60 | 16.51 | 5.57 | 12.12 | 5.98 |
| 对二异丁基苯 | 17.75 | 8.43 | 1.65 | 2.22 | 5.24 |
| 顺式环己烯 | 4.18 | 3.28 | 0.22 | 0.37 | 0.16 |
| 反式环己烯 | 65.54 | 7.97 | 2.05 | 19.09 | 36.26 |
| 反式 2,6-二甲基-二环[2,2,2]辛烷 | 22.38 | 8.15 | 1.18 | 1.52 | 11.40 |
| 顺式 2,6-二甲基-二环[2,2,2]辛烷 | 24.46 | 8.50 | 1.23 | 2.64 | 11.94 |

注：1kcal=4.184kJ。

# 4.5　晶体结构的计算

Lii 与 Allinger 成功地发展出利用分子力学计算分子晶体堆积的方法。依照热力学，分子的吉布斯自由能（Gibbs free energy）可表示为：

$$G=U(r_{ij})+K+pV-TS \tag{4-29}$$

式中，$U(r_{ij})$ 为分子处于平衡状态的势能；$K$ 为分子的动能；$T$ 为热力学温度；$S$ 为熵（entropy)。若考虑晶体的吉布斯自由能，则可将自由能更细分为分子内自由能（intra free energy）与分子间自由能（inter free energy）两部分。

$$G_{crystal}=E_{inter}+E_{intra}+pV-T(S_{inter}+S_{intra}) \tag{4-30}$$

式中，分子内能 $E_{intra}$，与分子间内能 $E_{inter}$，分别为：

$$E_{inter}=U_{inter}(r_{ij})+K_{inter}$$

$$E_{intra}=U_{intra}(r_{ij})+K_{intra} \tag{4-31}$$

在一般的条件下，假设分子形成晶体时，其分子结构的形变可忽略，因此可将其视为刚体分子。并且，由晶格振动所提供的分子的零点能与热能为一固定值。基于这两点假设，分子的内能、熵值与分子间的动能及熵值均为常数。对刚体分子而言，平衡的分子间势能由分子的质心在晶体中的位置$\vec{R}_i$与分子的位向 $\theta_k$ 所决定。因晶体的 $PV$ 值远小于其分子间的势能，故可将其自式(4-31) 中忽略。综合上述，分子晶体的吉布斯自由能可简化为：

$$G_{\text{crystal}} = U_{\text{inter}}(\vec{R}_i, \theta_k) + \text{常数} \tag{4-32}$$

依此式，分子间的作用势能为影响晶体吉布斯自由能最重要的因素。因为已假设晶体中的分子为刚体分子，故分子间势能为分子所含原子间范德瓦耳斯作用的总和。合并式(4-30) 和式(4-31)，可得分子晶体的升华热(heat of sublimation) $\Delta H_{\text{sub}}$的关系为：

$$\Delta G_{\text{sub}} = G_{\text{gas}} - G_{\text{crystal}} = \Delta H_{\text{sub}} + \Delta(pV) \tag{4-33}$$

最简单的分子晶体为晶胞中仅含一个分子，另外 8 个分子位于晶胞的格角上。这样的系统需要 24 个坐标以定义这些分子的质心在晶胞的位置。其中 12 个坐标由晶胞的对称性决定，剩下的坐标为 6 个晶胞参数及 3 个晶胞转动坐标与 3 个晶胞移动坐标。因晶胞的移动与转动并不改变分子在晶胞间的相对位置，故仅需 6 个晶胞参数以定义分子质心的位置。综合上述，定义此分子晶体系统的几何参数为 6 个晶胞参数，($a$，$b$，$c$，$\alpha$，$\beta$，$\gamma$) 及 3 个欧拉角 ($\theta_x$，$\theta_y$，$\theta_z$)。

执行晶体计算通常输入已经过能量最优化得到的分子的几何坐标及猜测的晶胞参数与欧拉角的值。依据这些输入的资料，建立一含有 27 个晶胞 (3×3×3) 的晶体。然后以分子力学方法改变参数求取晶体能量的最小值。对应于晶体能量最小值的结构应为晶体的最佳结构。由此结构可计算晶体的升华热与实验对比。图 4-9 为 Lii 与 Allinger 等所发展的晶体计算程序“CRSTL”的流程图。

表 4-3 为利用此程序与 MM2 及 MM3 力场所计算的正庚烷晶体的结果。表中百分误差为由 MM3 力场计算的结果与实验值相比较的结果。表中显示，引用 MM3 力场计算的结果优于 MM2 力场的结果。计算的几何参数与 X 射线衍射实验所得的数据相近，升华热与热力学实验值的误差亦小于 10%。可以说是一个利用分子力学成功的计算。

**表 4-3　利用 CRSTL 程序与 MM2 及 MM3 力场所计算的正庚烷晶体的结果**

| 参数 | MM2 | MM3 | 实验值 | 误差/% |
|---|---|---|---|---|
| $a$ | 4.14 | 4.34 | 4.15±0.01 | 4.6 |
| $b$ | 4.44 | 4.65 | 4.69±0.01 | −0.9 |
| $c$ | 10.12 | 20.23 | 19.97±0.02 | 1.3 |
| $\alpha$ | 91.6 | 91.4 | 91.3±0.2 | 0.1 |
| $\beta$ | 94.3 | 93.5 | 94.9±0.2 | −1.5 |
| $\gamma$ | 102.9 | 103.7 | 105.7±0.2 | −1.9 |
| vol① | 359.4 | 395.6 | 373.0 | 6.1 |
| $E_{\text{sub}}$② | 22.01 | 13.45 | 12.60 | 6.7 |

① vol：晶胞的体积，Å$^3$。

② $E_{\text{sub}}$：计算的升华热，kcal/mol。

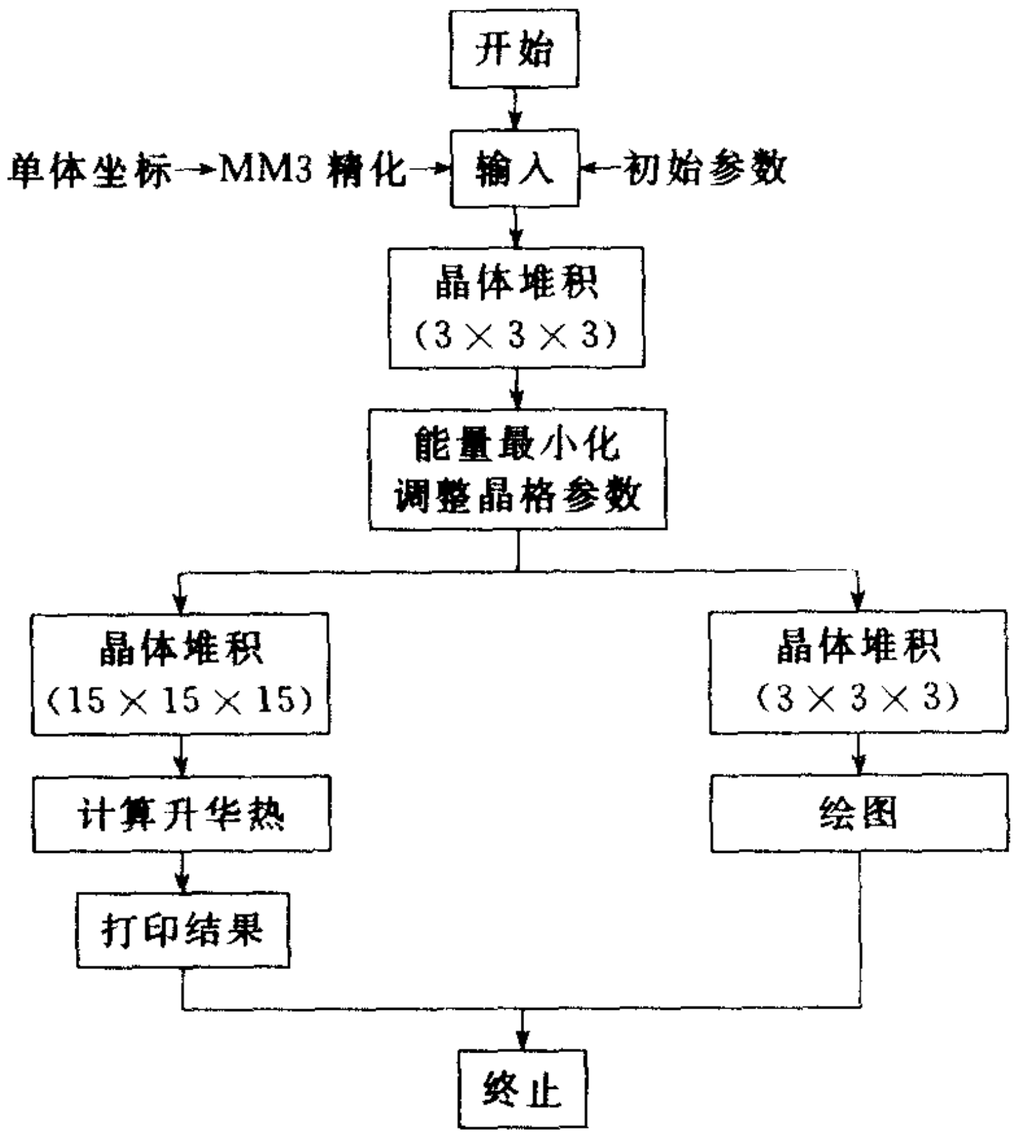



**图 4-9　CRSTL 程序之计算流程**

图 4-10　显示利用 MM3 力场及 CRSTL 计算所得晶体正庚烷的堆积情形。

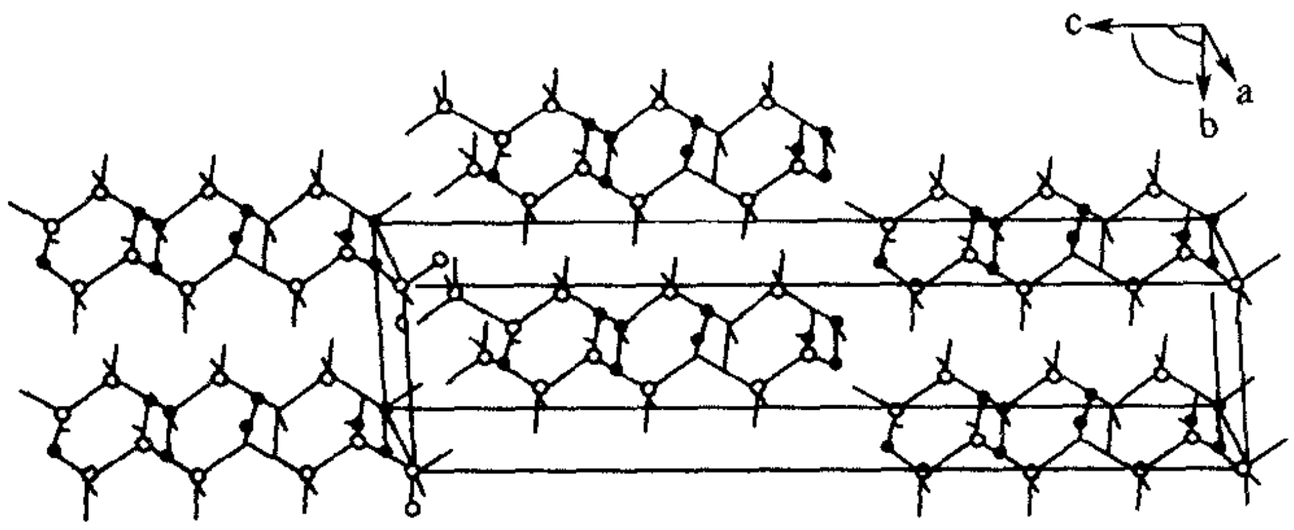



**图 4-10　利用 MM3 力场计算所得正庚烷晶体的堆积图**

# 4.6　立体效应

分子力学方法亦可用以研究化学反应中的立体效应（steric effect）。Taft 等曾研究以酸催化酯的水解反应，认为酯的立体效应因子 $E_s$，及各

种酯水解的速率常数 $k$，皆与乙酸乙酯水解速率常数 $k_0$ 的比值相关，彼此之间具有如下的关系：

$$\ln \frac{k}{k_0} = -E_s \tag{4-34}$$

这种关系式称为 Taft 方程式（Taft equation）。DeTar 与 Tenpas 以分子力学方法研究 $RCOOCH_2CH_3$ 的水解反应。设酯水解的反应机制为：

$$RCOOCH_2CH_3 \longrightarrow RC(OH)_2OCH_2CH_3 \longrightarrow 产物 \tag{4-35}$$

利用分子力学计算反应物 $RCOOCH_2CH_3$ 与中间过渡产物 $RC(OH)_2OCH_2CH_3$ 的焓，并假设这两种焓的差即为其立体能量的差，即：

$$H[RC(OH)_2OCH_2CH_3] - H(RCOOCH_2CH_3) = SE \tag{4-36}$$

经由分子力学计算不同的 R 基团的酯的水解反应的立体能量差与实验的反应速率常数相比较，发现彼此之间符合如下的关系：

$$\ln k = 0.340 - 0.789SE \tag{4-37}$$

因为反应速率常数与自由能相关，故式(4-37) 表示生成物与反应物立体能量的差与自由能呈良好线性关系。依照此关系可得到：

$$\ln \frac{k}{k_0} = -0.789(SE - SE_0) = -E_s \tag{4-38}$$

此计算证实了 Taft 的 $E_s$ 值的确反映出分子的立体效应。

DeTar 等亦引用分子力学计算研究如下的 $S_N2$ 反应的立体效应。

$$MY + RX \longrightarrow YR + MX \tag{4-39}$$

式中，X，Y 为卤素原子。设 $i$ 为烷基 R 的标号，$j$ 为分子 MY 的编号，则可得到实验的速率常数具以下的关系：

$$\ln k_{ij} = a_j + b_j E_i \tag{4-40}$$

式中，$a_j$ 与 $b_j$ 为分子 MY 的参数；$E_i$ 为反应物与中间过渡产物的能差，约等于其自由能的差，$\Delta G^{\neq}$。DeTar 等利用分子力学方法计算溴化物的中间过渡产物与反应物的生成热差，$\Delta H^{\neq}$（$\Delta H^{\neq} \sim \Delta G^{\neq}$），并以式(4～41）估计其与立体能差的关系。

$$\Delta\Delta H^{\neq} = \Delta H_i^{\neq} - \Delta H_{ethyl}^{\neq} = (SE_i^{trans} - SE_i) - (SE_{ethyl}^{trans} - SE_{ethyl}) = E_{st} - E_{s,ethyl} \tag{4-41}$$

式中，“ethyl” 为乙烷。经过计算可得到与实验结果非常吻合的结果。以上的例子均为分子力学计算立体效应成功的应用。

## 参 考 文 献

1 Burkert U，Allinger N L. “Molecular Mechanics”，ACS Monograph 177，American Chemical

Society, Washington, D. C. USA, 1982
2 Osawa E, Collins J B, Schleyer P V R. Tetrahedron, 33, 2667, 1977
3 Wertz D H, Allinger N L. Tetrahedron, 30, 1579, 1974
4 Rao S N, Kollman P A. J. Am. Chem. Soc., 108, 3048, 1986
5 Weiner S J, Kollman P A, Case D A, Singh U C, Ghio C, Alagona G, Profeta S, Weiner P. J. Am. Chem. Soc., 106, 76, 1984
6 Kao J, Allinger N L. J. Am. Chem. Soc., 99, 975, 1977
7 Lii J H, Allinger N L. J. Am. Chem. Soc., 111, 8576, 1989
8 DeTar D F, McMullen D F, Lurthra N P. J. Am. Chem. Soc., 100, 2484, 1978
9 DeTar D F, Tenpas C J. J. Am. Chem. Soc., 98, 7903, 1976

# 第5章 分子动力模拟计算的原理

分子动力模拟（molecular dynamics simulation）计算，简称为 MD 计算，自 1966 年起发展至今约四十年。应用十分广泛，随着计算机的迅速发展，越来越多地被应用于各种体系的模拟。目前已有许多成熟的商用软件，使用方便，图形清晰，功能较强。但是，计算的成败很大程度上取决于力场的适用性、计算速度的快慢、计算方法的正确性以及起始结构的合理性。涉及力场的内容请查阅第 2 章的介绍。本章分两部分，正文部分比较系统地介绍了分子动力模拟计算的原理，5.10 给出了不同系综的分子动力计算公式。

## 5.1 分子动力计算的基本原理

考虑含有 $N$ 个分子或原子的运动系统，系统的能量为系统中分子的动能与系统总势能的总和。其总势能为分子中各原子位置的函数 $U(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_n)$。通常势能可分为分子间（或分子内）原子间的非键结范德瓦耳斯作用（VDW）与分子内部势能（int）两大部分，即：

$$U = U_{VDW} + U_{int} \tag{5-1}$$

范德瓦耳斯作用一般可将其近似为各原子对间的范德瓦耳斯作用的加成：

$$U_{VDW} = u_{12} + u_{13} + \cdots + u_{1n} + u_{23} + u_{24} \cdots = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} u_{ij}(r_{ij}) \tag{5-2}$$

式中，$r_{ij}$ 为二原子 $i$ 与 $j$ 间的距离。分子内势能则为各类型内坐标（如键伸缩、键角弯曲……）势能的和。

依照经典力学，系统中任一原子 $i$ 所受的力为势能的梯度：

$$\vec{F}_i = -\nabla_i U = -\left(\vec{i}\frac{\partial}{\partial x_i} + \vec{j}\frac{\partial}{\partial y_i} + \vec{k}\frac{\partial}{\partial z_i}\right)U \tag{5-3}$$

依此，由牛顿运动定律可得 $i$ 原子的加速度为：

$$\vec{a}_i = \frac{\vec{F}_i}{m_i} \tag{5-4}$$

将牛顿运动定律方程式对时间积分，可预测 $i$ 原子经过时间 $t$ 后的速度与位置。

$$\begin{aligned} \frac{d^2}{dt^2}\vec{r}_i &= \frac{d}{dt}\vec{v}_i = \vec{a}_i \\ \vec{v}_i &= \vec{v}_i^0 + \vec{a}_i t \\ \vec{r}_i &= \vec{r}_i^0 + \vec{v}_i^0 t + \frac{1}{2}\vec{a}_i t^2 \end{aligned} \tag{5-5}$$

式中，$\vec{r}$ 及 $\vec{v}$ 分别为粒子的位置与速度，上标“0”为各物理量的初始值。

分子动力计算的基本原理，即为利用牛顿运动定律。先由系统中各分子位置计算系统的势能，再由式(5-3) 和式(5-4) 计算系统中各原子所受的力及加速度，然后在式(5-5) 中令 $t=\delta t$，$\delta t$ 表示一非常短的时间间隔，则可得到经过 $\delta t$ 后各分子的位置及速度。重复以上的步骤，由新的位置计算系统的势能，计算各原子所受的力及加速度，预测再经过 $\delta t$ 后各分子的位置及速度……如此反复循环，可得到各时间下系统中分子运动的位置、速度及加速度等资料。一般将各时间下的分子位置称为运动轨迹(trajectory)。图 5-1 为简单的三原子系统的运动图，图中显示各原子的位置及标示原子的速度。

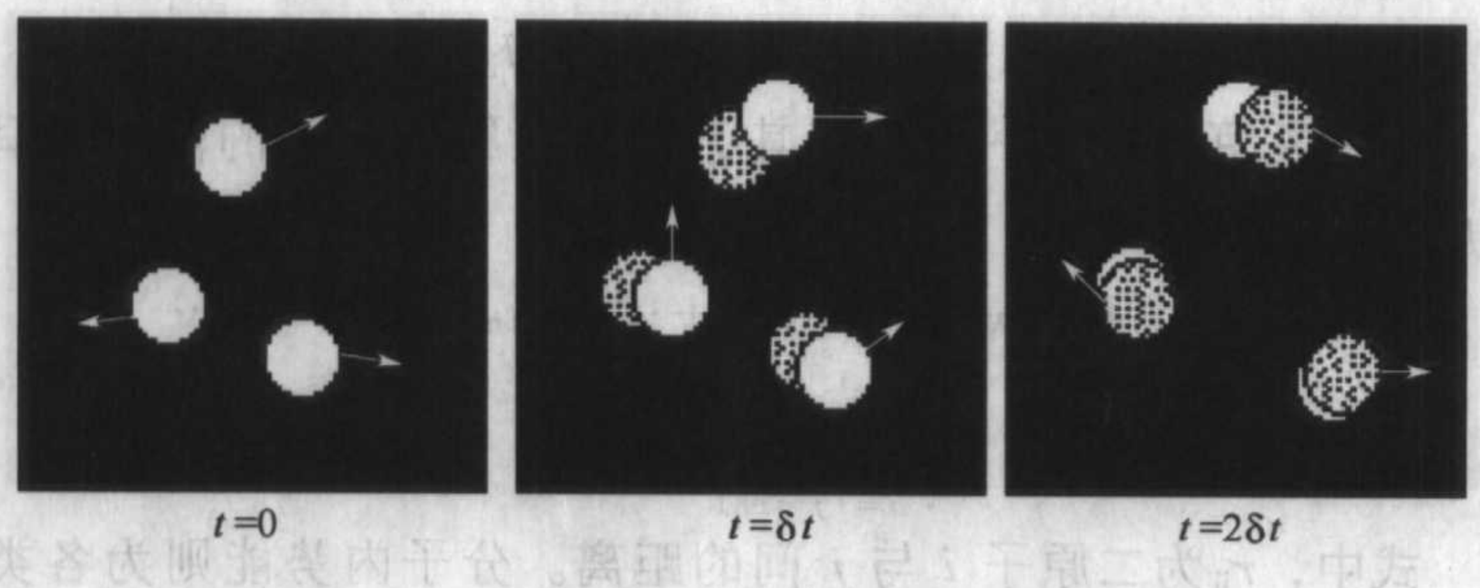



**图 5-1　三原子系统的运动图**

## 5.2 牛顿运动方程式的数值解法

在分子动力计算中必须解式(5-5) 的牛顿运动方程式以计算速度与位置。一般在分子动力计算中，最常用的方法为 Verlet 所发展的数值解法。最早的 Verlet 方法，将粒子的位置以泰勒式展开，即：

$$r(t+\delta t)=r(t)+\frac{\mathrm{d}}{\mathrm{d}t}r(t)\delta t+\frac{1}{2!}\frac{\mathrm{d}^2}{\mathrm{d}t^2}r(t)(\delta t)^2+\cdots \tag{5-6}$$

将式中的 $\delta t$ 换为 $-\delta t$，得：

$$r(t-\delta t)=r(t)-\frac{\mathrm{d}}{\mathrm{d}t}r(t)\delta t+\frac{1}{2!}\frac{\mathrm{d}^2}{\mathrm{d}t^2}r(t)(\delta t)^2+\cdots \tag{5-7}$$

将式(5-6) 与式(5-7) 相加，得：

$$r(t+\delta t)=-r(t-\delta t)+2r(t)+\frac{\mathrm{d}^2}{\mathrm{d}t^2}r(t)(\delta t)^2 \tag{5-8}$$

因 $\frac{\mathrm{d}^2}{\mathrm{d}t^2}r(t)=a(t)$，故依据式(5-8)，可由 $t$ 及 $t-\delta t$ 的位置预测 $t+\delta t$ 时的位置。将式(5-5) 与式(5-6) 相减，得速度式，即：

$$v(t)=\frac{\mathrm{d}r}{\mathrm{d}t}=\frac{1}{2\delta t}[r(t+\delta t)-r(t-\delta t)] \tag{5-9}$$

此式表示，时间 $t$ 时的速度可由 $t+\delta t$ 及 $t-\delta t$ 的位置得到。Verlet 式的缺点在于式(5-9) 中含有 $1/\delta t$ 项，由于实际计算中通常选取很小的 $\delta t$ 值（$\delta t$ 约 $10^{-15}$ s)，容易导致误差。为了矫正此缺点，Verlet 发展出另一种称为跳蛙方法（leap frog method) 的计算式。此方法计算速度与位置的数学式为：

$$\vec{v}_i\left(t+\frac{1}{2}\delta t\right)=\vec{v}_i\left(t-\frac{1}{2}\delta t\right)+\vec{a}_i(t)\delta t$$

$$\vec{r}_i(t+\delta t)=\vec{r}_i(t)+\vec{v}_i\left(t+\frac{1}{2}\delta t\right)\delta t \tag{5-10}$$

计算时假设已知 $\vec{v}_i\left(t-\frac{1}{2}\delta t\right)$ 与 $\vec{r}_i(t)$，则由 $t$ 时的位置 $\vec{r}_i(t)$ 计算质点所受的力与加速度 $\vec{a}_i(t)$。再依式(5-10) 预测时间为 $t+\frac{1}{2}\delta t$ 时的速度 $\vec{v}_i\left(t+\frac{1}{2}\delta t\right)$，依此类推。时间为 $t$ 时的速度可由式 (5-11) 算出，即：

$$\vec{v}_i(t)=\frac{1}{2}\left[\vec{v}_i(t+\frac{1}{2}\delta t)+\vec{v}_i(t-\frac{1}{2}\delta t)\right] \tag{5-11}$$

利用 Verlet 跳蛙法计算仅需储存 $\vec{v}_i\left(t-\frac{1}{2}\delta t\right)$ 与 $\vec{r}_i(t)$ 两种资料，可

节省储存空间。此外，这种方法使用简便且准确性及稳定性均高，至今已被广泛地采用。以 $x$ 轴上的两个粒子为例，其初始位置分别为 $x_1(0)=0.0$，$x_2(0)=2.0$；初速度分别为 $v_1=1.0\times10^{-3}$，$v_2=-0.5\times10^{-3}$。设粒子的质量均为 $m=1.0$，积分时间间隔 $\delta t=0.01$，其范德瓦耳斯势能参数 $\varepsilon=1.0$，$\sigma=1.0$。在任一瞬间二粒子的距离为：

$$r=|x_1-x_2|=[(x_1-x_2)^2]^{1/2}$$

所受的力分别为：

$$F_1=-\frac{\mathrm{d}U}{\mathrm{d}x_1}=-\frac{\mathrm{d}U}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}x_1}=4\varepsilon\left[12\left(\frac{\sigma^{12}}{r^{13}}\right)-6\left(\frac{\sigma^6}{r^7}\right)\right]\left(\frac{\mathrm{d}r}{\mathrm{d}x_1}\right)=4\varepsilon\left(\frac{12\sigma^{12}}{r^{13}}-\frac{6\sigma^6}{r^7}\right)\left(\frac{x_1}{r}\right)$$

$$F_2=-F_1 \tag{5-12}$$

表 5-1 列出利用 Verlet 跳蛙方法计算此二粒子运动 20 步的位置$\vec{r}(t+\delta t)$、速度$\vec{v}\left(t+\frac{1}{2}\delta t\right)$及作用势能 $U$。

**表 5-1　由 Verlet 跳蛙方法计算两个 LJ 粒子的运动**

| 步数($\delta t$) | $x(1)$ | $x(2)$ | $v(1)$ | $v(2)$ | $U$ |
|---|---|---|---|---|---|
| 1 | 0.0019 | 1.9981 | 0.0192 | −0.0187 | −0.0615 |
| 2 | 0.0057 | 1.9944 | 0.0376 | −0.0371 | −0.0622 |
| 3 | 0.0113 | 1.9888 | 0.0564 | −0.0559 | −0.0636 |
| 4 | 0.0189 | 1.9813 | 0.0761 | −0.0756 | −0.0658 |
| 5 | 0.0286 | 1.9717 | 0.096 | −0.0962 | −0.0688 |
| 6 | 0.0405 | 1.9598 | 0.1188 | −0.1183 | −0.0729 |
| 7 | 0.0548 | 1.9456 | 0.1428 | −0.1423 | −0.0784 |
| 8 | 0.0717 | 1.9287 | 0.1694 | −0.1689 | −0.0856 |
| 9 | 0.0916 | 1.9088 | 0.1994 | −0.1989 | −0.0952 |
| 10 | 0.1150 | 1.8855 | 0.2340 | −0.2335 | −0.1080 |
| 11 | 0.1426 | 1.8580 | 0.2752 | −0.2747 | −0.1257 |
| 12 | 0.1751 | 1.8255 | 0.3258 | −0.3253 | −0.1508 |
| 13 | 0.2142 | 1.7865 | 0.3906 | −0.3901 | −0.1882 |
| 14 | 0.2620 | 1.7387 | 0.4783 | −0.4778 | −0.2473 |
| 15 | 0.3225 | 1.6783 | 0.6048 | −0.6043 | −0.3486 |
| 16 | 0.4023 | 1.5985 | 0.7981 | −0.7976 | −0.5404 |
| 17 | 0.5038 | 1.4970 | 1.0154 | −1.0149 | −0.8993 |
| 18 | 0.3324 | 1.6685 | −1.7141 | 1.7146 | 0.1751 |
| 19 | 0.1815 | 1.8195 | −1.5094 | 1.5099 | −0.5796 |
| 20 | 0.0374 | 1.9636 | −1.4414 | 1.4419 | −0.1964 |

除了 Verlet 跳蛙法外，另一种较为常见的 Beeman 方法（Beeman Method），其积分公式为：

$$\vec{r}_i(t+\delta t)=\vec{r}_i(t)+\vec{v}_i(t)\delta t+\frac{1}{6}[4\vec{a}_i(t)-\vec{a}_i(t-\delta t)]\delta t^2$$

$$\vec{v}_i(t+\delta t)=\vec{v}_i(t)+\frac{1}{6}[2\vec{a}_i(t+\delta t)+5\vec{a}_i(t)-\vec{a}_i(t-\delta t)]\delta t \tag{5-13}$$

此方法需储存$\vec{r}_i(t)$、$\vec{v}_i(t)$ 与$\vec{a}_i(t-\delta t)$，储存量大于 Verlet 的跳蛙法。但优点在于可以引用较长的积分间隔 $\delta t$。Levitt 与 Meirovitch 曾比较 Beeman 与 Verlet 跳蛙法的差异。他们研究发现，Beeman 方法所引用的积分步长 $\delta t$ 可为 Verlet 方法的 3～4 倍，而具有相同的准确性。

除了上述两种方法外，还有一种 Gear 所提出利用数值解的方法，称为校正预测法（predictor-corrector method)。将其简介如下。因为经典运动粒子的轨迹为连续的，故于时间 $t+\delta t$ 时的位置、速度等可由时间 $t$ 的泰勒展开式预测得：

$$\vec{r}^{\,p}(t+\delta t)=\vec{r}(t)+\vec{v}(t)\delta t+\frac{1}{2}\vec{a}(t)\delta t^2+\frac{1}{6}\vec{b}(t)\delta t^3+\cdots$$

$$\vec{v}^{\,p}(t+\delta t)=\vec{v}(t)+\vec{a}(t)\delta t+\frac{1}{2}\vec{b}(t)\delta t^2+\cdots$$

$$\vec{a}^{\,p}(t+\delta t)=\vec{a}(t)+\vec{b}(t)\delta t+\cdots$$

$$\vec{b}^{\,p}(t+\delta t)=\vec{b}(t)+\cdots \tag{5-14}$$

式中，$\vec{v}(t)$、$\vec{a}(t)$、$\vec{b}(t)$ 为$\vec{r}(t)$ 的 1 次、2 次及 3 次微分。式(5-14) 所产生的速度、加速度等并非完全正确，因为这些物理量来自泰勒展开式，而非由解牛顿运动方程式而来。可由所预测的位置$\vec{r}^{\,p}(t+\delta t)$ 计算所受的力及正确的加速度$\vec{a}^{\,c}(t+\delta t)$。设正确的加速度与预测的加速度之间的误差为：

$$\Delta\vec{a}(t+\delta t)=\vec{a}^{\,c}(t+\delta t)-\vec{a}^{\,p}(t+\delta t) \tag{5-15}$$

以此为依据，可得各量之校正式为：

$$\vec{r}^{\,c}(t+\delta t)=\vec{r}^{\,p}(t+\delta t)+c_0\Delta\vec{a}(t+\delta t)$$

$$\vec{v}^{\,c}(t+\delta t)=\vec{v}^{\,p}(t+\delta t)+c_1\Delta\vec{a}(t+\delta t)$$

$$\vec{a}^{\,c}(t+\delta t)=\vec{a}^{\,p}(t+\delta t)+c_2\Delta\vec{a}(t+\delta t)$$

$$\vec{b}^{\,c}(t+\delta t)=\vec{b}^{\,p}(t+\delta t)+c_3\Delta\vec{a}(t+\delta t) \tag{5-16}$$

式中，$c_0$、$c_1$、$c_2$、$c_3$ 均为常数。以上所述为 Gear 的一次预测校正法，

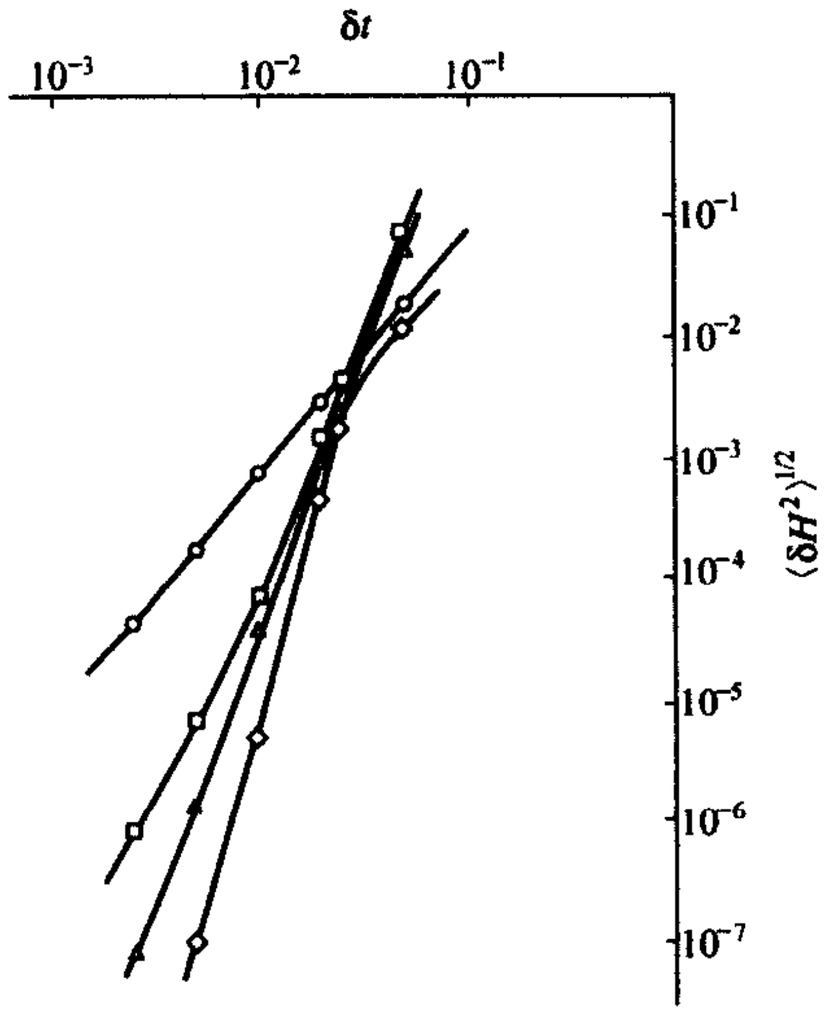



**图 5-2　Verlet 跳蛙法与 Gear 预测校正法的比较**

Verlet 跳蛙法（圆圈），Gear 4 次预测校正法（方块），5 次预测校正法（三角形），6 次预测校正法（菱形）的比较

亦可将此计算法推展至更高次的校正法。图 5-2 比较 Verlet 跳蛙法与 Gear 预测校正法的结果。计算系统为仅有范德瓦耳斯作用的粒子，密度与温度分别为 $\rho^* = 0.6$，$T^* = 1.06$（* 表示简化单位），比较的物理量为均方根内焓 $\langle \delta H^2 \rangle^{1/2}$。由此比较图可知，当 $\delta t$ 减小时，Gear 的高次预测校正法误差较小，但 $\delta t$ 增大时，Verlet 方法较佳。

## 5.3 周期性边界条件与最近镜像

执行分子动力计算通常选取一定数目 $N$ 的分子，将其置于一立方体的盒子中。设盒子的边长为 $L$，则其体积为 $V = L^3$。若分子的质量为 $m$，则系统的密度为：

$$d = \frac{Nm}{L^3} \tag{5-17}$$

计算系统的密度应等于实验所测定的密度。例如欲执行 1000 个水分子的分子动力模拟计算，水的密度为 $1.00 g/cm^3$，依式(5-17)，则有：

$$1.00 g/cm^3 = \frac{1000 \times (18.0/6.20 \times 10^{23}) g}{L^3}$$

$$L = 31.04 \times 10^{-8} cm = 31.04 Å \tag{5-18}$$

模拟计算中，为使计算中系统的密度维持恒定，通常采用周期性边界条件（periodic boundary condition)。以二维的计算系统为例，图 5-3 显示二维盒中系统粒子的排列及移动方向。

图中位于中央的盒子表示所模拟的系统，其周围的盒子与模拟系统具有相同的排列及运动，称为周期性镜像（periodic mirror image）系统。当计算系统中任一粒子移出盒外，则必有一粒子由相对的方向移入，如图中的第 2 个粒子。这样的限制条件使得系统中的粒子数维持恒定，密度不变，符合实际的要求。

计算系统中分子间的作用力时，采取最近镜像方法（nearest mirror image)。如图 5-4 所示，计算分子 1 与 3 的作用力，是取与分子 1 和其最近的距离镜像分子 3。在所有镜像系统中，分子 1，3 距离最近的为计算系统的分子 1 与 D 盒中的分子 3，而非计算系统中的分子 1 与分子 3。同样的，计算分子 3 与分子 1 的作用为取计算系统的分子 3 与 E 盒中的分子 1。

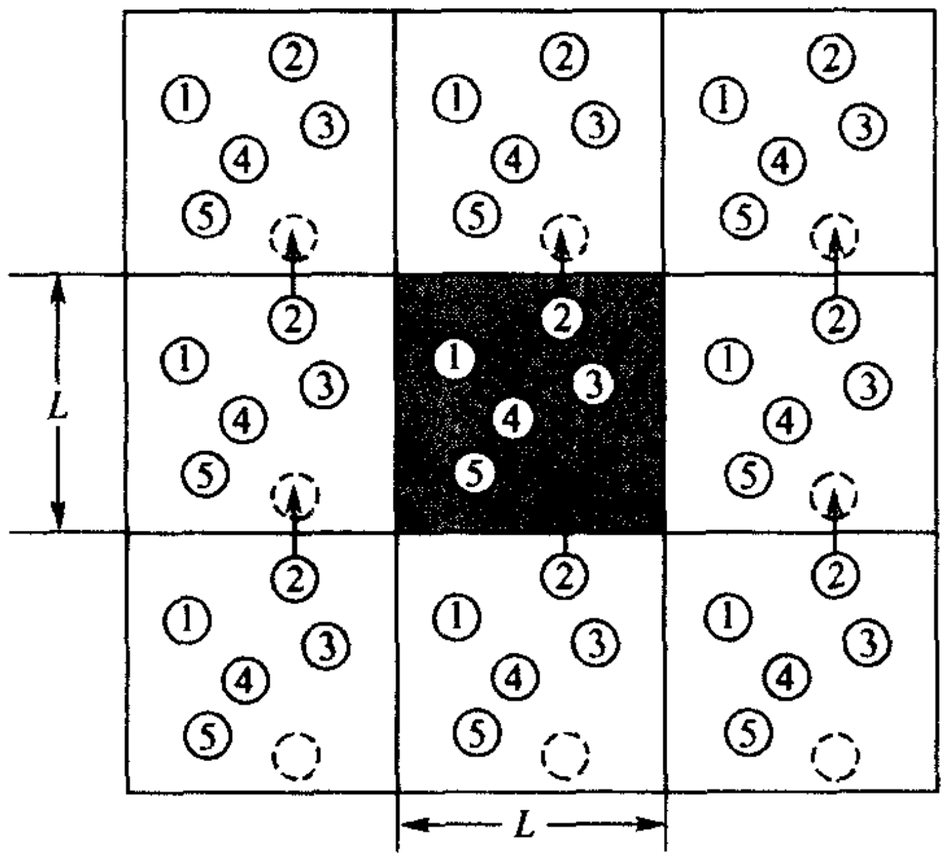



**图 5-3　二维周期性系统的粒子排列与移动**

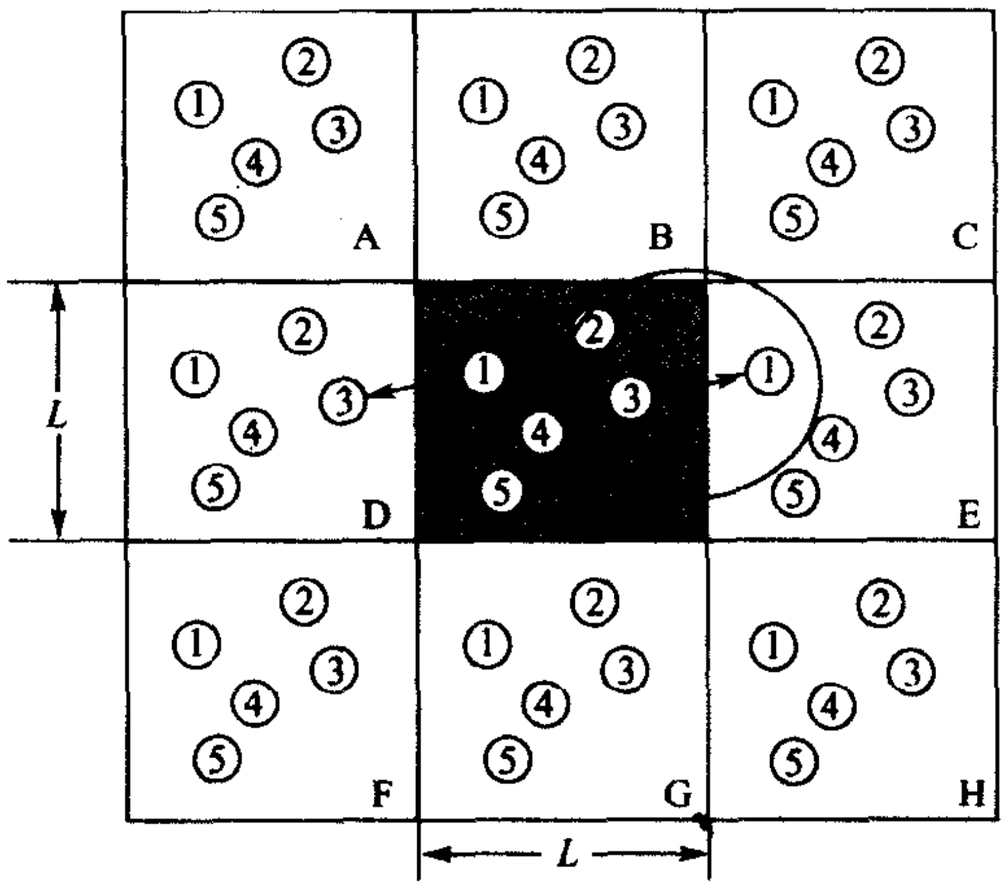



**图 5-4　粒子的最近影像**

因为在计算中利用最近镜像的观念，因此需采用截断半径（cut-off radius）的方法计算非键结的远程作用力，否则会因重复计算粒子所受的力而导致不正确的结果。图 5-5 为一般的范德瓦耳斯势能图。图中显示，当 $r>r_c$ 时，势能的值已非常趋近于零，表示分子间的范德瓦耳斯作用力已可忽略不计，$r_c$ 称为截断半径。在实际的分子动力模拟过程中，若分子间的距离大于截断半径，则将其作用视为零。截断半径最大不能超过盒长的一半，即 $r_c \leqslant L/2$。一般的原子所选取的截断半径约为 10～15Å。

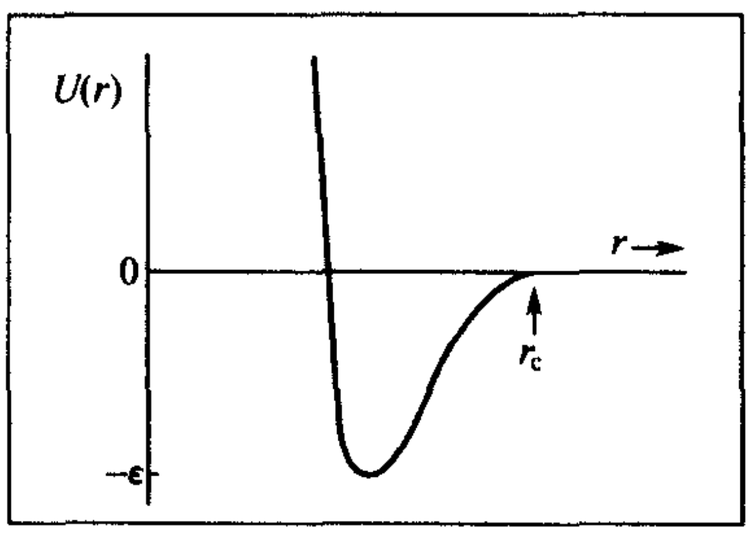



**图 5-5　范德瓦尔斯作用势能图**

一般的非键结势能函数于所选取的截断半径处其势能值虽然很小但并不为零。分子动力计算中，如将其完全忽略不计，会造成能量的不连续情形。处理这种问题的方法，通常将势能函数乘上一个开关函数（switching function）以弥补这一缺点。设截断半径为 $R_c$，开关函数的起始点为 $R_s$，则该函数 $S(r)$ 的形式为：

$$S(r)=1，且\frac{dS(r)}{dr}=0，当 r=R_s$$

$$S(r)=0，且\frac{dS(r)}{dr}=0，当 r=R_c$$

当 $R_s \leqslant r \leqslant R_c$ 时，此函数及其一次微分式分别为：

$$S(r)=\frac{(R_c-r)^2(R_c^2+2r-3R_s)}{(R_c-R_s)^3}$$

$$\frac{dS(r)}{dr}=\frac{6(R_c-r)(R_s-r)}{(R_c-R_s)^3} \tag{5-19}$$

图 5-6 显示利用开关函数所调整的非键结势能图。

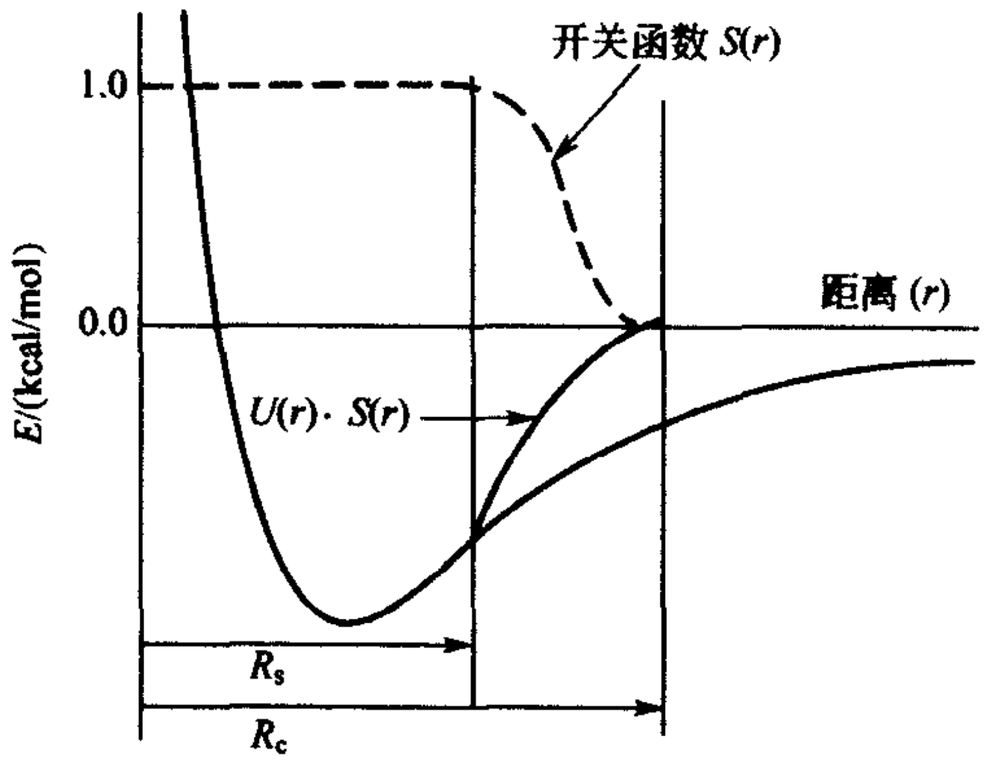



**图 5-6　利用开关函数调节非键结势能函数**

由此图可知，利用开关函数可将由 $R_s \rightarrow R_c$ 区间的非键结势能函数逐渐调整为零，形成连续且符合计算要求的势能。通常 $R_s$ 不宜过小，一般约取 $0.9R_c$。

## 5.4 积分步程的选取

在分子动力计算中，最重要的工作为如何选取适当的积分步程（integration time step）$\delta t$，以节省计算的时间而不失去其精准性。通常的原则为积分步程应小于系统中最快运动周期的 1/10。以水分子的动力计算为例，水分子的内运动为键长的变化，键角的变化；其分子间的运动为质心的移动与分子的转动。分子间的运动源自范德瓦耳斯作用，一般较慢，但分子内的运动则较快。由红外光谱的证据可知其最大的振动频率约为 $1.08\times10^{14}\,s^{-1}$，亦即最快的运动为每秒振动 $1.08\times10^{14}$ 次，其振动周期为 $T=1/v=0.92\times10^{-14}\,s$。因此，分子动力计算的积分步程最大 $\delta t$ 约为 $0.9\times10^{-16}\,s$。再以氩原子的分子动力计算为例，氩的原子量为 39.95 原子质量单位，其 12-6 作用势能的参数为 $\sigma=3.504\text{Å}$，$\epsilon=0.24\text{kcal/mol}$（1kcal=4.184kJ）。因为势能函数最低点的两次微分相当于力常数 $k$，即：

$$U''(r)=\frac{d}{dr}U'(r)=\frac{d}{dr}[4\epsilon(-12\sigma^{12}/r^{13}+6\sigma^{6}/r^{7})]$$
$$=4\epsilon(156\sigma^{12}/r^{14}-42\sigma^{6}/r^{8}) \tag{5-20}$$

当势能有极小值时，$r_{min}=\sqrt[6]{2}\sigma$，将此代入式(5-20) 得：

$$U''(r_{min})=57.14\epsilon/\sigma^{2} \tag{5-21}$$

由氩原子的势能参数得 $U''(r_{min})=1.117\text{kcal/(mol}\cdot\text{Å}^2)$。令力常数 $U''(r_{min})=k$，借由简谐振子的关系式知：

$$k=4\pi^2 v^2\mu \tag{5-22}$$

式中，$\mu$ 为简化质量（reduced mass）。氩原子的简化质量为：

$$\mu=\frac{m_{Ar}\times m_{Ar}}{m_{Ar}+m_{Ar}}=19.98\times1.662\times10^{-22}(g) \tag{5-23}$$

由此得频率：

$$v=\left[\frac{57.14\times0.24\times6.9446\times10^{-14}\,erg}{(3.504\times10^{-8}\,cm)^2\times4\pi^2\times19.8\times1.662\times10^{-22}\,g}\right]=7.72\times10^{10}\,s^{-1} \tag{5-24}$$

故其周期为 $T=1.3\times10^{-11}\,s$，积分步程可取为 $\delta t$ 约 $1.3\times10^{-12}\,s$。比较以上的结果可知以分子动力方法计算氩原子的积分步程可取为计算水分子积分步程的 1000 倍。分子动力计算中，积分步程愈大则可研究的时间范围愈长。例如，欲研究时间范围为 $10^{-9}\,s$ 的氩原子系统，仅仅需要 800 步的分子动力计算，但研究相同时间范围内的水分子系统则至少需 $1.05\times10^{6}$

步的分子动力计算。以上由 12-6 势能估计原子系统的积分步程的方法亦可应用于不同原子的混合系统或多原子分子系统。通常取 12-6 势能中 $\epsilon$ 值最大者以估计原子间运动最快者的积分步程大小。例如水分子间的范德瓦耳斯作用包括氧原子-氧原子、氧原子-氢原子、氢原子-氢原子等原子对间的作用。因氧原子-氧原子的 $\epsilon$ 值最大，故应以此为估计积分步程的标准。表 5-2 为各种模拟系统所建议采取的积分步长。

**表 5-2　各种模拟系统所建议采取的积分步长**

| 系　统 | 运 动 型 态 | 积分步长/s |
|---|---|---|
| 简单原子体系 | 移动 | $10^{-14}$ |
| 刚性分子 | 移动与转动 | $5\times10^{-15}$ |
| 软性分子，限制键长 | 移动，转动，扭动 | $2\times10^{-15}$ |
| 软性分子，无键长限制 | 移动，转动，扭动，振动 | $10^{-15}$ 或 $5\times10^{-16}$ |

## 5.5　简化单位

分子动力方法计算系统中原子或分子的运动。若采用 cgs 制，原子质量以 g 为单位，则原子质量的量纲为 $10^{-22}$（g），位置以 cm 为单位之量纲为 $10^{-8}$（cm），积分步程以秒为单位之量纲约为 $10^{-13}\sim10^{-16}$（s）。这些量的量纲均很小，可能在模拟中导致计算的误差。实际执行分子动力计算时则采用简化单位（reduced unit）以减少误差。表 5-3 列出各种物理量简化单位的转换式。

**表 5-3　各种物理量简化单位的转换式**

| 简化物理量 | | 转换式 | 简化物理量 | | 转换式 |
|---|---|---|---|---|---|
| 数目密度 | $\rho^*$ | $\rho\times\sigma^3$ | 时间 | $t^*$ | $(\epsilon/m\sigma^2)^{1/2}t$ |
| 温度 | $T^*$ | $k_BT/\epsilon$ | 力 | $f^*$ | $f\sigma/\epsilon$ |
| 能量 | $E^*$ | $E/\epsilon$ | 力矩 | $\tau^*$ | $\tau/\epsilon$ |
| 压力 | $P^*$ | $\rho\sigma^3/\epsilon$ | | | |

表中，$\sigma$ 与 $\epsilon$ 为 12-6 势能参数。质量的简化单位为原子的质量，长度的简化单位为 $\sigma$。以氩分子系统为例，氩分子之原子量为 39.95 原子质量单位，熔点为 84K，12-6 势能参数为 $\sigma=3.504$Å，$\epsilon=0.24$kcal/mol。温度为 84K 时，氩原子系统为液体，系统的密度为 1.784g/cm$^3$。设分子动力

计算采用的积分步程为 $10^{-12}$s，系统的温度为 84K，压力为 1atm。依表(5-3)，将各相关的物理量以简化单位表示，得：

密度 $d=M/V=1.784\text{g/cm}^3$

数目密度为 $\rho=(1.784\text{g}/78\text{g})\times 6.02\times 10^{23}/\text{cm}^3=0.13769\times 10^{23}/\text{cm}^3$

简化密度为 $\rho^*=\rho\times\sigma^3=(0.13769\times 10^{23}/\text{cm}^3)\times(3.504\times 10^{-8}\text{cm})^3=0.5924$

温度 $T=84\text{K}$

简化温度为 $T^*=(1.3806\times 10^{-16}\text{erg/K})\times 84\text{K}/[(0.24\text{kcal/mol})\times 6.9446\times 10^{-14}\text{erg}/(\text{kcal}\cdot\text{mol})]=6.958\times 10^{-1}$

压力 $P=1\text{atm}=101325\text{bar}=101325\text{kg}/(\text{m}\cdot\text{s}^2)=1.101325\times 10^6\text{g}/(\text{cm}\cdot\text{s}^2)$

简化压力为 $P^*=1.01325\times 10^6\text{g}/(\text{cm}\cdot\text{s}^2)\times(3.5\times 10^{-8}\text{cm})^3/(0.24\times 6.9446\times 10^{-14}\text{erg})=0.0026$

积分步程 $\delta t=10^{-12}\text{s}$

简化积分步程为 $\delta t^*=\{0.24\text{kcal/mol}\times 6.9446\times 10^{-14}\text{erg}/[39.95\times 1.66\times 10^{-22}\text{g}\times(3.504\times 10^{-8}\text{cm})^2]\}^{1/2}\times 10^{-12}\text{s}=0.0452$

若是系统中含有不只一种原子，则选取其中一种原子的 LJ 势能参数 $\sigma$ 与 $\epsilon$ 为简化单位的标准。例如水分子系统含有 O 与 H 两种原子，通常取 O 原子的 $\sigma$ 与 $\epsilon$ 为转换简化单位的标准。

## 5.6 分子动力计算流程

真正执行分子动力计算前，必须先行估计计算的可行性。一般分子动力计算的容量约为数千个原子的系统，如果系统过大，则超过计算的范围。含 $N$ 个原子的系统，每一计算步骤需要计算 $(1/2)N(N+1)$ 组远程作用力，以此部分的计算最为耗时。若原子数目增加 1 倍，则计算时间增加 4 倍。除了系统的容量外，亦需考虑研究对象的时间范围。通常分子动力计算所选取的积分步程约为飞秒 (femto second，fs)($1\text{fs}=10^{-15}\text{s}$)。若以目前一般的工作站或较强的个人计算机从事 1000 个原子系统的计算，累积一百万步，相当于 $10^{-9}$s 的系统时间的轨迹约需两星期的计算时间。因此，由实际的角度看来，分子动力计算适合于研究较 $10^{-9}$s 更快的运动，而不适合研究较慢的运动。例如蛋白质的折叠运动通常需要毫秒级

(纳秒，ns)（1ns=$10^{-3}$s)，若利用分子动力计算研究折叠运动，则需要非常长的时间。

执行分子动力计算的起点，将一定数目的分子置于立方体的盒中，使其密度与实验的密度相符，再选定计算的温度，即可着手开始计算。计算时，必须知道系统中分子的初始位置与速度。通常可将分子随机置于盒中，或是取其结晶形态的位置排列为初始位置。系统中所有原子运动的动能总和应满足式(5-25) 的条件。

$$\text{K.E} = \sum_{i=1}^{N} \frac{1}{2} m_i (|\vec{v}_i|)^2 = \frac{3}{2} N k_B T \tag{5-25}$$

式中，K.E 为系统的总动能；$N$ 为总原子数；$k_B$ 为玻耳兹曼常数；$T$ 为热力学温度。原子运动的初速度可依此关系式产生。例如，可取一半的原子向右运动，而另外一半向左运动；而原子的总动能为 $3/2Nk_BT$。或是令原子的初速度呈高斯分布（Gaussian distribution)，而总动能为$3/2Nk_BT$。产生原子的起始位置与初速度后，则可进行分子动力计算。详细的初始位置与速度的产生方法将于下节中讨论。

由初始位置与速度开始，计算的每一步产生新的速度与位置，由新产生的速度可计算系统的温度 $T_{cal}$ 为：

$$T_{cal} = \frac{\sum_{i=1}^{N} m_i (v_{i,x}^2 + v_{i,y}^2 + v_{i,z}^2)}{3Nk_B} \tag{5-26}$$

若系统的计算温度与所定的温度相比过高或过低，则需校正速度。一般允许的温度范围为：

$$0.9 \leqslant \frac{T_{cal}}{T} \leqslant 1.1 \tag{5-27}$$

若计算的温度超过此允许范围，则将所有原子的速度乘以一校正因子，即：

$$f = \sqrt{\frac{T}{T_{cal}}} \tag{5-28}$$

使得系统的计算温度重新调整为：

$$\frac{\sum_{i=1}^{N} m_i (v_{i,x}^2 + v_{i,y}^2 + v_{i,z}^2) \times f^2}{3Nk_B} = \frac{T}{T_{cal}} \cdot T_{cal} = T \tag{5-29}$$

将计算的温度校正回到系统的设定温度。实际执行分子动力计算的过程，于计算开始时每隔数步即需校正速度；随后校正的间隔增长；每隔数

百步或数千步才需校正。直至原子的速度不需再校正，而系统的总动能在 $3/2Nk_BT$ 上下呈现约 10% 涨落，此时系统称为已达热平衡（thermal equilibrium）状态。在到达热平衡状态前的轨迹与速度不需保存，因其物理意义不够周严，仅当系统达平衡后，才开始储存计算的轨迹与速度。通常的分子动力计算需要累积数百万步的运动轨迹以供分析，因而导致储存的问题。以 1000 个原子的系统为例，储存每步轨迹与速度约需 10kB 的硬盘容量，储存 10000 步则需 100MB 的容量。因为分子动力计算中，积分步程很小，每一步原子移动的幅度有限，故常每隔 10～20 步存取一次轨迹与速度以节省硬盘的容量。图 5-7 为 Verlet 跳蛙法的分子动力计算的流程图。这种计算方法称为一般性分子动力计算（conventional molecular dynamics），计算中系统的原子数 $N$、体积 $V$ 与总能量 $E$ 维持不变，相当于统计力学中的微正则系综（micro-canonical ensemble），记为（$N$，$V$，$E$）系综。

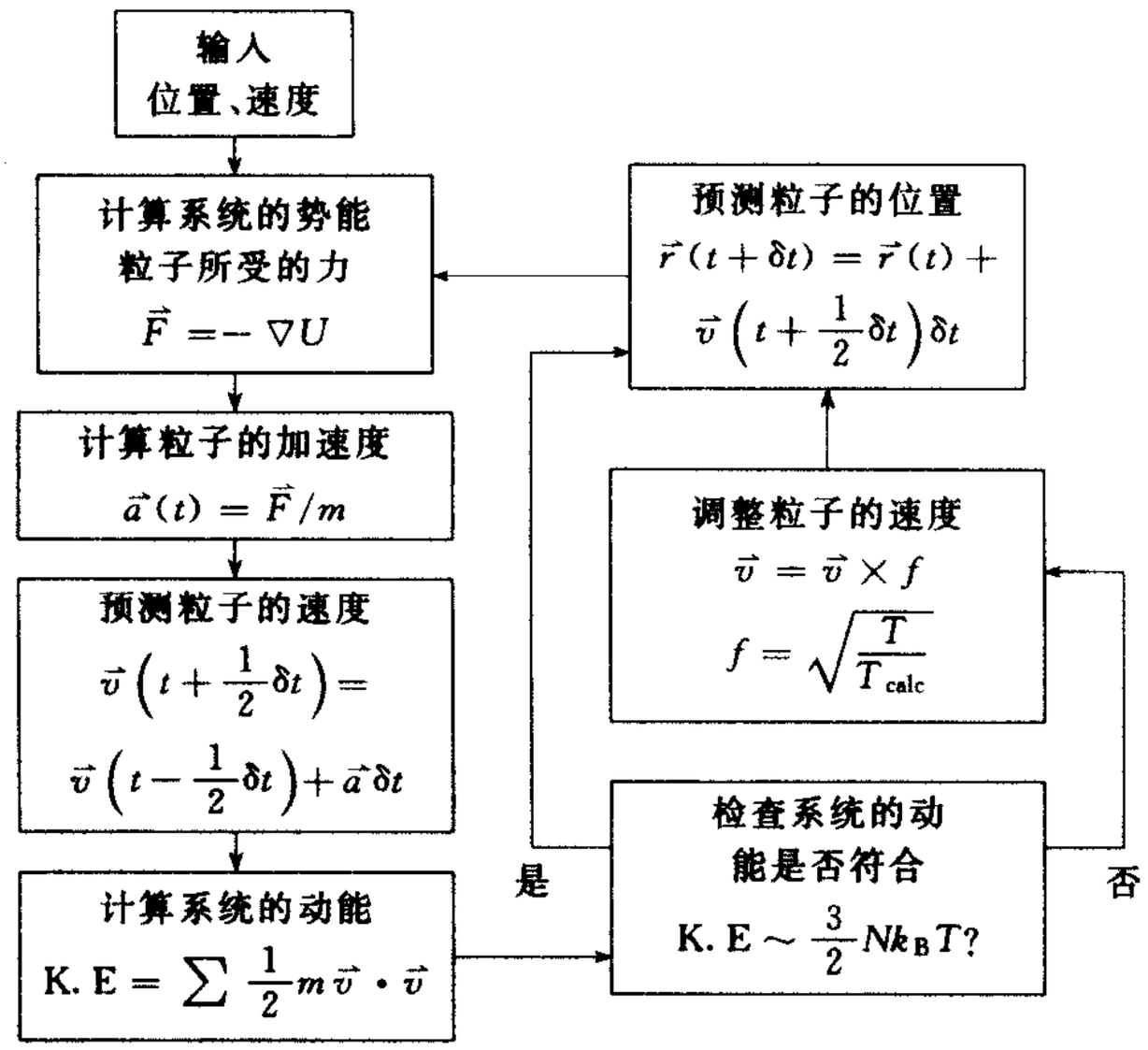



**图 5-7　Verlet 跳蛙法分子动力计算流程**

## 5.7　分子动力计算的初始设定

执行分子动力计算必须选取适当的初始条件，如起始位置、速度、执

行温度、积分步长等。若初始条件选择不当，则往往需要浪费相当长的计算时间才能达到系统的热平衡，或甚至根本无法平衡。初始结构的选取，愈近似模拟系统的结构愈佳。例如欲模拟面心立方的固体系统而选择体心立方作为初始结构即为不明智的选择。通常模拟时最好执行一些试验性的预算，以确定设定的系统中没有太高能量的作用，否则会导致模拟的不稳定。有时可利用能量最小化的方法，以最低能量的结构为模拟的起点，以避免高能作用的产生。

均匀相的液体系统的初始位置通常定为其晶体的结构。若模拟的系统已有实验的 X 射线衍射晶体结构，则建议以此结构为计算的起点。若无此实验的结构可参照，则通常以均匀分布的面心立方结构为起点，如图 5-8 所示。

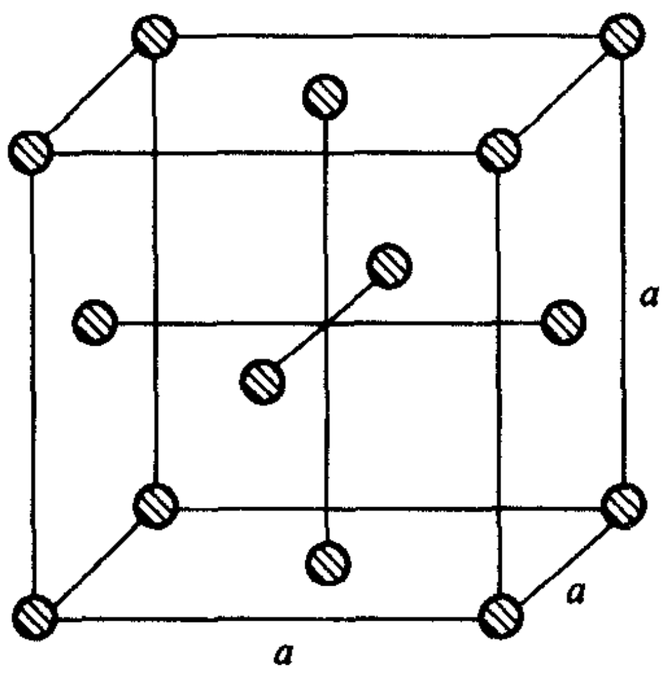



**图 5-8　面心立方的结构**

这种面心立方的立体结构所含的粒子数为 $4M^3$，$M=2$，3，4，…对应于 32，108，256，…个粒子。模拟分子液体的初始结构，通常将分子的质心置于上述的面心立方点上，而分子的位向则视模拟的系统而定。若分子的体积很小，如 $CO_2$，$H_2O$ 等，则可随意指定其初始位向，但若分子的体积很大（如戊烷），则位向的排列以分子间不产生体积的重叠为原则。初始位向愈近似系统的平衡结构愈佳，例如模拟杆状液晶分子的初始位向最好将分子平行排列。

模拟复杂系统如溶液，聚合物的初始结构最好能参照 X 射线衍射或核磁共振 NMR 实验的结果。若无这些资料可依据，则可由分子力学能量最小化的结果为模拟起点。特殊的系统往往需要特殊的方法产生起始结构，例如 Chen 等利用逐步缩小体积方法以产生无定形（amorphous）固态聚合物的初始结构。

通常系统粒子初速度的产生，可于 $-1\sim1$ 间选取一合于高斯分布的随机数，将此随机数乘以粒子的平均速率，即：

$$\bar{v}=\frac{k_B T}{m} \tag{5-30}$$

得到符合麦克斯韦-玻耳兹曼分布（Maxwell-Boltzmann distribution）的粒子速度，即：

$$p(v_\alpha)=\left(\frac{m}{2\pi k_B T}\right)^{1/2}\exp\left[-\frac{1}{2}\frac{mv_\alpha^2}{k_B T}\right]\quad \alpha=x,y,z \tag{5-31}$$

通常计算前检查粒子的速度分布以使各方向的总动量为零，否则计算的系统本身产生移动而导致总能量的不稳定。

## 5.8 定温计算法

以上所介绍的分子一般性动力计算适用于（$N$，$V$，$E$）系综，计算的系统总能量恒定，温度于其平衡值附近扰动。除了（$N$，$V$，$E$）系综外，分子动力计算亦可处理其他类型的系综，如（$N$，$V$，$T$）、（$N$，$P$，$T$）、（$N$，$E$，$H$）等。使用何种系综进行计算得视实际的体系需要而定。例如研究材质的相变化，多采用（$N$，$P$，$T$）系综；在定压下计算各温度时系统的能量或比热容，由此判断相变化的温度。而研究一般溶液的特性，如水合能（hydration energy），解离效应等则多采用正则系综。此处简单介绍两种定温计算法，有兴趣的读者可参考其原始文献，或参考 Daan Frenkel 与 Berend Smit 所著“Understand Molecular Simulation”一书。以 Anderson 推导的（$N$，$V$，$T$）系综为例，系统中粒子的位置与速度分别为：

$$\vec{r}(t+\delta t)=\vec{r}(t)+\vec{v}(t)\delta t+\frac{\vec{F}(t)}{2m}\delta t^2$$

$$\vec{v}(t+\delta t)=\vec{v}(t)+\frac{\vec{F}(t+\delta t)+\vec{F}(t)}{2m}\delta t \tag{5-32}$$

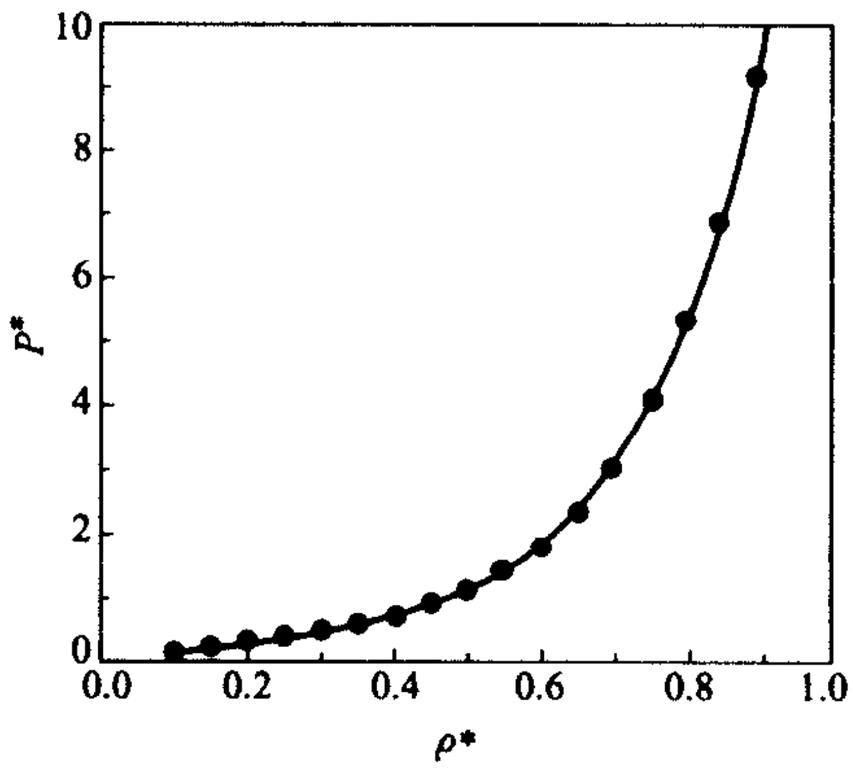



**图 5-9 Anderson 定温计算法 108 个 LJ 粒子的压力与密度关系**

图 5-9 为由 Anderson 定温方法计算得到的系统的简化压力与简化密度的关系图。计算的样品含 108 个粒子，粒子间的作用为范德瓦耳斯作用，势能型式为 Lennard-Jones（LJ）势能，系统的简化温度为 $T^*=2.0$；图中的实线为由热力学系统的状态方程式计算所得，黑点为利用 Anderson 定温分子动力模拟的结果，两者十分吻合。

除了 Anderson 的方法外，Nosé 与 Hoover 推导出另一种恒温计算法（Nosé-Hoover thermostat method）。这

方法是以一变量 $s$ 调节过高或过低的动量 $\vec{p}$，使系统的温度维持不变。调节后的物理量为：

$$\vec{v}' = \vec{v}$$
$$\vec{p}' = \vec{p}/s$$
$$s' = s$$
$$\delta t' = \delta t/s \tag{5-33}$$

式中，上标（′）表示调整后的量。由此可得运动方程式为：

$$\frac{\mathrm{d}\vec{r}_i}{\mathrm{d}t} = \frac{\vec{p}_i}{m_i}$$
$$\frac{\mathrm{d}\vec{p}_i}{\mathrm{d}t} = -\nabla_i U - \xi \vec{p}_i$$
$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = \left(\sum_{i=1}^{N} \frac{p_i^2}{m_i} - \frac{3N}{k_B T}\right)/Q$$
$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)/s = \xi \tag{5-34}$$

式中，$U$ 为系统的势能函数；$Q$ 称为有效质量（effective mass），为人为设定值，通常选为 5 或 10；$\xi$ 为热力摩擦系数（thermodynamic friction coefficient）。执行分子动力计算时可将以上的运动方程式积分计算各时间 $t$ 的位置，动量与 $s$ 值。Nosé-Hoover 的定温计算法较 Andersen 的计算法稳定，可用于调整未平衡体系的温度。而 Anderson 的定温计算则较适用于已达热平衡的体系。图 5-10 为 Nosé-Hoover 计算的 LJ 粒子体系

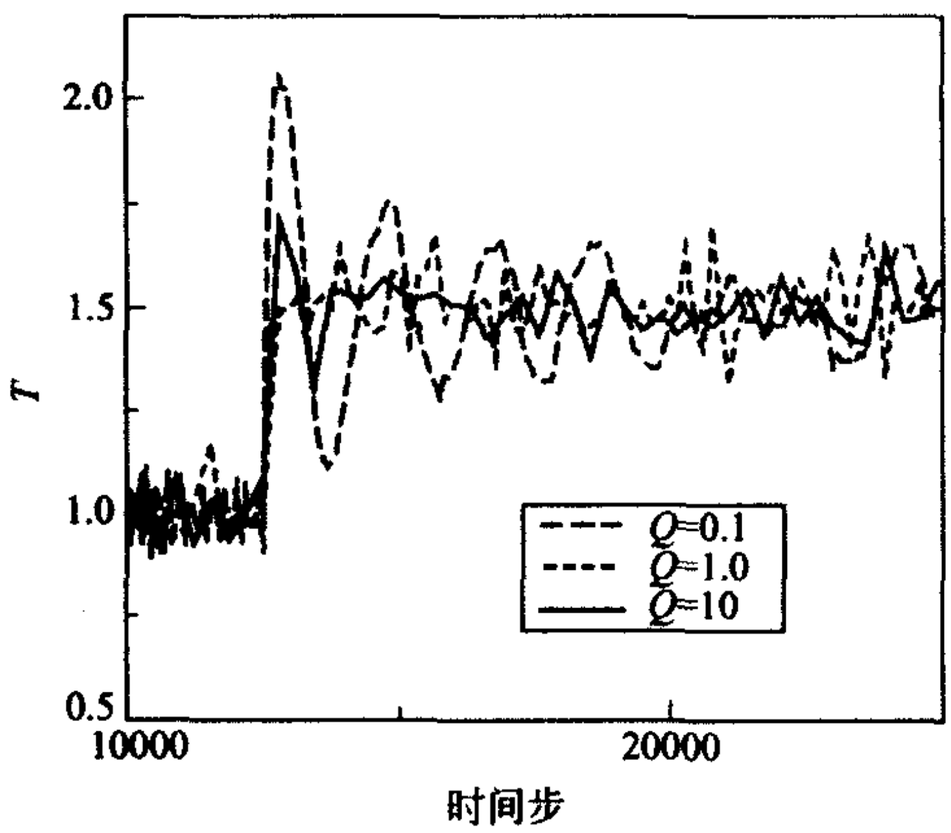



**图 5-10　Nosé-Hoover 计算系统温度变化与有效质量的关系**

（$N=256$，$\rho=0.75$）当温度突然提升时，系统的温度扰动与有效质量的关系图。此图显示，$Q$ 值越大，系统的温度愈容易趋于平衡。

其余由 Brown 及 Clark 所测试的各种系综，利用 Verlet 跳蛙方法的分子动力计算法列于本章的附录中。

## 5.9 系统形状的改变计算

一般的分子动力模拟计算是将分子置于立方体的盒子中，但依据实际的模拟情形，往往需要改变盒子的形状。例如，计算高分子体系常利用长方形的盒子。Parrinello 与 Rahman 提出处理这种形状改变的通用方法，此方法利用度量化坐标（scaled coordinates），即：

$$\vec{r}=\widetilde{H}\,\vec{s} \tag{5-35}$$

式中，$\widetilde{H}=(\vec{h}_1,\ \vec{h}_2,\ \vec{h}_3)$ 为转换矩阵，其中的行向量 $\vec{h}$ 表示盒子的边长向量，如图 5-11 所示。

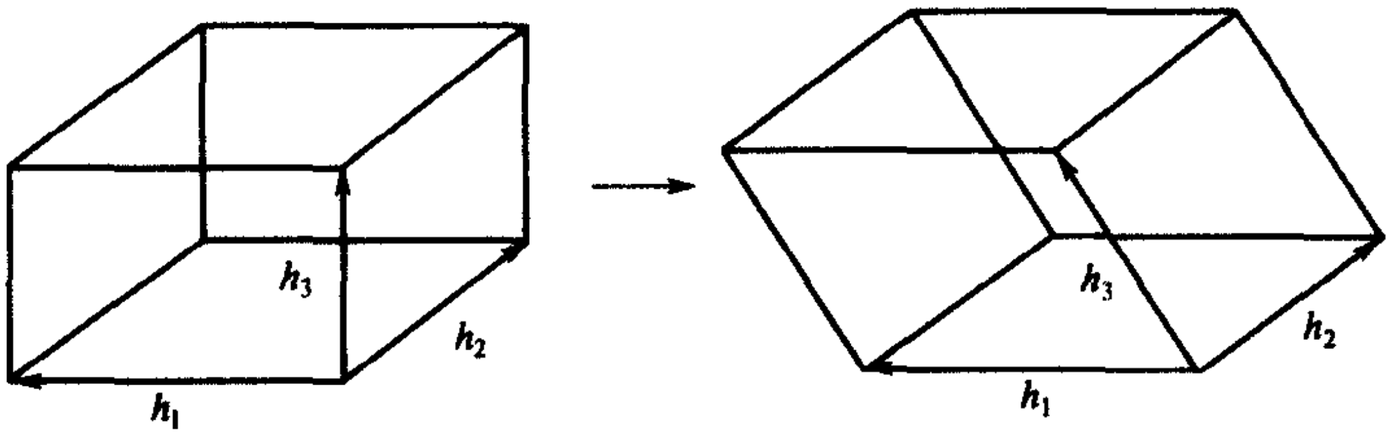



**图 5-11　系统形状改变的示意**

盒子的体积为：

$$V=|\widetilde{H}|=\vec{h}_1\cdot\vec{h}_2\times\vec{h}_3 \tag{5-36}$$

与盒子体积相关的势能为：

$$U_V=PV \tag{5-37}$$

与盒子体积变化相关的动能为：

$$K_V=\frac{1}{2}Q\sum_{\alpha}\sum_{\beta}\dot{H}_{\alpha\beta}^2 \tag{5-38}$$

式中，$Q$ 为盒子的“质量”。利用拉格朗日乘子 $L=K+K_V+U+U_V$ 可得运动方程式，即：

$$m\ddot{s} = \tilde{H}^{-1}\vec{f} - m\tilde{G}^{-1}\dot{\tilde{G}}\dot{s}$$

$$Q\ddot{\tilde{H}} = (\tilde{P} - \tilde{1}P)V(\tilde{H}^{-1})^T \tag{5-39}$$

式中，$\tilde{G}=\tilde{H}^T\tilde{H}$ 为度规张量（metric tensor）；上标“$T$”为转置矩阵的符号；$\tilde{P}$ 为压力张量（pressure tensor），即：

$$P_{\alpha\beta} = \frac{1}{V}\left[\sum_i m_i(\tilde{H}\dot{s}_i)_\alpha^T(\tilde{H}\dot{s}_i)_\beta + \sum_i\sum_{j>1}(\tilde{H}s_{ij})_\alpha(\vec{f}_{ij})_\beta\right] \tag{5-40}$$

$\vec{f}_{ij}$为未度量化坐标系中 $i$ 粒子所受到 $j$ 粒子的力。运动方程式可利用预测-校正方法求解。改变形状的模拟计算多见于固体的体系，尤其是系统受到压力而产生形变。

## 5.10 各种系综的分子动力计算方法

此节以 Verlet 的跳蛙法为例列举各种不同系综的分子动力计算法。

**(1)** $(N, V, E)$ **系综**

$$\vec{v}_i\left(t+\frac{1}{2}\delta t\right) = \vec{v}_i\left(t-\frac{1}{2}\delta t\right) + \frac{F_i(t)}{m}\delta t$$

$$\vec{r}_i(t+\delta t) = \vec{r}_i(t) + \vec{v}_i\left(t+\frac{1}{2}\delta t\right)\delta t$$

$$\vec{v}_i(t) = \frac{1}{2}\left[\vec{v}_i\left(t-\frac{1}{2}\delta t\right) + \vec{v}_i\left(t+\frac{1}{2}\delta t\right)\right] \tag{5-41}$$

**(2)** $(N, V, T)$ **系综**

① 速度校正法（rescaling velocity method）

$$\vec{v}_i\left(t+\frac{1}{2}\delta t\right) = \vec{v}_i\left(t-\frac{1}{2}\delta t\right)\beta + \frac{\vec{F}_i(t)}{m}\delta t$$

$$\beta^2 = \frac{[3(N-1)k_B T_v/m]}{\sum_{i=1}^{N}\left[\vec{v}_i\left(t-\frac{1}{2}\delta t\right)\right]^2} \tag{5-42}$$

式中，$\beta$为速度校正因子；$T_v$ 为所设定的温度；$k_B$ 为玻耳兹曼常数。

此方法的$\vec{r}_i$ $(t+\delta t)$ 与$\vec{v}_i(t)$ 的计算法与方法（1）相同。

② 阻尼力方法（damped force method） 此方法将哈密顿运动方程式修正为：

$$\dot{\vec{p}}=\vec{F}-\alpha\frac{\vec{p}}{m}\ \text{或}\ \dot{\vec{v}}=\frac{\vec{F}-\alpha\vec{v}}{m} \tag{5-43}$$

式中，$\alpha$ 为一常数。实际计算时并不必直接计算 $\alpha$，而可由式(5-44)计算速度。

$$\vec{v}_i\left(t+\frac{1}{2}\delta t\right)=\vec{v}_i\left(t-\frac{1}{2}\delta t\right)(2\beta-1)+\beta\frac{\vec{F}_i(t)}{m}\delta t \tag{5-44}$$

式(5-44) 中，

$$\beta^2=\frac{[3(N-1)k_B T_v/m]}{\sum_{i=1}^{N}\vec{v}'(t)^2}$$

$$\vec{v}_i'(t)=\vec{v}_i\left(t-\frac{1}{2}\delta t\right)+\frac{\vec{F}_i(t)}{2m}\delta t \tag{5-45}$$

此方法的 $\vec{r}_i(t+\delta t)$ 与 $\vec{v}_i(t)$ 的计算法与方法（1）相同。

**(3)（$N$，$P$，$H$）系综** 此方法多应用于计算发生体积变化的系统。计算中，系统的体积 $V$ 并非定值。定义度量化的速度为：

$$\vec{q}_i=\frac{\vec{r}_i}{V^{1/3}},\ \dot{\vec{q}}=\frac{\vec{p}_i}{mV^{1/3}} \tag{5-46}$$

则对应的哈密顿运动方程式为：

$$\ddot{\vec{q}}=\frac{\vec{F}_i}{mV^{1/3}}-\frac{2}{3}\frac{\dot{\vec{q}}\cdot\dot{V}}{V}$$

$$\ddot{V}=\frac{P-P_E}{M} \tag{5-47}$$

式中，$M$ 为系统所有粒子的质量和；$P$ 为系统所计算的压力，即：

$$P=\left(\frac{1}{3V}\right)\left[\sum_{i=1}^{N}m\left(\frac{\vec{p}_i}{m}\right)\cdot\left(\frac{\vec{p}_i}{m}\right)+\sum_{i=1}^{N}\sum_{j>i}^{N}\vec{r}_{ij}\cdot\vec{F}_{ij}\right] \tag{5-48}$$

式中，$\vec{r}_{ij}=\vec{r}_i-\vec{r}_j$；$\vec{F}_{ij}$ 为第 $i$ 个质点所受来自第 $j$ 个质点的力。

$P_E$ 为外压，即计算所设定的压力值。速度的计算式为：

$$\vec{v}_i\left(t+\frac{1}{2}\delta t\right)=\vec{v}_i\left(t-\frac{1}{2}\delta t\right)+\left\{\frac{\vec{F}_i(t)}{m}+\frac{1}{3}\frac{\vec{r}_i(t)}{V(t)}\left[\ddot{V}(t)-\frac{2}{3}\left(\frac{\dot{V}(t)}{V(t)}\right)^2\right]\right\}\delta t \tag{5-49}$$

此方法的 $\vec{r}_i(t+\delta t)$ 与 $\vec{v}_i(t)$ 的计算法与方法（1）相同。

必须注意此计算法中系统的体积随时间而变，若引用周期性边界条件，则第 $i$ 个原子的镜像原子 $x$ 坐标的微分式（$y$，$z$ 类似）为：

$$\dot{x}_i' = \dot{x}_i \pm \frac{1}{3}\frac{\dot{V}}{V}L \tag{5-50}$$

式中，$L$ 为立方体盒子的边长。

**(4)** （$N$，$P$，$T$）**系综**　此系统的计算结合方法（2）②与方法(3)。其速度式为：

$$\vec{v}_i\left(t+\frac{1}{2}\delta t\right)=\vec{v}_i\left(t-\frac{1}{2}t\right)(2\beta-1)+\left[\frac{2}{3}\frac{\vec{r}_i(t)\dot{V}(t)}{V(t)}\right](1-\beta)+$$

$$\beta\left\{\frac{\vec{F}_i(t)}{m}+\frac{1}{3}\frac{\vec{r}_i(t)}{V(t)}\left[\ddot{V}(t)-\frac{2}{3}\left(\frac{\dot{V}(t)}{V(t)}\right)^2\right]\right\}\delta t \tag{5-51}$$

式(5-51) 中，

$$\beta^2=\frac{[3(N-1)k_B T_v/m]}{\sum_{i=1}^{N}\left(\frac{\vec{p}_i'(t)}{m}\right)^2}$$

$$\frac{\vec{p}_i'(t)}{m}=\vec{v}_i'(t)-\frac{1}{3}\frac{\dot{V}(t)}{V(t)}$$

$$\vec{v}_i'(t)=\vec{v}_i\left(t-\frac{1}{2}\delta t\right)+\frac{1}{3}\left(\frac{\vec{r}_i(t)}{V(t)}\right)\left[\ddot{V}(t)-\frac{2}{3}\left(\frac{\dot{V}(t)}{V(t)}\right)^2\right]\times\frac{1}{2}\delta t \tag{5-52}$$

**(5) 各系综计算结果的比较**　Brown 与 Clark 等以 256 个氩原子液体为例，对上述各系统进行分子动力计算以检验其结果。氩原子间的作用势能为：

$$U(r)=4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right],\ r<r_C$$

$$U(r)=0, r\geqslant r_C \tag{5-53}$$

式中，$\sigma=3.405\text{Å}$，$\epsilon/k_B=119.8\text{K}$，$r_C=2.5\sigma$，氩原子质量为 $6.63382\times10^{-23}\text{g}$。其简化分子密度 $\rho^*=N\sigma^3/V$ 为 0.83134，简化温度 $T^*=Tk_B/\epsilon$ 为 0.722。各系统皆执行 5000 步的分子动力计算，每步的时间为 $2.0\times10^{-14}\text{s}$。图 5-12 为计算的均方根位移。

$$\langle R^2\rangle(t)=\langle N^{-1}\sum_{i=1}^{N}[\vec{r}_i(t)-r_i(0)]^2\rangle \tag{5-54}$$

由此，可计算扩散系数［方法（1）］，即：

$$D=\lim_{t\to\infty}\langle R^2\rangle(t)/6t \tag{5-55}$$

图 5-13 为各种系综计算的速度相关函数。

$$C_v(t)=\langle\vec{v}_i(0)\cdot\vec{v}_i(t)\rangle \tag{5-56}$$

由此亦可计算扩散系数［方法（2）］，即：

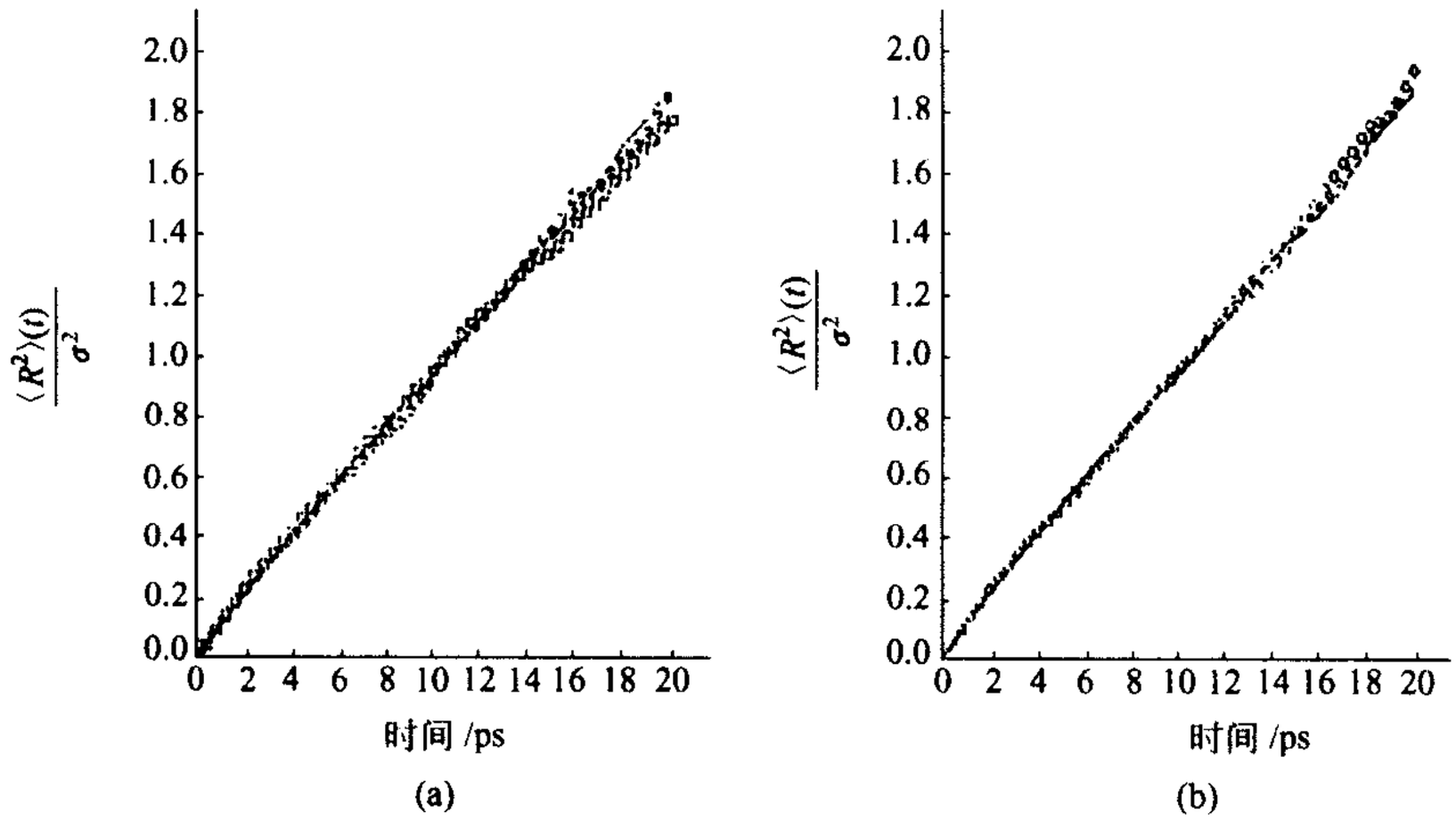



**图 5-12　利用各种系综计算 LJ 系统的均方根位移**

（a）中实线为（$N$，$V$，$E$），倒三角为（$N$，$V$，$T$）$^a$，方块为（$N$，$V$，$T$）$^b$ 系综；

（b）中，菱形为（$N$，$P$，$H$），圆形为（$N$，$P$，$T$）系综

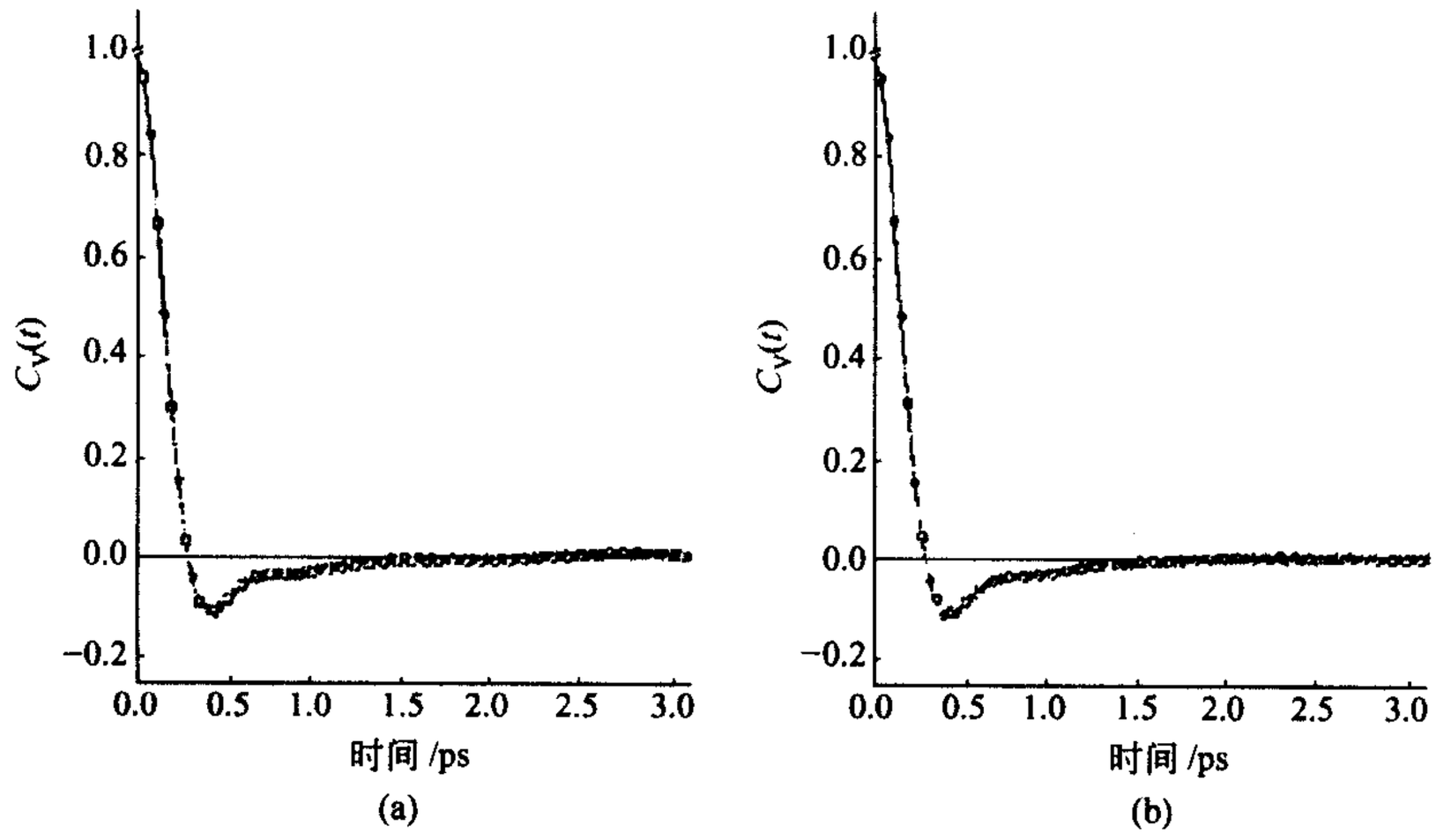



**图 5-13　利用各种系综计算 LJ 系统的速度相关函数**

（a）中实线为（$N$，$V$，$E$），倒三角为（$N$，$V$，$E$）$^a$，方块为（$N$，$V$，$T$）$^b$ 系综；

（b）中，菱形为（$N$，$P$，$H$），圆形为（$N$，$P$，$T$）系综的计算结果

$$D = \frac{1}{3}\int_0^\infty C_v(t)\mathrm{d}t \tag{5-57}$$

表 5-4 列出由各种系综计算的能量及各种热力性质与其均方根误差。值得注意的是（$N$，$P$，$H$）计算所得压力的误差 $\langle \Delta P^2 \rangle^{1/2}$ 比其余各种系综计算的误差更大。

**表 5-4　各种系综计算的能量及各种热力性质与其均方根误差**

| 项　目 | $(N,V,E)$ | $(N,V,T)^a$ | $(N,V,T)^b$ | $(N,P,H)$ | $(N,P,T)$ |
|---|---|---|---|---|---|
| $\langle U \rangle$/(J/mol) | −4914.6 | −4897.4 | −4906.9 | −4914.7 | −4898.4 |
| $\langle \Delta U^2 \rangle^{1/2}$/(J/mol) | 1.5 | 46.8 | 47.7 | 2.8 | 82.9 |
| $\langle H \rangle$/(J/mol) | — | — | — | −4996.4 | −4980.3 |
| $\langle \Delta H^2 \rangle^{1/2}$/(J/mol) | — | — | — | 1.5 | 82.2 |
| $\langle \Phi \rangle$/(J/mol) | −5988.9 | −5977.7 | −5981.1 | −5979.8 | −5972.6 |
| $\langle \Delta \Phi^2 \rangle^{1/2}$/(J/mol) | 34.8 | 46.8 | 47.7 | 51.7 | 83.1 |
| $\langle P \rangle / \times 10^5$Pa | −29.1 | −1.5 | −4.5 | −29.2 | −29.1 |
| $\langle \Delta P^2 \rangle^{1/2} / \times 10^5$Pa | 69.2 | 101.6 | 100.4 | 124.5 | 122.5 |
| $\langle v \rangle \sigma^3$ | 307.937 | 307.937 | 307.937 | 308.583 | 308.883 |
| $\langle \Delta V^2 \rangle^{1/2} \sigma^3$ | — | — | — | 3.223 | 4.088 |
| $\langle T(t) \rangle$/K | 86.47 | 86.96 | 86.47 | 85.73 | 86.47 |
| $\langle \Delta T(t)^2 \rangle^{1/2}$/K | 2.77 | 0.87 | 0.00 | 4.08 | 0.00 |

表 5-5 为由式(5-55) 及式(5-57) 两种不同方法计算的扩散系数的比较。由表中数据可知，不论采用何种系综，由式(5-55) 及式(5-57) 两种方法计算的扩散系数结果非常一致。

**表 5-5　各种系综由两种计算方法所计算的扩散系数 $D$（$\times 10^{-9}\mathrm{m}^2/\mathrm{s}$）的比较**

| 方法 | $(N,V,E)$ | $(N,V,T)^a$ | $(N,V,T)^b$ | $(N,P,H)$ | $(N,P,T)$ |
|---|---|---|---|---|---|
| 1(A-15) | 1.76 | 1.75 | 1.71 | 1.78 | 1.80 |
| | ±0.09 | ±0.10 | ±0.14 | ±0.09 | ±0.09 |
| 2(A-17) | 1.83 | 1.79 | 1.71 | 1.79 | 1.81 |
| | ±0.08 | ±0.08 | ±0.08 | ±0.08 | ±0.09 |

以上，以简单系统的计算显示各种系统的计算方法。在实际处理问题时，应视需要，选取适当的方法。譬如，在相变化时系统的温度与压力为定值，因此最适当的系统应为（$N$，$P$，$T$)。

## 参　考　文　献


Verlet 速度计算法

1　Swope W C，Anderson H C，Berens P H，Wilson K R. J. Chem.. Phys.，76，637，1982

2　Van Gunsteren W F，Berendsen H J C. Mol. Phys.，34，1311，1977

Verlet 计算法

3 Verlet L. Physical Review, 165, 201, 1967

4 Gear C W. "Numerical initial value problems in ordinary differential equations" Prentice-Hall, Englewood Cliffs, NJ. USA, 1971

跳蛙计算法

5 Hockney R W. Comput. Phys., 9, 136, 1970

6 Fincham D, Heyes D M. CCP5 Quarterly, 6, 4, 1982

分子动力模拟系统与方法

7 van Gunsteren W F, Berendsen H J C. Ahgew. Chem. Int. Ed., 29, 990, 1990

8 Leach A R. "Molecular Modeling", Addison Wesley Longman Limitted, England, 1996

9 Allen M P, Tildesley D J. "Computer Simulation of Liquids", Clarendon Press, Oxford, England, 1987

10 Frenkel D, Smit B. "Understanding Molecular Simulation", Academic Press, San Diego, USA, 1996

11 Chen C L, Lee C L, Chen H L, Shih J H. Macromolecules, 27, 7872, 1994

12 Chen C L, Chen H L, Lee C L, Shih J H. Macromolecules, 27, 2087, 1994

定温定压计算法

13 Anderson H C. J. Chem. Phys., 72 (4), 2384, 1980

14 Anderson H C, Allen M P, Bellemans A, Board J, Clarke J H R, Ferrario M, Halle J M, Nosé S, Opheusden J V, Ryckaert J P. Rapport d'activité scientifique du CECAM, pp. 82, 1984

15 Nosé S, Klein M L. Mol. Phys., 50, 1055, 1983

16 Nosé S. Mol. Phys., 52. 255, 1984a

17 Nosé S J. Chem. Phys., 81, 511, 1984b

18 Nosé S. Prog. Theoret. Phys. Supplement, 103, 1, 1991

19 Brown D, Clark J H R. Mol. Phys., 51, 1243, 1984

形变

20 Parrinello M, Rahman A. J. Appl. Phys., 52, 7182, 1981

# 第6章 特殊形态的分子动力计算

第5章叙述了分子动力模拟的一般方法，虽然分子动力模拟的计算能力随着计算机速度与日俱增，但通常能够计算只有数纳秒内上万个分子体系的变化，与实际的体系（$10^{20}$个分子，微秒或更长的变化时间）相距甚远，因此，改进计算方法为分子动力模拟中的最重要的问题之一。本章介绍了多种特殊形态的分子动力模拟方法，利用它们可以大大简化计算量，从而提高分子动力模拟的实用性。

## 6.1 刚性线形分子的计算

执行分子动力计算时，若系统的热力学性质或宏观性质为主要的研究对象，则可忽略分子内部的各种运动。在此种条件下，常采取刚性分子(rigid molecule) 的计算法。刚性计算方法将分子中原子的相对位置固定，分子内部无任何键长或键角的变化。此种计算法的好处在于容许选取较大的积分步长、增加计算的效率。

刚性分子的运动可分为质心的移动与分子整体的转动运动两部分。分子质量中心（center of mass，CM）的坐标 $\vec{R}_{cm}$ 为：

$$\vec{R}_{cm} = \sum_{a=1}^{n} \frac{m_a \vec{r}_a}{M} \tag{6-1}$$

式中，$n$ 为分子所含的原子数；$m_a$ 与 $\vec{r}_a$ 分别为 $a$ 原子的质量与其位置向量；$M$ 为分子的总质量。对线形的刚性分子（如 HCN、$CO_2$）而言，其角速度与力矩 $\vec{\tau}^s$ 必垂直其分子轴。设 $\vec{e}^s$ 为沿着分子轴的单位向量，则其与力矩的关系为：

$$\vec{\tau}^s = \vec{e}^s \times \vec{g}^s \tag{6-2}$$

式中，$\vec{g}^s$ 取决于分子间的相互作用力。分子中的原子相对其质心的位置为：

$$\vec{d}_a^s = d_a \vec{e}^s \tag{6-3}$$

式中，$d_a$ 为 $a$ 原子相对其质心的距离。则：

$$\vec{g}^s = \sum_a d_a \vec{f}_a^s \tag{6-4}$$

$\vec{f}_a^s$ 为 $a$ 原子所受的力。可将式(6-4) 中的向量 $\vec{g}$ 分解为垂直与平行于分子轴的两个分量，而仅有垂直于分子轴的分量与分子的力矩有关，即：

$$\vec{\tau}^s = \vec{e}^s \times \vec{g}' \tag{6-5}$$

此处垂直于分子轴的分量 $\vec{g}'$ 可由式(6-6) 求出。

$$\vec{g}' = \vec{g}^s - (\vec{g}^s \cdot \vec{e}^s)\vec{e}^s \tag{6-6}$$

分量 $\vec{g}'$ 具有 $|\vec{g}'|^2 = |\vec{\tau}^s|^2$ 的特性，可利用此关系计算分子的力矩平方的平均值。

根据以上的关系，线形刚性分子的转动运动可写成两个一次微分运动方程式，即：

$$\dot{\vec{e}}^s = \vec{\mu}^s$$

$$\dot{\vec{\mu}}^s = \vec{g}'/\tilde{I} + \lambda \vec{e}^s \tag{6-7}$$

式中，$\vec{\mu}$ 为角速度；$\tilde{I}$ 为分子的转动惯量（moment of inertia)；$\lambda$ 为拉格朗日未定乘子（Lagrange undetermined multiplier)。利用 $\vec{e}^s(t) \cdot \vec{\mu}^s(t) = 0$ 与 $\vec{e}^s(t) \cdot \vec{g}'(t) = 0$ 的性质，Fincham 以 Verlet 的跳蛙方法解此微分方程式，推出如下的计算式，即：

$$\vec{\mu}^s\left(t + \frac{1}{2}\delta t\right) = \vec{\mu}^s\left(t - \frac{1}{2}\delta t\right) + \dot{\vec{\mu}}^s(t)\delta t$$

$$\vec{e}^s(t + \delta t) = \vec{e}^s(t) + \vec{\mu}^s\left(t + \frac{1}{2}\delta t\right)\delta t \tag{6-8}$$

分子的移动为其质心的移动。质心所受的力为分子的原子所受力的总和，即：

$$\vec{f}_{cm} = \sum_a \vec{f}_a \tag{6-9}$$

质心的运动方程式以 Verlet 跳蛙法解得：

$$\vec{v}_{cm}\left(t + \frac{1}{2}\delta t\right) = \vec{v}_{cm}\left(t - \frac{1}{2}\delta t\right) + \frac{\vec{f}_{cm}(t)}{M}\delta t$$

$$\vec{R}_{cm}(t+\delta t)=\vec{R}_{cm}(t)+\vec{v}_{cm}\left(t+\frac{1}{2}\delta t\right)\delta t \tag{6-10}$$

式中，$M$ 为分子的总质量。

## 6.2 刚性非线形分子的计算

对非线形多原子分子而言，造成分子转动的力矩为：

$$\vec{\tau}=\sum_a(\vec{r}_a-\vec{R}_{cm})\times\vec{f}_a=\sum_a\vec{d}_a\times\vec{f}_a \tag{6-11}$$

分子的位向定义分子固定轴（molecular-fixed axis，$x^by^bz^b$）与空间固定轴（space-fixed axis，$xyz$）间之关系。一般分子（固定）轴的选取方法为使分子的转动惯量的张量呈对角线化形态（$I_{xx}$，$I_{yy}$，$I_{zz}$）的轴。设空间中的一个单位向量于分子轴坐标系的位置为 $\vec{e}^b$，在空间固定轴坐标系的位置为 $\vec{e}^s$。两者间的关系为：

$$\vec{e}^b=\tilde{A}\cdot\vec{e}^s \tag{6-12}$$

式中，旋转矩阵 $\tilde{A}$ 的 9 个元素与欧拉角（Euler angle）$\phi$、$\theta$、$\psi$ 相关。图 6-1 为欧拉角的定义。

$\tilde{A}$ 矩阵的元素为

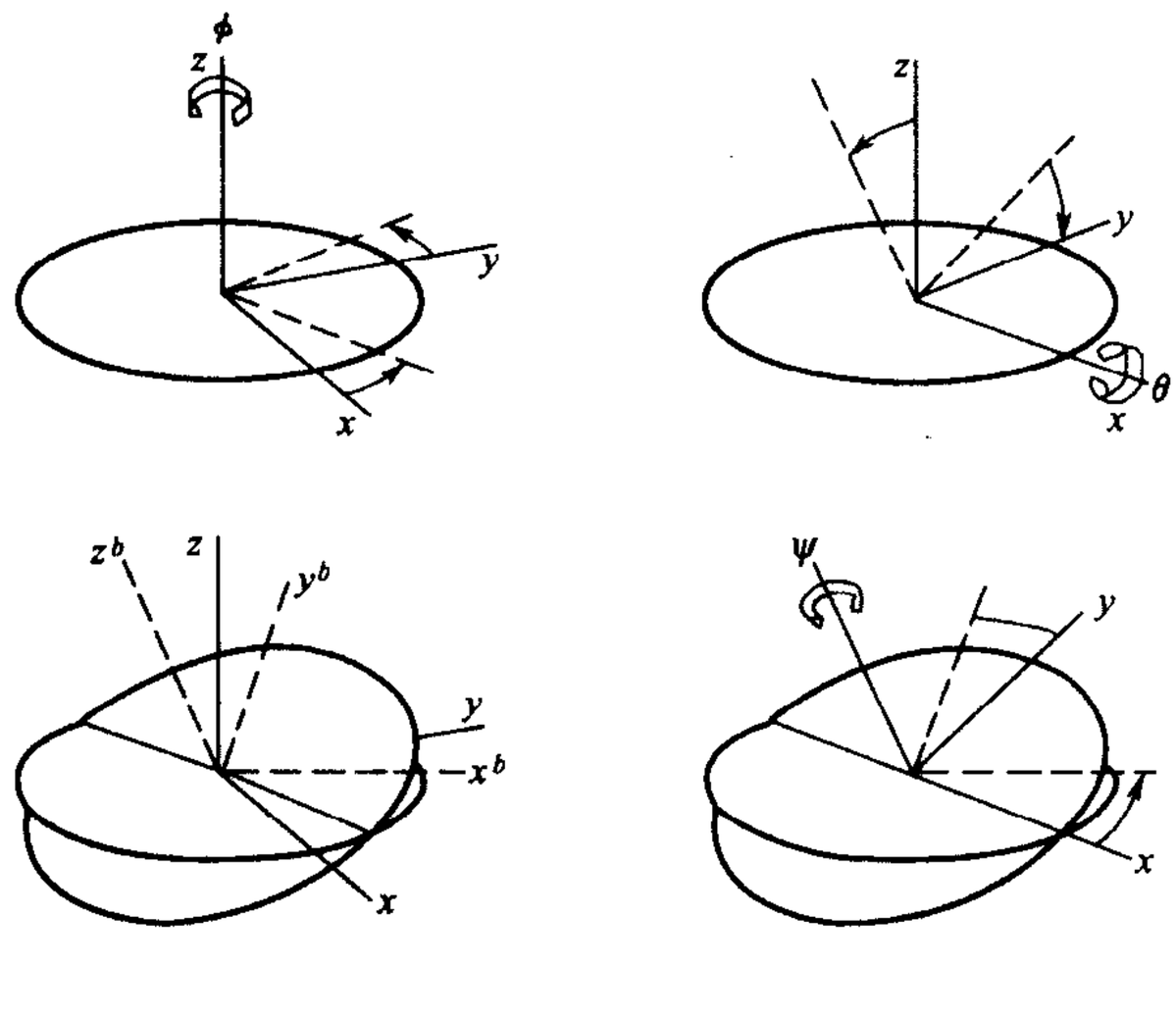



**图 6-1　欧拉角的定义**

$$\widetilde{A}=\begin{pmatrix}\cos\phi\cos\psi-\sin\phi\cos\theta\sin\psi & \sin\phi\cos\psi+\cos\phi\cos\theta\sin\psi & \sin\theta\sin\psi\\ -\cos\phi\sin\psi-\sin\phi\cos\theta\cos\psi & -\sin\phi\sin\psi+\cos\phi\cos\theta\cos\psi & \sin\theta\cos\psi\\ \sin\phi\sin\theta & -\cos\phi\sin\theta & \cos\theta\end{pmatrix} \tag{6-13}$$

根据以上的推导可得分子固定坐标轴的角速度的运动方程式为：

$$\begin{aligned}\dot{\omega}_x^b&=\frac{\tau_x^b}{I_{xx}}+\left(\frac{I_{yy}-I_{zz}}{I_{xx}}\right)\omega_y^b\omega_z^b\\ \dot{\omega}_y^b&=\frac{\tau_y^b}{I_{yy}}+\left(\frac{I_{zz}-I_{xx}}{I_{yy}}\right)\omega_z^b\omega_x^b\\ \dot{\omega}_z^b&=\frac{\tau_z^b}{I_{zz}}+\left(\frac{I_{xx}-I_{yy}}{I_{zz}}\right)\omega_x^b\omega_y^b\end{aligned} \tag{6-14}$$

式中，$I_{xx}$、$I_{yy}$与 $I_{zz}$为分子固定轴的 3 个转动惯量。

分子固定轴与空间固定轴二坐标系统的力矩与角速度间的关系为：

$$\begin{aligned}\vec{\tau}^b&=\widetilde{A}\cdot\vec{\tau}^s\\ \vec{\omega}^s&=\widetilde{A}^{-1}\cdot\vec{\omega}^b=\widetilde{A}^T\cdot\vec{\omega}^b\end{aligned} \tag{6-15}$$

式中，$\widetilde{A}^{-1}=\widetilde{A}^T$。依此，可得到欧拉角的运动方程式，即：

$$\begin{aligned}\dot{\phi}&=-\omega_x^s\frac{\sin\phi\cos\theta}{\sin\theta}+\omega_y^s\frac{\cos\phi\cos\theta}{\sin\theta}+\omega_z^s\\ \dot{\theta}&=\omega_x^s\cos\phi+\omega_y^s\sin\phi\\ \dot{\psi}&=\omega_x^s\frac{\sin\phi}{\sin\theta}-\omega_y^s\frac{\cos\phi}{\sin\theta}\end{aligned} \tag{6-16}$$

由式(6-14) 与式(6-16) 可解得瞬时的角速度与欧拉角的变化，而得到分子的位向与转动的速度。虽然，原则上这样的计算可行，但实际上，因式(6-16) 的分母有 $\sin\theta$ 项，故当 $\theta\sim0$ 或 $\theta\sim\pi$ 时易造成很大的计算误差。为改进此缺点，Evans 与 Murad 发展出新的计算法，其方法简述如下。

定义 $\vec{Q}$ 为含有 4 个分量的向量，即：

$$\vec{Q}=(q_0,q_1,q_2,q_3) \tag{6-17}$$

此 4 个分量合于如下的限制条件，即：

$$q_0^2+q_1^2+q_2^2+q_3^2=1 \tag{6-18}$$

此 4 个分量与分子的位向角相关，其关系为：

$$
\begin{aligned}
q_0 &= \cos\frac{1}{2}\theta\cos\frac{1}{2}(\phi+\psi)\\
q_1 &= \sin\frac{1}{2}\theta\cos\frac{1}{2}(\phi-\psi)\\
q_2 &= \sin\frac{1}{2}\theta\sin\frac{1}{2}(\phi-\psi)\\
q_3 &= \cos\frac{1}{2}\theta\sin\frac{1}{2}(\phi-\psi)
\end{aligned}
\tag{6-19}
$$

将式(6-13) 的转动关系矩阵以此 4 个分量表示为：

$$
\tilde{A}=\begin{bmatrix}
q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2+q_0q_3) & 2(q_1q_3-q_0q_2)\\
2(q_1q_2-q_0q_3) & q_0^2-q_1^2+q_2^2-q_3^2 & 2(q_2q_3+q_0q_1)\\
2(q_1q_3+q_0q_2) & 2(q_2q_3-q_0q_1) & q_0^2-q_1^2-q_2^2+q_3^2
\end{bmatrix}
\tag{6-20}
$$

每一分子的 $\vec{Q}$ 向量满足如下的运动方程式，即：

$$
\dot{\vec{Q}}=\frac{1}{2}\tilde{q}\cdot\vec{\omega}^b,\quad
\begin{pmatrix}\dot{q}_0\\ \dot{q}_1\\ \dot{q}_2\\ \dot{q}_3\end{pmatrix}=\frac{1}{2}\begin{pmatrix}q_0 & -q_1 & -q_2 & -q_3\\ q_1 & q_0 & -q_3 & -q_2\\ q_2 & q_3 & -q_0 & -q_1\\ q_3 & -q_2 & q_1 & q_0\end{pmatrix}\begin{pmatrix}0\\ \omega_x^b\\ \omega_y^b\\ \omega_z^b\end{pmatrix}
\tag{6-21}
$$

利用 Verlet 跳蛙方法解此微分式的步骤如下。设已知分子于时间 $t$ 时在空间固定坐标系统的位置，则可计算此时的 $\tilde{A}$ 与 $\vec{\tau}^s$。若已知时间为 $t-\frac{1}{2}\delta t$ 时的角动量，则利用跳蛙方法可计算分子在 $t$ 时的角动量，即：

$$
\vec{l}^s(t)=\vec{l}^s\left(t-\frac{1}{2}\delta t\right)+\vec{\tau}^s(t)\delta t
\tag{6-22}
$$

此定义于分子轴的角动量可利用 $\vec{l}^b=\tilde{A}\cdot\vec{l}^s$、$\tilde{I}^b=\tilde{A}\tilde{I}^s\tilde{A}^T$ 的关系转换为定义于空间固定轴的角动量。再利用 $\vec{\omega}^b=\tilde{A}\cdot\vec{\omega}^s$ 的关系，可得定义于分子轴的角速度 $\vec{\omega}^b(t)$，即：

$$
\begin{aligned}
&\vec{l}^s(t)=\tilde{I}^s(t)\cdot\vec{\omega}^s(t)\\
&\tilde{A}(t)\cdot\vec{l}^s(t)=\tilde{A}(t)\tilde{I}^s(t)\tilde{A}^T(t)\tilde{A}(t)\cdot\vec{\omega}^s(t)\\
&\tilde{A}(t)\cdot\vec{l}^s(t)=\tilde{I}^b\cdot\vec{\omega}^b(t)\\
&\vec{\omega}^b(t)=(\tilde{I}^b)^{-1}\tilde{A}(t)\cdot\vec{l}^s(t)
\end{aligned}
\tag{6-23}
$$

将此代入式(6-21) 可求得向量 $Q$ 于时间 $t$ 的微分 $\dot{\vec{Q}}(t)$。再由跳蛙法求得：

$$\vec{Q}\left(t+\frac{1}{2}\delta t\right),\vec{Q}\left(t+\frac{1}{2}\delta t\right)=\vec{Q}(t)+\frac{1}{2}\dot{\vec{Q}}(t)\delta t \tag{6-24}$$

这种计算方法可改进前述式(6-16) 方法的缺点，已被广泛地应用于刚体分子的分子动力计算。图 6-2 为非线形刚性分子计算的流程图。

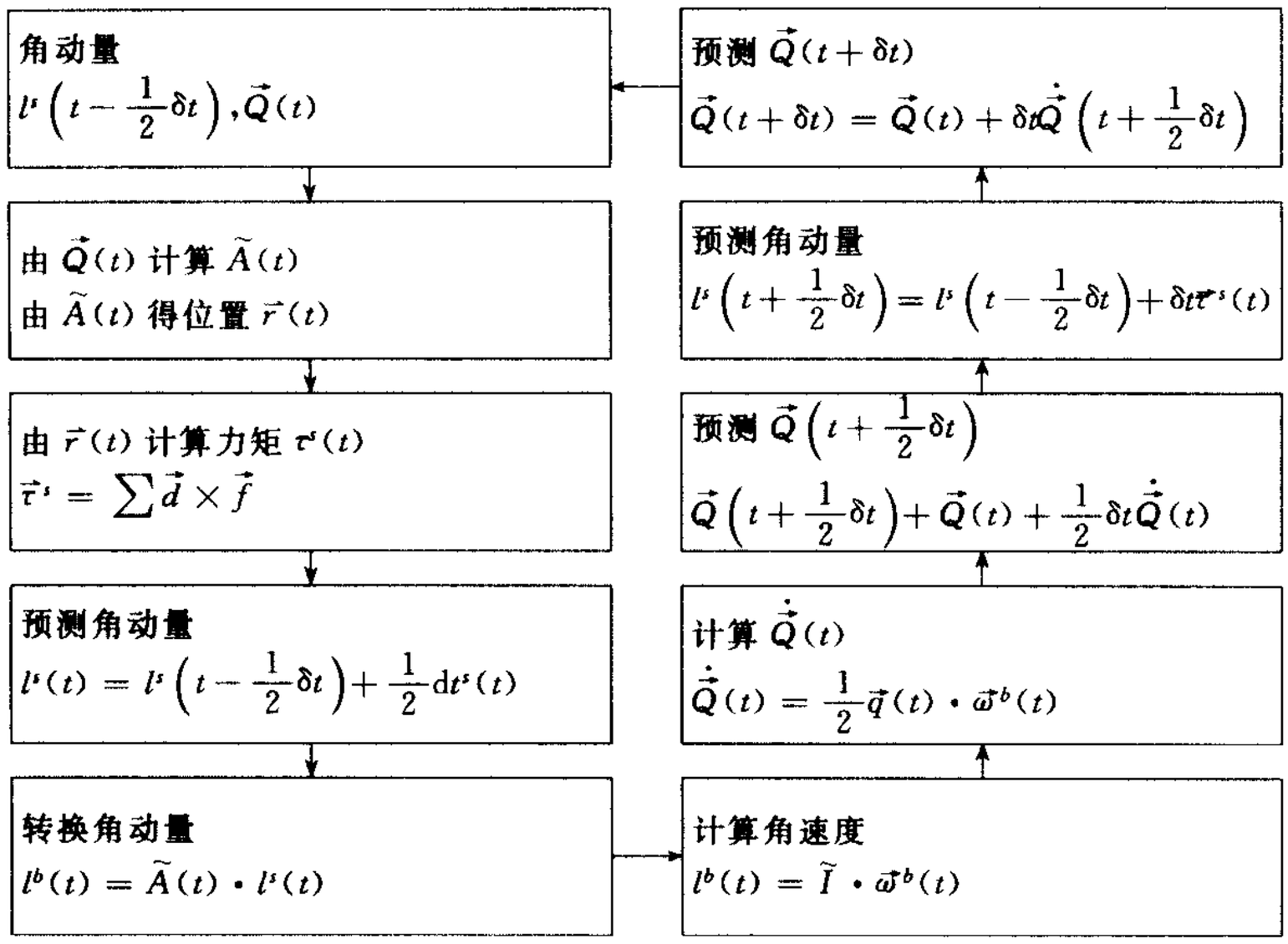



**图 6-2　非线形刚性分子计算的流程**

利用刚性分子计算方法可节省许多计算的时间，例如一般水分子系统的计算。由于水分子的 O—H 键伸缩振动频率很大，约 $3660\mathrm{cm}^{-1}$，因此仅能采用较小的积分步长，$\delta t\sim5\times10^{-16}\mathrm{s}$。但若采用刚性分子的模型，则积分步长可加大至 $6\times10^{-15}\mathrm{s}$，即增加约 10 倍的计算效率。基于这个原因，现今生化体系的分子动力模拟多采用刚性的水分子模型以节省计算的时间。因为生化体系中都含有大量的水分子，并且水分子的运动对生化分子的性能有非常重要的影响。

## 6.3　键长限制法

分子的各种内运动中，以键伸缩与键角弯曲最快。但多数的分子其键

长与键角在其平衡值附近来回呈小幅度振动，键长的差异约为3%，键角的变化约为5°～10°。事实上许多由分子动力计算的特性与分子内键长与键角变化的关系不大。因此，若能在计算中将键长键角维持固定不变，则可选用较长的积分步程，获得更长时间的动力资料。将键长键角维持固定的分子动力计算属于限制计算法（constraint dynamics）的一种。

最通常的限制分子键长的方法为 Ryckaert 等所设计的。这种限制分子键长的方法为在势能中加入一限制项。设欲限制原子 $i$ 与 $j$ 间的键长 $r_{ij}$ 为固定长度 $d_{ij}$，则势能为：

$$U=U^0+\frac{1}{2}\lambda_{ij}(r_{ij}^2-d_{ij}^2) \tag{6-25}$$

式中，$U^0$ 为分子的一般势能项，第二项为键长限制项，$\lambda_{ij}$ 为拉格朗日未定乘子，若 $r_{ij}=d_{ij}$，则此项为零。由牛顿运动定律，各原子所受的力为：

$$\vec{F}_i=m_i\ddot{\vec{r}}_i=-\nabla_i U=-\nabla_i U^0+\lambda_{ij}r_{ij}\nabla_i(r_{ij})=\vec{F}_i^0-\lambda_{ij}r_{ij}\nabla_i(r_{ij})$$

$$\vec{F}_j=m_j\ddot{\vec{r}}_j=\vec{F}_j^0-\lambda_{ij}r_{ij}\nabla_j(r_{ij}) \tag{6-26}$$

式中，$\vec{F}_0$ 为无限制项时原子所受的力。将式(6-26) 积分，依照 Verlet 方法得 $i$ 原子于时间 $t+\delta t$ 时的位置为：

$$\vec{r}_i(t+\delta t)=\vec{r}_i(t)+\vec{r}_i\left(t-\frac{1}{2}\delta t\right)\delta t+\frac{F_i^0}{m_i}\delta t^2-\frac{1}{m_i}\lambda_{ij}r_{ij}(t)\nabla_i[r_{ij}(t)]$$

$$\vec{r}_i(t+\delta t)=\vec{r}_i^0(t+\delta t)+\frac{1}{m_i}\lambda_{ij}r_{ij}(t)\nabla_i[r_{ij}(t)] \tag{6-27}$$

式中，$\vec{r}_i^0(t+\delta t)$ 为无限制项时所得到的位置。同理得：

$$\vec{r}_j(t+\delta t)=\vec{r}_j^0(t+\delta t)-\frac{1}{m_j}\lambda_{ij}r_{ij}(t)\nabla_j[r_{ij}(t)] \tag{6-28}$$

但依键长限制的条件，不论何时，原子 $i$ 与 $j$ 间的键长均需等于 $d_{ij}$，即：

$$r_{ij}(t+\delta t)^2=d_{ij}^2 \tag{6-29}$$

将式(6-28) 的结果代入式(6-29) 得：

$$\vec{r}_{ij}(t+\delta t)\cdot\vec{r}_{ij}(t+\delta t)=$$

$$[\vec{r}_i(t+\delta t)-\vec{r}_j(t+\delta t)]\cdot[\vec{r}_i(t+\delta t)-\vec{r}_j(t+\delta t)]=d_{ij}^2$$

$$\vec{r}_i(t+\delta t)-\vec{r}_j(t+\delta t)=$$

$$\vec{r}_{ij}^0(t+\delta t)-\lambda_{ij}r_{ij}(t)\left[\frac{\nabla_i r_{ij}(t)}{m_i}-\frac{\nabla_j r_{ij}(t)}{m_j}\right]\delta t^2 \tag{6-30}$$

利用梯度的关系式 $\nabla_i r_{ij}(t)=\frac{\vec{r}_{ij}(t)}{r_{ij}(t)}$，$\nabla_j r_{ij}=\frac{\vec{r}_{ij}(t)}{r_{ij}(t)}$ 可得：

$$\vec{r}_{ij}(t+\delta t)=\vec{r}^{\,0}_{ij}(t+\delta t)-\lambda_{ij}\left(\frac{1}{m_i}+\frac{1}{m_j}\right)\vec{r}_{ij}(t)\delta t^2 \tag{6-31}$$

将此代入式(6-30)，整理后得：

$$r^0_{ij}(t+\delta t)^2-2\lambda_{ij}\left(\frac{1}{m_i}+\frac{1}{m_j}\right)\delta t^2\,\vec{r}^{\,0}_{ij}(t+\delta t)\cdot\vec{r}_{ij}(t)+O(\delta t^4)=d^2_{ij} \tag{6-32}$$

式中，$O(\delta t^4)$ 表示 $\delta t$ 的 4 次方项，因 $\delta t$ 很小，所以通常此项可忽略。将式(6-32) 整理得：

$$\lambda_{ij}\left(\frac{1}{m_i}+\frac{1}{m_j}\right)\vec{r}^{\,0}_{ij}(t+\delta t)\cdot\vec{r}_{ij}(t)=\frac{r^0_{ij}(t+\delta t)^2-d^2_{ij}}{2\delta t^2} \tag{6-33}$$

根据此式，可以解出 $\lambda_{ij}$。将此 $\lambda_{ij}$ 代入式(6-31) 可得 $t+\delta t$ 时原子 $i$ 与 $j$ 的正确位置。

## 6.4 限制键长与键角的计算

除了限制键长外，亦可限制分子的键角。以非线形的三原子分子为例(如水分子)；将相连的原子 1,2,3 的键长 $r_{12}$、$r_{23}$ 及键角 $\theta_{123}$ 同时限制住的计算法，可令限制项为：

$$\frac{1}{2}\lambda_{12}(r^2_{12}-d^2_{12})+\frac{1}{2}\lambda_{23}(r^2_{23}-d^2_{23})+\frac{1}{2}\lambda_{13}(r^2_{13}-d^2_{13}) \tag{6-34}$$

图 6-3 显示非线形三原子分子限制力的关系，任一瞬间限制力的合力必为零。

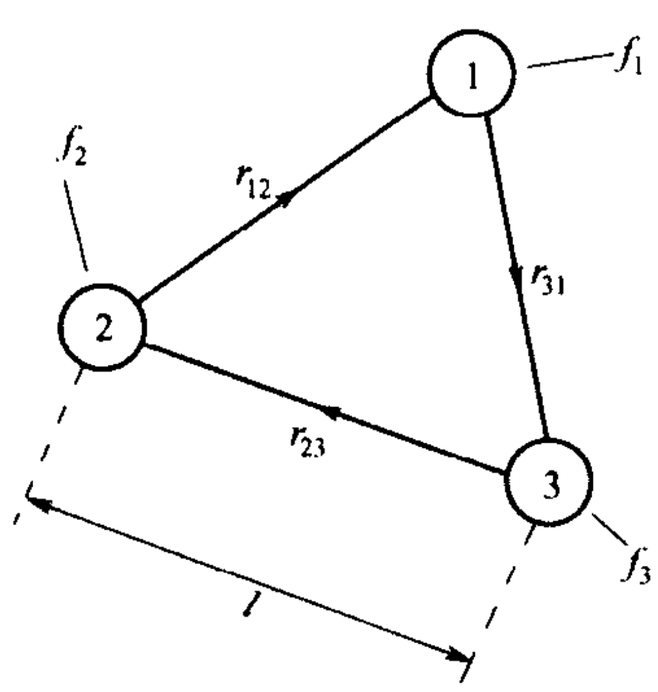



**图 6-3　非线形三原子分子的限制力的关系**

依上节的方法，由式(6-34) 可导出如下的关系式，即：

$$\vec{r}_1(t+\delta t)=\vec{r}^{\,0}_1(t+\delta t)-\lambda_{12}\frac{\vec{r}_{12}(t)}{m_1}\delta t^2-\lambda_{13}\frac{\vec{r}_{13}(t)}{m_1}\delta t^2$$

$$\vec{r}_2(t+\delta t)=\vec{r}^{\,0}_2(t+\delta t)+\lambda_{12}\frac{\vec{r}_{12}(t)}{m_2}\delta t^2-\lambda_{23}\frac{\vec{r}_{23}(t)}{m_2}\delta t^2 \tag{6-35}$$

$$\vec{r}_3(t+\delta t)=\vec{r}^{\,0}_3(t+\delta t)+\lambda_{13}\frac{\vec{r}_{13}(t)}{m_3}\delta t^2+\lambda_{23}\frac{\vec{r}_{23}(t)}{m_3}\delta t^2$$

因为 $r_{12}(t+\delta t)^2=d^2_{12}$，$r_{23}(t+\delta t)^2=d^2_{23}$，$r_{13}(t+\delta t)^2=d^2_{13}$，将式

(6-35) 的三项代入，并仿前将 $O(\delta t^4)$ 项略去，得联立方程组：

$$2\lambda_{12}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\delta t^2\vec{r}^{\,0}_{12}(t+\delta t)\cdot\vec{r}_{12}(t)+2\lambda_{23}\frac{\delta t^2}{m_2}\vec{r}^{\,0}_{12}(t+\delta t)\cdot$$

$$\vec{r}_{23}(t)+2\lambda_{13}\frac{\delta t^2}{m_1}\vec{r}^{\,0}_{12}(t+\delta t)\cdot\vec{r}_{13}(t)=r^0_{12}(t+\delta t)^2-d^2_{12}\cdot$$

$$2\lambda_{12}\frac{\delta t^2}{m_2}\vec{r}^{\,0}_{23}(t+\delta t)\cdot\vec{r}_{12}(t)+2\lambda_{23}\left(\frac{1}{m_2}+\frac{1}{m_3}\right)\delta t^2\vec{r}_{23}\cdot$$

$$(t+\delta t)\cdot\vec{r}_{23}(t)+2\lambda_{13}\frac{\delta t^2}{m_2}\vec{r}_{23}(t+\delta t)\cdot\vec{r}_{13}(t)=r^0_{23}(t+\delta t)^2-d^2_{23}\cdot$$

$$2\lambda_{12}\frac{\delta t^2}{m_3}\vec{r}^{\,0}_{13}(t+\delta t)\cdot\vec{r}_{13}(t)+2\lambda_{23}\frac{\delta t^2}{m_3}\vec{r}^{\,0}_{13}(t+\delta t)\cdot\vec{r}_{23}(t)+2\lambda_{13}\cdot$$

$$\left(\frac{1}{m_1}+\frac{1}{m_3}\right)\delta t^2\vec{r}^{\,0}_{13}(t+\delta t)\cdot\vec{r}_{13}(t)=r^0_{13}(t+\delta t)^2-d^2_{13}$$

(6-36)

以上的联立方程组可以写成矩阵的型式，即：

$$\begin{bmatrix}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\vec{r}^{\,0}_{12}(t+\delta t)\cdot\vec{r}_{12}(t) & \frac{1}{m_1}\vec{r}^{\,0}_{12}(t+\delta t)\cdot\vec{r}_{23}(t) & \frac{1}{m_1}\vec{r}^{\,0}_{12}(t+\delta t)\cdot\vec{r}_{13}(t)\\ \frac{1}{m_2}\vec{r}^{\,0}_{23}(t+\delta t)\cdot\vec{r}_{12}(t) & \left(\frac{1}{m_2}+\frac{1}{m_3}\right)\vec{r}^{\,0}_{23}(t+\delta t)\cdot\vec{r}_{23}(t) & \frac{1}{m_2}\vec{r}^{\,0}_{23}(t+\delta t)\cdot\vec{r}_{13}(t)\\ \frac{1}{m_3}\vec{r}^{\,0}_{13}(t+\delta t)\cdot\vec{r}_{12}(t) & \frac{1}{m_3}\vec{r}^{\,0}_{13}(t+\delta t)\cdot\vec{r}_{23}(t) & \left(\frac{1}{m_1}+\frac{1}{m_3}\right)\vec{r}^{\,0}_{13}(t+\delta t)\cdot\vec{r}_{13}(t)\end{bmatrix}$$

$$\begin{pmatrix}\lambda_{12}\\ \lambda_{23}\\ \lambda_{13}\end{pmatrix}=\begin{bmatrix}\frac{r^0_{12}(t+\delta t)^2-d^{\,2}_{12}}{2\delta t^2}\\ \frac{r^0_{23}(t+\delta t)^2-d^{\,2}_{23}}{2\delta t^2}\\ \frac{r^0_{13}(t+\delta t)^2-d^{\,2}_{13}}{2\delta t^2}\end{bmatrix} \tag{6-37}$$

将以上的矩阵方程式简写成：

$$\widetilde{R}\vec{\lambda}=\vec{d} \tag{6-38}$$

解出以上的矩阵方程式可以求得 $\lambda_{12}$、$\lambda_{23}$、$\lambda_{13}$ 的值。一般求解的方法为利用反矩阵，即：

$$\widetilde{R}^{-1}\widetilde{R}\vec{\lambda}=\widetilde{R}^{-1}\vec{d},\ \vec{\lambda}=\widetilde{R}^{-1}\vec{d} \tag{6-39}$$

式中，$\widetilde{R}^{-1}$ 为矩阵 $\widetilde{R}$ 的反矩阵。求得各 $\lambda$ 值后，由式(6-35) 可求出各原子于 $t+\delta t$ 时的正确位置。

除了上述的方法外，还有一种较为简便解方程式(6-38) 的方法，为利用迭代计算法（iterative method)。常用的迭代计算法为 SHAKE 方法，此方法将矩阵方程式(6-37) 中第一个矩阵的所有非对角线元素忽略，由

此很容易解出 $\lambda_{12}$、$\lambda_{23}$、$\lambda_{13}$，并将此代入式(6-35)，得 $\vec{r}_1(t+\delta t)$、$\vec{r}_2(t+\delta t)$ 与 $\vec{r}_3(t+\delta t)$。然后，检查由此新得到的原子位置所计算的键长是否满足 $r_{12}^2=d_{12}^2$、$r_{23}^2=d_{23}^2$、$r_{13}^2=d_{13}^2$的关系。若所计算的键长关系不合，则将 $\vec{r}_1(t+\delta t)$、$\vec{r}_2(t+\delta t)$、$\vec{r}_3(t+\delta t)$视为新的 $\vec{r}_1^0(t+\delta t)$、$\vec{r}_2^0(t+\delta t)$、$\vec{r}_3^0(t+\delta t)$，再代入式(6-37) 的矩阵方程式中计算。如此，不断重复，直至所得到的键长满足要求的长度。实际的分子动力计算通常需要 7～10 次左右的迭代计算才能找出原子适当的位置。迭代计算方法的优点在于不用求式(6-39) 中的反矩阵，因而适用于较多限制的情形。最先发展的 SHAKE 方法为利用 Verlet 的计算法，以后逐渐推展至 Verlet 的跳蛙计算法等其他的方法。除了 SHAKE 方法外，Anderson 所发展的 RATTLE 方法，利用 Verlet 的速度算法，亦为常用的限制计算法。

## 6.5 聚合物的限制动力计算

利用限制动力计算，可节省许多的计算时间。Toxvard 以类正癸烷液体系统为例，检验限制动力计算的正确性。此计算将正癸烷分子的甲基及亚甲基均视为粒子，这种计算方法称为联合原子模型（united atom model）如图 6-4 所示。每根正癸烷链含有 10 个这样的粒子，以符号 $\overline{C}$ 表示，称为类正癸烷。

Toxvard 的分子动力模拟计算采用 Ryckaert 等所发展烷类聚合物的简化力场，即：

$$U=U_{VDW}+U_s(r)+U_b(\theta)+U_t(\phi) \tag{6-40}$$

式中，$U_{VDW}$为范德瓦耳斯作用项；$U_s$ 为 $\overline{C}$-$\overline{C}$ 键伸缩项；$U_b$ 为 $\overline{C}$-$\overline{C}$-$\overline{C}$ 键角弯曲项；$U_t$ 为 $\overline{C}$-$\overline{C}$-$\overline{C}$-$\overline{C}$ 二面角 $\phi$ 的扭转项。

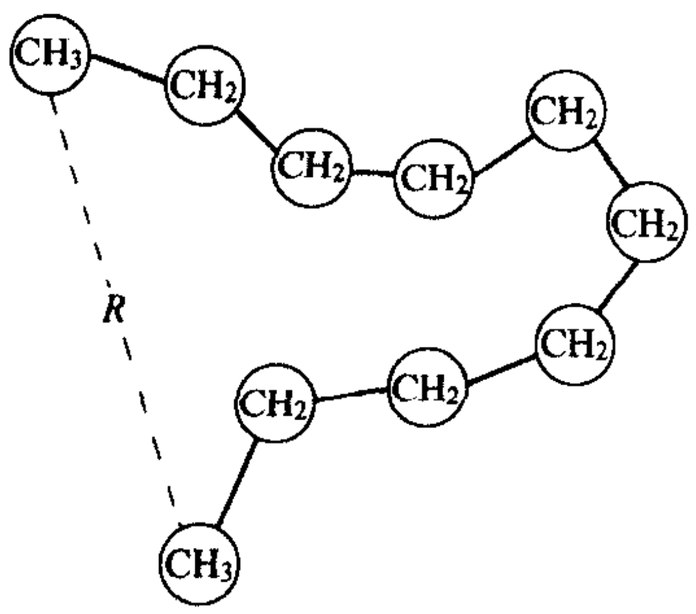



图 6-4 类正癸烷分子的示意

分子动力计算分为全无限制系统与限制系统，无限制系统考虑上述势能的每一项，而限制系统分为两类；第一类系统限制所有的键长，第二类系统除键长外亦限制所有的键角。计算系统的密度定为 0.63g/mL，温度

为 481K，全无限制系统的计算时间为 20ns，加入限制系统的计算时间为 6ns。分子动力计算引用 Verlet 的跳蛙方法，积分步程均为 $\delta t = 2 \times 10^{-3}$ps。聚合物的特性可由一些物理性质显示，一般较重要的为：①末端距 (end-to-end distance) 平方的平均值 $\langle R^2 \rangle$，$R$ 为链分子头端基团与末端基团的距离，如图 6-4 所示；②回转半径 (gyration radius) $\langle S^2 \rangle$，回转半径的计算式为：

$$S^2 = \frac{\sum_{i=1}^{n} m_i d_i^2}{\sum_{i=1}^{N} m_i} \tag{6-41}$$

式中，$m_i$ 为每个粒子的质量 ($CH_3 = 15$amu，$CH_2 = 14$amu)；$d_i$ 为各粒子与正癸烷分子质量中心的距离。除了上述两个物理量外，还需计算聚合物的转动惯量 $I$，此物理量的形式为：

$$\tilde{I}_{xyz} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \tag{6-42}$$

式中，转动惯量矩阵的各元素为：

$$I_{xx} = \sum_{i=1}^{n} m_i (y_i^2 + z_i^2), I_{yy} = \sum_{i=1}^{n} m_i (x_i^2 + z_i^2), I_{zz} = \sum_{i=1}^{n} m_i (x_i^2 + y_i^2)$$

$$I_{xy} = I_{yx} = \sum_{i=1}^{n} m_i x_i y_i, I_{yz} = I_{zy} = \sum_{i=1}^{n} m_i y_i z_i, I_{xz} = I_{zx} = \sum_{i=1}^{n} m_i x_i z_i \tag{6-43}$$

式中，$m_i$ 为链中第 $i$ 个粒子的质量；$x_i$，$y_i$，$z_i$ 为粒子的笛卡儿坐标。通常可选取另外的坐标系统 $(a,b,c)$，使转动惯量矩阵呈对角线化的形式，即：

$$I_{abc} = \begin{pmatrix} I_{aa} & 0 & 0 \\ 0 & I_{bb} & 0 \\ 0 & 0 & I_{cc} \end{pmatrix} \tag{6-44}$$

则此矩阵对角线上的三个分量相当于绕着 $a,b,c$ 轴的转动惯量。上述过程相当于将矩阵 $I_{xyz}$ 对角线化，即：

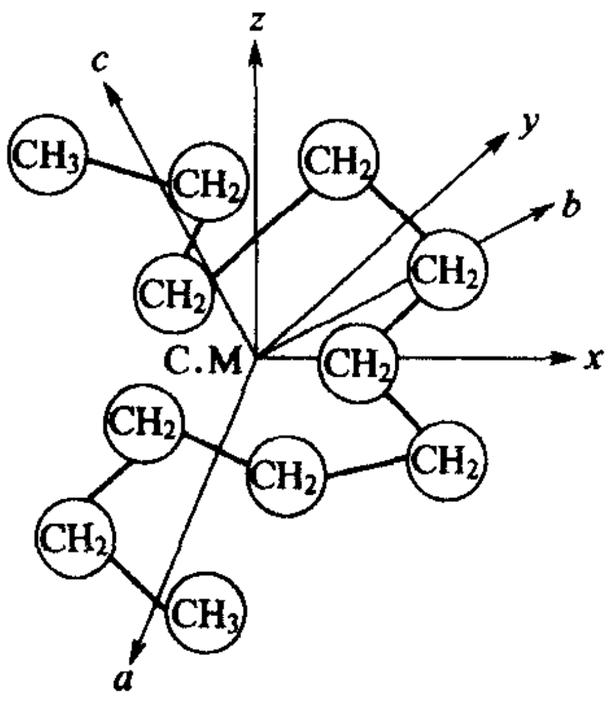



**图 6-5　类正癸烷分子的坐标系统**

$$I_{abc} = \widetilde{C}^T \widetilde{I}_{xyz} \widetilde{C} \tag{6-45}$$

式中的 $C$ 矩阵为特征向量矩阵。矩阵中的每一行即为对应于 $\vec{a}$，$\vec{b}$，$\vec{c}$ 轴的单位向量。图 6-5 例示类正癸烷分子的坐标系统。

由分子动力模拟所得的末端距平方及回转半径的平均值 $\langle R^2 \rangle$、$\langle S^2 \rangle$ 与转动惯量 $I_{aa}$，$I_{bb}$，$I_{cc}$ 可以反映聚合物的形状。“〈〉”符号表示平均值，设系统含有 $N$ 根聚合物链，则任何物理量 $A$ 的平均值为：

$$\langle A \rangle = \frac{\sum_{t=1}^{T} \sum_{j=1}^{N} A_{j,t}}{T \times N} \tag{6-46}$$

式中，$T$ 为计算的总时间（步数）；$A_{j,t}$ 为第 $j$ 根链在时间 $t$ 时的 $A$ 值。图 6-6 为非限制与限制条件下计算所得的扭转角分布的比较。

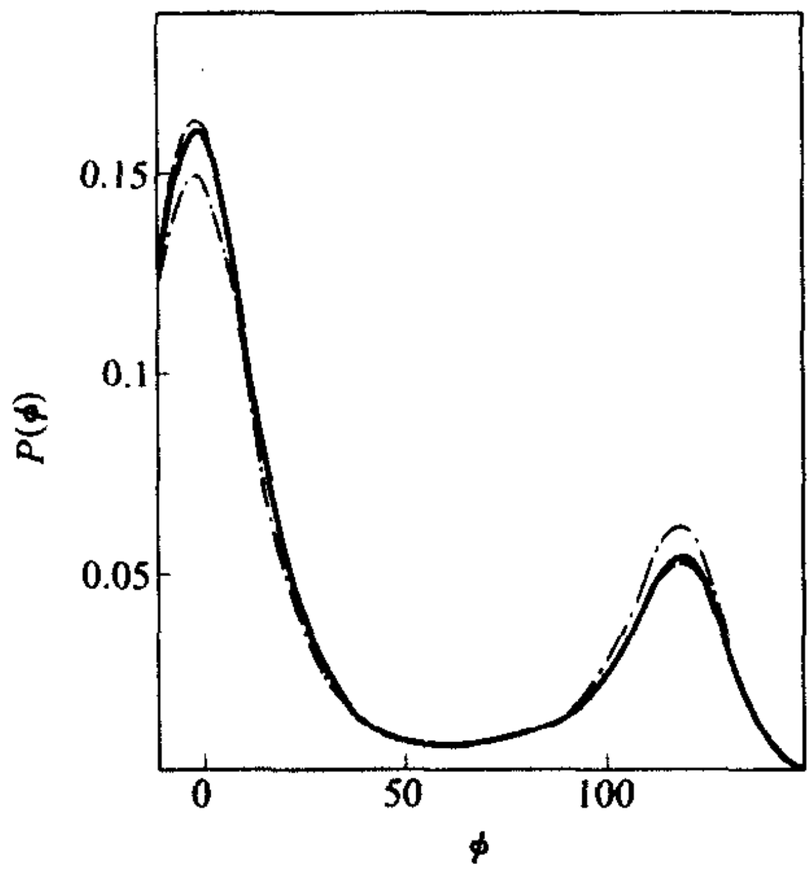



**图 6-6　各种限制条件下计算所得的扭转角分布的比较**

**实线为未限制系统，虚线为限制 C-C 键长，点虚线为限制键长与键角系统**

图中显示，若仅限制键长，则计算的扭转角分布与未限制系统的结果十分相近。若加上键角的限制，则误差较大，但仍属于相似的分布。

图 6-7 比较各种限制条件下的末端距。亦显示增加键角的限制会导致较大的偏差。聚合物的每组 $\overline{C}$-$\overline{C}$-$\overline{C}$-$\overline{C}$ 扭转角均有邻位交叉式（gauche）与反式（trans）的构象，图 6-8 所示。

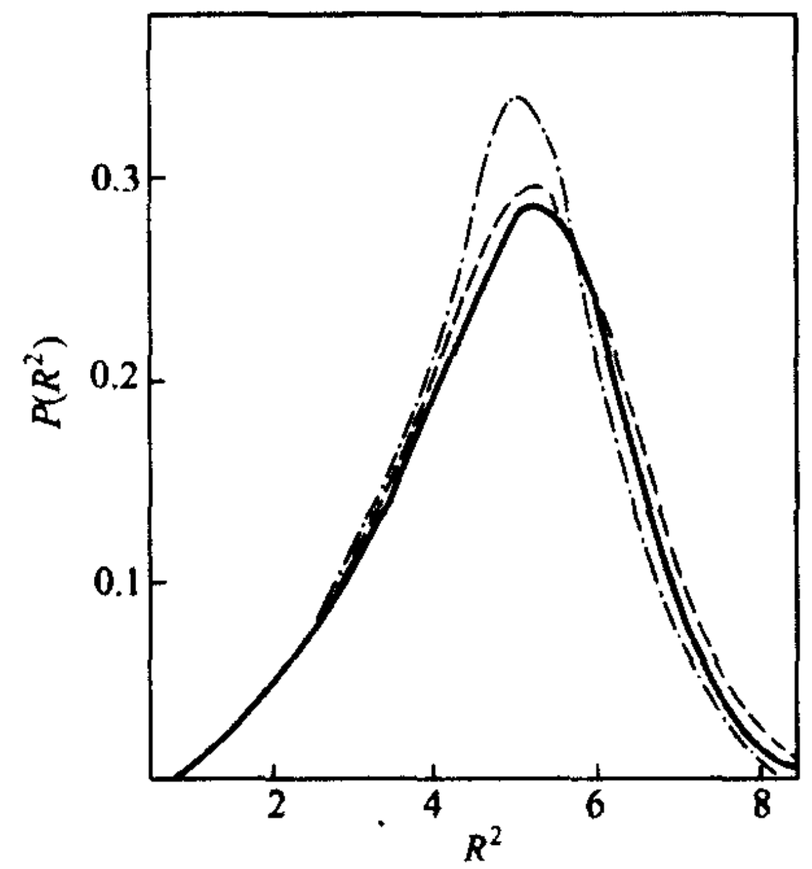
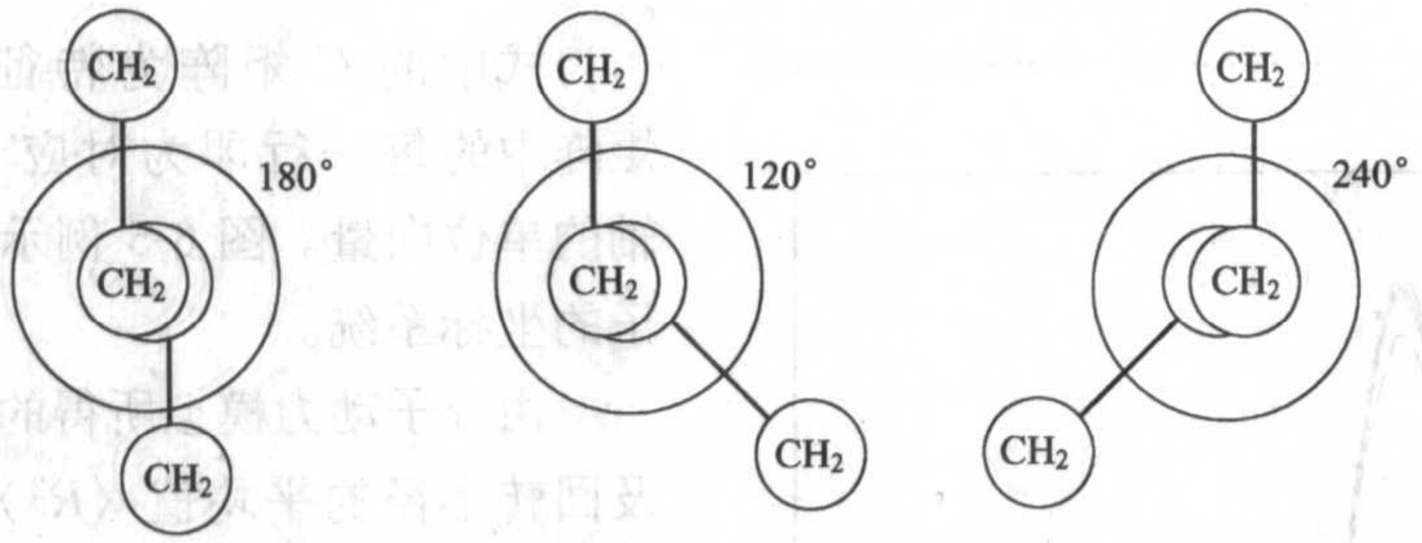



**图 6-7　各种限制条件下计算所得的末端距的比较**

**实线为未限制系统，虚线为限制 C-C 键长，点虚线为限制键长与键角系统**

邻位交叉式共有两种，分别标示为 $g^+$ 与 $g^-$。反式为最稳定的构象，邻位交叉式则为次稳定的构象。令 $n_g$ 与 $n_t$ 分别表示邻位交叉式与反式构象的含量，其平均值的比为 $\langle n_g \rangle / \langle n_t \rangle$。




反式 (t)　　邻位交叉式 ($g^+$)　　邻位交叉式 ($g^-$)

**图 6-8　反式与邻位交叉式的示意**

此比值的大小反映聚合物的稳定性，比值愈小，反式构象愈多，聚合物愈稳定。表 6-1 比较所计算各系统的这些特性及扭转势能的平均值 $\langle U_t(\phi)\rangle$，及范德瓦耳斯势能的平均值 $\langle U_{VDW}\rangle$。同样的，所有的计算量均显示增加键角的限制会导致较大的偏差。

表 6-2 比较各系统中聚合物正癸烷的形状。表中，$I_{aa}$、$I_{bb}$、$I_{cc}$为式 (6-45) 的方法所计算出聚合物转动惯量的本征值。

**表 6-1　各种限制系条件的系统在 481K 的分子动力计算结果**

| 计算类型 | $\langle R^2\rangle$ | $\langle S^2\rangle$ | $\langle n_g\rangle/\langle n_t\rangle$ | $\langle U_t(\phi)\rangle$ | $\langle U_{VDW}\rangle$ |
|---|---|---|---|---|---|
| 无限制 | 5.06±0.03 | 0.632±0.002 | 0.602±0.008 | 44.21±0.13 | −7.55±0.001 |
| 限制键长 | 5.00±0.04 | 0.622±0.003 | 0.588±0.013 | 43.61±0.29 | −7.73±0.001 |
| 限制键长键角 | 4.86±0.04 | 0.609±0.002 | 0.72±0.02 | 45.09±0.30 | −8.41±0.06 |

注：$\langle R^2\rangle$ 与 $\langle S^2\rangle$ 的单位为 $\sigma^2=(3.8A)^2$，$\langle U_t(\phi)\rangle$ 与 $\langle U_{VDW}\rangle$ 的单位为 $\varepsilon=72K$。

**表 6-2　各种限制系条件的系统聚合物形状的比较**

| 项　　目 | $\langle I_{aa}^2\rangle/\langle S^2\rangle$ | $\langle I_{bb}^2\rangle/\langle S^2\rangle$ | $\langle I_{cc}^2\rangle/\langle S^2\rangle$ |
|---|---|---|---|
| 无限制 | 0.889±0.001 | 0.0897±0.0011 | 0.0210±0.001 |
| 限制键长 | 0.891±0.002 | 0.0844±0.0015 | 0.0209±0.0004 |
| 限制键长键角 | 0.886±0.002 | 0.0914±0.0019 | 0.0230±0.0003 |

由表 6-2 中的数据可以看出，$\langle I_{bb}^2\rangle/\langle S^2\rangle$与$\langle I_{cc}^2\rangle/\langle S^2\rangle$均远小于 $\langle I_{aa}^2\rangle/\langle S^2\rangle$，这表示聚合物的平均形状约略为扁盘形（orblate），见图 6-9。

大体说来，由以上的分子动力计算可知，限制系统与无限制系统的趋

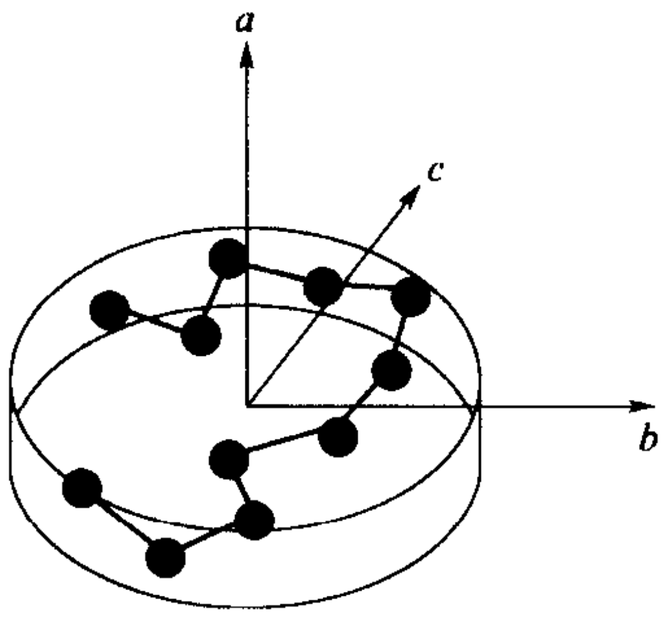



**图 6-9　扁盘形聚合物示意**

势一致。仅限制键长系统的计算结果无论是分子形状、角度分布、物理性质等都非常接近无限制系统的计算结果。但加入键长的限制可容许在分子动力计算中引用较长的积分步程，可节省实际计算时间与累积更长时间的计算轨迹。因此，若研究较长时间的动力行为，则限制计算为重要的简化方法。尤其是只限制键长的计算法，更是值得引用。对于较复杂的分子，如含有环状基团的糖类分子、聚苯乙烯等，同时限制键长与键角并非易事，仍有待发展适当的数学方法。

## 参　考　文　献

1　Fincham D. CCP5 Quarterly, 12, 47, 1984

2　Evans D J, Murad S. Mol. Phys., 34, 327, 1977

3　Potter D. "Computational Physics" Wiley, New York. USA, 1972

4　Ryckaert J P, Ciccotti G, Berendsen H J C. J. Comput. Phys., 23, 327, 1977

5　Anderson H C. J. Comput. Phys., 52, 1, 1983

6　Toxvard S. J. Chem. Phys., 87, 6140, 1987

7　Ryckaert J P, Ciccotti G, Berendsen H J. J. Compuy. Phys., 23, 327, 1977

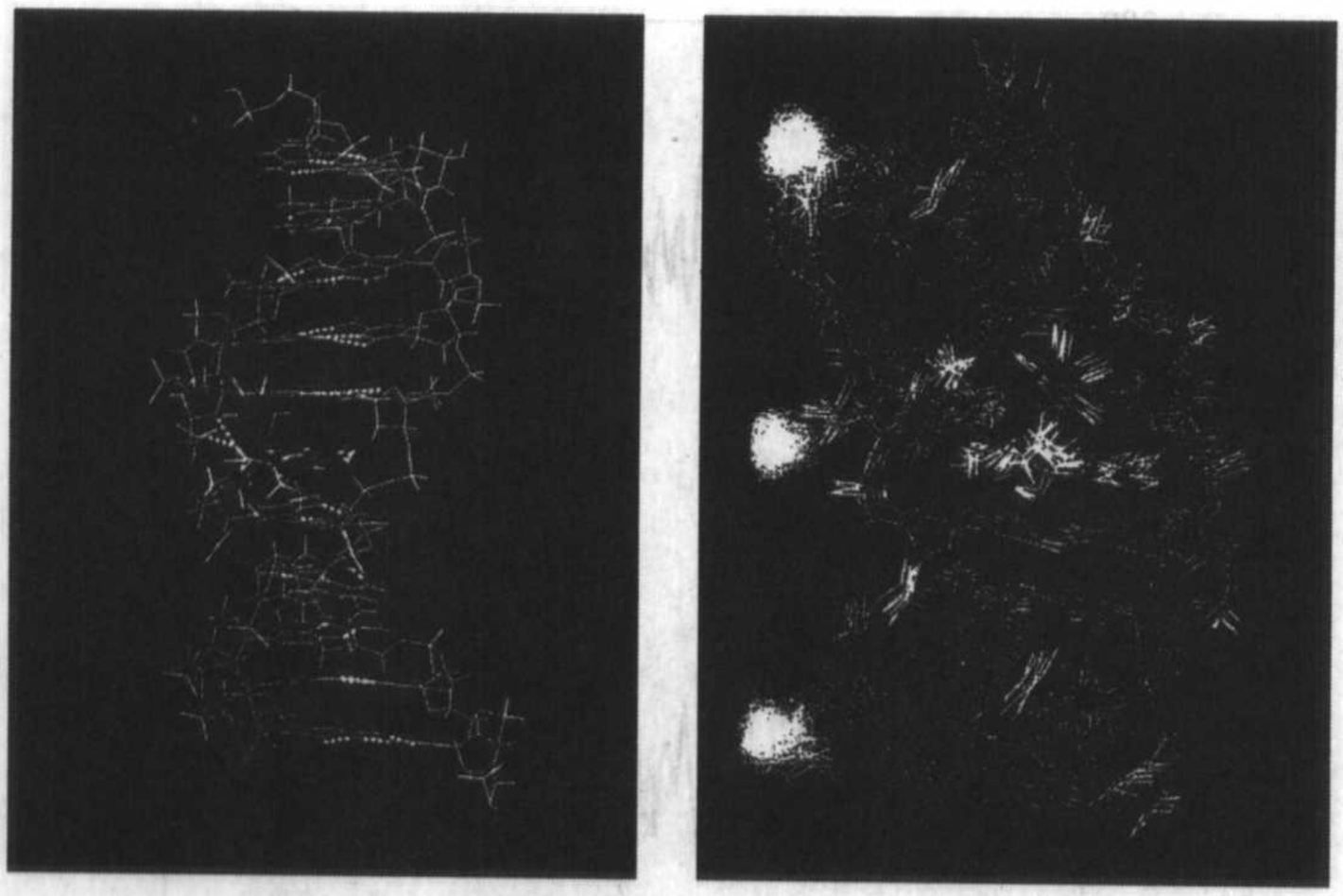
# 第7章 分子动力计算的应用

第5、6两章介绍分子动力计算的基本方法及特殊形态的分子动力计算。本章介绍了从分子动力模拟获得系统中粒子位置与速度的轨迹后，通过分析这些轨迹获得各种的性质和参数的方法，可以帮助解决生化体系等各种实际问题。本章涉及的体系性质和参数包括：各种热力学的特性与光谱性质、系统自由能变化。

## 7.1 运动轨迹分析

分子动力计算可得到系统内所有原子（或基团）在每一步的坐标及速度。这些随着时间变异的原子坐标反映出系统中原子运动的途径。这些原子运动的途径，称为运动轨迹（trajectory）。分子动力计算中原子的运动速度反映出原子运动的快慢与方向，通常计算中每10步或20步存取系统中所有原子的坐标及速度以作为分析之用，称为存取轨迹。利用所存取的轨迹可以计算相关的物理量而得到许多有用的热力学、统计学及动态的信息。最直接显示计算结果的方法为将所有计算的轨迹以图形的方式显现。图7-1为利用分子动力计算核糖体的结果图7-1(a)为起始的结构，图7-1(b)为将模拟时间在35～75ps（pico second，ps）（$1ps=1.0\times10^{-12}s$）的区间内，每5ps模拟所得的原子位置重合的结果。比较此二图可发现分子真正的运动途径。

另一种常用的轨迹表示法为将特定的原子或特定的内坐标的运动轨迹与时间的关系以图形表示，用以研究特殊形态的运动。以聚合物聚醚醚酮（polyether ether ketone，PEEK）的计算为例，聚醚醚酮为重要的安全玻璃的材料，具有高熔点与高玻璃转化点，是韧性强的透明材质。图7-2为



(a)　　(b)

**图 7-1　分子动力计算所模拟的核糖体的原子轨迹**

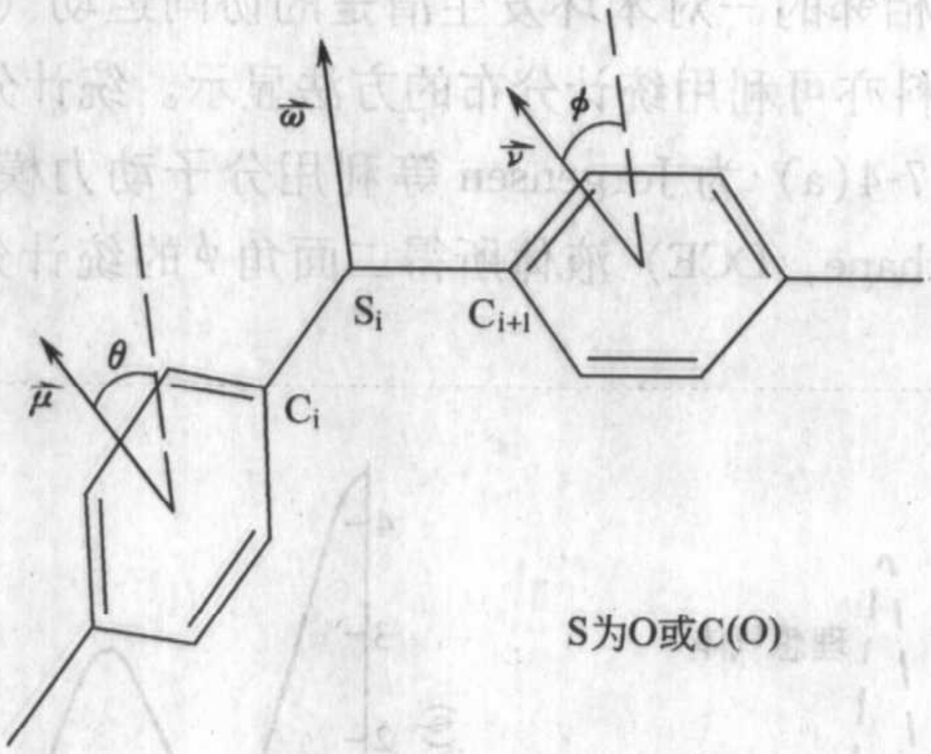



**图 7-2　聚醚醚酮聚合物坐标的定义**

PEEK 聚合物与苯环有关的二面角的定义图。图中，$\vec{\mu}$ 与 $\vec{\nu}$ 为苯环的法向量，$\vec{\omega}$ 位于 $C_i$、$S_i$ 与 $C_{i+1}$三原子的平面上。

Chen 等利用分子动力模拟计算非晶相 PEEK 聚合物的运动。图 7-3 为计算所得的一对特定的扭转角 $(\theta,\phi)$ 的运动轨迹。

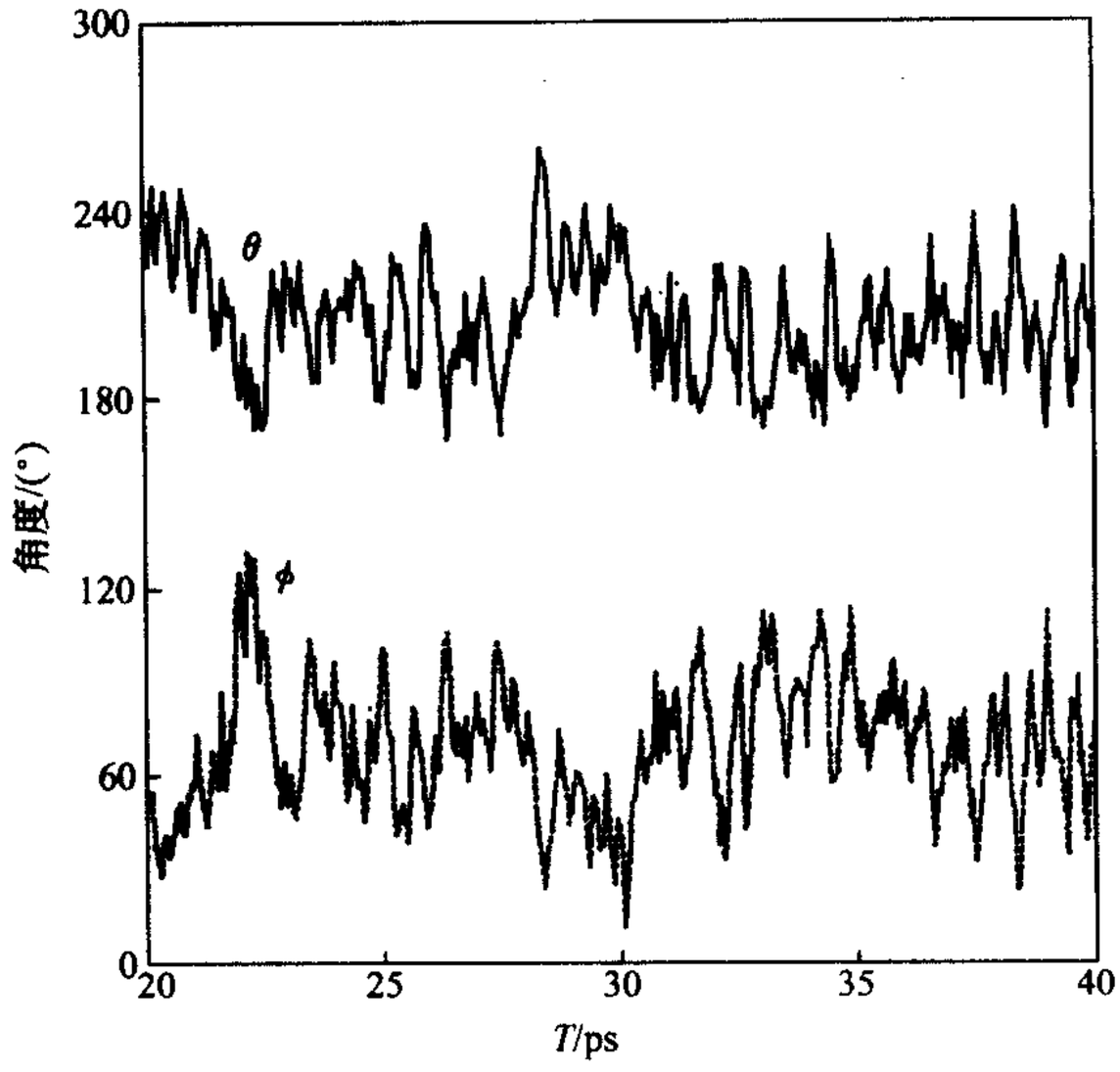
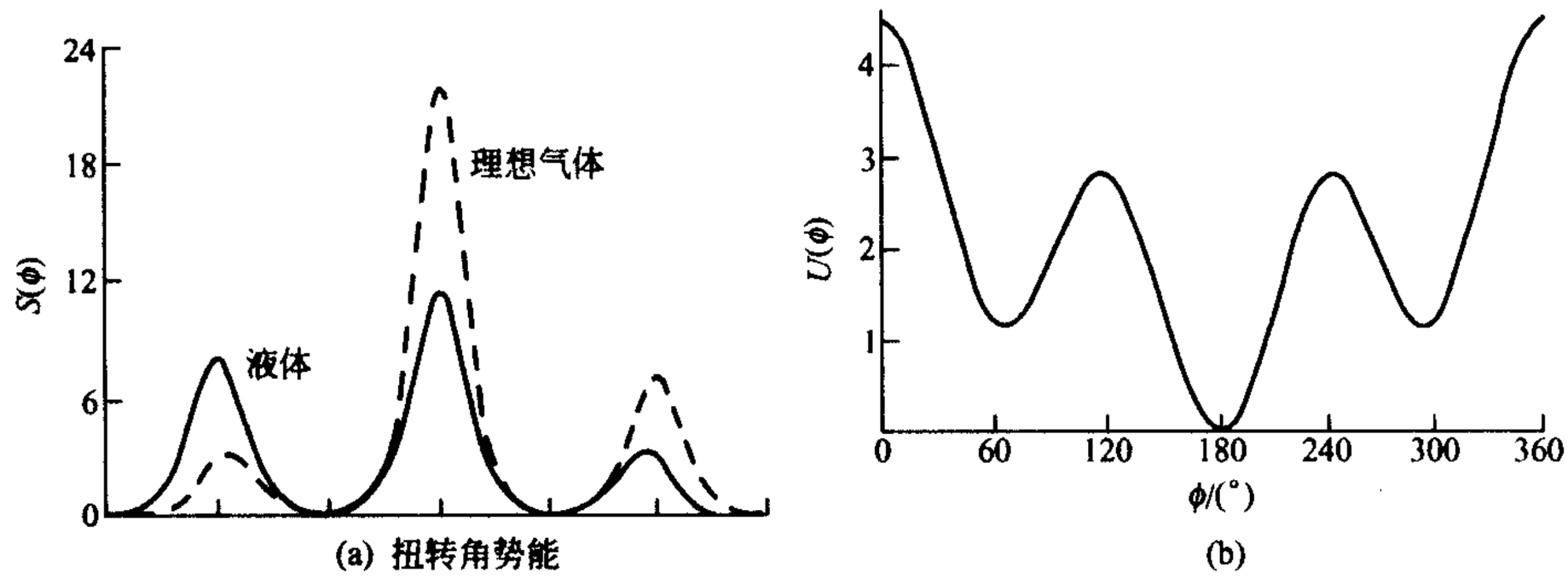



**图 7-3　聚醚醚酮的扭转角运动轨迹**

图中，可发现 $\theta$ 增加时 $\phi$ 减少；反之 $\theta$ 减少时 $\phi$ 增加。此二扭转角同时变化，亦即相邻的一对苯环发生清楚的协同运动（cooperative motion）。

轨迹的资料亦可利用统计分布的方法显示。统计分布的大小反映出现的概率大小。图 7-4(a) 为 Jorgensen 等利用分子动力模拟计算 1,2-二氯乙烷 (1,2-dichloroethane，DCE) 液体所得二面角 $\phi$ 的统计分布图，图 7-4(b) 为




(a) 扭转角势能　　(b)

**图 7-4　DCE 分子的扭转角统计分布 S(φ) 图与丁烷分子的势能 U(φ) 图**

$S(\phi)$ 单位：分布次数；$U(\phi)$ 单位：相对能量（kJ/mol）

所对应的丁烷分子扭转角势能 [$U(\phi)$] 图。

由此统计分布图可知，DCE 分子的扭转角主要分布于 60°、180°、300°等三个角度附近，相当于邻位交叉构象（60，300）与反式构象(180)。与 DCE 分子对应的丁烷分子，其扭转角势能于此三个角度附近最小，表示这些构象较为稳定，出现的概率最大。但气态的 DCE 分子其反式构象的概率为任一邻位交叉构象的 5 倍，而液态的 DCE 分子其反式构象的概率仅为任一邻位交叉构象的 1.5 倍。这表示液体的 DCE 分子间具有很强的作用，导致构型分布的改变。

除了以统计的方式将计算的结果显现外，亦可将轨迹以动态的图形方式显现（如 Cerius2、Materials Studio 等软件均有此功能），则更有助于了解系统的运动过程。分子动力计算可用以研究各种不同的系统（液体、固体、聚合物、生化分子等），各系统均有其独特的物理特性，计算的物理量也因系统不同而异。第 8 章中，将举出一些针对特殊系统的分子动力计算实例。以下介绍一些通常的物理量的计算方法及应用。

## 7.2 热力学特性的计算

计算分子动力模拟系统动能的平均值可以得到系统的温度，即：

$$T=\frac{\langle 2K\rangle}{f} \tag{7-1}$$

式中，$K$ 为动能；$T$ 为温度；$f$ 为系统的自由度。对含有 $N$ 个原子的系统，每个原子具有 3 个移动自由度，但因为系统的动量守恒，故须减去 3 个自由度，故系统的总自由度为 $f=3N-3$。若系统加入限制条件，则每个限制减少一个自由度。如 $N$ 个双原子分子的系统，若限制各分子的键长，则自由度为 $f=N\times2\times3-N-3$。

若计算的 $(N,V,E)$ 系统含有 $N$ 个原子，且原子间的作用为成对加成（pair-wise additive），即系统的势能为：

$$U=\sum_{i=1}^{N}\sum_{j>i}^{N}U_{ij} \tag{7-2}$$

则系统的压力为：

$$P=\rho k_B T+\frac{1}{3V}\left\langle \sum_{i<j}\vec{F}_{ij}(\vec{r}_{ij})\cdot\vec{r}_{ij}\right\rangle \tag{7-3}$$

式中，$\rho$ 为系统的密度；$\vec{F}_{ij}$ 为原子 $i$ 与 $j$ 间的作用力；$\vec{r}_{ij}$ 为其距离向量。

**（1）动能的统计涨落** 动能的统计涨落（fluctuation）定义为

$$(\delta K)^2 = \langle K^2 \rangle - \langle K \rangle^2 \tag{7-4}$$

由此物理量，可计算系统的定容比热容（heat capacity at constant volume），$C_V$，即：

$$\langle K^2 \rangle - \langle K \rangle^2 = \frac{3k_B^2 T^2}{2N}\left(1 - \frac{3k_B}{2C_V}\right) \tag{7-5}$$

式中，$N$ 为系统的原子总数；$K$ 为系统的动能。关于统计涨落的应用实例将在后面举出。图 7-5 为 Frenkel 等针对仅有 Lennard-Jones 作用的单原子分子系统的能量计算结果。系统的初始温度为 $T^* = 0.728$（简化单位），密度为 $\rho^* = 0.8442$（简化单位），计算所收集的轨迹为 2500 步。

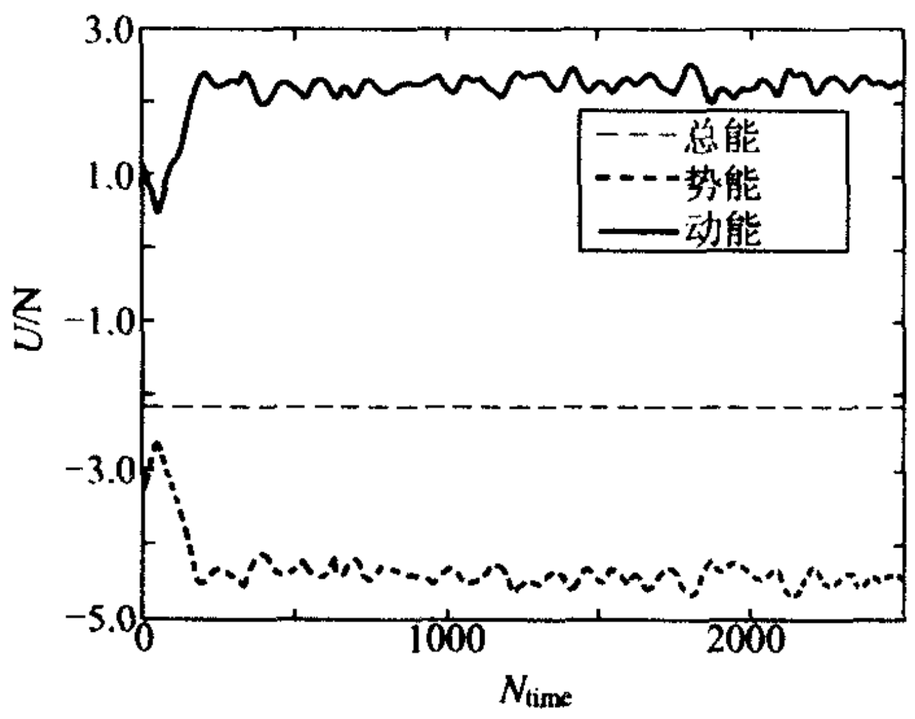



**图 7-5 分子动力学计算 LJ 粒子的各种能量（简化单位）**

图中显示，欲取得动能及势能的正确平均值，约需 1000 步的分子动力计算以平衡系统。由图可得系统中原子的平均动能与势能（以简化单位表示）分别为 $K = 2.2564 \pm 0.0012$，$U = -4.4190 \pm 0.0012$。由动能的平均值计算的温度（以简化单位表示）为 $T = 1.5043 \pm 0.0008$，由动能的统计涨落可以计算系统的定容比热容。

**（2）压力与应力** 压力的定义为单位面积所受的力。标准大气压为 1.013bar（$1\text{bar} = 10^5\text{Pa}$）。事实上，系统在各方向所受的力并非一定相同，因此，压力的正确定义应为张量，即：

$$\tilde{P}=\begin{pmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{pmatrix} \tag{7-6}$$

对于均匀相系统，如气体、液体，压力在各方向均相等，故：

$$\tilde{P}=P\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7-7}$$

式中，$P$ 为纯量，相当于流体的静压（hydrostatic pressure）。但在一些系统中，尤其是固态的材料，压力张量的非对角线元素不全为零。此时压力张量称为应力张量（stress tensor），其对角线元素称为拉伸应力（tensile stress），而非对角线元素则称为剪切应力（shear stress）。

常用的压力与应力的单位为 GPa($1\text{bar}=10^5\,\text{N/m}^2$，$1\text{GPa}=10^9\,\text{N/m}^2$，$1\text{GPa}=10^4\,\text{bar}$)。定义正的压力值为压缩系统向内的力所致，正的应力值为拉伸系统向外之力所致，两者采取的标准相反。系统的压力可由体积、温度与位力（virial）计算，即：

$$PV=Nk_BT+\frac{2}{3}\langle W\rangle \tag{7-8}$$

位力的定义为：

$$W=\frac{1}{2}\sum_{j=1}^{N}\vec{r}_j\cdot\vec{F}_j \tag{7-9}$$

分子动力计算中，压力为瞬间压力 $P$ 的平均值，即：

$$P=\frac{Nk_BT}{V}+\frac{2}{3}\frac{W}{V} \tag{7-10}$$

式中，$T$ 与 $W$ 为系统的瞬间温度与位力。瞬间温度可由瞬间动能 $K$ 计算，即：

$$T=\frac{2}{3Nk_B}K \tag{7-11}$$

将式(7-11) 代入式(7-10)，得：

$$P=\frac{2}{3V}(K+W) \tag{7-12}$$

因压力为一张量，故式(7-12) 可改写成：

$$P=\frac{1}{V}\left(\sum_{j=1}^{N}m_j\vec{v}_j\vec{v}_j^{\,T}+\sum_{j=1}^{N}\vec{r}_j\vec{F}_j^{\,T}\right) \tag{7-13}$$

式中，上标“$T$”为转置向量的符号。右边两项为：

$$\sum_{j=1}^{N} m_j \vec{v}_j \vec{v}_j^{T} = \begin{pmatrix} \sum_j m_j v_{jx} v_{jx} & \sum_j m_j v_{jx} v_{jy} & \sum_j m_j v_{jx} v_{jz} \\ \sum_j m_j v_{jy} v_{jx} & \sum_j m_j v_{jy} v_{jy} & \sum_j m_j v_{jy} v_{jz} \\ \sum_j m_j v_{jz} v_{jx} & \sum_j m_j v_{jz} v_{jy} & \sum_j m_j v_{jz} v_{jz} \end{pmatrix}$$

$$\sum_{j=1}^{N} \vec{r}_j \vec{f}_j^{T} = \begin{pmatrix} \sum_j x_j F_{jx} & \sum_j x_j F_{jy} & \sum^{j} x_j F_{jz} \\ \sum_j y_j F_{jx} & \sum_j y_j F_{jy} & \sum_j y_j F_{jz} \\ \sum_j z_j F_{jx} & \sum_j z_j F_{jy} & \sum_j z_j F_{jz} \end{pmatrix} \tag{7-14}$$

计算中依式(7-14) 算出瞬间压力张量，瞬间静压为 1/3 乘以上张量的迹 (trace)，即：

$$P = \frac{1}{3}(P_{xx} + P_{yy} + P_{zz}) \tag{7-15}$$

对材质而言，可以在不同的压力或应力的条件下从事分子动力计算，获得体积与单位晶格参数在不同压力或应力的改变量，建立应力-应变曲线 (stress-strain curve)。

分子动力计算中因采用周期性边界条件的关系，除了系统中粒子的相互作用力之外，尚有计算系统与其影像系统间的作用力。故计算位力时必须考虑影像系统的作用。若势能为成对加成，则位力可改写为：

$$W = \frac{1}{2}\sum_{i<j} \vec{r}_{ij} \cdot \vec{F}_{ij} \tag{7-16}$$

分子动力计算可控制系统的压力或应力，常用的有 Berendrsen、Andersen 或 Parrinello-Rahman 的方法。

## 7.3 径向分布函数

径向分布函数 (radial distribution function，RDF) 的物理意义可由图 7-6 显示。图中，黑球为流体系统中的一个分子，称其为“参考分子”，与其中心的距离由 $r \to r+dr$ 间的分子数目为 $dN$。

定义径向分布函数 $g(r)$ 为：

$$\rho g(r) 4\pi r^2 = dN \tag{7-17}$$

式中，$\rho$ 为系统的密度。若系统的分子数目为 $N$，则由以上的关系

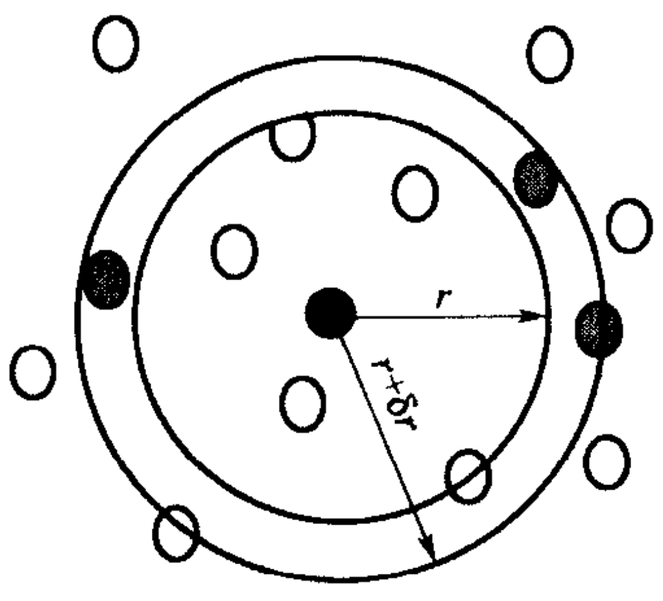



**图 7-6　径向分布示意**

可得：

$$\int_0^{\infty} \rho g(r) 4\pi r^2 \mathrm{d}r = \int_0^N \mathrm{d}N = N \tag{7-18}$$

由式(7-17) 得径向分布函数与 $\mathrm{d}N$ 的关系，即：

$$g(r) = \frac{\mathrm{d}N}{\rho 4\pi r^2 \mathrm{d}r} \tag{7-19}$$

径向分布函数可以解释为系统的区域密度（local density）与平均密度（bulk density）的比。参考分子的附近（$r$ 值小）区域密度不同于系统的平均密度，但与参考分子距离远时区域密度应与平均密度相同，即当 $r$ 值大时径向分布函数的值应接近 1。分子动力计算径向分布函数的方法为：

$$g(r) = \frac{1}{\rho 4\pi r^2 \delta r} \frac{\sum_{t=1}^{T} \sum_{j=1}^{N} \Delta N(r \to r + \delta r)}{N \times T} \tag{7-20}$$

式中，$N$ 为分子的总数目；$T$ 为计算的总时间（步数）；$\delta r$ 为设定的距离差；$\Delta N$ 为介于 $r \to r + \delta r$ 间的分子数目。图 7-7 为 Verlet 在不同温度下，计算 Lennard-Jones（LJ）作用的单原子分子流体的径向分布函数。

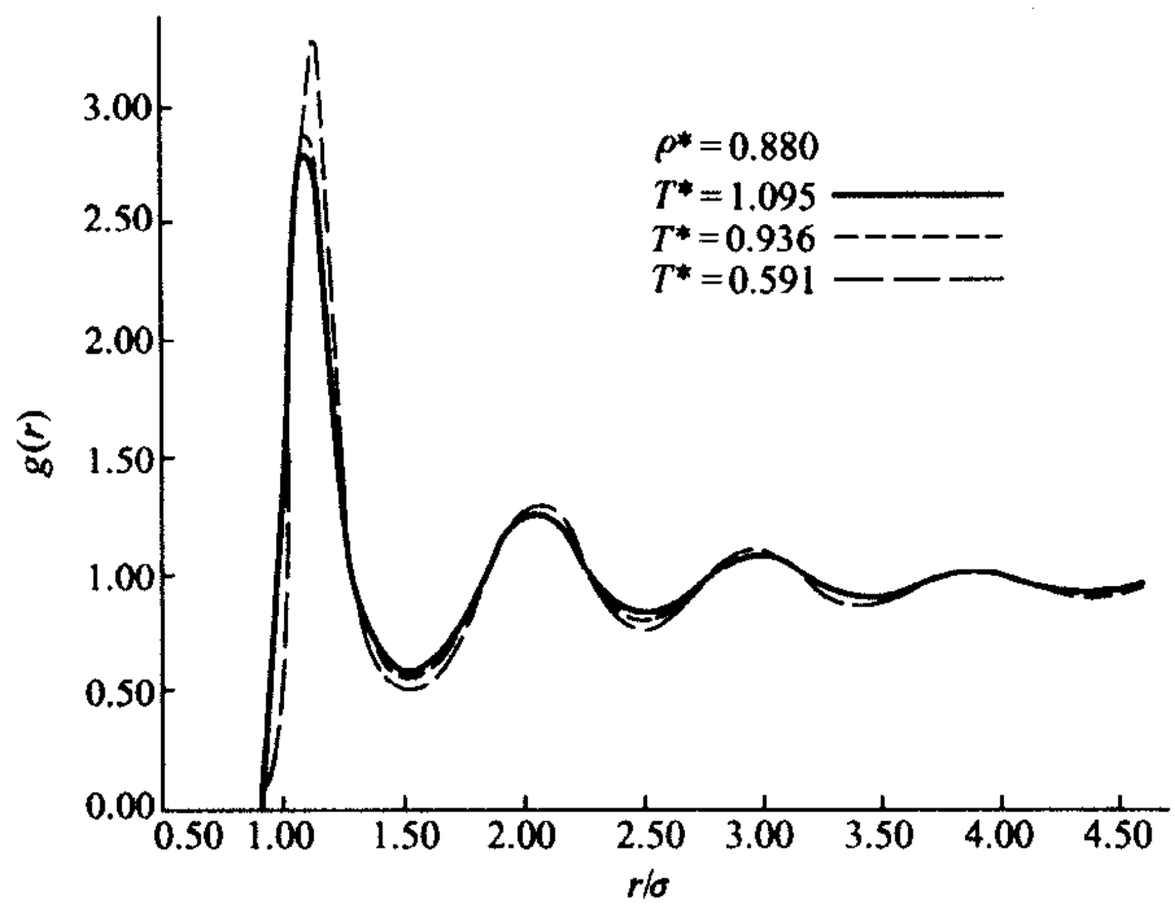



**图 7-7　LJ 粒子的径向分布函数图**

图 7-7 中的 $\sigma$ 为 LJ 势能的参数。此图显示，当 $r>4.5\sigma$ 时径向分布函数的值趋近于 1，当 $r<0.85\sigma$ 时径向分布函数均为零。这表示在所计算的流体系统中分子与分子最近的距离不会小于 $0.85\sigma$，平均看来，当分子间的距离超过 $4.5\sigma$ 即与均匀的液体性质相同。图 7-7 中，径向函数的最大峰出现于 $r_{max}$ 约 $1.2\sigma$ 处，峰强度的范围约为 $2.8\sim3.4\sigma$，这表示平均看来，离分子 $1.2\sigma$ 处附近出现其他分子的概率最大，此处的区域密度最大。图中亦显示其他的峰，在峰的位置附近，区域密度均大于平均密度。径向分布函数反映出液体中分子聚集的特性，可借此了解液体的“结构”。将 $g(r)$ 对 $r=0\rightarrow r=r_{max}$ 积分，即求图 7-7 中在此范围内的面积，其结果为在此范围内的平均分子数目。利用此方法可以求得水溶液中分子的水合数目（hydration number）。

径向分布函数的应用很广，除了上述关于液体结构外，还可用以计算系统的平均势能及压力，如：

$$\frac{U}{N}=\frac{1}{2}\rho\int_0^\infty u(r)\cdot g(r)4\pi r^2\,\mathrm{d}r \tag{7-21}$$

$$P=\rho T-\frac{1}{3}\,\frac{1}{2}\rho^2\int_0^\infty \frac{\mathrm{d}u(r)}{\mathrm{d}r}\cdot r\cdot g(r)4\pi r^2\,\mathrm{d}r \tag{7-22}$$

式中，势能为成对加成形式 $U=\sum_{i<j}u(r_{ij})$；$P$ 为压力；$\rho$ 为系统的密度。由此 $g(r)$ 所计算 LJ 液体的平均分子势能及压力分别为 $U/N=-4.419$，$P=5.181$ 与 7.2 节中直接由统计的平均值计算的结果十分吻合。

# 7.4 与时间相关物理量的计算

分子动力仿真方法除了可以计算系统的统计平均值之外，最重要的是，可以计算系统各种动态的特性。这些动态特性反映出系统中粒子的运动状态，可借以了解各种形态的运动。

## 7.4.1 均方位移

分子动力计算系统中的原子由起始位置不停移动，每一瞬间各原子的位置皆不相同。以 $\vec{r}_i(t)$ 表示时间 $t$ 时粒子 i 的位置。粒子位移平方的平均值称为均方位移（mean square displacement，MSD），即：

$$\mathrm{MSD}=R(t)=\langle|\vec{r}(t)-\vec{r}(0)|^2\rangle \tag{7-23}$$

式中，括号表示平均值。依据统计原理，只要分子数目够多，计算时

间够长，系统的任一瞬间均可当作时间的零点，所计算的平均值应相同。因此，由储存的轨迹计算均方位移应将各轨迹点视为零点。设分子动力计算共收集了 $n$ 步轨迹，各步的位置向量分别为 $\vec{r}(1)$，$\vec{r}(2),\cdots,\vec{r}(n)$，通常将此轨迹分为相等数目的两部分，计算均方位移时，每次计算 $R(t)$ 皆取 $n/2$ 组数据的平均。将轨迹分为：

$$\vec{r}(1),\vec{r}(2),\cdots,\vec{r}(n/2) \quad 及 \ \vec{r}(n/2+1),\vec{r}(n/2+2),\cdots,\vec{r}(n)$$

设步数的时间间隔为 $\delta t$，因为任一瞬间均可视为零点，故均方位移为：

$$R(\delta t)=\frac{|\vec{r}(2)-\vec{r}(1)|^2+|\vec{r}(3)-\vec{r}(2)|^2+\cdots+|\vec{r}(n/2+1)-\vec{r}(n/2)|^2}{n/2}$$

$$R(2\delta t)=\frac{|\vec{r}(3)-\vec{r}(1)|^2+|\vec{r}(4)-\vec{r}(2)|^2+\cdots+|\vec{r}(n/2+2)-\vec{r}(n/2)|^2}{n/2}$$

$$\cdots$$

$$R(m\delta t)=\frac{|\vec{r}(m+1)-\vec{r}(1)|^2+|\vec{r}(m+2)-\vec{r}(2)|^2+\cdots+|\vec{r}(m+n/2)-\vec{r}(n/2)|^2}{n/2}$$

$$\cdots$$

$$R(n\delta t/2)=\frac{|\vec{r}(n/2+1)-\vec{r}(1)|^2+|\vec{r}(n/2+2)-\vec{r}(2)|^2+\cdots+|\vec{r}(n)-\vec{r}(n/2)|^2}{n/2} \tag{7-24}$$

式(7-24) 为计算某一个粒子的均方位移，如计算系统中所有粒子的均方位移则需再对粒子数平均。图 7-8 为 LJ 液体系统的均方位移。

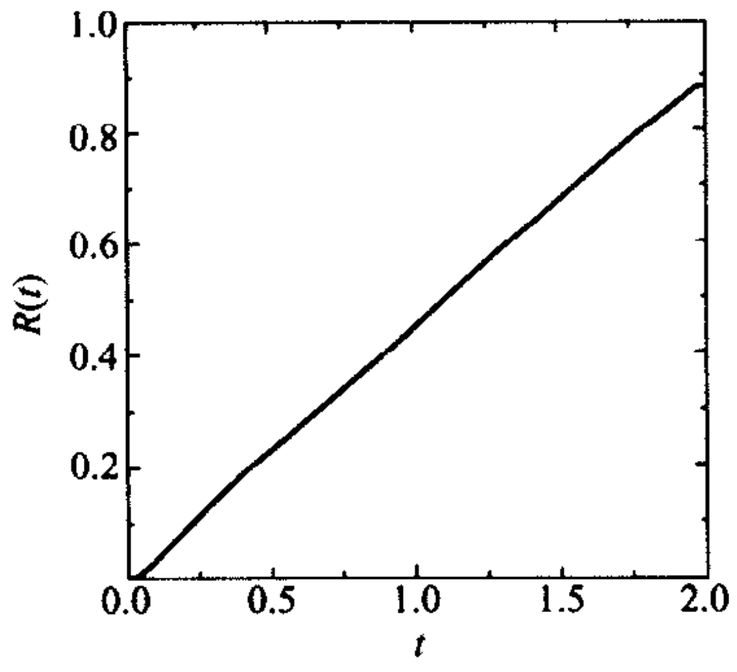



**图 7-8　LJ 粒子系统的均方位移图**

图 7-8 显示，当 $t$ 值小时，$R(t)$ 呈指数增加；当 $t$ 值大时 $R(t)$ 近似直线。根据爱因斯坦的扩散定律，有：

$$\lim_{t\to\infty}\langle|\vec{r}(t)-\vec{r}(0)|^2\rangle=6Dt \tag{7-25}$$

式中，$D$ 为粒子的扩散系数（diffusion constant）。因此，当时间很长时，均方位移对时间曲线的斜率即为 $6D$。由分子动力计算所得到 LJ 粒子系统的扩散系数为 $D=2.43\times10^{-5}\,\mathrm{cm^2/s}$。

## 7.4.2　相关函数

设 $A$，$B$ 分别表示系统的两个不同的物理量（如位置、速度、角度……），

则 $A$ 或 $B$ 的自身相关函数（auto correlation function）的表示式为：

$$C_A(t)=\langle A(t)\cdot A(0)\rangle=\langle A(T+t)\cdot A(T)\rangle$$
$$C_B(t)=\langle B(t)\cdot B(0)\rangle=\langle B(T+t)\cdot B(T)\rangle \tag{7-26}$$

相关函数的物理意义为一个物理量随时间改变后与其起始时的相关性。因为物理量的平均值不随选择时间的起点而改变，故式(7-26）中，$T$ 为任意的起始时间。$A$ 与 $B$ 的交互相关函数（cross correlation function）的表示式为：

$$C_{AB}(t)=\langle A(t)\cdot B(0)\rangle=\langle A(T+t)\cdot B(T)\rangle \tag{7-27}$$

相关函数（correlation function）表示物理量与物理量间与时间的相关性。通常定义归一化（normalized）的自身相关函数为：

$$\overline{C}_A(t)=\frac{C_A(t)}{C_A(0)}=\frac{\langle A(t)\cdot A(0)\rangle}{\langle A(0)\cdot A(0)\rangle} \tag{7-28}$$

由分子动力模拟所储存的轨迹计算相关函数的方法与计算均方位移一样，应将轨迹的各点均视为时间的起始点。设分子动力计算系统含有 $N$ 个分子，共收集了 $n$ 步轨迹，则：

$$C_A(t)=\frac{\sum_{i=1}^{N}[A(t+1)\cdot A(1)+A(t+2)\cdot A(2)+\cdots A(t+n/2)\cdot A(n/2)]}{N\times(n/2)} \tag{7-29}$$

图 7-9 例示储存轨迹中 $t=4\delta t$ 的相关性。

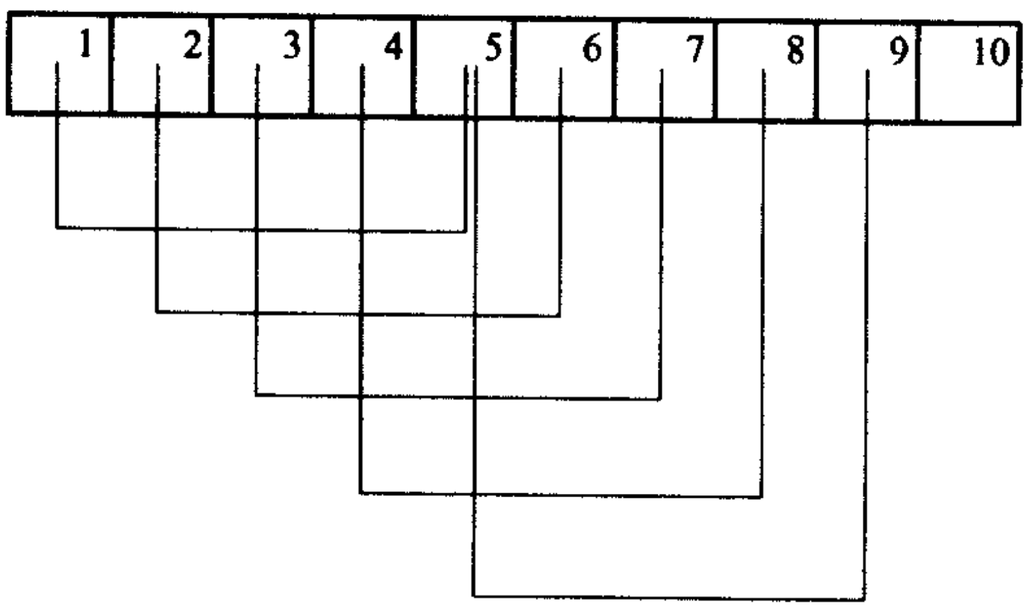



**图 7-9　A(t) · A(0), t= 4δt 的相关性**

由以上的讨论可知，相关函数可由所储存的轨迹直接计算。设积分步长为 $\delta t$，则 $t=\tau\delta t$ 的相关函数计算表示为：

$$C_{\mathrm{A}}(\tau)=\langle A(\tau)\cdot A(0)\rangle=\frac{1}{\tau_{\max}}\sum_{\tau_0=1}^{\tau\max}A(\tau_0)\cdot A(\tau_0+\tau) \tag{7-30}$$

式中，$\tau_{\max}$为所取的平均总数。以上直接计算相关函数的方法，若储存的轨迹点太多，则计算时需占去许多空间。除了直接计算相关函数外，最有效计算的计算方法为利用快速傅里叶变换（fast Fourier transform，FFT)。若不考虑归一化系数，则式(7-30）可写为

$$C'_{\mathrm{A}}(\tau)=\sum_{\tau_0=1}^{2\tau_{\mathrm{run}}}A(\tau_0)A(\tau_0+\tau)\quad(0\leqslant\tau<2\tau_{\mathrm{run}}) \tag{7-31}$$

式中，上标（′）表示未含归一化系数，$2\tau_{\mathrm{run}}$表示计算了 2 倍所储存的资料点。实际操作上，可于一组资料点后加上一组零，以满足此要求。为方便起见，将式(7-31）改写为：

$$C'_{\mathrm{A}}(\tau)=\sum_{\tau_0=0}^{2\tau_{\mathrm{run}}-1}A(\tau_0)A(\tau_0+\tau)\quad(0\leqslant\tau\leqslant2\tau_{\mathrm{run}}) \tag{7-32}$$

根据 FFT 的卷积定理（convolution theory)，若：

$$C(\tau)=\sum_{\tau'=0}^{n-1}A(\tau')B(\tau+\tau') \tag{7-33}$$

则：

$$\hat{C}(\nu)=\hat{A}^{*}(\nu)\hat{B}(\nu) \tag{7-34}$$

式中，符号（^）为由时间空间至频率空间 $\tau\rightarrow\nu$ 的傅里叶变换，即：

$$\hat{C}(\nu)=\sum_{\tau=0}^{n-1}C(\tau)\exp(-i2\pi\nu\tau/n)\quad(\nu=0,1,\cdots,n-1) \tag{7-35}$$

而由 $\nu\rightarrow\tau$ 的反傅里叶变换为

$$C(\tau)=\frac{1}{n}\sum_{\nu=0}^{n-1}\hat{C}(\tau)\exp(i2\pi\nu\tau/n)\quad(\tau=0,1,\cdots,n-1) \tag{7-36}$$

因此，相关函数可转换为：

$$\hat{C}'_{\mathrm{A}}(\nu)=\hat{A}^{*}(\nu)\hat{A}(\nu)=|\hat{A}(\nu)|^{2}\quad(\nu=0,1,\cdots,2\tau_{\mathrm{run}}-1) \tag{7-37}$$

此处，$\nu$ 为断续的频率指标。$C'_{\mathrm{A}}(\tau)$ 可由反转换得到：

$$C'_{\mathrm{A}}(\tau)=\frac{1}{2\tau_{\mathrm{run}}}\sum_{\nu=0}^{2\tau_{\mathrm{run}}-1}|\hat{A}(\nu)|^{2}\exp(i2\pi\nu\tau/2\tau_{\mathrm{run}}) \tag{7-38}$$

总结如上的讨论，利用傅里叶变换计算相关函数的步骤为：

① 将 $\tau_{\mathrm{run}}$个 MD 资料加上 $\tau_{\mathrm{run}}$个零。

② 利用 FFT 方法将 $A(\tau)$ 转换为 $\hat{A}(\nu)$。

③ 计算 $\hat{C}'_A(\nu)=|\hat{A}(\nu)|^2$。

④ 利用反傅里叶变换计算 $\hat{C}'_A(\nu)\rightarrow\hat{C}'_A(\tau)$。

⑤ 乘上归一化常数 $(\tau_{run}-\tau)^{-1}$，求得 $C_A(\tau)$。

图 7-10 为 LJ 液体系统计算所得的速度自身相关函数。图中显示，当时间长时（$t$ 值很大时），粒子的速度与其初速度已无关联性，故相关函数值趋近于零。反之，时间短时粒子速度与其初速度的关联性强，相关函数值较大。因为粒子的速度为向量，故图中的负值表示与初速度的方向相反。相关函数为统计力学中最重要的一种函数，可借由不同形式的相关函数计算各种与时间有关物理量的平均值。

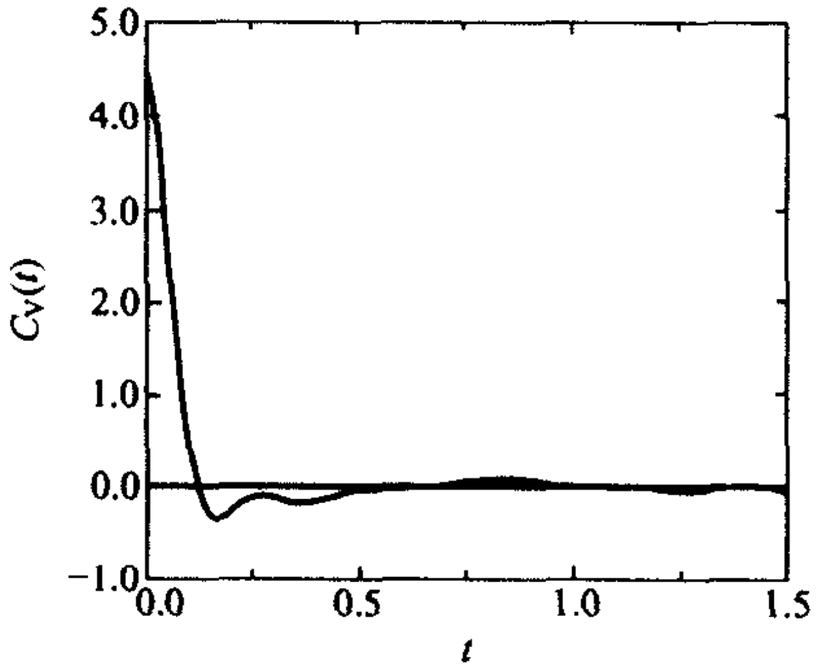



图 7-10　LJ 粒子系统的速度相关函数

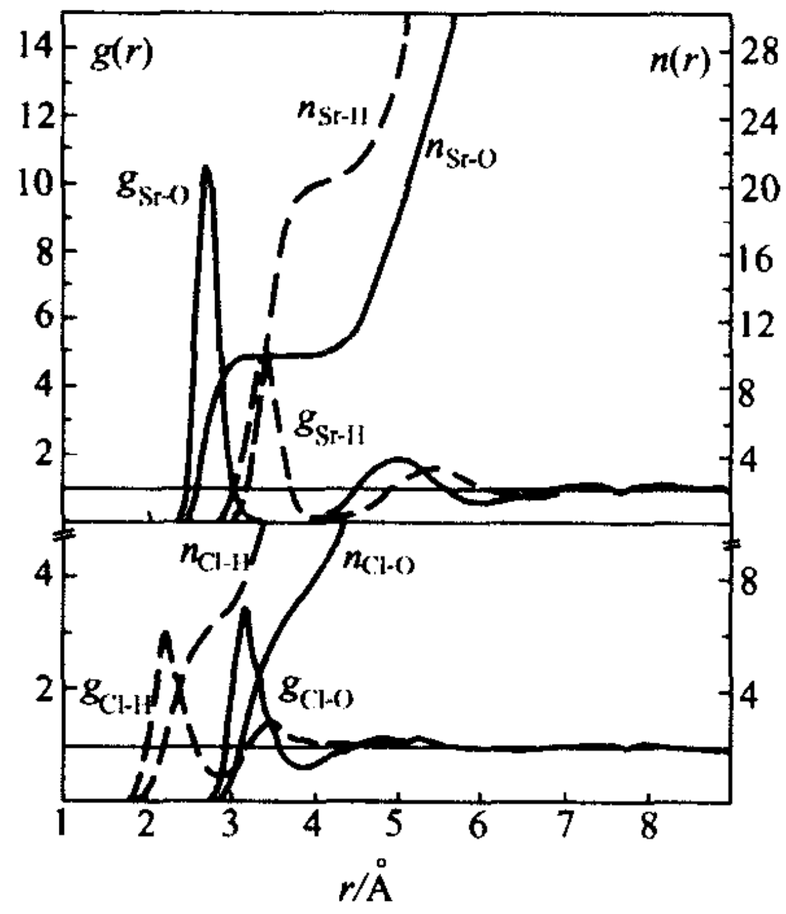



图 7-11　$SrCl_2$ 水溶液的径向分布函数

速度的自身相关函数可用以计算粒子的扩散系数，其关系为：

$$D=\frac{1}{3}\int_0^\infty\langle\vec{v}(t)\cdot\vec{v}(0)\rangle\mathrm{d}t \tag{7-39}$$

下例为 Sophr 等对 $Sr^{2+}$ 与 $Cl^-$ 水溶液的分子动力模拟计算的结果。计算系统包含 200 个水分子、4 个 $Sr^{2+}$ 阳离子与 8 个 $Cl^-$ 阴离子，系统的密度为 1.145g/cm³，相当于重量摩尔浓度 1.1mol/L。图 7-11 为各原子对的径向分布函数，

$$g_{AB}=\frac{\Delta n}{4\pi(r_{AB})^2\Delta r} \tag{7-40}$$

式中，$r_{AB}$为原子对的距离；$\Delta n$ 为以 A 为中心在 $r_{AB} \rightarrow r_{AB}+\Delta r$ 范围内平均的 B 原子数目。

图中并显示利用径向分布函数的积分所得的水合数目 $n_{A-O}$与 $n_{A-H}$，即：

$$n_{AB} = \rho_0 \int_0^r g_{AB}(r') 4\pi r'^2 \mathrm{d}r' \tag{7-41}$$

式中，$\rho_0$ 为 B 的数目密度。由此积分式，可得到 $Sr^{2+}$ 第一水合层 (hydration shell) 的水合数目为 9.8，与 X 射线衍射实验的结果相近。图 7-12 为各离子与原子速度相关函数的计算。

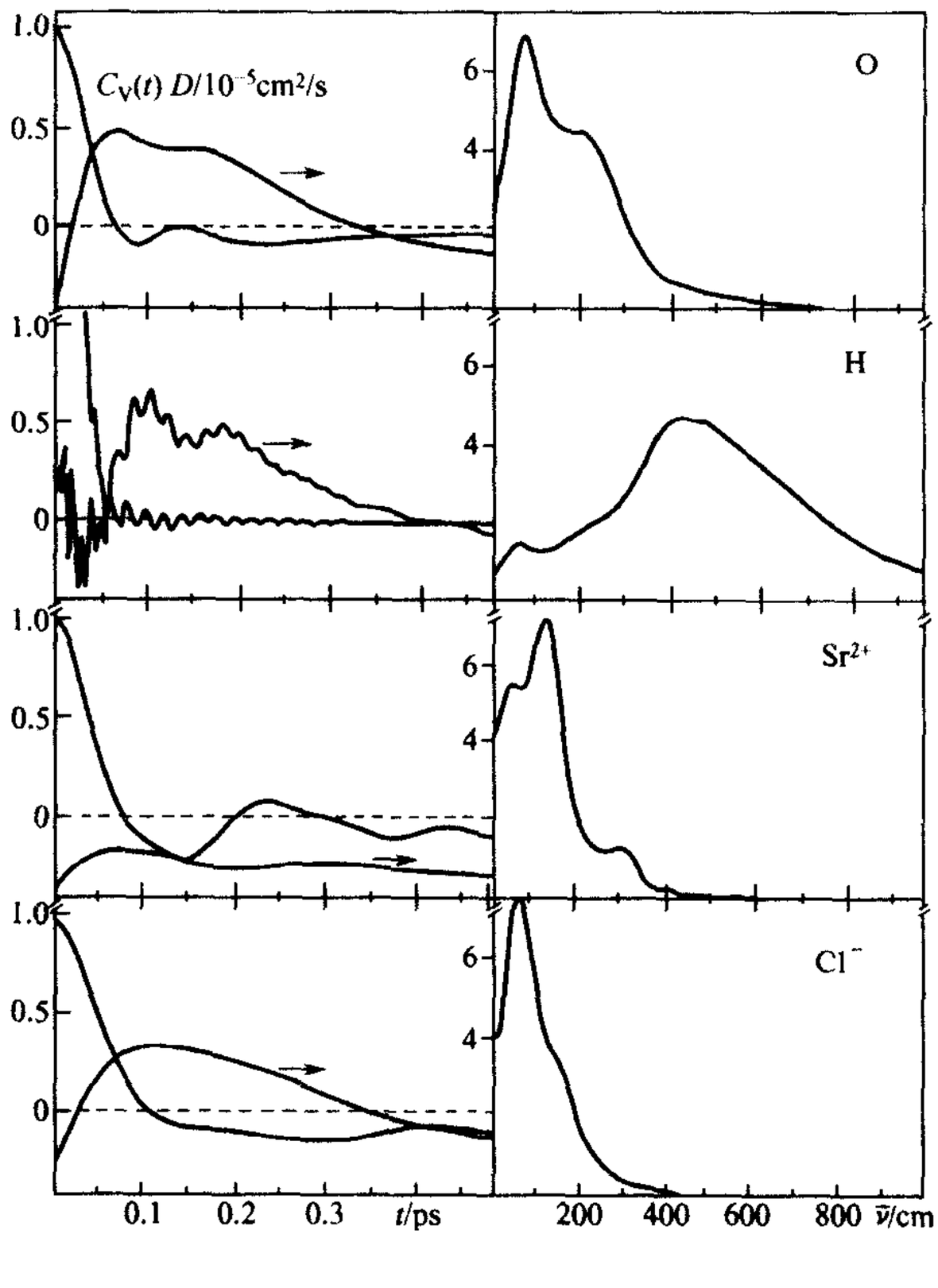



**图 7-12　$SrCl_2$ 水溶液的速度相关函数及其功率谱**

图的左边显示各原子与离子的归一化速度相关函数 Cv($t$)，与 1/3×[Cv($t$)曲线下的面积]。由计算所得的离子在水中的扩散系数为：$Cl^-$ $(1.2\pm0.3)\times10^{-5}cm^2/s$；$Sr^{2+}$ $(0.6\pm0.1)\times10^{-5}cm^2/s$。这计算的结果

表示离子水溶液中，阴离子的运动较阳离子更快。

速度相关函数除了可用以计算扩散系数外还有一重要的应用。流体系统中的原子不停地运动，与原子运动相关的运动模式各有其运动的频率。以水分子中的氢原子为例，与其相关的运动包括水分子的移动，转动与各种内振动。分子的移动与转动频率较小，而分子内的振动频率较大，通常属于红外光的侦测范围（400～4000cm$^{-1}$）。与原子有关运动模式的频率可经由速度相关函数的傅里叶转换计算，即：

$$I(\omega)= \mathrm{Re}\left[\frac{1}{2\pi}\int_0^{\infty}\langle\vec{v}(t)\cdot\vec{v}(0)\rangle e^{i\omega t}\,dt\right] \tag{7-42}$$

式中，$\omega$ 为频率；Re 表示取括号中的实数部分；$I(\omega)$ 称为该频率的图谱密度（spectra density）。根据计算的轨迹将式(7-42) 转成：

$$I(\omega)\propto \lim_{N_T\to\infty}\sum_{n=0}^{N_T} C_V(n\Delta t)\cos(\omega\cdot n\Delta t)\Delta t \tag{7-43}$$

式中，$N_T$ 为计算的总步数。将各频率的图谱密度绘出所得的图谱称为功率谱（power spectrum）。图 7-12 的右边为由速度相关函数所计算各离子或原子的功率谱。由功率谱可检视各种运动的频率。图中显示溶液中的离子与分子的“阻碍移动”（hindered translation）运动。阻碍移动的意义为在溶液中离子或分子的移动受到周围环境的阻碍，结果类似在一区间内来回移动。如 $Cl^-$ 的功率谱中最大峰位于 50cm$^{-1}$处，表示 $Cl^-$ 来回移动的频率约为 50cm$^{-1}$。$Sr^{2+}$ 功率谱的最大峰位于 125cm$^{-1}$ 处，并于 50cm$^{-1}$与 300cm$^{-1}$处有两个较低的峰。这表示 $Sr^{2+}$ 在溶液中具有移动较 $Cl^-$ 快的运动，在 300cm$^{-1}$的峰可能是 $Sr^{2+}$ 与其水合分子间的$Sr^{2+}$…O振动。检视 O 与 H 的功率谱，在 50cm$^{-1}$的峰可能是水分子间氢键的弯曲振动，200cm$^{-1}$的峰可能是水分子间 O…O 的伸缩振动。依此，由计算的功率谱可分析系统中各种特殊形态的运动及其运动的频率，对了解分子间的作用有很大的帮助。

分子偶极的相关函数可用以计算红外线光谱。光谱的线状函数（line-shape function）为：

$$I(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty} dt e^{-i\omega t}\langle\vec{M}(t)\cdot\vec{M}(0)\rangle \tag{7-44}$$

式中，$\vec{M}$ 为分子的偶极矩。红外线光谱所侦测的分子内振动的频率与由功率谱所得到的振动频率应相同。故比对计算的功率谱与实验的红外光谱可用以检验分子动力计算的准确性。图 7-13 为商业软件 CERIUS 2 的 DYNAMICS 模块中利用分子动力计算与 UNIVERSAL 力场计算所得的聚

甲基异丁烯酸盐［poly（methyl methacrylate），PMMA］的红外光谱与实验的红外光谱的比较。

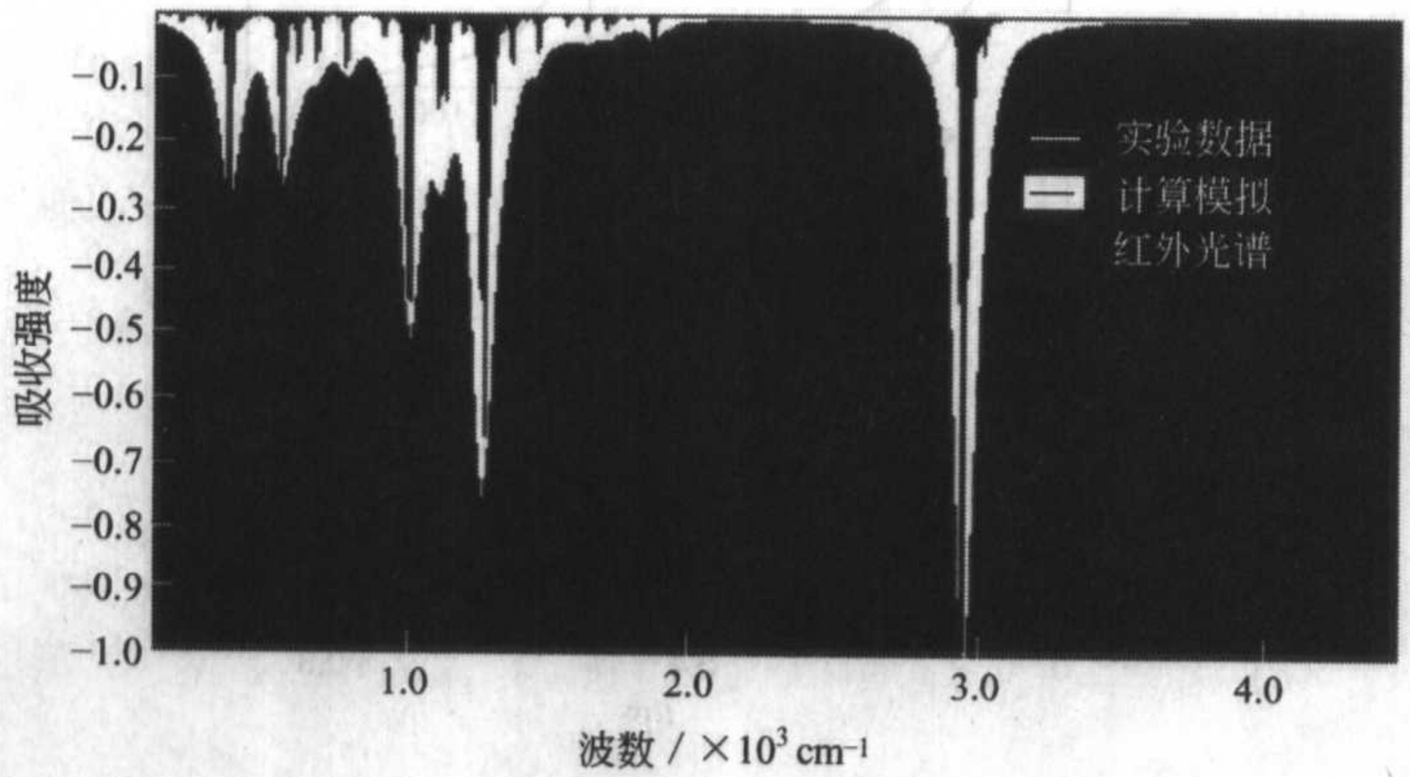



**图 7-13　Cerius 2 软件模拟所得 PMMA 的红外光谱**

图中，黑色的峰为计算所得；白色为实验所得；纵轴的坐标为吸收强度。此模拟与实验的红外光谱不论吸收频率或吸收强度均非常相近。

除了红外光吸收光谱的计算以外，利用相关函数亦可计算散射（scattering）光谱。拉曼光谱的线状函数为：

$$I(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{C}(t)\mathrm{e}^{-i\omega t}\,\mathrm{d}t \tag{7-45}$$

式中，$\hat{C}(t)$ 为归一化的转动相关函数（rotational correlation function）。分子全对称伸缩（totally symmetric stretch）振动模式的 $\hat{C}(t)$ 为：

$$\hat{C}(t)=\langle P_2[\vec{u}(t)\cdot\vec{u}(0)]\rangle \tag{7-46}$$

式中，$\vec{u}$ 为分子的偶极矩；$P_2$ 为勒让德多项式（Legendre polynomial），即：

$$P_2(x)=\frac{1}{2}(3x^2-1) \tag{7-47}$$

图 7-14 为 Böhm 等利用分子动力模拟方法计算所得液态 $CH_3CN$ 系统的转动相关函数。图中显示计算的相关函数与由实验的拉曼及红外光谱所度量的相关函数的比较，小图为计算的拉曼光谱。

相关函数的应用范围很广，许多系统动态的特性可由选择适当的相关函数计算。关于相关函数的计算可参考 McQuarrie 的著作。除了以上的各

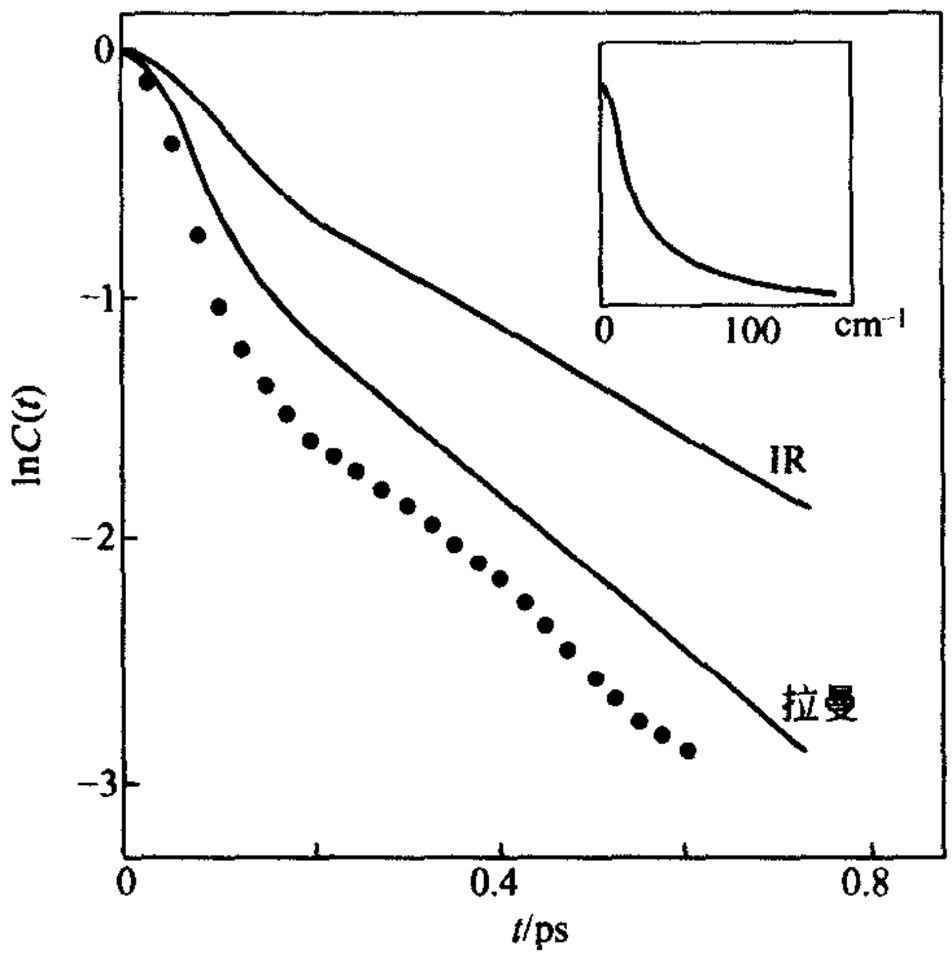



**图 7-14　液态 $CH_3CN$ 分子系统的转动相关函数与光谱的比较**

项性质外，较重要的如黏度（viscosity）与热导率（thermal conductivity coefficient）的计算。黏度 $\eta$ 的算式为：

$$\eta = \frac{1}{Vk_B T}\int_0^\infty \langle J(t) \cdot J(0)\rangle \mathrm{d}t \qquad (7\text{-}48)$$

$$J = \sum_{j=1}^{N}\left(\frac{p_{xj}p_{zj}}{m} + z_j F_{jx}\right)$$

式中，$p$ 为粒子的动量；$m$ 为质量；$F$ 为所受的力。热导率 $\lambda$ 的算式为：

$$\lambda = \frac{1}{3Vk_B T}\int_0^\infty \langle \vec{S}(t) \cdot \vec{S}(0)\rangle \mathrm{d}t \qquad (7\text{-}49)$$

式中，$\vec{S}(t)$ 为通量的向量式（vectorial flux），即：

$$S = \sum_{j=1}^{N}\left\{\left(\frac{\vec{p}_j^{\,2}}{2m} - \langle h_j\rangle\right)\frac{\vec{p}_j}{m} + \frac{1}{2}\sum_{j\neq i}[\vec{r}_{ij} \cdot \vec{F}_{ij} + U(r_{ij})I]\right\} \cdot \frac{\vec{p}_j}{m} \qquad (7\text{-}50)$$

式中，$\langle h_j\rangle$ 为分子的焓（enthalpy）。

一般以分子动力计算模拟系统的黏度与热导率与实验值偏差较大，这是由于所模拟的体系太小所致。分子动力计算的系统仅含有数千个粒子，而实际的体系则含有约 $10^{20}$ 个粒子，因此并不能完全反映系统宏观的动力特性。

## 7.5 自由能的计算

### 7.5.1 自由能的变化

自由能（free energy）是影响系统变化最重要的因素。依据统计热力学，$(N,V,T)$ 系统的自由能与系统的配分函数 $Q$（partition function）相关，即：

$$A=-k_{\mathrm{B}}T\ln Q \tag{7-51}$$

依此式，自由能的变化为：

$$\Delta A=A_2-A_1=-k_{\mathrm{B}}T\ln\frac{Q_2}{Q_1} \tag{7-52}$$

式(7-52) 中，下标 1、2 为起始与终结状态。理论上，配分函数可由经典力学配分函数的定义直接计算，得：

$$Q=(1/N!h^{3N})\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp[-H(\vec{p},\vec{q})/k_{\mathrm{B}}T] \tag{7-53}$$

式中，$H(\vec{p},\vec{q})$为系统的哈密顿量(Hamiltonian)，为动能与势能的总和，即：

$$H(\vec{p},\vec{q})=\mathrm{K.E}(\vec{p})+U(\vec{q}) \tag{7-54}$$

$\vec{p}$ 与 $\vec{q}$ 分别表示系统中粒子的动量与坐标。但实际上，若由分子动力模拟方法直接计算系统的配分函数需要非常长的时间才能收敛，不符合经济的效益。因此，必须寻找别的方法计算自由能。由式(7-51) 及式(7-53) 得：

$$\begin{aligned}\Delta A&=-k_{\mathrm{B}}T\ln\frac{Q_2}{Q_1}=-k_{\mathrm{B}}T\ln\left\{\frac{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp(-H_2/k_{\mathrm{B}}T)}{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp(-H_1/k_{\mathrm{B}}T)}\right\}\\&=-k_{\mathrm{B}}T\ln\left\{\frac{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp(-H_2/k_{\mathrm{B}}T)\exp(H_1/k_{\mathrm{B}}T)\exp(-H_1/k_{\mathrm{B}}T)}{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp(-H_1/k_{\mathrm{B}}T)}\right\}\\&=-k_{\mathrm{B}}T\ln\left\{\frac{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp[-(H_2-H_1)/k_{\mathrm{B}}T]\exp(-H_1/k_{\mathrm{B}}T)}{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\exp(-H_1/k_{\mathrm{B}}T)}\right\}\end{aligned} \tag{7-55}$$

依照统计力学平均值的关系式（将常数忽略），平衡系统的任一物理量 $B$ 的平均值为：

$$\langle B\rangle=\frac{\iint \mathrm{d}\vec{p}\,\mathrm{d}\vec{q}\,B\cdot \mathrm{e}^{-H/k_{\mathrm{B}}T}}{Q} \tag{7-56}$$

根据此关系，可将式(7-55) 写成：

$$\Delta A = A_2 - A_1 = -k_B T \ln\langle \exp(-\Delta H / k_B T)\rangle_1 = -k_B T \ln\langle \exp(-\Delta E_{12} / k_B T)\rangle_1 \quad (7\text{-}57)$$

式中，$\Delta H = H_2 - H_1 = \Delta E_{12}$ 为由状态 1 改变为状态 2 系统的能量变化，“〈　〉”为平均值。式中，下标 1 表示对状态 1 所取的平均值。同样的，仿照如上的推导，亦可得到：

$$\Delta A = -k_B T \ln\langle \exp[-(H_2 - H_1)/k_B T]\rangle_2 = -k_B T \ln\langle \exp(-\Delta E_{12} / k_B T)\rangle_2 \quad (7\text{-}58)$$

式中，下标 2 表示对状态 2 所取的平均值。上式计算自由能的变化只取决于起始状态与终结状态的能量，与反应进行的途径无关，为 Zwanzig 所提出。利用式(7-58) 必须先定义 $H_0$ 与 $H_1$，然后执行状态 1 的计算，以求出系综平均值 $\exp(-\Delta E/k_B T)$；或者执行状态 2 的计算求出对应于式(7-58)的系综平均值。以计算乙醇（$CH_3CH_2OH$）及硫醇（$CH_3CH_2SH$）在水溶液中自由能的变化为例。定状态 1 为乙醇，状态 2 为硫醇，则自由能的变化可根据式(7-57)，由计算乙醇的水溶液获得。实际计算的方法为：计算中所产生的每一瞬间构象（configuration)，将乙醇的氧原子以硫取代，计算其两者的能量差。或者，以同样的方法，依式(7-58) 计算硫醇的水溶液，于每一构象将硫醇中的硫原子以氧原子取代，并计算其二者的能量差。当自由能的变化小于 $2k_B T$ 时这种计算的结果非常精准，但若自由能的变化超过此范围，则导致统计的不准度。即若状态 1 与状态 2 于相空间（phase space）的重叠性高（分子大致类似)，则可利用这种方法计算自由能的变化；反之，若重叠性很低，则不易得到正确的自由能变化值。

为了克服重叠性低的障碍，可利用分段的方法。设状态 $i$ 介于 1→2 状态间，则可将由 1～2 的变化分为由 1～$i$ 与由 $i$～2，即：

$$\begin{aligned}\Delta A_{12} &= A_2 - A_1 = (A_2 - A_i) + (A_i - A_1) \\ &= -k_B T \ln\left(\frac{Q_2}{Q_i} \cdot \frac{Q_i}{Q_1}\right) \\ &= -k_B T \ln\langle \exp(-\Delta E_{i2}/k_B T)\rangle - k_B T \ln\langle \exp(-\Delta E_{1i}/k_B T)\rangle \end{aligned} \quad (7\text{-}59)$$

依此原则，可将 1→2 状态的变化分成许多自由能变化不大的阶段，分别计算各阶段自由能的变化，再将其相加得到结果；

$$\Delta A_{12} \sim (\Delta A_{1m} + \Delta A_{mm+1} + \Delta A_{m+1m+2} + \cdots + \Delta A_{n2})$$

$$\begin{aligned}\Delta A_{12} &= -k_B T \ln\left(\frac{Q_2}{Q_n} \cdot \frac{Q_n}{Q_{n-1}} \cdots \frac{Q_{m+1}}{Q_m} \cdot \frac{Q_m}{Q_1}\right) \\ &= -k_B T \ln\langle \exp(-\Delta E_{n2}/k_B T)\rangle - k_B T \ln\langle \exp(-\Delta E_{n-1,n}/k_B T)\rangle + \cdots \\ &\quad - k_B T \ln\langle \exp(-\Delta E_{1m}/k_B T)\rangle \end{aligned} \quad (7\text{-}60)$$

以上这种计算自由能变化的方法，称为热力微扰法（thermodynamic perturbation method)。

## 7.5.2 微扰计算

执行式(7-60) 的计算，需计算两状态的能量差。因为动能的平均值皆为$\frac{3}{2}Nk_BT$，故一般可将动能项由是将由式(7-60) 中消去。所以计算能量差简化为计算势能差。通常将状态 1 改变为状态 2 的势能变化分成许多中间状态，每个状态的能量可写成参数 λ 的函数，即：

$$U_m(\lambda_m)=U_1(1-\lambda_m)+U_2\lambda_m \tag{7-61}$$

λ 的范围为 0→1。依此关系式，若 $\lambda=0$，则 $U_m=U_1$；若 $\lambda=1$，则 $U_m=U_2$。

由状态 $U_m$ 变换为状态 $U_{m'}$ 的配分函数的比为：

$$\frac{Q(U_{m'})}{Q(U_m)}=\frac{\int e^{-U_{m'}/k_BT}d\vec{q}}{\int e^{-U_m/k_BT}d\vec{q}}=\int e^{-(U_{m'}-U_m)/k_BT}d\vec{q}\left(\frac{e^{-U_m/k_BT}}{\int e^{-U_m/k_BT}d\vec{q}}\right)$$

$$=\langle e^{-(U_{m'}-U_m)/k_BT}\rangle \tag{7-62}$$

式中，〈 〉为系综的平均值，亦可视为长时间的平均值。由式(7-52) 及式(7-62) 得自由能的改变为：

$$\exp[-\Delta A(\lambda_m\rightarrow\lambda_{m'})/k_BT]=\langle\exp[-(U_{m'}-U_m)/k_BT]\rangle_m \tag{7-63}$$

〈 〉$_m$ 表示于 $U_m$ 点计算的平均值。若 λ 的变化很小，则将式(7-63) 中 exp 式展开，比较等式左右两端的个别项，得：

$$\Delta A(\lambda_m\rightarrow\lambda_{m'})=\langle U_{m'}-U_m\rangle_m,(\lambda_m-\lambda_{m'})\rightarrow 0 \tag{7-64}$$

此式表示，微扰（perturbation）自由能的变化，可由计算两状态势能差异的平均值得到。依此，若将 0→1 的范围分为 $n$ 份，则整体自由能的变化为：

$$\Delta A(U_1\rightarrow U_2)=\sum_{m=0}^{n-1}\Delta A(\lambda_m-\lambda_{m+1}) \tag{7-65}$$

这种计算自由能的方法均称为微扰计算法（perturbation method），λ 称为耦合系数（coupling constant）。例如，以原子在溶剂中带电量改变过程为例，King 与 Warshel 等利用上述微扰法计算的自由能变化见图 7-15。

图 7-15 显示，此过程的自由能变化量约为－100kcal/mol，与由 Born 方程式，即：

$$\Delta A(q)=-166(q^2/\bar{a})(1-1/d) \tag{7-66}$$

与预测的值相同。其中，$q$ 为电荷量；$\bar{a}$ 为溶剂分子形成的平均孔洞

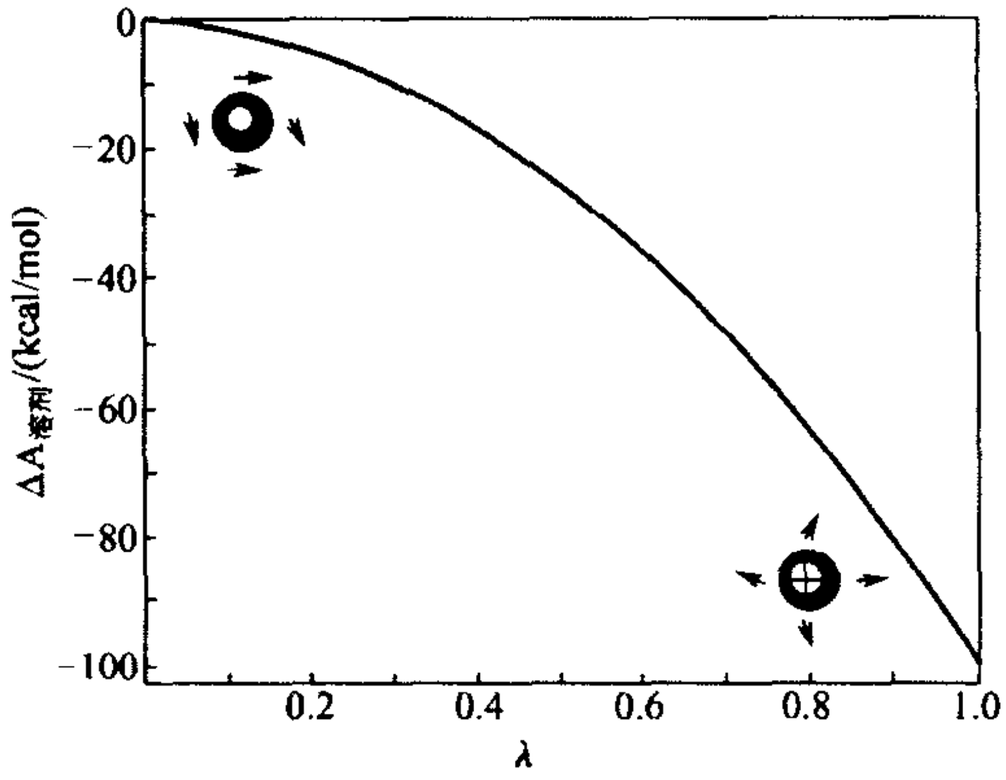



**图 7-15　微扰方法计算的自由能变化**

半径；$d$ 为溶剂的介电常数（dielectric constant）。

## 7.5.3　耦合系数

由式(7-61)，计算自由能变化均需先决定何者为耦合系数。将系统的势能定为耦合系数的函数，再利用连续变化耦合系数计算自由能的变化。一般可将力场的参数写成耦合常数的函数，如此，则以此力场参数计算的系统势能即为此耦合常数的函数。以简单的简谐振子为例，振子的势能为力常数 $K$ 与平衡键矩 $r_0$ 的函数，即：

$$E(r)=K(r-r_0)^2 \tag{7-67}$$

设计算自由能的变化为化学键由一种形式 $A$ 变为另一种形式 $B$，各不同的形式均有其特定的力常数与平衡键矩，则可定能量式为：

$$E(r,\lambda)=[\lambda K_A+(1-\lambda)K_B]\{r-[\lambda r_0^A+(1-\lambda)r_0^B]\}^2 \tag{7-68}$$

依此，当 λ 逐渐由 0 改变为 1，此化学键由 $A$ 状态逐渐转换为 $B$ 状态。以乙醇/硫醇的水溶液为例。系统的势能为：

$$U(\vec{r}^N)=\sum_{\text{bond}}\frac{1}{2}k_i(l_i-l_i^0)^2+\sum_{\text{angle}}\frac{1}{2}k_\theta^a(\theta_i-\theta_i^0)^2+\sum_{\text{torsion}}\frac{V_n}{2}\cdot$$
$$[1+\cos(n\omega-\gamma)]+\sum_{i=1}^{N}\sum_{j>i}^{N}\left\{4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]+\frac{q_iq_j}{4\pi\varepsilon_0 r_{ij}}\right\} \tag{7-69}$$

当 λ 由 0 逐渐变化为 1，系统的哈密顿量由 $H_1$ 变化为 $H_2$。变化过程中对应于 λ 的势能项中的每一项，均可比照式(7-68) 写成状态 1 与状态 2

的线性组合：

① 键伸缩项
$$k(\lambda)=\lambda k_2+(1-\lambda)k_1$$
$$l_0(\lambda)=\lambda l_0(2)+(1-\lambda)l_0(1)$$

② 键角弯曲
$$k_\theta(\lambda)=\lambda k_\theta(2)+(1-\lambda)k_\theta(1)$$
$$\theta_0(\lambda)=\lambda\theta_0(2)+(1-\lambda)\theta_0(1)$$

③ 扭转角
$$V_n(\lambda)=\lambda V_n(2)+(1-\lambda)V_n(1)$$

④ 静电作用
$$q_i(\lambda)=\lambda q_i(2)+(1-\lambda)q_i(1)$$

⑤ 范德瓦耳斯作用
$$\epsilon(\lambda)=\lambda\epsilon(2)+(1-\lambda)\epsilon(1)$$
$$\sigma(\lambda)=\lambda\sigma(2)+(1-\lambda)\sigma(1) \qquad (7\text{-}70)$$

于每一不同的 $\lambda$ 值，引用如上的势能参数，执行分子动力模拟。模拟的第一步为设 $\lambda=0$，将系统计算至达热平衡；然后依照式(7-63) 和式(7-65)，逐步计算 $-k_BT\ln\langle\exp(-\Delta H_i/k_BT)\rangle$，$\Delta H_i=H_{i+1}-H_i$ 得到 $\Delta A(\lambda_i\rightarrow\lambda_{i+1})$。由此，$\lambda$ 由 0→1 的自由能变化为各步骤自由能变化的和，如图 7-16 曲线下的面积。

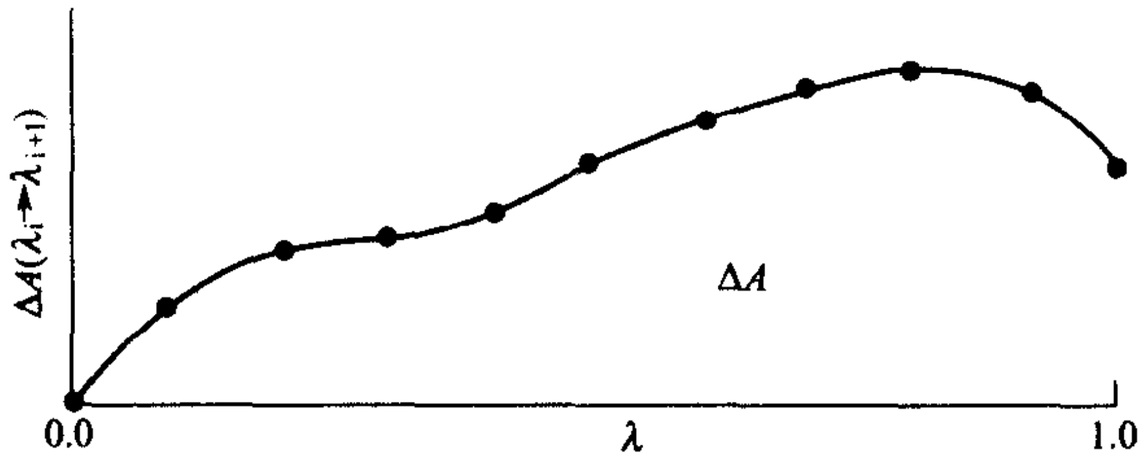



**图 7-16　微扰计算自由能图**

以上的计算方法为向前取样法（forward sampling），因为自由能由 $\lambda_i\rightarrow\lambda_{i+1}$ 的取样方法计算。相对的亦可由向后取样法（backward sampling）计算，即计算 $\lambda_i\rightarrow\lambda_{i-1}$ 的变化。不论何种取样法，$\lambda$ 的变化均为 0→1。图 7-17 为双向取样（double-wide sampling）的示意图。例如于图中的 $B$ 点同时计算 $\lambda_i\rightarrow\lambda_{i+1}$ 与 $\lambda_i\rightarrow\lambda_{i-1}$ 自由能的变化。这种方法经测试后，证实较单向取样法更为有效，计算效率为前者的 2 倍。

## 7.5.4　热力积分方法

除了前述的微扰计算法外，还有其他计算自由能的方法。自由能的变

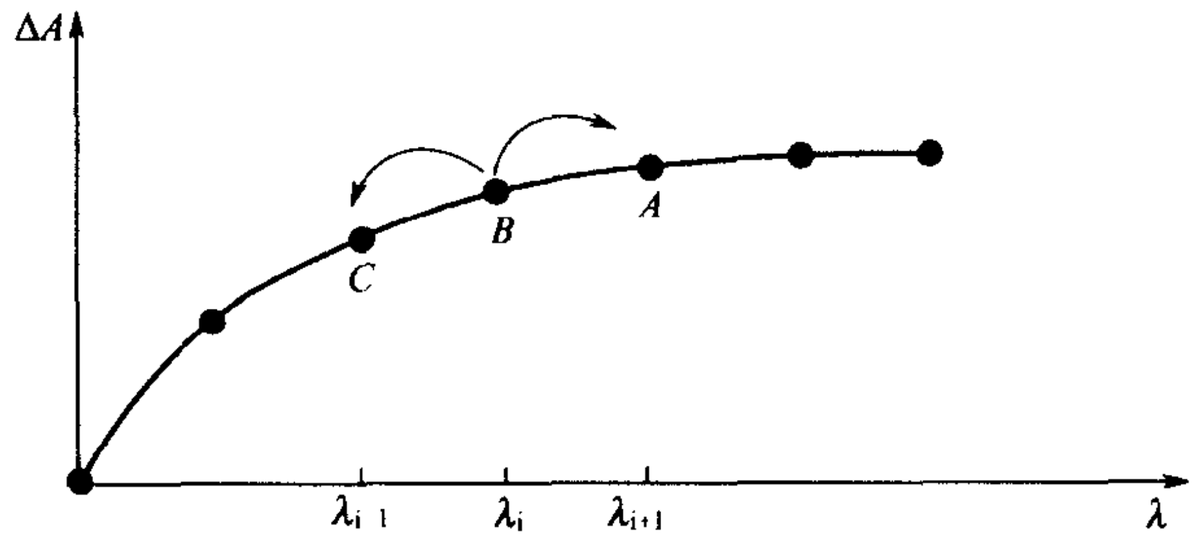



**图 7-17　双向取样法示意**

化可以积分式表示为：

$$\Delta A = \int_0^1 \frac{\delta A(\lambda)}{\delta \lambda} \mathrm{d}\lambda \tag{7-71}$$

式中，λ 为耦合系数，或解释为变化的程度。如上热力学的自由能变化积分式可用数值积分的方式计算，称为有限差分热力积分法（finite difference thermodynamic integration，FDTI）。依照式(7-71)，利用自由能与配分函数的关系，即：

$$A(\lambda) = -k_{\mathrm{B}} T \ln Q(\lambda) \tag{7-72}$$

得：

$$\Delta A = -k_{\mathrm{B}} T \int_0^1 \left[ \frac{\partial \ln Q(\lambda)}{\partial \lambda} \right] \mathrm{d}\lambda = \int_0^1 \frac{-k_{\mathrm{B}} T}{Q(\lambda)} \cdot \frac{\partial Q(\lambda)}{\partial \lambda} \mathrm{d}\lambda \tag{7-73}$$

由经典力学的配分函数的表示式，即：

$$Q = \frac{1}{N!} \frac{1}{h^{3N}} \iint \mathrm{d}\vec{p}\, \mathrm{d}\vec{r} \exp\left(-\frac{H}{k_{\mathrm{B}} T}\right) \tag{7-74}$$

可得配分函数的微分式为：

$$\frac{\partial Q(\lambda)}{\partial \lambda} = \frac{1}{N!} \frac{1}{h^{3N}} \iint \mathrm{d}\vec{p}\, \mathrm{d}\vec{r}\, \frac{\partial}{\partial \lambda} \exp\left(-\frac{H}{k_{\mathrm{B}} T}\right) = -\frac{1}{N!} \frac{1}{h^{3N}} \frac{1}{k_{\mathrm{B}} T} \iint \mathrm{d}\vec{p}\, \mathrm{d}\vec{r}\, \frac{\partial H}{\partial \lambda} \exp\left(-\frac{H}{k_{\mathrm{B}} T}\right) \tag{7-75}$$

将此微分式代入式(7-73)，得：

$$\frac{\partial A(\lambda)}{\partial \lambda} = \iint \mathrm{d}\vec{p}\, \mathrm{d}\vec{r}\, \frac{\partial H}{\partial \lambda} \left\{ \frac{\exp(-H/k_{\mathrm{B}} T)}{Q(\lambda)} \right\} = \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \tag{7-76}$$

将此关系代入式(7-71)，得：

$$\Delta A = \int_0^1 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \mathrm{d}\lambda \tag{7-77}$$

此式表示，计算自由能的变化可利用对应于每一 λ 值对哈密顿量 $H$

的微分求得。若 $\lambda$ 的改变很小，则哈密顿量 $H$ 的微分可以用数值方法计算，得：

$$\frac{\partial H(\lambda)}{\partial \lambda}=\frac{H(\lambda\pm\delta\lambda)-H(\lambda)}{\pm\delta\lambda} \tag{7-78}$$

实际计算时，选取 0→1 间的 $k$ 个 $\lambda$ 值，则由 $\lambda_i$ 变化为 $\lambda_i\pm\delta\lambda$ 时自由能的变化依式(7-77) 及式(7-78) 为：

$$\Delta A_i=A(\lambda_i)-A(\lambda_i\pm\delta\lambda)=-k_B T\ln\langle\exp\{-[E(\lambda_i)-E(\lambda_i\pm\delta\lambda)]/k_B T\}\rangle_i \tag{7-79}$$

依此，计算各点的 $\Delta A_k$，由数值方法得式(7-72) 的积分为：

$$\Delta A=-k_B T\sum_{i=1}^{k}\frac{\ln\langle\exp\{-[E(\lambda_i)-E(\lambda_i\pm\delta\lambda)]/k_B T\}\rangle_i}{\delta\lambda}\Delta\lambda \tag{7-80}$$

而自由能的变化为积分曲线所包含的面积如图 7-18 所示。

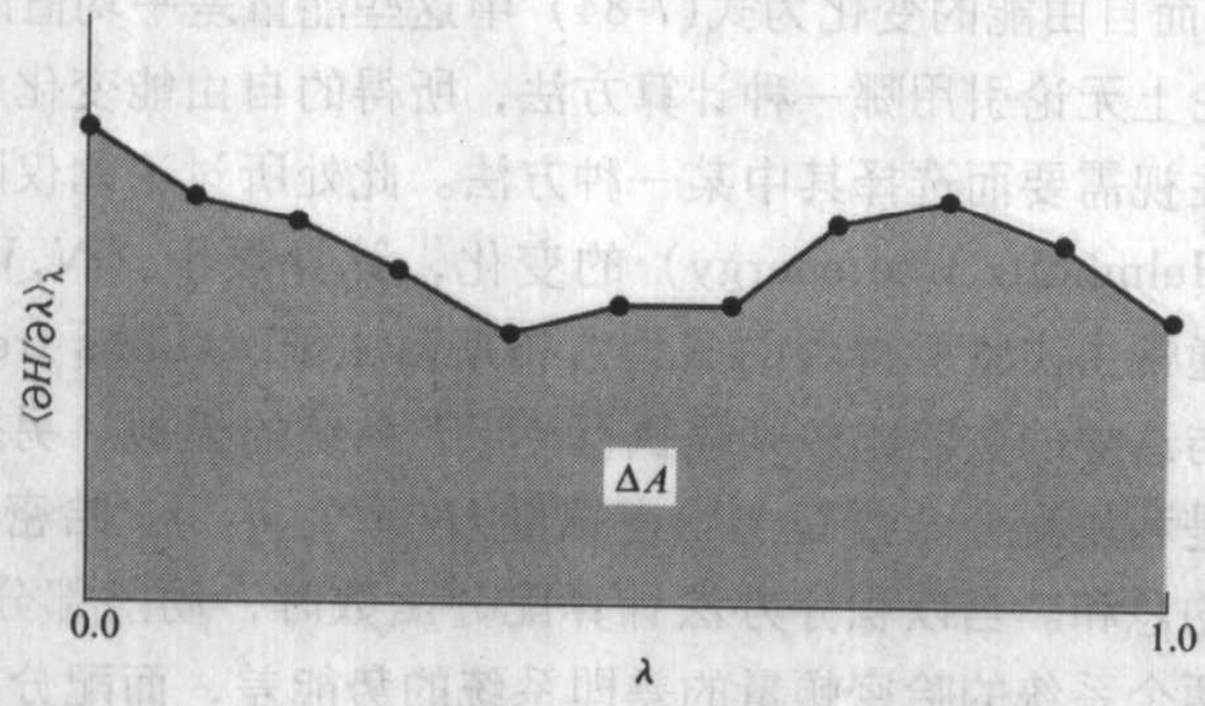



**图 7-18　热积分计算自由能**

通常在计算中，同时计算 $\lambda_i$ 的正向扰动 $\lambda_i+\delta\lambda$ 与负向扰动 $\lambda_i-\delta\lambda$。适当的选取扰动的幅度 $\delta\lambda$，应使由此两点计算出的 $\Delta A_i$ 值相同。或者在两点差异不大的情况下取其平均值。$\delta\lambda$ 的选取有两种常用的方法：①Trapezodial 等分法，即将 0→1 的区间分为相等的间隔，各点的 $\delta\lambda$ 均相等；②Gaussian-Legendre方法，选取适当的积分点，则各 $\delta\lambda$ 值不同。

## 7.5.5　缓慢成长法

除了上述的微扰方法与热力积分法以外，计算自由变化的第三种通用方法为缓慢成长法（slow growth method)。由微扰方法的式(7-65)，即：

$$\Delta A = -k_B T \sum_{i=0}^{N-1} \ln\langle \exp\{-[H(\lambda_{i+1}) - H(\lambda_i)]/k_B T\}\rangle \tag{7-81}$$

将式(7-81) 以泰勒级数展开，则自由能的变化可近似为：

$$\Delta A \approx -k_B T \sum_{i=0}^{N-1} \ln\langle 1 - [H(\lambda_{i+1}) - H(\lambda_i)]/k_B T + \cdots\rangle \tag{7-82}$$

式中，若 $H$ 的变化很小，则利用：

$$\ln(1+x) \approx x \tag{7-83}$$

将式(7-82) 转换为：

$$\begin{aligned}\Delta A &\approx -k_B T \sum_{i=0}^{N-1} \ln\left\{1 - \frac{\langle [H(\lambda_{i+1}) - H(\lambda_i)]\rangle}{k_B T} + \cdots\right\} \\ &\approx \sum_{i=0}^{N-1} \langle H(\lambda_{i+1}) - H(\lambda_i)\rangle\end{aligned} \tag{7-84}$$

这种计算方法将 $\lambda$ 值“缓慢”改变，每一步的 $H(\lambda_{i+1})$ 都非常接近 $H(\lambda_i)$。而自由能的变化为式(7-84) 中这些能量差平均值的总和。

理论上无论引用哪一种计算方法，所得的自由能变化均应相同。实际上，往往视需要而选择其中某一种方法。此处所讨论的仅限于亥姆霍兹自由能（Helmholtz free energy）的变化，计算限于（$N,V,T$）系统的模拟。但通常由实验所得到的多为吉布斯自由能（Gibbs free energy）的变化，若与这样的实验比较则需执行 $NPT$ 系统的模拟。另外，计算自由能的变化是利用配分函数 $Q$ 与哈密顿量 $H(\vec{p}^N, \vec{q}^N)$。哈密顿量为系统动能与势能的总和。当以积分方法计算配分函数时，动能部分可以互相消去。所以，两个系统的哈密顿量的差即系统的势能差，而配分函数可以以结构积分（configuration integral，$Z$）取代，即：

$$Z = \int d\vec{r} \exp[-U(\vec{r})/k_B T] \tag{7-85}$$

式中，$U(\vec{r})$ 为系统的势能。

## 7.6 相对自由能的变化

考虑如图 7-19 的乙烷与甲醇的溶剂化热力循环。

图 7-19 的循环中，各步骤自由能变化的关系为：

$$\Delta A_4 - \Delta A_3 = \Delta A_2 - \Delta A_1 \tag{7-86}$$

式(7-86) 的右边为甲醇与乙烷溶剂化自由能的差异，可由实验量出

$$\begin{array}{ccc} CH_3CH_3(g) & \xrightarrow{\Delta A_1} & CH_3CH_3(aq) \\ \Delta A_3 \downarrow & & \downarrow \Delta A_4 \\ CH_3OH(g) & \xrightarrow{\Delta A_2} & CH_3OH(aq) \end{array}$$

**图 7-19　$CH_3CH_3 \rightarrow CH_3OH$ 的热力循环**

此自由能的变化。上式的左边为在气态与溶液中乙烷转换为甲醇的自由能变化的差异，可利用分子动力计算得到。依照上节的 FDTI 方法，将 $\lambda=0$ 的状态定为乙烷，$\lambda=1$ 的状态定为甲醇，逐步计算其间各点自由能的变化，即可得到适当的结果。图 7-20 显示于不同的 $\lambda$ 值计算玻耳兹曼因子 $\exp(-\Delta E/k_B T)$ 的结果，此结果可用以检验自由能变化的计算是否收敛。

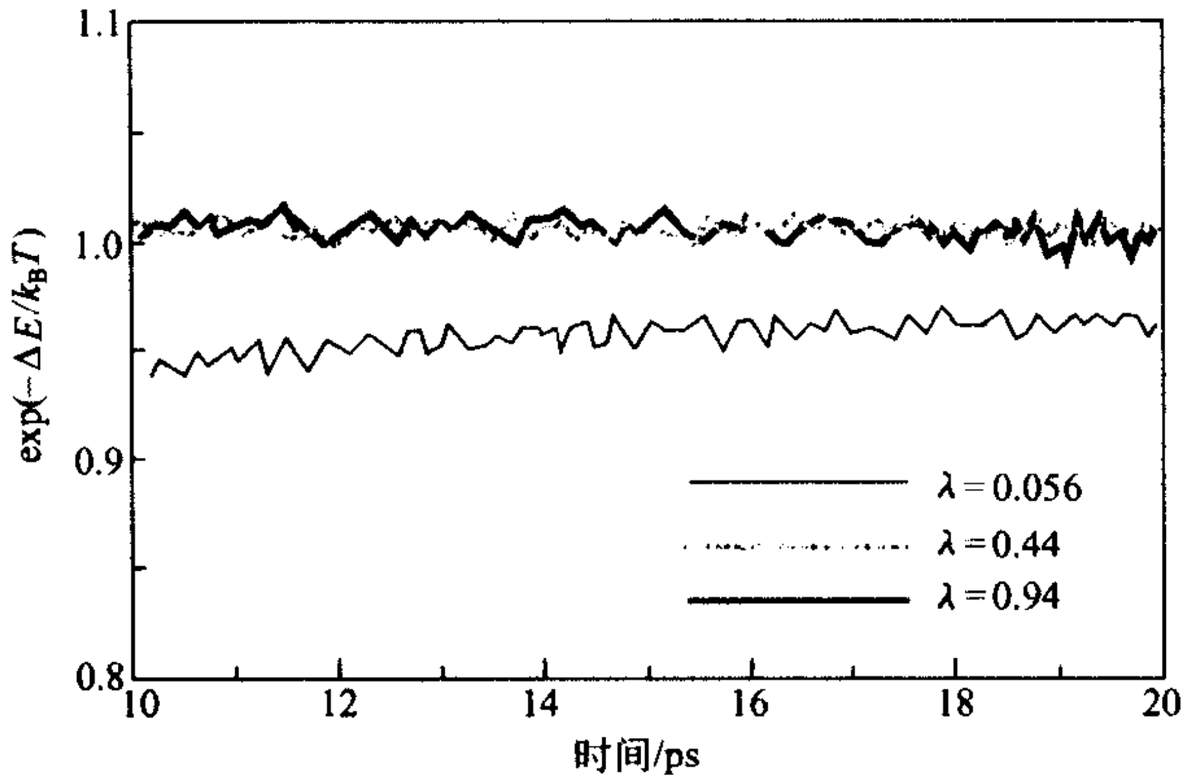



**图 7-20　分子动力计算玻耳兹曼因子 $\exp(-\Delta E/k_B T)$**

图中的每一点来自 10 段计算的平均，每段计算的时间为 100fs。图 7-20显示各 $\lambda$ 值计算的玻耳兹曼因子在 20ps 时均已收敛。Jorgensen 与 Ravimohan 利用此方法计算由乙烷转换为甲醇的自由能变化值为 $\Delta A_4 - \Delta A_3 =$ (6.75±0.2)kcal/mol，与实验度量的 $\Delta A_2 - \Delta A_1 = 6.93$kcal/mol 极为吻合。

## 7.7　构型自由能

除了化学变化以外，化合物本身亦可能由一种状态转换为另一种状态(如构象的变化)。计算化合物各构象的绝对自由能（absolute free energy），或称为构形自由能（conformational free energy），可了解转换的可能性，并

且不需要找出特殊的转换途径或过渡状态，亦可避免计算相对自由能的收敛问题。事实上，由于系统中原子间复杂的相互作用，不可能得到自由能的分析式。因此，仍需依照上节所提到的 FDTI 方法计算系统自由能的变化，即：

$$A_1 - A_0 = \int_0^1 \left\langle \frac{\partial H(p,r,\lambda)}{\partial \lambda} \right\rangle d\lambda \tag{7-87}$$

因此，自由能的变化可由计算哈密顿量微分的平均值计算。上式中，哈密顿量为耦合系数的函数，可将其写成：

$$H(\lambda) = \text{K. E}(\vec{p}) + (1-\lambda)U_0 + \lambda U_1 \tag{7-88}$$

式中，K. E($\vec{p}$) 为动能项；$U_0$ 与 $U_1$ 则分别为两个状态（构象）的势能。以此哈密顿量的表示式代入式(7-87)，得：

$$A_0 = A_1 - \int_0^1 \left\langle \frac{\partial}{\partial \lambda}[\text{K. E}(\vec{p}) + (1-\lambda)U_0 + \lambda U_1] \right\rangle d\lambda \tag{7-89}$$

因动能的平均值只与温度有关而与 $\lambda$ 的变化无关，$U_0$ 可设为参考结构的势能，为一固定的常数，通常选系统最稳定结构的势能为此参考点。依此将式(7-89) 化为：

$$A_0 = A_1 + \int_0^1 \langle U_0 - U_1 \rangle d\lambda \tag{7-90}$$

式(7-90) 提供了一个计算构型自由能的简便方法。利用许多特定 0→1 间 $\lambda$ 值计算的 $U_0 - U_1$ 的平均值可得到系统的自由能。而势能 $U_0 - U_1$ 的平均值为系统中所有结构的势能平均值。

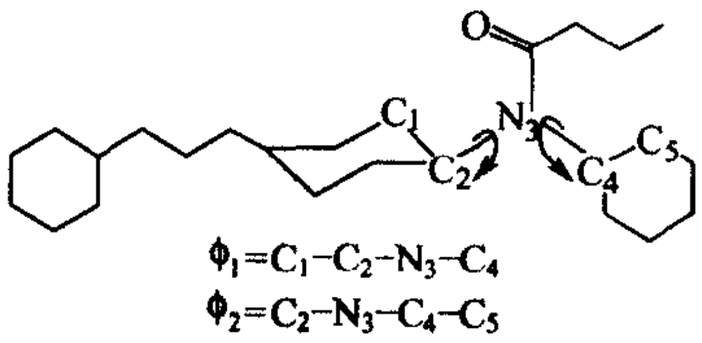



**图 7-21　fentayl 分子的结构**

以图 7-21 所示的 fentayl 分子为例，说明构型自由能的计算。此分子的构型差异主要来自图中二扭转角的旋转。

图 7-22 为 Tolleaneare 等利用能量最优化方法计算的 fentayl 分子的等高线势能图。图中的 a，b 两点为能量最小值点，其两点的能量差为 1kcal/mol，两点间的能量障碍约为 10kcal/mol；c 点为鞍点，对应于过渡态（transition state)。

图中的 c 点与最低能量点 a，b 的能量差即为由 a 或 b 的构型自由能。图 7-23 为以分子动力模拟方法所得到的 $\phi_1$ 与 $\phi_2$ 的运动轨迹。此计算选取的 $\lambda$ 值为 0.9，因所选的 $\lambda$ 值很大，故计算的角度均出现于能量最小区附近。倘若选取的 $\lambda$ 值很小，则 $\phi_1$ 与 $\phi_2$ 可能出现于其他的高能区。将此段计算的能量平均，可得对应于此 $\lambda$ 值的 $\langle U_0 - U_1 \rangle$，若选取各 $\lambda$ 值及重复分子动力计算，则可依式(7-90) 得到由 a 结构转换为 b 结构的自由能变化。

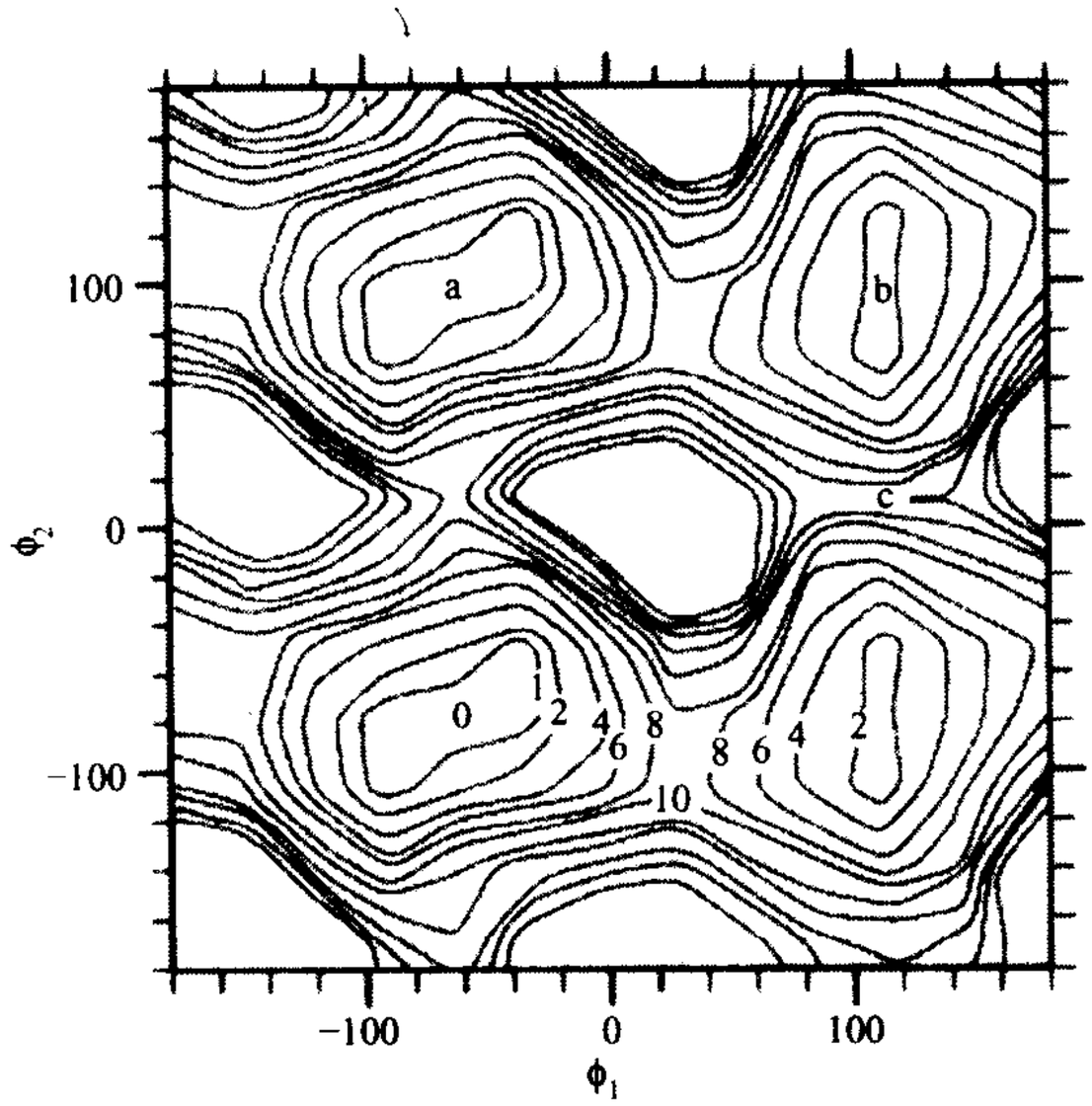



**图 7-22　fentayl 分子的等高线势能图**

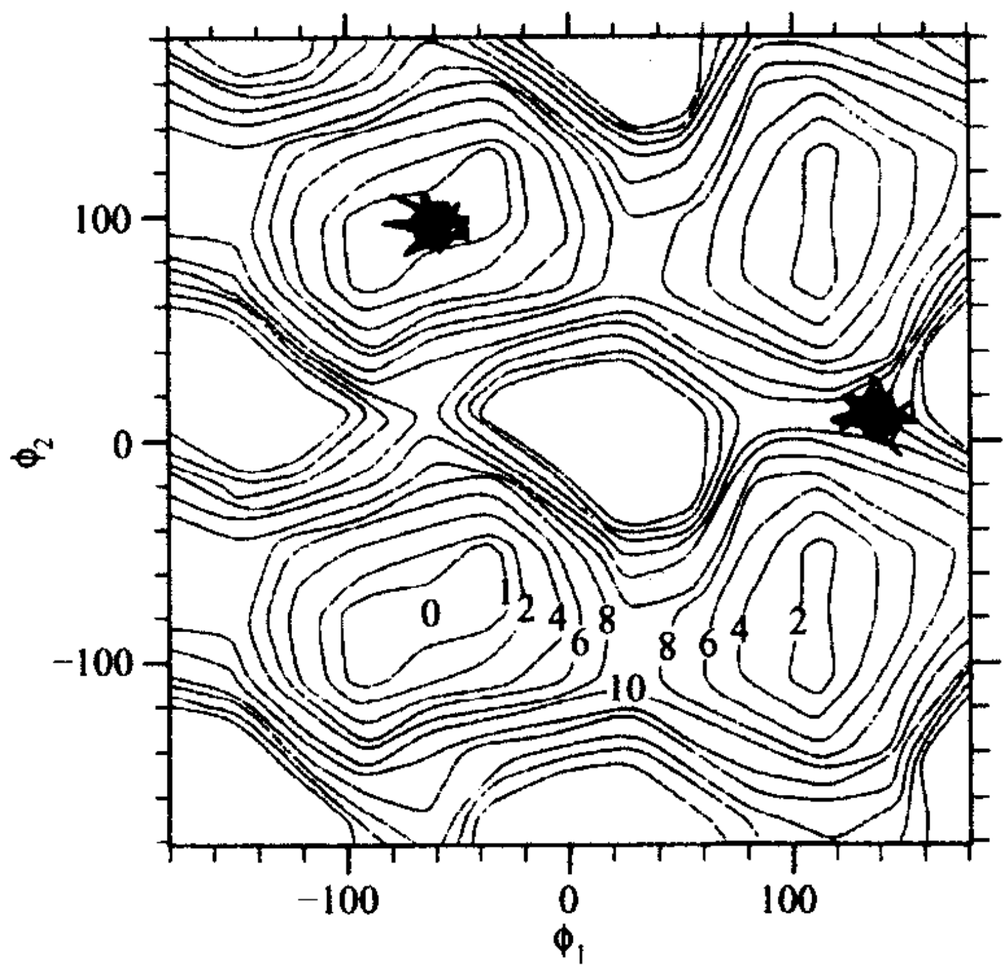



**图 7-23　分子动力模拟的 fentayl 分子运动轨迹**

# 参 考 文 献


1 Chen C L. Review Article, Chemistry, 2002

2 Chen H L, Lee C L, Shih J H, Chen C L. Comput. Polym. Sci., 4, 67, 1994

3 Jorgensen W L, Binning Jr R C, Bigot B. J. Am. Chem. Soc., 103, 4393, 1981

4 Landau L D, Liftshitz E M. "Statistical mechanics" (Course of Theoretical Physics, 5), Pergamon Press, Oxford, 1980

5 Lebowitz J L, Percus J K, Verlet L. Phys. Rev., 153, 250, 1967

6 Allen M P, Tildesley D J. Procedings of the NATO ASI on "Computer Simulation in Chemical Physics", Kluwer, Dordrecht, Holland, 1993

7 Anderson H C, Allen M P, Bellemans A, Board J, Clarke J H R, Ferrario M, Halle J M, Nosé S, Opheusden J V, Ryckaert J P. Rapport d'activité scientifique du CECAM, pp. 82, 1984

8 Berendsen H J C, Postma J P M, van Gunsteren W F, DiNola A, Haak J R, J. Chem. Phys., 81, 3684, 1984

9 Parrinello M, Rahman A. J. Appl. Phys., 52, 7182, 1981

10 Spohr E, et al. J. Phys. Chem., 92, 6754, 1988

11 Cerius 2 Softuare Package, Accelry Inc., San Diego, USA

12 Böhm H J, Lynden-Bell R M, Madden P A, McDonald I R. Mol. Phys., 51, 761, 1984

13 McQuarrie D A. "Statistical Mechanics", Harper and Row, New York, USA, 1976

14 Zwanzig R W. J. Chem. Phys., 22, 1420, 1954

15 Verlet L. Phys. Rev., 165, 201, 1958

16 King G, Warshel A. J. Chem. Phys., 91, 3647, 1989

17 Jorgensen W, Ravimohan C. J. Chem. Phys., 83, 3050, 1985

18 Tolleaneare, et al. "Forcefield-based Simulations" MSI, San Diego, USA, 1997

# 第8章 分子动力计算的应用实例

在第5、6章介绍原理的基础上，第7章介绍了获取参数的方法，本章举出几个实际体系的分子动力模拟例子，包括：最常见的溶剂-水分子体系，只有单纯的范德瓦耳斯力互相作用、存在多构象变化的烷类分子体系，正十六烷链状分子体系，链状双性体的分子体系和玻璃材质聚合物体系。这些体系具有重要的代表性，有关性质可以与实验对比，体现了分子模拟方法在实际工作中的应用。

## 8.1 水分子系统的计算

水分子为最常见的溶剂分子，许多重要的系统都含有水分子。因此，研究这些系统时，不得不考虑水分子对溶质的影响。在分子动力计算的发展过程中，水分子系统的计算是一个重要的关键。一则有充分的实验资料以检验计算的重要性；再则水分子本身具有多原子分子的运动模式，分子间又有特殊的氢键存在，若计算能成功地验证这些特性，即意味可将此研究推广至各种系统。目前，已成功地开发出一些水分子的力场，并广泛地应用于各种化学与生化领域的研究上。

### 8.1.1 简单的水分子的力场

为了研究水的特殊性质，针对水分子的特性，发展出一些水的力场。常用的水分子力场约有下列几种：①中心力力场（central force，CF）；②可转移的分子间势能（transferable intermolecular potential，TIP）；③简单点电荷力场（simple point charge，SPC）。

**(1) 中心力力场** 中心力力场为早期由 Stillinger 等所提出的。其形式为：

$$U_{HH}(r)=\frac{36.1345}{r}+\frac{18}{1+\exp[40(r-2.05)]}-17\exp[-7.61277(r-1.45251)^2]$$

$$U_{OH}(r)=-\frac{72.269}{r}+\frac{6.23403}{r^{9.19912}}-\frac{10}{1+\exp[40(r-1.05)]}-$$

$$\frac{4}{1+\exp[5.49305(r-2.2)]}$$

$$U_{OO}(r)=\frac{144.538}{r}+\frac{26758.2}{r^{8.8591}}-0.25\exp[-4(r-3.4)^2]-$$

$$0.25\exp[-1.5(r-4.5)^2] \qquad (8\text{-}1)$$

式中，力场的单位为 kcal/mol；距离的单位为 Å；$r$ 为分子内或分子间原子对的距离。水分子本身为刚性分子。图 8-1 为 Stillinger 与 Rahman 引用 CF 力场，对密度为 1.0g/cm³ 含 216 个水分子系统在 29.5℃的条件下执行分子动力计算所得氢-氢原子对的径向分布函数的结果。图中，在氢-氢距离为 1.5Å 处有最大的峰，此距离相当于水分子内氢-氢原子间的距离。在氢-氢距离为 2.25Å 处有另一明显的峰，此为水分子形成氢键时氢-氢原子的距离。此与 X 射线衍射实验所观测的结果相吻合。

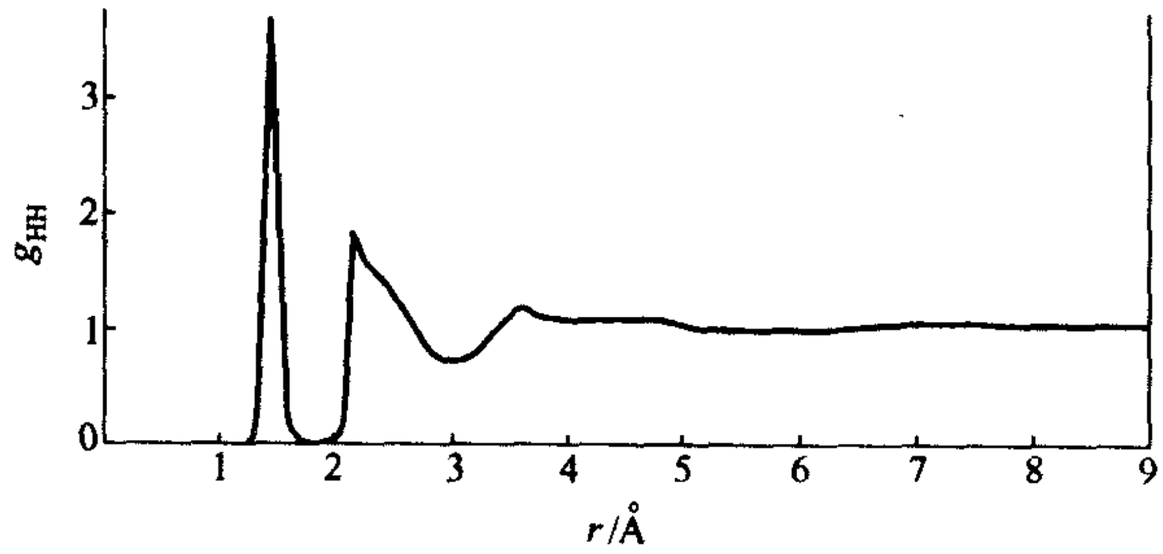



**图 8-1 应用 CF 力场计算的水分子系统氢-氢原子径向分布函数**

**(2) 可转移的分子间势能** 可转移的分子间势能（transferable intermolecular potential，TIP）的水分子力场是 Jorgensen 利用蒙地卡罗计算方法与精确的量子力学计算结果相比较所推导出的。除各种热力学的性质外，最重要的是比较水分子二聚体（dimmer）的结构。此力场将水分子视为刚体分子，即其键长键角的值恒定。此种形式的力场因其形式简单，并且不用考虑分子中各种高频率的振动，因此常与其他力场合并应用于生化系统的计算中。TIP 力场的一般形式为：

$$U_{AB} = \sum_{a}^{A} \sum_{b}^{B} \left( \frac{q_a q_b e^2}{r_{ab}} + \frac{A_a A_b}{r_{ab}^{12}} - \frac{C_a C_b}{r_{ab}^{6}} \right) \tag{8-2}$$

式中，$a$，$b$ 为不同水分子上的原子。$q_O = -0.80$，$q_H = 0.40$ 分别为水分子中氧原子与氢原子所带的部分电荷。式中参数 $A_O^2 = 5.8 \times 10^{-1}$ kcal/mol，$A_H = 0.0$；$C_O^2 = 5.25 \times 10^2$ kcal/mol，$C_H = 0.0$。为了使计算的径向分布函数与实验一致，Jorgensen 将原先 TIPS 力场中水分子的 3 个作用点改为 4 个作用点（TIPS2）。将第 4 作用点定为水分子的质心（$M$），而力场的形式与原先的 TIPS 一致。TIPS2 力场的各参数为：$q_O = 0.0$，$q_H = 0.535$，$q_M = -1.070$；$A_O^2 = 6.95 \times 10^{-3}$ kcal · Å$^{12}$/mol，$A_H = A_M = 0.0$，$C_O^2 = 600$kcal · Å$^6$/mol，$C_H = C_M = 0.0$。利用此力场计算水线型二聚体的最佳结构能为 $-6.20$kcal/mol，近似于实验值（$-5.44 \pm 0.7$）kcal/mol。继此之后，Jorgensen 等陆续发展出 3 作用点的 TIP3P 与 4 作用点的 TIP4P 力场。图 8-2 为由 TIPS2 计算 O-O 的径向分布函数与 X 射线衍射实验的比较。

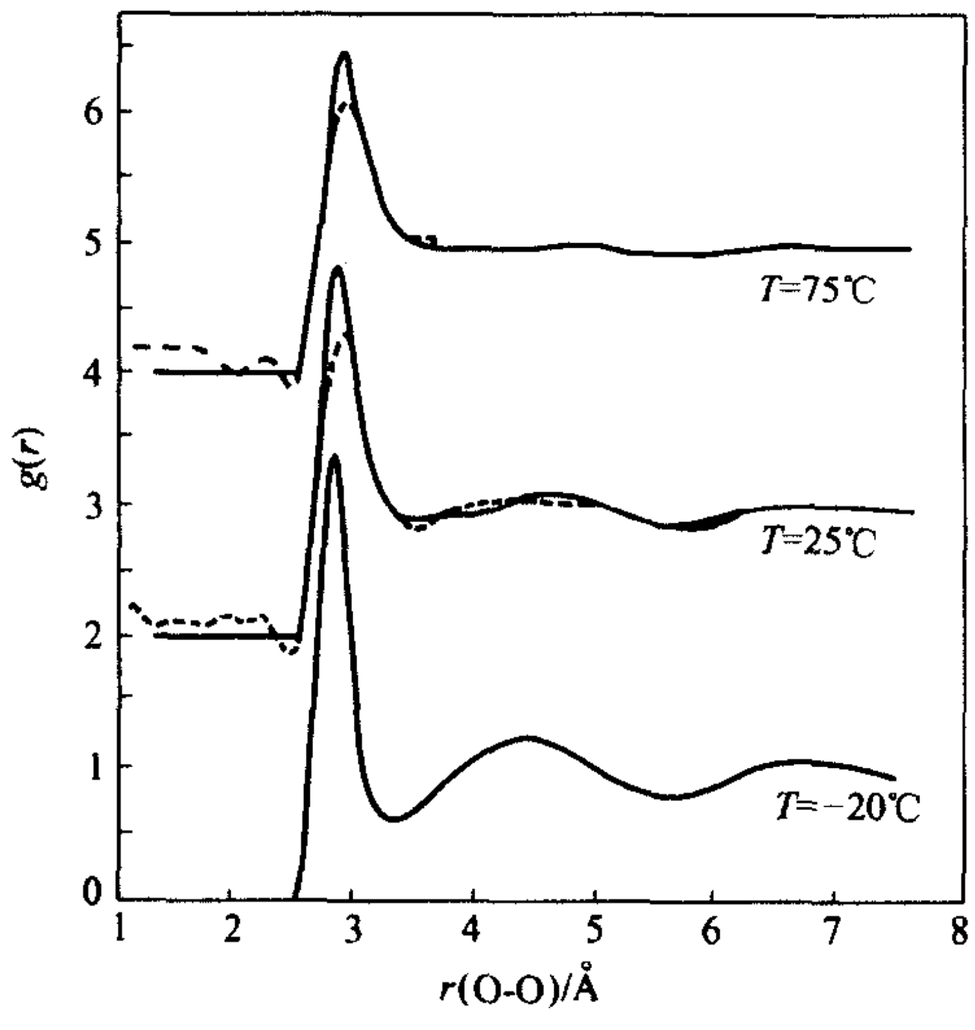



**图 8-2　利用 TIPS2 计算所得各温度下水的 O-O 径向分布函数**

图中实线为由 TIPS2 力场计算所得，虚线为 X 射线衍射实验所得

**（3）简单点电荷力场**　简单点电荷力场（simple point charge，SPC）为 Berendsen 等所提出的。此力场亦将水分子视为刚体分子，其键长与键角的值固定。此力场将二水分子的作用分为范德瓦耳斯作用与库仑

作用。仅氧-氧原子间有范德瓦耳斯作用。水分子的每个原子均带有电荷，分子间原子对有库仑静电作用。SPC 力场的形式为：

$$U_{OO}(r_{OO})=\frac{q_O q_O e^2}{r_{OO}}+\frac{A}{r_{OO}^{12}}-\frac{C}{r_{OO}^{6}}$$

$$U_{OH}(r_{OH})=\frac{q_O q_H e^2}{r_{OH}}$$

$$U_{HH}(r_{HH})=\frac{q_H q_H e^2}{r_{HH}} \tag{8-3}$$

表 8-1 列出 SPC 与 TIPS 力场的各种参数。利用 SPC 力场及分子动力计算可得到合理的水分子径向分布函数与各种热力学性质。但以此力场计算的扩散系数则与实验值有较大的偏差（误差为 200%～300%）。各种水分子的力场中应用最广的为 TIP3P，尤其是一些生化系统的计算，如 AMBER、CHARM 等均引用此力场计算水分子溶剂的运动。

**表 8-1　水分子各力场的分子几何结构及力场参数**

| 项　　目 | SPC | TIP3P | TIPS2 | TIP4P |
|---|---|---|---|---|
| $r$(OH)/Å | 1.0 | 0.9572 | 0.9572 | 0.9572 |
| $\angle HOH$/(°) | 109.47 | 104.52 | 104.52 | 104.52 |
| $A\times10^{-3}$/(Å$^{12}$ · kcal/mol) | 629.4 | 582.0 | 695.0 | 600.0 |
| $C$/(Å$^6$ · kcal/mol) | 625.5 | 595.0 | 600.0 | 610.0 |
| $q_O$ | −0.82 | −0.834 | 0.0 | 0.0 |
| $q_H$ | 0.41 | 0.417 | 0.535 | 0.52 |
| $q_M$ | 0.0 | 0.0 | −1.07 | −1.04 |
| $r$(OM)/Å | 0.0 | 0.0 | 0.15 | 0.15 |

## 8.1.2　ab initio 水分子力场

Nieser、Corongiu 与 Clementi 等于1990年利用精确的量子力学 ab initio 的方法，发展出适用于液体与二聚集体及三聚集体等团簇体的水分子力场，称为 NCC 力场。此力场与 TIP4P 力场一样，氢原子带负电荷，另一带正电荷的点 $P$ 位于 HOH 键角的平分线上。各原子的定义如图 8-3。

其力场的形式为：

$$U=q^2\left(\frac{1}{R_{13}}+\frac{1}{R_{14}}+\frac{1}{R_{23}}+\frac{1}{R_{34}}\right)$$

$$+\frac{4q^2}{R_{78}}-2q^2\left(\frac{1}{R_{81}}+\frac{1}{R_{82}}+\frac{1}{R_{73}}+\frac{1}{R_{74}}\right)$$

$$+A_{OO}e^{-B_{OO}R_{56}}+A_{HH}(e^{-B_{HH}R_{13}}+e^{-B_{HH}R_{14}}+e^{-B_{HH}R_{23}}+e^{-B_{HH}R_{24}})$$

$$+A_{OH}(e^{-B_{OH}R_{53}}+e^{-B_{OH}R_{54}}+e^{-B_{OH}R_{61}}+e^{-B_{OH}R_{62}})$$
$$+A'_{OH}(e^{-B'_{OH}R_{53}}+e^{-B'_{OH}R_{54}}+e^{-B'_{OH}R_{61}}+e^{-B'_{OH}R_{62}})$$
$$+A_{PH}(e^{-B_{PH}R_{73}}+e^{-B_{PH}R_{74}}+e^{-B_{PH}R_{81}}+e^{-B_{PH}R_{82}})$$
$$+A_{PO}(e^{-B_{PO}R_{76}}+e^{-B_{PO}R_{85}}) \tag{8-4}$$

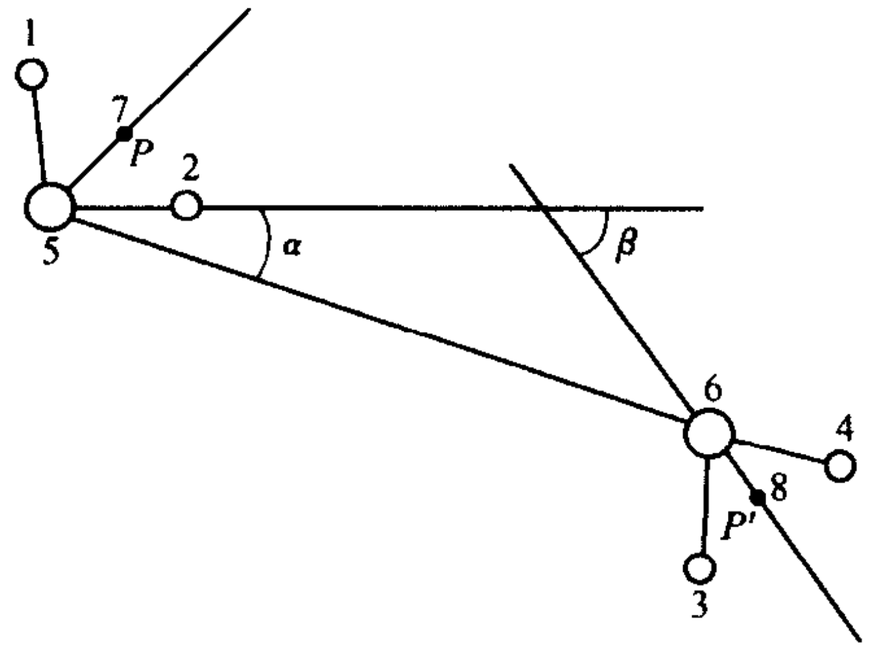



**图 8-3　NCC 力场各参数的定义**

式中，7，8 两点即 $P$ 与 $P'$点。利用 NCC 力场可准确的计算液体水及团簇体的各种性质。Corongiu 改进 NCC 力场的刚体水分子模型，加入水的内振动项：

$$U_{intra}=\frac{1}{2}f_{RR}(\delta R_1^2+\delta R_2^2)+\frac{1}{2}f_{\theta\theta}(\delta\rho^2)+f_{RR'}\delta R_1\delta R_2+f_{R\theta}(\delta R_1+\delta R_2)\delta\rho$$
$$+\frac{1}{R_e}[f_{RRR}(\delta R_1^3+\delta R_2^3)+f_{\theta\theta\theta}\delta\rho^3+f_{RRR'}(\delta R_1+\delta R_2)\delta R_1\delta R_2$$
$$+f_{RR\theta}(\delta R_1^2+\delta R_2^2)\delta\rho+f_{RR'\theta}\delta R_1\delta R_2\delta\rho+f_{R\theta\theta}(\delta R_1+\delta R_2)\delta\rho^2]$$
$$+\frac{1}{R_e^2}[f_{RRRR}(\delta R_1^4+\delta R_2^4)+f_{\theta\theta\theta\theta}\delta\rho^4+f_{RRRR'}(\delta R_1^2+\delta R_2^2)\delta R_1\delta R_2$$
$$+f_{RRR'R'}\delta R_1^2\delta R_2^2+f_{RRR\theta}(\delta R_1^3+\delta R_2^3)\delta\rho]$$
$$+\frac{1}{R_e^2}[f_{RRR'\theta}(\delta R_1+\delta R_2)\delta R_1\delta R_2\delta\rho+f_{RR\theta\theta}(\delta R_1^2+\delta R_2^2)\delta\rho^2$$
$$+f_{RR'\theta\theta}\delta R_1\delta R_2\delta\rho^2+f_{R\theta\theta\theta}(\delta R_1+\delta R_2)\delta\rho^3] \tag{8-5}$$

式中，$\delta R_1=R_1-R_e$，$\delta R_2=R_2-R_e$，$\delta\rho=R_e(\theta-\theta_e)$。应用此力场可准确地计算水的振动光谱。

### 8.1.3　水分子自由能的计算

超额自由能（excess free energy）的定义为真实溶液的自由能与理想

溶液的自由能之差。依照第 7 章所讨论的计算自由能的方法，系统的超额自由能可由式(8-6) 计算。

$$\Delta A = A(\lambda = 1) - A(\lambda = 0) = \int_0^1 \mathrm{d}\lambda \langle \partial U / \partial \lambda \rangle \tag{8-6}$$

式中，$U$ 为系统的势能；$\lambda$ 为耦合参数。耦合参数的意义为：$\lambda=1$ 对应于液体分子在全力场作用的状态；$\lambda=0$ 对应于液体的理想状态，即分子间无任何作用存在的状态。依此式，欲计算 $\Delta A$，需计算 $\langle \partial U/\partial \lambda \rangle$ 的积分。通常可将此积分分成 $N_{\mathrm{sim}}$ 个相等的区间 ($10 \leqslant N_{\mathrm{sim}} \leqslant 20$)，对应于 $\lambda$ 由 0～1 的变化。分别求得对应于每一 $\lambda$ 值的 $\langle \partial U/\partial \lambda \rangle$ 项，以计算 $\Delta A$。

$$\Delta A \approx \sum_{i=1}^{N_{\mathrm{sim}}} \langle \partial U / \partial \lambda \rangle_i \Delta \lambda \tag{8-7}$$

式中，$\Delta\lambda = 1/N_{\mathrm{sim}}$。

此例以简单点电荷力场（SPC）计算以水分子系统的超额自由能。Quintana 等将 SPC 力场化为如下的形式：

$$u_{\mathrm{OO}}(r) = \lambda^5 \frac{B}{r^{12}} + \lambda^3 \left( -\frac{A}{r^6} + \frac{C_1}{r} \right)$$

$$u_{\mathrm{OH}}(r) = \lambda^3 \frac{C_2}{r}$$

$$u_{\mathrm{HH}}(r) = \lambda^3 \frac{C_3}{r} \tag{8-8}$$

式中，$A$、$B$、$C$ 为常数。

Quintana 等所执行的分子动力模拟的系统含 108 个水分子，密度为 $1\mathrm{g/cm^3}$，温度为 298K 与 373K。计算的积分步长为 $\Delta t = 0.001\mathrm{ps}$。先执行各系统于 $\lambda=1$ 条件下的计算。当其达平衡后，逐次降低 $\lambda$ 值至 0。每次降低的幅度为 $\Delta\lambda=0.10$。对应于每一个 $\lambda$ 值，执行分子动力计算使系统达于平衡。再执行 2ps 的分子动力计算以求取 $\langle \partial U/\partial \lambda \rangle$ 的平均值。表 8-2 列出在不同温度与不同截断半径下计算所得的超额自由能与实验值的比较。由表中可知，SPC 力场计算的结果与实验值十分接近。

**表 8-2　SPC 力场计算水分子系统的超额自由能与实验值的比较**

| 温度/K | 截断半径/Å | $\Delta A$ 计算值 /(kcal/mol) | $\Delta A$ 实验值 /(kcal/mol) |
|---|---|---|---|
| 298 | 7 | 5.64 | 5.74 |
| 298 | 6 | 5.51 | 5.74 |
| 373 | 6 | 4.61 | 4.74 |

## 8.2 正戊烷构象的研究

聚合物分子的许多构象主要因其单键所对应的二面角旋转所造成的。研究聚合物分子的构象，可将每一个二面角由 0°～360°转动一定的量，而得到不同的构象。若一分子含有 $n$ 个二面角，则其主要的构象为这些二面角的函数，$\{\phi_1, \phi_2, \cdots, \phi_n\}$。若每一个二面角每次转动的量为 $\Delta\phi$，则由转动所产生的构象共有 $(360°/\Delta\phi)^n$ 个。依照分子力学，每一种构象必须将其附近最优化的结构定出（如最佳键长、最佳键角等），以决定此构象所对应的能量。以某含有 10 个二面角的烷类分子为例，若每一个二面角每次转动的量为 60°，则需求 $6^{10}$～$6\times10^7$ 个构象的最优化结构。假设求最优化结构每次需要 30s，则完成此计算需要超过 60 年的时间。因此，这样的研究并不实际，必须找出其他的方法。

旋转正戊烷的两根单键所对应二面角可产生许多构象。若仅考虑每次转动量为 120°的情形，共可产生 $3^2=9$ 种构象。定义这些构形二面角 180°为反式（t），60°为正邻位交叉式［g(＋)］，－60°为负邻位交叉式［g(－)］。则此 9 种构象如表 8-3 所示。

**表 8-3　正戊烷的 9 种构象**

| 序号 | 二面角 | | 构　象 |
|---|---|---|---|
| | C1-C2-C3-C4 | C2-C3-C4-C5 | |
| 1 | 180 | 180 | tt |
| 2 | 60 | 180 | g(＋)t |
| 3 | 180 | 60 | tg(＋) |
| 4 | －60 | 180 | g(－)t |
| 5 | 180 | －60 | tg(－) |
| 6 | 60 | 60 | g(＋)g(＋) |
| 7 | －60 | －60 | g(－)g(－) |
| 8 | 60 | －60 | g(＋)g(－) |
| 9 | －60 | 60 | g(－)g(＋) |

表 8-3 所列正戊烷的构象，其中有几种为等效构象（equivalent conformation），(2≡3，4≡5，8≡9)。其余的为对映构象（enantiomer），(2/4,3/5,6/7)。因此这 9 种构象事实上仅有 4 种不同的情形，其中有些为外消旋对，有些具有 $C_2$ 对称性（对称数 $\sigma=2$），如表 8-4 所示。

表 8-4　正戊烷的 4 种不同构象

| 标　识 | 构　象 | 二面角 $\sigma$ | $d,l$ 对 |
|---|---|---|---|
| C1 | tt | (180,180)2 | |
| C2 | g(+)t,tg(+) | (±60,180)1 | 存在 |
| C3 | g(+)g(+),g(−)g(−) | (±60,±60)2 | 存在 |
| C4 | g(+)g(−),g(−)g(+) | (±60,∓60)1 | |

此 4 种不同的构象可利用分子力学的方法，求取其附近的最优化结构，即能量最小值的结构，称为构象异构体。Mencarelli 引用 MM+力场，并分别以最速下降法及对角线块状牛顿-拉森法计算各构象的最优化结构。所计算的结果如表 8-5。

表 8-5　正戊烷各构象的最优化结果（$E_s$ 为相对于 C1 的构象能）

| 标识 | 构象异构体 | $E_s$ /(kcal/mol) | 二面角 /(°) | 键角 /(°) | 键长 /Å |
|---|---|---|---|---|---|
| C1 | tt | 0.00 | 180.0/ 180.0 | 111.8/112.0/111.8 | 1.535/1.537/1.537 /1.535 |
| C2 | g(+)t,tg(+) | 0.897 | ±65.4/ ±175.6 | 113.6/113.7/111.8 | 1.534/1.538/1.537 /1.535 |
| C3 | g(+)g(+),g(−)g(−) | 1.623 | ±59.6/ ±59.6 | 113.6/115.5/113.6 | 1.534/1.538/1.538 /1.534 |
| C4 | g(+)g(−),g(−)g(+) | 3.216 | ±94.8/ ∓62.6 | 113.7/115.1/114.5 | 1.536/1.539/1.139 /1.533 |

图 8-4 为正戊烷各种构象异构体的结构图，其中，C2，C3，C4 为外消旋对（racemic pair）。

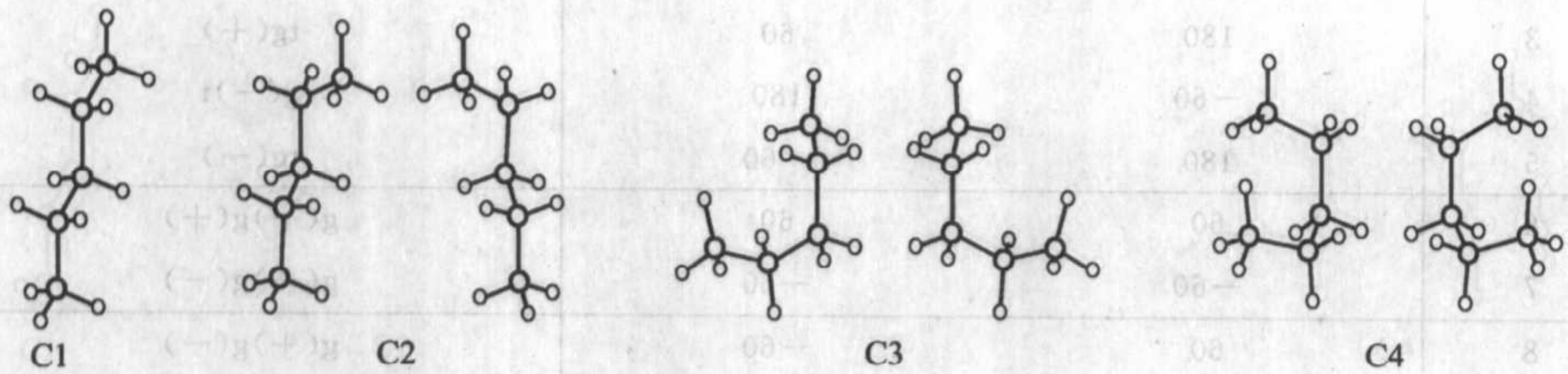



图 8-4　分子力学所计算的正戊烷的构象异构体

分子动力学计算产生分子运动的轨迹。任一时间分子在空间的坐标相当于分子的一个构象。因此可仿照分子力学，由求取能量的最小值方法找出该构象附近的构象异构体。若计算的时间够长，且温度够高，则分子的运动可能翻越所有的势能障而产生所有对应于势能极小值的构象异构体。

Mencarelli 应用分子动力计算求取正戊烷的构形异构体的分布。计算的条件为 300K 与 600K，时间间隔为 0.5fs（femto second，fs）。当分子系统平衡后，执行 500ps（pico second，ps）的分子动力计算，并储存每皮秒的分子运动轨迹。将此储存的分子运动轨迹视为分子的一种构象，以能量最小化法求得构象异构体的结构与能量。依照玻耳兹曼分布定律，平衡系统中分子任一构象异构体出现的概率为：

$$x_i = \frac{e^{-(G_i/RT)}}{\sum_i e^{-(G_i/RT)}} \tag{8-9}$$

式中，$G_i = H_i - TS_i$ 为吉布斯自由能。由分子力学知 $H_i$ 约等于构型能 $E_s$。分子的熵则可由其对称数（symmetry number）$\sigma$，算出：

$$S_i = -R\ln(\sigma_i) \tag{8-10}$$

若考虑外消旋对的因素，则：

$$S_i = -R\ln\left(\frac{\sigma_i}{2}\right) \tag{8-11}$$

图 8-5 为两种温度下，计算的构象异构体的概率与实验的比较。

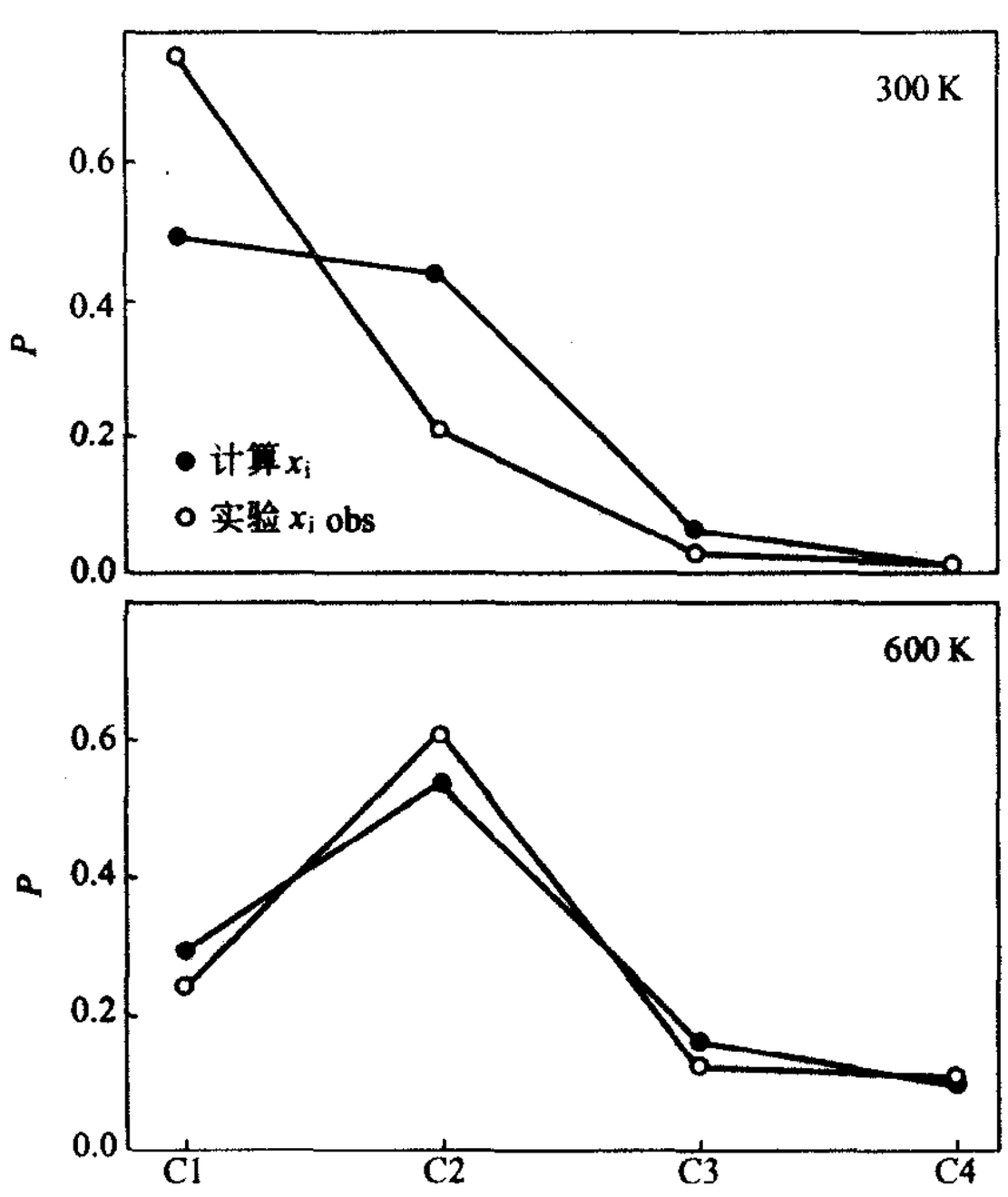



**图 8-5　分子动力计算正戊烷的构象异构体之概率 P**

分析分子动力计算的结果，应可得到所有先前由分子力学找出的构象异构体。由图 8-5 知，当计算温度为 600K 时，构象异构物的出现概率与实验值相近。因此该温度的分子动力计算已可成功地探测正戊烷的所有构象异构物。但在较低的温度（300K）条件下，趋势虽然一致但分布有些不同。这表示分子动力计算的时间不够长，并未涵盖所有的构象空间。表 8-6 列出分子动力计算的各种结果。

**表 8-6　正戊烷分子动力计算的结果**（$G_i$ 为相对于 C1 的值）

| 标识 | 构象异构体 | $\sigma$ | $d,l$ | $G_i^{300}$ | $G_i^{600}$ | $x_{i,calc}^{300}$ | $x_{i,exp}^{300}$ | $x_{i,calc}^{600}$ | $x_{i,exp}^{600}$ |
|---|---|---|---|---|---|---|---|---|---|
| C1 | tt | 2 | | 0.000 | 0.000 | 0.491 | 0.752 | 0.273 | 0.222 |
| C2 | g(+)t,tg(+) | 1 | 存在 | 0.071 | −0.756 | 0.436 | 0.208 | 0.514 | 0.586 |
| C3 | g(+)g(+),g(−)g(−) | 2 | 存在 | 1.210 | 0.797 | 0.064 | 0.028 | 0.140 | 0.104 |
| C4 | g(+)g(−),g(−)g(+) | 1 | 存在 | 2.390 | 1.563 | 0.009 | 0.012 | 0.073 | 0.088 |

## 8.3　链状分子的计算

最常见的链状分子为烷类分子。含有不同甲基数目的烷类，其性质有非常大的差异。一般的油品分子多为含碳数目大于 10 的烷类。在油品分子系统中掺入其他的分子可以改变其性能。图 8-6 为利用 MATERIALS

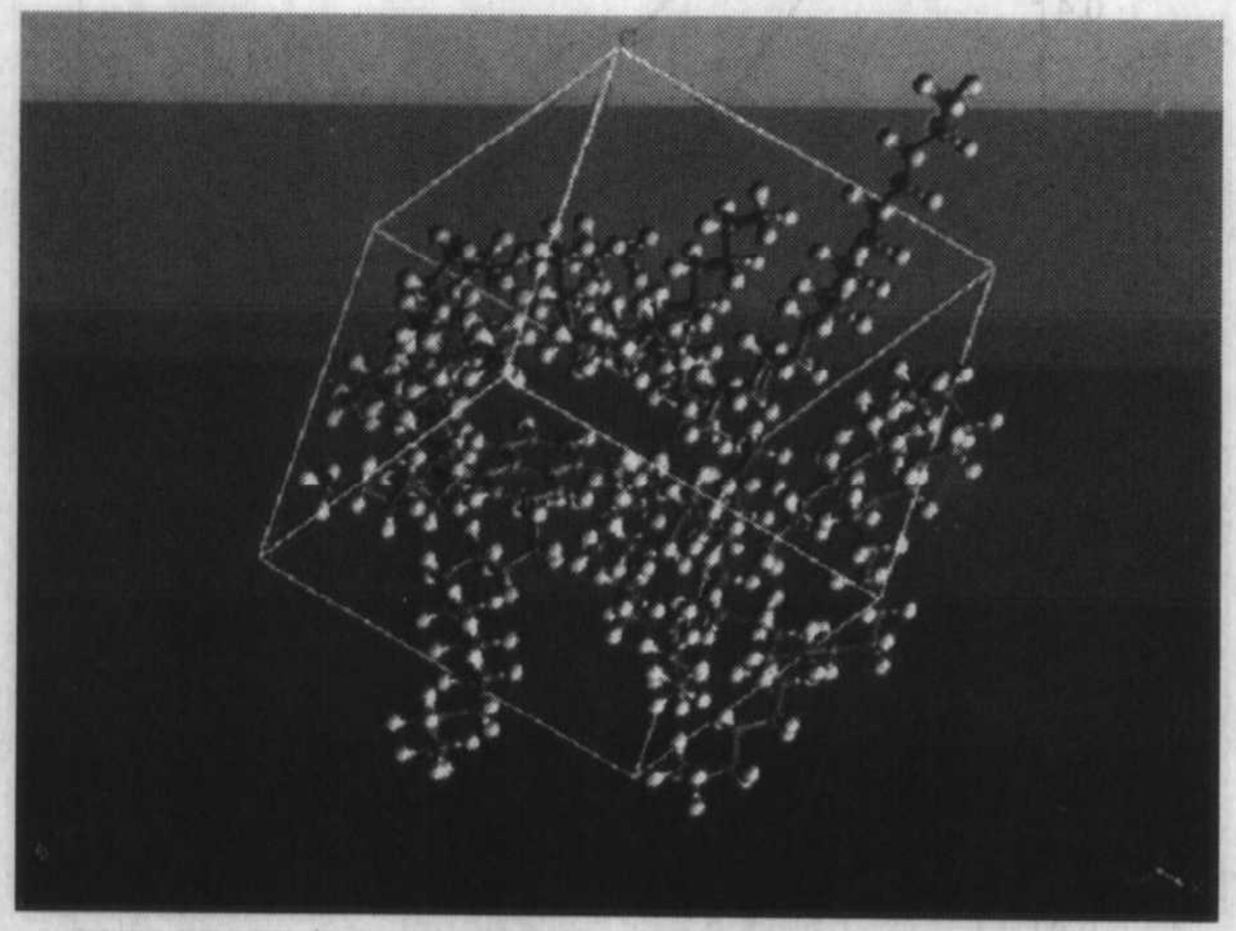


**图 8-6　正十六烷与 α-甲基萘的混合体系**

STUDIO 商业软件所计算的链状液态正十六烷（*n*-hexadecane）分子与添加剂 α-甲基萘（α-methyl naphthalene，NPH）的系统。正十六烷分子为柴油的主要成分。柴油系统中加入 α-烷基萘可增加流动性或降低固化点（solidify point）。这两种性质在油品工业的实际应用上非常重要。

图 8-7 为 Ying 等所计算的正十六烷与 α-烷基萘混合系统的热容。热容的计算方法为式(7-5)，其单位为 kJ/(mol · K)。图 8-7 中的曲线均呈现明显的转折点，这种转折点间的区间代表系统的相转换点（phase-transition point）。图中的虚线为实验所测得的正十六烷及与 α-正十二烷基萘的混合

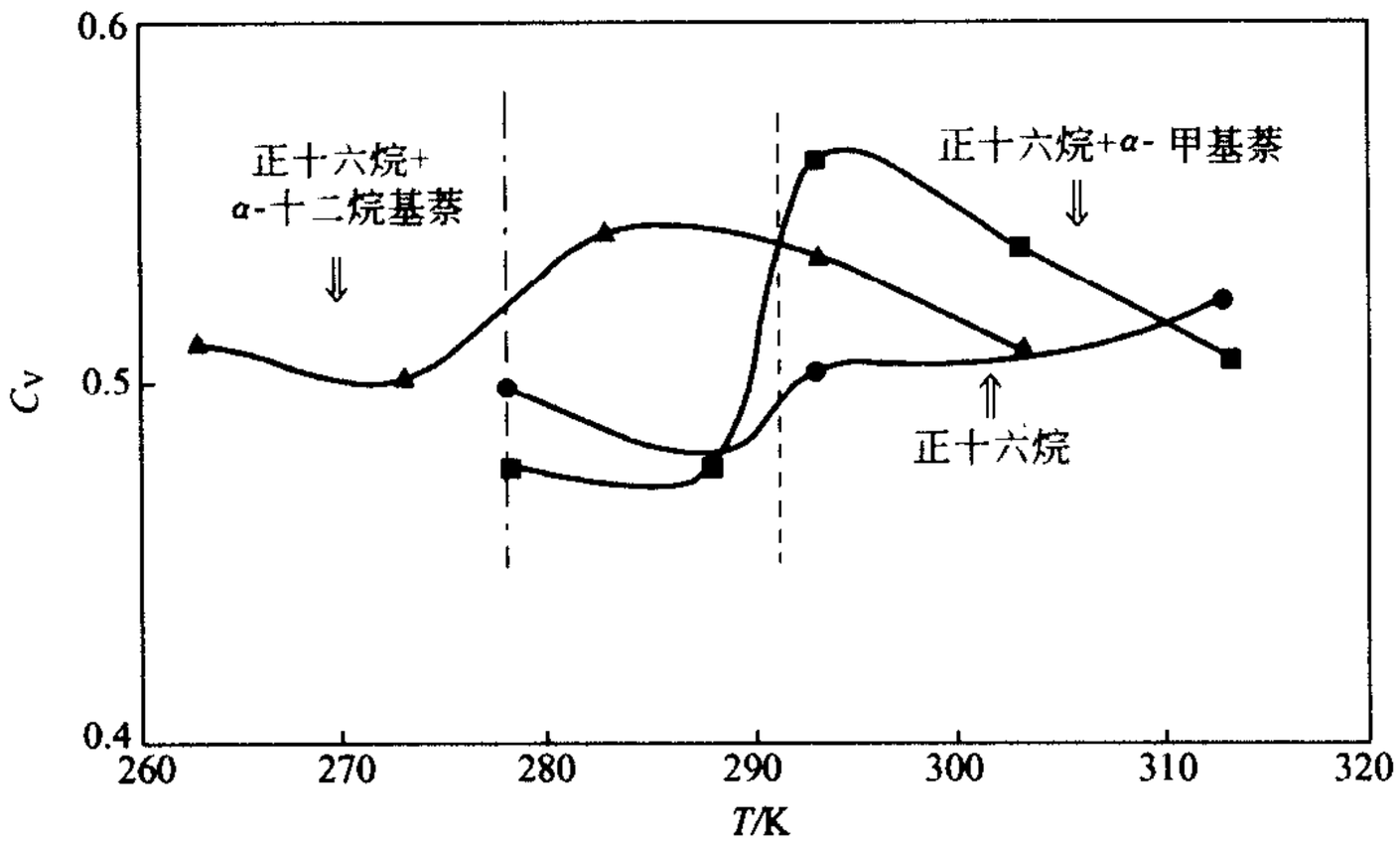



**图 8-7　计算的 α-烷基萘与正十六烷系统的定容比热容**

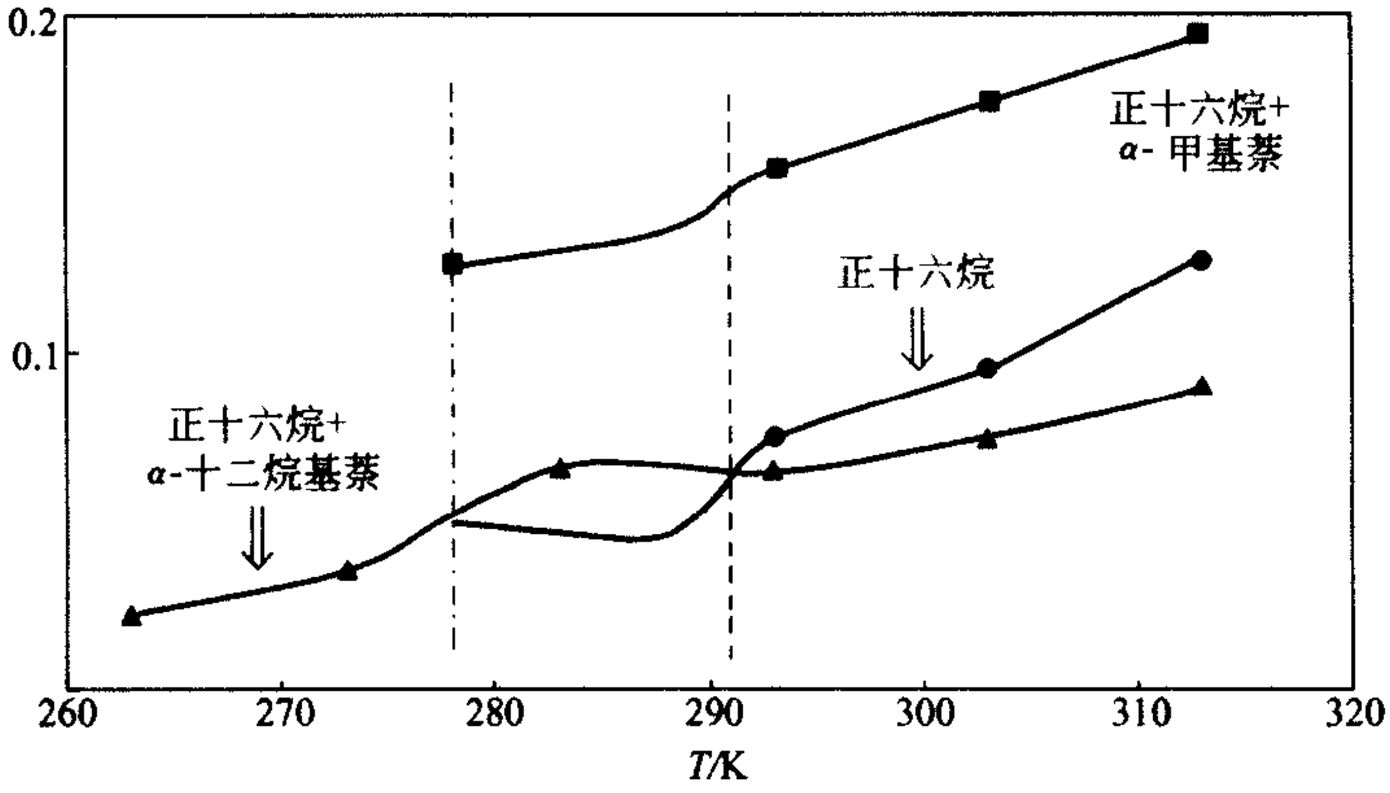


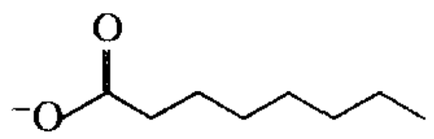

**图 8-8　由 MATERIALS STUDIO 计算的 α-烷基萘与正十六烷系统的原子扩散系数**

体系的固化点。显示计算与实验的一致性。

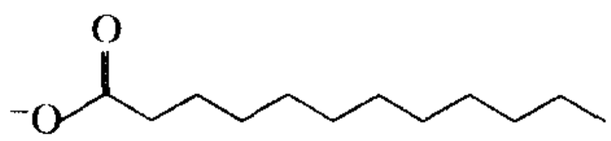
图 8-8 为利用分子动力模拟计算的均方位移所得到的扩散系数 $D$。图中显示，系统中加入 α-甲基萘后，扩散系数较正十六烷增加约 3 倍。若系统中加入 α-正十二烷基萘（α-dodecyl NPH），则固化点降低约 18K。

## 8.4 链状双性体的分子动力计算

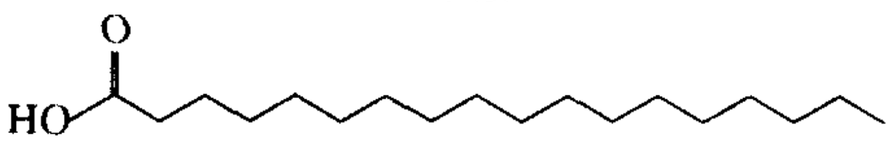
链状双性体（amphiphiles）分子为一根或数根碳氢链连上头端的极性官能团所形成的分子。这种分子常见于生物体与材料，图 8-9 为一些典型的链状双性体分子。



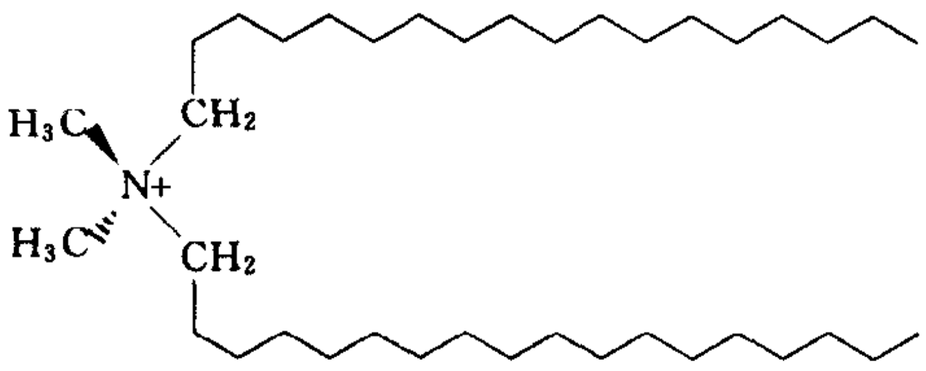
辛酸



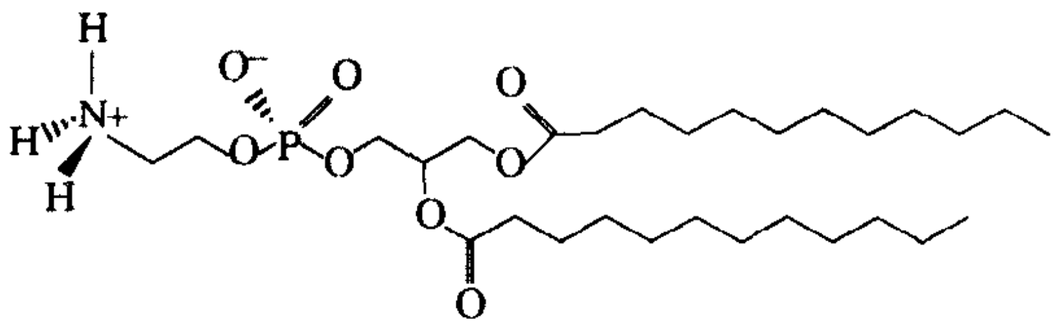
十二烷酸



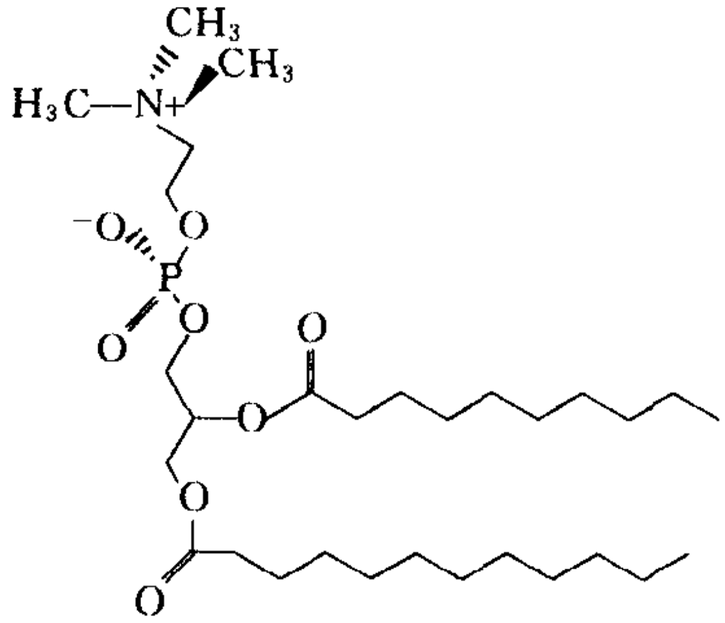
十八烷酸



二甲基-二(十八烷基)胺



二(十二酰基)磷脂酰乙醇胺



二(十四酰基)磷脂酰胆碱

图 8-9 典型的链状双性分子

双性体分子极性的头端具强的亲水性（hydrophilic），而其碳氢烷链则一般存在于疏水性（hydrophobic）的环境中。因此，这样的分子存在于水/油界面的两相中。双性分子最重要的特性是可以形成延伸的层状结构。在水/空气界面的单层结构称为朗格缪尔薄膜（Langmuir film）；在固体基

材表面形成的薄膜称为朗格缪尔-布罗杰特膜（Langmuir-Blodgett，LB）。LB 膜可能为单层、双层或多层的结构。实验室中可控制薄膜的厚度与双性分子的排列向性，可作为半导体材料的绝缘体、过滤组件、抗反射镀膜等重要的应用。双性分子在生物系统中有非常重要的功能，例如细胞膜即是由双性的双层脂质分子（lipid）所构成的。在高浓度的条件下，一些双性分子于水溶液中可形成微胞（micelles），为分子的头基朝外、碳链朝内的团状结构。图 8-10 显示一些由双性分子所构成的特殊系统。

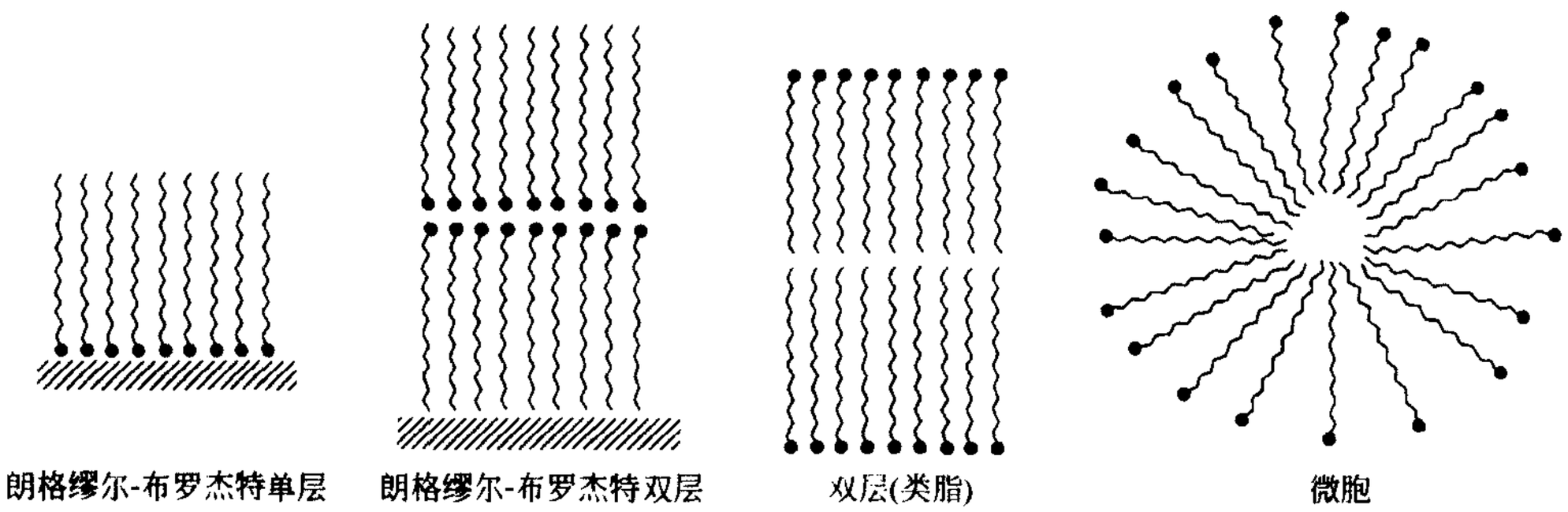



**图 8-10　双性分子所构成的特殊系统**

双性分子通常具有很复杂的相区，且常发现其有液晶（liquid crystal）相的存在。所谓的液晶相表示分子的排列在特定的方向具长距离的秩序性，并且形成层状的结构。通常系统的结构可以 X 射线衍射与核磁共振 NMR 实验侦测。由 NMR 实验所侦测的碳链的秩序性可以由秩序参数（order parameter）表示，其定义为：

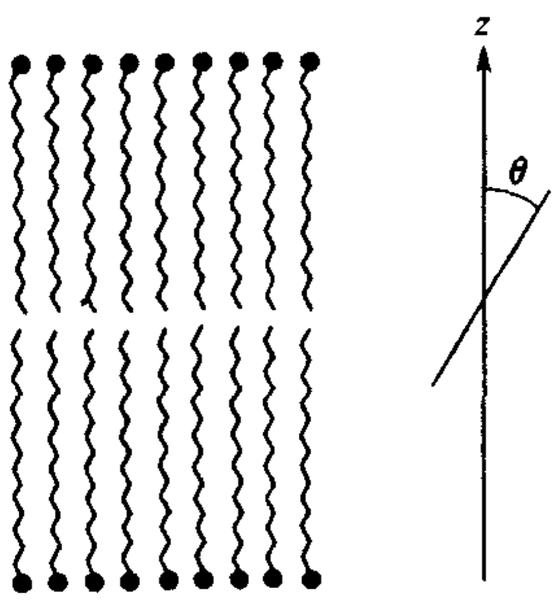



**图 8-11　双层薄膜的参考轴方向**

$$S=\frac{1}{2}(3\cos\theta_i\cos\theta_j-\delta_{ij}) \quad (8\text{-}12)$$

式中，$\theta_i$ 为第 $i$ 根碳链轴与参考轴的夹角，参考轴的方向为所有碳链轴平均的方向，通常将参考轴定为 $z$ 轴；$\delta_{ij}$ 为 Kronecker delta 函数（若 $i=j$，$\delta_{ij}=1$；若 $i\neq j$，$\delta_{ij}=0$）。图 8-11 显示一般双层结构薄膜的参考轴与碳链的夹角的定义。

由 NMR 实验所测得秩序参数的范围为 1.0（所有碳链同向排列）～—0.5（所有碳链垂直于双层膜的法向量）；若秩序参数为 0 表示碳链呈均向性的排列。Essex 等利用联合原子的简化模型，以分子动力计算模拟双性分子的双层结构。因在联合原子模型中，碳链上的 $CH_2$ 视为一基团，故模拟的结果欲与 NMR 所量测的 C—D 键的次序参数相比较必须做适当的修正。图 8-12 为联合原子模型中分子轴的示意。

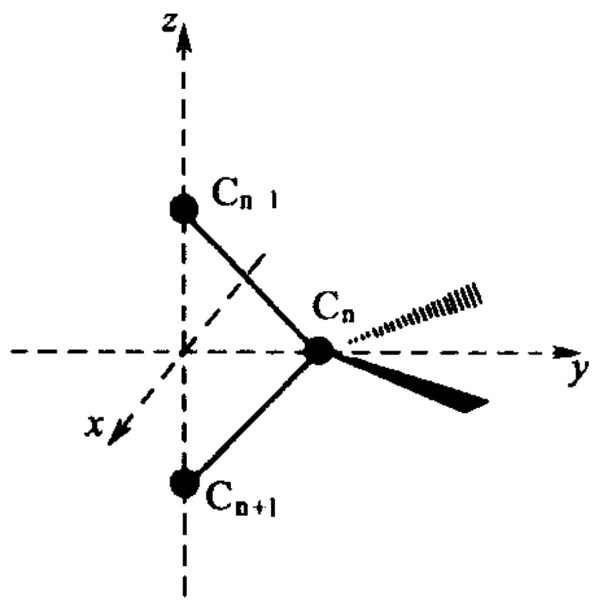



**图 8-12　碳链坐标轴的定义**

根据图 8-12 碳链上第 $n$ 个 $CH_2$ 基团所定义的分子轴向量如下所述。

$z$ 向量：$C_{n-1}$至 $C_{n+1}$的向量；

$y$ 向量：垂直 $x$ 轴且包含于 $C_{n-1}$，$C_n$，$C_{n+1}$所构成的平面上；

$x$ 向量：垂直 $y$ 与 $z$。

依据如上的定义，可由分子动力模拟的结果计算双层膜的碳链各方向的秩序参数。秩序参数为所有链向量，所有计算步数的平均，如 $z$ 方向的秩序参数 $S_{zz}$为分子碳链轴与双层法向量夹角的平均值。NMR 实验所测得的秩序参数与计算的次序参数间的关系为：

$$S_{CD}=\frac{2}{3}S_{xx}+\frac{1}{3}S_{yy} \tag{8-13}$$

Essex 等利用此关系式比较以简化联合原子模型计算的秩序参数与实验的量测值。两者十分吻合，成功地缩减了分子动力计算的工作。

除了计算朗格缪尔薄膜外，Kim 等利用分子动力计算模拟 LB 薄膜，计算的系统为 $CH_3(CH_2)_{16}COOH$ 于石墨表面的吸附。因为这样的体系并无适当的力场，Kim 等于是根据一些实验的数据设计出特定的势能函数。石墨表面的作用势能与原子（α）离表面的距离（$z_\alpha$）相关，其形式为：

$$U_{\alpha s}(z_\alpha)=\frac{2\pi\rho}{3}\epsilon_{ss}\left[\frac{2}{15}\left(\frac{\sigma_{\alpha s}}{z_\alpha}\right)^9-\left(\frac{\sigma_{\alpha s}}{z_\alpha}\right)^3\right] \tag{8-14}$$

式中，$\rho$ 为固体表面的数值密度（number density）；$\epsilon_{ss}$与 $\sigma_{\alpha s}$为范德瓦耳斯势能参数。双性分子极性头基与其镜像间的库仑作用势能为：

$$U_{ic}(z)=\frac{1}{2}\frac{\epsilon-\epsilon'}{\epsilon+\epsilon'}\left[\frac{q_\alpha^2}{8\pi\epsilon_0(z-z_{ip})}\right] \tag{8-15}$$

式中，$\epsilon'=4.0$ 为固体的相对穿透系数（relative permittivity）；$\epsilon=1.0$ 为表面上的穿透系数，镜像平面的位置为 $z_{ip}=\sigma_{ss}/2$。

Kim 等以全原子模型的分子动力计算模拟 64 根链在石墨上的 LB 薄膜。其计算的结果显示，双性分子呈约 20°集合性的倾斜，如图 8-13。

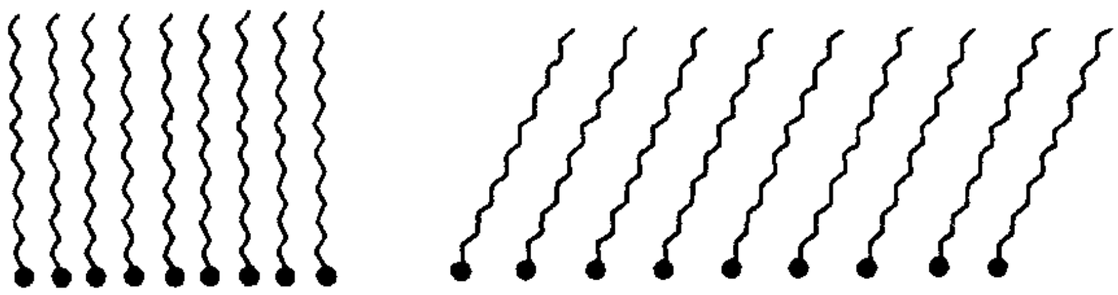


**图 8-13　LB 薄膜的分子轴排列**

此种倾斜是因头基（极性基团）面积的增加所致。计算亦发现碳链反式的比例随头基面积的增加而减少；头基面积为 20.6Å$^2$ 时，97.7%的碳链呈反式构型；头基面积为 21.2Å$^2$ 时反式的构型减少为 66.9%。计算中并发现与头基相连的键具相当程度的转动无序性，这表示碳链的秩序参数值很小。以上的分子动力计算成功地检验了 LB 薄膜的结构，并探讨分子链倾斜的原因，非常有助于实验设计新的薄膜材质。

## 8.5　聚合物的特殊运动

由分子动力计算所获得的运动轨迹可用以检视分子的运动。此处应用分子动力计算检视聚合物中特殊的片段运动。一般安全玻璃的材料均为含有苯环的聚合物，最典型的安全玻璃聚合物为双酚 A 聚碳酸酯 {bisphenol-A polycarbonate，BPAPC，$\{C(O)O-C_6H_4-C(CH_3)_2-O\}_n$} 与聚醚醚酮 {PEEK，$\{O-C_6H_4-O-C_6H_4-C(O)\}_n$}。这些聚合物均为韧性的材料。常温下，BPAPC 与 PEEK 均呈非晶形的结构，为透明的材质。其熔点与玻璃转换温度（glass transition temperature）都在 400K 以上。图 8-14 为 PEEK 聚合物的单体结构。

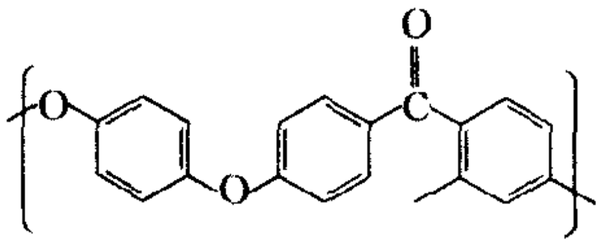


**图 8-14　PEEK 聚合物的单体结构**

由于聚合物中含有大量的苯环，且苯环的旋转不会改变聚合物的整体结构，因此一般认为 BPAPC 或 PEEK 等聚合物材质遭受撞击时系统所接受的能量可用于苯环的旋转。由这种运动消耗接受的能量，而不会导致材质

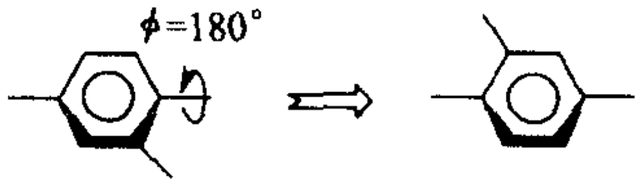



**图 8-15　苯环 π 翻转示意**

的碎裂，这是其为安全玻璃的原因。Schaefer 等由固态核磁共振光谱的证据推断 BPAPC 中苯环的主要运动为 180°的旋转，称为“π 翻转”（pi-flip），见图 8-15。旋转前后不改变聚合物的任何结构，因此可借此翻转消耗多余的能量。但是苯环的翻转到底能消耗多少能量却有不同的推测，综合各种实验的结果，一般任为翻转的能量应在 3～12kcal/mol。

Chen 等首先提出利用分子动力模拟的方法检视安全玻璃聚合物材质中苯环的翻转运动。Chen 等的分子动力模拟引用 TRIPOS 5.2 力场，该力场的形式为：

$$U=U_{VDW}+U_l+U_\theta+U_\phi+U_d \tag{8-16}$$

式中，$U_{VDW}$为范德瓦耳斯作用项；$U_l$、$U_\theta$、$U_\phi$ 分别为键长伸缩、键角弯曲、二面角扭曲项；$U_d$ 为离平面振动项。计算的温度分别为 300K 及 410K，模拟的 PEEK 链含有 61 个苯环。计算采用 Beeman 方法，积分步程为 1fs，共收集 60ps 的轨迹作为分析。

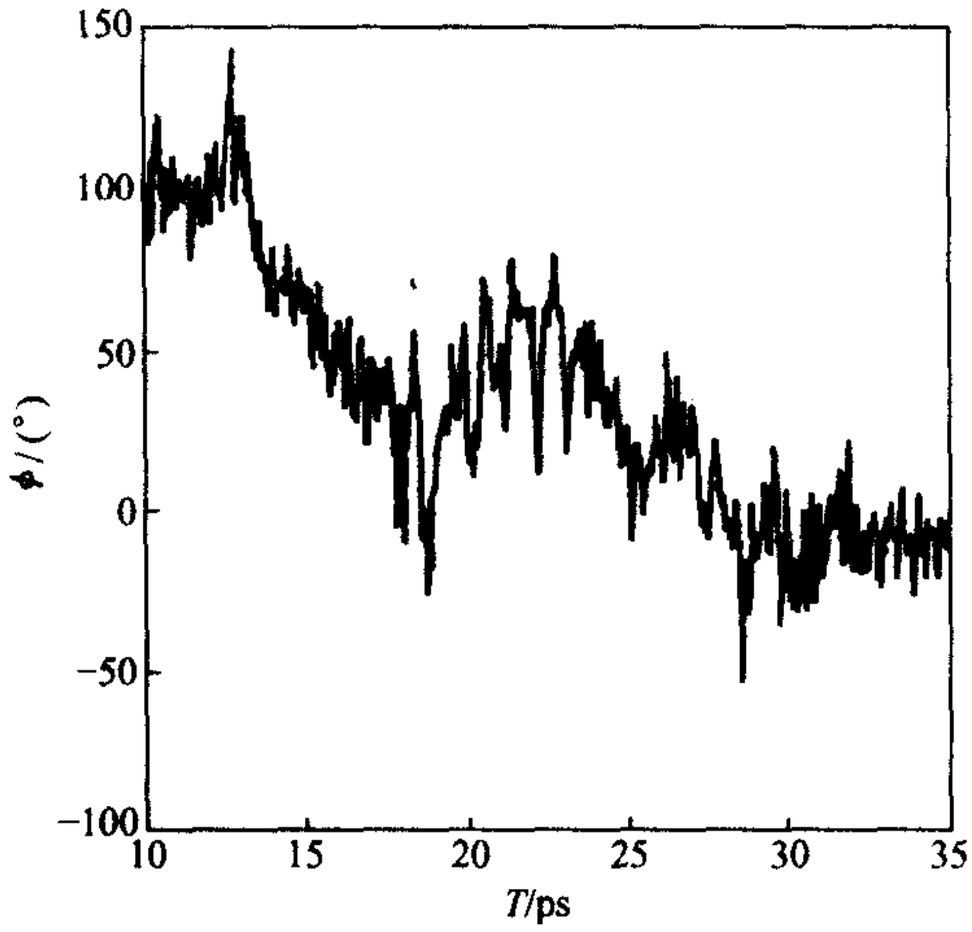



**图 8-16　PEEK 中苯环扭转角的运动轨迹**

图 8-16 为计算系统中，一个苯环的扭转角的运动轨迹。图中，此苯环的扭转角在时间范围 13～28ps 内改变 180°，明显地证实了 π 翻转运动的发生。计算显示在 300K 时系统内仅有两个苯环发生 π 翻转，但在 410K

时共有 9 个苯环发生 π 翻转。故温度愈高，愈容易发生 π 翻转；或解释为外界供给系统的能量愈多，愈容易发生 π 翻转。

以如图 8-17 的坐标系统定义苯环的翻转运动。图中，坐标的原点为发生 π 翻转苯环的几何中心；$z$ 轴沿着连接苯环的键，$x$ 轴位于如图所示的 $L_1$-$L_2$-$C_4$ 平面上。

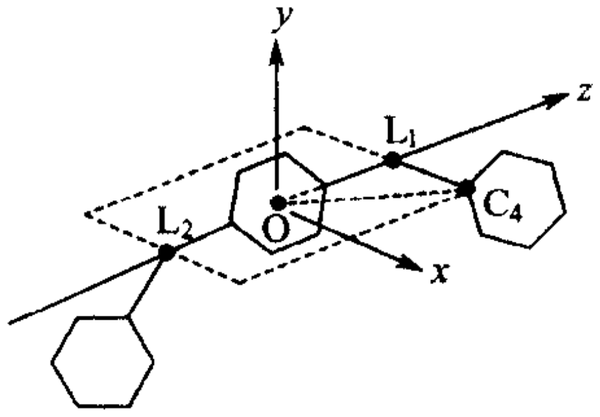



**图 8-17　苯环翻转坐标的定义**

依此坐标系统，苯环绕着 $z$ 轴的角动量分量为：

$$J_z = \sum_{i=1}^{n} (x_i p_{y,i} - y_i p_{x,i}) \quad (8\text{-}17)$$

式中，$x_i$，$y_i$ 为环上第 $i$ 个原子的坐标；$p_{x,i}$，$p_{y,i}$ 为其沿 $x$，$y$ 方向的线性动量。苯环可以绕着 $z$ 轴以顺时针或逆时针方向旋转，顺时针旋转时 $J_z$ 为正值，逆时针时为负值。苯环旋转的动能为：

$$E_{rot} = \frac{J_z^2}{I_{zz}} \quad (8\text{-}18)$$

式中，$I_{zz}$ 为 $z$ 方向的转动惯量，即：

$$I_{zz} = \sum_{i=1}^{n} m_i (x_i^2 + y_i^2) \quad (8\text{-}19)$$

图 8-18 为 Tasi 等模拟的 BPAPC 系统中发生 π 翻转的一个苯环的扭转角轨迹，$J_z$，及转动能。图 8-18(b) 中显示转动的角动量正负值兼具，表示苯环在 π 翻转时不停地摆动，并非沿单一方向的翻转。将图 8-18(b) 与图 8-18(a) 对比，可以得到由 $J_z=0$ 到下一次 $J_z=0$ 间的转动能量。令顺时针方向旋转的能量为 $E(+)$，逆时针方向旋转的能量为 $E(-)$。苯环 π 翻转所需的能量可估计为：

$$\Delta E = |\sum E_{max}(+) - \sum E_{max}(-)| \quad (8\text{-}20)$$

式中，下标“max”表示此段能量的极大值。式(8-20) 的物理意义为顺时针方向的转动能之总和与逆时针方向转动能总和的差值，即为过剩的动能，也就是翻转所需要克服的能量障碍。

表 8-7 列出利用这种分析方法，计算所得 BPAPC 系统 π 翻转的时间及对应翻转的能障。由表 8-7 中所列出的数据知温度愈高翻转的苯环数目愈多，或可解释为外界供给的能量愈大，愈容易造成材质中苯环的翻转。此外，表中所列计算所得的翻转能障范围为 2.8～12.8kcal/mol，与实验

的测量值十分吻合。此例利用分子动力模拟的方法，借着分析扭转角的轨迹，成功地检测出苯环的π翻转运动及与温度的关系，并借着角动量的计算估算出翻转所需的能障。此种研究模式可应用于检测如聚合物的片段运动、糖分子的环转动等特殊单位的运动。

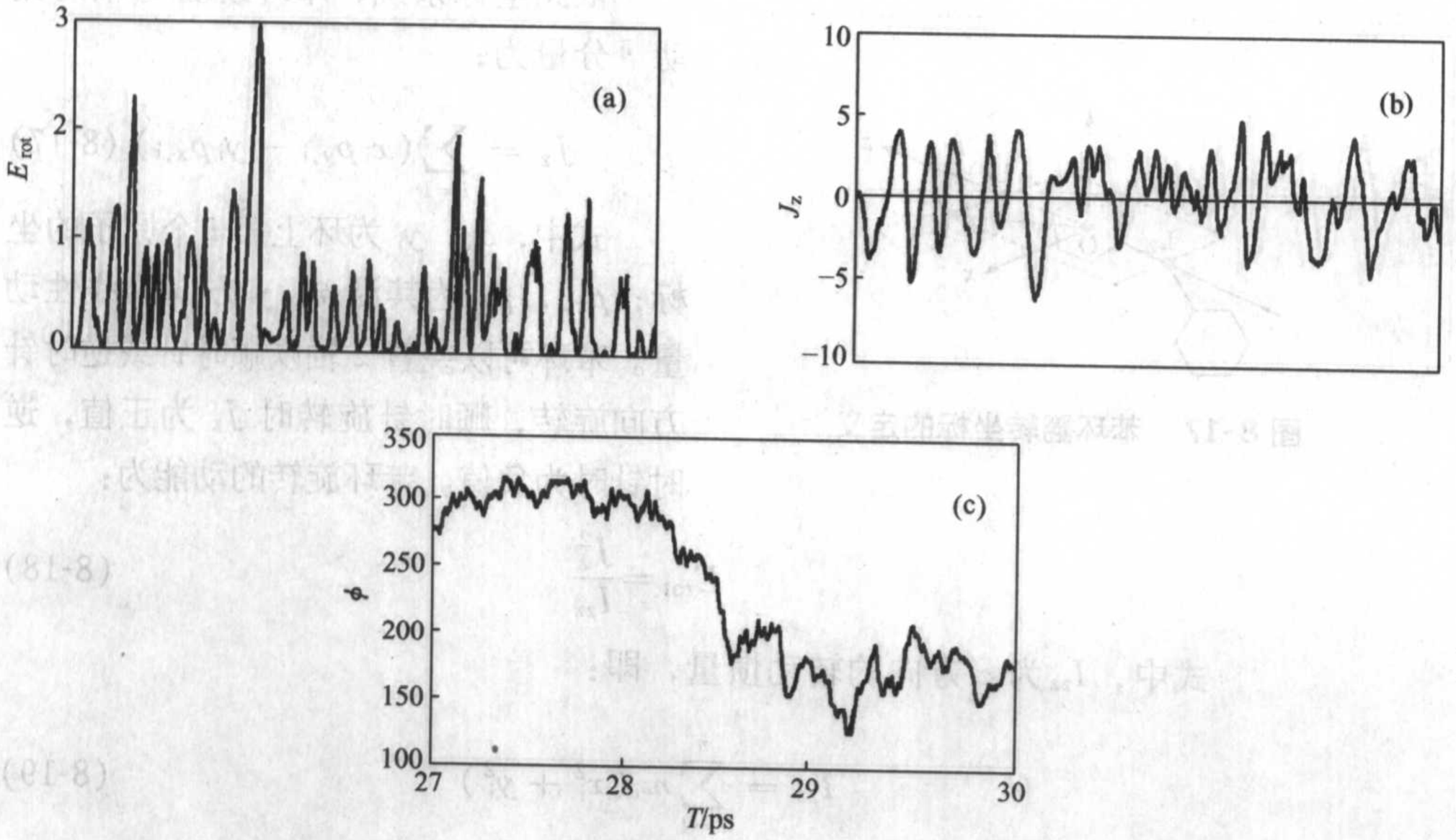



图 8-18　聚合物 BPAPC 中苯环翻转的转动能、 角动量及运动轨迹

表 8-7　BPAPC 中苯环翻转的时间与翻转的能量障碍

| π翻转苯环序号 | π翻转时间/ps | π翻转能障/(kcal/mol) |
|---|---|---|
| | 温度 410K | |
| 1 | 1.5 | 2.8 |
| 23 | 8.7 | 9.9 |
| 28 | 4.3 | 7.6 |
| 28 | 2.2 | 7.4 |
| 28 | 9.1 | 9.8 |
| 28 | 1.8 | 4.1 |
| 39 | 16.0 | 7.9 |
| 42 | 20.9 | 12.8 |
| 50 | 41.6 | 7.4 |
| | 温度 300K | |
| 2 | 17.2 | 7.2 |
| 53 | 8.0 | 4.6 |

## 参 考 文 献


1 Stillinger F H. Adv. Chem. Phys., 81, 1, 1975

2 Stillinger F H, Rahman A. J. Chem. Phys., 68, 666, 1978

3 Jorgensen W L. J. Am. Chem. Soc., 103, 335, 1981 (TIPS)

4 Jorgensen W L. J. Chem. Phys., 77, 4156, 1981 (TIPS2)

5 Jorgensen W L, Chandrasehkar J, Madura J D, Impey R W, Klein M L. J. Chem. Phys., 79, 926, 1983 (TIP3P, TIP4P)

6 Berendsen H C, Postma J P M, van Gusteren W F, Hermans J. "Intermolecular Forces" Dordrecht, Reidel 331, 1981

7 Niesar U, Corongiu G, Clementi E, Keller G R, Bhattacharya D K. J. Am. Chem. Soc., 94, 7949, 1990

8 Corongiu G J. Quant. Chem., 42, 1209, 1992

9 Quintana J, Haymet A D J. Chem. Phys. Lett., 189, 273, 1992

10 Mencarelli P. J. Chem. Edu., 72, 511, 1995

11 Yin K L, et al. Acta. Phys. -Chin. Sin., 20, 302, 2004

12 Essex J W, Hann M M, Richards W G, Phil. Trans. Royal Soc. London B, 344, 239, 1994

13 Kim K S, Moller M A, Tildesley D J, Quirke N. Molecular Simulation, 13, 77, 1994

14 Kim K S, Tildesley D J, Quirke N. Molecular Simulation, 13, 101, 1994

15 Schaefer J, Stejskal E O, McKay R A, Dixon W T. Macromolecules, 19, 1356, 1986

16 Chen C L, Lee C L, Chen H L, Shih J H. Macromolecules, 27, 7872, 1994

17 Tsai S F, Lan I K, Chen C L. Computation And Theoretical Polymer Science, 8, 283, 1998

# 第9章 长程作用力的计算

分子力学、分子动力和蒙地卡罗计算中常常需要考虑带电荷的原子、离子或分子间的作用。电荷作用项随着 $r^{-n}$ 衰减，称为长程作用项。许多体系需要计算准确的长程作用项才能得到正确的相关性质，如计算系统的介电常数（dielectric constant）、光谱等，因此，远程作用力的正确处理是各种分子模拟方法中又一个重要问题。本章介绍了分子动力学计算中长程作用力的处理方法和溶剂介电常数的模型化方法。

## 9.1 中央多极展开法

两分子间（或同分子相距一定距离的两个基团间）的静电作用可视为所有原子核上的点电荷对的作用总和。依照库仑作用定律，可写为：

$$U_{\text{elec}} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{9-1}$$

式中，$N_A$，$N_B$ 为两分子所有点电荷的数目。一般将原子核上的电荷称为部分原子电荷（partial atomic charge）或净原子电荷（net atomic charge）。点电荷间的静电作用势能收敛缓慢，当距离很远时才趋近于零。例如在空间中相距 100Å 带 +0.5e 与−0.5e 的两个电荷，其间的作用势能为：

$$U = 332 \times \frac{0.5 \times 0.5}{100} = 0.83 \text{ (kcal/mol)} \tag{9-2}$$

依此，欲使其收敛至小数点后第 2 位，两个电荷必须相距 1000Å。此意味着计算中必须考虑较大的体系，否则计算的能量不稳定，不能收敛。但计算系统增大，则增加许多的计算量。例如，水的密度为 $1\text{g/cm}^3$，边长 1000Å 的立方体含有：

$$1\text{g/cm}^3\times(1000\times10^{-8}\text{cm})^3=1.0\times10^{-15}\text{g}\rightarrow\frac{1.0\times10^{-15}}{18/(6.02\times10^{23})}$$

$$=3.3\times10^7\text{（个分子）}$$

而通常的分子动力计算系统约含数千个至数万个粒子，这样庞大的体系是远远超出计算的能力范围。为了有效地计算分子间的静电作用，通常引用中央多极展开法（central multiple expansion）。这种方法将分子视为一个整体，利用分子的电次极计算相互间的作用。分子的各电极（pole）分别为：零极（分子所带的电荷量）、二极（偶极，dipole）、四极（quadrupole）及八极（octapole）等。通常，最重要的是分子最低不为零的电极。非中性的离子或分子，如 $Na^+$、$Cl^-$、$NH_4^+$、$CH_3CO_2^-$ 等均带有电荷。中性的分子如水、$NH_3$、$CH_3Cl$ 等最低的电极为偶极。如 $CO_2$、$N_2$ 等分子最低的电极为四极。$CH_4$ 与 $CCl_4$ 等分子最低的电极为八极。各种电极可用适当的电荷排列表示，如偶极可表示为两个带相反电量的电荷排在适当的距离，四极可用四个电荷排列表示，八极则可用八个电荷的排列表示。图 9-1 为四极的各种表示法。

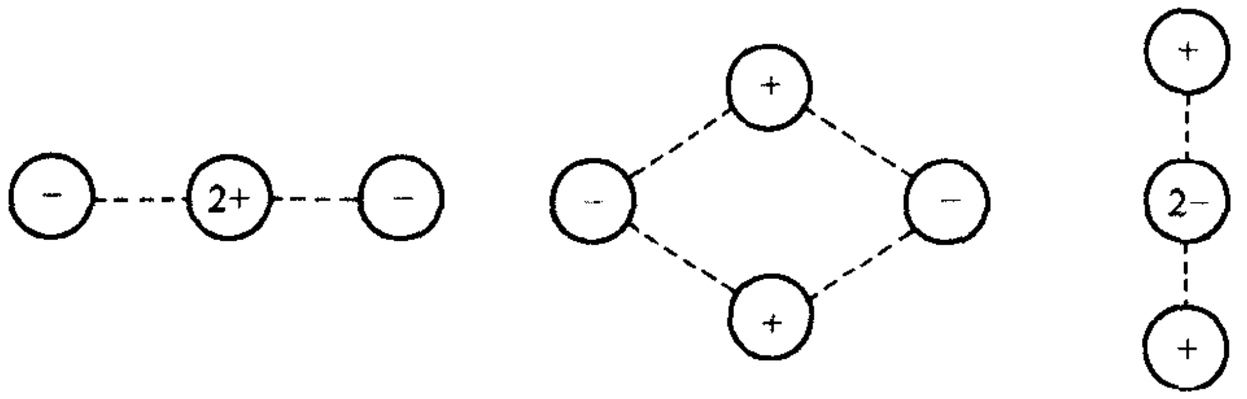



**图 9-1　四极的各种表示方法**




**图 9-2　二点电荷所产生的电势**

讨论分子电荷间的作用力，最简单的模型为双原子分子。设一双原子分子的原子分别带有 $q_1$ 及 $q_2$ 的电荷，位于 $-z_1$ 及 $+z_2$ 的位置，如图 9-2 所示。

则位于图中 $P$ 点的势能为：

$$\phi(r)=\frac{1}{4\pi\varepsilon_0}\left(\frac{q_1}{r_1}+\frac{q_2}{r_2}\right)\tag{9-3}$$

由余弦定律，可得：

$$\phi(r)=\frac{1}{4\pi\epsilon_0}\left(\frac{q_1}{\sqrt{r^2+z_1^2+2rz_1\cos\theta}}+\frac{q_2}{\sqrt{r^2+z_2^2-2rz_2\cos\theta}}\right) \tag{9-4}$$

若 $r\gg z_1$，且 $r\gg z_2$，则式(9-4) 可展开为：

$$\phi(r)=\frac{1}{4\pi\epsilon_0}\left[\frac{q_1}{r\sqrt{1+(z_1/r)^2+2(z_1/r)\cos\theta}}+\frac{q_2}{r\sqrt{1+(z_2/r)-2(z_2/r)\cos\theta}}\right]$$

$$=\frac{1}{4\pi\epsilon_0}\left[\frac{q_1+q_2}{r}+\frac{(q_2z_2-q_1z_1)\cos\theta}{r^2}+\frac{(q_1z_1^2+q_2z_2^2)(3\cos^2\theta-1)}{2r^3}+\cdots\right]$$

$$=\frac{1}{4\pi\epsilon_0}\left[\frac{q}{r}+\frac{\mu\cos\theta}{r^2}+\frac{\Theta(3\cos^2\theta-1)}{2r^3}+\cdots\right] \tag{9-5}$$

式中，$q=q_1+q_2$ 为分子的总电荷；$\mu=(q_2z_2-q_1z_1)$ 为分子的偶极；$\Theta=(q_1z_1^2+q_2z_2^2)$ 为分子的四极，依此类推。由以上的展开式可知，$P$ 点的势能可展开成分子多极的多项式，这种方法称为中央多极法。此展开式中，仅第 1 个不为零的次极项与坐标无关。若分子为中性分子（$q_1+q_2=0$），则其偶极矩与位置无关。例如此例中，若原点移至 $z'$，则偶极为：

$$\mu'=q_2(z_2+z')-q_1(z_1-z')=\mu+qz'=\mu \tag{9-6}$$

因此，若总电荷为零，则偶极不变。同理可得出若总电荷与偶极均为零，则四极与选择的位置无关。通常在计算中，选取电荷分布的中心为原点。

电极的正确数学式为张量（tensor），电荷为 0 阶张量，偶极为 1 阶张量，四极为 2 阶张量等。阶数为 $n$ 的张量含有 $3^n$ 个元素。对空间中的电荷排列，其电偶极为：

$$\vec{\mu}=\sum q_i\vec{r}_i \tag{9-7}$$

依此式，偶极矩在笛卡儿坐标轴的分量为 $\mu_x=\sum q_ix_i$，$\mu_y=\sum q_iy_i$，$\mu_z=\sum q_iz_i$。四极的张量形式为：

$$\Theta=\begin{pmatrix}\sum q_ix_i^2 & \sum q_ix_iy_i & \sum q_ix_iz_i\\ \sum q_iy_ix_i & \sum q_iy_i^2 & \sum q_iy_iz_i\\ \sum q_iz_ix_i & \sum q_iz_iy_i & \sum q_iz_i^2\end{pmatrix} \tag{9-8}$$

由式(9-8)，四极与电荷在坐标轴的排列与位向有关。另一种定义四极矩的方法为：

$$\Theta=\frac{1}{2}\begin{pmatrix}\sum q_i(3x_i^2-r_i^2) & 3\sum q_ix_iy_i & 3\sum q_ix_iz_i\\ 3\sum q_ix_iz_i & \sum q_i(3y_i^2-r_i^2) & 3\sum q_iy_iz_i\\ 3\sum q_iz_ix_i & 3\sum q_iz_iy_i & \sum q_i(3z_i^2-r_i^2)\end{pmatrix} \tag{9-9}$$

式中，$r_i^2=x_i^2+y_i^2+z_i^2$。这样的定义可了解电荷排列与球形对称分布

的差异，若为球形对称分布，则：

$$\sum q_i x_i^2 = \sum q_i y_i^2 = \sum q_i z_i^2 = \frac{1}{3}\sum q_i r_i^2 \tag{9-10}$$

故球形对称的分子，上张量式中对角线的元素皆为零。另一种四极表示的方法为计算其主轴（principal axes）$\alpha$，$\beta$，$\gamma$ 的张量。主轴为三根互相垂直的轴，其四极的张量除主轴上的分量外均为零，即：

$$\Theta = \begin{pmatrix} \Theta_{\alpha\alpha} & 0 & 0 \\ 0 & \Theta_{\beta\beta} & 0 \\ 0 & 0 & \Theta_{\gamma\gamma} \end{pmatrix} \tag{9-11}$$

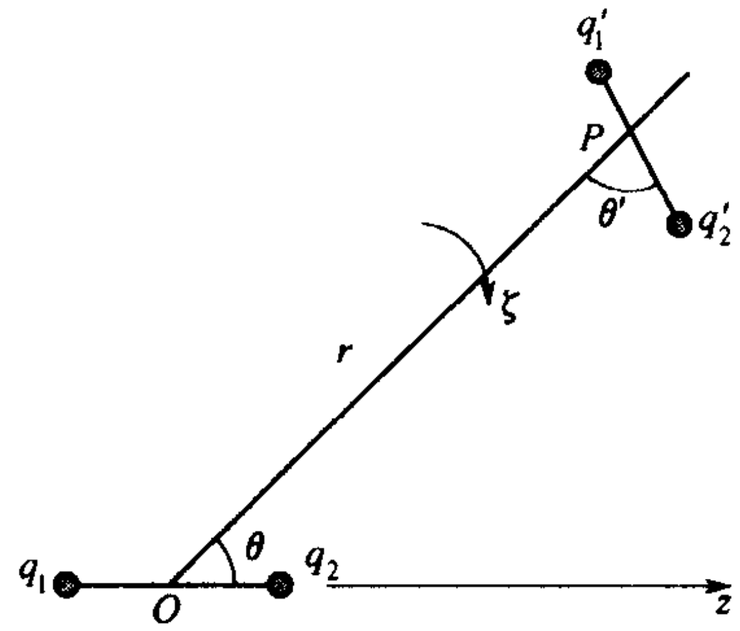



**图 9-3　二偶极排列的相对位置**

图 9-3 考虑在此两电荷 $q_1$ 与 $q_2$ 外，有另一线性的电荷分布排列于 $P$ 点的位置（电荷为 $q_1'$ 与 $q_2'$）。图中，$P$ 点为此两电荷的中心。

利用中央多极法计算此二偶极的作用势能为：

$$U(q,q') = \frac{1}{4\pi\epsilon_0}\left[\begin{array}{l} \dfrac{qq'}{r} \\ +\dfrac{1}{r^2}(q\mu'\cos\theta + q'\mu\cos\theta') \\ +\dfrac{\mu\mu'}{r^3}(2\cos\theta\cos\theta' + \sin\theta\sin\theta'\cos\zeta) \\ +\dfrac{1}{2r^3}[q\Theta'(3\cos^2\theta' - 1) + q'\Theta(3\cos^2\theta - 1)] \\ +\dfrac{3}{2r^2}\left[\begin{array}{l} \mu\Theta'\{\cos\theta(3\cos^2\theta' - 1) + 2\sin\theta\sin\theta'\cos\theta'\cos\zeta\} \\ +\mu'\Theta\{\cos\theta'(3\cos^2\theta - 1) + 2\sin\theta'\sin\theta\cos\theta\cos\zeta\} \end{array}\right] \\ +\dfrac{3\Theta\Theta'}{4r^5}\left[\begin{array}{l} 1 - 5\cos^2\theta - 5\cos^2\theta' + 17\cos^2\theta\cos^2\theta' \\ +2\sin^2\theta\sin^2\theta'\cos^2\zeta + 16\sin\theta\sin\theta'\cos\theta\cos\theta'\cos\zeta \end{array}\right] \\ +\cdots \end{array}\right] \tag{9-12}$$

此式表示作用的势能为一无穷展开式，包含分子的电荷-电荷、电荷-偶极、偶极-偶极、电荷-四极、偶极-四极等作用项。各项与 $r$ 的 $n$ 次方的倒数相关。若分子为中性分子（$q=q'=0$），则作用项中第一不为零的项为偶极-偶极的作用，其大小反比于 $r^3$。此结果非常重要，在相同的作用范围内，

偶极-偶极间的作用（$r^{-3}$）远小于电荷-电荷间的库伦作用（$r^{-1}$）。图 9-4 显示此两作用项的比较。

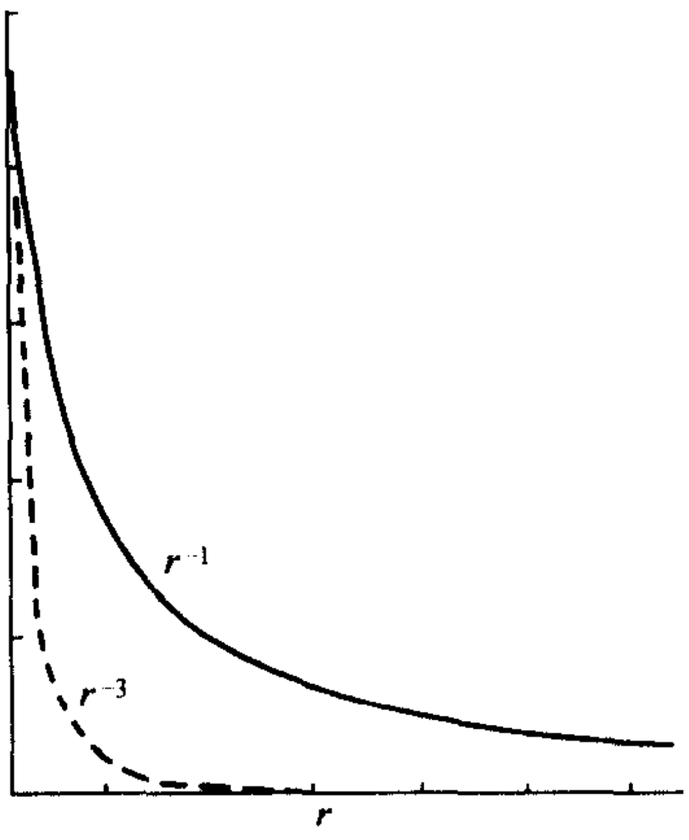



**图 9-4　点电荷库仑作用与偶极-偶极作用的比较**

由图 9-4 可知，当 $r$ 值增大时，偶极矩-偶极矩的作用项衰减的很快，因此无需计算长距离的作用。中央多极展开法适用于 $r$ 值远大于分子尺度的状况。通常，$n$ 次与 $m$ 次极矩间之作用项随 $r^{-(n+m+1)}$ 而衰减。

利用中央多极展开法计算静电作用的优点在于其具有较高的效率。以苯分子为例，若以原子所带点电荷模型计算两个苯分子的静电作用，需要计算 144 个作用项。但苯为无电荷且无偶极的分子，最低次不为零的电极为四极。因此，以中央多极展开法计算两个苯分子间的静电作用仅需计算一个四极-四极作用项。当两个分子间的距离若与分子的尺度相当，则不可引用中央多极展开法。通常当分子间的距离必须大于各个分子的中心至其最远的电荷距离的和时，才能应用中央多极展开法。

## 9.2　Ewald 求和法

Ewald 求和法为 Ewald 于 1921 年所提出，最初用以计算离子晶体的能量。此方法选定一计算系统盒子，盒中的粒子除了与盒中其他的粒子作用外，亦与所有其镜像系统中的粒子作用。镜像系统与计算系统的结构完全相同。图 9-5 显示计算的作用系统，中央的黑色区域为计算的系统盒子，周围的镜像系统随距离的增加以渐淡的颜色表示，其极限的作用区域如同一球，称为 Ewald 球体（Ewald sphere）。

设作用系统的边长为 $L$，含有 $N$ 个粒子，则镜像系统的位置可以 $(\pm iL, \pm jL, \pm kL)$；$i,j,k=0,1,2,3,\cdots$ 表示。计算系统盒子中所有的点电荷间的作用为：

$$U=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (i\neq j) \tag{9-13}$$

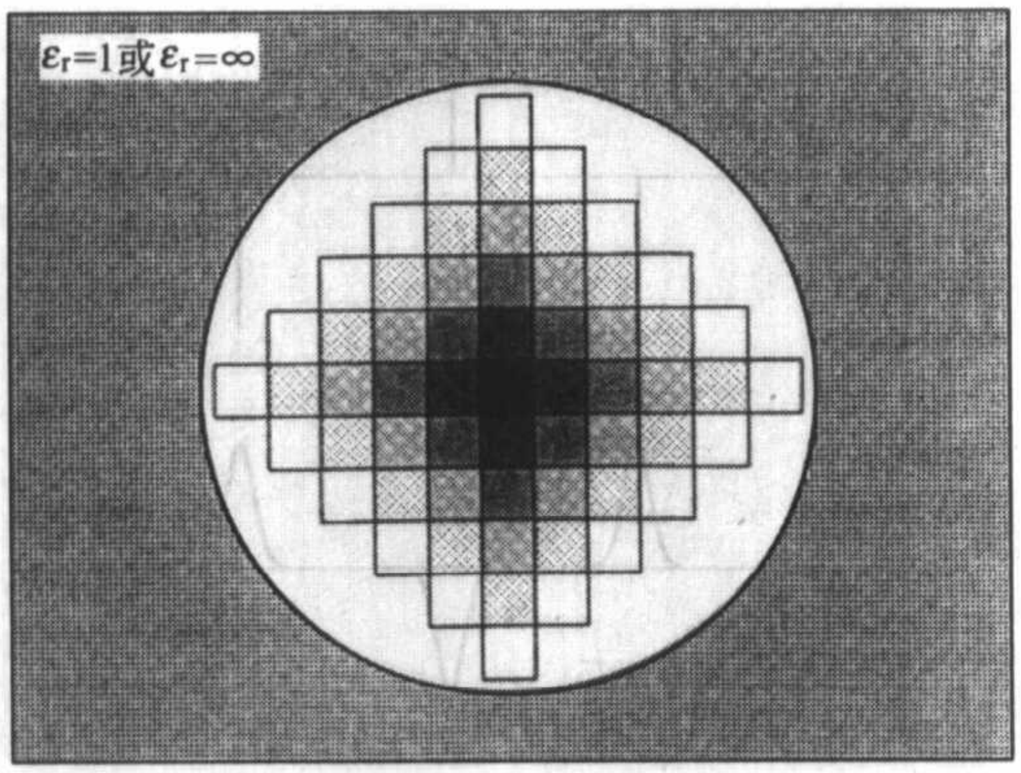



**图 9-5　建立 Ewald 求和法的周期作用区**

式中，$r_{ij}$为 $i$ 与 $j$ 间的最短距离。与计算盒子系统最近的盒子共有 6 个（图 9-5 的平面系统中只显示其中的 4 个），其位置 $\vec{r}_{box}$分别为 $(0,0,L)$、$(0,0,-L)$、$(0,L,0)$、$(0,-L,0)$、$(L,0,0)$、$(-L,0,0)$。计算系统盒子与此 6 个邻近的镜像系统点电荷间的作用为：

$$U=\frac{1}{2}\sum_{box=1}^{6}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\epsilon_0\mid\vec{r}_{ij}+\vec{r}_{box}\mid} \tag{9-14}$$

计算系统盒子的镜像位置可用向量 $\vec{n}=(n_xL,n_yL,n_zL)$ 表示，$n_x$，$n_y$，$n_z$ 均为整数。依此式，电荷间的作用可表示为：

$$U=\frac{1}{2}\sum_{\vec{n}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\epsilon_0\mid\vec{r}_{ij}+\vec{n}\mid} \tag{9-15}$$

向量 $\vec{n}$ 的绝对值$|\vec{n}|=1,\sqrt{2},\cdots$若加入计算系统盒子中的电荷作用（$|\vec{n}=0$），作用项可一并写为：

$$U=\frac{1}{2}\sum_{|\vec{n}|=0}^{\infty}{}'\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\epsilon_0\mid\vec{r}_{ij}+\vec{n}\mid} \tag{9-16}$$

式中，上标（′）表示$|\vec{n}|=0$时不计算 $i=j$ 的作用。实际计算式(9-16)的求和项，其收敛的速度非常缓慢。Ewald 的求和方法（Ewald summation）为将式(9-16) 以两个收敛速度较快的数列取代。Ewald 将每个电荷视为被同量反相呈中和（neutralising）的电荷分布所围绕，如图 9-6 所示。

Ewald 将电荷分布以高斯分布的函数形式表示，即：

$$\rho_i(\vec{r})=\frac{q_i\alpha^3}{\pi^{3/2}}\exp(-\alpha^2 r^2) \tag{9-17}$$

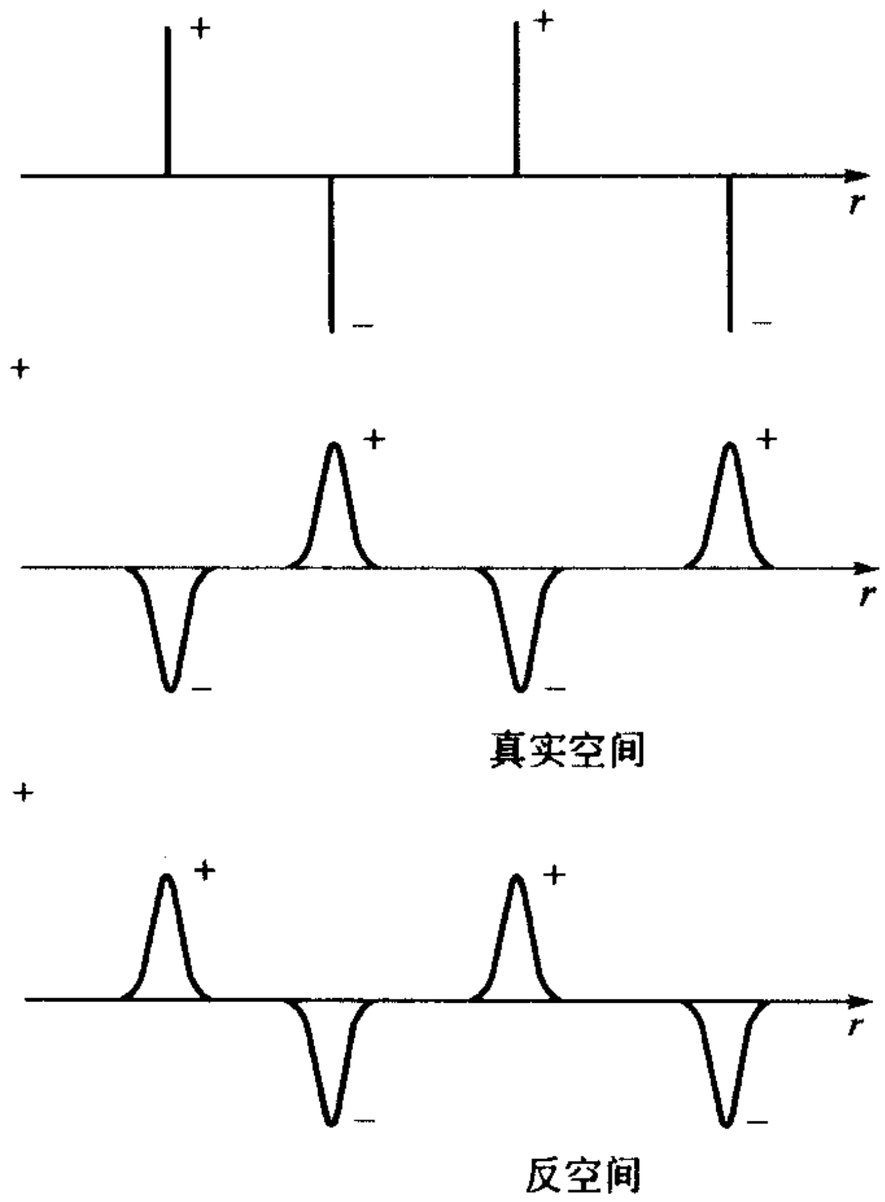



**图 9-6　Ewald 方法的电荷分布**

利用这种电荷分布式，可将点电荷间的加成项转换为与电荷及中性电荷分布间的作用。Ewald 方法的第一个数列为由真实空间（real space）所得的加成项，即：

$$U = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{|\vec{n}|=0}^{\infty}{}' \frac{q_i q_j}{4\pi\epsilon_0}\frac{\mathrm{erfc}(\alpha\,|\,\vec{r}_{ij}+\vec{n}\,|)}{|\,\vec{r}_{ij}+\vec{n}\,|} \tag{9-18}$$

式中，erfc 为互补错误函数（complementary error function）。

$$\mathrm{erfc}(x)=\frac{2}{\sqrt{\pi}}\int_x^{\infty}\exp(-t^2)\mathrm{d}t \tag{9-19}$$

此式最重要之处，在于含有收敛速度极快的错误函数。因此加成式中通常仅需计算几个$|\vec{n}|$值即可收敛。加成式收敛的速度取决于高斯函数中的 $\alpha$。$\alpha$ 值愈大，加成项的收敛速度愈快。实际计算上通常选取适当的 $\alpha$ 值使得式（9-18）的加成式仅需计算$|\vec{n}|=0$ 的电荷作用。Ewald 方法的第二个数列为由反空间（reciprocal space）所得的加成项，其电荷分布与第一个数列中的电荷分布等量而异号，如图 9-6 中所示。第二个数列的加成式为：

$$U = \frac{1}{2}\sum_{k\neq 0}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{\pi L^3}\frac{q_i q_j}{4\pi\epsilon_0}\frac{4\pi^2}{k^2}\exp\left(-\frac{k^2}{4\alpha^2}\right)\cos(\vec{k}\cdot\vec{r}_{ij}) \tag{9-20}$$

式中，$\vec{k}$ 为晶体的反向量（reciprocal vector），$\vec{k}=2\pi\vec{n}/L^2$。此加成式的收敛速度亦远快于点电荷加成式，但其加成项 $k$ 正比于 $\alpha$。实际计算时，通常选取 $\alpha=5/L$，计算约 100～200 个 $\vec{k}$ 值的加成项。

计算电荷作用的加成，在式(9-20) 的加成式中必须除去真实空间中高斯函数自身的作用项，即：

$$U=-\frac{\alpha}{\sqrt{\pi}}\sum_{k=1}^{N}\frac{q_k^2}{4\pi\epsilon_0} \tag{9-21}$$

另一校正项与环绕 Ewald 球介质的性质有关，若介质为完全导电体（穿透系数 $\epsilon$ 无限大），则不需加入此校正项；但若介质为真空（穿透系数 $\epsilon$ 为 1），则需加入校正项为：

$$U_{\text{correction}}=\frac{2\pi}{3L^3}\left|\sum_{i=1}^{N}\frac{q_i}{4\pi\epsilon_0}\vec{r}_i\right|^2 \tag{9-22}$$

综合上述，利用 Ewald 方法计算的作用势能的表示式为：

$$U=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\{\begin{aligned}&\sum_{|\vec{n}|=0}^{\infty}{}'\frac{q_iq_j}{4\pi\epsilon_0}\frac{\operatorname{erfc}(\alpha|\vec{r}_{ij}+\vec{n}|)}{|\vec{r}_{ij}+\vec{n}|}\\&+\sum_{k\neq 0}\frac{1}{\pi L^3}\frac{q_iq_j}{4\pi\epsilon_0}\frac{4\pi^2}{k^2}\exp\left(-\frac{k^2}{4\alpha^2}\right)\cos(\vec{k}\cdot\vec{r}_{ij})\\&-\frac{\alpha}{\sqrt{\pi}}\sum_{k=1}^{N}\frac{q_k^2}{4\pi\epsilon_0}\\&+\frac{2\pi}{3L^3}\left|\sum_{k=1}^{N}\frac{q_k}{4\pi\epsilon_0}\vec{r}_k\right|^2\end{aligned}\right. \tag{9-23}$$

Ewald 方法为各种计算静电作用的方法中准确性最高最为可行的方法。通常运用于计算离子性液体、固体等高带电量的系统，和一些静电作用明显的体系，如酯类、蛋白质及核酸等。Ewald 方法比一般的分子动力计算费时，因为需要计算许多加成项，其速度约为计算无静电作用系统的 1/10。

## 9.3 作用场与镜像电荷计算法

除了 Ewald 方法外，还有一些较简便的计算静电作用的方法，例如作用场（reaction field）方法。其相关工作可参考 Friedman 的论文。此方法计算截断半径内所有电荷间的库仑静电作用，而将截断半径外的库仑作用

视为介电常数为 $\epsilon_s$ 的均匀介质对电荷的作用，如图 9-7 所示。

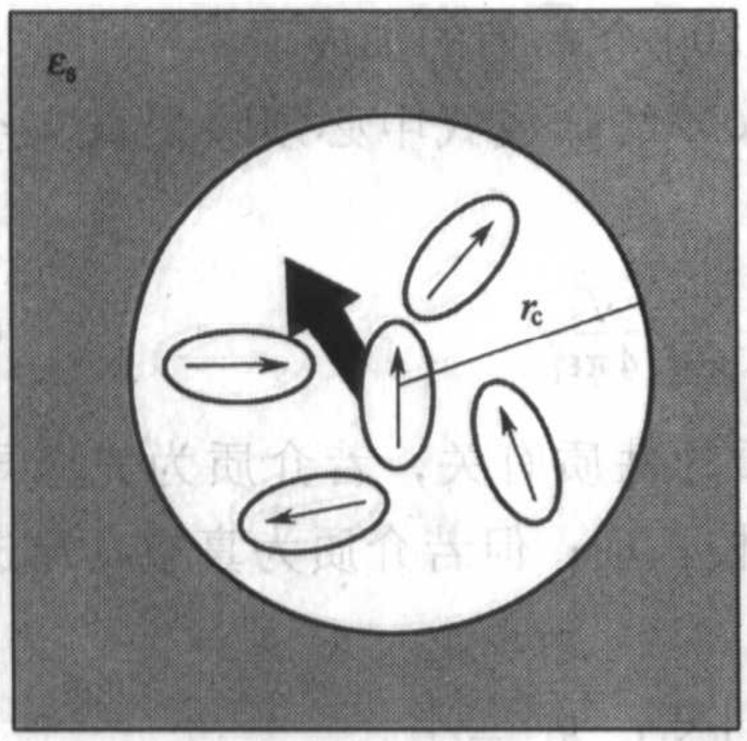



**图 9-7　作用场方法示意**

依照作用场方法，均匀介质所形成的电势为：

$$\vec{E}_i = \frac{2(\epsilon_s - 1)}{\epsilon_s + 1}\left(\frac{1}{r_c^3}\right)\sum_{j, r_{ij} \leqslant r_c} \vec{\mu}_j \tag{9-24}$$

式中，$\vec{\mu}_j$ 为在作用半径内，围着 $i$ 分子的偶极，图 9-7 中的黑色箭号表示所有这些电偶极的和。因此，$i$ 分子与作用场间的作用势能为：

$$U_{field} = \vec{E}_i \cdot \vec{\mu}_i \tag{9-25}$$

作用场计算的缺点为于截断半径内的分子数目改变时，计算的能量或作用力会出现不连续的情形，通常利用开关函数控制靠近截断半径边界的分子电偶极以避免计算中产生这样的错误。

另外一种类似作用场的计算法称为镜像电荷（image charge）计算法。这种计算法以一个半径为 $R$ 的球形的界面包裹计算的系统，界面中电荷的位置若为 $\vec{r}_i$，则界面外于位置 $(R/r_i)^2\vec{r}_i$ 处镜像电荷的荷电量为：

$$q_{im} = -\frac{(\epsilon_s - \epsilon_r)}{(\epsilon_s + \epsilon_r)}\frac{q_i R}{r_i} \tag{9-26}$$

式中，$\epsilon_s$ 与 $\epsilon_r$ 为球面内与球面外的介电常数，如图 9-8 所示。

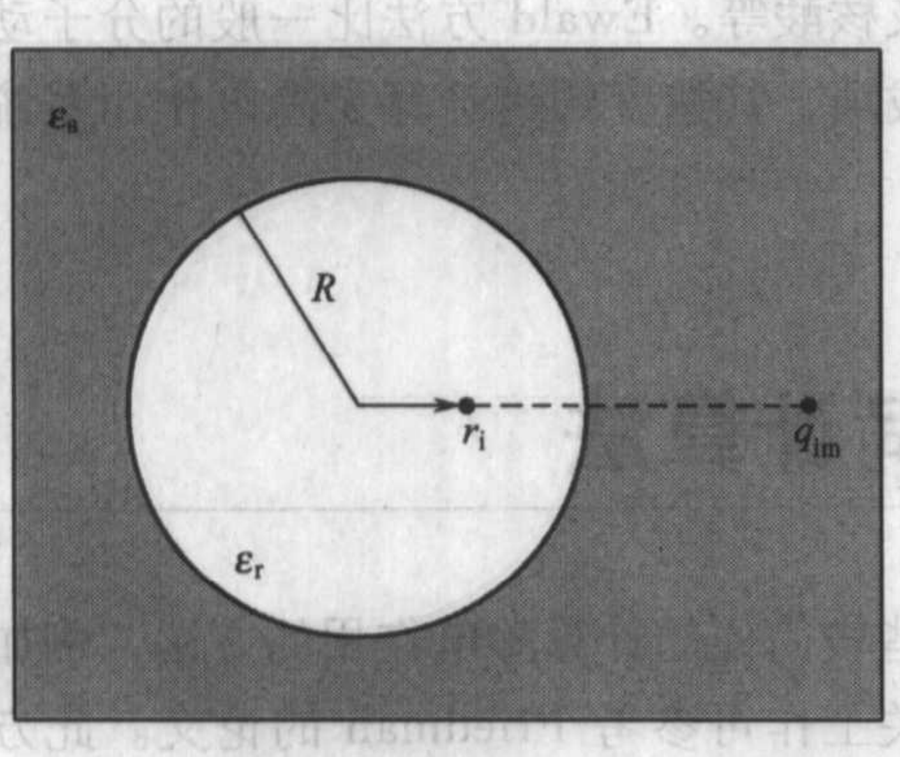



**图 9-8　镜像电荷系统示意**

此假设适用的条件为 $\epsilon_s \gg \epsilon_r$，其缺点为当电荷离球的界面很近，即其镜像电荷离球面很近时，此假设不再适用。

作用场方法与镜像电荷法均将截断半径以外的系统视为均匀介质，简化了计算工作。此计算方法中，通常最重要的是必须符合 $\epsilon_r \leqslant \epsilon_s \leqslant \infty$ 的条件。介电常数通常可由计算系统的平均总偶极矩 $\langle M^2 \rangle$ 求得。由作用场理论可推得：

$$\frac{4\pi}{9}\frac{\langle M^2\rangle}{Vk_BT}=\frac{\epsilon_r-1}{3}\frac{2\epsilon_s+1}{2\epsilon_s+\epsilon_r} \tag{9-27}$$

式中，$V$ 为计算系统的体积。以上此节所讨论的方法，均为非常简化的计算法，主要的用处为计算水溶液系统。其结果通常仅能定性地反映出系统中溶质分子的运动情形，而不能定量地与实验对比。但其优点在于能快速的得到一些结果，因此在适当的状况下，仍值得采用。

## 9.4 非键结作用项的格胞多极计算方法

格胞多极法（cell multipole method，CMM），又称为快速多极法（fast multipole method）。此方法的优点在于能将所有 $N(N-1)$ 成对作用项简化为 $N$ 项的作用。可简化的加成项的形式如：

$$\sum_i \sum_{j>i} \frac{q_i q_j}{|\vec{r}_i-\vec{r}_j|^p} \tag{9-28}$$

因此，库仑静电作用与 LJ 作用项均可视为这样的形式。计算静电作用时 $p=1$；计算 LJ 作用时 $p=12$，$p=6$。格胞多极法将计算系统分为许多体积相同的格胞（cell），由格胞内所有的原子计算每一格胞的多极（电荷、偶极、四极……）。格胞与格胞间的作用则可利用中央多极展开法计算。因中央多极展开法的条件为格胞间的距离必须大于一定的距离，故距离大于一个格胞以上的以多极展开法计算格胞间的作用；而相邻的格胞间的作用则以一般的成对原子方法计算。

如图 9-9 所示，中心的黑色区域为格胞 $C_0$，$C_0$ 中的原子与其自身及其邻近格胞（$C_n$，$C_f$）中原子的作用依一般成对作用计算。$C_0$ 与其余格胞间的作用则以多极展开方法计算。

在 CMM 方法中，每一格胞与格胞外原子的作用可利用多极展开法写成：

$$U_A^{pole}(r)=\frac{Z}{R^p}+\frac{\mu_\alpha R_\alpha}{R^{p+2}}+\frac{Q_{\alpha\beta}R_\alpha R_\beta}{R^{p+4}}+\frac{O_{\alpha\beta\gamma}R_\alpha R_\beta R_\gamma}{R^{p+6}}+\cdots \tag{9-29}$$

式中，$\vec{R}=\vec{r}-\vec{r}_A$，$\vec{r}_A$ 为格胞中心的位置向量，$\vec{r}$ 为格胞外任一原子的位置向量，$\alpha=x,y,z$。式中所有的项均为隐含 $\sum\alpha\alpha$ 等的加成项。最低次的零极项为 $Z=\sum_i q_i$，偶极项为 $\mu_\alpha=p\sum_i q_i r_{i\alpha}$，四极项为 $Q_{\alpha\beta}=$

$\sum_i q_i \frac{p}{2}[(p+2)r_{i\alpha}r_{i\beta} - \delta_{\alpha\beta}r_i^2]$，八极项为 $O_{\alpha\beta\gamma} = \sum_i q_i \frac{p(p+2)}{6}[(p+4)r_{i\alpha}r_{i\beta}r_{i\gamma} - (r_{i\alpha}\delta_{\beta\gamma} + r_{i\beta}\delta_{r\alpha} + \gamma_{i\gamma}\delta_{\alpha\beta})r_i^2]$。此处，$r_{i\alpha}$为 $i$ 原子至格胞 A 的中心的位置向量的 $\alpha$ 分量。CMM 展开项中通常可将作用分为近程作用与远程作用两类。如图 9-9 中的格胞 $C_0$ 其周围的 26 个 $C_0$ 格胞为近程作用，除此以外的所有其他作用（如图中的 $C_f$ 格胞）为远程作用。近程的作用直接由式(9-28) 计算，远程作用则以式(9-29) 多极展开法计算。

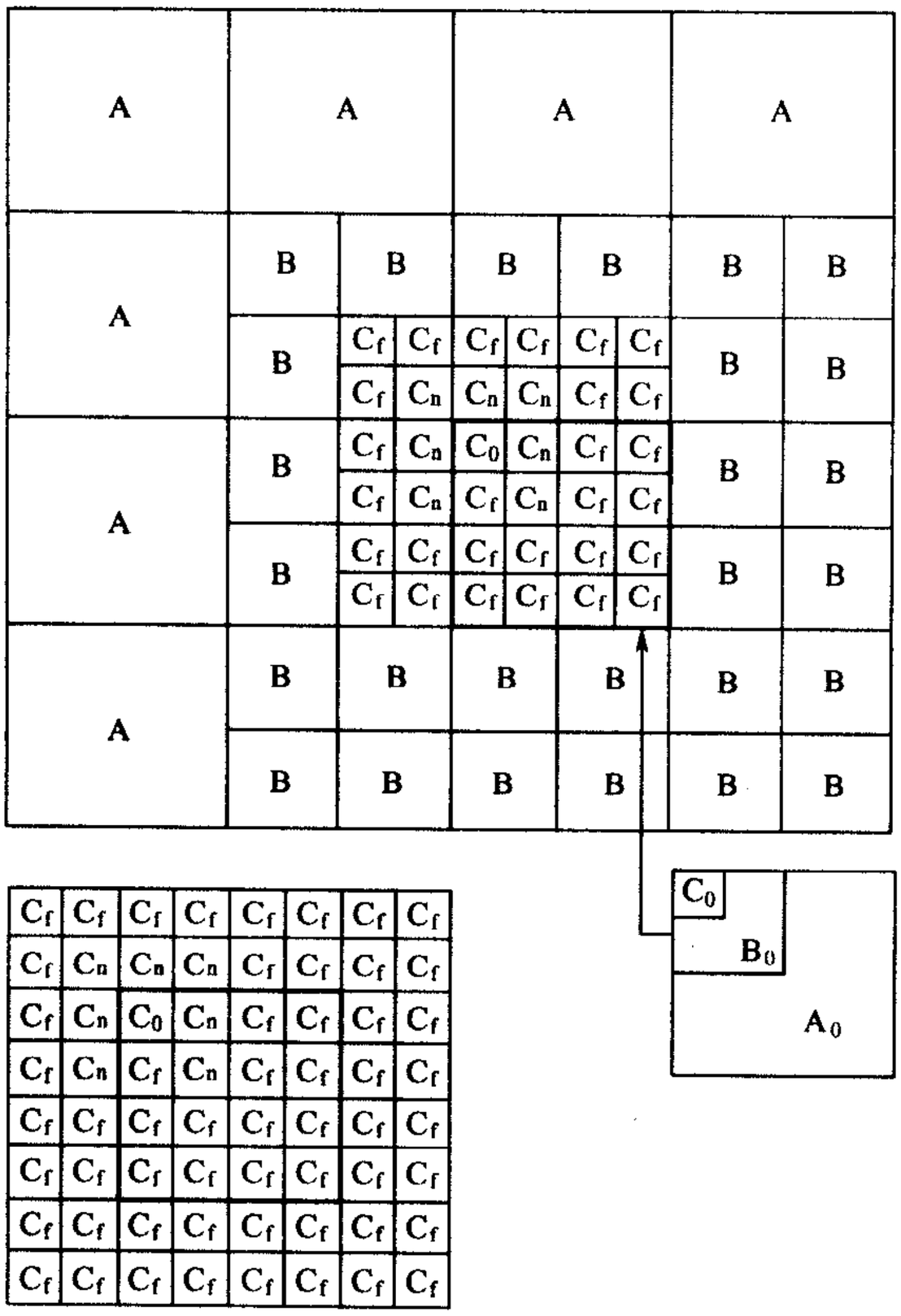



**图 9-9　格胞多极展开法示意图**

$$U(r_i) = U_{far}(r_i) + U_{near}(r_i) = \sum_{A \in far} U_A^{pole}(\vec{r}_i - \vec{r}_A) + \sum_{j \in near} \frac{q_j}{|\vec{r}_i - \vec{r}_j|^p} \tag{9-30}$$

在远程的格胞（$C_{far}$）中，格胞中原子与格胞中心的距离远小于格胞

中心间（$C_0-C_{far}$）的距离，因此可将作用项以泰勒级数展开为：

$$\sum_A U_A^{pole}(\vec{r}-\vec{r}_A)=U^{(0)}+U_\alpha^{(1)}r_\alpha+U_{\alpha\beta}^{(2)}r_\alpha r_\beta+\cdots\equiv U^T(\vec{r}) \quad (9\text{-}31)$$

将式(9-31）的系数可与式(9-29）中的系数的展开项相较，得零极项为：

$$\frac{Z}{|\vec{r}-\vec{r}_A|^p}=\frac{Z}{r_A^p}+\frac{pZr_A}{r_A^{p+2}}+\cdots \quad (9\text{-}32)$$

偶极项为：

$$\frac{\mu\cdot(\vec{r}-\vec{r}_A)}{|\vec{r}-\vec{r}_A|^{p+2}}=\frac{-\mu\cdot\vec{r}_A}{r_A^{p+2}}+\left[\frac{\mu}{r_A^{p+2}}-\frac{(p+2)(\mu\cdot\vec{r}_A)\vec{r}_A}{r_A^{p+2}}\right]\cdot\vec{r}+\cdots \quad (9\text{-}33)$$

余类推。故得 $U^{(0)}$ 为所有如式(9-32)、式(9-33）的展开项的常数项的和，$U^{(1)}$ 则为展开项的线性项的和。

由上述的方法，基本的格胞通常选取为含有 4 个原子的小区域（C区），对 $C_0$ 而言，每个较远格胞对 $C_0$ 的作用势能可以泰勒级数对 $C_0$ 的中心展开。若计算的系统共分成 $M$ 个格胞，则每个格胞共有 $M-27$ 个远格胞的计算，整个系统共需计算 $M\times(M-27)$ 个格胞-格胞的远程作用。以每个格胞含 4 个原子计算，设系统共含有 $N$ 个原子，则共需计算 $N^2/16$ 个作用项，即作用项的数目与 $N^2$ 成正比。

如上的计算，远程的格胞与近处的格胞计算作用项的方法相同。事实上，依照泰勒展开式，愈远的格胞展开式所导致的误差愈小。因此，依其与 $C_0$ 的距离将小格胞集合成较大的格胞，距离愈远集合的格胞愈大，如图 9-9 中的 A 与 B。若集合格胞的大小与离 $C_0$ 距离的比例相同，则以泰勒展开式计算所得的误差亦大致相同。格胞的多极可由其组成小格胞的多极矩加成而成。多极展开法与泰勒展开式的准确性与其截断项（truncation）有关，因此大的格胞必须在多极展开式中引用较多的电极。由于格胞多极计算法需要不断地检查格胞的大小及位置，实际计算时这种检查工作需要耗费许多计算时间，对某些粒子数目的系统，其计算效率反不如直接计算 $N^2$ 个作用项。但自系统粒子数目约为 300～3000 起，其计算速度即降至与 $N$ 成正比。对含数千个粒子的系统，或只要计算的系统不含数百万个粒子，格胞多极展开法为非常有效的计算方法。图 9-10 为 Ding 等利用格胞多极展开法比较不同数目的水分子所得的计算结果。

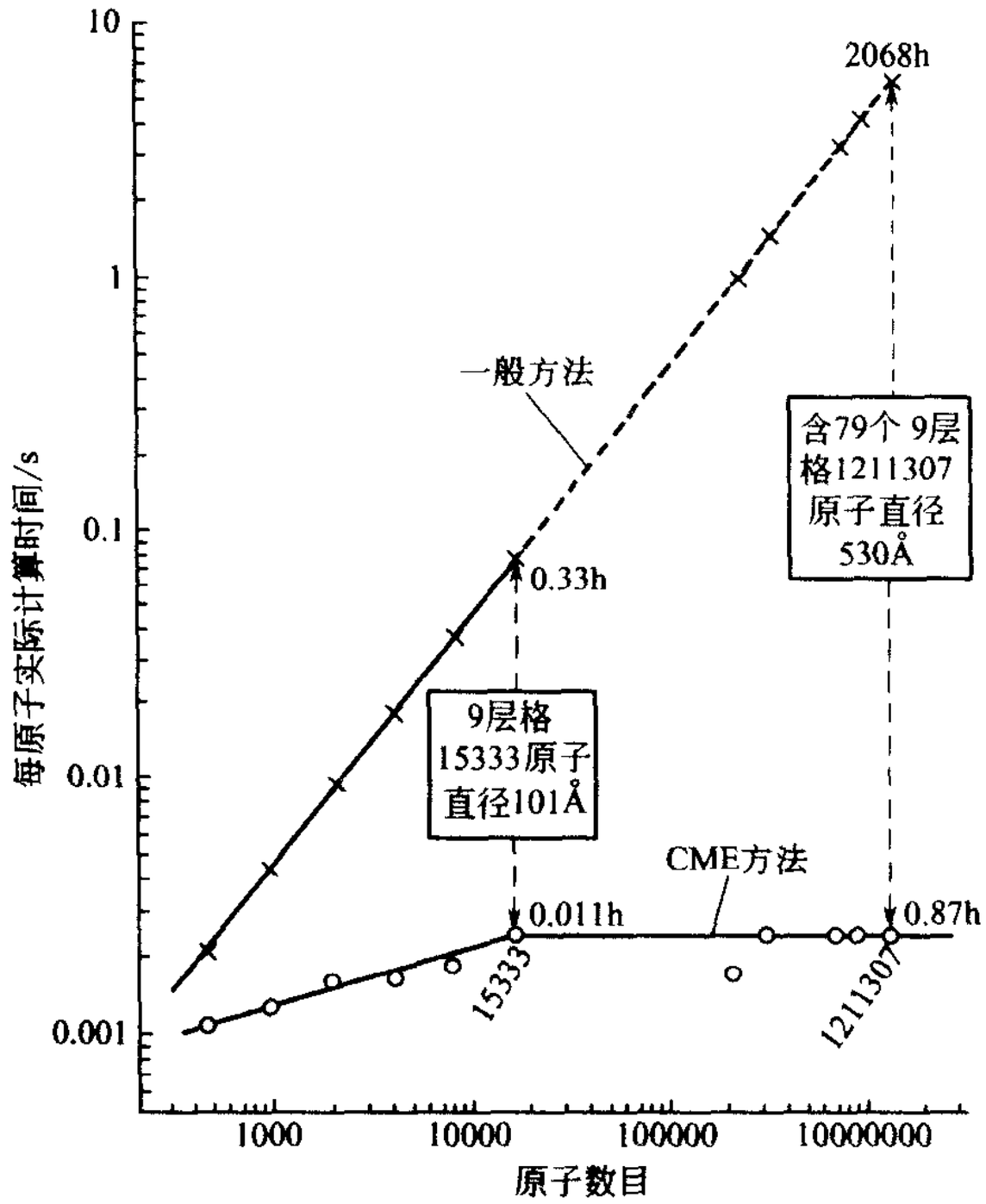



图 9-10　格胞多极展开法计算水系统的比较

# 9.5　溶剂介电质模型

前述各类关于库仑作用的能量，作用力与势能的表示式均为带电质点在真空穿透系数 $\epsilon_0$（permittivity in vacuum）的条件下的作用。意即，对两个带电质点 $q_i$，$q_j$ 其库仑作用的势能为：

$$U(r)=\frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{r} \tag{9-34}$$

式中，$r$ 为两个质点的距离。但对于许多庞大的系统，执行全原子计算往往超出一般计算机的负荷量。因此，若能简化这样的作用能可以节省许多的计算量。Smith 与 Pettitt 提出所谓的溶剂介电模型（solvent dielectric model）以简化溶液系统中带电质点间的相互作用。在此模型中，设介电常数 $\epsilon$（dielectric constant）为一有效介电常数 $\epsilon_{eff}$，其大小与质点

的距离相关。距离愈近，溶剂分子的效应愈低，介电常数值与真空介电常数值愈相近。距离愈远，溶剂分子的效应愈显著，介电常数近似于溶液体系介电常数（bulk dielectric constant）。Smith 等建议以式(9-35) 表示距离相关的介电常数。

$$\epsilon_{\text{eff}}(r)=\epsilon_{\text{r}}-\frac{\epsilon_{\text{r}}-1}{2}[(rS)^2+2rS+2]\text{e}^{-rS} \tag{9-35}$$

式中，$\epsilon_{\text{r}}$ 为溶液体系的介电常数；$\epsilon_{\text{eff}}$ 的最小值为 1。$S$ 为参数，随溶液不同而异，通常选取值为 0.15～0.3Å。图 9-11 为水溶液的有效介电常数值的变化。图中显示当距离小于 2Å 时有效介电常数的值趋近于 1，当距离大于 30Å 时有效介电常数值趋近于一般水溶液的介电常数值 80。

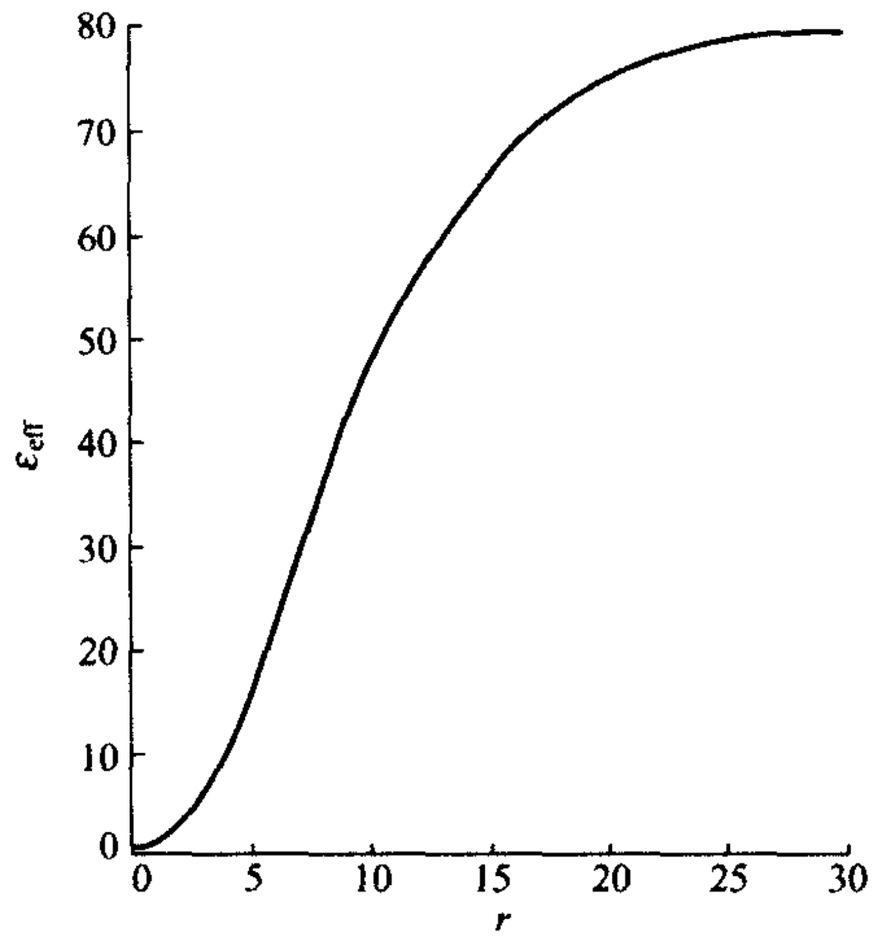



**图 9-11　水溶液中有效介电常数与距离的关系**

这种距离相关的介电常数模型适用于像水溶液等介电常数很大的系统。因为介电常数愈大，带电质点间的库仑作用力愈小，远程作用力愈不重要。例如分别带 $\pm 0.5e$ 的两个质点，当其在水中的距离为 30Å 时，其库仑作用势仅为：

$$U(r)=332.0\times\frac{1}{80}\times\frac{(0.5)(-0.5)}{30}=-0.0345\ \text{(kcal/mol)}$$

因此，利用这种方法可以大量简化系统中远程作用力的计算。如图 9-10 中显示，当距离大于 15Å 时，有效介电常数值已趋近于 70，此时带电质点间的库仑作用力已非常微弱而不重要了。但因为此种方法较缺乏理

论的依据，除非系统太过庞大，否则应尽量避免采用。

## 参 考 文 献


1 Ewald P. Ann. Phys., 64, 253, 1921

2 Friedman H L. Mol. Phys., 29, 1533, 1975

3 Ding H Q, Karasawa N, Goddard Ⅲ W A. Chem. Phys. Lett., 196, 6, 1992

4 Ding H Q, Karasawa N, Goddard Ⅲ W A. J. Chem. Phys., 97, 4309, 1992

5 Smith P E, Pettitt B M. J. Phys. Chem., 98, 9700, 1994

# 第10章 蒙地卡罗计算方法

第3～9章介绍了分子力学与分子动力模拟两种计算方法，分子力学依据力场及能量最小化法则，分子动力模拟依据牛顿运动力学法则。本章介绍的是另一种常见模拟方法：蒙地卡罗计算法，以随机取样为主，利用简化的力场进行模拟。该方法的发展始自20世纪50年代，迄今已50余年，常用以计算液体分子系统、合金系统、聚合物系统、生化系统等，应用十分广泛。本章介绍了蒙地卡罗计算的基本原理以及各类系综的计算方法，其应用与计算实例将在第11、12章中讨论。

## 10.1 基本原理

所谓的蒙地卡罗（Monte Carlo，MC）计算泛指凡是在计算中引用随机数（random number）的计算法。下例为简单的蒙地卡罗计算，设欲求图10-1中以单位长度为半径的1/4圆的面积。

利用蒙地卡罗方法计算此面积，可随意选取范围为（$0 \leqslant x \leqslant 1$，$0 \leqslant y \leqslant 1$）的一组随机数点 $P(x_i, y_i)$，若此点符合 $y_i \leqslant \sqrt{1-x_i^2}$ 的关系，即：

$$\sqrt{x_i^2+y_i^2} \leqslant 1 \tag{10-1}$$

则 $P$ 点为被接受的点，反之则被舍弃。此相当于向图10-1中曲线内的标靶射镖，若射中标靶则被接受，反之则被舍弃。设共随意选取了 $N$ 个随机点，其中有 $n$ 点被接受，则所求的1/4圆的面积为 $P$ 点落于1/4圆区间之概率乘以矩形的面积，即：

$$A = \frac{n}{N} \times 1^2 \tag{10-2}$$

所求的面积的真正值为 $\pi(1)^2/4 = 0.785398$。表10-1列出以这种随意选取

点的方法计算面积的结果。

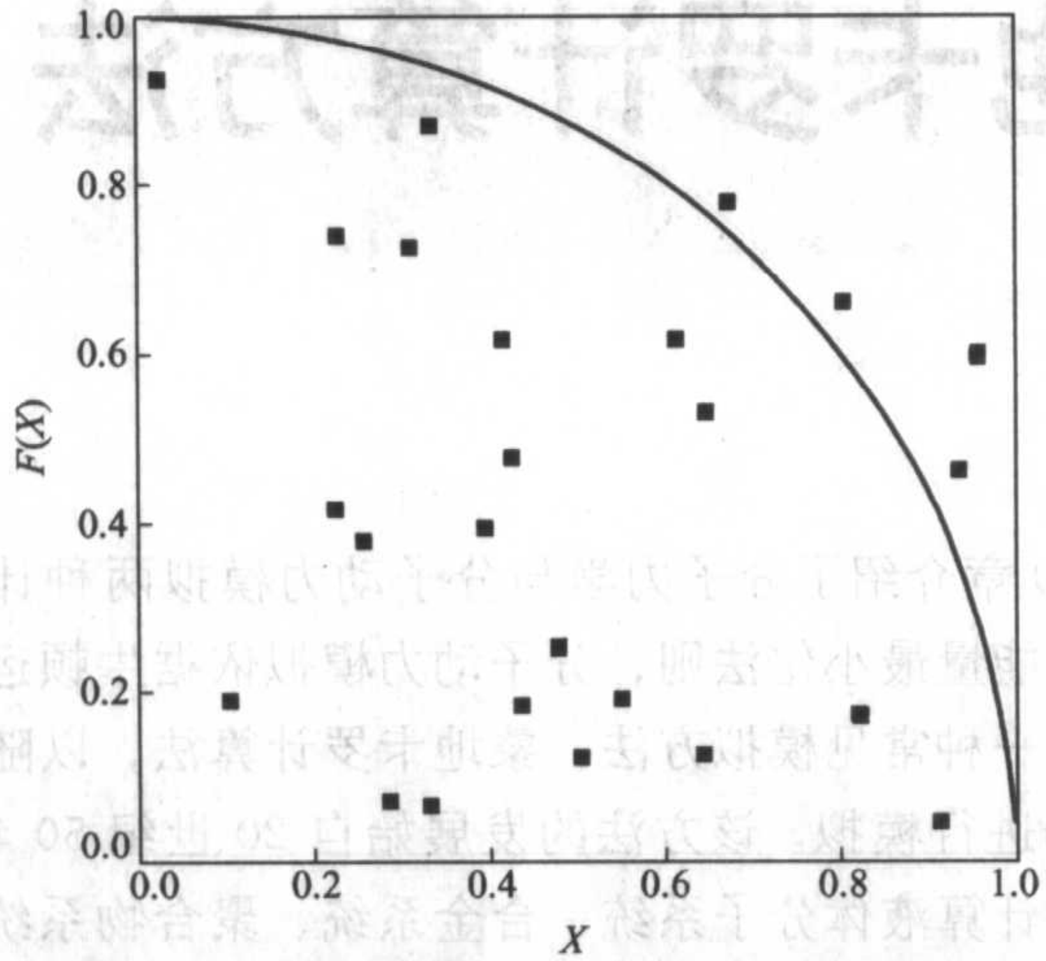



**图 10-1　向 1/4 单位圆随意射镖的示意**

**表 10-1　以蒙地卡罗方法计算 1 / 4 单位圆面积的结果**

| 选取的随机数数目 | 随机数的平均值 | 计算的面积 |
|---|---|---|
| 100 | 0.46431 | 0.8600 |
| 400 | 0.48547 | 0.7980 |
| 2000 | 0.50020 | 0.7888 |
| 2500 | 0.49945 | 0.7830 |

由表中的数据可知，当所选的随机数数目愈多，则计算的结果愈接近真实的值。此外，随机数的平均值愈接近 0.5，则计算的面积愈接近真实的面积。以上的问题相当于利用蒙地卡罗方法求式(10-3) 的积分。

$$A = \int_0^1 \sqrt{1-x^2}\,\mathrm{d}x \tag{10-3}$$

将这种方法推广，可应用于处理多重积分的问题。若一般的多重积分不易得到数理值时，可用蒙地卡罗方法计算：

$$I = \int_0^1\int_0^1 \cdots \int_0^1 f(x_1, x_2, \cdots x_N)\mathrm{d}x_1\,\mathrm{d}x_2 \cdots \mathrm{d}x_N \tag{10-4}$$

计算此积分的蒙地卡罗方法为在各 $x_i$ 可能的区间内（0→1）任意选取 $n$ 组点 $\{x_1^{(i)}, x_2^{(i)}, \cdots, x_N^{(i)}\}$，计算其对应的函数 $f$ 的值。则此积分可近似为：

$$I = \frac{1}{n}\sum_{i=1}^{n} f[x_1^{(i)}, x_2^{(i)}, \cdots, x_N^{(i)}] \tag{10-5}$$

这种计算积分方法的误差约为 $n^{-1/2}$，故选取的点数愈多，$n$ 值愈大，误差愈小。

考虑如下的积分式：

$$F = \int_{x_1}^{x_2} f(x)\mathrm{d}x \tag{10-6}$$

可将此式改写为：

$$F = \int_{x_1}^{x_2} \left[\frac{f(x)}{p(x)}\right] p(x)\mathrm{d}x \tag{10-7}$$

式中，$p(x)$ 为一种概率的分配函数。设在范围 $[x_1, x_2]$ 间依 $p(x)$ 的分配选取 $n$ 个随机数 $\xi_x$，则依照统计学积分式(10-7) 可写成：

$$F = \left\langle \frac{f(\xi_x)}{p(\xi_x)} \right\rangle \tag{10-8}$$

式中，〈〉为选取 $n$ 次随机数的平均值。利用此方法可计算许多统计力学中的积分，例如在正则系综中，任一特殊坐标的函数 $F$ 的平均值为：

$$\bar{F} = \int\cdots\int F(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) p(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) \mathrm{d}\vec{r}_1 \mathrm{d}\vec{r}_2 \cdots \mathrm{d}\vec{r}_N \tag{10-9}$$

其中：

$$p(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) = \frac{\exp\left[-\dfrac{U(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N)}{k_B T}\right]}{\int\cdots\int \exp\left[-\dfrac{U(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N)}{k_B T}\right] \mathrm{d}\vec{r}_1 \mathrm{d}\vec{r}_2 \cdots \mathrm{d}\vec{r}_N} \tag{10-10}$$

为系统于特殊结构$(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N)$的概率。定义结构积分(configuration integral)为：

$$Z_N = \int\cdots\int \exp\left[-\frac{U(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N)}{k_B T}\right] \mathrm{d}\vec{r}_1 \mathrm{d}\vec{r}_2 \cdots \mathrm{d}\vec{r}_N \tag{10-11}$$

则可将式(10-9)化为：

$$\langle F \rangle = \frac{\int\cdots\int F(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) \exp\left(-\dfrac{U}{k_B T}\right) \mathrm{d}\vec{r}_1 \mathrm{d}\vec{r}_2 \cdots \mathrm{d}\vec{r}_N}{Z_N} \tag{10-12}$$

因此，理论上，若任意选取大量的结构点，则可依蒙地卡罗方法计算式(10-9) 之积分。其计算的步骤如下所述。

① 随机产生系统中 $N$ 个原子的坐标。

② 计算此构型（configuration）的势能 $U(\vec{r}^N)$ 及函数 $F$。

③ 计算玻耳兹曼因子 $\exp\left(-\frac{U}{k_B T}\right)$。

④ 累加玻耳兹曼因子，并将函数与玻耳兹曼因子相乘后累加，回复步骤①。

⑤ 当计算 $N_t$ 循环后，求出 $F$ 的平均值为：

$$\langle F\rangle = \frac{\sum_{i=1}^{N_t} F_i \cdot \exp[-U_i(\vec{r}^N)/k_B T]}{\sum_{i=1}^{N} \exp[-U_i(\vec{r}^N)/k_B T]} \tag{10-13}$$

虽然理论上可以用蒙地卡罗方法计算统计的平均值，但若引用此法计算如液体系统的积分时必须选取大量的样品点，而这些样品点所对应的结构多数因为能量太高，因此出现的概率太低。此外，依式(10-11)若需计算样品点出现的概率必须先计算式中的结构积分，但实际上此积分值不容易收敛。基于这些原因，事实上直接以上述的蒙地卡罗方法计算系统的平均值是不恰当的。为了提高蒙地卡罗方法的效率，Metropolis 于 1953 年发展出选取重要取样法（importance sampling）的计算技巧。

## 10.2 重要取样法

欲引用随机数计算统计平均值，通常采用重要取样法。例如欲求取一 $(N,V,T)$ 系综的某一物理量 $A$ 的平均值，统计力学算式为：

$$\langle A\rangle = \int d\vec{r}^N \rho_{NVT}(\vec{r}^N) A(\vec{r}^N) \tag{10-14}$$

式中，$\rho_{NVT}$ 为概率；$\vec{r}^N$ 为对应的结构。由统计力学可知有的结构出现的概率大（重要，trial），有的结构其出现的概率小（不重要）。重要取样法计算平均值时设法选择重要的 $\rho(\vec{r}^N)$，使所求的积分仅为这些具有重要贡献的取样结构的和，即：

$$\langle A\rangle_{NVT} = \langle A\rangle_{trial} = \sum_{i=1}^{n} \rho(\vec{r}_i^N) A(\vec{r}_i^N) \tag{10-15}$$

式中，$n$ 为重要取样点的数目。Metropolis，Rosenbluth，Teller 等于 1953 年最先提出选择重要取样点的方法，此方法现今称为 Metropolis 法，为蒙地卡罗计算方法的基础。Metropolis 方法的重心在于建立系统的马克

夫链（Markov chain），使最终合于重要取样点的要求。所谓马克夫链需合于下列两个条件。

① 设系统各状态所成的有限集合为$\{\vec{r}_1^N, \vec{r}_2^N, \cdots, \vec{r}_m^N, \vec{r}_n^N, \cdots\}$，则每次由选择的随机数取样点所产生的状态$\vec{r}^N$，必为此集合的一个状态。

② 每次所选的取样点仅与其前一个取样点有关，而与由何处开始取样无关。

马克夫链属于一种遍历链（ergodic chain），即最终状态皆可由任一起始状态达到。兹以下例说明马克夫链的应用，假设你拥有的计算机其运行状态符合如下的规则：若正常运转一天则次日正常运转的概率为60%，若该日当机则次日接连当机的概率为70%。考虑此问题可将计算机分为正常运转（状态1），与当机（状态2）两种状态。由一种状态$i$转换为另一种状态$j$的概率为$\pi_{ij}$，$\pi_{11}=0.6$为由状态1转换为状态1的概率，即第一天为正常第二天亦正常的概率；$\pi_{22}=0.7$代表第一天当机第二天亦当机的概率；$\pi_{12}$代表第一天正常第二天当机的概率，$\pi_{12}$代表第一天当机第二天正常的概率。不论第一天计算机的状态为何，第二天转换为状态1与状态2的概率总和必为1，即符合：

$$\begin{aligned}\pi_{11}+\pi_{12}&=1\\ \pi_{21}+\pi_{22}&=1\end{aligned}\tag{10-16}$$

将$\pi_{11}$与$\pi_{22}$的值代入式(10-16)，得$\pi_{12}=0.4$，$\pi_{21}=0.3$。依此规则可建立转换矩阵（transition matrix）为：

$$\tilde{\pi}=\begin{pmatrix}\pi_{11} & \pi_{12}\\ \pi_{21} & \pi_{22}\end{pmatrix}=\begin{pmatrix}0.6 & 0.4\\ 0.3 & 0.7\end{pmatrix}\tag{10-17}$$

设$\rho_1$表示系统于状态1的概率，$\rho_2$表示系统处于状态2的概率，则系统的状态概率可用一个概率向量$\vec{\rho}$表示：

$$\vec{\rho}=(\rho_1,\rho_2)\tag{10-18}$$

设$\vec{\rho}^{(1)}$为系统起始状态的概率向量，$\vec{\rho}^{(2)}$为系统经过第一次转换后的概率向量，则：

$$\vec{\rho}^{(2)}=\vec{\rho}^{(1)}\tilde{\pi}\tag{10-19}$$

再经一次转换的概率向量为：

$$\vec{\rho}^{(3)}=\vec{\rho}^{(2)}\tilde{\pi}\tag{10-20}$$

余类推。若计算机第一天正常运转与当机的概率相同，则，$\vec{\rho}^{(1)}=(0.5,0.5)$，依式(10-19)第二天计算机状态的概率向量为：

$$\vec{\rho}^{(2)}=\vec{\rho}^{(1)}\times\hat{\pi}=(0.45,0.55)\tag{10-21}$$

意即，计算机于第二天正常运转的概率为 45%，当机的概率为 55%。第三天计算机状态的概率向量为：

$$\vec{\rho}^{(3)}=\vec{\rho}^{(2)}\times\hat{\pi}=\vec{\rho}^{(1)}\times\hat{\pi}\times\hat{\pi}=\vec{\rho}^{(1)}\hat{\pi}^2=(0.435,0.565)$$

经过许多天后计算机状态的概率向量的值达到一极限值 (limiting value)：

$$\vec{\rho}^{\infty}=\lim_{\tau\to\infty}\vec{\rho}^{(1)}\hat{\pi}^{\tau},\ \vec{\rho}^{\infty}=(0.4286,0.5714) \tag{10-22}$$

若概率向量已达其极限值，则再乘以转换矩阵必得到相同的概率向量，如：

$$\vec{\rho}^{\infty}\times\hat{\pi}=\vec{\rho}^{\infty}$$

$$(0.4286,0.5714)\begin{pmatrix}0.6 & 0.4\\ 0.3 & 0.7\end{pmatrix}=(0.4286,0.5714) \tag{10-23}$$

这表示当达极限情形后，状态的概率分配值不再变化，系统已达平衡 (equilibrium)。此极限概率向量的概率值又称为平衡状态的概率分配值。

以上的例子可推广为具有 $N$ 种可能状态的系统，则 $\vec{\rho}$ 为 $N$ 维向量，转换矩阵 $\tilde{\pi}$ 为 $N\times N$ 矩阵。$\tilde{\pi}$ 矩阵每列元素的和均为 1，即：

$$\sum_{n}\pi_{mn}=1 \tag{10-24}$$

满足极限概率分配值的条件为：

$$\sum_{m}\rho_{m}^{\infty}\pi_{mn}=\rho_{n}^{\infty} \tag{10-25}$$

实际应用上，设一系综含有 $N$ 种不同的状态（结构），则当其达到平衡时（即各状态的概率值不变），任何一次改变状态（结构）的转换，依然维持系综的平衡。

Metropolis 等所设计的蒙地卡罗方法，需建立一随机选择矩阵 (stochastic matrix)，此随机选择矩阵的元素以 $\alpha_{mn}$ 表示，为选择转换状态 $m\to n$ 的概率。若所选择的转换状态可被接受的概率为 $\rho_{mn}$，则由状态 $m$ 转换至状态 $n$ 的概率 $\pi_{mn}$ 为：

$$\pi_{mn}=\alpha_{mn}\rho_{mn} \tag{10-26}$$

通常 $\hat{\alpha}$ 为一对称矩阵，即 $\alpha_{mn}=\alpha_{nm}$，其意义为随机选择由 $m\to n$ 的概率等于随机选择由 $n\to m$ 的概率。设 $\rho_{m}^{\infty}$ 与 $\rho_{n}^{\infty}$ 分别表示状态 $m$ 与状态 $n$ 的极限概率，若已达极限平衡状态，且 $\rho_{n}^{\infty}>\rho_{m}^{\infty}$，则由 $m$ 转换为 $n$ 的概率等于随机选择 $m\to n$ 的概率，即 $\pi_{mn}=\alpha_{mn}$。若 $\rho_{n}^{\infty}<\rho_{m}^{\infty}$，则由 $m$ 转换为 $n$ 的概率为随机选择 $m\to n$ 的概率乘以 $n$ 状态的概率与 $m$ 状态的概率比。综合将以上的情形写为：

$$\pi_{mn}=\alpha_{mn}\qquad(\rho_{n}^{\infty}\geqslant\rho_{m}^{\infty}\qquad m\neq n)$$

$$\pi_{mn} = \alpha_{mn}(\rho_n^{\infty}/\rho_m^{\infty}) \qquad (\rho_n^{\infty} < \rho_m^{\infty} \quad m \neq n) \tag{10-27}$$

式(10-27) 符合微观可逆 (microscopic reversibility) 的条件，即：

$$\rho_m^{\infty}\pi_{mn} = \rho_n^{\infty}\pi_{nm} \tag{10-28}$$

状态无变化的转换概率为：

$$\pi_{mm} = 1 - \sum_{n \neq m}\pi_{mn} \tag{10-29}$$

根据以上的讨论，欲建立含有许多粒子系统的马克夫链则必须建立符合遍历性的转换矩阵。因多粒子系统有许多不同的状态，故其转换矩阵十分庞大。所幸由统计力学可得知概率的极限状态，对任一系统的结构 $\vec{r}_m^N$ 而言，其极限的概率为 $\rho_m^{\infty} = \rho_{NVT}(\vec{r}_m^N)$，故：

$$\frac{\rho_n^{\infty}}{\rho_m^{\infty}} = \frac{\exp[-U(\vec{r}_n^N)/k_B T]}{\exp[-U(\vec{r}_m^N)/k_B T]} = \exp(-\Delta U_{mn}/k_B T) \tag{10-30}$$

式中，$\Delta U_{mn}$ 为二状态 $n$ 与 $m$ 的势能差。由此极限概率的比值可决定符合式(10-27) 要求的转换矩阵。若 $n$ 状态的势能低于 $m$ 状态的势能，依式(10-27)，此结构的转换可被接受。若 $n$ 状态的势能高于 $m$ 状态的势能，则需依式(10-27) 的概率接受此转换移动。实际运用上，可将此转换的玻耳兹曼因子 $\exp(-\Delta U_{mn}/k_B T)$ 与一介于 0 与 1 间的随机数比较；若玻耳兹曼因子大于此随机数，则此转换可被接受 (accept)。反之，若玻耳兹曼因子小于此随机数，则排除 (reject) 此转换。此取样方法可以图 10-2 表示如下。

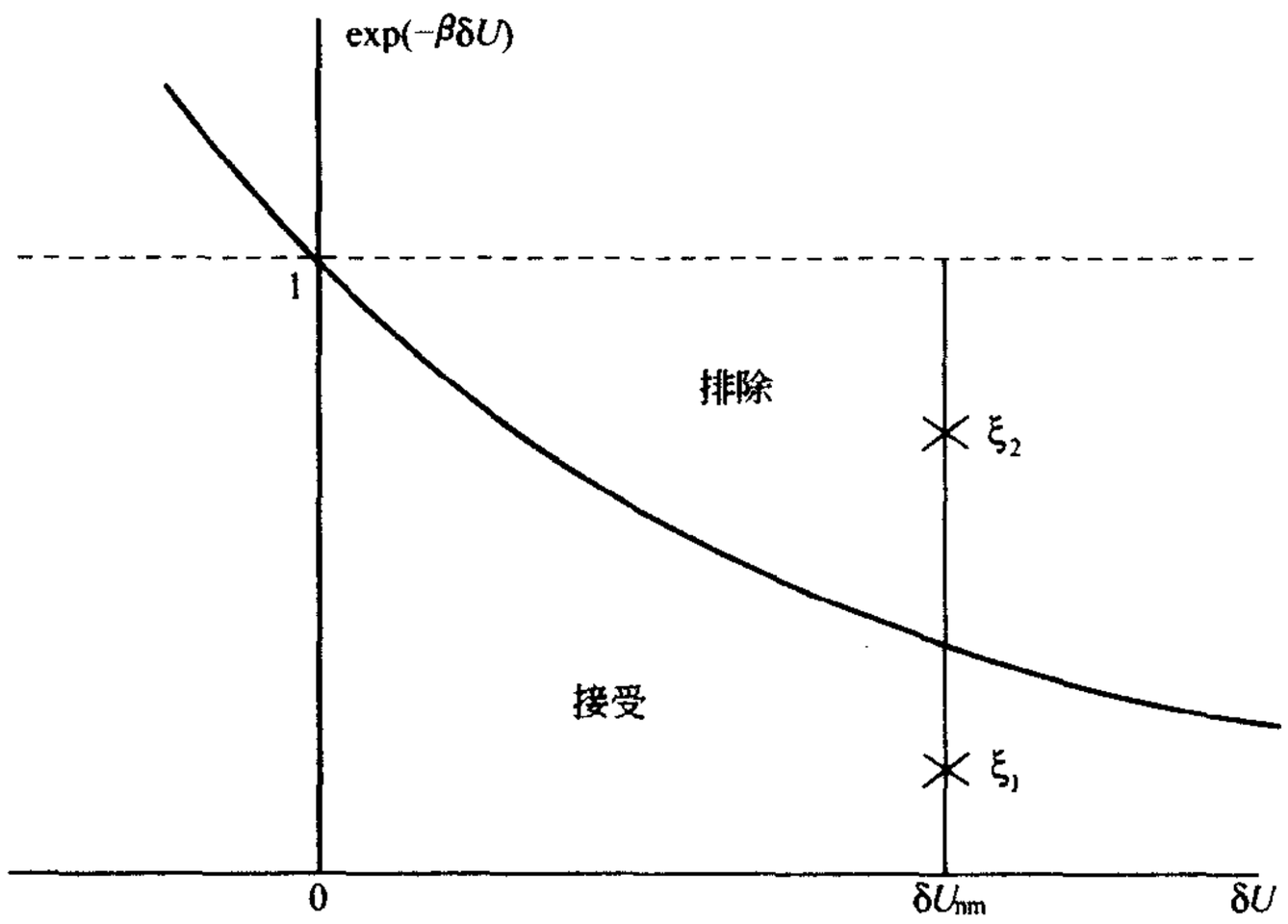



**图 10-2　Metropolis 的取样方法**

依图 10-2，若新状态 $n$ 的势能与旧状态 $m$ 的势能接近，玻耳兹曼因子的值接近 1，即使 $U_n > U_m$，此转换仍非常可能被接受。反之，若二状态的能差很大，则玻耳兹曼因子的值接近 0，此转换不大可能被接受。利用这样的筛选过程，将不可能的转换排除，即可达到重要取样的目的。依照 Metropolis 方法，可由系统的任意起始状态开始，由转换矩阵产生一系列的结构，或称产生一系列的相空间轨迹（phase space trajectories），而计算其所对应性质的平均值。若计算的平均值不再变化，则系统已达平衡状态。转换矩阵元素 $\pi_{mn}$ 依系综的性质而不同。下列各节将讨论如何决定这些元素及执行 Metropolis 方式的蒙地卡罗计算。

## 10.3 Metropolis 计算法

Metropolis 计算法在每次取样（产生随机数）时考虑其是否合于所设定的物理要求。如此可避免一些不“重要”的取样，以增加计算的效率。欲执行此计算需先决定对称的 $\alpha$ 矩阵，其要求为 $\alpha_{mn} = \alpha_{nm}$。以图 10-3 二度空间的系统为例，此系统含有数个粒子，处于状态 $m$。

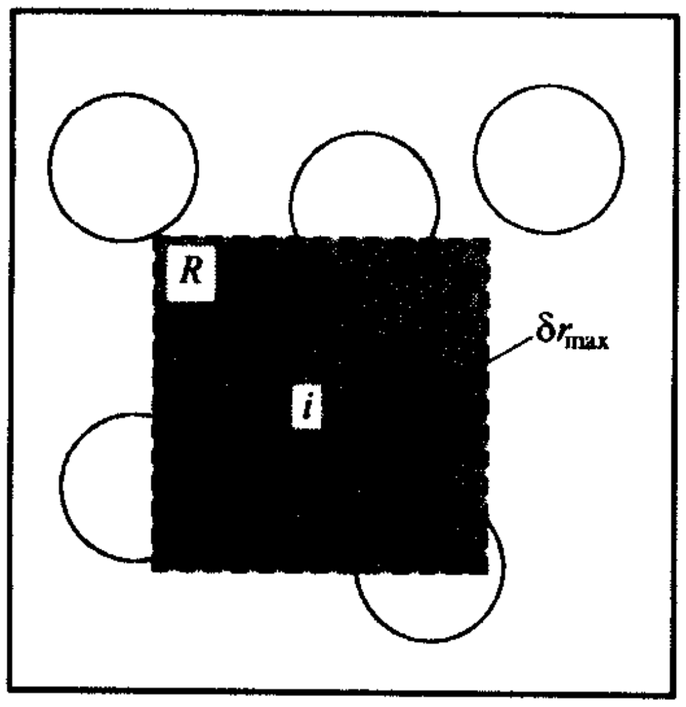



**图 10-3　随机选取粒子的移动**

产生另一状态的方法为随机选取系统中的某一粒子，$i$，将其由原来的位置，$\vec{r}_i^m$，以相等的概率移动至范围 $R$ 内的任意一位置，$\vec{r}_i^n$。范围 $R$ 的中心为 $\vec{r}_i^m$，边长为 $2\delta r_{max}$。若系统为三度空间的盒子则 $R$ 为一边长为 $2\delta r_{max}$ 的立方体。设这样的移动可产生 $N_R$ 个新的位置，即产生 $N_R$ 个新的状态，因此，选择 $m \to n$ 的概率为 $1/N_R$。可定义 $\alpha_{mn}$ 为：

$$\begin{aligned} \alpha_{mn} &= 1/N_R \qquad \text{若 } \vec{r}_i^n \in R \\ \alpha_{mn} &= 0 \qquad \text{若 } \vec{r}_i^n \notin R \end{aligned} \tag{10-31}$$

这样的定义即可满足式 $\alpha_{mn} = \alpha_{nm}$ 的要求。下一步骤为计算移动的玻耳兹曼因子。如图 10-4 所示，设选择的粒子 $i$ 由原先的 $\vec{r}_i^m$ 移动至新的位置 $\vec{r}_i^n$。设系统中粒子的作用仅为 LJ 作用，则：

$$\Delta U_{mn} = U(\vec{r}_n^N) - U(\vec{r}_m^N) = \left[\sum_{j=1}^{N} v_{LJ}(r_{ij}^n) - \sum_{j=1}^{N} v_{LJ}(r_{ij}^m)\right] \qquad (10\text{-}32)$$

换句话说，计算粒子移动前后系统的势能差，并不需要计算状态 $m$ 与 $n$ 所有的能量，仅需计算所有与 $i$ 粒子相关的能量变化即可。

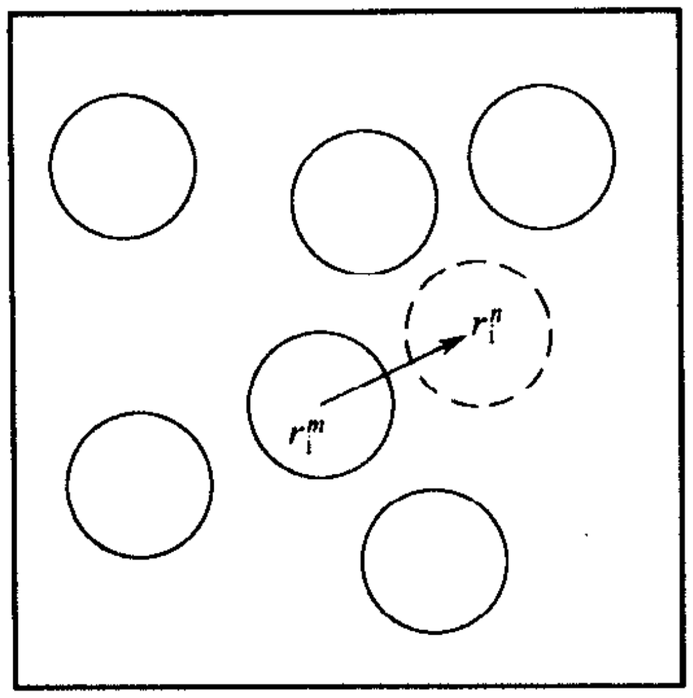



**图 10-4　蒙地卡罗粒子移动示意**

若 $\Delta U_{mn} \geqslant 0$，则将玻耳兹曼因子与随机数 $0 \leqslant \xi \leqslant 1$ 相比，以决定是否接受此粒子的移动。若玻耳兹曼因子大于此随机数，则此移动被接受，反之，则此移动被拒绝。若此移动被接受，则以此新位置为 $i$ 粒子的位置，否则回到原来的位置。再随意选择系统的任一粒子，并重复如上的移动步骤。这样的计算方法，即为 Metropolis 计算法。

对于含有 $N$ 个粒子系统的（$N$, $V$, $T$）系综，Metropolis 计算法的实际操作步骤如下。

① 任意挑选一个粒子，产生任意方向的位移。位移的大小为一设定值范围内的随意值。设该粒子原坐标为（$x_{old}$, $y_{old}$, $z_{old}$），移动后的新坐标为（$x_{new}$, $y_{new}$, $z_{new}$）则：

$$x_{new} = x_{old} + \text{Rand}[-1,1] \times \delta r_{max}$$

$$y_{new} = y_{old} + \text{Rand}[-1,1] \times \delta r_{max}$$

$$z_{new} = z_{old} + \text{Rand}[-1,1] \times \delta r_{max}$$

式中，Rand[−1,1] 表示范围为−1～1 的随机数；$\delta r_{max}$ 为计算中所设定的最大位移。

② 设移动前系统的势能为 $U_{old}$，移动后系统的势能为 $U_{new}$，即：

$$\Delta U = \delta U_{nm} = U_{new} - U_{old}$$

若 $\Delta U \leqslant 0$，则此移动为可完全被接受的移动，被接受的概率为 1，此粒子的新坐标为（$x_{new}$, $y_{new}$, $z_{new}$）。这相当于式(10-28) 中 $\rho_n^\infty \geqslant \rho_m^\infty$ 的情况。若 $\Delta U > 0$，则相当于式(10-28) 中 $\rho_n^\infty < \rho_m^\infty$ 的情况，此种移动被接受的概率为 $\rho_n^\infty / \rho_m^\infty$。由统计力学此比值为：

$$\frac{\rho_n^{\infty}}{\rho_m^{\infty}}=\frac{Z_{NVT}^{-1}\exp\left(-\frac{U_n}{k_BT}\right)}{Z_{NVT}^{-1}\exp\left(-\frac{U_m}{k_BT}\right)}=\frac{\exp\left(-\frac{U_n}{k_BT}\right)\exp\left(-\frac{\delta U}{k_BT}\right)}{\exp\left(-\frac{U_n}{k_BT}\right)}=\exp\left(-\frac{\delta U_{nm}}{k_BT}\right) \tag{10-33}$$

式中，$k_B$ 为玻耳兹曼常数；$T$ 为系统的设定温度。$\Delta U>0$ 时，可另外产生一个范围为 0～1 的随机数 $p$，若 $\exp(-\Delta U_{nm}/k_BT)\geqslant p$，则此移动可被接受，其被接受的概率为 $\exp(-\Delta U_{nm}/k_BT)$，粒子的新坐标为（$x_{new}$，$y_{new}$，$z_{new}$）。反之，若 $\exp(-\Delta U_{nm}/k_BT)<p$，则该移动无效，粒子的坐标不变，为（$x_{old}$，$y_{old}$，$z_{old}$）。

这整个步骤可综合记为移动被接受的概率为：

$$\min\left[1,\exp\left(-\frac{\delta U_{nm}}{kT}\right)\right] \tag{10-34}$$

重复①和②的步骤，收集被接受的结构，并于每一被接受的结构计算一些热力或结构的物理量的平均值，如总能量、均方根位移、次序参数等。若上述这些物理量的平均值收敛，则谓产生可靠的平衡系统。利用这种方法可将许多不重要（造成过高能量状态）的结构省略，增加计算的效率。通常执行计算时，调整 $\delta r_{max}$ 的值，使所产生的随意移动约有 50%可被接受。

传统的 Metropolis 计算方法，每次仅选择一个移动粒子。Chapman 等于 1985 年提出以含有 32 个粒子的系统模拟液体的蒙地卡罗计算。计算中每次系统中的所有原子皆同时移动。在 $\delta r_{max}$ 约 $0.3\sigma$ 的条件下，移动的被接受率约为 30%，而此系统达成平衡的速度较选择单一移动粒子的计算系统快了许多。

## 10.4 各类系综的蒙地卡罗计算

### 10.4.1 定温定压系综

蒙地卡罗方法可依实际的需要，用于计算各种不同的系综。定温定压系综（isothermal-isobaric ensemble），（$N,P,T$）系综为分子数目，温度与压力皆为定值的系统集合。系统的体积在计算中可更改，达平衡时系统的平均密度趋于一定值。设（$N,P,T$）系综为设定压力下一含有 $N$ 个粒子的立方体盒子，任何物理量 $A$ 的平均值为：

$$\langle A\rangle_{\mathrm{NPT}}=\frac{\int_0^{\infty}\mathrm{d}V\exp\left(-\frac{PV}{k_{\mathrm{B}}T}\right)V^N\int\mathrm{d}\vec{s}\,A(\vec{s})\exp\left[\quad\frac{U(\vec{s})}{k_{\mathrm{B}}T}\right]}{Z_{\mathrm{NPT}}} \tag{10-35}$$

式中，$Z_{\mathrm{NPT}}$为对应于单位立方体的结构积分；$V$ 为系统的体积；$\vec{s}=(s_1,s_2,\cdots,s_{\mathrm{N}})$ 为一组度量化坐标（scaled coordinates），$\vec{s}=L^{-1}\vec{r}$。$L=V^{1/3}$为系统的边长。系统的任一结构，$\vec{s}$，其极限概率依照统计力学表示式为：

$$\exp\left[-\frac{PV+U(\vec{s})}{k_{\mathrm{B}}T}+N\ln V\right] \tag{10-36}$$

以 Metropolis 的方法产生（$N,P,T$）系综的马克夫链的计算步骤如下所述。

① 随意移动一个粒子的坐标，或随意改变系统的体积，或两种步骤皆进行，即：

$$\vec{s}_{\mathrm{i,new}}=\vec{s}_{\mathrm{i,old}}+\delta\vec{s}_{\max}(2\vec{\xi}-\vec{1})$$

$$L_{\mathrm{new}}=L_{\mathrm{old}}+\delta L_{\max}(2\xi-1)$$

式中，$\xi$为 1 与 0 间的随机数；$\vec{\xi}$ 为由此类随机数所组成的向量（$\xi_1,\xi_2,\cdots,\xi_n$），$\vec{1}=(1,1,\cdots,1)$。$\delta\vec{s}_{\max}$与 $\delta L_{\max}$控制向量化坐标与系统的盒长的最大改变量。一般计算中选择的 $\delta\vec{s}_{\max}$与 $\delta L_{\max}$值使新位置的接受率为 30%～50%。

② 产生新坐标后，计算系统焓的改变量 $\delta H_{\mathrm{nm}}$，即：

$$\delta H_{\mathrm{nm}}=\delta U_{\mathrm{nm}}+P(V_{\mathrm{new}}-V_{\mathrm{old}})-Nk_{\mathrm{B}}T\ln\frac{V_{\mathrm{new}}}{V_{\mathrm{old}}} \tag{10-37}$$

若 $\delta H_{\mathrm{nm}}\leqslant 0$，则此移动可完全被接受。若 $\delta H_{\mathrm{nm}}>0$，则依 Metropolis 法比较 $\exp(-\delta H_{\mathrm{nm}}/k_{\mathrm{B}}T)$ 与 0～1 间的随机数 $p$，以决定此粒子的移动或体积的改变是否被接受。将此综合记为粒子移动被接受的概率为：

$$\min\left[1,\exp\left(-\frac{\delta H_{\mathrm{nm}}}{k_{\mathrm{B}}T}\right)\right] \tag{10-38}$$

重复①、②的步骤，并计算系统体积、焓的平均值，以检验其是否收敛。由体积、焓的统计涨落（fluctuation）可计算一些与其相关的性质。例如，由计算系统的位力（virial）可得到系统的压力，此计算的压力应与设定的压力相同，通常由此检验蒙地卡罗计算的正确性。

### 10.4.2 巨正则系综

巨正则系综（grand canonical ensemble），$\Xi(V,T,\mu)$，为体积、温度与化学势能 $\mu$（chemical potential）皆为定值的系统集合。系统的粒子数为

可变化的参数。任一物理量 $A$ 的平均值的表示式为：

$$\langle A\rangle_{\mu VT}=\frac{\sum_{N=0}^{\infty}(N!)^{-1}V^{N}z^{N}\int d\vec{s}\,A(\vec{s})\exp\left[-\frac{U(\vec{s})}{k_{B}T}\right]}{Q_{\mu VT}} \tag{10-39}$$

式中，$z=\exp\left(\frac{\mu}{k_{B}T}\right)/\Lambda^{3}$，为系统的活性（activity）；$\Lambda=\left(\frac{h^{2}}{2\pi mk_{B}T}\right)^{1/2}$ 为得布罗意波长（de Broglie wavelength）；$h$ 为普朗克常数；$\vec{s}$ 为一组度量化的坐标。

$$Q_{\mu VT}=\frac{V^{N}}{N!\Lambda^{3N}} \tag{10-40}$$

为系统的配分函数（partition function）。依照式(10-40) 系统物理量的平均值与分布于状态的粒子数目 $N$ 有关。但此粒子数目并非一常数值，故式(10-39) 中为加成符号而非积分。系统的任一结构，$\vec{s}$，其极限概率为：

$$\exp\left[-\frac{U(\vec{s})-N\mu}{k_{B}T}-\ln N!-3N\ln\Lambda+N\ln V\right] \tag{10-41}$$

此类系统的计算最常用的为 Norman 与 Filinov 所发展的方法。以此方法产生马克夫链的计算步骤如下所述。

① 任意移动系统中的一个粒子，

$$\vec{s}_{i,new}=\vec{s}_{i,old}+\delta\vec{s}_{max}(2\vec{\xi}-\vec{1})$$

并依照一般的 Metropolis 方法决定此移动是否被接受。

② 任意消除系统中的一个粒子，消除前后系统的极限概率比为：

$$\frac{\rho_{n}^{\infty}}{\rho_{m}^{\infty}}=\exp\left(-\frac{\delta U_{nm}}{k_{B}T}\right)\exp\left(-\frac{\mu}{k_{B}T}\right)\frac{N\Lambda^{3}}{V} \tag{10-42}$$

将式(10-42) 转换为活性，得：

$$\frac{\rho_{n}^{\infty}}{\rho_{m}^{\infty}}=\exp\left(-\frac{\delta U_{nm}}{k_{B}T}+\ln\frac{N}{zV}\right)=\exp\left(-\frac{\delta D_{nm}}{k_{B}T}\right) \tag{10-43}$$

式中，定义 $\delta D_{nm}$ 为消除函数（destruction function）。若 $\delta D_{nm}\leqslant 0$，则此消除步骤可完全被接受。若 $\delta D_{nm}>0$，则依前比较 $\exp(-\delta D_{nm}/k_{B}T)$ 与 0～1 间的随机数 $p$，以决定此粒子的移动或体积的改变是否被接受。将此综合记为消除粒子被接受的概率为：

$$\min\left[1,\exp\left(-\frac{\delta D_{nm}}{k_{B}T}\right)\right] \tag{10-44}$$

③ 随意产生一个新的粒子，产生前后系统的极限概率比为：

$$\frac{\rho_{n}^{\infty}}{\rho_{m}^{\infty}}=\exp\left(-\frac{\delta U_{nm}}{k_{B}T}+\ln\frac{zV}{N+1}\right)=\exp\left(-\frac{\delta C_{nm}}{k_{B}T}\right) \tag{10-45}$$

式中定义 $\delta C_{nm}$ 为新生函数（generation function）。由此函数依 Metropolis 方法决定此步骤是否被接受。

重复①、②、③的步骤，并计算系统的构型能量（configuration energy）、压力及密度以判定计算的收敛性。此种蒙地卡罗方法可用以直接计算系统的自由能 $A$，即：

$$\frac{A}{N}=\mu-\frac{\langle P\rangle_{\mu VT}V}{\langle N\rangle_{\mu VT}} \tag{10-46}$$

并由自由能计算系统其他相关的统计性质。

上述计算方法中的消除与产生步骤，亦可由如下的比较式决定。设 $\xi$ 为 0～1 间的随机数。

若 $\left[1+\frac{N+1}{zV}\exp\left(\frac{\delta U_{nm}}{k_B T}\right)\right]^{-1}\geqslant\xi$，则产生步骤可被接受；

若 $\left[1+\frac{zV}{N+1}\exp\left(\frac{\delta U_{nm}}{k_B T}\right)\right]^{-1}\geqslant\xi$，则消除步骤可被接受。 (10-47)

化学势能在此系统中为一常数的输入值，Adams 将其分为理想气体的化学位能 $\mu^{id}$ 与超额化学势（excess chemical potential）$\mu^{ex}$ 两部分，系统的势能为此两部分势能的和，即：

$$\begin{aligned}\mu&=\mu^{ex}+\mu^{id}\\&=(\mu^{ex}+k_B T\ln\langle N\rangle_{\mu VT})+k_B T\ln\left(\frac{\Lambda^3}{V}\right)\\&=k_B TB+k_B T\ln\left(\frac{\Lambda^3}{V}\right)\end{aligned} \tag{10-48}$$

Adams 引用式(10-48) 执行在固定 $(B,V,T)$ 下的蒙地卡罗计算。计算中 $B$ 为一个人为输入的值，而由计算 $\langle N\rangle_{\mu VT}$ 可得到超化学位能 $\mu^{ex}$ 及化学位能 $\mu$。此种计算条件相当于固定 $(z,V,T)$ 的计算。

## 10.5 进阶蒙地卡罗计算法

在一般传统的蒙地卡罗计算中，每次仅随机选取一个粒子随意移动，因此系统中所有粒子的移动概率皆相等。但此种移动方式未必为最有效率的移动方式。例如在系统中常遇到某部分粒子的移动概率大于其他部分的情形，或是移动粒子时最好能使粒子的移动遵行其应该依循的方向。

### 10.5.1 优先取样蒙地卡罗计算法

优先取样蒙地卡罗（preferential sampling MC）计算法为 Owicki 与 Scherage 等所提出，为一种改进取样点使其更为有效的计算方法。以稀薄的离子水溶液为例，此系统中最重要的作用为溶质（离子）与溶剂（水）间的作用。溶剂间彼此的作用较不重要，溶质间彼此作用的重要性更低。若考虑溶质与溶剂的作用，则以第一媒合层（primary solvation shell）的溶剂分子与溶质分子的作用最为重要。在此层以外的溶剂分子，因其与溶质的距离较远，其重要性较低。逻辑上，处于此第一媒合层的溶剂分子应较层外的分子移动频繁，溶质分子的运动主要来自这些溶剂分子与溶质分子的作用。图 10-5 显示溶液中的溶质、溶剂分子的移动。

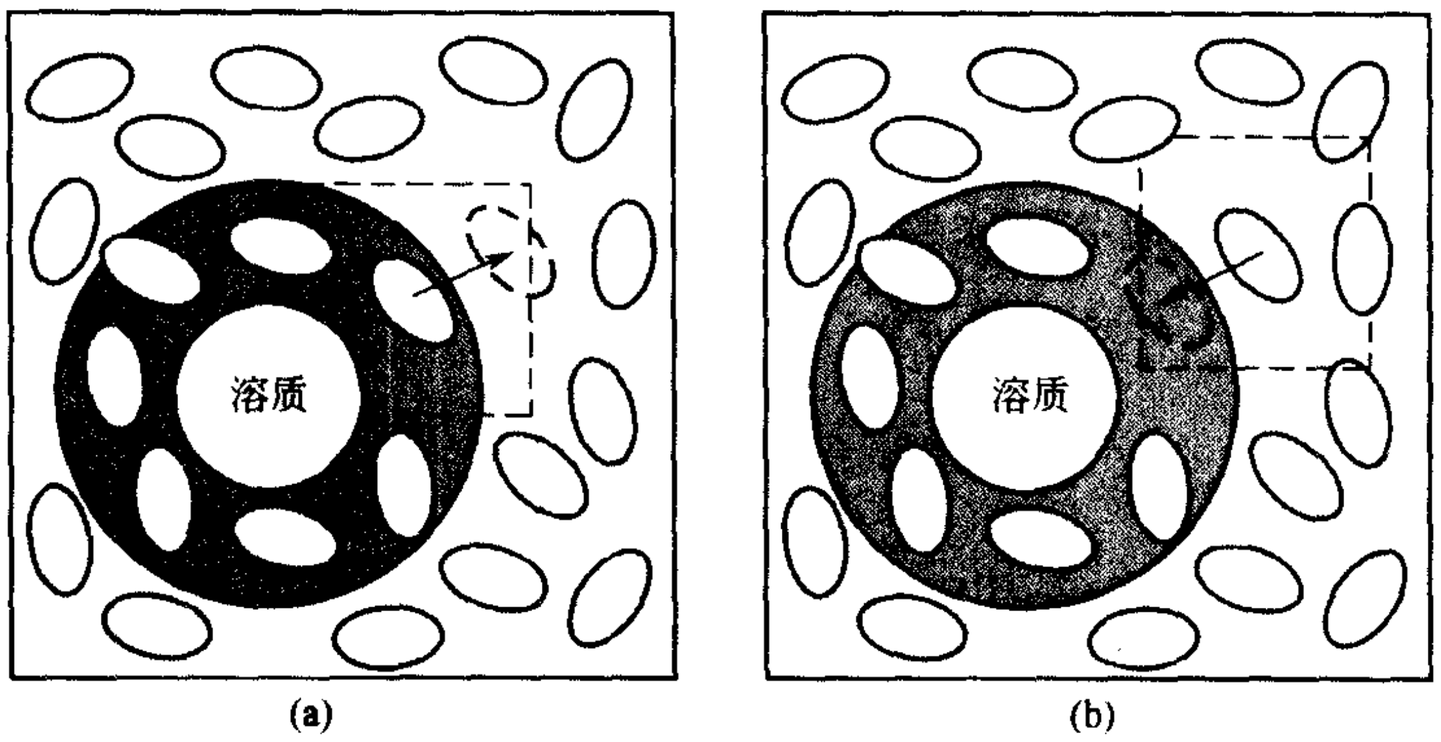



**图 10-5　溶液中溶质与溶剂分子移动的示意**

依此图，若定义围绕溶质分子的特定区间为 $R_{sol}$（如图中所标示的区域），以此分辨处于此区间内外的溶剂分子。则可定义一介于 0 与 1 间的参数 $p$。参数 $p$ 用以决定移动处于区间内分子与移动区间外分子的相较频率，$p$ 值愈小表示移动区内分子的概率愈大。以上假设所执行的蒙地卡罗计算步骤如下。

① 在系统中随机选取一溶剂分子。

② 若所选取的分子处于 $R_{sol}$ 区内，则可以产生移动。

③ 若所选取的分子处于 $R_{sol}$ 区外，则产生一介于 0 与 1 间的随机数 $\xi$。若所设定的 $p$ 值大于或等于所产生的随机数 $\xi$，则移动所选定的分子。否

则，分子不移动。

④ 若分子不移动，则回复步骤①。

依此，$p$ 值愈靠近零，外层分子愈不移动，内层分子移动的概率远大于外层分子。分子移动被接受的概率为：

$$\min\left(\frac{1,\alpha_{nm}\rho_n}{\alpha_{mn}\rho_m}\right) \tag{10-49}$$

设系统的状态 $m$ 含有 $N_{in}$ 个由外层移入内层的分子，含有 $N_{out}$ 个由内层移出的分子。则选择往内移分子的机会为：

$$p_{ij}=\frac{N_{in}}{N}+(1-p)\frac{N_{out}}{N}\frac{N_{in}}{N}+\left[(1-p)\frac{N_{out}}{N}\right]^2\frac{N_{in}}{N}+\cdots=\frac{N_{in}}{N'} \tag{10-50}$$

式中，$N'=pN+(1-p)N_{in}$。对应于选择一个内层的粒子移往外层的 $\alpha_{mn}$ 元素为

$$\alpha_{mn}=\frac{1}{N_{in}}\frac{N_{in}}{N'}\frac{1}{N_R}=\frac{1}{N'N_R} \tag{10-51}$$

式中，$N_R$ 为处于图 10-5 中虚线方块区所有可能的位置。依此，可证明对应于所有移动的方法的 $\alpha_{mn}$ 元素为：

$$\begin{aligned}
&&(m\rightarrow n)\\
\alpha_{nm}/\alpha_{mn}&=p[1-(1-p)/N']^{-1} &(\text{内}\rightarrow\text{外})\\
\alpha_{nm}/\alpha_{mn}&=1 &(\text{外}\rightarrow\text{外})\\
\alpha_{nm}/\alpha_{mn}&=1 &(\text{内}\rightarrow\text{内})\\
\alpha_{nm}/\alpha_{mn}&=\{p[1+(1-p)/N']\}^{-1} &(\text{外}\rightarrow\text{内})
\end{aligned} \tag{10-52}$$

这类以作用的重要性为依据，决定移动分子的频率的蒙地卡罗计算法较一般随机取样的 Metropolis 方法有效得多，计算平均值的收敛速度也大为提升。

## 10.5.2 力场导向蒙地卡罗计算法

力场导向蒙地卡罗（force bias MC）计算法为 Pangali 等所提出的。此方法是以粒子在系统中所受的力的方向为移动粒子的依据。一般的 Metropolis 计算随机决定所选取粒子的移动方向。因此，由于所选的方向往往造成过高的能量而不被接受，以致效率低。以含有 $N$ 个粒子的 LJ 作用系统为例，设粒子 $i$ 与 $j$ 间的作用势能为 $u_{ij}$，则力场导向的方法如下所述。

① 随机选取移动的粒子。

② 计算此粒子所受的力及方向，

$$\vec{f}_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} f_{ij} - \sum_{\substack{j=1 \\ j\neq i}} \nabla_i u_{ij} \tag{10-53}$$

③ 由步骤②中 $\vec{f}_i$ 的方向决定所选粒子的移动方向。依此新的位移为：

$$\vec{r}_{i,new} = \vec{r}_{i,old} + \text{Rand}[-1,1] \times \delta r_{max} \times \frac{\vec{f}_i}{|\vec{f}_i|} \tag{10-54}$$

式中，Rand[−1,1] 为介于−1 与 1 之间的随机数；$\delta r_{max}$为计算所设定的最大移动幅度。

粒子移动被接受的概率为：

$$\min(1, \alpha_{nm}\rho_n / \alpha_{mn}\rho_m)$$

$$\frac{\alpha_{nm}\rho_n}{\alpha_{mn}\rho_m} = \exp\{-[\Delta U_{nm} + \lambda \delta \vec{r}_i^{mn} \cdot (\vec{f}_i^m + \vec{f}_i^n) + \delta W^{FB}]/k_B T\} \tag{10-55}$$

式中，$\delta W$ 为系统位力的变化量。若位移 $\delta r_{max}$很小，则可得到如下的结果，即：

$$\delta W^{FB} \approx \frac{1}{6}\lambda^2 \frac{\delta r_{max}^2 [(\delta \vec{f}_i^{mn})^2 + 2\delta \vec{f}_i^{mn} \cdot \vec{f}_i^m]}{k_B T} \tag{10-56}$$

式中，$\lambda$ 及 $\delta r_{max}$ 为可调变量，用以产生最大的均方根位移（root mean square displacement，RMSD）以增加计算效率，$\vec{f}_i^{mn} = \vec{f}_i^n - \vec{f}_i^m$。力场导向蒙地卡罗计算法对于计算含有氢键作用的分子系统特别有效。因为此类系统中由于分子间具有一定方向性的作用力，故引用力场导向计算法可节省许多取样的浪费。Rao 等将此种方法应用于计算纯水系统，并证实力场导向蒙地卡罗计算结果较一般的蒙地卡罗计算精准，其结果与由分子动力计算所得的结果相近。

## 10.5.3 明智的蒙地卡罗计算法

蒙地卡罗计算是基于随机取样与随机移动的原理。前面讨论了力场导向蒙地卡罗计算法合理地决定粒子移动的方向。移动的幅度由计算中所设定的 $\delta r_{max}$及随机数决定。若是将系统的一些其余的特性也一并考虑，则更可增加计算的精准性。布朗动力（Brownian dynamics）计算法将溶质分子的运动想象为分子自身力场所造成的运动与因溶剂的影响而导致的布朗运动（Brownian motion）。此种方法是利用系统的扩散常数（diffusion constant）与随机作用力效应（stochastic effect）所执行的计算。同样的原

理亦可应用于蒙地卡罗计算法中，称为明智蒙地卡罗（smart Monte Carlo）计算法。此种计算方法为 Rossky 等所提出的，计算中将系统中粒子的移动定为：

$$\delta\vec{r}_i = \vec{r}_{i,new} - \vec{r}_{i,old} = \frac{D^{SMC}}{k_B T}\vec{f}_i + \delta\vec{r}_i^{G} \tag{10-57}$$

式中，$D^{SMC}$为一与系统的扩散常数相关的可调参数；$\delta\vec{r}_i^{G}$为随机移动位移，其各分量由符合高斯分配的随机数决定，并且具备两个条件：

$$\langle \delta r_{ix}^{G} \rangle = \langle \delta r_{iy}^{G} \rangle = \langle \delta r_{iz}^{G} \rangle = 0$$

$$\langle (\delta r_{ix}^{G})^2 \rangle = \langle (\delta r_{iy}^{G})^2 \rangle = \langle (\delta r_{iz}^{G})^2 \rangle = 2D^{SMC} \tag{10-58}$$

这两个限制条件，为使粒子的移动符合系统的扩散性质。粒子移动被接受的概率为：

$$\min\left[1, \exp\left(-\frac{A^{SMC}}{k_B T}\right)\right]$$

式中：

$$A^{SMC} = \delta U_{nm} + \frac{1}{2}(\vec{f}_i^{n} + \vec{f}_i^{m}) \cdot \delta\vec{r}_i^{nm} + \delta W^{SMC}$$

$$\delta W^{SMC} = \frac{D^{SMC}}{4k_B T}[(\delta\vec{f}_i^{nm})^2 + 2\vec{f}_i^{n} \cdot \vec{f}_i^{nm}] \tag{10-59}$$

明智的蒙地卡罗方法与力场导向法相似，皆为在计算中考虑粒子的受力及整体环境的影响（如调整 $D^{SMC}$）。此两种方法的计算结果远优于传统的蒙地卡罗计算，但是计算的复杂度却几乎与分子动力计算或布朗动力计算相当，所需耗费的计算时间亦与这些动力计算相近。但由于蒙地卡罗只能获得统计的平均值，而无法得到系统动态的资料。但这些统计的资料亦可由分子动力计算或布朗动力计算获得，且后者还可得到各种动力的资料。因此，在研究上是否必须从事明智的蒙地卡罗计算，还是从事分子动力模拟是必须审慎考虑的。关于蒙地卡罗方法的介绍可参考 Binder 编辑的“Monte Carlo Methods in Statistical Physics”与“Applications of Monte Carlo Method in Statistical Physics”等书。

## 参 考 文 献


1 Metropolis N, Ulam S. J. Am. Stat. Ass., 44, 335, 1949

2 Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H, Teller E. J Chem. Phys., 21, 1087, 1953

3 Chapman W, Quirke N. Proceedings of the NATO Workshop on the Simulation of Condensed Matter. Physica B, 131, 34, 1985

4 Norman G E, Filnov V S. High Temp (USSR) 7, 216, 1969

5 Adams D J. Mol. Phys., 28, 1241, 1974

6 Owicki J C, Scherage H A. Chem. Phys. Lett., 47, 600, 1977

7 Pangali C, Rao M, Berne B J. Chem. Phys. Lett., 55, 413, 1978

8 Rao M, Pangali C, Berne B J. Mol. Phys., 37, 1773, 1979

9 Rossky P J, Doll J D, Friedman H L. J. Chem. Phys., 69, 4628, 1978

10 Binder K. "Monte Carlo Methods in Statistical Physics", 2nd Ed. Topics in Current Physics, Vol. 7, Springer-Verlag, New York, 1979

11 Binder K. "Applications of Monte Carlo Method in Statistical Physics", 2nd Ed. Topics in Current Physics, Vol. 36, Springer-Verlag, New York, 1987

# 第11章 蒙地卡罗计算的应用

第10章介绍了蒙地卡罗计算方法原理，此方法最初仅限于计算一些简单分子体系的热力学性质，经过许多改进，目前的蒙地卡罗计算已被应用于计算如蛋白质分子、高聚物等巨大的体系。本章举出了蒙地卡罗计算过去至今一些重要的发展例子和算法改进，为第12章介绍计算实例提供了基础。

## 11.1 压力与自由能的变化

计算系统自由能的变化是蒙地卡罗方法的一项重要的应用。早先此类的计算皆是以仅含有较少数量粒子（32～108）的（$N,V,T$）系综为模拟的对象。设系统的总势能为成对加成形式：

$$U=\sum_{i=1}^{N-1}\sum_{j>i}^{N}u_{ij}(r) \tag{11-1}$$

依照统计力学，可推导出（$N,V,T$）系综的状态方程式，即：

$$\frac{PV}{Nk_BT}=1-\frac{N}{6Vk_BT}\int 4\pi r^3 g(r)\frac{du(r)}{dr}dr \tag{11-2}$$

式(11-2)中 $g(r)$ 为径向分布函数，$P$ 为压力。径向分布函数可由计算距离为 $r\to r+\delta r$ 间的平均粒子对数目得到。设平均粒子对数目为 $\langle N\rangle$，则：

$$g(r)=\frac{3\langle N\rangle}{2\pi N\rho[(r+\delta r)^3-r^3]} \tag{11-3}$$

利用如上的关系式可由计算所得系统的径向分布函数计算系统的

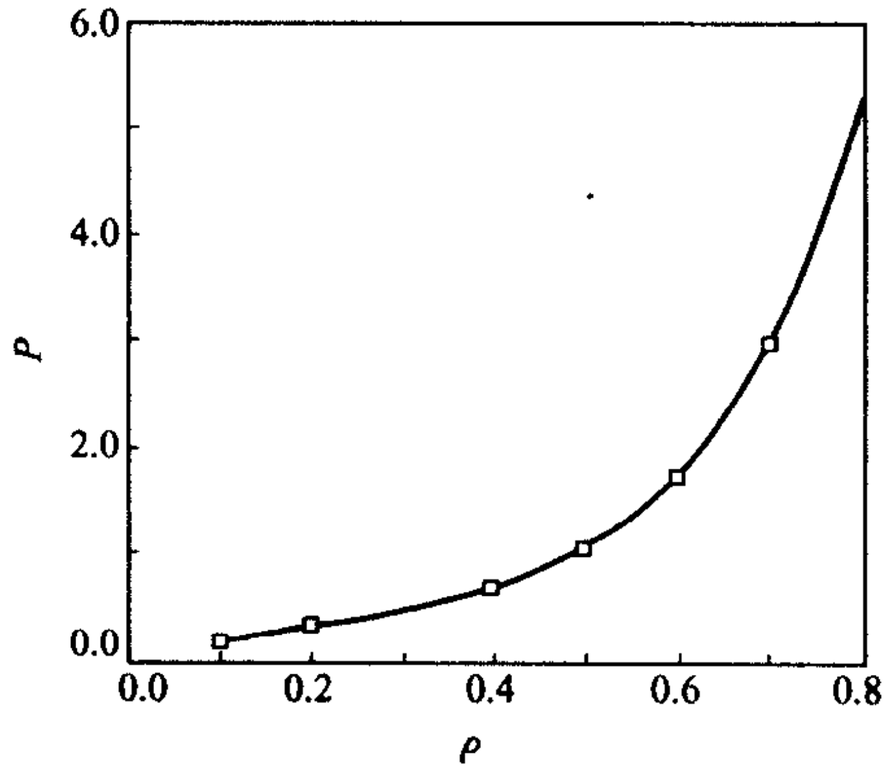



**图 11-1　MC 方法计算 LJ 系统的压力**

压力。

图 11-1 为利用蒙地卡罗方法计算 108 个 LJ 粒子的结果。计算系统的简化温度定为 $T=2.0$。图中的实线为由 Johnson 等所推导的系统的状态方程式得到的结果，方块为 MC 的计算。此图显示计算与理论的一致。利用计算的压力可计算其相关的热力学特性，例如亥姆霍兹自由能（Helmholtz free energy）。系统的亥姆霍兹自由能 $A$ 与压力的关系为：

$$P=-\left(\frac{\partial A}{\partial V}\right)_T \tag{11-4}$$

将式(11-4) 积分得：

$$A(V_2,T)-A(V_1,T)=-\int_{V_1}^{V_2}P\mathrm{d}V \tag{11-5}$$

应用 $\rho=N/V$ 将式(11-5) 转换为以数目密度为变量，得：

$$A(\rho_2,T)-A(\rho_1,T)=N\int_{\rho_1}^{\rho_2}\frac{P(\rho)}{\rho^2}\mathrm{d}\rho \tag{11-6}$$

依此式，可计算系统在定温条件下各不同数目密度的亥姆霍兹自由能的变化。另外，亦可由式(11-7) 计算同样数目密度的系统在不同温度的条件下亥姆霍兹自由能的变化，即：

$$\frac{A(\rho,T_2)}{T_2}-\frac{A(\rho,T_1)}{T_1}=\int_{1/T_1}^{1/T_2}\left(\langle U\rangle+\frac{3}{2}Nk_\mathrm{B}T\right)\mathrm{d}\left(\frac{1}{T}\right) \tag{11-7}$$

式中，$\langle U\rangle+\frac{3}{2}Nk_\mathrm{B}T$ 为系统的内能。通常计算式(11-6) 和式(11-7) 的积分时，采取已知的状态（如理想气体或爱因斯坦固体）为其积分的下限。此外，必须注意，以上的积分公式仅能用以计算积分范围内无相变化系统的自由能变化。通常引用蒙地卡罗方法计算自由能的变化十分费时，因为必须计算许多不同密度或温度状态的压力或内能。近年来已发展出一些改进的方法，希望能更直接的计算任一状态的自由能。各种计算自由能的方法与分子动力学计算自由的方法相同，见于前述的相关章节。

## 11.2 分子液体

早期的蒙地卡罗计算多用以计算分子液体的热力学特性。简单的分子液体如单原子分子，多采用 LJ 势能。这样的计算系统简称为 LJ 液体。较复杂的液体如甲烷、苯等，则采用刚性分子的计算模式。对于刚性分子的计算，一般产生 0→1 的随机数 $\xi$，其欧拉角的直接采样方法为：

$$\begin{aligned}\phi_{new} &= \phi_{old} + (2\xi - 1)\delta\phi_{max} \\ \cos\theta_{new} &= \cos\theta_{old} + (2\xi - 1)\delta(\cos\theta)_{max} \\ \psi_{new} &= \psi_{old} + (2\xi - 1)\delta\psi_{max}\end{aligned} \tag{11-8}$$

式(11-8) 中，若直接于 $\theta$ 的范围内采样，不能得到均一分布（uniform distribution），因此必须于 $\cos\theta$ 的范围内采样。上式的下标“max”表示最大变化幅度。除用欧拉角直接采样外，亦可比照第 6 章的方法，由 $q$ 的四个分量采样，其计算过程如下。

① 于−1→1 的范围内产生一对随机数（$\xi_1$，$\xi_2$）直至符合 $S_1 = \xi_1^2 + \xi_2^2 < 1$。

② 仿照以上的方法产生（$\xi_3$，$\xi_4$），直至 $S_2 = \xi_3^2 + \xi_4^2 < 1$。

③ 由上所产生的随机数，形成四维向量 $[\xi_1, \xi_2, \xi_3 \sqrt{(1-S_1)/S_2}, \xi_4 \sqrt{(1-S_1)/S_2}]$，以此四维向量为（$q_0, q_1, q_2, q_3$）的采样。

除了产生转动的采样外，还需要分子质心移动的采样，通常利用反复试验（trial and error）方法以决定符合预设的接受率（约 50%）最好的采样参数。

对于较软的分子，即含有较多运动自由度的分子，执行一般的蒙地卡罗计算较为困难。通常计算仅限于小分子系统，或是将一些自由度限制住，或是针对系统采取特殊的方法。一般的计算中产生新的构形的方法通常是随机选取原子，将其沿笛卡儿坐标方向移动。但是因为键伸缩势能的关系，通常这样的移动造成较高的势能，以致被接受的概率很低。为了弥补此一缺点，计算中常将键长及键角等易造成高势能变化的内坐标限制住。这种方法可用以计算一些小分子系统，如丁烷、1,2-二氯乙烷等。但以此限制方法计算大分子，即使绕着某根键转动很小的幅度也

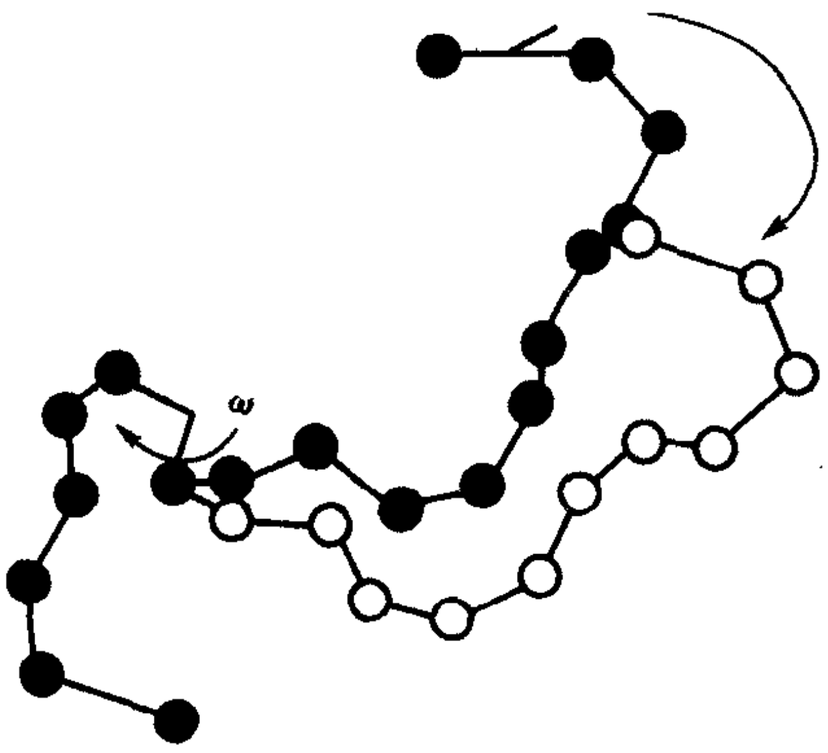



**图 11-2　分子中的键旋转导致势能的大幅升高**

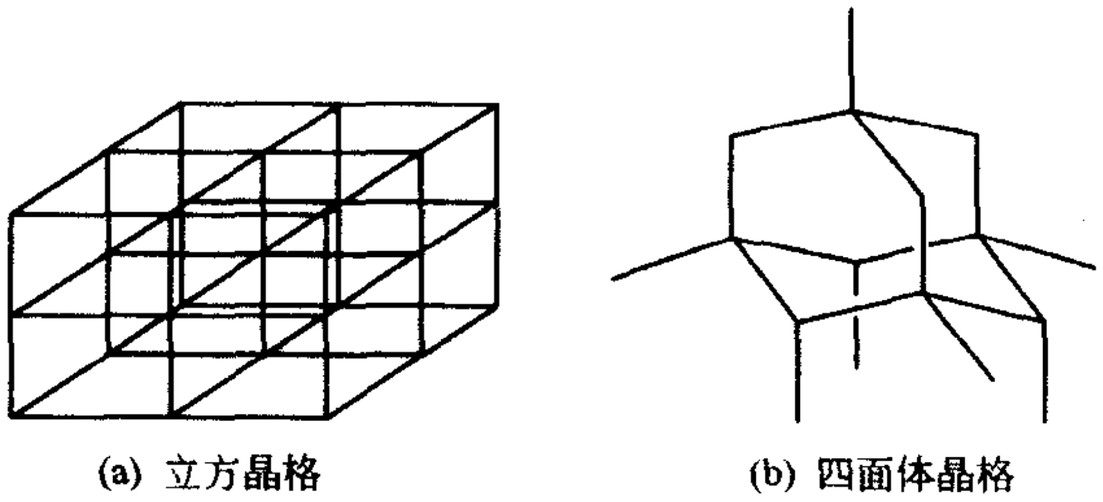
容易造成势能的大幅度升高，如图 11-2 所示。

基于如上的原因，采用键长及键角限制方法的计算必须谨慎，有时亦须注意是否因为增加了这些限制而导致其余的内运动自由度的不均匀采样。

# 11.3 聚合物晶格模式计算

## 11.3.1 晶格模式

蒙地卡罗计算可用于计算聚合物的结构及各种特性。最简单的计算模式是将聚合物考虑为一系列联结的格子点，这种模式称为晶格模式（lattice model），如图 11-3 所示。图中显示两种常用的晶格模型，即三维的立方晶格（cubic lattice）与四面体晶格（或称为钻石晶格，diamond lattice）。



(a) 立方晶格　　(b) 四面体晶格

**图 11-3　两种常用的晶格模型**

建立聚合物的晶格模型时，将聚合物的每一基团视为坐落于晶格上不重复的晶格点。一般尽量将晶格设计为接近真实分子的空间排列。如图 11-4 为描述聚合物分子聚乙烯的键扰动晶格模型；图中将聚合物中所标示的碳原子的中心置于三维的晶格点上，每根键则表示相隔两个亚甲基的碳碳间的位置。

## 11.3.2 晶格模式计算的执行步骤

此例介绍如何应用简单的晶格模型与蒙地卡罗方法计算聚合物构型的各种统计资料。在此简单的计算范例中，系统仅含一根聚合物链，其余则

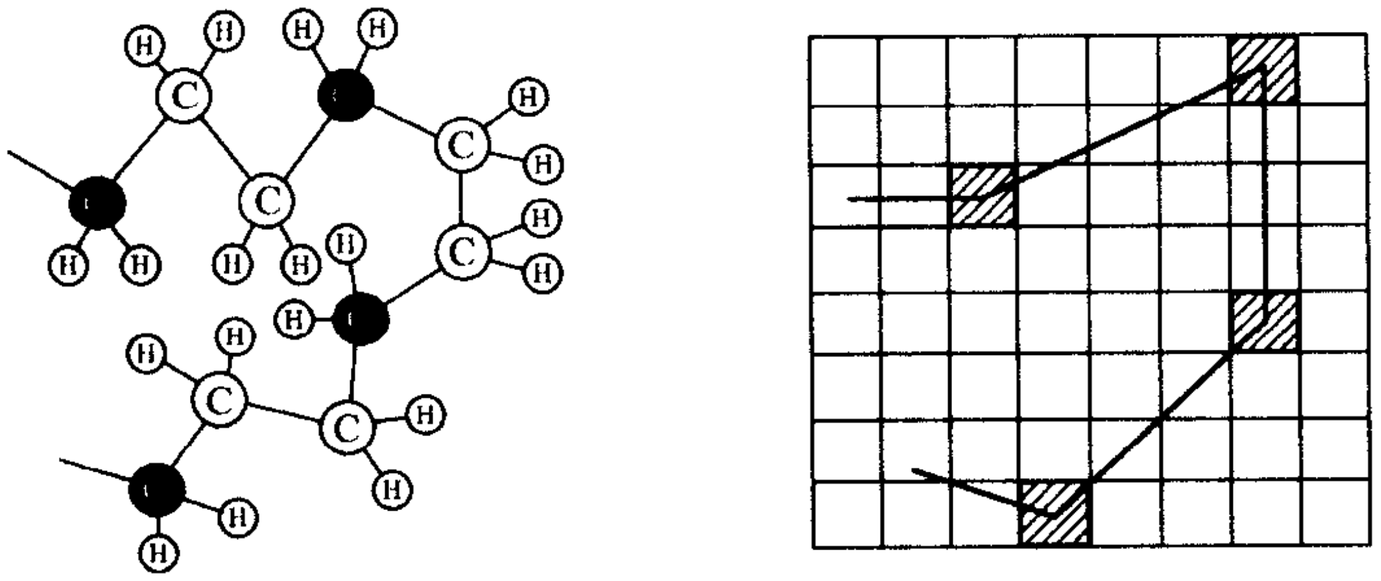


图 11-4　模拟聚合物聚乙烯的键扰动晶格模型

假设为溶剂分子。将聚合物假想为由 $N$ 根键的 $N+1$ 个二维格子点所串联而成，如图 11-5 所示。

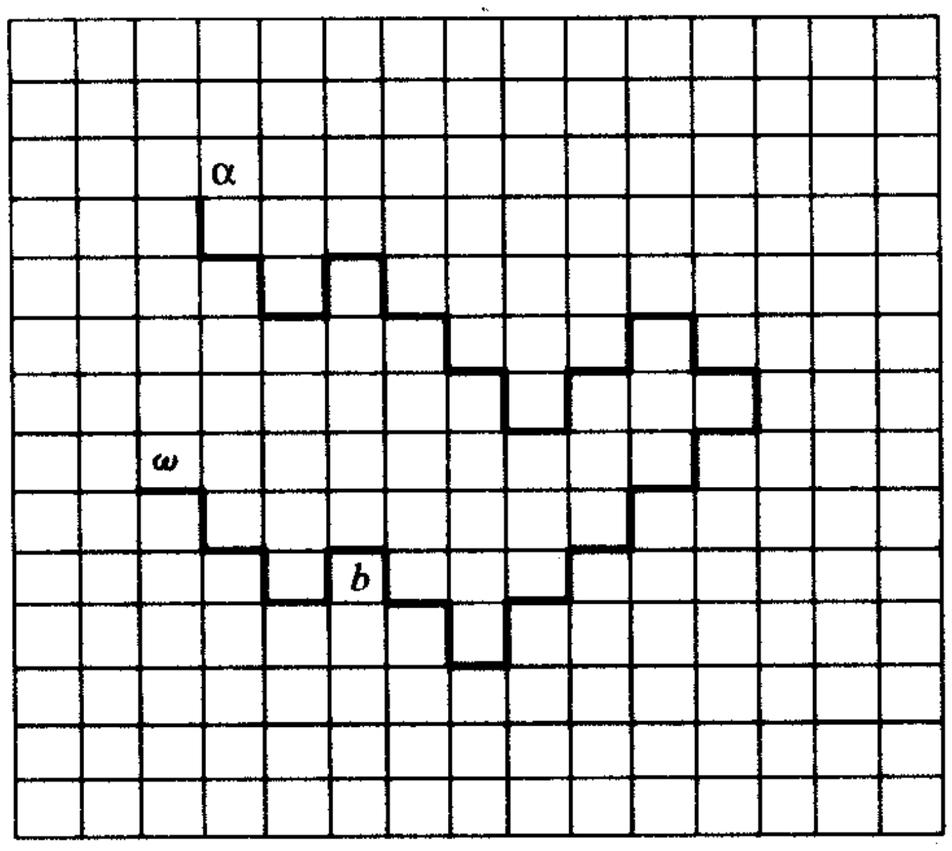



图 11-5　单链二维晶格的计算模式

图中每一点代表组成聚合物的一个单体（monomer），$\alpha$ 与 $\omega$ 分别为聚合物的起始基团与终端基团，图 11-5 面上其余的格子点代表溶剂分子。连接聚合物单体间的键长为 $b$，而此模型中可能的键角为 90°及 180°。设每一聚合物的单体，除了与其键结的格子点上的单体无作用外，最多只与其邻近的 8 个格子点上的单体或溶剂分子作用。依此假设，共需考虑三种型式的作用能：$E(m-m)$ 为非键结的单体与单体间之作用；$E(m-s)$ 为单体与溶剂间之作用；$E(s-s)$ 为溶剂与溶剂间之作用。可定义作用参数 $\omega$ 为：

$$\omega = E(m-s) - \frac{1}{2}[E(m-m)+E(s-s)] \tag{11-9}$$

执行蒙地卡罗计算的步骤为，随机选取聚合物的任一单体，检视是否可将此单体移至相邻近的任一格子点上。若是移动后的格子点并无任何其他的单体，则此移动步骤可被接受。反之，若是移动后的格子点原先已被其他的单体占据，这种移动称为违背了不重复原则，此移动步骤被拒绝而不被接受。依照 Metropolis 的方法，设移动前后的能量分别为 $E$(old) 与 $E$(new)。若 $E$(new)$\leqslant E$(old)，则此移动可被接受，单体移至新的位置，原先的位置被溶剂分子所取代。若 $E$(new)$>E$(old)，则计算玻耳兹曼因子，$\exp(-\Delta E/k_B T)$，$\Delta E = E(\text{new}) - E(\text{old})$。若此因子大于一个0～1 的随机数，则此移动被接受，否则此移动不被接受，单体回至原先的位置。

### 11.3.3 研究聚合物的物理量

P. J. Flory 曾致力于发展各种聚合物的构型统计学，提出许多有用的关于聚合物结构的理论，并于 1974 年以此获得诺贝尔奖。对一含有由 $N$ 根键串联 $N+1$ 个单体的聚合物，定义 $\vec{r}_i$ 为第 $i$ 个单体相对于聚合物质心的径向量；$\vec{b}_i = \vec{r}_{i+1} - \vec{r}_i$ 为第 $i$ 个键向量，$\theta_i$ 为第 $i$ 个键的键角。聚合物的形状与其平均末端距（end-to-end distance）相关。末端距即聚合物的头基与尾基间的距离。在“优良”溶剂（good solvent）中，聚合物因受溶剂分子的碰撞而倾向于链的伸展。一完全伸展的链，其末端距的理论值应为 $N \cdot b$。若溶剂为“劣质”溶剂（bad solvent），则聚合物形成卷曲形，而与溶剂分离。很明显的，在此种溶剂系统中，聚合物的末端距将小于完全伸展的末端距。通常可由度量末端距平方的平均值推测聚合物的形状，即：

$$\langle R^2 \rangle = \langle |\vec{r}_{N+1} - \vec{r}_1|^2 \rangle \tag{11-10}$$

式中，$R$ 为末端距。另一种研究聚合物大小的物理量，为其回转半径（radius of gyration）平方的平均值，$\langle S^2 \rangle$。其值可由光散射实验得出。$\langle S^2 \rangle$ 的定义为：

$$\langle S^2 \rangle = \frac{\sum_{i=1}^{N+1} m_i r_i^2}{M} \tag{11-11}$$

式中，$m_i$ 为第 $i$ 个单体的质量；$M$ 为聚合物的总质量。研究聚合物的运动，最粗略的模式为随意乱动模式（random walk)。依照随意乱动模式，聚合物的运动无任何扭转角的限制。因此，可由统计力学推导出如下

的关系式，即：

$$\langle R^2\rangle = N \cdot b^2$$

$$\langle S^2\rangle = (N+1)(N+2)\frac{\langle R^2\rangle}{6} \tag{11-12}$$

研究聚合物的柔软度（flexibility）的物理量为驻留长度（persistence length），$\langle a\rangle$。其定义为：

$$\langle a\rangle = \frac{1}{b_1}\sum_{i=1}^{N}\langle \vec{b}_1 \cdot \vec{b}_i\rangle \tag{11-13}$$

其意义为所有聚合物的键向量对第一根键向量投影的总和。对可随意运动的聚合物而言，其柔软度最大，$\langle a\rangle = b$。因在随意乱动模式中，键与键彼此无任何关联。故可得到关系式：

$$\langle \vec{b}_1 \cdot \vec{b}_i\rangle = |b|^2\langle \cos\theta_{i1}\rangle = 0 \tag{11-14}$$

若聚合物为完全伸展的棒形，则所有的键向量均平行于 $\vec{b}_1$，故 $\langle a\rangle = N \cdot b$。

研究聚合物的形状亦可计算其惯性矩的张量，$\tilde{I}$。

$$I_{\alpha\beta} = \sum_{i=1}^{N+1} m_i(r_i^2\delta_{\alpha\beta} - r_{i\alpha}r_{i\beta}) \qquad (\alpha,\beta = x,y) \tag{11-15}$$

式中，$\delta_{\alpha\beta}$为 Kronecker 的三角形符号，即当 $\alpha=\beta$ 时 $\delta_{\alpha\beta}$等于 1，否则为 0。惯性矩的张量的迹（trace），即张量所构成矩阵对角线元素的和，等于回转半径的平方乘以聚合物的质量。

$$\sum_{\alpha=x,y,z} I_{\alpha\alpha} = MS^2 \tag{11-16}$$

对无方向性随机运动的聚合物而言，其惯性矩对角线元素的平均值必相同。模拟过程中，对每一聚合物的新构型，计算其惯性矩，并将 $\tilde{I}$ 对角线化，可得到该构型的主转动惯量（principal moment of inertia）。其对角线化的过程为：

$$\tilde{I}_P = \tilde{C}^T \cdot \tilde{I} \cdot \tilde{C}$$

$$\tilde{I}_P = \begin{pmatrix} I_{aa} & 0 & 0 \\ 0 & I_{bb} & 0 \\ 0 & 0 & I_{cc} \end{pmatrix} \tag{11-17}$$

将对角线化后所得的较大与较小的主转动惯量分别相加，并分别求得其平均值，再除以聚合物的总质量，则可得到物理量 $\Gamma_1$（对应于较大的主转动惯量），$\Gamma_2$（对应于较小的转动惯量）。由这些物理量可判定聚合物的形状。又因对角线化并不改变张量的迹，故 $\Gamma$ 的和必等于 $M\langle S^2\rangle$。

聚合物中单体间的相互作用可分为短距作用与长距作用。短距作用决定聚合物的局部结构（local structure），如键长、键角、扭转能障等。长距作用来自库仑静电作用与排除体积（excluded volume）效应。一般难以用一个固定的物理量描述排除体积效应。此处以式(11-18) 显示排除体积效应的影响，即：

$$\langle R^2 \rangle = f(N) \cdot \langle R^2 \rangle_0 \tag{11-18}$$

式中，下标 0 表示无排除体积效应的理想值。$f(N)$ 为一与聚合物组成单体个数有关的函数，通常符合如下的指数律，即：

$$f(N) = A \cdot N^{v-1} \tag{11-19}$$

将此式与随意运动模式的结果，$\langle R^2 \rangle_0 = Nb^2$，代入式(11-17)，得：

$$\langle R^2 \rangle = A \cdot N^v b^2 \tag{11-20}$$

式中，常数 $A$ 与 $v$ 可由计算的结果推出。

### 11.3.4 MC 计算的结果

以下的范例为 Engström 等以蒙地卡罗的二维晶格计算模式模拟 4 个聚合物系统的结果。系统中构成聚合物键的数目分别为 5、9、13 与 17。对每一系统，先计算约 $(N+1) \cdot 2000$ 种构型，以使系统达到所谓的“平衡”状态。然后，再计算 $(N+1) \cdot 5000$ 个构型，以求取各物理量的平均值。表 11-1 列出由不同系统计算所得的结果。此计算中，将作用参数 $\omega$ 设定为零。由表中的计算结果可知，末端距平方的平均值随键数而增加，此为预期的结果。将计算的 $\langle R^2 \rangle$ 值代入式(11-20)，可以利用线性回归法 (linear fit) 求出 $A=0.81$，$v=1.50\pm0.01$。Bishop 与 Michels 曾以复杂的球与弹簧模式计算聚合物的排除体积效应。其计算所得的结果为 $v=1.48\pm0.03$，与此简单模式的计算结果十分相近。

**表 11-1　聚合物各种物理量于 $\omega = 0$ 条件下的计算**

| $N$ | $\langle R^2 \rangle$ | $\langle S^2 \rangle$ | $\langle a \rangle$ | $\langle \Gamma_1 \rangle / \langle S^2 \rangle$ | $\langle \Gamma_2 \rangle / \langle S^2 \rangle$ |
|---|---|---|---|---|---|
| 5 | 9(1) | 1.8(0.0) | 1.6(0.0) | 0.85 | 0.15 |
| 9 | 22(3) | 3.5(0.3) | 1.7(0.3) | 0.86 | 0.14 |
| 13 | 38(7) | 5.8(0.5) | 1.9(0.4) | 0.85 | 0.14 |
| 17 | 56(16) | 8.4(1.2) | 1.9(0.5) | 0.86 | 0.14 |

注：表中的（　）显示计算的误差。

当聚合物的链达到足够的长度时，由随意运动模式所得的理论值为 $\langle R^2\rangle/\langle S^2\rangle \approx 6$。表 11-1 中，此两个物理量的比值均接近 6，与此结果一致。表中所显示的驻留长度 $\langle a\rangle$，其值随聚合物的键数而增加，但均小于 2 倍的键长。这表示聚合物的柔软度甚高。此外，此计算所得较大与较小主惯性矩的平均值分别接近 0.86 与 0.14，显示聚合物的形状几乎呈椭圆体状。此点亦与 Bishop 等计算键数为 63 的聚合物的结果几乎一致 (0.855 与 0.145)。

若计算所设定的作用参数改变，则会导致不同的结果。表 11-2 列出针对键数为 17 的聚合物系统，由不同的作用参数，$\omega = -0.20k_BT \sim 0.20k_BT$ 计算所得的结果。当作用参数增加时，不利于聚合物的单体与溶剂之作用。这种情形下，聚合物倾向于形成卷曲的构型，以避免与溶剂分子作用。因此造成 $\langle R^2\rangle$，$\langle S^2\rangle$ 与 $\langle a\rangle$ 值的下降，如表 11-2 所示。此外比较表中的结果可知，当作用参数增加时，聚合物的形状稍有改变。其形状略倾向于球形，以避免单体与溶剂的作用。以上的结果显示利用简单的晶格模型可以得到聚合物的一些结构的统计值。

**表 11-2　含 17 个键的聚合物各种物理量于 $\omega$ 变化条件下的计算**

| $\omega$ | $\langle R^2\rangle$ | $\langle S^2\rangle$ | $\langle a\rangle$ | $\langle \Gamma_1\rangle/\langle S^2\rangle$ | $\langle \Gamma_2\rangle/\langle S^2\rangle$ |
|---|---|---|---|---|---|
| −0.20 | 80(21) | 11.0(1.8) | 2.8(1.2) | 0.87 | 0.13 |
| −0.10 | 72(19) | 8.9(1.3) | 2.8(0.6) | 0.87 | 0.13 |
| 0 | 56(16) | 6.4(1.2) | 1.9(0.5) | 0.86 | 0.14 |
| +0.10 | 48(16) | 7.2(1.3) | 1.6(0.6) | 0.86 | 0.14 |
| +0.20 | 29(16) | 5.4(0.9) | 0.9(0.8) | 0.83 | 0.14 |

注：表中的（　）显示计算的误差。

## 11.4　改良型晶格计算

虽然晶格计算方法的原理十分简单，但事实上因为不重复原则的限制，欲产生足够的链组态非常不易，效率很低。意即所计算的统计的平均值不易收敛。针对此缺点，Rosenbluth 等提出改进取样的方法，以增加取样的效率。设由三度空间的原点 (0,0,0) 开始，每次移动一格子的长度，经过 $m$ 次移动构成一根链。若第一次移动至晶格点 (1,0,0)，其被接受的统计比重 (statistical weight) 为 $W_1 = 1$，此点移动至晶格点 ($x_m$, $y_m$,

$z_m$)，则下一步移动可能的晶格点为（$x_m \pm 1, y_m \pm 1, z_m \pm 1$)。但这六个晶格点中有一点与第$m-1$步移动的坐标重复，不被接受。其余五个移动位置其被接受的概率如下。

① 若此 5 个位置均无其他单体存在，意即不违背不重复原则，则第 $m+1$ 步被接受的统计比重与第 $m$ 步被接受的统计比重相同，为 $W_{m+1} = W_m$，可在这 5 个位置随意选取一个。

② 若此 5 个位置中仅有 $n$ 个无其他单体存在，则第 $m+1$ 步被接受的统计比重为 $W_{m+1} = \frac{1}{5} n W_m$，可在此 $n$ 个位置中任选一个。

依此原则移动 $N$ 步，产生了聚合物链，此链的总统计比重为：

$$W'_N = W_1 \times W_2 \times \cdots \times W_N \tag{11-21}$$

任何物理量 $A$ 的平均值为

$$\langle A \rangle = \frac{\sum_N W'_N A_N}{\sum_N W'_N} \tag{11-22}$$

利用这样的方法，Rosenbluth 以二维晶格模型与三维晶格模型计算了总长为 65 个单体内的聚合物系统，其计算聚合物头尾距离平方的平均值的结果为：

$$\langle R^2 \rangle = 0.917 N^{1.45} \qquad \text{（二维晶格）}$$

$$\langle R^2 \rangle = 1.0 N^{1.22} \qquad \text{（三维晶格）}$$

Wall 与 Erpenbeck 亦提出链含过程（chain-enrichment process）的取样方法，使原先接受率为 20%提升至 80%。其计算的结果为：

$$\langle R^2 \rangle = 0.80 N^{1.50}, \quad \langle S^2 \rangle = 0.145 \langle R^2 \rangle, \quad \text{（二度空间晶格）}$$

$$\langle R^2 \rangle = 1.40 N^{1.18}, \quad \langle S^2 \rangle = 0.157 \langle R^2 \rangle, \quad \text{（三度空间晶格）}$$

理论上，三维晶格的模型应较二维晶格的模型接近真实的系统，但上述计算末端距与键数的关系指数值均非常接近，其值均介于 1.2～1.5 的范围内。

## 11.5 蛇形滑动模式

上述各晶格模式方法所产生聚合物的链长不一，聚合物的结构由不违背不重复原则的随意移动所产生。另一种研究含有固定链长的聚合物系统的方法为 Wall 与 Mandel 所发展的蛇形滑动模式（slithering snake

model)。在真实的聚合物液体中，聚合物链以滑动的方式运动，即当链的首端单体向一新位置移动时，链的第二单体则递补原先首端单体的位置，链的第三单体则递补原先第二单体的位置，依此类推。因此整根链的运动好像蛇的滑动。图 11-6 显示一根含有六个单体的聚合物链在二度空间格子上的蛇形运动。若将键角限制为 90°则共有五种可能的结构，分别以 $A$，$B$，$C$，$D$ 与 $E$ 表示。图中，以箭号与圆圈分别代表链的首端与末端。

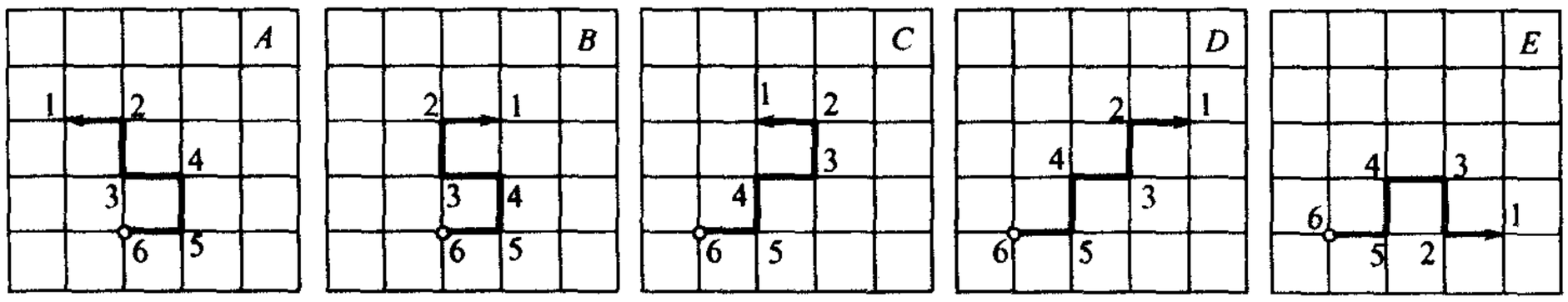



**图 11-6　含有六个单体的聚合物链在二度空间格子上的蛇形运动**

设以上任一种构形的势能皆相同，因此若产生大量的分子链，则任一种构型出现的概率应为 1/5。计算产生新构形的方式为，让链的首端随机移动一格子点，其余的链以蛇形滑动，如 2 移至原先 1 的位置，3 移至原先 2 的位置……。图 11-7 列举 $A$ 链移动的情形，当首端移动后，可产生 $C$ 链或 $D$ 链。

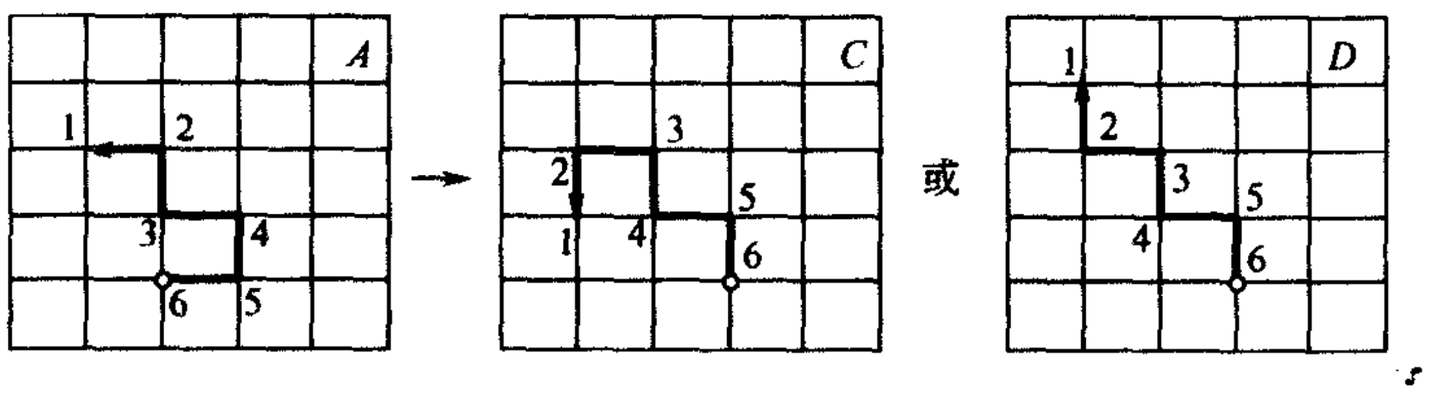



**图 11-7　A 链移动的情形**

图 11-8 显示 $B$ 链的移动情形；当 $B$ 链的首端移动后，可产生 $E$ 链或是违背不重复原则的结构。

由此种滑动产生的 $E$ 链可被接受，而违背不重复原则的结构则被拒绝。此时，将原来的首端重新定义为末端而将原先的末端重新定义为首端，以产生新的结构。可证明利用这种方法以产生新链的过程符合标准的

统计取样法。图 11-9 显示 $B$ 链互换首尾的情形，此时，结构仍旧为 $B$ 链，无新的结构产生。

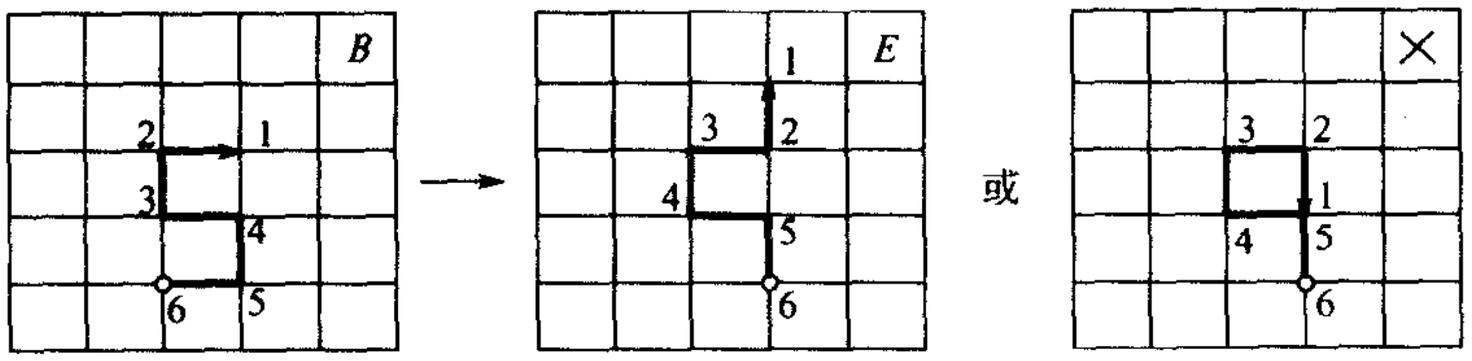



**图 11-8　B 链移动的情形**

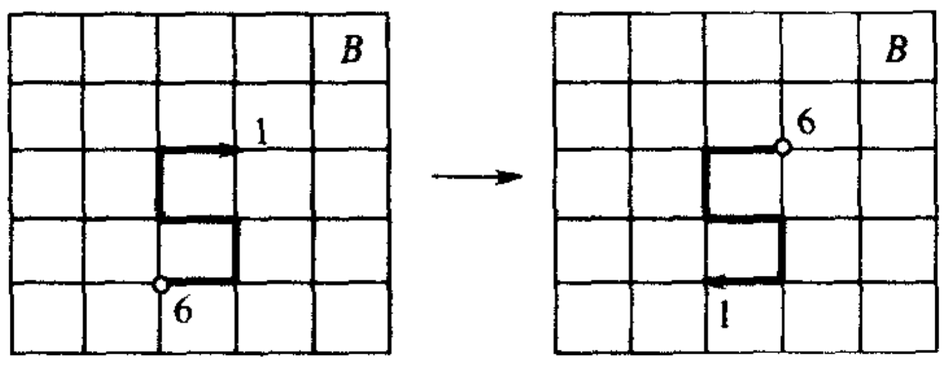



**图 11-9　B 链首尾互换以产生新的构象**

但若将 $C$ 链的首尾互换，则不会形成 $C$ 链，而产生了一根新的 $A$ 链。图 11-10 为由 $C$ 链转为 $A$ 链的状况。

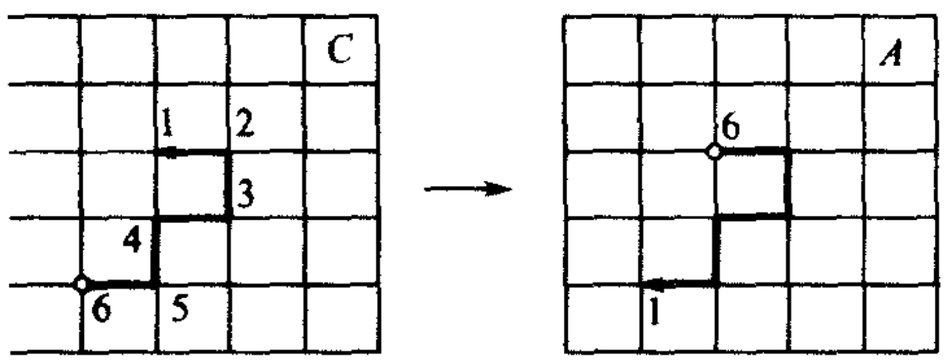



**图 11-10　由 C 链转为A 链的状况**

此处为了举例故将键角限制在 90°，事实上在 Wall 与 Mandel 的早先计算中，键角亦可能为 180°。利用他们的方法，先任意产生一根聚合物链，再以蛇形滑动法产生大量链的结构，由此可计算聚合物的各种平均值。Mandel 应用此种方法，计算含有 21 个单体的聚合物链的头尾平均距离，其结果与实验所得非常相近。

## 11.6 聚合物的改进计算模式

### 11.6.1 球与棒模式

在聚合物的晶格模式计算法中，单体间的作用只简单的以不重复原则为唯一的考虑。较精确的计算为球与棒模式 (bead-rod model)。此模式将含有 $N+1$ 个单体的聚合物视为由 $N$ 个长度为 $l$ 的键联结 $N+1$ 个球所成。球与球间的作用为简单的 LJ 势能，即：

$$u_{ij}^{LJ}=\varepsilon\left[\left(\frac{r_0}{r_{ij}}\right)^{12}-\left(\frac{r_0}{r_{ij}}\right)^{6}\right] \tag{11-23}$$

式中，$r_{ij}=|\vec{r}_i-\vec{r}_j|$，为第 $i$ 个球与第 $j$ 个球间的距离；$\varepsilon$ 为势能常数；$r_0$ 为常数，一般将其定义为与键长 $l$ 相同。球与棒模式的蒙地卡罗计算由任一组结构 $G(\vec{r}_1,\vec{r}_2,\cdots,\vec{r}_{N+1})$ 开始，依如下规则进行。随意选取一个球 $i$，绕着由 $\vec{r}_{i-1}$ 与 $\vec{r}_{i+1}$ 所连成的轴任意旋转至一新的位置 $\vec{r}_i'$，如图 11-11 所示。

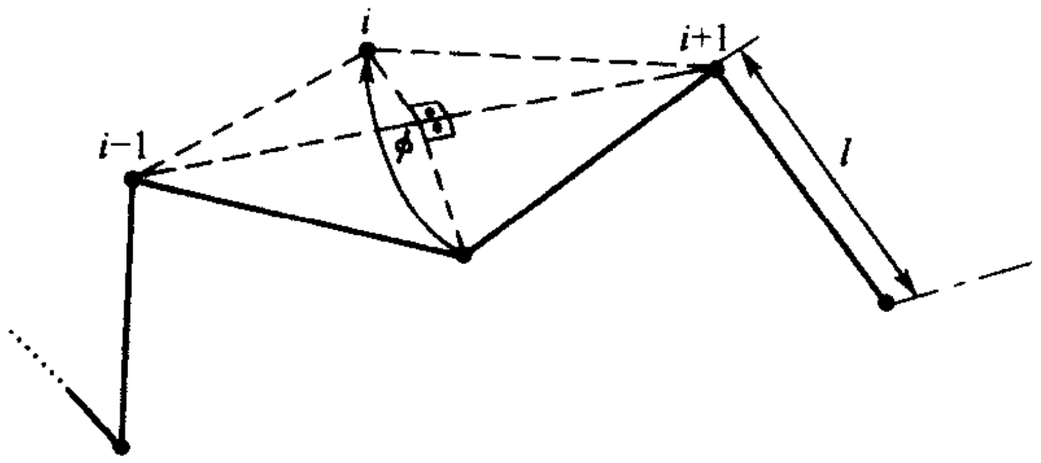



**图 11-11　球与棒模式的绕轴旋转**

依此法旋转后，得到新的构型。此新的构形并无键长与键角的变化。产生新的构型后，依照 Metropolis 取样方法，若新构型的能量 $U_i'$，低于原构型的能量 $U_i$，则此步骤可被接受。若 $U_i'>U_i$，则比较 $\exp[(U_i-U_i')/kT]$ 与一个 0～1 的随机数。若指数项大于此随机数，则此移动可被接受，否则不被接受，而重新由 $G$ 选取另一个球，再进行同样的移动试验。

Baumgärtner 引用球与棒计算模式成功的计算了一些聚合物的性质。因为在 LJ 势能项中含有排斥项与吸引项两部分。在低温时，作用以吸引项为主，聚合物分子应处于低能量的较稳定球状结构 (globular struc-

ture)。在高温时，作用以排斥项为主，聚合物应处于高能量的散乱状态(random-coil)。可定义一温度$\theta$以描述聚合物分子由稳定结构至散乱状态的变化。若聚合物呈散乱状态，则其末端距呈发散状态，依此可将$\theta$定义为当聚合物的末端距合于式(11-24) 时的温度，即：

$$\lim_{N\to\infty}\frac{\langle R_N^2\rangle}{N}\bigg|_{T=\theta}=\text{常数} \tag{11-24}$$

或将此温度定义为当聚合物液体其二次位力系数 (virial coefficient) 为零时的温度，此定义相当于即符合式(11-25) 条件的温度，即：

$$\lim_{N\to\infty}\left\langle\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=i+1}^{N}\left[1-\exp\left(\frac{u_{ij}}{k_BT}\right)\right]\right\rangle\bigg|=0 \tag{11-25}$$

将 LJ 势能及式(11-24) 的条件代入式(11-25)，可推得 $k_B\theta/\epsilon=3.70$ 的关系。并得到当聚合物液体的密度为$\rho$时，其末端距与回转半径符合如下关系，即：

$$\langle R_N^2\rangle\propto\langle S_N^2\rangle\propto N^{2v},\ \rho\propto N^{1-3v} \tag{11-26}$$

并且，可以得到参数$v$与系统温度$T$及$\theta$的关系，即：

若 $T>\theta$，$v\approx0.59$

若 $T=\theta$，$v=0.5$

若 $T<\theta$，$v=1/3$ (11-27)

Baumgärtner 应用蒙地卡罗方法及球与棒模式计算含有 64 个单体的聚合物。当 $k_BT/\epsilon>\theta$ 时其计算所得的末端距与回转半径的关系与式(11-26) 非常吻合，计算所得的 $v$ 值约等于 0.59。当 $k_BT/\epsilon<\theta$ 时，计算所得的 $v$ 值远小于理论值 1/3。这是由于计算的聚合物链并非无限长 ($N\to\infty$) 所造成的差异。但此计算成功的检验了变化温度$\theta$，并得到高温时与理论值一致的结果。

Webman 等引用所谓“reptation 蒙地卡罗”法改进上述模型的缺陷。此计算依然采用球与棒模式，其势能为 LJ 形式，但加入截断半径的观念。

$$u_{ij}=\begin{cases}4\epsilon\left[\left(\dfrac{\sigma}{r_{ij}}\right)^{12}-\left(\dfrac{\sigma}{r_{ij}}\right)^{6}\right]+\epsilon & (r_{ij}\leqslant r_c)\\ 0 & (r_{ij}>r_c)\end{cases} \tag{11-28}$$

此处截断半径为 $r_c=2^{1/6}\sigma$。依此式，仅当球与球的距离小于截断半径 $r_c$ 时彼此才有互相排斥作用，若大于截断半径则彼此间无作用力。设含有 $N+1$ 个球链的起始结构为 $\vec{X}=(\vec{r}_1,\vec{r}_2,\cdots,\vec{r}_{N+1})$，新的结构 $\vec{X}'$可由下述方式产生：随意选择首端的第一个球或是末端的第 $N+1$ 个球，依任意选择的方向移动一键长的距离，链上其余的球则依蛇形滑动方式移动至新的

位置。将此种移动以数学式表示如下所述。

当选择移动末端球时：

$$\vec{r}_i'=\vec{r}_{i+1} \qquad (i=1,\cdots,N)$$

$$\vec{r}'_{N+1}=\vec{r}_{N+1}+\vec{s} \tag{11-29}$$

当选择移动首端球时：

$$\vec{r}_i'=\vec{r}_{i-1} \qquad (i=2,\cdots,N+1),\vec{r}_1'=\vec{r}_1+\vec{s} \tag{11-30}$$

此处 $\vec{s}$ 为随意方向长度为一个键长之向量。依照蒙地卡罗计算 Metropolis 法则，每一移动步骤前后的能量若符合 $U'(\vec{X}')<U(\vec{X})$，则此移动可被接受，否则计算玻耳兹曼因子 $\exp[-\Delta U/k_B T]$，并将其与 0～1 的随机数 $\eta$ 比较，若指数项大于 $\eta$ 则此移动可被接受，否则返回原结构 $\vec{X}$。Webman 等由此方法计算很短的 $N$ 个键（$N\leqslant 10$）聚合物链的末断距 $\langle R_N^2\rangle$，得到如下的关系，即：

$$\langle R_N^2\rangle \propto N^{2v},2v=1.19 \tag{11-31}$$

由式(11-31) 中所得的指数几乎与式(11-26) 中的理论值 $2\times 0.59=1.18$ 相同，成功地检验了此计算模式的正确性。

## 11.6.2 球与弹簧模式

前述的球与棒模式仅考虑了聚合物单体间的范德华作用。较精准的计算模式为 Bishop 所提出的球与弹簧模式（bead-spring model)。此模式将聚合物视为一串由弹簧联结的球，除了考虑范德瓦耳斯作用外并考虑相邻两单体间的振动。若将振动视为简谐振动，则相邻两球 $i$，$j$ 间的振动势能项可定为：

$$u_{ij}^H=\begin{cases}-0.5k_B r_0^2\ln\left[1-\left(\dfrac{r_{ij}}{r_0}\right)^2\right] & (0\leqslant r_{ij}\leqslant r_0)\\ 0 & (r_{ij}>r_0)\end{cases} \tag{11-32}$$

式中，$r_{ij}$ 为相邻二球间的距离，若 $r_{ij}<r_0$，则两球间呈简谐振动。Bishop 设定 $r_0=1.95\sigma$，$k_B=20\epsilon/\sigma^2$。引用此势能项可避免聚合物链与链间的缠绕（chain crossing）现象。Bishop 等应用此模式及 reptation 蒙地卡罗方法计算多聚合物链系统。计算的过程如下。

① 聚合物链依次移动首端或末端的球，沿着 $\vec{s}$ 方向至新的位置 $\vec{r}'$，$\vec{s}$ 为一任意选择的向量。

② 由伸缩量 $\vec{s}$ 所导致的振动势能的变化量为 $\Delta u^H$。

③ 若 $\Delta u^H\leqslant 0$，则此 $\vec{s}$ 可被接受。否则，计算 $\exp(-\Delta u^H/k_B T)$ 并

与一介于 0～1 的随机数比较。若指数项大于此随机数，则此 $\vec{s}$ 被接受，否则被排斥。

④ 由步骤③所得的聚合物链的新组态为 $\vec{X}'=(\vec{r}_1+\vec{s},\vec{r}_1,\vec{r}_2,\cdots,\vec{r}_N)$，或为 $\vec{X}'=(\vec{r}_2,\vec{r}_3,\cdots,\vec{r}_{N+1},\vec{r}_{N+1}+\vec{s})$。为决定此新的组态是否被接受，需计算范德瓦耳斯作用势能，若势能变化大于 0，则计算玻耳兹曼乘子 $\exp\left[-\sum_{j=2}^{N}\frac{u^{LJ}(\vec{r}_j-\vec{r}')-u^{LJ}(\vec{r}_j-\vec{r}_1)}{k_BT}\right]$ 并将此指数项与一介于 0 与 1 的随机数比较。若指数项大于此随机数，则此新组态可被接受，否则被排斥。

⑤ 重复①～④的步骤。

Bishop 等利用球与弹簧模式计算含有 5、10、20、32、50 与 70 个球的单链聚合物系统，得到与前节相同的结果。除了单链系统外，亦计算了密度为 0.1、0.3 与 0.5 的多链系统。此处密度为 $n(N+1)/L^3$，$n$ 为链的数目，每根链均含有 $N+1$ 个球，$L$ 为系统立方盒子的边长。计算显示当系统密度由 0.1 增加至 0.3 时，$\langle R_N^2\rangle$ 与 $\langle S_N^2\rangle$ 的指数律中的指数，即 $\langle R_N^2\rangle\propto\langle S_N^2\rangle\propto N^{2v}$ 中的 $v$ 值，呈现递减的趋势。此计算结果亦与式(11-26) 中，$\rho$ 增加则 $v$ 下降的趋势相同。

## 11.7 随机搜寻最佳结构方法

在复杂大分子的结构最优化过程中，常因所定的初始结构与能量最小结构相差太远，以致结构最优化得到的仅为区域能量最小的结构。利用随机搜寻（random search）方法，或称为随机乱踢（random kick）方法，可将构型由势能面的一个区域移动至完全不相关的另一个区域。通常可由移动原子的笛卡儿坐标或是旋转分子内某根键改变分子的构型。图 11-12 为利用随机搜寻方法求分子能量最小构型的流程。

该方法中，每次迭代计算（iteration）得到新的构象后，再以此构象为新的起点，随机产生另一构象，并以此构象为能量最小化的起点，执行结构最优化。若结构最优化所得的构象不同于先前任何的构象，则储存此构象，并重复如上的步骤，直至设定的计算终点为止。由随机移动原子的笛卡儿坐标的方法产生新的结构的方法较为简单及容易程序化。但因其容易产生高势能构型，以致结构最优化过程造成分子的扭曲与变形。另一

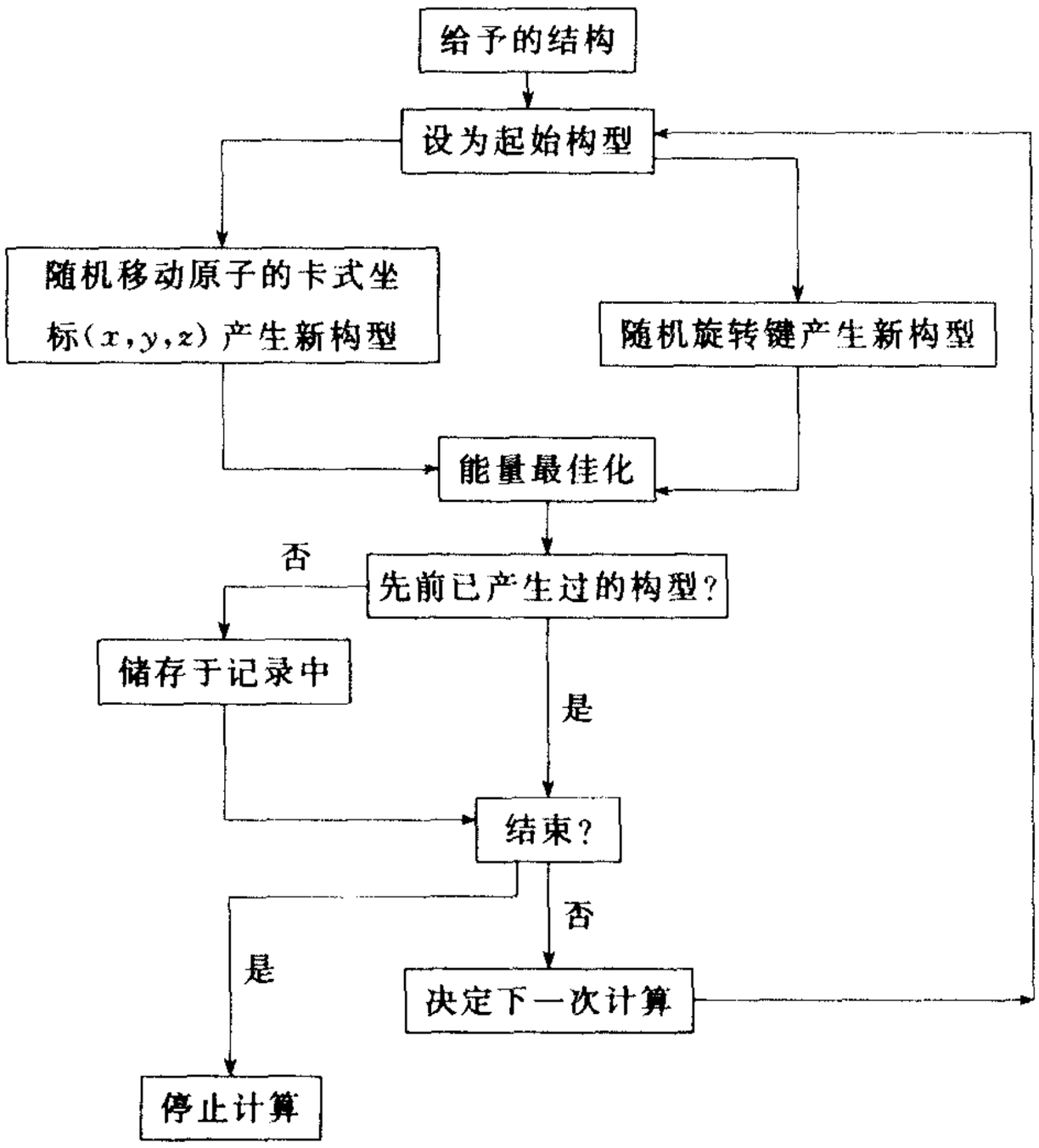



**图 11-12　随机搜寻方法求分子能量最小构型的流程**

种利用绕键转动改变扭转角而产生新构象的方法，必须选择适当的坐标，较为麻烦。尤其是当分子含有环状单位时，绕键转动易造成环的破裂，如图 11-13 所示。

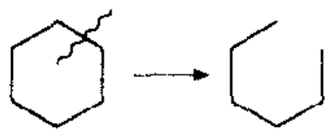


**图 11-13　绕键转动产生环的破裂**

此时必须在结构最优化过程中检查所有与环相关的内坐标以确定不会造成非环状的构象。如环已烷的转动必须检查接合处的键长与键角等内坐标，以确定无环破裂的现象。

寻找能量最低构象时，通常可由前一个被接受的结构为起点，按照上述的方法产生新的随机构形。或是由先前所产生的构象中，随机选取一个作为起点；亦可以选取次数的多寡决定统计比重。

第三种方法为由所产生的构象中选取能量最低的作为起点，并且利用 Metropolis 的蒙地卡罗方法。每一个随机产生并经过结构最优化的新构象，若其能量 $U(\text{new})$ 低于起始构象的能量 $U(\text{old})$，则此步骤被接受；若其能量高于起始构型能量，但玻耳兹曼乘子 $\exp(-\Delta U/k_{\mathrm{B}}T)$ 大于 0→1

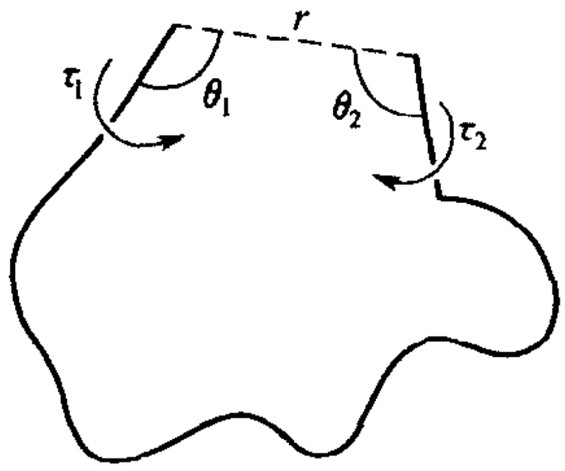



**图 11-14　环己烷接合处的内坐标**

的随机数，则此步骤可被接受；若能量较起始构象高，且玻耳兹曼因子又小于随机数，则此步骤被拒绝，由原先的起始结构重新产生新的随机结构。上述的这些方法，并无一定的优劣分别，需视实际的分子系统而定。环己烷接合处的内坐标如图 11-14 所示。

一般标准的求能量最小构象的方法是考虑所有可转动键造成的构象，而由其中找出最小的能量，有固定的计算终点。但以随机搜寻最佳结构方法找寻能量最小构象却无固定的终点。通常采用的方法为不断产生构象直至没有新的结构产生为止。因此，会造成不断产生重复的构象，而在相同的结构空间中来回探寻的现象。

利用结构最优化的方法可找出大分子的稳定构型。对非常复杂的大分子（如蛋白质、聚合物……）有时可利用退火淬火（annealing-quenching）的方法。这种方法由随机选取的结构为出发点，将模拟的温度提升，进行结构的最优化。当找出最优化结构后，以此结构为新的起点，将系统的温度降低，再执行蒙地卡罗模拟，以找出另一新的最优化结构。如此慢慢将系统的温度降低，并不断以蒙地卡罗方法找寻最优化结构，直至温度达到所要求的温度为止。这种方法的优点在于可避免产生的结构陷于区域极小值而无法翻越能量障碍的情形。

## 参　考　文　献


1　Johnson J K, Zollweg J A, Gubbins K E. Mol. Phys. 78, 591, 1993

2　Flory P J. "Statistical Mechanics of Chain Molecules", Interscience, Pub., New York, 1969

3　Engström W, Lindberg M. J. Chem. Edu., 65, 973, 1988

4　Bishop M, Michels J P J. J. Chem. Phys., 83, 4791, 1985

5　Rosenbluth M N, Rosenbluth A W. J. Chem. Phys., 23, 356, 1955

6　Wall F T, Erpenbeck J J. J. Chem. Phys., 30, 634, 1959

7　Wall F T, Mandel F. J. Chem. Phys., 63, 4592, 1975

8　Baumgärtner A. J. Chem. Phys., 72, 871, 1980

9　Webman I, Lebowitz J L, Kalos M H. J. Physique, 41, 579, 1980

# 第⑫章
# 蒙地卡罗计算的实例

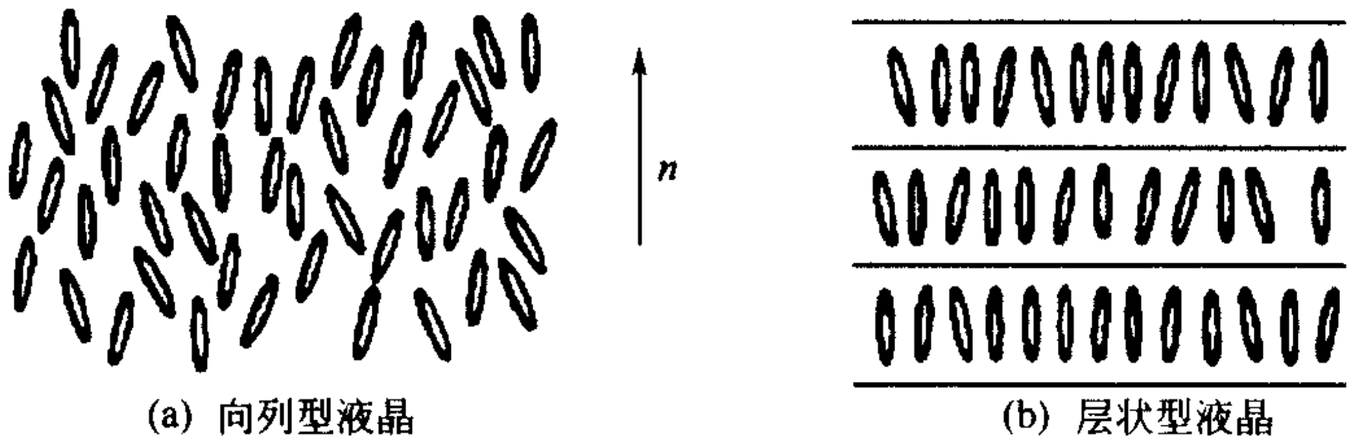
前面两章讨论蒙地卡罗方法的基本原理和算法，本章举出蒙地卡罗方法在几个实际体系的典型应用，包括：液晶系统、多相体系、气体分子在沸石的吸附与等温线的计算，以及蛋白质分子的折叠计算。蒙地卡罗方法可应用的系统很多，实用性很强。有兴趣的读者可参考 K. Binder 所编辑的“Monte Carlo Methods in Statistical Physics”与“Applications of Monte Carlo Method in Statistical Physics”等著作。

## 12.1 液晶的蒙地卡罗计算

液晶（liquid crystal）在化学材质中自成一类，约有 5% 的有机化合物具有液晶的特性。工业上液晶的应用很广，例如用以制造电子表的显示板、计算机的荧屏等。液晶的性质介于有序（ordered）的固体与无序的液体间。图 12-1 为两种常见的液晶形态。图 12-1(a) 中，分子多数沿 $n$ 的方向排列，称为向列型液晶（nematic liquid crystal）；图 12-1(b) 中，分子形成层状，每层沿同一方向排列，称为层状型液晶（sematic liquid crystal）。

由分子的角度看来，向列型液晶可视为具有长距离的排列秩序与短距离的移动秩序的系统。向列型分子一般为具有朝方向轴（$\bar{n}$）排列倾向的杆状分子，其液晶的秩序性不同于层状型液晶，向列型分子的质心较易沿方向轴方向呈不规则的移动。

Hemenway 等以简单的二维晶格蒙地卡罗计算模拟研究向列型液晶系统。图 12-2 为简单的向列型液晶系统的计算模型。此模型中以向量表示向列型分子，分子的质心固定于三角形的晶格点上，分子仅能转动而不能移动。设分子间的作用势能为：

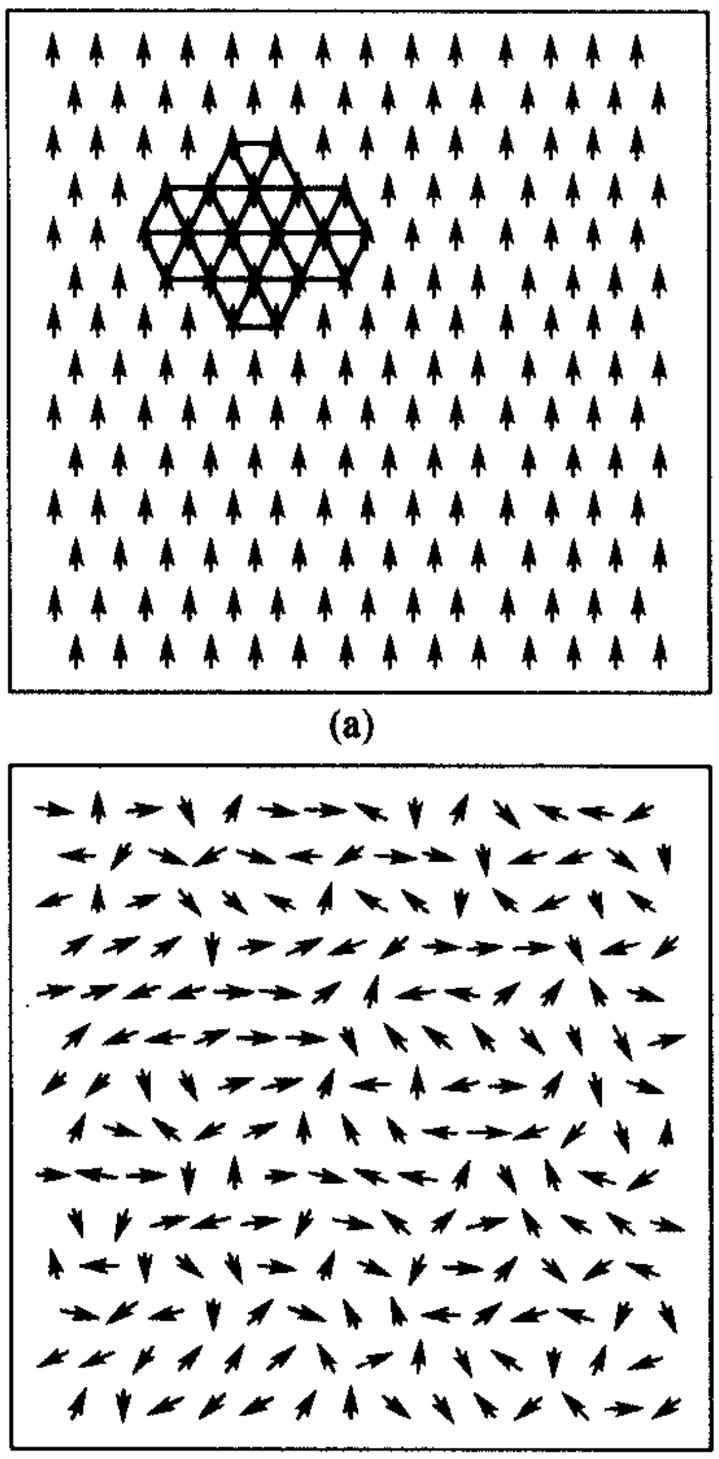



(a) 向列型液晶　　(b) 层状型液晶

**图 12-1　向列型与层状型液晶的示意**

$$u_{ij} = -\varepsilon(\cos 2\theta_{ij} + \mu\cos\theta_{ij}) \tag{12-1}$$

式中，$\theta_{ij}=\theta_i-\theta_j$，$\theta_i$ 与 $\theta_j$ 为第 $i$ 个分子向量与第 $j$ 个分子向量与方向轴的夹角；$\varepsilon$ 为设定的作用参数；$\mu$ 为介于0～1 的参数。$\cos 2\theta_{ij}$ 项显示杆状分子的非极性作用，含 $\mu$ 的项反应分子间的偶极作用；$\mu=0$ 表示非极性分子，$\mu=1$ 表示极性分子。



(a)

(b)

**图 12-2　二维液晶系统的计算模型**

设分子只与其最邻近的分子作用，则每个分子有 6 个最邻近的分子，其作用参数 $\varepsilon$ 为一正值常数。除了最邻近的分子外，其余的作用参数均为零。Hemenway 等模拟的系统共含 $14\times14$ 个晶格点，起始时将分子的向量朝同一方向排列，如图 12-2(a) 所示，此方向定为固定实验轴方向。

以 Metropolis 方法执行蒙地卡罗模拟的计算过程如下所述。

① 随机选择系统中的任一分子，并令其转动。

② 此转动分子与方向轴的角度为：

$$\theta_{new}=\theta_{old}+(\mathrm{Rand1}\times\delta\theta)$$

式中，$\theta_{new}$ 与 $\theta_{old}$ 分别表示此分子转动后与原先未转时与方向轴的夹角；Rand1 为介于 $-1$ 与 1 间的随机数，$\delta\theta$ 为转动的幅度。通常选取适当的 $\delta\theta$，使得产生新构型的接受率约为 50%。

③ 计算转动前后的势能差，

$$\Delta U=U(\theta_{new})-U(\theta_{old})$$

若 $\Delta U<0$，则转动被接受，否则产生一介于 0～1 的随机数 Rand2。

若：

$$\exp\left(\frac{-\Delta U}{k_B T}\right)>\text{Rand2}$$

则此转动被接受，否则回复原先的状态。

④ 若所产生的转动被接受，则计算系统的能量，比热容等物理量。检查这些计算的物理量的平均值是否收敛。

⑤ 若计算的物理量平均值未收敛，则回到步骤①继续模拟。

利用上述方法计算的物理量包括平均的分子内能、定容比热容、极性长距离秩序参数与非极性短距离秩序参数。设模拟系统共含有 $N$ 个分子，则系统的平均分子内能$\overline{U}$为：

$$\overline{U}=\frac{U_T}{M} \tag{12-2}$$

式中，$U_T$ 为系统的势能，即：

$$U_T=\sum_{i=1}^{M-1}\sum_{j>i}u_{ij} \tag{12-3}$$

除了计算系统内能的平均值外，亦可计算内能的统计涨落（fluctuation)。根据统计力学，由内能的统计涨落可计算系统的定容比热容，$C_V$，即：

$$C_V=\frac{\overline{U}_i^2-(\overline{U}_i)^2}{k_B T^2} \tag{12-4}$$

秩序参数(order parameter)可用以研究向列型液晶的方向性，此参数随温度的不同而改变。秩序参数通常分成非极性长距离(apolar long-range)与极性长距离(polar long-range)两种。计算秩序参数必须先得到系统分子的排列方向(director)。系统分子的排列方向与实验固定轴的夹角 $\alpha$ 的计算式为：

$$\tan 2\alpha=\frac{\sum_{i=1}^{M}\sin 2\theta_i}{\sum_{i=1}^{M}\cos 2\theta_i} \tag{12-5}$$

式中，$M$ 为液晶分子的总数。非极性长距离秩序参数与 $\alpha$ 的关系为：

$$\overline{T}_2=\frac{1}{M}\sum_{i=1}^{M}\cos 2(\theta_i-\alpha) \tag{12-6}$$

极性轴的位向可由计算位向角（$\alpha'$）求得，即：

$$\tan\alpha' = \frac{\sum_{i=1}^{M}\sin\theta_i}{\sum_{i=1}^{M}\cos\theta_i} \tag{12-7}$$

对应的极性长距离秩序参数为：

$$\overline{T}_1 = \frac{1}{M}\sum_{i=1}^{M}\cos(\theta_i - \alpha') \tag{12-8}$$

理论上，$\alpha$ 与 $\alpha'$的值应相同。若系统中的分子完全朝向同一方向排列，则$\overline{T}_1 = \overline{T}_2 = 1$；若系统中的分子无任何排列秩序，则$\overline{T}_2$ 与$\overline{T}_2$ 均趋近于零。亦有可能得到$\overline{T}_1 = 0$，$\overline{T}_2 < 0$ 的结果。图 12-3 为利用此种蒙地卡罗方法于各温度计算所得的平均分子势能，图中温度与能量均以简化单位表示，即：

$$\overline{U}^* = \frac{\overline{U}}{\epsilon}, \quad T^* = \frac{k_B T}{\epsilon} \tag{12-9}$$

图 12-3 显示平均的分子势能于 $2.0 < T^* < 3.0$ 的区间内有一剧烈的变化，这表示系统于 $T^* < 2.0$ 与 $T^* > 3.0$ 的两个区间所处的“相”(phase) 不同。

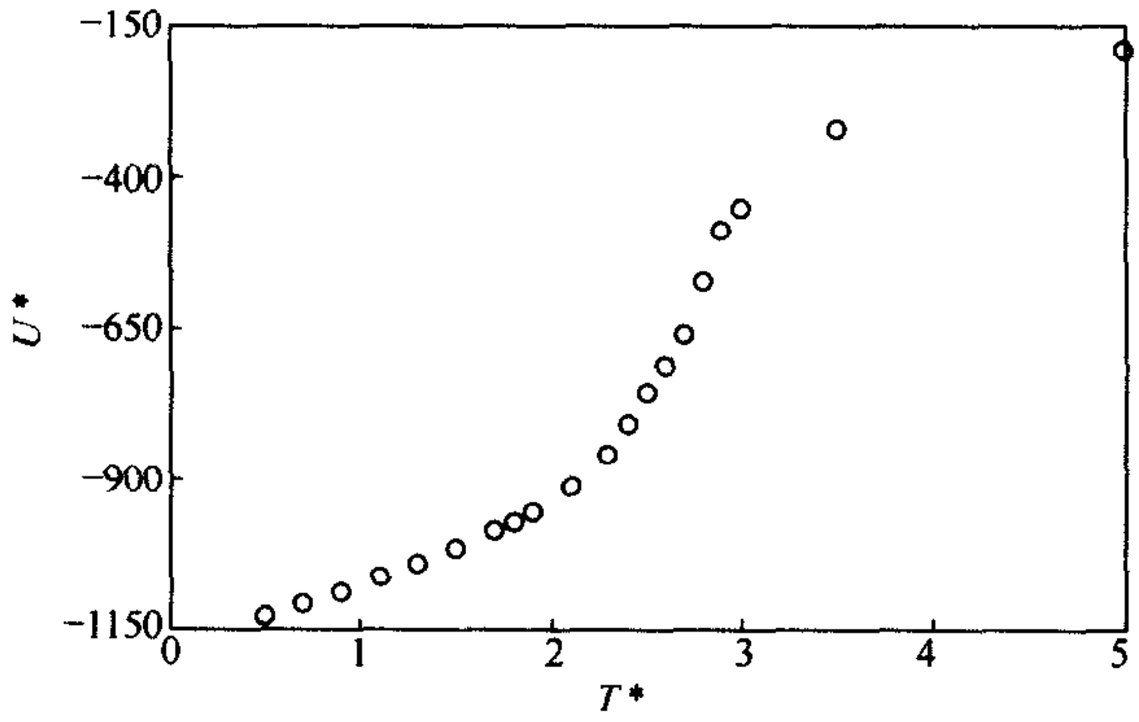



**图 12-3　计算液晶分子系统的平均势能 （坐标为简化单位）**

图 12-4 为计算的定容比热容与温度的关系图，定容比热容简化单位表示为：

$$C_V^* = \frac{C_V}{k_B} \tag{12-10}$$

图中清楚地显示 $T^* < 2.0$ 与 $T^* > 3.0$ 表示两种不同的相。

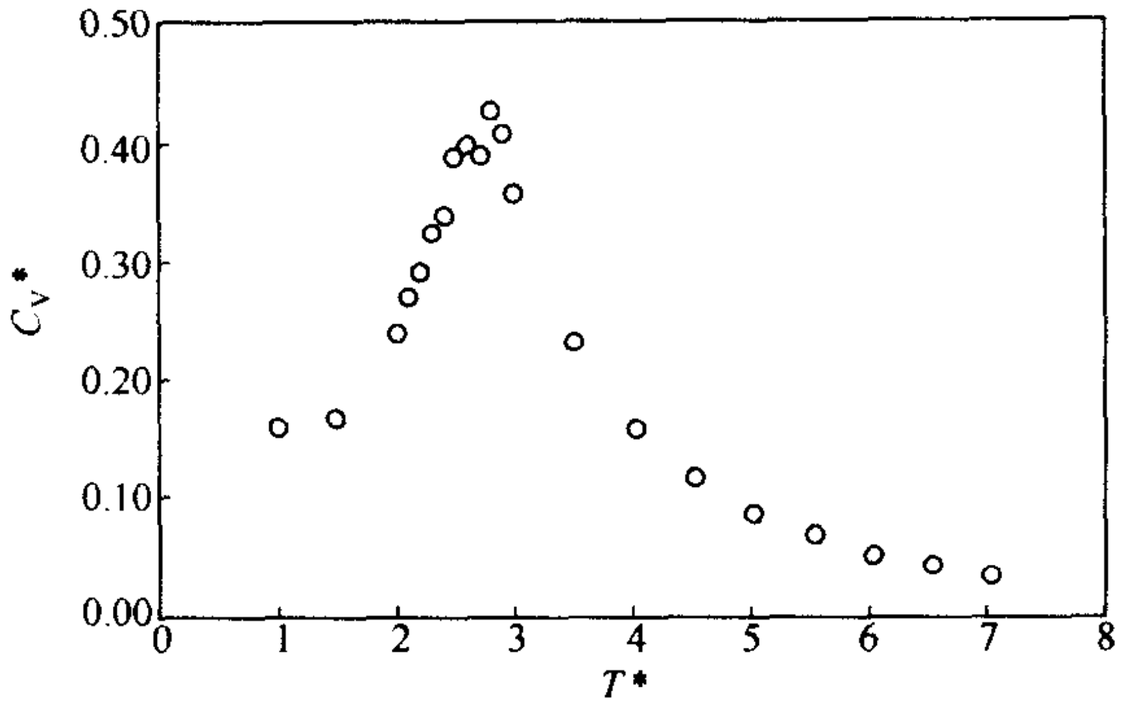



**图 12-4　计算液晶分子系统的定容比热容（坐标为简化单位）**

图 12-5 为计算的极性与非极性长距离秩序参数。图中显示，当简化温度小于 2.0 时，$\overline{T}_1$ 与 $\overline{T}_2$ 的值存在非常明显的差异，但是，当简化温度大于 3.0 时，$\overline{T}_1$ 与 $\overline{T}_2$ 的值几乎完全相同。这表示高温时，系统的秩序降低，即使是极性分子，其秩序参数一如非极性分子，此为液体的特性。但当低温时，极性秩序参数大于非极性秩序参数，分子排列有序，此为晶体的特性。因此，此系统随温度的增加，由晶体转换为液体，显现液晶材质的标准特性。而 $2.0 < T^* < 3.0$ 的区间为相转换温度区间。

以上的范例利用蒙地卡罗方法及简单的模型模拟液晶系统的特性，得到与实验一致的结论。若适当的调整计算中所用的参数，如 $\epsilon$、$\mu$ 等，或引用更精确的计算模型，如推广至三维空间或考虑分子的移动，则模拟的结果应更接近真实分子的实验结果。

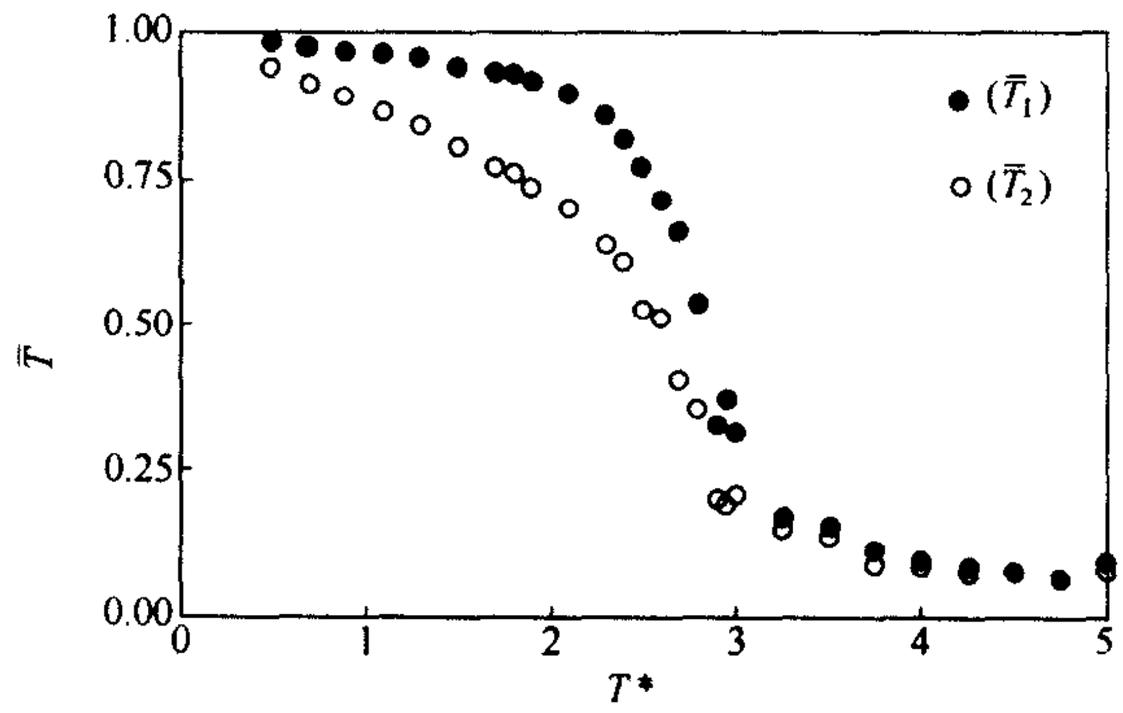



**图 12-5　计算液晶分子系统的秩序参数**

## 12.2 两相并存系统的蒙地卡罗计算

单一成分或混合成分的流体系统，其相与相间的平衡性质为工业上分离过程的重要依据。若能建立流体系统的相图（phase diagram），则可适当地控制系统的温度与压力等达到分离的目的，或者可以由物质在各相的成分比例预测分离的效率。两相共存系统的蒙地卡罗模拟方法是将系统假想为两个区域，如图 12-6 所示。

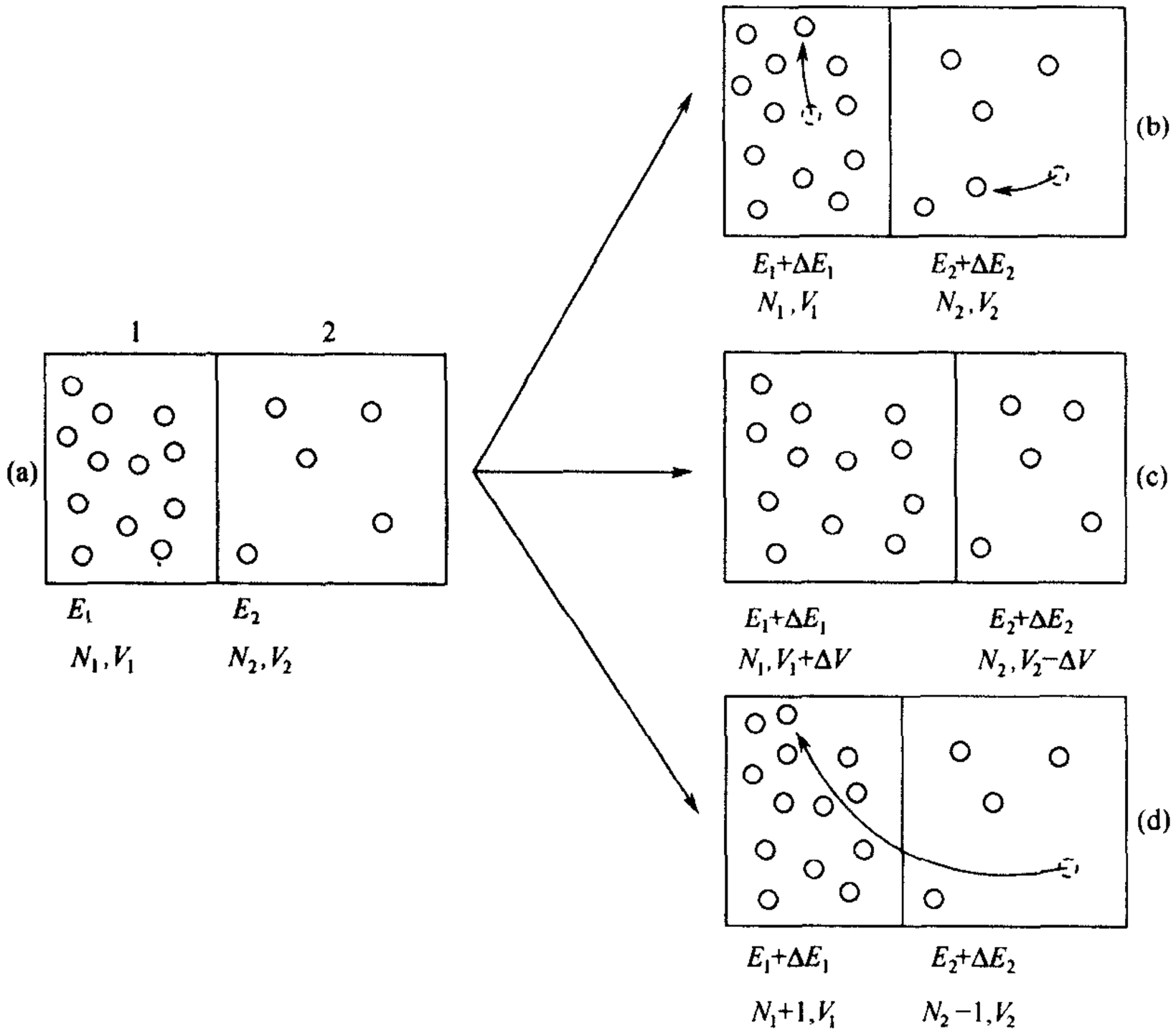



**图 12-6　二相并存系统计算示意**

图 12-6 中，(a) 为模拟系统的起始结构，图中的 1，2 两个区域分别表示系统的两个相。将整个系统视为 $(N,V,T)$ 系综，图中分离二区间的界线仅为人为的分隔而无重要的物理意义。

由图 12-6(a)→图 12-6(b) 的计算为一般的 $(N,V,T)$ 条件的计算，由随机数选取法移动二区间的粒子，但各区间的体积与粒子数目不

变。对于这样的封闭系统（closed system），粒子移动被接受的概率为：

$$P \propto \exp(\Delta S_t / k_B) = \exp(-\Delta E_{min} / k_B T) \tag{12-11}$$

式中，$\Delta S_t$ 为系统熵（entropy）的改变量；$\Delta E_{min}$ 为由平衡状态至变化状态的可逆变化所需最小的功。依此，在热平衡条件下，状态 $x$ 与状态 $x'$ 互相变换的概率比为：

$$\frac{P(x \to x')}{P(x' \to x)} = \exp(-\Delta E_{min} / k_B T) \tag{12-12}$$

根据式(12-12)，执行蒙地卡罗模拟时，可将粒子移动被接受的概率设为：

$$P = \min[1, \exp(-\Delta E_{min} / k_B T)] \tag{12-13}$$

故粒子移动后，系统能量的变化为：

$$\Delta E_{min} = \Delta E_1 + \Delta E_2 \tag{12-14}$$

这样的计算结果与分别计算系统的两个区域所得到的结果相同，为两个独立的 $(N,V,T)$ 系统的计算。

由图 12-6(a)→图 12-6(c) 的计算，当粒子移动时，系统的总体积不变，但容许各区域的体积变化。由此假设，区域 1 体积的变化与区域 2 体积的变化等值而异号。若仅考虑移动区间 1 的粒子，则移动粒子后，各区域的结构根据其新的体积而定标（scaling）。此方法如同前述 $(N,P,T)$ 系综的模拟。设二区域能量的变化分别为 $\Delta E_1$ 与 $\Delta E_2$，则：

$$\Delta E_{min} = \Delta E_1 + \Delta E_2 - N_1 k_B T \ln \frac{V_1 + \Delta V}{V_1} - N_2 k_B T \ln \frac{V_2 - \Delta V}{V_2} \tag{12-15}$$

式(12-15) 的后两项为理想气体体积变化熵对能量的贡献。因系统的温度、体积与粒子数均维持不变，故移动粒子的接受概率仍为式(12-13)。这样的计算，系统中的二区域借体积的变化以维持压力的平衡，此压力有别于外加的压力，而应视为系统的内压。

此方法亦可用于 $(N,P,T)$ 系综的计算，若区域 2 的体积远大于区域 1 的体积，且其性质不受体积改变的影响（如气体），则依式(12-15)，移动粒子可接受的概率为：

$$P = \min\left\{1, \exp\left[-(\Delta E_1 - N_1 k_B T \ln \frac{V_1 + \Delta V}{V_1} + P_{ex} \Delta V) / k_B T\right]\right\} \tag{12-16}$$

式中，$P_{ex}$ 为外压。对区域 1 而言，如此的计算相当于 $(N,P,T)$ 的计算。

由图 12-6(a)→图 12-6(d) 的计算为将区域 2 的粒子移往区域 1（或由区域 1 移往区域 2）。实际的模拟方法为在区间 2 中任意选取一个粒子使其消失，而于区间 1 中于任意位置产生一个新的粒子。这样的移动方法所对

应的最小能量变化为：

$$\Delta E_{\min}=\Delta E_1+\Delta E_2+N_1 k_B T\ln\frac{N_1+1}{N_1}+N_2 k_B T\ln\frac{N_2-1}{N_2}$$

$$+k_B T\ln\frac{V_2}{N_2-1}-k_B T\ln\frac{V_1}{N_1+1} \tag{12-17}$$

而粒子移动被接受的概率仍为式(12-16)。这样的计算，系统中的二区域维持压力与化学势能（chemical potential）的平衡。此方法亦可用于多成分系统的模拟，计算中维持式(12-18）的平衡条件。

$$P_1=P_2$$

$$\mu_{i,1}=\mu_{i,2} \tag{12-18}$$

式中，$\mu_i$ 为第 $i$ 种成分物种的化学势能。这样的系综特称为吉布斯系综（Gibbs ensemble)。图 12-7 为 Panagiotopoulos 以这种方法在简化温度为 1.15 的条件下，对总数为 300 个粒子系统模拟的结果。二区域的起始简化密度均为 $\rho^*=0.2$，粒子间的作用力均为 LJ 作用。

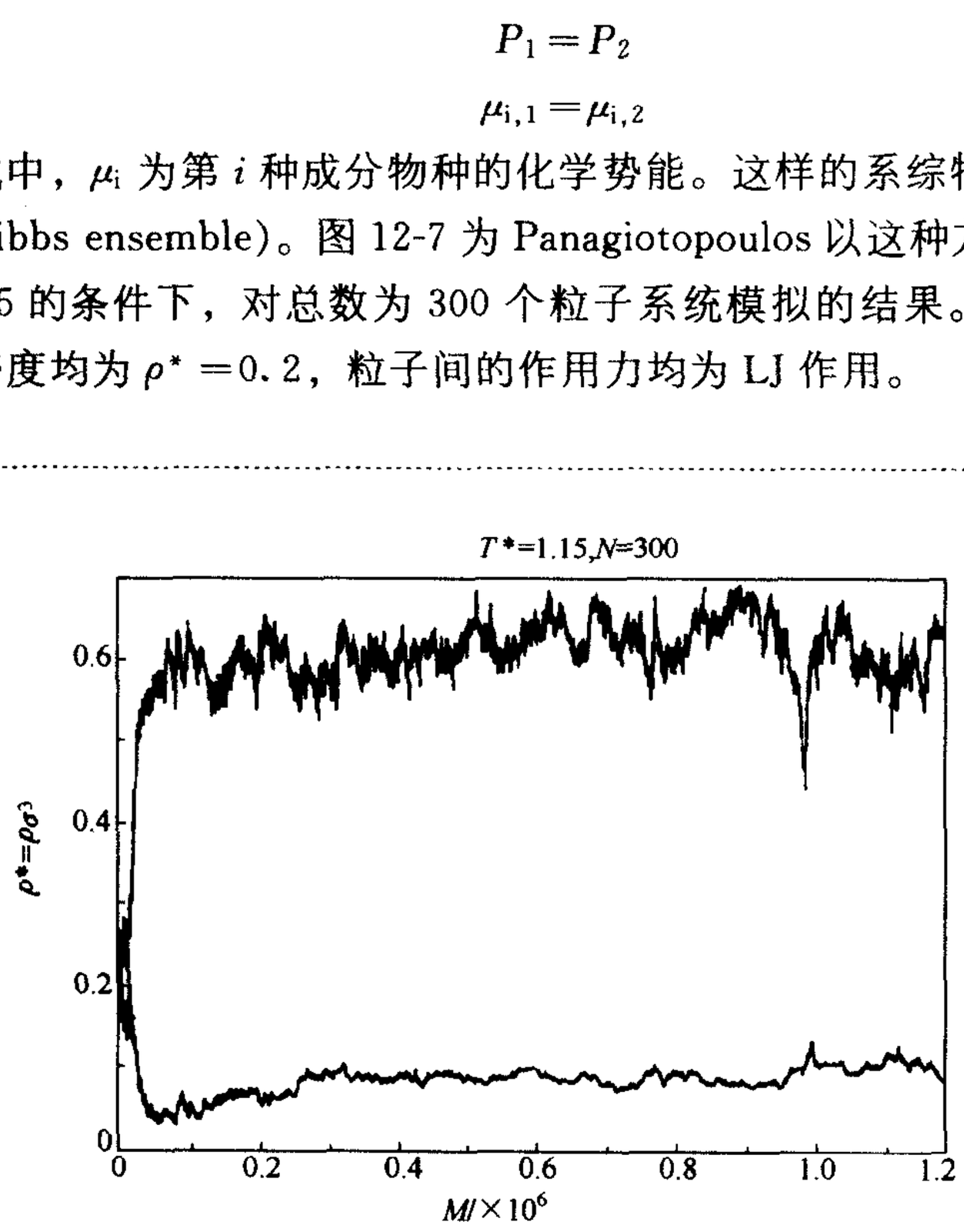



**图 12-7　蒙地卡罗计算两相系统的密度变化（坐标为简化单位）**

图中，$M$ 为模拟所产生的结构数目。图 12-7 显示于此条件下二区域呈现两种不同的密度，密度低者为气相，密度高者为液相。此为液相与气相共存系统的模拟。图 12-8 为 Panagiotopoulos 以不同数目（300，500）的 LJ 粒子模拟不同温度的二相共存系统的结果。图中并将结果与 Adam 利用（$\mu VT$）系综的蒙地卡罗模拟的结果互相比较。

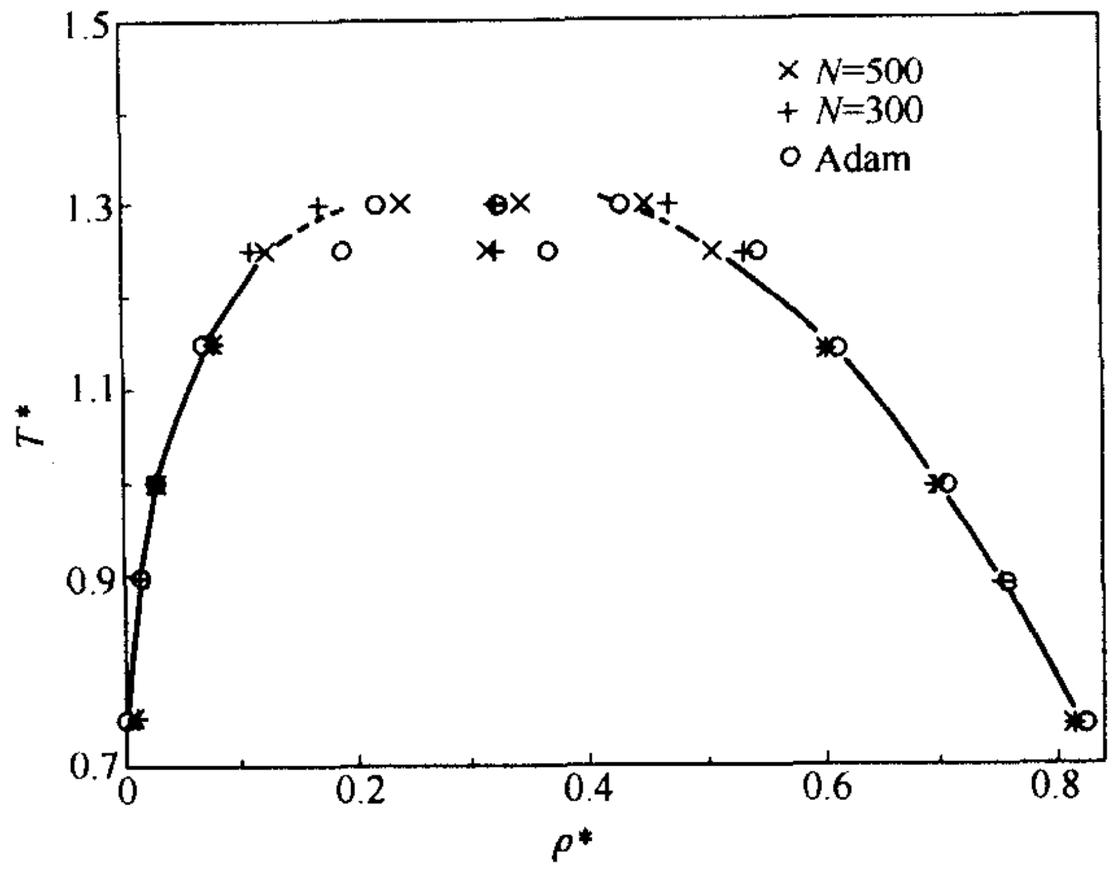



**图 12-8　两相系统简化温度与简化密度的关系（坐标为简化单位）**

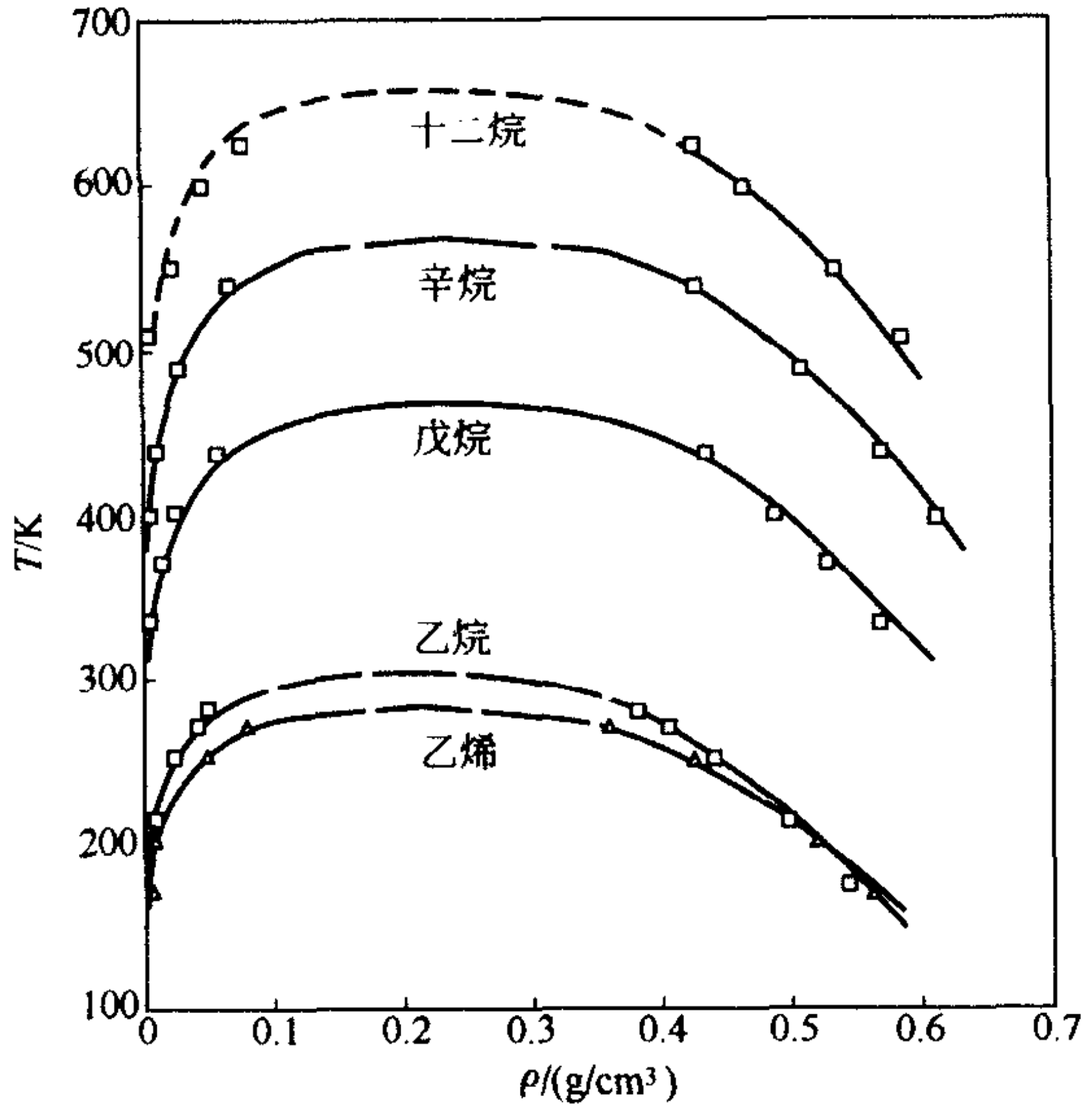



**图 12-9　各种碳氢化合物模拟的相图**

当两相的密度相同时的温度为临界温度（critical temperature），此时的密度为临界密度（critical density）。依图 12-8，临界温度的简化值约为 1.28～1.3，临界密度的简化值约为 0.31±0.02。此例显示，利用简单的模拟方法即可定出系统的临界条件，对一些混合物系统具有高度的应用价值。图 12-9 为 Nath 等利用 NERD 力场（S. K. Nath，F. A. Escobedo，J. J. de Poblo）与吉布斯系综模拟各种碳氢化合物所得的二相图。图中实线部分为实验所得，虚线部分为由已知的资料外插所得。由此图可得各碳氢化合物系统的临界温度与密度。

## 12.3 沸石等温吸附线的计算

沸石（zeolites）为无机聚合物分子所形成的固体，通常含有许多孔洞（micropores），如图 12-10 所示。这些沸石的孔洞可容纳各种分子，称为客体分子（guest molecules）。沸石具有很大的内表面积，其体系稳定度高，并且含有数以千计的酸性基点（acid sites）。这些因素使得沸石成为石化工业重要的催化剂。为了充分发挥沸石的功用，必须了解吸附于沸石孔洞中分子的各种行为。实验上，通常以核磁共振光谱方法侦测沸石中吸附分子的行为，但这样的侦测至今仍存在许多困难。理论计算乃成为解决此问题可行的方法。Stroud 等首先研究沸石中的吸附分子的热力学性质；其余的以模拟方法研究沸石中分子的吸附行为可参考 Catlaw 所编辑的“Modeling Structure and Reactivity in Zeolite”一书。

利用巨正则系综的蒙地卡罗方法可计算体积小的客体分子（如甲烷、惰性原子）于沸石中的等温吸附线（isotherm）。图 12-11 为 Goodbody 等以此方法计算甲烷于硅酸沸石（zeolite silicalite）的等温吸附线。图中的纵轴为分子的吸附量，横轴为外施压力，黑色的符号为实验所量测的吸附量，空心圆则为蒙地卡罗计算的吸附量。此图显示计算与实验的结果十分吻合，可经由计算预测类似的小分子的吸附。

对长链的烷类（如丁烷或更高的烷）而言，欲成功产生合理的计算条件十分困难。通常将分子插入沸石的孔洞中，客体分子的原子总会与沸石原子重合。因此，需要花费许多计算时间以避免这样的情况出现，计算效率很低。解决这样的问题，则需利用特殊的方法，其中一种计算法为利用 Rosenbluth 等早期所提出的构型选择（configurational-bias）蒙地卡罗

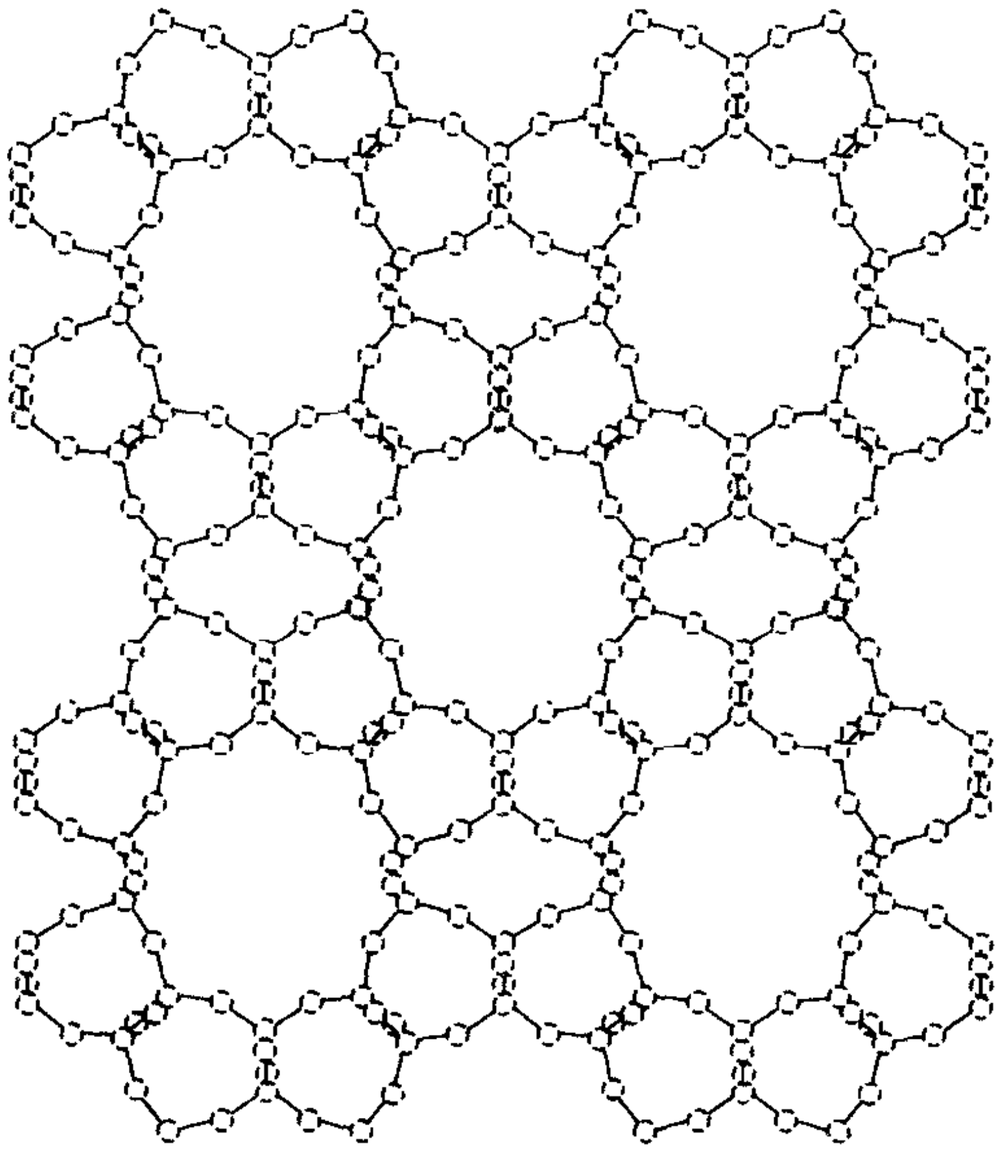


**图 12-10　沸石结构示意**

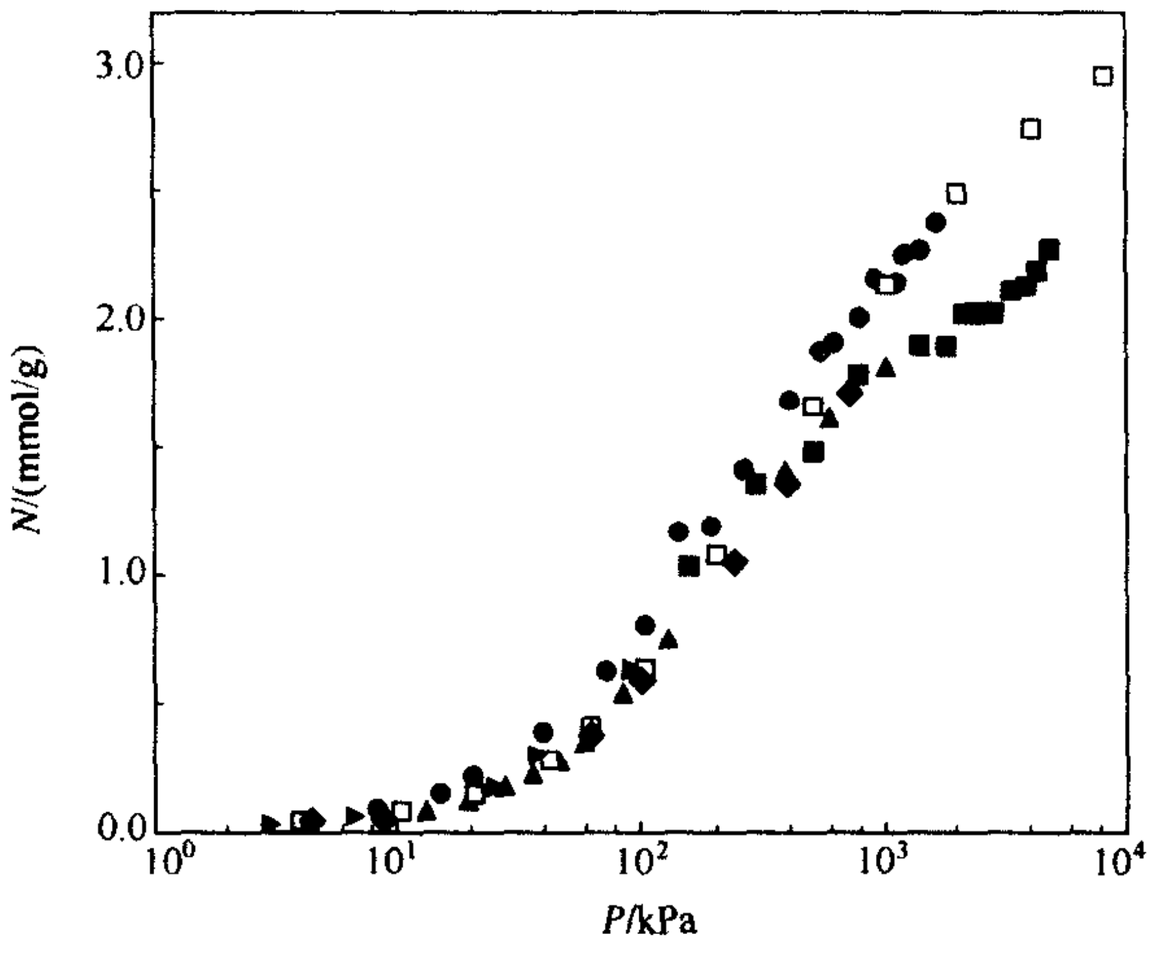



**图 12-11　甲烷于沸石的吸附量与压力关系图**

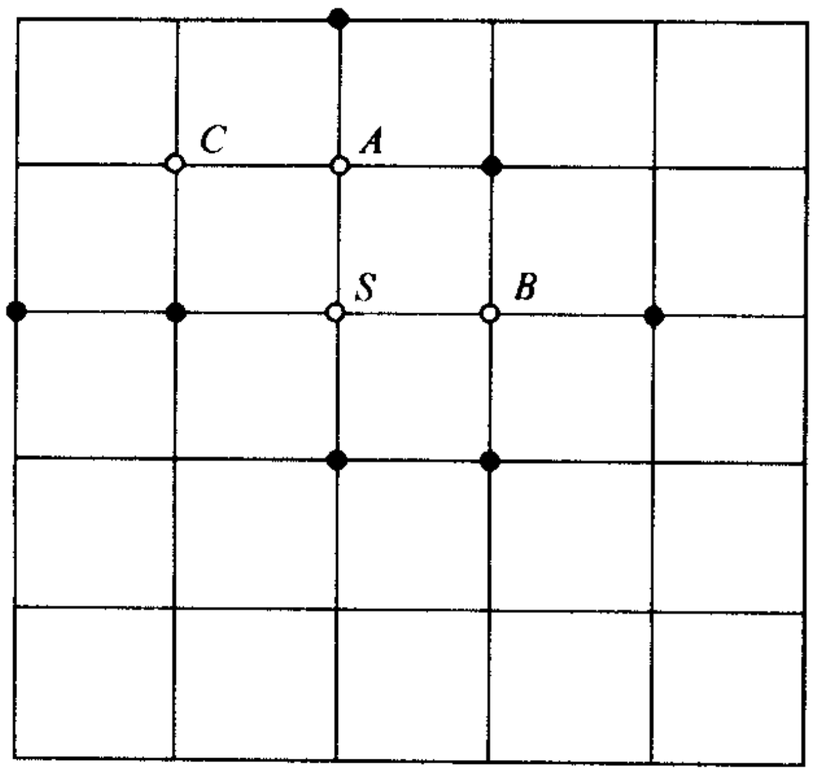



**图 12-12　二维晶格模式构形选择 MC 图**

计算。这种方法的重心是当模拟产生分子时，选择可能的构型而非随意产生。此处以二维的晶格模式说明这种计算法。

假设在如图 12-12 的二维晶格面上欲产生一个三单位的分子，一般的蒙地卡罗方法为先随意产生第一个粒子，如图中的 $S$。然后随意选取 $S$ 邻近四个格子点上的任一个位置。此时这四个位置中已有两个被别的粒子所占据，而仅有 $A$、$B$ 可选择。这种状况下产生第二个粒子被接受的概率为 50%。但若选择了位置 $B$，则因为所有邻近的位置均已被占据，$n$ 无法再产生第三个粒子。假设选取的位置为 $A$，则其邻近的三个位置中仅有 $C$ 未被占据，可被接受。此例显示一般的蒙地卡罗方法，由 $S$ 产生三单位分子被接受的概率仅为 1/12。构形选择的蒙地卡罗方法是先随意产生第一个粒子，设仍位于 $S$。其次，在 $S$ 的邻近位置中先选择可能的位置 $A$ 与 $B$（而非随意选择 $S$ 的邻近位置），然后在 $A$ 与 $B$ 中随意选取一个作为第二个粒子的位置。此时计算所谓的 Rosenbluth 统计比重（Rosenbluth statistical weight）

$$W_i = \frac{n'}{n} W_{i-1} \tag{12-19}$$

$W_{i-1}$ 为前一步的统计比重，$W_0 = 1$；$n'$ 为所有可能产生粒子的位置数，$n$ 为不含前一步粒子所有可能的位置数。此例中，$W_1 = 2/4 = 1/2$。若选择了 $B$ 位置，则不可能产生第三个粒子，此移动必须被排除。若选择位置 $A$，则检查其邻近的自由位置，仅有一个位置可用以产生第三个粒子。此步骤的统计比重为 $W_2 = (1/3) \times (1/2) = 1/6$。总的统计比重为各产生步骤统计比重的乘积，$(1/2) \times (1/6) = 1/12$，与一般蒙地卡罗方法的统计比重相同。所不同的是，一般蒙地卡罗方法需要产生 12 次位置，其中仅一次被接受，而构型选择法仅需产生两个位置，其中一个可被接受。

将此种构型选择方法应用于链分子的产生，并考虑产生粒子链本身的能量与粒子与环境的作用。设占据特殊位置 $\Gamma$ 的第 $i$ 个粒子其能量为 $U_\Gamma(i)$，则在此位置产生此粒子的概率为：

$$p_\Gamma(i)=\frac{\exp[-U_\Gamma(i)/k_BT]}{Z_i} \tag{12-20}$$

式中，$Z_i$ 为产生粒子所有可能位置 $b$ 的玻耳兹曼乘子的和，即：

$$Z_i=\sum_{\Gamma=1}^{b}\exp[-U_\Gamma(i)/k_BT] \tag{12-21}$$

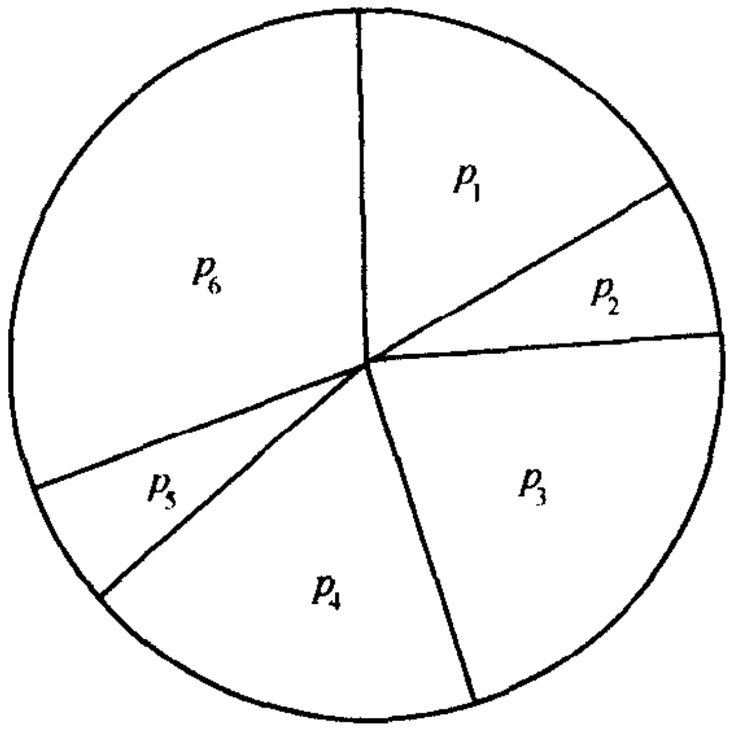



**图 12-13　以 roulett-wheel 的方法产生位置的概率分布示意**

通常可由所谓的 roulett-wheel 方法产生新的位置。此方法将 0～1 的区间分成 $b$ 个相邻的部分，每部分的大小正比于其概率 $p_1(i),p_2(i),\cdots,p_b(i)$，如图 12-13 所示。所产生 0～1 间的随机数若属于图中的某个区间，则产生的粒子于此区间所对应的位置。以此方法依较高的玻耳兹曼统计比重产生链上的各个粒子。整根链（$l$ 个粒子）的 Rosenbluth 统计比重为：

$$W_l=\exp[-U_{total}(l)/k_BT]\prod_{i=2}^{l}\frac{Z_i}{b} \tag{12-22}$$

式中，$U_{total}$ 为整根链的能量，等于各粒子的能量和 $U_\Gamma(i)$。

依照统计力学，统计比重的平均值正比于超化学位能（excess chemical potential），即：

$$\mu_{ex}=k_BT\ln\langle W_l\rangle \tag{12-23}$$

在应用 roulett-wheel 方法时，若产生的粒子其 Rosenbluth 比重为零，则此产生的步骤终结。但在计算比重的平均值时，却必须将这样的链计入以得到正确的超化学位能。若考虑真实的系统（如沸石、表面等），产生粒子的位向为 $\Gamma$，其能量为 $U_\Gamma(i)$，则其概率为：

$$p_\Gamma(i)=\frac{\exp[-U_\Gamma(i)/k_BT]}{\sum_{\Gamma=1}^{k}\exp[-U_\Gamma(i)/k_BT]} \tag{12-24}$$

而对应的 Rosenbluth 比重因子为：

$$W_i=W_{i-1}\frac{1}{k}\sum_{\Gamma=1}^{k}\exp[-U_\Gamma(i)/k_BT] \tag{12-25}$$

构型选择的蒙地卡罗方法以逐步产生或消除链上的粒子模拟将分子插入（insert）或移除（remove）的步骤。将链分子的势能分为键结势能与作用势

能两部分，前者表示分子内部的作用，后者则表示分子链间、远程作用等分子外的作用。定义理想链为仅有内部作用的链。每次产生或消减粒子时均计算其键结势能与作用势能以决定步骤是否被接受。当完全产生或消减链上的全部粒子后，依一般的蒙地卡罗方法决定此链被插入或移除的接受概率。图12-14 为 Smit 等利用这种方法计算烷类分子在硅酸沸石中的等温吸附线。

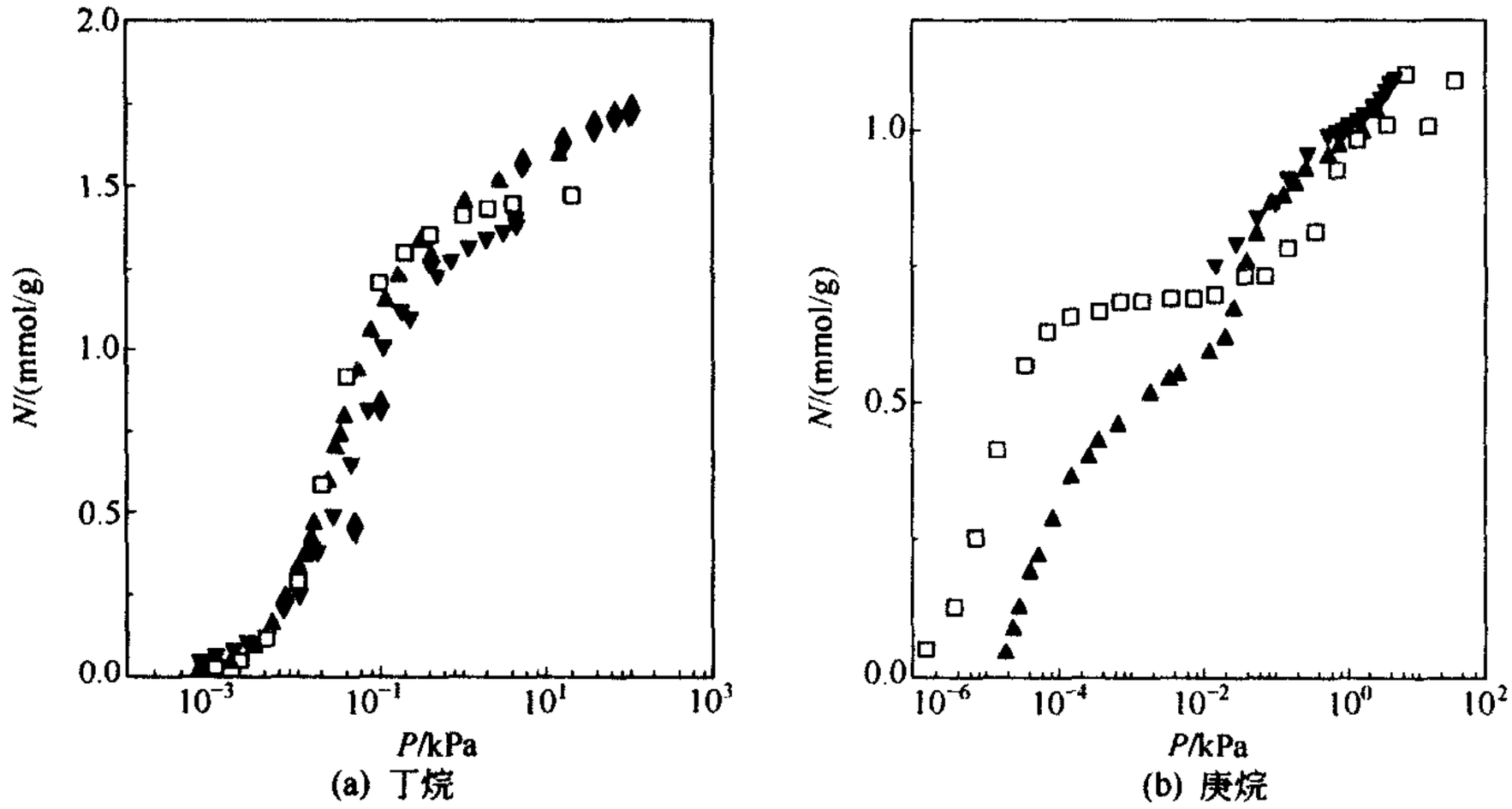



**图 12-14　烷类分子在沸石的吸附量对压力图**

□为模拟所得，▲为实验所得

图 12-14 中，图(a) 为丁烷分子的吸附，图(b) 为庚烷分子的吸附。图(a) 显示计算丁烷的吸附曲线与实验非常吻合；但图(b) 显示计算庚烷的吸附曲线仅于高压时方与实验吻合。这表示这种简单的计算模型仍有需加强改进之处。

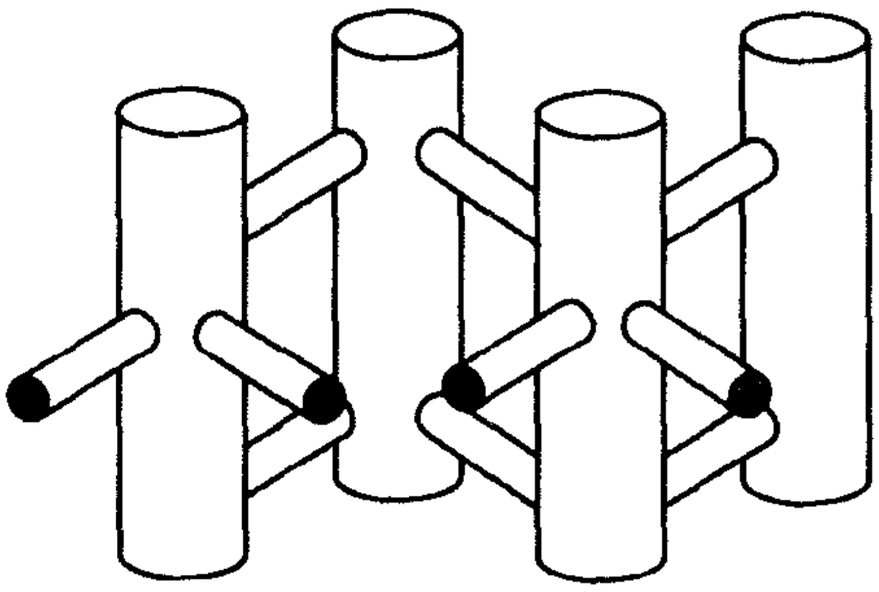


**图 12-15　沸石孔洞管道示意**

硅酸沸石有圆柱状与折曲形两种孔洞管道，彼此贯连，如图 12-15 所示。

分析 Smit 等的计算结果可发现短的烷类分子（如丁烷）在这两种孔洞管道间的分布概率约略相同，但较长的烷类分子则较易分布于直桶状的孔洞管道中。正已烷分子的长度约等于折曲状孔洞管道的大小，在低压的状况下，正已烷分子可在折曲孔洞管道中自由运动，

且有部分时间停留于两种孔洞管道的交接处。因此，在沸石中填入如正己烷的长链分子，应先将其填入折曲孔洞管道而非交接处。这种位置的限制使系统的熵降低，化学位能升高，因而在图 12-15 的吸附曲线上产生扭结的现象。

## 12.4 蛋白质的折叠

蛋白质折叠（protein folding）为生化上非常重要的问题。折叠的动力

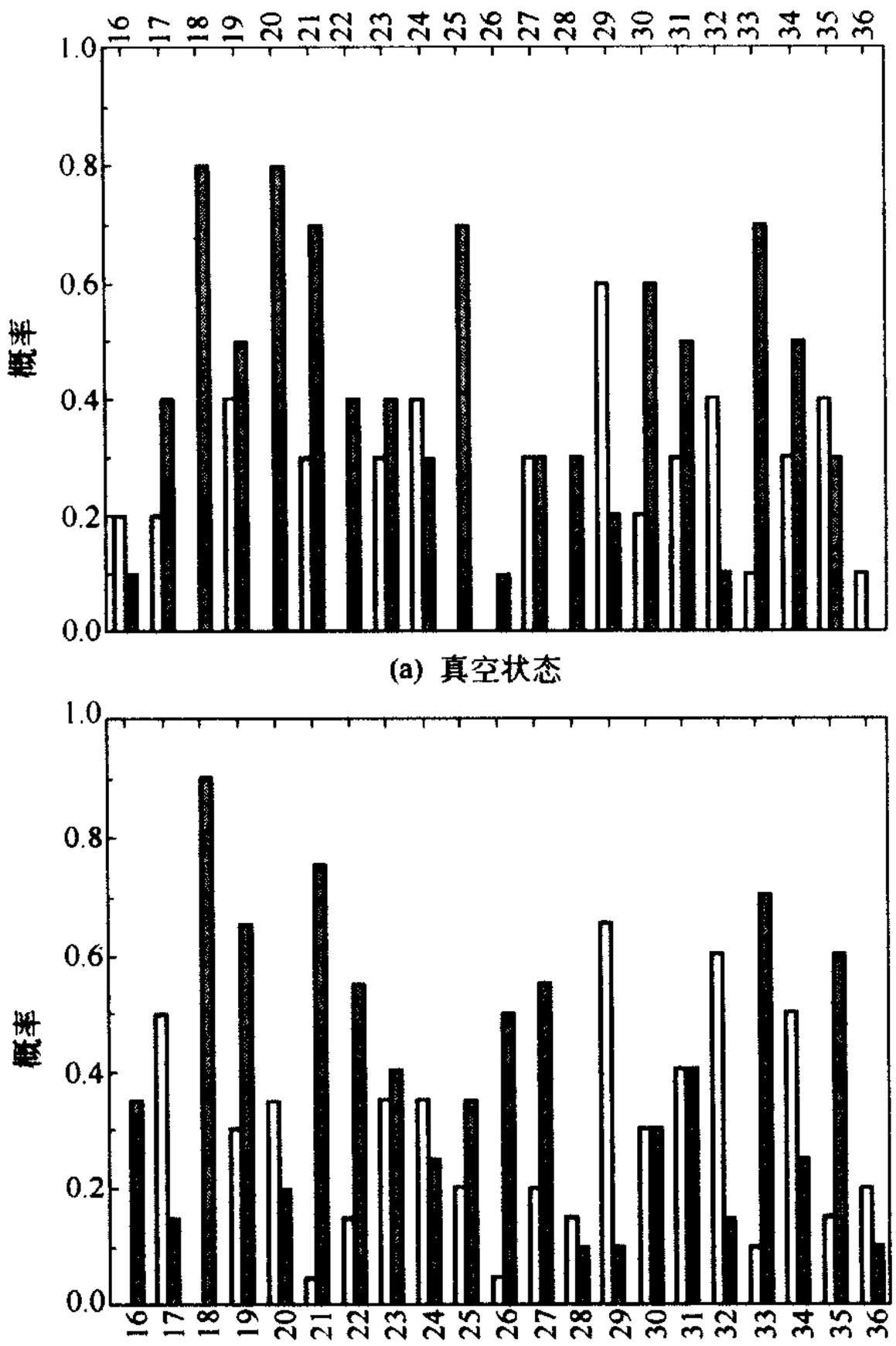



**图 12-16 BPTI (16-36) 蛋白质片段形成 α 螺旋与 β 折叠的倾向性**

空白的块状为形成 α 螺旋的倾向性，填满的块状表示形成 β 折叠的倾向性

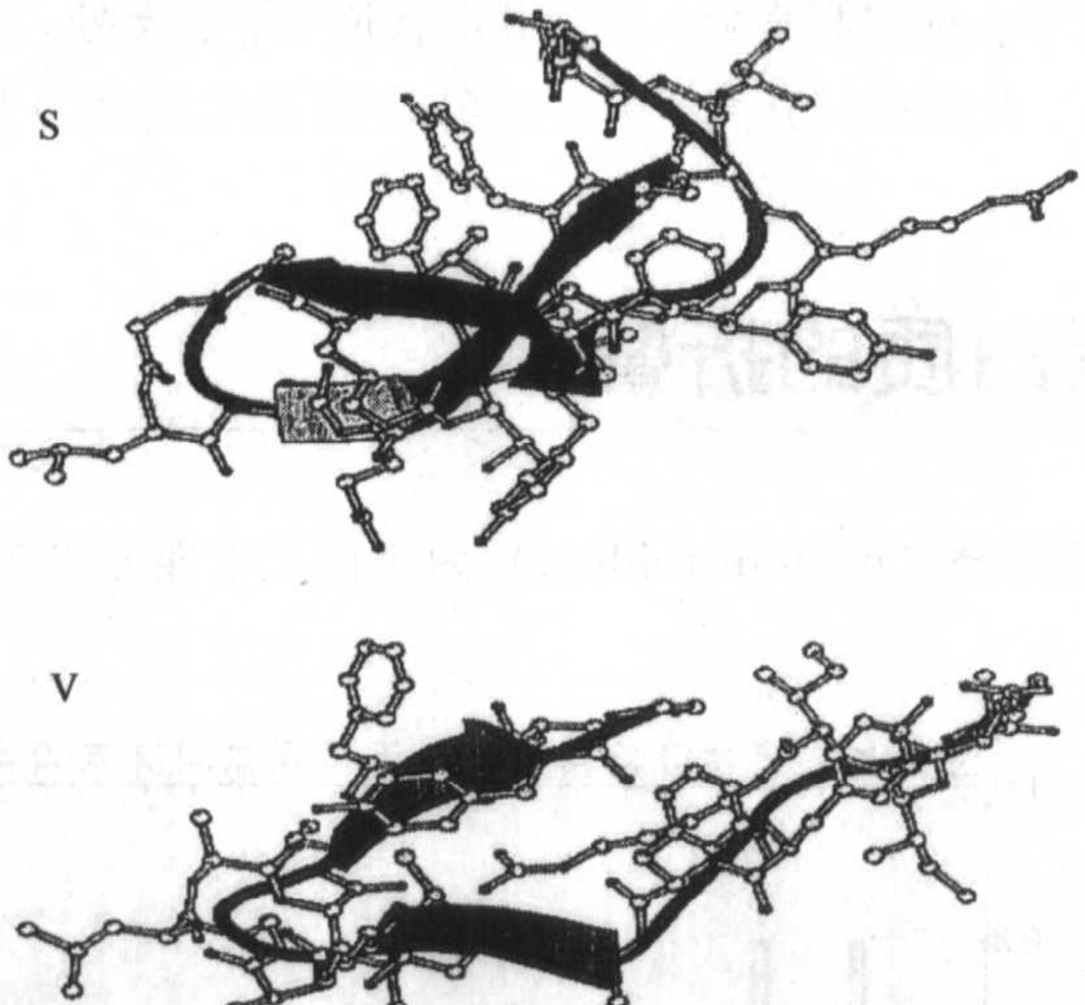
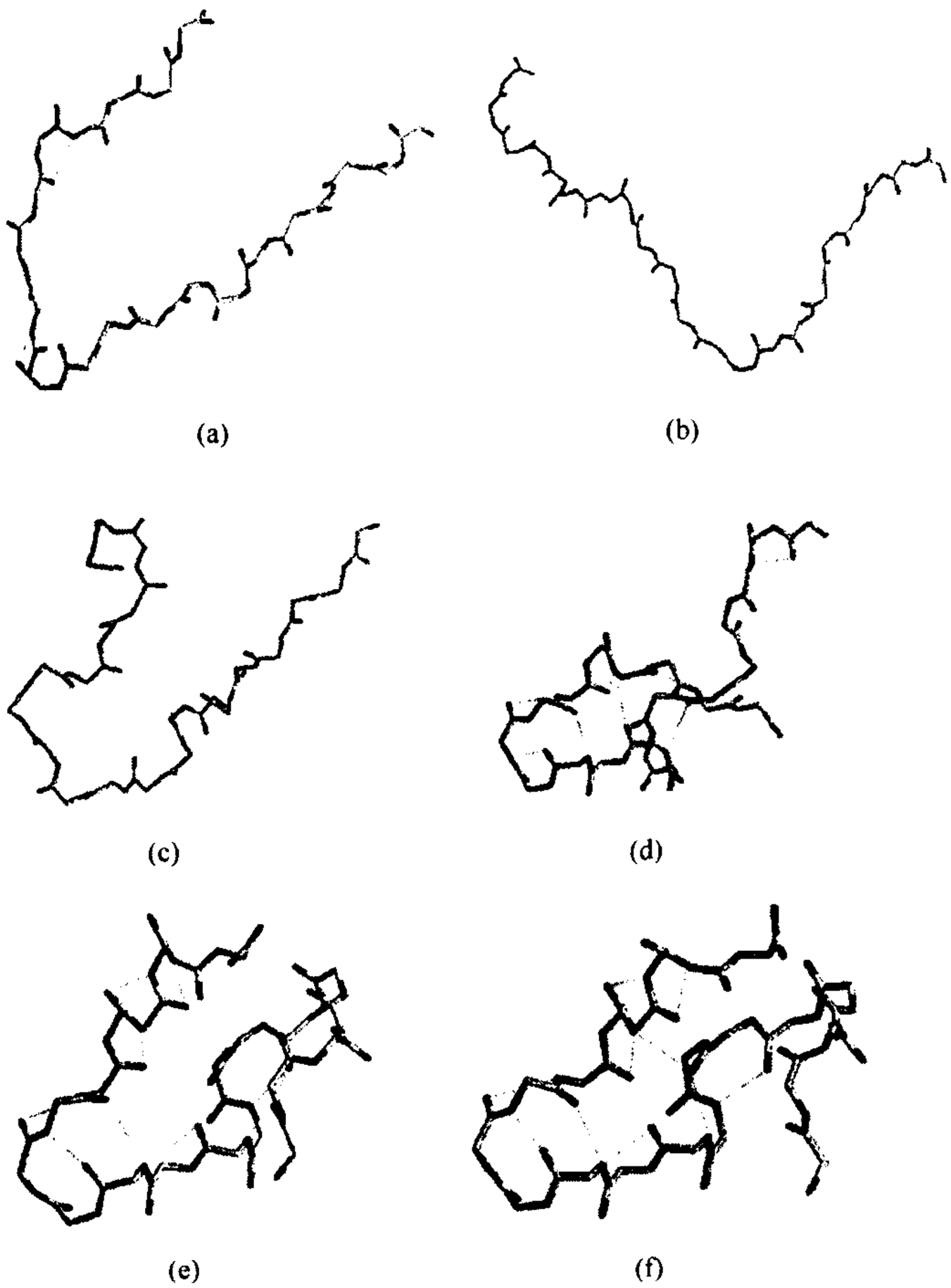



**图 12-17　模拟的 BPTI 片段与 X 射线衍射实验结构的比较**

S：在水溶液中；V：真空中；X：X 射线结构

过程与各种热力学的特性为现今最热门的研究领域。蛋白质分子由一系列的氨基酸按照一定的排列顺序所组成，不同的排列顺序构成不同的蛋白质。蛋白质分子通常拥有非常复杂的空间立体结构，必须利用各种侦测的方法（如 NMR、X 射线衍射等）以确定其结构。但一般的蛋白质分子其分子量非常庞大，即使利用先进的仪器与技术亦需很久的时间才能定出确切的结构。利用分子模拟的方法探讨蛋白质的结构与折叠的过程，可以节省许多时间与工作，为非常重要的研究工具。

Okamoto 等应用退火淬火（annealing-quenching）的蒙地卡罗方法模拟特殊的蛋白质分子片段 BPTI（16-36）的结构与折叠的过程。这种蛋白



(a) (b) (c) (d) (e) (f)

**图 12-18　由随机选取的结构形成 S 状折叠的过程**

质分子的氨基酸序列为：

Ala-Arg-He-He-Arg$^{+}$-Tyr-Phe-Tyr-Asn-Ala-Lys$^{+}$-Ala-Gly-Leu-Cys-Gln-Thr-Phe-Val-Tyr-Gly

其中的残基（basic residues）Arg 与 Lys 带有正电荷。模拟过程中将系统的温度升至 1000K，再逐渐降至 250K。模拟由随机选取的结构开始，重复这样的退火淬火过程 20 次。图 12-16 显示 BPTI（16-36）蛋白质片段形成 $\alpha$ 螺旋与 $\beta$ 折叠的倾向性。此图的（a）为真空状态的模拟结果，(b) 则为在溶液状态的模拟结果。图中的倾向性是由模拟的结构中，氨基酸间的二面角的范围所定出的。Okamoto 等确定 $\phi = -60° \pm 45°$，$\psi =$

$-50°\pm45°$为形成$\alpha$螺旋的残基（residue），$\phi=-105°\pm50°$，$\psi=135°\pm50°$为形成$\beta$折叠的残基。

图 12-17 为模拟的结构与 X 射线衍射实验的比较。图中显示在水溶液中模拟的最优化结构形成$\beta$折叠状，与观测的结构一致。但真空系统的模拟则仅形成弯曲转折（turn），但不形成$\beta$折叠的结构。

图 12-18 显示由随意选取的结构开始，逐步形成 S 结构的过程。图中的虚线表示形成氢键的残基对。综合以上的结果，Okamoto 等认为 BPTI 蛋白质形成折叠的$\beta$折叠状结构的原因主要为受溶液中水分子的影响。这些水分子迫使 BPTI 蛋白质分子的带电侧链互相吸引，形成氢键，而导致分子的折叠。

## 参 考 文 献


1 Hemenway E C, Wade A P. J. Chem. Edu., 69, 123, 1992

2 Panagiotopoulos A Z. Mol. Phys., 61, 813, 1987

3 Adams D J. Mol. Phys., 32, 647, 1976

4 Adams D J. Mol. Phys., 37, 211, 1979

5 Nath S K, Escobedo F A, de Pablo J J, Patramal I. Ind. Eng. Chem. Res., 37, 3195, 1998

6 Nath S K, Escobedo F A, de Poblo J J. J. Chem. Phys, 102, 905, 1998

7 Stroud H J F, Richards E, Limcharoen P, Parsonage N G. J. Chem. Soc. Faraday Trans. I, 72, 942, 1976

8 Catlow C R A. "Modeling of Structure and Reactivity in Zeolites", Academic Press, London, England, 1992

9 Goodbody S J, Watanabe K, MacGowan D, Walton J P R B, Quirke N. J. Chem. Soc., Faraday Trans., 87, 1951, 1991

10 Rosenbluth M N, Rosenbluth A W. J. Chem. Phys., 23, 356, 1955

11 Smit B, Maesen T L M. Nature, 374, 42, 1995

12 Smit B, Siepmann J I. Science, 264, 1118, 1994

13 Okamoto Y, Masuya M, Nabeshima M, Nakazawa T. Chem. Phys. Lett., 299, 17, 1999

# 第13章 布朗动力计算方法

以分子动力计算方法研究较慢运动体系，需要很长的计算时间，庞大的计算量远远超出一般计算机的负担。布朗动力计算（Brownian dynamics，BD）中引入了摩擦力与随机力的概念，朗之万方程式为描述布朗粒子的运动方程式，解此方程式可得到布朗粒子的位置与速度。BD最大的优势在于节省大量的计算时间，可以研究较长时间范围内的运动行为，通常用于布朗动力计算研究高分子与生化分子的运动。本章内容分两部分，第一部分从13.1～13.4节，介绍了方法原理和计算公式，13.5～13.7节是应用例子，13.8节还提供了朗之万方程的推导过程。

## 13.1 布朗粒子运动力学

1827年生物学家Robert Brown发现若观察一溶液中粒子的运动，对观察者而言，粒子运动的轨迹像是杂乱无章的乱动。事实上这所谓的乱动是因为粒子与其周围溶液中的分子产生不时的碰撞所造成的。这样乱动的粒子被称为布朗粒子（Brownian particle），如图13-1所示。

分析影响粒子运动的力可分为粒子与其周围液体的摩擦力（frictional force）与溶液施与粒子的随机力（stochastic force）两部分。设粒子运动的速度为$\vec{v}$，则由斯托克斯定律（Stokes' law），粒子所受的摩擦力，$\vec{f}_f$，为：

$$\vec{f}_f = -\gamma'\vec{v}$$
$$\gamma' = 6\pi a\eta \tag{13-1}$$

式中，$\gamma'$为溶液的摩擦系数；$a$为粒子的半径；$\eta$为溶液介质的黏度（viscosity）。设粒子所受的力为来自溶剂分子的随机力，$\vec{A}'(t)$，这种力与

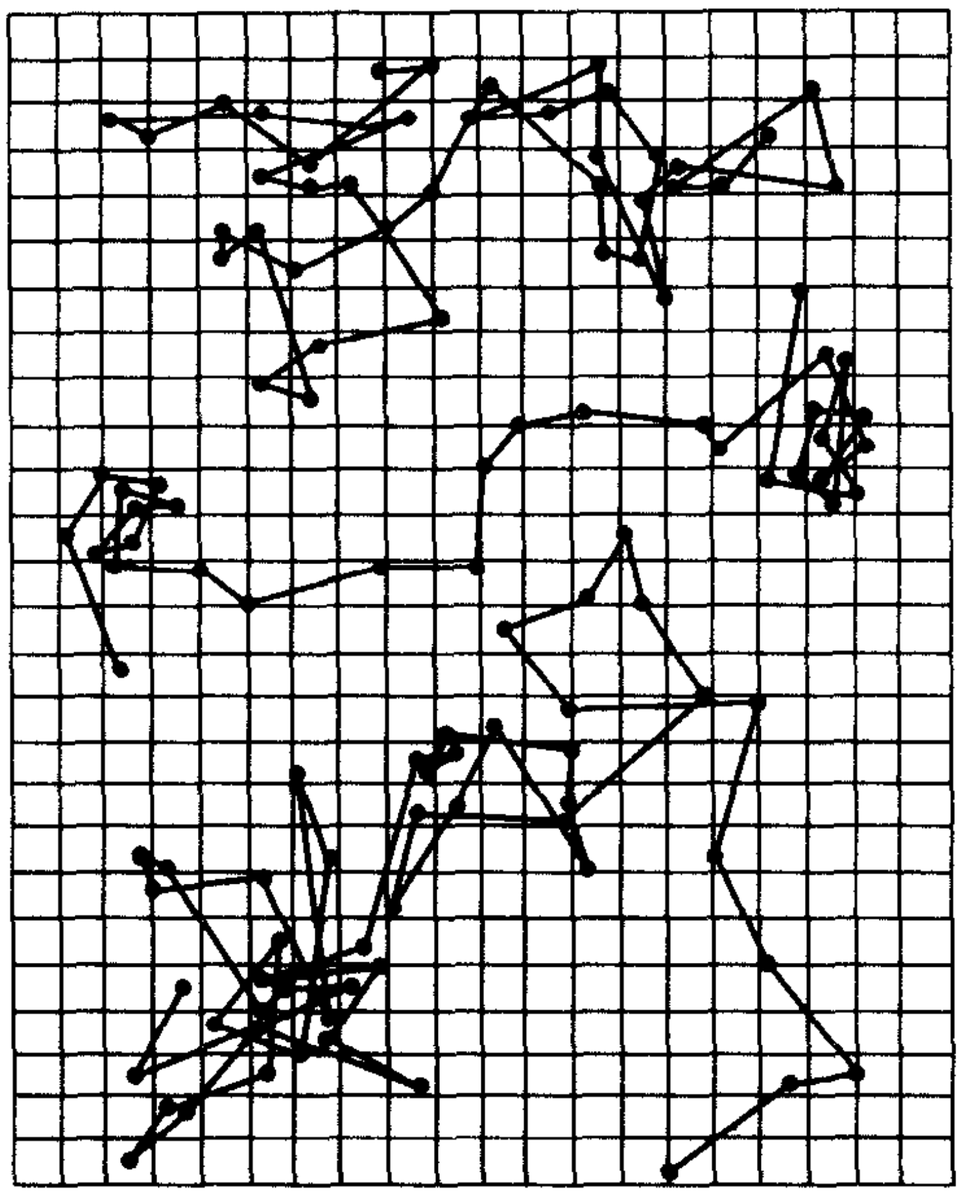


图 13-1　布朗运动粒子的轨迹

粒子的速度无关。并且，瞬间的随机力彼此间无任何的相关性，即 $\vec{A}'(t)$ 与 $\vec{A}'(t+\Delta t)$ 无关。则布朗粒子的运动可以以朗之万（Langevin）方程式描述，即：

$$m\frac{\mathrm{d}\vec{v}}{\mathrm{d}t}=-\gamma'\vec{v}+\vec{A}'(t) \tag{13-2}$$

或将此方程式表示为更方便的形式为：

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t}=-\zeta\vec{v}+\vec{A}(t) \tag{13-3}$$

式中，$m$ 为粒子的质量；$\vec{A}(t)=\vec{A}'(t)/m$；$\zeta=\gamma'/m$。凡如此类的运动方程式，其中含有随机力项的，皆称为随机力微分方程式（stochastic differential equation）。

将式(13-3) 积分，可得朗之万方程式的一般解为：

$$\vec{v}(t)=\vec{v}(0)\mathrm{e}^{-\zeta t}+\mathrm{e}^{-\zeta t}\int_0^t \mathrm{e}^{-\zeta t'}\vec{A}(t')\mathrm{d}t' \tag{13-4}$$

利用随机力平均值为零的关系，$\langle\vec{A}(t)\rangle=0$，可推得：

$$\langle\vec{v}\rangle=\vec{v}(0)\mathrm{e}^{-\zeta t} \tag{13-5}$$

式中，“〈〉”为系统的平均值。设粒子的概率函数为 $W[\vec{v},t,\vec{v}(0)]$，当 $t\to\infty$时此概率函数值必趋近于符合麦克斯韦分配（Maxwell distribution）关系的值，即：

$$W[\vec{v},t,\vec{v}(0)]\to\left(\frac{m}{2\pi k_B T}\right)^{3/2}\exp\left(-\frac{mv^2}{2k_B T}\right)\quad(t\to\infty)\tag{13-6}$$

若以笛卡儿坐标描述粒子的运动，并结合式(13-6)，则可得到随机力间的相关关系式为：

$$\langle A_\alpha(t)A_\alpha(0)\rangle=2mk_B T\zeta\delta(t)\tag{13-7}$$

式中，$\alpha=x,y,z$，$k_B$ 为玻耳兹曼常数，$\delta$ 为 Kronecker delta 函数。同样的推导，亦可得到速度间的相关式为：

$$\langle v_\alpha(t)v_\alpha(0)\rangle=\langle v_\alpha^2\rangle e^{-\zeta t}\tag{13-8}$$

由以上的关系式可推导出当运动时间很长时，粒子的位置与其初始位置的关系为：

$$\langle|\vec{r}-\vec{r}_0|^2\rangle=\frac{6k_B T}{m\zeta}t=6Dt\tag{13-9}$$

式中，$D=k_B T/m\zeta$，为爱因斯坦所导出液体的扩散系数。此扩散系数正比于温度，反比于粒子的质量。

## 13.2 布朗动力计算

以上所介绍的朗之万方程式为描述布朗粒子的运动方程式，解此方程式可得到布朗粒子的位置与速度。这种计算方法称为布朗动力计算(Brownian dynamics，BD)，或称为随机动力计算（stochastic dynamics，SD)。若欲执行布朗动力计算则必须先完全定义式(13-3）中的随机力项 $\vec{A}(t)$。在BD 方法中，一般均视此随机力项为一高斯随机运动过程(Gaussian random process)。设 $f(t)$ 为符合高斯随机运动过程的任一函数，则变量：

$$\delta A_\alpha^G=\int_t^{t+\delta t}f(t')A_\alpha(t')\mathrm{d}t'\tag{13-10}$$

亦为符合高斯分配的随机数，即符合如下的概率分配式，即：

$$\rho(\delta A_\alpha^G)=\frac{1}{\sigma_A(2\pi)^{1/2}}\exp\left[-\frac{1}{2}(\delta A_\alpha^G/\sigma_A)^2\right]\tag{13-11}$$

式中，$\rho$ 为概率。符合此种概率分配的随机力其平均值为 0，且标准误差（variance）为：

$$\sigma_A^2 = 2\zeta m k_B T \int_t^{t+\delta t} f^2(t')\mathrm{d}t' \tag{13-12}$$

以上这些结果有助于建立朗之万方程式的数理积分方法。下面将讨论如何比照一般的分子动力计算方法，将积分化为 $\delta t$ 连续步骤的 BD 计算。

在实际的布朗动力计算中，对系统中全不相关的粒子（如非常稀薄溶液中的分子），可直接应用朗之万方程式。有时粒子与粒子间的做用不能忽略（如浓度大的液体），或是分子内很重要的作用（如生化分子自身的分子内作用）则应用布朗动力计算时，除了粒子所受的摩擦力与随机力外亦应将这部分的作用力一并考虑。此时将式(13-3) 改写为：

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\zeta\vec{v}(t) + \frac{\vec{f}(t)}{m} + \vec{A}(t) \tag{13-13}$$

式中，$\vec{f}(t)$ 为由粒子的作用势能所推导出的力，即来自势能函数的梯度。实际上，不同的分子其摩擦系数不同，且分子的结构不同其摩擦系数亦不相同。但在 BD 模拟中通常将摩擦效应简化视为各向同性（isotropic）效应。因此 $\zeta$ 视为一纯量而非向量。此外，作用于不同分子的随机力，$\vec{A}(t)$，彼此间亦无任何关系，即符合：

$$\langle \vec{A}_i(t)\vec{A}_j(t)\rangle = 0 \qquad (i \neq j) \tag{13-14}$$

式中，$i$，$j$ 为不同的分子。

一般解朗之万方程式有三种情况，视积分步程的相对大小与速度松弛的时间而定。速冻松弛时间为摩擦系数的倒数，$1/\gamma$。第一种情况的积分步程时间小于速度松弛时间，$\gamma\delta t \ll 1$。此状况下，摩擦系数很小，溶剂分子对目标分子的影响很小。在极限条件 $\gamma \to 0$（$\zeta \to 0$）时，速溶剂分子对目标分子无任何作用时，朗之万方程式(13-13) 简化为牛顿运动方程式。另一种状况为速度松弛时间远小于积分步程时，摩擦系数很大，此相当于扩散区间（diffusive regime），目标分子的运动不断地受溶剂的影响而减缓。第三种状况则介于上述的两种状况之间。van Gunsteren 等所推导的第一种状况 $\gamma\delta t \ll 1$，式(13-13) 的解为：

$$\vec{r}(t+\delta t) = \vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}(\delta t)^2\{-\gamma\vec{v} + m^{-1}(\vec{F}+\vec{A})\}$$

$$\vec{v}(t+\delta t) = \vec{v}(t) + (\delta t)\{-\gamma\vec{v} + m^{-1}(\vec{F}+\vec{A})\} \tag{13-15}$$

式中随机力 $\vec{A}$ 符合高斯分布，$\langle A^2\rangle = 2mk_B T\gamma(\delta t)^{-1}$。Brunger 等提出另一种解朗之万运动方程式的方法，依据如下的有限差分（finite difference）近似法，即：

$$\frac{\mathrm{d}^2\vec{r}(t)}{\mathrm{d}t^2}\approx\frac{[\vec{r}(t+\delta t)-2\vec{r}(t)+\vec{r}(t-\delta t)]}{(\delta t)^2}$$

$$\frac{\mathrm{d}\vec{v}(t)}{\mathrm{d}t}\approx\frac{\vec{r}(t+\delta t)-\vec{r}(t-\delta t)}{2\delta t} \tag{13-16}$$

将式(13-15) 和式(13-16) 代入式(13-13) 可得位置 $\vec{r}(t+\delta t)$，即：

$$\vec{r}(t+\delta t)=\vec{r}(t)+[\vec{r}(t)-\vec{r}(t-\delta t)]\frac{1-\frac{1}{2}\gamma\delta t}{1+\frac{1}{2}\gamma\delta t}+\left(\frac{\delta t^2}{m}\right)\frac{\vec{F}+\vec{A}}{1+\frac{1}{2}\gamma\delta t} \tag{13-17}$$

第二种情况，$\gamma\delta t\gg 1$，则在积分步程的时间范围内，分子间的作用力可设为常数，van Gunsteren 等所推导的解为：

$$\vec{r}(t+\delta t)=\vec{r}(t)+\frac{\vec{F}}{m\gamma}\delta t+\vec{A}(\delta t) \tag{13-18}$$

式中，$\vec{A}$ 为符合高斯分布的随机力项，其平均值为零，平方误差为 $2k_{\mathrm{B}}T(m\gamma)^{-1}=2D\delta t$。此解法的进一步修正为将作用力设为与积分步程呈线性关系，得：

$$\vec{r}(t+\delta t)=\vec{r}(t)+[\vec{F}(t)+\frac{1}{2}\dot{\vec{F}}(t)\delta t]\frac{\delta t}{m\gamma}+\vec{A}(t) \tag{13-19}$$

式中，$\dot{\vec{F}}(t)$ 为时间 $t$ 时作用力的微分。由其线性关系，可推得关系式：

$$\dot{\vec{F}}(t)=\frac{\vec{F}(t)-\vec{F}(t-\delta t)}{\delta t} \tag{13-20}$$

第三种情况，$\gamma\delta t$ 介于上述两种极限条件之间，并无明显的限制。依 van Gunsteren 与 Berendsen 所推导的解法为：

$$\begin{aligned}\vec{r}(t+\delta t)=&\vec{r}(t)+\frac{\vec{v}(t)}{\gamma}[1-\exp(-\gamma\delta t)]\\&+\frac{\vec{F}(t)}{m\gamma}\left[\delta t-\frac{1-\exp(-\gamma\delta t)}{\gamma}\right]\\&+\frac{\int_t^{t+\delta t}\{1-\exp[-\gamma(t+\delta t-t')]\}\vec{R}(t')\mathrm{d}t'}{mg}\end{aligned}$$

$$\begin{aligned}\vec{v}(t+\delta t)=&\vec{v}(t)\exp(-\gamma\delta t)+\frac{\vec{F}(t)}{m\gamma}[1-\exp(-\gamma\delta t)]\\&+\frac{\int_t^{t+\delta t}\exp[-\gamma(t+\delta t-t')]\vec{R}(t')\mathrm{d}t'}{m}\end{aligned} \tag{13-21}$$

式中均含有对随机力 $\vec{R}$ 的积分，此二积分项彼此相关联且符合双变

量（bivariant）的高斯分布。这表示第二个变量的值与第一个变量值相关。这种限制增加了随机数取样的困难，van Gunsteren 与 Berendsen 等将此二方程式重组，推出可用两个互相独立的高斯分布随机数解朗之万方程式的方法。13.8 节列出 Ermark 所推导朗之万方程式的一般解。

## 13.3 布朗动力计算的技巧

执行布朗动力计算必须决定碰撞频率的摩擦系数 $\gamma$。对简单的分子而言，摩擦系数可由实验度量溶液的扩散系数得到。若执行刚性分子的布朗动力计算，则摩擦系数亦可由分子动力计算的结果获得。一般最通常的状况为需要分子中每一种原子的摩擦系数。简单的分子如丁烷，可假设每个原子的摩擦系数均相同。通常尝试不同系数的布朗动力计算以找出模拟结果与实验最吻合的系数值，或是利用分子动力计算的方法找出平均的摩擦系数。对大分子而言，原子的摩擦系数取决于其与溶剂分子接触的程度，依此通常将原子的摩擦系数假设正比于原子的可用表面积（accessible surface area）。

执行布朗动力计算最大的优势在于节省大量的计算时间，可以研究较长时间范围内的运动行为。例如 Widmalm 与 Pastor 利用布朗动力计算研究 ethylene glycol 与 259 个水分子溶液于 1ns 范围内的运动。若用分子动力模拟的方法研究相同的系统约需 300h 的计算时间。但执行此布朗动力计算仅需 24min。节省的时间是因为 BD 模拟仅需计算溶质分子 ethylene glycol 的运动，并且可引用较长的积分步程。通常利用布朗动力计算研究高分子与生化分子的运动，因为这些分子许多重要的运动均属较长的时间范围。但执行布朗动力计算必须注意计算系统中是否存在特殊的溶剂与溶质分子的作用力。以下用一实际例子说明此项影响因素的重要性。

Yun-yu S. 与 van Gunsteren 利用分子动力计算与布朗动力计算研究药物分子 cyclosporin 于四氯化碳溶剂与水分子溶液的运动行为。图 13-2 为 cyclosporin 分子的结构图。由两种不同的计算方法分别得到分子的平均结构、扭转角转动轨迹等资料。计算的结果显示在四氯化碳溶剂中，由分子动力计算与布朗动力计算所得的结构极为相近。但在水溶液中，两种计算所得的结构却有很大的差异。这是因为布朗动力计算太过强调分子内部的氢键作用，而分子动力计算却显示分子与溶剂形成了大量的氢键。

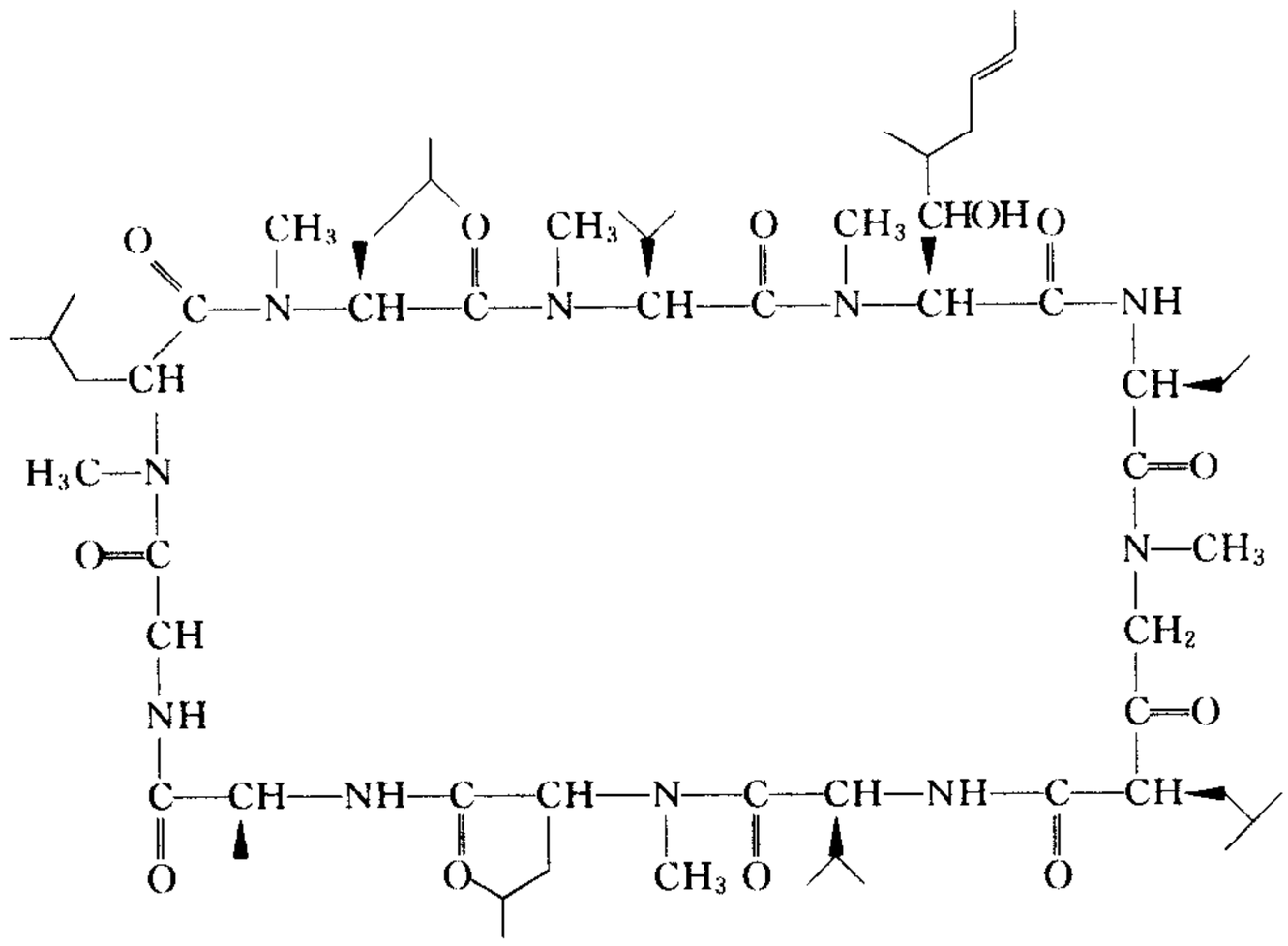


图 13-2　cyclosporin 分子的结构

## 13.4　流动效应

上节所讨论的布朗动力计算式，在溶液中其溶质分子所受的力 $\vec{f}(t)$（包括分子内作用力与溶质分子间作用力）是由势能函数推导而来。但未考虑不同的溶质分子其附近溶剂分子的流动不同所造成的影响。这称为溶剂分子的流动效应（hydrodynamic effect）。若将此效应考虑在内，则式(13-3) 可改写为：

$$\dot{\vec{r}}(t)=\frac{\widetilde{D}(t)}{k_{\mathrm{B}}T}\cdot\vec{f}(t)+\nabla\vec{r}\cdot\widetilde{D}(t)+\dot{\vec{A}}^{G}(t) \tag{13-22}$$

或者，对第 $i$ 个分子而言，其速度式为：

$$\dot{\vec{r}}_{\mathrm{i}}(t)=\sum_{j}\frac{\widetilde{D}_{\mathrm{ij}}(t)}{k_{\mathrm{B}}T}\cdot\vec{f}_{\mathrm{j}}(t)+\nabla\vec{r}_{\mathrm{j}}\cdot\widetilde{D}_{\mathrm{ij}}(t)+\dot{\vec{A}}_{\mathrm{i}}^{G}(t) \tag{13-23}$$

此处，扩散系数项 $\widetilde{D}$ 为一个 $3N\times3N$ 的张量或矩阵，$\widetilde{D}_{\mathrm{ij}}$ 为一个 $3\times3$ 的张量矩阵，由分子的位置决定。由此，位置的积分式为：

$$\vec{r}_{\mathrm{i}}(t+\delta t)=\vec{r}_{\mathrm{i}}(t)+\sum_{j}\frac{\widetilde{D}_{\mathrm{ij}}(t)}{k_{\mathrm{B}}T}\cdot\vec{f}_{\mathrm{j}}(t)\delta t+\nabla\vec{r}_{\mathrm{j}}\cdot\widetilde{D}_{\mathrm{ij}}(t)\delta t+\delta\vec{A}_{\mathrm{i}}^{G} \tag{13-24}$$

式中，$\delta\vec{A}_i^G$ 随机力其平均值为零，具高斯分布。且符合式(13-25) 的关系：

$$\langle \delta\vec{A}_i^G \delta\vec{A}_j^G \rangle = 2\,\widetilde{D}_{ij}\,\delta t \tag{13-25}$$

注意，此处不同分子的 $\delta\vec{A}_i^G$ 不再完全无关。至今仅知扩散张量矩阵取决于分子位置，但彼此间的关系到底为何却无定论。一般在计算中采用奥森氏（Oseen）由流体力学所导出的张量关系，即：

$$\begin{aligned}\widetilde{D}_{ij} &= \frac{k_B T}{6\pi\eta r_0}\cdot\widetilde{1} \quad (i=j)\\ &= \frac{k_B T}{8\pi\eta r_0}\left(\widetilde{1}+\frac{\vec{r}_{ij}\vec{r}_{ij}}{r_{ij}^2}\right) \quad (i\neq j)\end{aligned} \tag{13-26}$$

式中，$\eta$ 为液体的黏度（viscosity）；$r_0$ 为估计的分子有效半径；$\widetilde{1}$为单位矩阵。因为此张量具有$\nabla\vec{r}\cdot\widetilde{D}=0$ 的性质，故式(13-24) 可简化为：

$$\vec{r}_i(t+\delta t) = \vec{r}_i(t) + \sum_j \frac{\widetilde{D}_{ij}(t)}{k_B T}\cdot\vec{f}_j(t)\delta t + \delta\vec{A}_i^G \tag{13-27}$$

关于应用流动效应的 BD 模拟，可参考 Ermark 与 McCammon 等的工作。

## 13.5 布朗动力计算在生化系统的应用

此节介绍应用布朗动力计算研究 DNA 中碱基对的张开现象。许多实验显示双螺旋核酸（double-strand nucleic acid）的结构并非固定不变。核酸中的各种原子或基团以不同形式及不同的速度运动。有些运动非常快速，运动的时间约为纳秒范围，可利用核磁共振（NMR）或荧光去极化(fluorescence depolarization) 技术侦测。但有些运动非常缓慢，例如碱基对的张开运动，运动的时间约在微秒的范围。这种碱基对张开的运动虽可利用氢原子的交换实验侦测，以决定运动的时间范围，但实验无法侦测张开的路径与所需克服的能量障碍。

Briki 等以布朗动力计算的方法研究 DNA 片段胸腺嘧啶双螺旋中 $(dA)_5$-$(dT)_6$ 碱基对张开的运动。张开的运动路径以碱基与垂直于碱基对平面轴的夹角 $\theta$ 描述。对应于此的朗之万方程式为：

$$I\frac{d^2\theta}{dt^2} = -\zeta\frac{d\theta}{dt} + \Gamma_{eq} + \Gamma_{rand} \tag{13-28}$$

式中，$I$ 为转动碱基的转动惯量。Briki 等将此惯量设定为 $I=2.3\times10^{-44}\,\mathrm{kg\cdot m^2}$；$\zeta$ 为对应的黏度；$\Gamma_{eq}$ 为系统平衡的恢复力矩；$\Gamma_{rand}$ 则为随机力矩。依据 Ermark 与 McCammom 的方法，可解式(13-28)，得任一瞬间转动的角度为：

$$\Delta\theta=\left(\frac{D}{k_B T}\right)\Gamma_{eq}(\theta)\delta t+R(\delta t) \tag{13-29}$$

式中，$D$ 为转动扩散系数；$R(\delta t)$ 为符合高斯分布的随机数，即：

$$\langle R\rangle=0,\langle R^2\rangle=2D\delta t \tag{13-30}$$

欲引用式(13-29)，计算所选取的积分步程 $\delta t$ 必须远大于 $DI/k_B T$。平衡力矩与张开的势能 $U(\theta)$ 的关系为：

$$\Gamma_{eq}(\theta)=-\frac{\partial U(\theta)}{\partial\theta} \tag{13-31}$$

Ramstein 与 Lavery 将势能函数定为 $U(\theta)=C\theta$，$C=0.4\times10^{-15}\,\mathrm{J/(°)}$；$0°<\theta<15°$。因碱基需转动超过 15°方能断开联结的氢键，因此将之定为碱基对张开的下限。转动扩散系数可由黏度的关系式定出，即：

$$D=k_B T/\zeta \tag{13-32}$$

于 300K 所度量的黏度为 $\zeta=2\times10^{-29}\,\mathrm{m^2\cdot kg/s}$，对应的扩散系数为 $D=2\times10^{8}\,\mathrm{s^{-1}}$。其余温度下的扩散系数，可假设 $D\propto T/\eta$；$\eta$ 为该温度下水的黏度。计算选取的积分步程为 $\delta t=2\times10^{-12}\,\mathrm{s}$，远大于一般分子动力计算所选的积分步程（约 $10^{-15}\,\mathrm{s}$）。此研究总共执行 $6\times10^{9}$ 步的布朗动力计算。

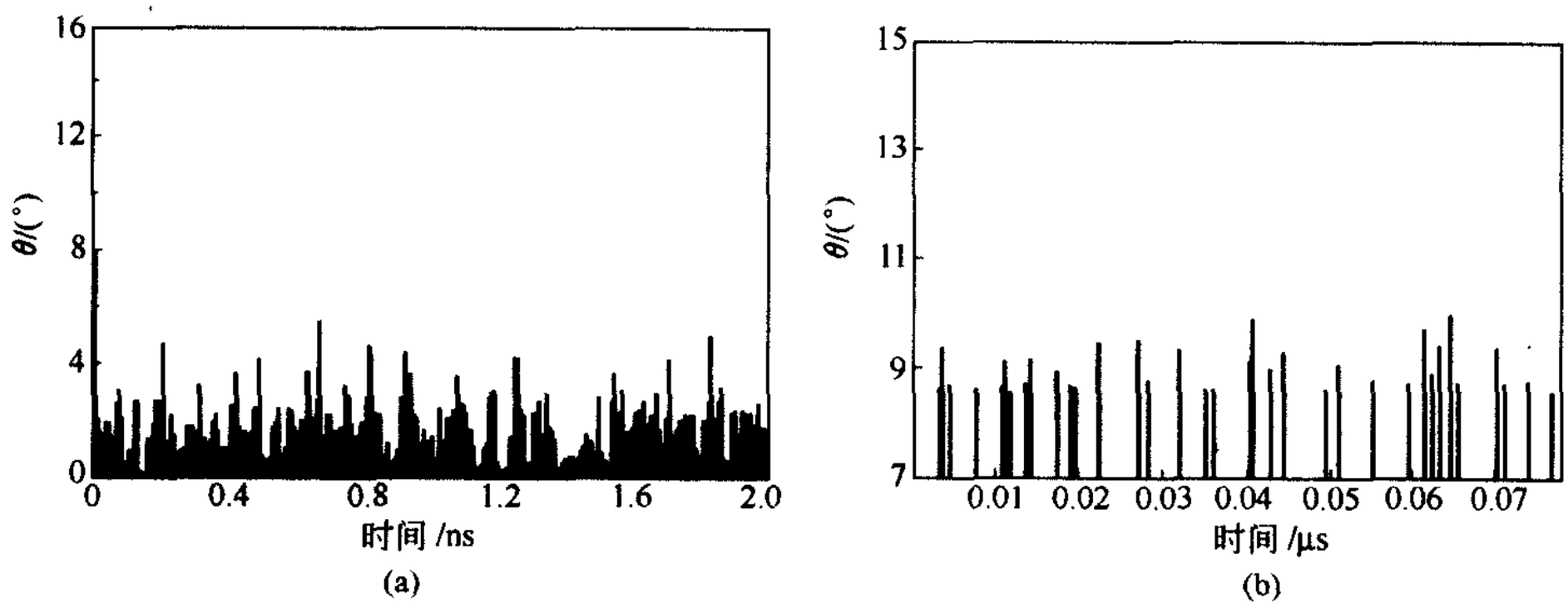



**图 13-3　2ns 与 0.08μs 执行布朗动力计算的角度 θ 的变化**

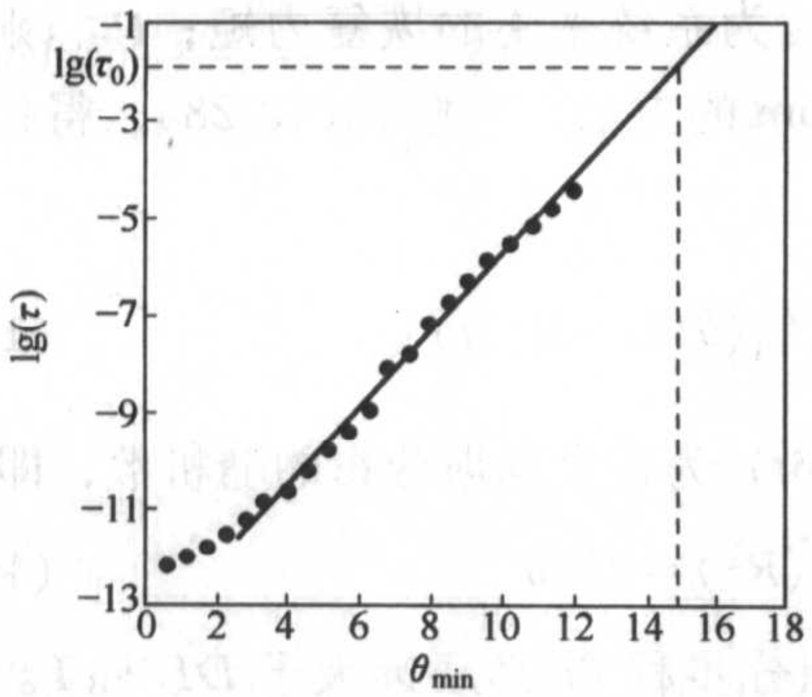



**图 13-4　lg(τ)对 $\theta_{min}$ 的关系**

图 13-3 中的（a）和（b）分别为 2ns 与 0.08μs 执行布朗动力计算的角度 $\theta$ 的变化图，图 13-3(b) 只显示超过 8°的角度变化。图中显示当计算的时间增长为微秒范围时，角度变化的幅度增加；2ns 范围最大的变化幅度为 5°，而 0.08μs 范围内最大的变化幅度超过 10°。

由此结果可计算转动角超过某一特殊值 $\theta_{min}$ 的平均间隔时间 $\tau$。图 13-4 为 300K 温度下 lg($\tau$) 对 $\theta_{min}$ 的关系图。

图 13-4 显示，当 $\theta_{min}>3°$，lg($\tau$) 对 $\theta_{min}$ 呈良好的线性关系，由此可由外插法推断当 $\theta_{min}=15°$ 时，$\tau_0=15ms$。意即在 300K，至少需执行 15μs 的布朗动力计算方可直接由计算的结果显示碱基对的完全张开。表 13-1 为各温度条件计算的 $\tau_0$ 值，图 13-5 为依照此表所得的阿仑尼乌斯(Arrhenius) 图。

**表 13-1　各温度条件下计算的 $\tau_0$ 值**

| T/K | $\tau_0$/ms | T/K | $\tau_0$/ms |
|---|---|---|---|
| 318 | 2.2 | 288 | 180 |
| 300 | 15 | 273 | 1000 |

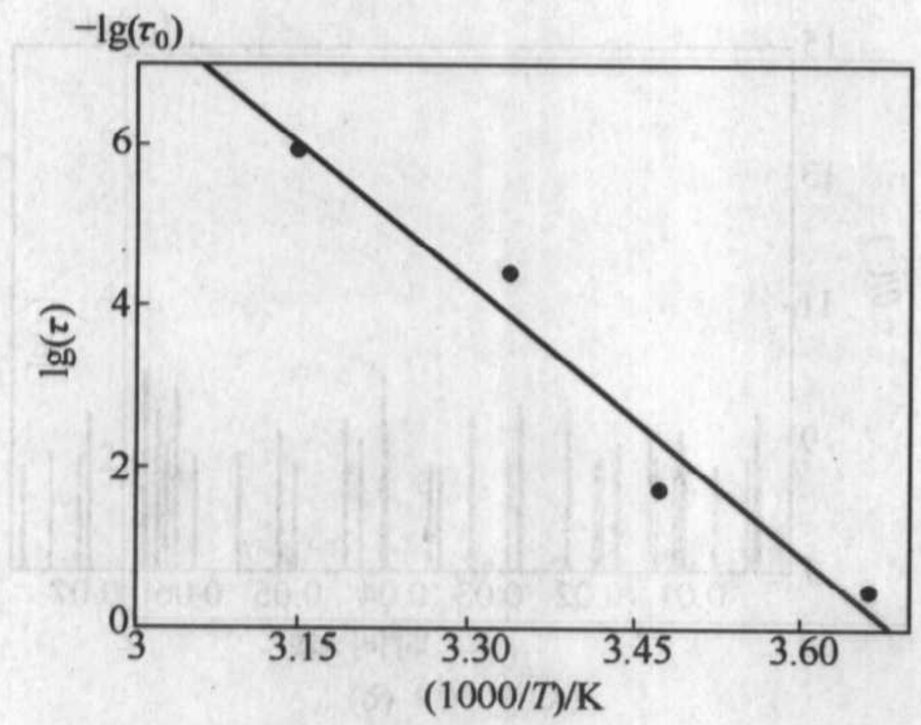



**图 13-5　lg(1/$\tau_0$)对温度导数的阿仑尼乌斯图**

由图 13-5 中的斜率可推断碱基张开的活化能障约为 20kcal/mol。

与实验相较，由氢原子的交换反应所度量类似的 DNA 分子碱基对的生命期（life time），即张开的时间间隔均在微秒范围。例如以 NMR 于 293K 测量 poly（rA)-poly(rU）碱基对的生命期为 5μs。除此以外，由 NMR 实验所量测的一些类似的双螺旋 DNA 分子［如 poly(rA)-poly(rU)］碱基对张开的活化能均为 20kcal/mol，皆与此计算的结果相吻合。

## 13.6 布朗动力计算在高分子溶液的应用

此节中介绍利用布朗动力计算对高分子溶液所进行的研究。设所研究的高分子溶液为高黏度状态。此系统的朗之万运动方程式可简化为：

$$\frac{\mathrm{d}\vec{r}_i}{\mathrm{d}t}=\frac{1}{\zeta_i}\nabla_i U+\vec{A}_i(t) \tag{13-33}$$

式中，下角 $i$ 为高分子链的第 $i$ 个成分粒子；$\zeta_i$ 为其摩擦系数；$U$ 为高分子链的势能；$\vec{A}_i(t)$ 为布朗随机力项。

$\vec{A}_i(t)$ 符合式(13-34) 的关系：

$$\langle\vec{A}_i(t)\vec{A}_j(t')\rangle=\frac{2k_B T}{\zeta_i}\delta_{ij}\delta(t-t') \tag{13-34}$$

式中，$k_B$ 为玻耳兹曼常数；$T$ 为热力学温度。

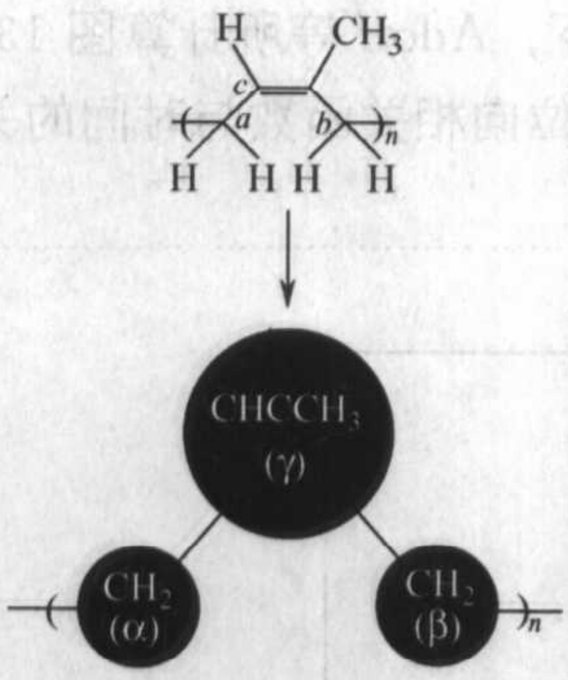



**图 13-6　聚异丙烯的单体结构与实际执行布朗动力计算所引用的简化构造**

Adolf 与 Ediger 等利用此种简单方法研究聚异丙烯（polyisoprene）在溶剂中的区域运动动力行为。为了更进一步简化计算，Adolf 等将聚异丙烯视为由联合的基团所组合成的分子。图 13-6 为聚异丙烯的单体结构与实际执行布朗动力计算所引用的简化构造。

图 13-6 中显示此计算将 $CHCCH_3$ 基团视为一混成之粒子。聚异丙烯共有两类粒子，即 $CHCCH_3$ 与 $CH_2$。每类粒子的摩擦系数可由其体积及 Stokes 定律(Stokes' law) 计算出。计算所采用的势能为简化的联合原子模型的势能，即：

$$U=U_b+U_\theta+U_\phi$$

$$U_b=\frac{1}{2}K_b(l-l_0)^2$$

$$U_\theta=\frac{1}{2}K_\theta(\cos\theta-\cos\theta_0)^2$$

$$U_\phi=\frac{1}{2}K_\phi\sum_{n=0}^{5}a_n\cos(n\phi) \tag{13-35}$$

式中，$K$ 及 $a$ 为力常数；$l$ 为键长；$\theta$ 为键角；$\phi$ 为连续四原子所形成

的扭转角。下标“0”表示平衡值。计算采取的积分步程 $\delta t$ 为 5fs，每种不同条件的 BD 计算约需收集 10ns 的轨迹。

实验上多以核磁共振 NMR 的方法研究溶液的动力行为。异丙烯溶液计算的结果与以$^{13}C$ NMR 所度量的松弛时间 $T_1$ 互相比较。$T_1$ 与链上 CH 键向量运动的位向相关函数（orientation auto-correlation function）$P_2$ 有关。图 13-6 中三种不同碳原子 $a$，$b$，$c$ 的相关时间（correlation time）均可由 NMR 度量。在聚异丙烯中，所观察亚甲基的相关时间为两个 CH 键相关时间的平均值。特殊的 CH 键所对应的位向相关函数为：

$$C(t)=\langle P_2[\vec{x}(0)\cdot\vec{x}(t)]\rangle=\frac{1}{2}\langle 3[\vec{x}(0)\cdot\vec{x}(t)]^2-1\rangle \quad (13\text{-}36)$$

式中，$P_2$ 为二次勒让德（Legendre）多项式；$\vec{x}(t)$ 为时间 $t$ 时 CH 键的单位向量。由 NMR 所测的相关时间（correlation time）与位向相关函数的关系为：

$$\tau_c=\int_0^\infty C(t)\mathrm{d}t \quad (13\text{-}37)$$

图 13-7 显示在温度为 425K 的条件下，Adolf 等所计算图 13-6 中聚异丙烯的 $c$ 所对应的 H—$CCCH_3$ 键所得的位向相关函数与时间的关系。

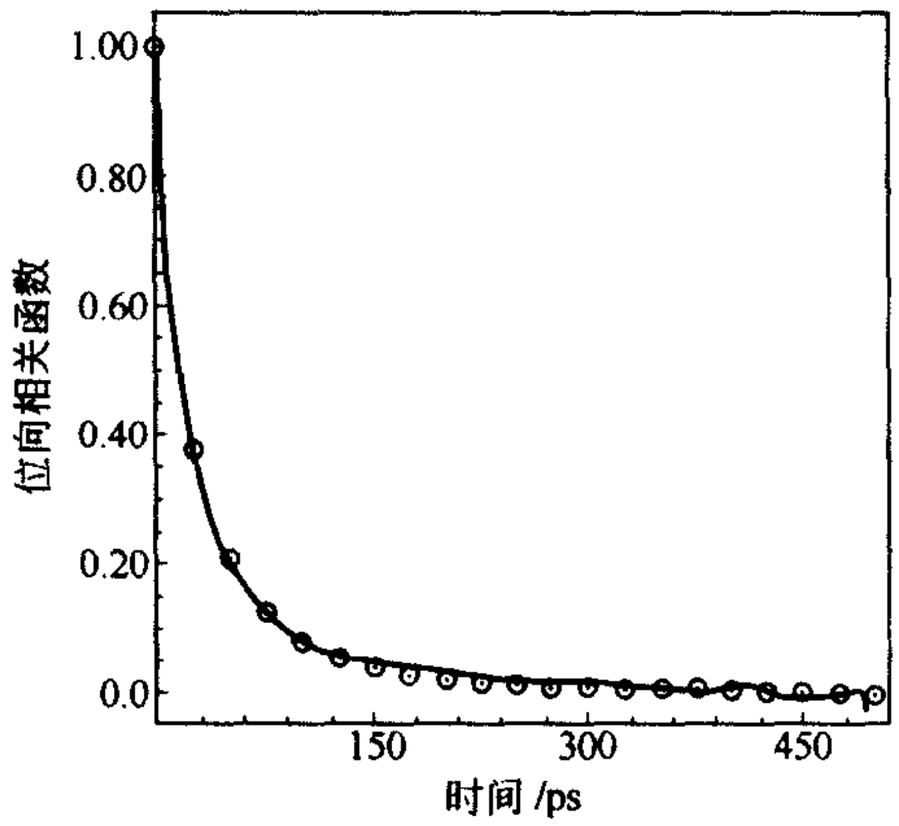



**图 13-7　H—$CCCH_3$ 键的位向相关函数**

NMR 实验可侦测出在聚异丙烯中不同种类的 C—H 键具有不同的相关时间。图 13-8 为 Adolf 等利用布朗动力计算所得的聚异丙烯的三种不同 CH 键的相关函数。每种相关函数皆对应于一种不同形态的运动。

图 13-8 中的——为 $c$ 原子，……为 $b$ 原子，—·—为 $a$ 原子的 CH 向量所对应的相关函数。计算所得的 $\tau_b : \tau_a$ 比为 $1.17 \pm 0.08$，$\tau_c : \tau_a$ 比为 $1.45 \pm 0.11$，实验所得对应的比值分别为 $1.19 \pm 0.12$ 与 $1.54 \pm 0.14$。Adolf 等所计算的相关函数均与实验值非常接近，为成功应用布朗动力计算探讨高分子内部运动的例子。

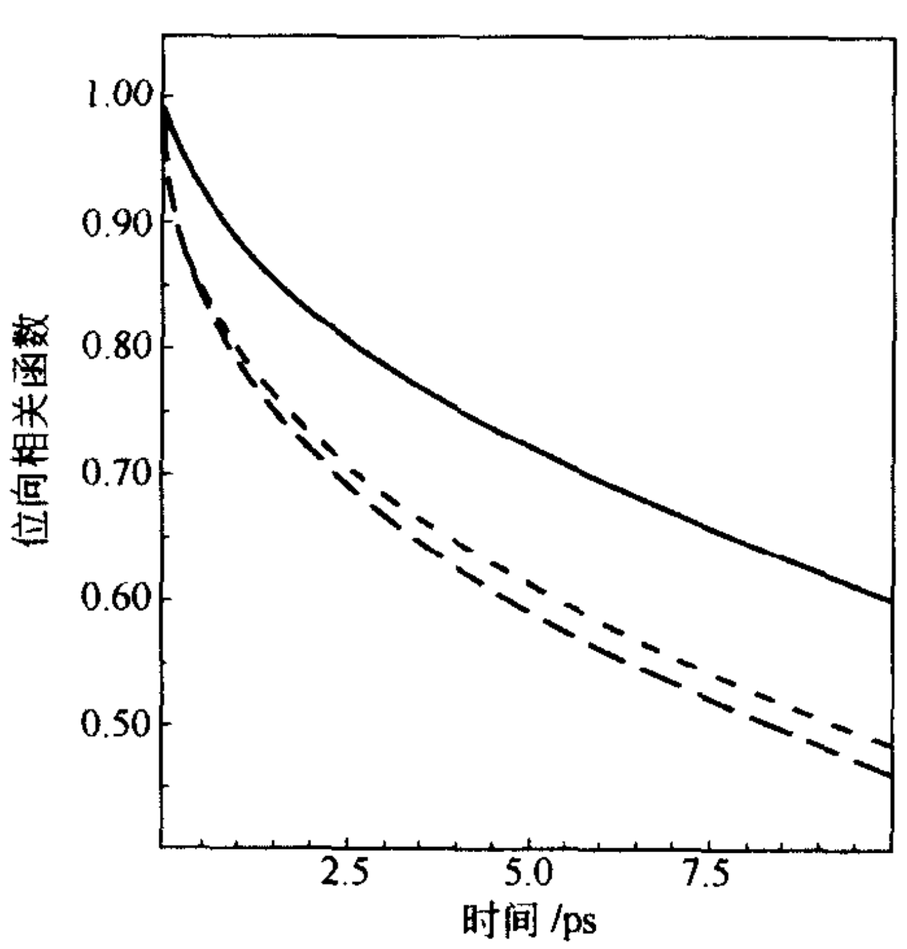



**图 13-8　聚异丙烯的三种不同 CH 键的位向相关函数**

# 13.7　分子动力计算与布朗动力计算的比较

Widmalm 与 Pastor 等比较分子动力模拟与布朗动力模拟的差异。本研究以乙二醇（EG），HO—$CH_2$—$CH_2$—OH 的水溶液系统为检验的对象。计算系统为一个 EG 分子与 259 个水分子。分子动力计算与布朗动力计算所选取的积分步程均为 $\delta t = 1.0$fs。布朗动力计算中，假设碳原子与氧原子的摩擦系数相同；计算中试验 $\gamma = 0.1 \sim 50$ 的范围；氢原子摩擦系数则设为 0。分子动力计算采用 CHARMM 力场与 TIP3P 水分子模型；由 5 种不同的初始速度分别开始计算。布朗动力计算则依 Brunger 所提出的方法执行。

EG 分子对其 O—C—C—O 扭转角形成反式（trans）与间式（gauch）的构象；由反式转换为间式的速率，$k_0$ 可由构象数目的相关函数求得。设

$\langle N_t \rangle$ 为反式构象的平均概率，$C_N(t)$ 为归一化（normalized）的数目相关函数，$\tau_N$ 为松弛时间，则其相互间的关系为：

$$\tau_N = \int_0^\infty C_N(t)\mathrm{d}t = \int_0^\infty \frac{\langle \delta N_t(t)\delta N_t(0)\rangle}{\langle \delta N_t^2 \rangle}\mathrm{d}t \tag{13-38}$$

式中，$\delta N_t(t) = N_t(t) - \langle N_t \rangle$。若为反式构形 $N_t = 1$；反之，$N_t = 0$。$N_t$ 的统计涨落为：

$$\langle \delta N_t^2 \rangle = \langle N_t^2 \rangle - \langle N_t \rangle^2 \tag{13-39}$$

利用此关系，可得由反式转换为间式的速率常数式为：

$$k_{tg} = \frac{1 - \langle N_t \rangle}{2\tau_N} \tag{13-40}$$

表 13-2 比较了布朗动力计算与分子动力计算的结果。

**表 13-2　EG 分子水溶液的布朗动力计算与分子动力计算的比较**

| 项　目 | 计算时间范围/ns | $\langle N_t \rangle$ | $k_{tg}$ /ns | 项　目 | 计算时间范围/ns | $\langle N_t \rangle$ | $k_{tg}$ /ns |
|---|---|---|---|---|---|---|---|
| BD 计算(γ) | | | | MD 计算 | | | |
| 0.1 | 50 | 0.211 | 8.1 | 1 | 1 | 0.182 | 123.9 |
| 1.0 | 50 | 0.211 | 28.8 | 2 | 1 | 0.244 | 82.2 |
| 2.5 | 10 | 0.227 | 39.0 | 3 | 1 | 0.193 | 56.0 |
| 7 | 10 | 0.198 | 62.7 | 4 | 1 | 0.176 | 108.4 |
| 10 | 10 | 0.211 | 65.8 | 5 | 1 | 0.186 | 61.3 |
| 11 | 10 | 0.207 | 59.2 | 平均 | | 0.196 | 86.4 |
| 12 | 100 | 0.210 | 66.9 | 标准误差 | | 0.027 | 29.2 |
| 13 | 10 | 0.201 | 70.1 | | | | |
| 25 | 100 | 0.213 | 55.3 | | | | |
| 50 | 20 | 0.222 | 34.7 | | | | |

由表中数据可知，不论摩擦系数多少，由布朗动力计算的 $\langle N_t \rangle = 0.212$ 值差异不大；由分子动力计算的 $\langle N_t \rangle = 0.196 \pm 0.027$ 亦在此范围内。因此可知水分子对于反式与间式的平衡值影响并不大。图 13-9 比较布朗动力计算与分子动力计算所得的数目相关函数。图中的实线为分子动力计算的结果，虚线为 5ns 布朗动力计算平均的结果。此图显示由布朗动力计算的相关函数呈指数函数的衰减，符合一般相关函数衰减的特性。这表示利用此相关函数计算的转移速度是可信的。另外，由分子动力计算所得的相关函数却不呈指数衰减的形式，这表示由此计算而得的转换速度并不可信。

除了以上的计算外，亦可检验转动的相关函数以推测反式-间式的转换。设 $\vec{\mu}$ 表示 EG 分子的 C—C、$C—O_1$ 或 $C—O_2$ 键的向量，则对应于各

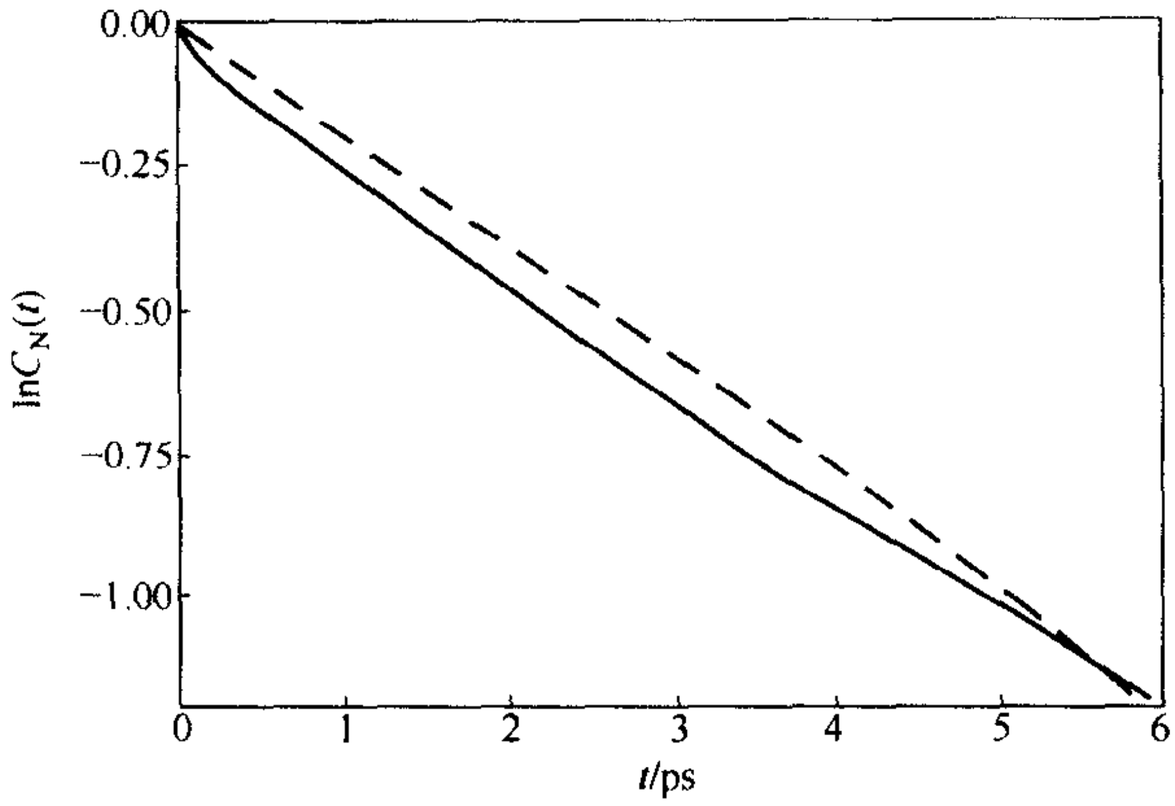



**图 13-9　布朗动力计算与分子动力计算所得的数目相关函数的比较**

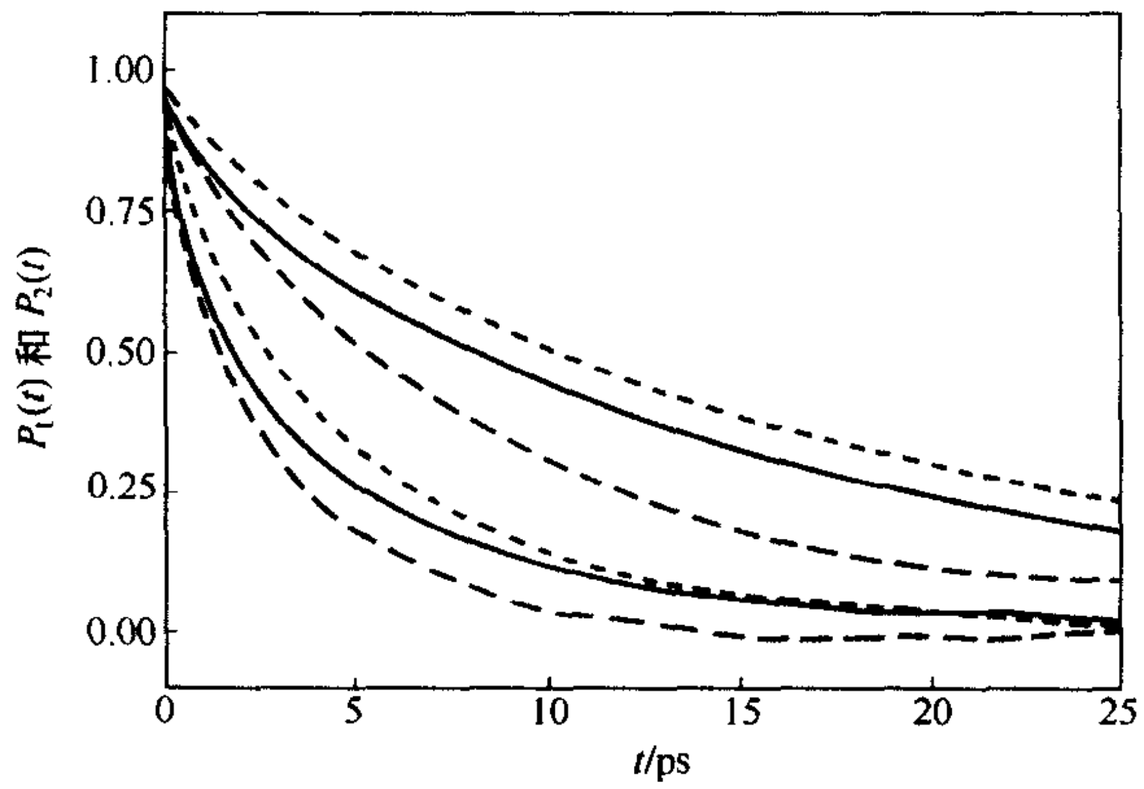



**图 13-10　分子动力计算所得 EG 分子三种键的转动相关函数**

键的转动相关函数为 $\langle P_1[\vec{\mu}(t)\cdot\vec{\mu}(0)]\rangle$ 与 $\langle P_2[\vec{\mu}(t)\cdot\vec{\mu}(0)]\rangle$，$P_1$ 与 $P_2$ 分别为一次及二次勒让德多项式（Legendre polynomial）。图 13-10 为由分子动力计算所得三种键的转动相关函数，图中实线对应 C—C 键，短虚线对应 C—$O_1$ 键，较长的虚线对应 $O_1$—$O_2$ 键；衰减较慢的为 $P_1(t)$，衰减较快的为 $P_2(t)$。

图 13-11 为利用布朗动力计算所得的转动相关函数图。图中各种线所对应的与上图相同。

依照式(13-41)，利用计算的转动相关函数可以求得转动的松弛时间。

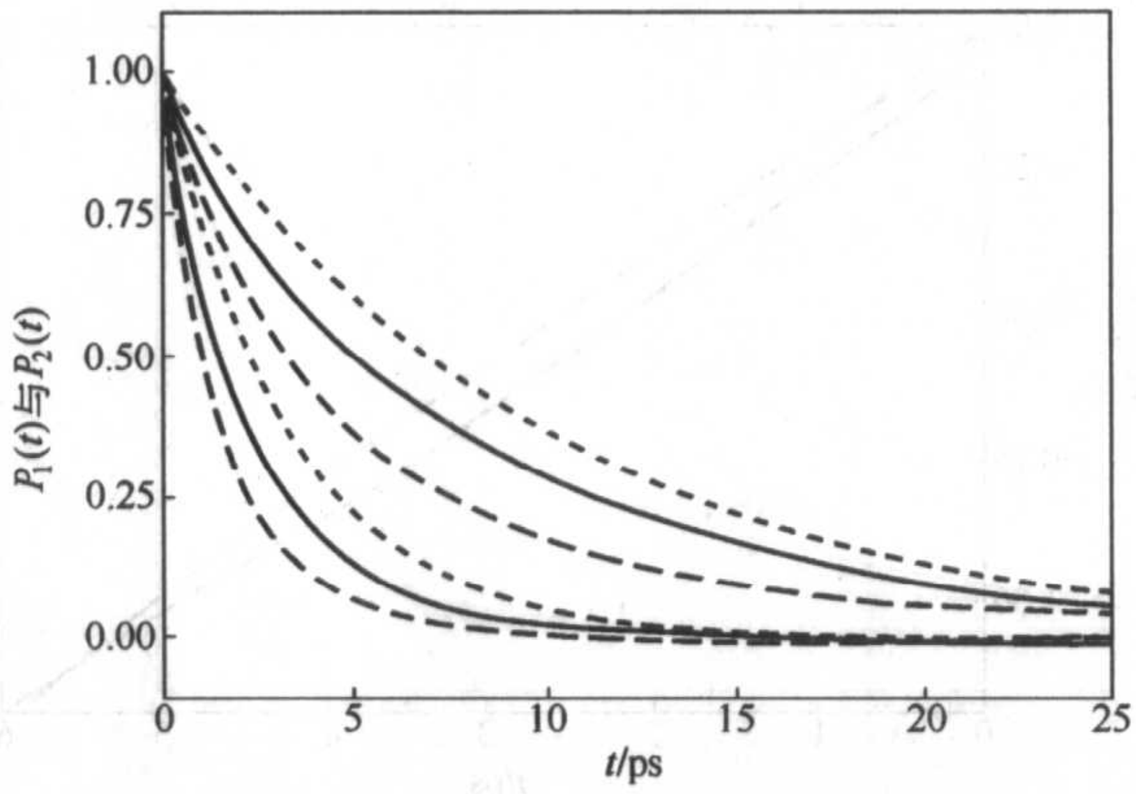



**图 13-11　利用布朗动力计算所得 EG 分子的转动相关函数**

$$P_1(t)=(1-A)\exp(-t/\tau_1)+A\exp(-t/\tau_2)$$

$$\tau=(1-A)\tau_1+A\tau_2 \tag{13-41}$$

表 13-3 比较根据分子动力计算及布朗动力计算分别得到的转动松弛时间。

**表 13-3　分子动力计算及布朗动力计算 EG 分子的转动松弛时间**

| 松弛时间 | $l=1$ | | | | $l=2$ | | | |
|---|---|---|---|---|---|---|---|---|
| | $\tau_1$ | $\tau_2$ | $A$ | $\tau$ | $\tau_1$ | $\tau_2$ | $A$ | $\tau$ |
| MD | | | | | | | | |
| C—C | 0.836 | 15.13 | 0.852 | 13.02 | 0.212 | 3.96 | 0.801 | 3.22 |
| C—O | 1.597 | 17.85 | 0.890 | 16.06 | 0.322 | 5.06 | 0.867 | 4.43 |
| O—O | 0.552 | 9.32 | 0.891 | 8.37 | 0.136 | 2.99 | 0.823 | 2.48 |
| BD | | | | | | | | |
| C—C | 1.769 | 9.21 | 0.857 | 8.15 | 0.856 | 2.75 | 0.816 | 2.40 |
| C—O | 2.446 | 10.45 | 0.980 | 10.28 | 0.0 | 3.34 | 1.00 | 3.34 |
| O—O | 2.053 | 6.63 | 0.738 | 5.43 | 0.394 | 1.73 | 0.908 | 1.60 |

由表中数据可知，转动松弛时间的趋势为 C—O>C—C>O—O；但分子动力计算的松弛时间约为布朗动力计算结果的1.5 倍。

此研究实例比较分子动力计算与布朗动力计算的差异性。虽然 EG 分子很小，但与溶剂水分子间有氢键的作用，导致分子的运动缓慢。类似的系统如生化分子的水溶液、聚合物等，主要的运动都比较缓慢。研究这种缓慢运动的分子体系，分子动力计算需要耗费许多计算时间才可得到满意

的结果，因此利用布朗动力计算应为较佳的选择。

## 13.8 朗之万方程式的一般解

Ermark 推导出利用数理方法解一般的布朗粒子运动方程式，即：

$$\vec{r}(t+\delta t)=\vec{r}(t)+c_1\vec{v}(t)\delta t+c_2\vec{a}(t)\delta t^2+\delta\vec{r}^G$$

$$\vec{v}(t+\delta t)=c_0\vec{v}(t)+c_1\vec{a}(t)\delta t+\delta\vec{v}^G \tag{13-42}$$

式中各常数项为：

$$c_0=\mathrm{e}^{-\xi\delta t}\approx 1-(\xi\delta t)+\frac{1}{2}(\xi\delta t)^2-\frac{1}{6}(\xi\delta t)^3+\frac{1}{24}(\xi\delta t)^4-\cdots$$

$$c_1=(\xi\delta t)^{-1}(1-c_0)\approx 1-\frac{1}{2}(\xi\delta t)+\frac{1}{6}(\xi\delta t)^2-\frac{1}{24}(\xi\delta t)^3+\cdots$$

$$c_2=(\xi\delta t)^{-1}(1-c_1)\approx \frac{1}{2}-\frac{1}{6}(\xi\delta t)+\frac{1}{24}(\xi\delta t)^2-\cdots \tag{13-43}$$

式(13-43) 中，位置与速度的随机数项为：

$$\delta\vec{r}^G=\int_t^{t+\delta t}\xi^{-1}[1-\mathrm{e}^{-\xi(t+\delta t-t')}]\vec{A}(t')\mathrm{d}t'$$

$$=\int_t^{t+\delta t}f_{\mathrm{r}}(t')\vec{A}(t')\mathrm{d}t'$$

$$\delta\vec{v}^G=\int_t^{t+\delta t}\mathrm{e}^{-\xi(t+\delta t-t')}\vec{A}(t')\mathrm{d}t'$$

$$=\int_t^{t+\delta t}f_{\mathrm{v}}(t')\vec{A}(t')\mathrm{d}t' \tag{13-44}$$

在计算中，随机数项的分配应合于高斯概率分配，即：

$$\rho(\delta r_\alpha^G,\delta v_\alpha^G)=\frac{1}{2\pi\sigma_{\mathrm{r}}\sigma_{\mathrm{v}}(1-c_{\mathrm{rv}}^2)^{1/2}}\times$$

$$\exp\left\{-\frac{1}{2(1-c_{\mathrm{rv}}^2)}\left[\left(\frac{\delta r_\alpha^G}{\sigma_{\mathrm{r}}}\right)^2+\left(\frac{\delta v_\alpha^G}{\sigma_{\mathrm{v}}}\right)^2-2c_{\mathrm{rv}}\left(\frac{\delta r_\alpha^G}{\sigma_{\mathrm{r}}}\right)\left(\frac{\delta v_\alpha^G}{\sigma_{\mathrm{v}}}\right)\right]\right\} \tag{13-45}$$

各随机数项的平均值为 0。其标准误差为：

$$\sigma_{\mathrm{r}}^2=\langle(\delta r_\alpha^G)^2\rangle=2\xi\frac{k_{\mathrm{B}}T}{m}\int_t^{t+\delta t}f_{\mathrm{r}}^2(t')\mathrm{d}t'=2\xi\frac{k_{\mathrm{B}}T}{m}\int_0^{\delta t}f_{\mathrm{r}}^2(t'')\mathrm{d}t''$$

$$=\delta t^2\frac{k_{\mathrm{B}}T}{m}(\xi\delta t)^{-1}[2-(\xi\delta t)^{-1}(3-4\mathrm{e}^{-\xi\delta t}+\mathrm{e}^{-2\xi\delta t})]$$

$$\sigma_{\mathrm{v}}^2=\langle(\delta v_\alpha^G)^2\rangle=2\xi\frac{k_{\mathrm{B}}T}{m}\int_t^{t+\delta t}f_{\mathrm{u}}^2(t')\mathrm{d}t'=2\xi\frac{k_{\mathrm{B}}T}{m}\int_0^{\delta t}f_{\mathrm{u}}^2(t'')\mathrm{d}t''$$

$$= \frac{k_B T}{m}(1 - e^{-2\xi\delta t}) \tag{13-46}$$

式(13-45) 中的相关系数 $c_{rv}$ 为:

$$c_{rv}\sigma_r\sigma_v = \langle \delta r_\alpha^G \delta v_\alpha^G \rangle = 2\xi \frac{k_B T}{m}\int_t^{t+\delta t} f_r(t') f_u(t') dt'$$

$$= 2\xi \frac{k_B T}{m}\int_0^{\delta t} f_r(t'') f_u(t'') dt'' = \delta t \frac{k_B T}{m}(\xi\delta t)^{-1}(1 - e^{-\xi\delta t})^2 \tag{13-47}$$

依此选择适当的 $\delta\vec{r}^G$ 与 $\delta\vec{v}^G$ 再依式(13-42) 计算 $\vec{r}(t+\delta t)$ 与 $\vec{v}(t+\delta t)$。Ermark 计算法在解 Langevin 方程式时，兼顾系统正常的动力行为与由随机力产生的动力行为。

若摩擦系数很低，$\xi \to 0$，则正常的动力行为为最重要的部分，Langevin 方程式近似于牛顿动力方程式。此时，比照分子动力计算的 Verlet 跳蛙法，可导出低摩擦系数系统的位置与速度的计算式为:

$$\vec{r}(t+\delta t) = \vec{r}(t) + c_1\vec{v}(t)\delta t + c_2\vec{a}(t)\delta t^2 + \delta\vec{r}^G$$

$$\vec{v}(t+\delta t) = c_0\vec{v}(t) + (c_1 - c_2)\vec{a}(t)\delta t + c_2\vec{a}(t+\delta t)\delta t + \delta\vec{v}^G \tag{13-48}$$

若仅需计算系统长时间范围的结构，则可应用简化的位置 Langevin 方程式，即:

$$\dot{\vec{r}}(t) = \frac{D}{k_B T}\vec{f}(t) + \dot{\vec{r}}^G(t) = \frac{f(t)}{m\xi} + \dot{\vec{r}}^G(t) \tag{13-49}$$

式中，$\dot{\vec{r}}^G(t)$ 为随机力速度，即:

$$\langle \dot{r}_\alpha^G(t)\dot{r}_\beta^G(0)\rangle = 2D\delta t\delta_{\alpha\beta} \tag{13-50}$$

依此，Ermark 导出位置的计算式为:

$$\vec{r}(t+\delta t) = \vec{r}(t) + \frac{D}{k_B T}\vec{f}(t)\delta t + \delta\vec{r}^G$$

$$= \vec{r}(t) + \frac{\vec{f}(t)}{m\xi}\delta t + \delta\vec{r}^G \tag{13-51}$$

式中，$\delta\vec{r}^G$ 为高斯分配的随机数，其平均值为 0，且$\langle(\delta r_\alpha^G)^2\rangle = 2D\delta t$。必须注意，此计算方法仅适用于研究系统长时间的结构行为，不适用于短时间范围的研究。

## 参 考 文 献


1 van Gunsteren W F, Berendsen H J C. Mol. Phys. 45, 637, 1982

2 van Gunsteren W F, Berendsen H J C, Rullmann J A C. Mol. Phys., 44, 69, 1981

3 Brunger A, Brooks C B, Karplus M. Chem. Phys. Lett., 105, 495, 1984

4 Ermark D L, Buckholtz H. J. Comput. Phys., 35, 169, 1980
5 Ermark D L, McCammon J A. J. Chem. Phys., 69, 1352, 1978
6 Ermark D L, Yeh Y. Chem. Phys. Lett., 24, 243, 1974
7 Widmalm G, Pastor R W. J. Chem. Soc. Faraday Trans., 88, 1747, 1992
8 Yun-yun S, Lu W, van Gunsteren W F. Mol. Simulation, 1, 369, 1988
9 Briki, et al. J. Chem. Soc. Faraday Trans, 88, 1746, 1992
10 Adolf D A, Ediger M D. in "Computer Simulation of Polymers", Prentice-Hall Int. Ed. New Jersey, USA, 1991

# 第14章 耗散粒子动力计算法

分子力学、分子动力学与蒙地卡罗等分子模拟方法适用于含有数万个粒子的系统，能够探讨的系统变化的时间范围为皮秒至纳秒。耗散粒子动力计算法为介于原子尺度与介观范围内的模拟方法，将分子视为一团或一堆，不但可以涵盖更多的粒子体系，而且可以研究较长时间内系统粒子的运动行为，如纳米级材料，自组装材料体系等。本章介绍近年来发展的一种新的模拟方法的原理和应用例子。耗散粒子动力计算的缺点则是不能探讨系统微观（micro-scale）的细部行为。

## 14.1 耗散粒子动力计算

耗散粒子动力计算（dissipative particle dynamics，DPD）为介于原子尺度与介观范围内的模拟方法。此种模拟方法的优点在于能有效地探讨分子的堆积与分散的问题，适用于研究混合体系的相行为。DPD 模拟的体系最多可包含 $10^6$ 个原子，并且可以研究微秒时间范围内的动力行为。其缺点则是只适用于探讨介观（meso-scale）特性，而不能探讨微观（micro-scale）的细部行为。在 DPD 的模拟方法中，将分子视为一团或一堆。依照牛顿运动方程式，即：

$$\frac{d\vec{r}_i}{dt}=\vec{v}_i,\quad \frac{d\vec{v}_i}{dt}=\vec{f}_i \tag{14-1}$$

每分子团所受的力为：

$$\vec{f}_i=\sum_{j\neq i}(\vec{F}_{ij}^{C}+\vec{F}_{ij}^{D}+\vec{F}_{ij}^{R}) \tag{14-2}$$

式中，$\vec{F}_{ij}^{C}$为分子团间的保守力（conservative force）；$\vec{F}_{ij}^{D}$为耗散力

(dissipative force)；$\vec{F}_{ij}^{R}$为随机力（random force)。Groot 与 Warren 建议作用力的形式为：

$$\vec{F}_{ij}^{C}=\begin{cases}a_{ij}(1-r_{ij})\hat{r}_{ij} & (r_{ij}<1)\\ 0 & (r_{ij}\geqslant 1)\end{cases} \tag{14-3}$$

式中，$a_{ij}$为 $i$ 团与 $j$ 团间的最大排斥力。$\vec{r}_{ij}=\vec{r}_i-\vec{r}_j$，$r_{ij}=|\vec{r}_{ij}|$，$\hat{r}_{ij}=\vec{r}_{ij}/|\vec{r}_{ij}|$。

耗散力与随机力为：

$$\vec{F}_{ij}^{D}=-\gamma w^{D}(r_{ij})(\hat{r}_{ij}-\vec{v}_{ij})\hat{r}_{ij}$$

$$\vec{F}_{ij}^{R}=\sigma w^{R}(r_{ij})\theta_{ij}\hat{r}_{ij} \tag{14-4}$$

式中，$\vec{v}_{ij}=\vec{v}_i-\vec{v}_j$，$w^D$ 与 $w^R$ 为权重因子（weighting factor)，为团间距离 $r$ 的函数。当 $r>r_C=1$ 时，$w^D$ 与 $w^R$ 均为零。$\theta_{ij}(t)$ 为高斯分布的随机函数，即：

$$\langle\theta_{ij}(t)\rangle=0$$

$$\langle\theta_{ij}(t)\theta_{kl}(t')\rangle=(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})\delta(t-t') \tag{14-5}$$

Español 与 Warren 建议权重因子的形式为：

$$w^{D}(r)=[w^{R}(r)]^2 \qquad \sigma^2=2\gamma k_B T \tag{14-6}$$

依此关系，可将权重因子简单表示为：

$$w^{D}(r)=[w^{R}(r)]^2=\begin{cases}(1-r)^2 & (r<1)\\ 0 & (r\geqslant 1)\end{cases} \tag{14-7}$$

分子团的运动方程式，可由类似 Verlet 的方法得解为：

$$\vec{r}_i(t+\Delta t)=\vec{r}_i(t)+\vec{v}_i(t)\Delta t+\frac{1}{2}\vec{f}_i(t)\Delta t^2$$

$$\vec{v}_i(t+\Delta t)=\vec{v}_i(t)+\lambda\vec{f}_i(t)\Delta t$$

$$\vec{f}_i(t+\Delta t)=\vec{f}_i[\vec{r}(t+\Delta t),\vec{v}(t+\Delta t)]$$

$$\vec{v}_i(t+\Delta t)=\vec{v}_i(t)+\frac{1}{2}[\vec{f}_i(t)+\vec{f}_i(t+\Delta t)]\Delta t \tag{14-8}$$

式中，$\lambda$ 值通常定为 1/2，$\lambda$ 的大小与所选取的积分步长 $\Delta t$ 相关。Groot 与 Warren 测试各种 $\lambda$ 值与 $\Delta t$，得到当 $\lambda=0.5$ 时，$\Delta t$ 应小于或等于 0.04。依此，计算的速度与温度满足式(14-9) 的关系。

$$k_B T=\langle v^2\rangle/3 \tag{14-9}$$

DPD 计算中，最重要的是选取团与团间的作用参数 $a_{ij}$。此参数可依据高分子系统的 Flory-Huggins（FH）相混理论求得。FH 理论适用于研

究液态-液态、液态-固态的高分子混合体系。依 FH 理论，两成分的混合系统，其平均自由能可表示为：

$$\frac{F}{k_B T}=\frac{\phi_A}{N_A}\ln\phi_A+\frac{\phi_B}{N_B}\ln\phi_B+\chi\phi_A\phi_B \tag{14-10}$$

式中，$\phi_A$ 与 $\phi_B$ 为成分 A 与 B 的体积分率，$\phi_A+\phi_B=1$。$N_A$ 与 $N_B$ 为成分 A 与 B 所含的片段（segment）数目。$\chi$ 为 A 与 B 间的作用参数，即：

$$\chi N_A=\frac{\ln[(1-\phi_A)/\phi_A]}{1-2\phi_A} \tag{14-11}$$

将 FH 理论与 DPD 方法比较，可推得参数 $a$ 与 $\chi$ 间的关系为：

$$\chi=\frac{2\alpha(a_{AB}-a_{AA})(\rho_A+\rho_B)}{k_B T} \tag{14-12}$$

式中，$\alpha$ 为校正系数，其值为 0.1012±0.001。$\rho_A$ 与 $\rho_B$ 为 A、B 的密度，$\rho_A=N_A/V$。设 $\rho=\rho_A+\rho_B$ 为系统粒子团的密度，对单一物种的系统，可得到式(14-13) 的关系

$$a_{AA}=75k_B T/\rho \tag{14-13}$$

对二成分体系，则可得到式(14-14) 良好的线性关系。

$$\begin{aligned} a_{AB}(T)&=a_{AA}+1.451\chi_{AB}(T) \quad (\rho=5)\\ a_{AB}(T)&=a_{AA}+3.497\chi_{AB}(T) \quad (\rho=3) \end{aligned} \tag{14-14}$$

## 14.2 离子型表面活性剂胶束的计算

水溶液中，双性（amphiphilic）聚合物的聚集现象是一种非常普遍的现象。例如许多生物膜、油性分子、接口活性剂等，都发现这种聚集体(aggregates)。双性聚合物聚集的形态和大小与其分子本身的特性及溶液的浓度相关。Mao 等利用核磁共振技术研究水溶液中离子型及非离子型表面活性剂 $n$-$CH_3(CH_2)_{12}OSO_3Na$（简称 STS）及 $n$-$CH_2(CH_2)_{13}O(C_2H_4O)_3H$（简称 $C_{14}E_3$）的聚集行为。实验结果显示，这两种活性剂溶液及其混合溶液均形成微胞（micelle）聚集体。实验量测出特殊碳原子的扩散系数，借以推断微胞可能的大小与构造。虽然，核磁共振实验可得到许多重要的信息，但无法彻底了解微胞的构造。为了弥补实验的不足，Kuo 等于是利用 DPD 计算检视此接口活性剂溶液。

执行 DPD 计算时 Kuo 等将 STS 及 $C_{14}E_3$ 皆分成体积约略相当的两段，STS 分子分为 C 及 S［$CH_3(CH_2)_7$ 和 $(CH_2)_6OSO_3^-$］两片段，$C_{14}E_3$ 则分

为 E 及 O [$CH_3(CH_2)_{11}$和 $(C_2H_4O)_4H$] 两片段；水则单独视为一个 DPD 分子。计算中将 S 片段视为带负电的离子片段。利用分子力学方法及 UNIVERSAL 力场先得到各分子片段的几何优化结构。根据这些结构，利用量子力学的半经验方法 INDO/2 可计算出各片段原子所带的部分电荷。依照 Ryikina 等所提出的方法，分子团间的平均混合能量可由式(14-15) 算出。

$$E_{AB}^{mix}=\frac{1}{2}[Z_{AB}\langle E_{AB}(T)\rangle+Z_{BA}\langle E_{BA}(T)\rangle-Z_{AA}\langle E_{AA}(T)\rangle-Z_{BB}\langle E_{BB}(T)\rangle] \tag{14-15}$$

式中，$Z$ 为平均的配位数 (coordination number)；$\langle E(T)\rangle$ 为平均的作用能。$Z_{AB}$表示以 A 为参考分子，B 对 A 的配位数。平均作用能可利用蒙地卡罗方法得到：

$$\langle E_{AB}(T)\rangle=\frac{\int dE_{AB}P(E_{AB})E_{AB}e^{-E_{AB}/k_BT}}{\int dE_{AB}P(E_{AB})e^{-E_{AB}/k_BT}} \tag{14-16}$$

由计算所得的平均混合能量，可得到 Flory-Huggins 的混合参数：

$$\chi_{AB}(T)=E_{AB}^{mix}/RT \tag{14-17}$$

得到 $\chi_{AB}$后，代入式(14-14) 可得各团间的作用参数 $a_{AB}$。表 14-1 为计算的 $a_{AB}$值，表中各参数均为 DPD 单位。

**表 14-1　模拟体系中各粒子的作用参数 $a_{AB}$**

| 项　目 | C | S | E | O | W |
|---|---|---|---|---|---|
| C | 15 | 69 | 79 | 74 | 81 |
| S | 69 | 15 | 53 | 32 | 0 |
| E | 79 | 53 | 15 | 81 | 105 |
| O | 74 | 32 | 81 | 15 | 85 |
| W | 81 | 0 | 105 | 85 | 15 |

以计算所得的作用参数，执行 STS、$C_{14}E_3$ 及其混合体系的结果如表 14-2 所示。表中列出不同浓度与不同系统所得聚集胶束的形状，约略大小与成分。表中各数据均为 DPD 的简化单位。

图 14-1 为模拟所得纯 STS 与 $C_{14}E_3$ 水溶液的聚集图。图中显示 STS 水溶液随浓度的不同形成不同的聚集相，$C_{14}E_3$ 水溶液则形成微胞状的聚集物。图 14-2 为模拟所得当 STS 与 $C_{14}E_3$ 物质的量相同时，在不同浓度下的聚集形态图。图中显示在低浓度时，形成微胞状的聚集物。微胞的结构外层均为 STS 分子，内层则为 $C_{14}E_3$ 分子。当浓度增高时，微胞与微胞互相聚集，系统进一步形成柱状的结构。图 14-3 为系统粒子密度的等高线图，借由此种

**表 14-2　不同浓度的 STS、$C_{14}E_3$ 水溶液所形成聚集体的形态、大小及聚集体的成分**

| STS：$C_{14}E_3$：$H_2O$ | 结　构 | 直　径 | 成　分 |
|---|---|---|---|
| 20：0：80 | 球形微胞 | 约 4～4.5 | 80～120 STS |
| 45：0：55 | 六角形排列 | 4.5(截面积) | |
| 70：0：30 | 层状 | 4(层间距) | |
| 0：20：80 | 球形微胞 | 6.6 | 500 $C_{14}E_3$ |
| 5：5：90 | 扁圆形微胞 | 8 | 外层 125 STS，内层 125 $C_{14}E_3$ |
| 10：10：80 | 细长扁圆形微胞 | 8.5 | 外层 250 STS，内层 250 $C_{14}E_3$ |
| 15：15：70 | 圆柱状微胞 | 8.5 | 外层 125 STS，内层 125 $C_{14}E_3$ |
| | 球形微胞 | 3.5 | 40 STS |
| 15：5：80 | 球形微胞 | 9 | 外层 275 STS，内层 125 $C_{14}E_3$ |
| | 球形微胞 | 4 与 2 | 80 STS 与 20 STS |
| 20：10：70 | 球形微胞 | 9 | 外层 380 STS，内层 250 $C_{14}E_3$ |
| | 球形微胞 | 约 4～4.5 | 120 STS |
| 25：5：70 | 球形微胞 | 5.5 | 外层 250 STS，内层 125 $C_{14}E_3$ |
| | 球形微胞 | 4.5 与 2 | 120 STS 与 20 STS |
| 30：5：65 | 球形微胞 | 5.5 | 外层 250 STS，内层 125 $C_{14}E_3$ |
| | 球形微胞 | 4.5 与 2 | 120 STS 与 20 STS |

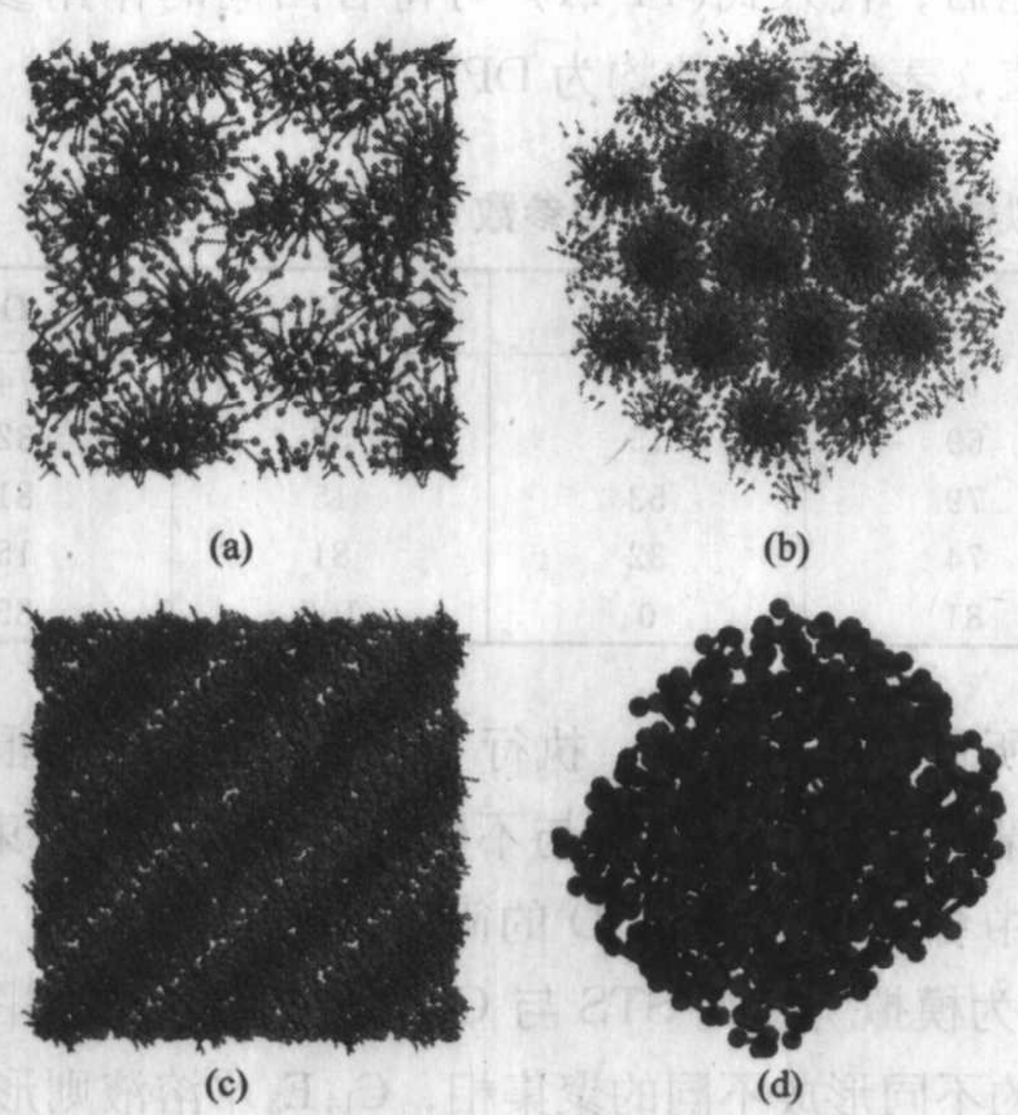



**图 14-1　纯 STS 与 $C_{14}E_3$ 水溶液的聚集形态**

(a) STS：$H_2O$=20：80；(b) STS：$H_2O$=45：55；
(c) STS：$H_2O$=70：30；(d) $C_{14}E_3$：$H_2O$=20：80

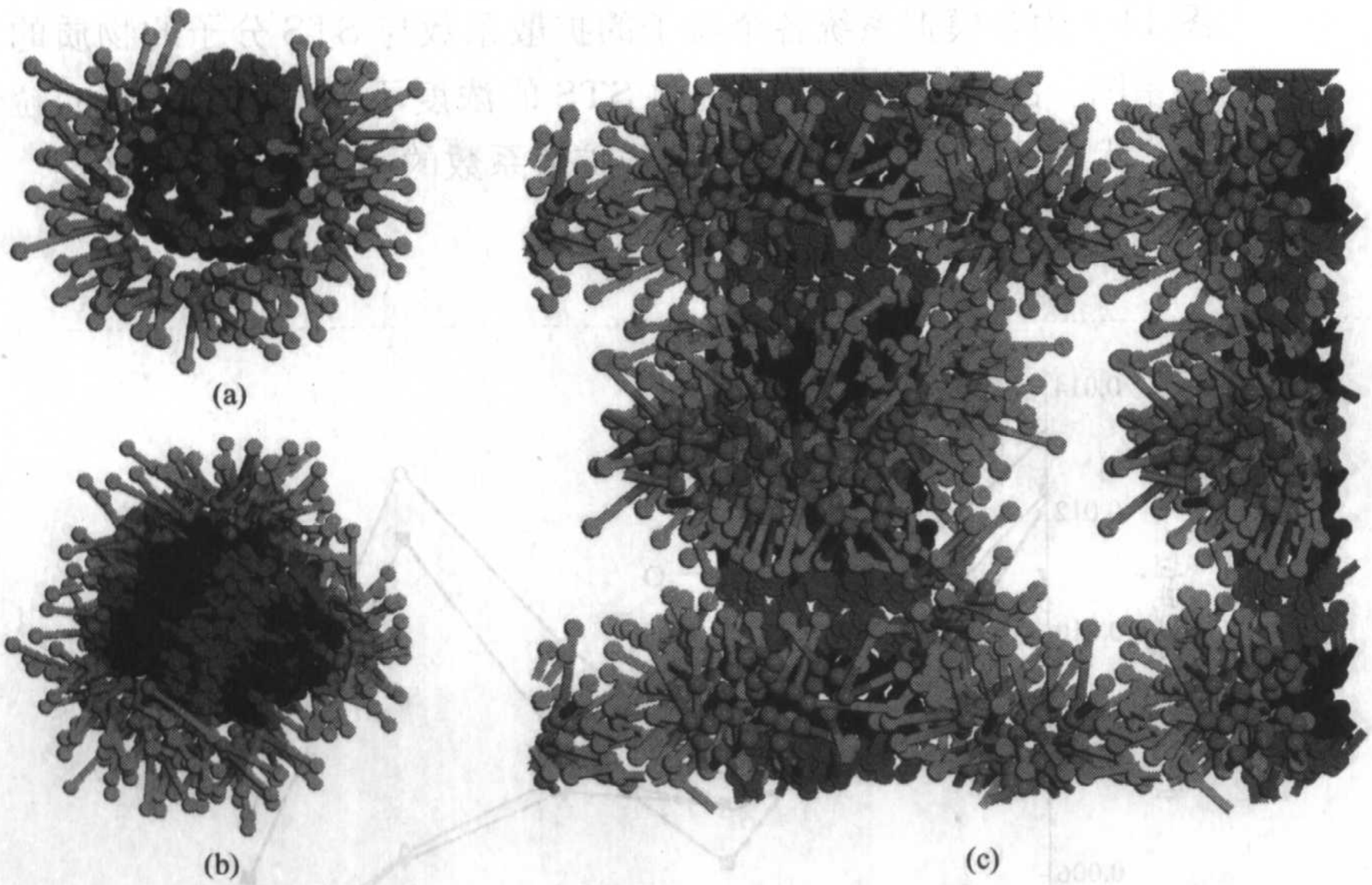



**图 14-2　相同物质的量之比的 STS 与 $C_{14}E_3$ 水溶液在不同浓度时的聚集形态**

(a) STS：$C_{14}E_3$：$H_2O$=5：5：90；(b) STS：$C_{14}E_3$：$H_2O$=10：10：80；

(c) STS：$C_{14}E_3$：$H_2O$=15：15：70

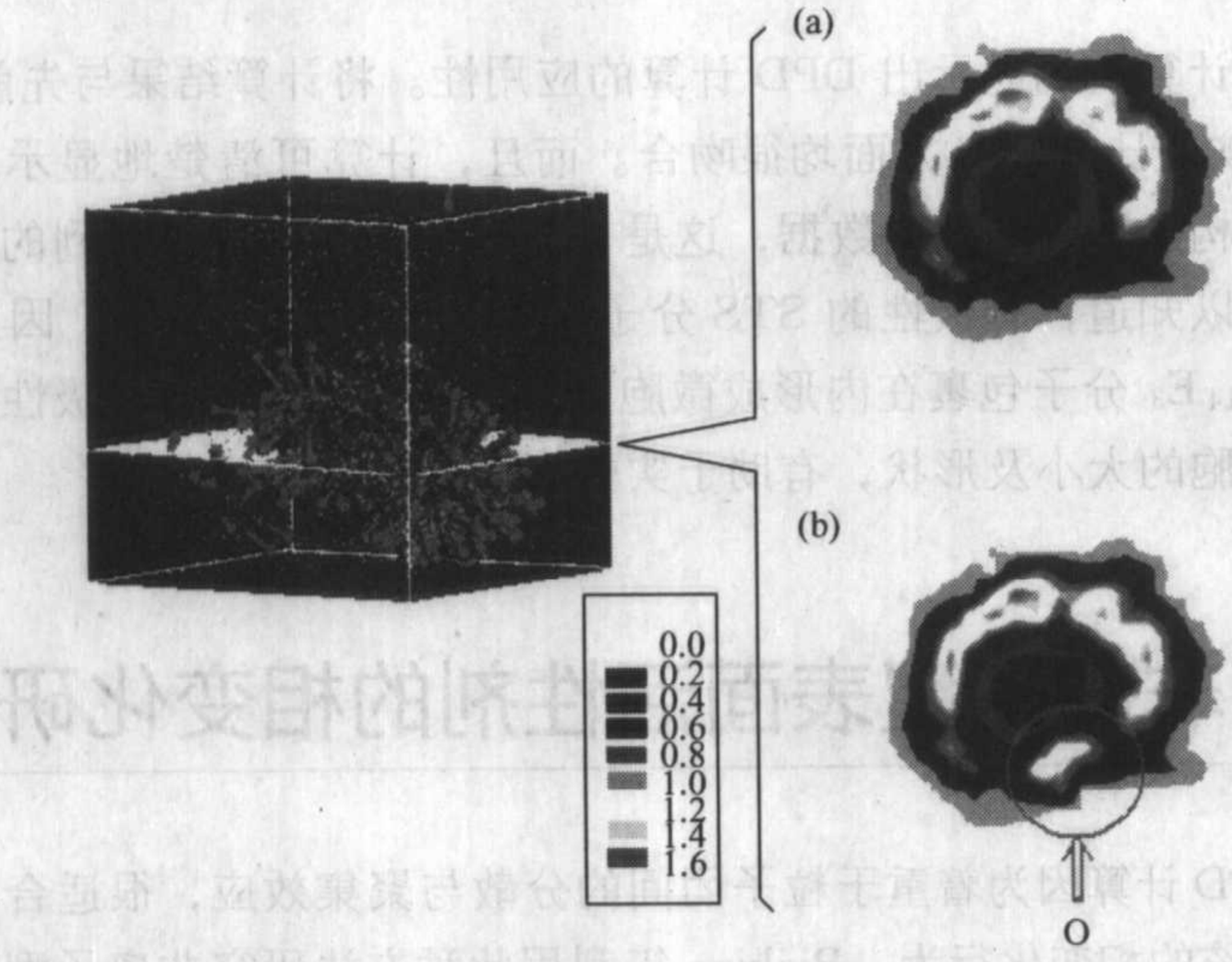



**图 14-3　S 及 O 的粒子密度等高线切面图**

计算的体系为 STS：$C_{14}E_3$：$H_2O$=5：5：90

粒子密度的等高线图可以估算出聚集体内各种成分原子的数目。

图 14-4 为各模拟系统各个粒子的扩散系数与 STS 分子的物质的量之比的关系图。图中显示随着系统中 STS 的浓度不同，呈现不同的粒子扩散系数。大体而言，约略可将系统依扩散系数的不同分成三个区间。表示系统可呈现三种不同的相变化。

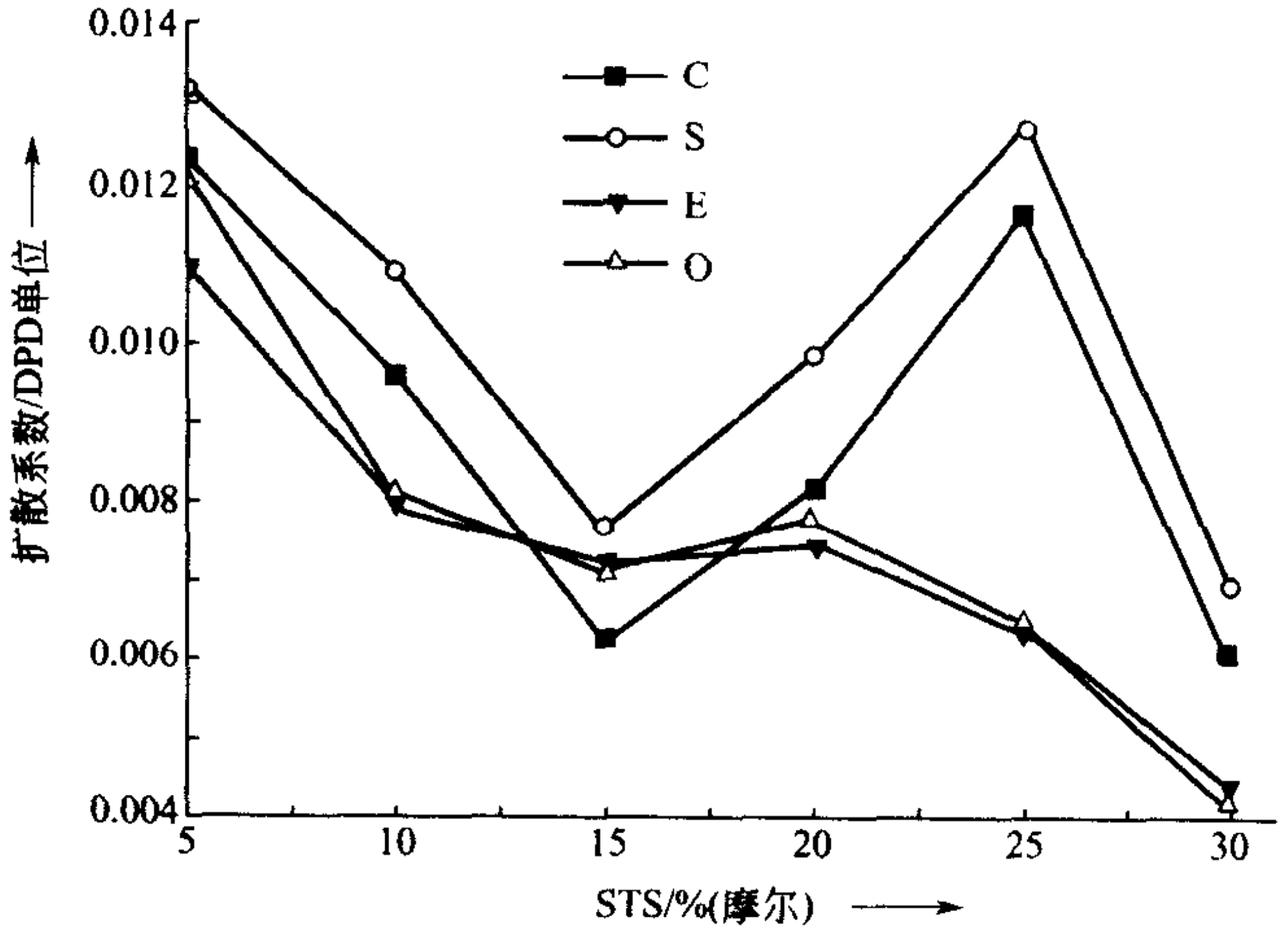



**图 14-4　粒子的扩散系数与溶液中 STS 物质的量之比的关系**

此计算实例显示出 DPD 计算的应用性。将计算结果与先前 Mao 等的 NMR 实验比较，各方面均很吻合。而且，计算可清楚地显示出如微胞的大小、构造等各细部的数据，这是一般的实验不能轻易得到的。从计算的结果可以知道，有极性的 STS 分子在水溶液中运动较快，因此易将非极性的 $C_{14}E_3$ 分子包裹在内形成微胞。利用分子的极性与非极性的关系可以控制微胞的大小及形状，有助于实验的设计。

## 14.3　非离子型表面活性剂的相变化研究

DPD 计算因为着重于粒子团间的分散与聚集效应，很适合用以研究多成分系统的相变化行为。Ryjkina 等利用此种方法研究非离子型表面活性剂 $C_{10}H_{25}N(CH_3)_2O$（简称 DDAO）水溶液的相变化行为。一般高分子溶液的相随其成分的不同而变化。图 14-5 为溶液的相与表面活性剂浓度关系的示意。

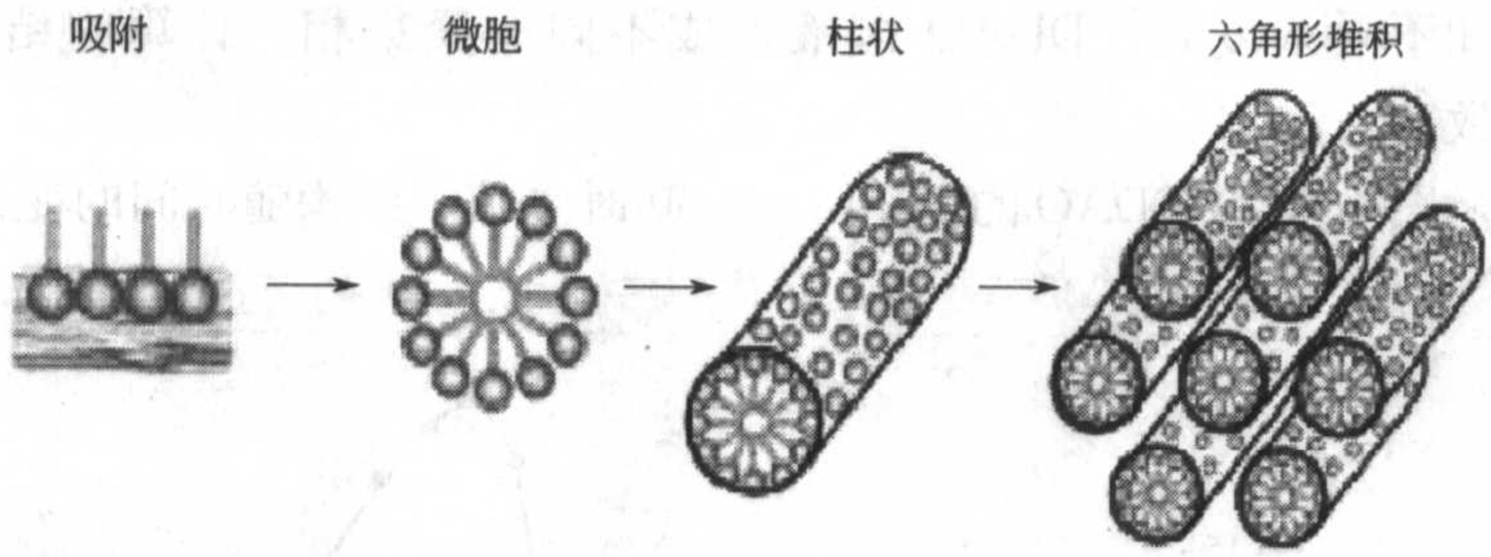



**图 14-5　表面活性剂系统浓度增加时系统的聚集形态**

大体而言，当浓度最低时形成吸附层；然后形成球状微胞，继而形成杆状微胞，而后为正六角相，当浓度再增加时形成多层状结构。DPD 计算中，将 DDAO 分子分成体积约略相当的两团，分别为具极性的 $C_5H_{10}N(CH_3)_2O$（N 片段）及无极性的 $C_7H_{15}$（C 片段）。图 14-6 及图 14-7 为实验与计算结果的比较。图中，DDAO 的物质的量之比分别为 0.15、0.50 及 0.80，温度为 303.15K。M1、H1、L1 为以杆状方式显示

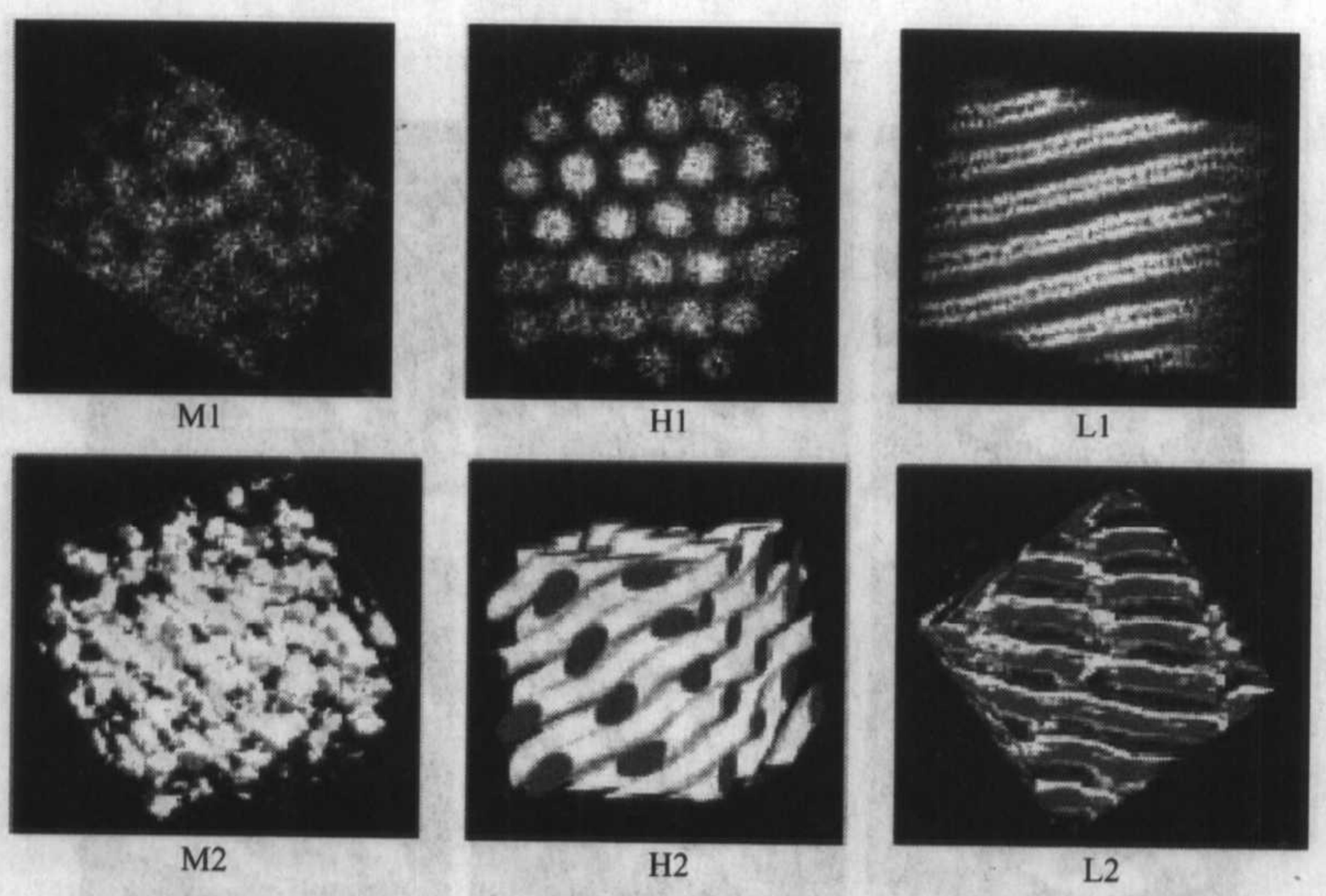



**图 14-6　以密度等高线图及球杆模型表示 DDAO 水溶液的相变化图**

M1，M2 为微胞相（$X_{DDAO}$= 0.15），H1，H2 为六角形堆积（$X_{DDAO}$=0.50），L1，L2 为层状结构（$X_{DDAO}$=0.80），计算温度皆为 303.15K

计算的结果，M2、H2、L2 则以密度等高线的方式显示计算的结果。由图可知不同浓度时，DDAO 溶液形成不同的聚集相，计算的结果与实验一致。

图 14-8 为 DDAO 的 $X_{DDAO}=0.50$ 时系统中结构随时间的变化过程。

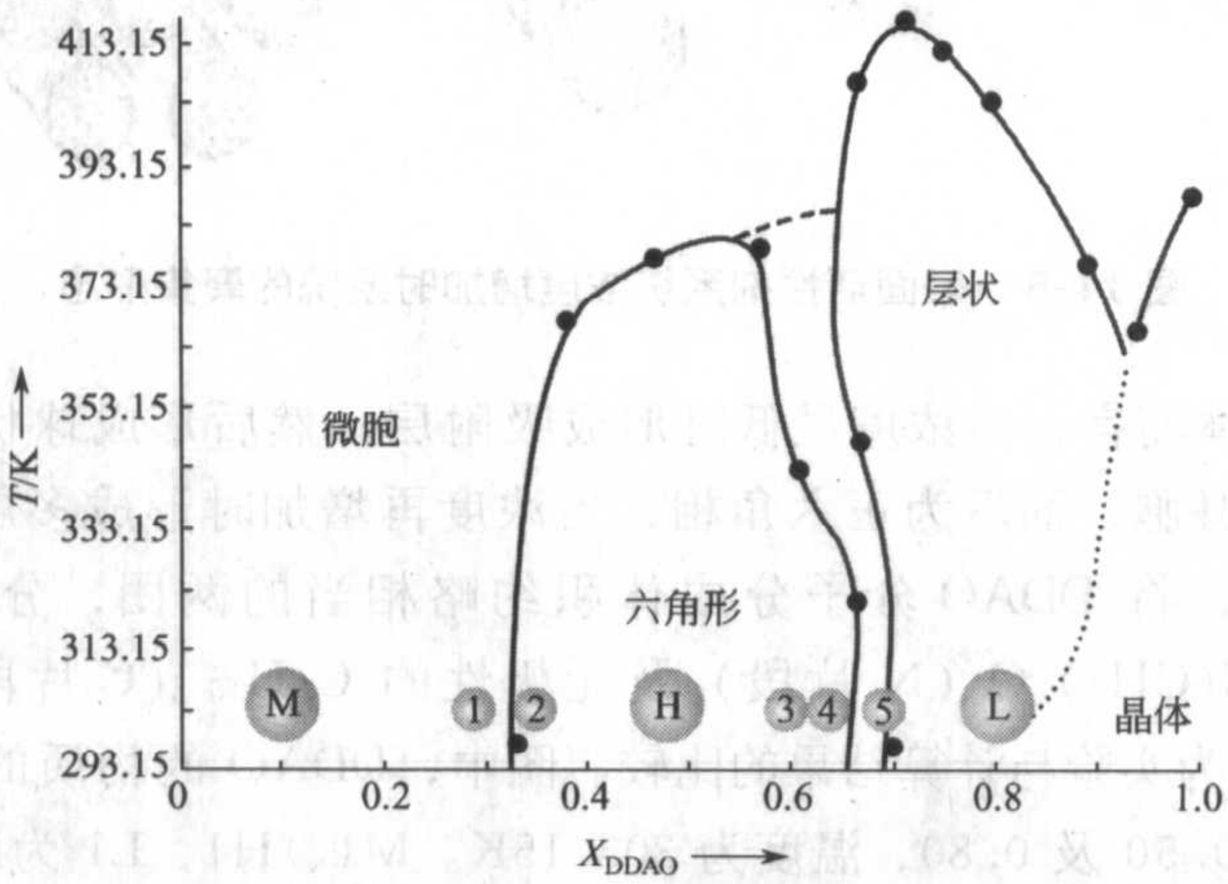



**图 14-7　实验所得 DDAO 水溶液于 303.15 K 时的相图与 DPD 计算的比较**

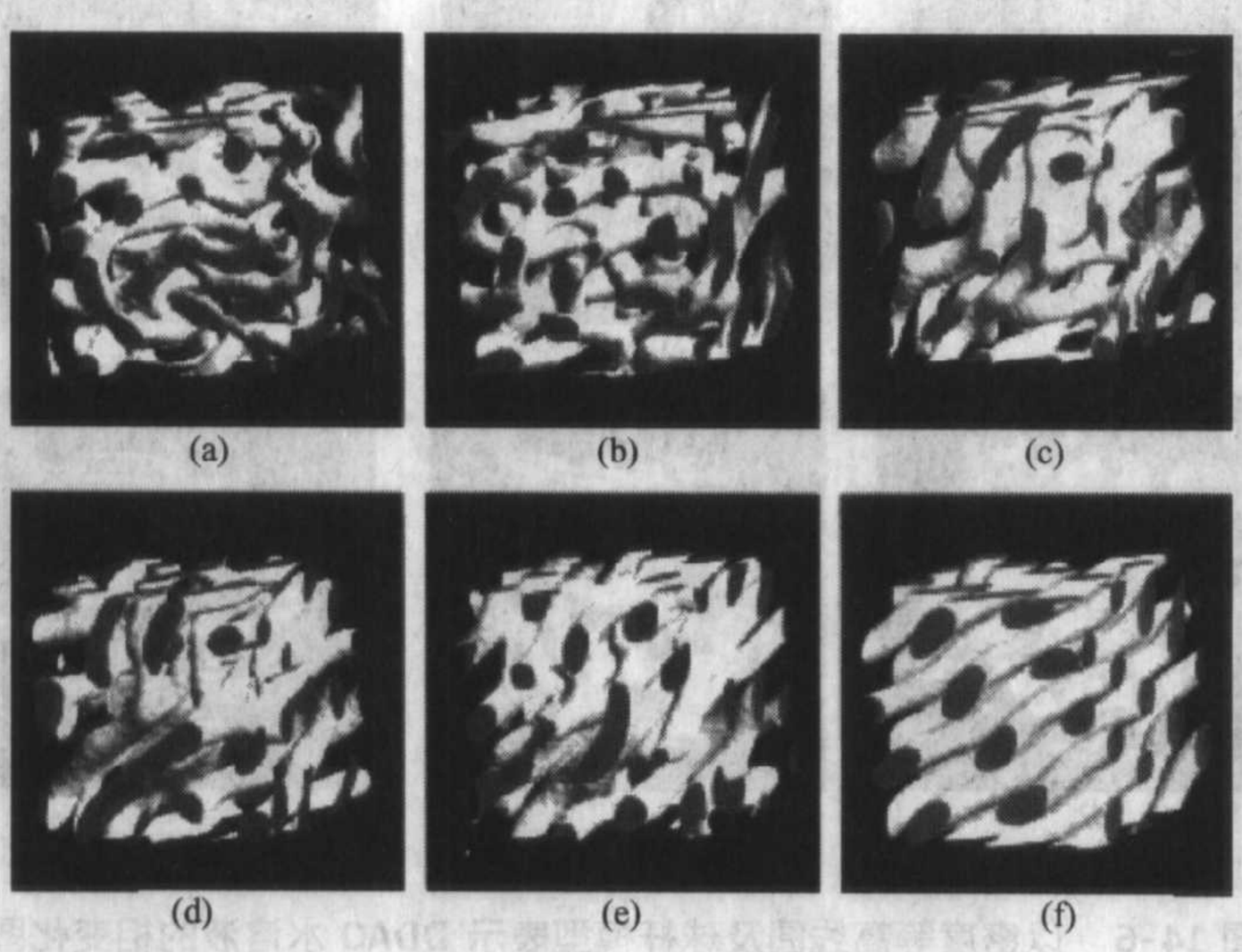



**图 14-8　在不同计算时间所得系统形成六角形堆积的过程**

$X_{DDAO}=0.50$；$T=303.15$K；(a) 1.8ns；(b) 9.3ns；(c) 18.6ns；(d) 22.3ns；(e) 27.9ns；(f) 35.7ns

由图可知，计算初始时系统为螺旋状（gyroid）的形态［图(a)、(b)］，随计算时间的增长，系统开始呈现对称性较高的排列［图(c)、(d)］至计算时间为 27.9ns 后系统呈现高对称性的六角形（hexagonal）柱状堆积［图(e)、(f)］。当计算时间超过 35.7ns，系统的结构不再发生任何变化，大体上维持高对称的堆积形态。

此例显示 DPD 计算在研究多成分系统相行为的实用性。其他的计算方法如分子动力计算、蒙地卡罗计算等比较，此种计算方法速度快了许多，研究的体系与时间范围都增加了两个数量级以上。由于计算中增加了耗散力与随机力，因此体系可适当的反映出堆积的形态。由于 DPD 计算发展的历史很短，目前仍有许多实际系统尚待测试。但毋庸置疑，此种计算方法具有相当的发展潜力，值得推广与探究。

## 14.4 脂质分子的相行为计算

粒子扩散动力计算法由于可包括大量的粒子，且可正确地反映出系统的相行为，已逐渐地被应用于研究生物分子体系的相行为。此处以 Li 等的工作列举应用粒子扩散动力学计算研究类似细胞膜的油脂分子（lipid）的组装与相变化与反相（inverted phase）变化。所谓的反相是指由溶剂分子而非溶质分子所形成的堆积相。在油脂分子的体系中，溶剂分子为水分子，溶质分子为油脂分子。当油脂分子的含量（%）高于水分子的含量（%）时，水分子堆积成不同的微胞相，称为“反相”。反相结构表示油脂分子形成的细胞膜中的孔道形状与大小。Li 等的计算，除了分子团间的作用力、粒子的耗散力，与随机力以外，计算中还考虑粒子间的键结作用力与键角弯曲作用力。图 14-9 为油脂分子及水分子简化的结构图。

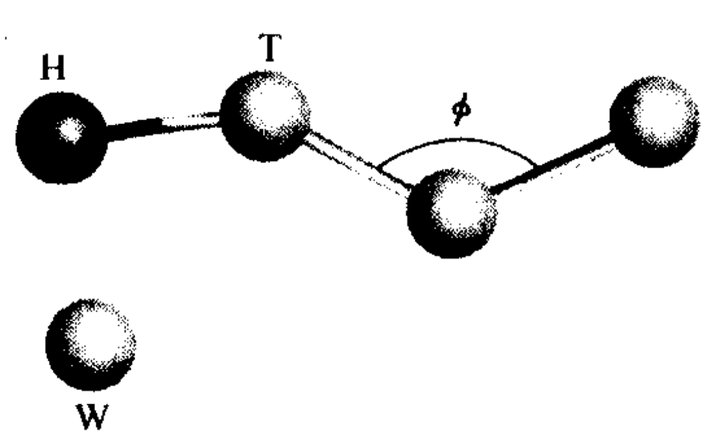



**图 14-9　油脂分子及水分子的简化结构**

图中“H”为油脂分子的头端亲水基团；“T”为疏水基团；“$\phi$”为弯曲键角。“HTN”表示一个亲水的头基与 N 个疏水的基团所组成的链状油脂分子。Li 等所使用的键结势能为：

$$U_{\text{bond}}(l)=\frac{1}{2}K(l-l_0)^2 \quad (14\text{-}18)$$

式中，$l$ 为键长；$K$ 为弹力常数；$l_0=0.5r_0$ 为平衡键长；$r_0$ 为 DPD 计算

的长度单位；$K=128k_BT/r_0^2$。

其疏水基团间键角弯曲势能为：

$$U_{angle}(\phi)=K_\phi[1-\cos(\phi-\phi_0)] \tag{14-19}$$

式中，力常数 $K_\phi=5k_BT$。

表 14-3 为计算所使用的作用常数（$a_{ij}$）、耗散力常数（$\gamma_{ij}$）与力常数。

**表 14-3　油脂分子的基团与水分子的作用常数、 耗散力常数与力常数**

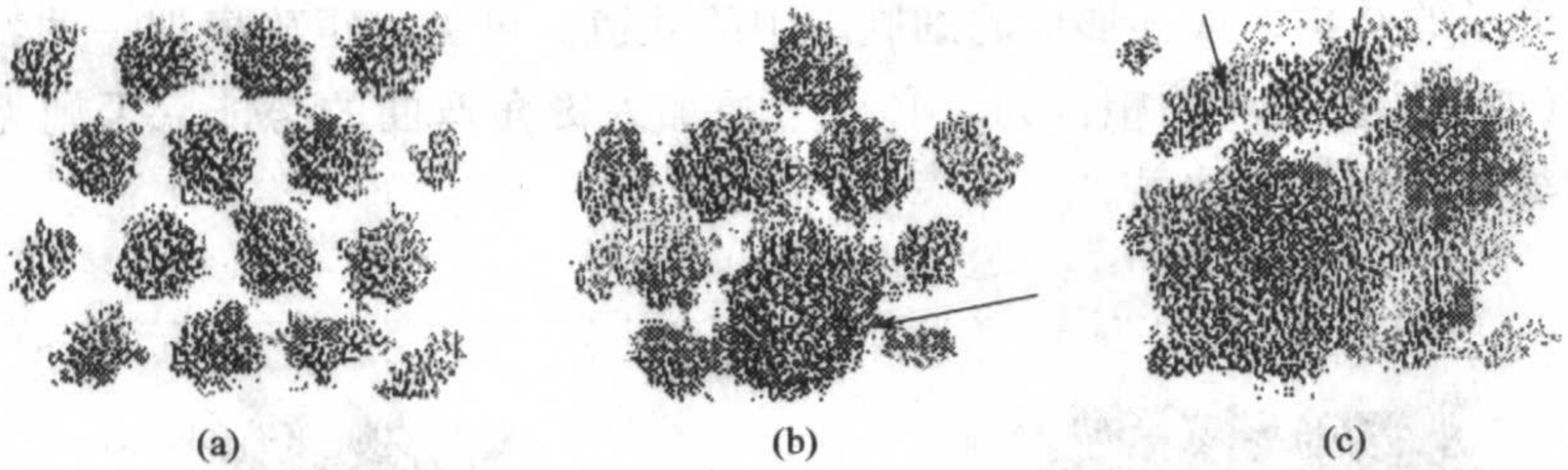
| 基团对 | $a_{ij}(k_BT/r_0)$ | $\gamma_{ij}(\sqrt{m_0k_BT/r_0^2})$ | $K(k_BT/r_0^2)$ | $l_0(r_0)$ |
|---|---|---|---|---|
| HH | 25 | 4.5 | | |
| TT | 25 | 4.5 | 128 | 0.5 |
| WW | 25 | 4.5 | | |
| HT | 95 | 20 | 128 | 0.5 |
| TW | 95 | 20 | | |
| WH | 25 | 4.5 | | |

Li 等的计算除了上述的各种力常数之外，还考虑了头端基团大小对各种相所造成的影响。符号“HTN($a$)”表示 HTN 分子的头基大小为 $ar_0$。表 14-4 为油脂分子各种头基大小所形成的微胞形状。

**表 14-4　油脂分子各种头基大小所形成的微胞形状**

| 头基大小 | HT1 | HT2 | HT4 | HT6 | HT10 | HT12 |
|---|---|---|---|---|---|---|
| 1.2 | 球状 | 圆盘状 | 圆盘状 | 反双锥体 | 反双锥体 | 反双锥体 |
| 1.1 | 柱状 | 圆盘状 | 反双锥体 | 反双锥体 | 反六角状 | 反六角状 |
| 1.0 | 柱状 | 圆盘状 | 反六角状 | 反六角状 | 反微胞状 | 反微胞状 |

计算的 HT1 与 HT2 系统共包括 8000 个 HT1 或 HT2 基团，HT3 与 HT4 系统包括 27000 个 HT3 或 HT4 基团，HT10 与 HT12 系统则包括 64000 个 HT10 或 HT12 基团。图 14-10 例示所谓的反相图。图 14-10(a)、(b)、(c) 显示计算系统中所有的水分子，所有的油脂分子基团则隐而不现。计算系统为 HT4(1.0)，图 14-10(a) 的系统浓度为 80%，计算的结果显示形成反六角形（inverted-hexagonal）的微胞；图 14-10(b) 的系统浓度为 70%，计算结果显示水分子聚集形成一个如箭头所指的较大微胞，其余的水分子则堆积形成类似六角形的微胞；图 14-10(c) 的计算系统浓度为 65%，计算结果显示水分子形成两个体积较大的微胞，其余的水分子则堆积成柱状，如图中的箭头所指。

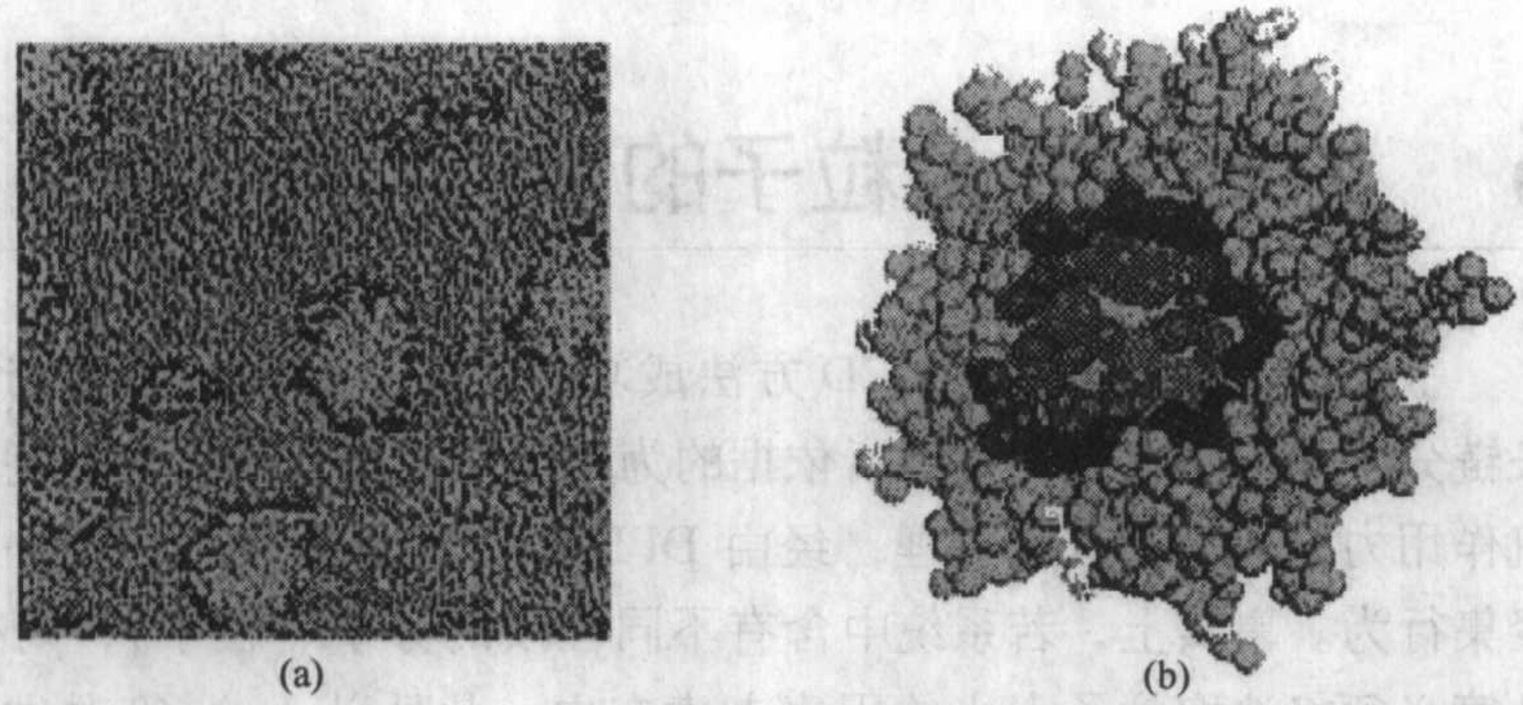


(a) (b) (c)

**图 14-10　HT4(1.0) 计算系统在各浓度所形成的反相图**

(a) 浓度为 80%；(b) 浓度为 70%；(c) 浓度为 65%

图 14-10 显示短链的油脂分子在不同的浓度形成不同的反相。图 14-11 为油脂分子的疏水链增长时，在不同的浓度条件下所形成反相结构。图 14-11(a) 为 HT12(1.0) 系统在浓度为 90% 的条件下所形成的反相。图 14-11(b) 为图 14-11(a) 中水分子微胞的结构，图中的黑球为油脂分子的头基，浅灰色球为油脂分子的疏水基团。



(a) (b)

**图 14-11　(a) HT12(1.0) 计算系统在浓度 90% 所形成的反相图；(b) 为 (a) 图中微胞的放大图**

图 14-12(a) 为 HT10(1.1) 系统在浓度为 90% 的条件下所形成的反相。图中显示，水分子形成柱状的堆集，柱状的水分子在系统中形成二度空间的六角形结构。图 14-12(b) 为图 14-12(a) 中水分子微胞的结构，图中的黑球为油脂分子的头基，浅灰色球为油脂分子的疏水基团。

Li 等所计算的油脂分子在水中所形成的相与反相结构，与实验所测量

的结果一致。此计算显示利用粒子扩散动力计算方法，可正确地模拟油脂分子的结构，或其形成的细胞膜中的孔道。可见，正确模拟一些特殊系统（如油脂分子）的相行为，必须考虑加入键角弯曲的势能，否则无法得到与实验相符的结果。

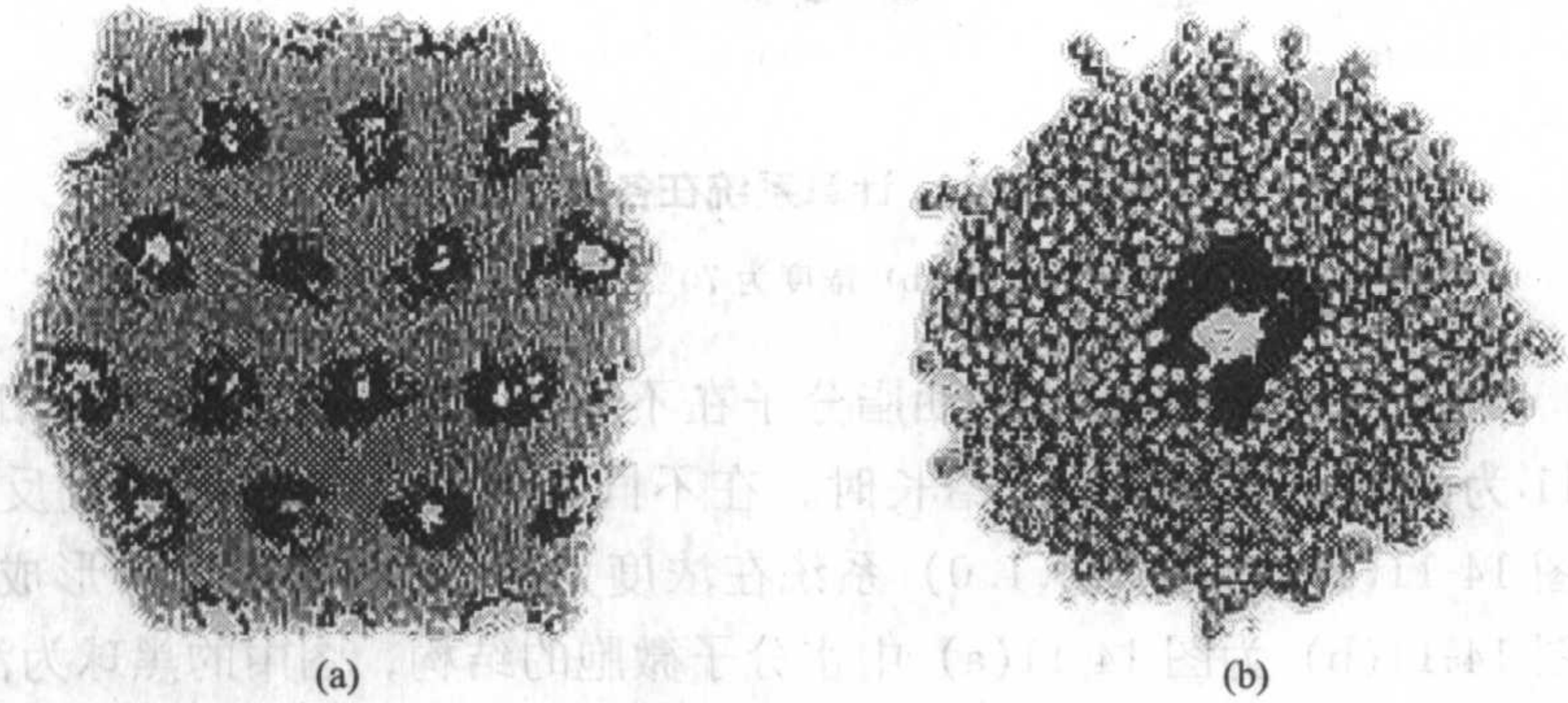
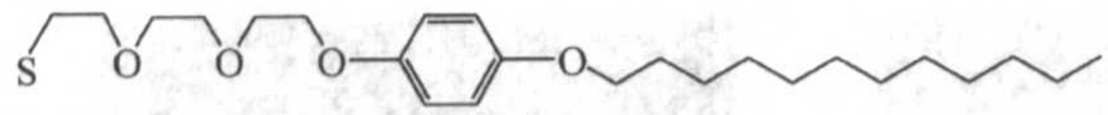



**图 14-12　(a) HT10（1.1）计算系统在浓度 90% 所形成的反相图；(b) 为 (a) 图中微胞的放大图**

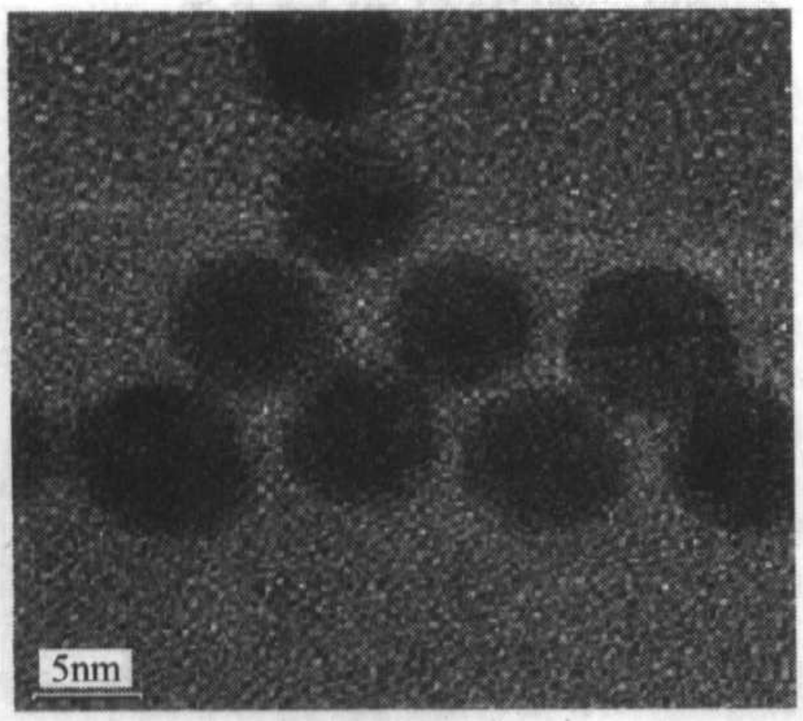
## 14.5　纳米尺寸金粒子的 DPD 模拟

上述的模拟实例，将 DPD 方法成功地应用于研究离子性与非离子性长链分子的聚集行为。模拟所依据的为粒子间的相互排斥力，耗散力与随机作用力的大小。依此原理，经由 DPD 的方法应可模拟任何一种系统的聚集行为。事实上，若系统中含有不同性质的分子（粒子），则执行 DPD 计算必须将选取粒子大小的因素考虑在内。此处以 Juan 等对纳米尺寸金粒子的 DPD 模拟为例，说明此种方法研究混合体系相行为的可行性及模拟中粒子大小选取的重要性。实验显示有机硫醇分子（thiol-terminated oligohydroquinoyl，TTOE）在甲苯溶剂中可吸附于金粒子（Au）的表面而形成纳米尺寸的球状聚集物。图 14-13 为 TTOE 的分子结构与利用穿透式电子显微镜（transmission electron microscopy，TEM）所观测的 TTOE-Au 所形成的金球粒子与粒子的结构示意图。TEM 的结果显示，所形成金球粒子的半径约为 5nm。

依照标准的 DPD 模拟方法，Juan 等将 TTOE 分子分为如图 14-14 所




TTOE分子



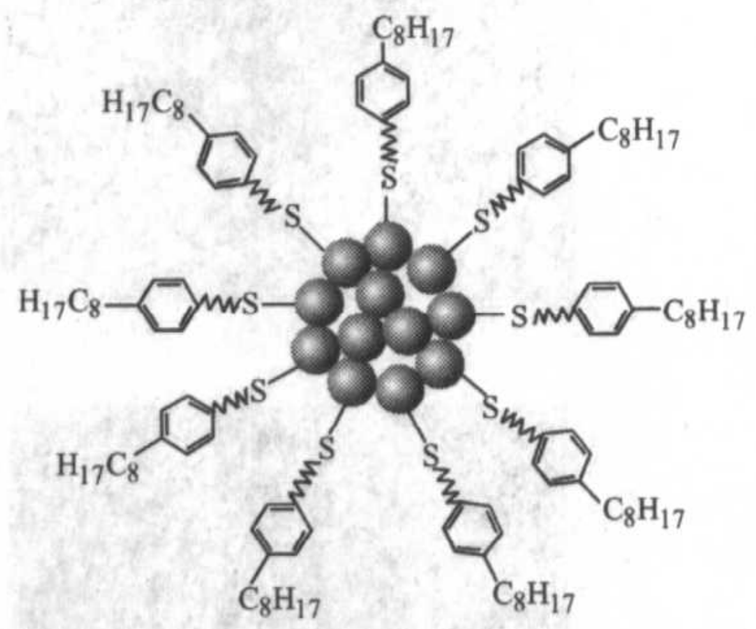




**图 14-13　穿透式电子显微镜所示 TTOE-Au 的球状粒子及结构示意**

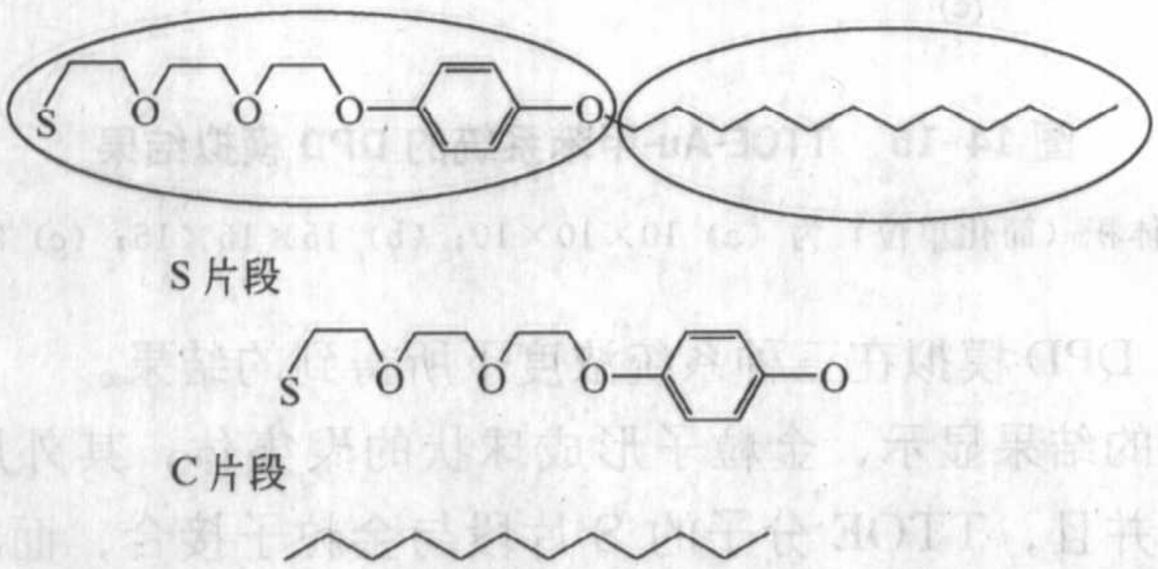
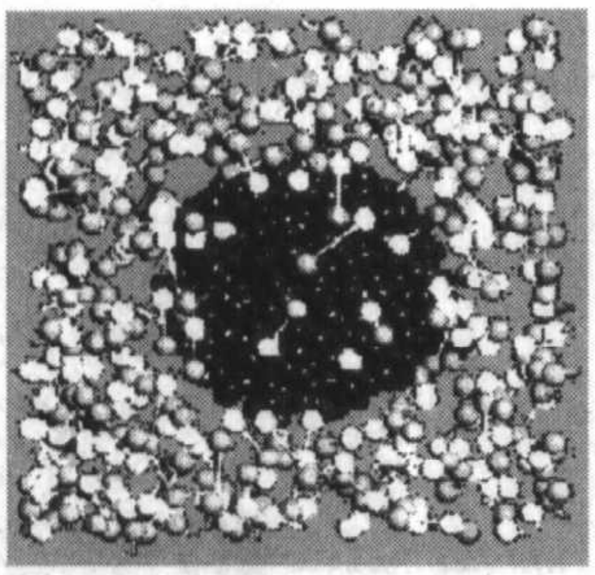


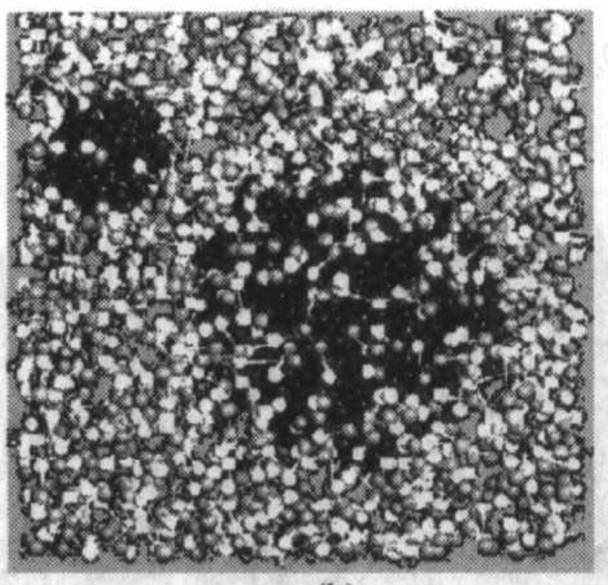

**图 14-14　TTOE 分子的 S 片段与 C 片段**

示含硫的 S 片段与烷链的 C 片段。

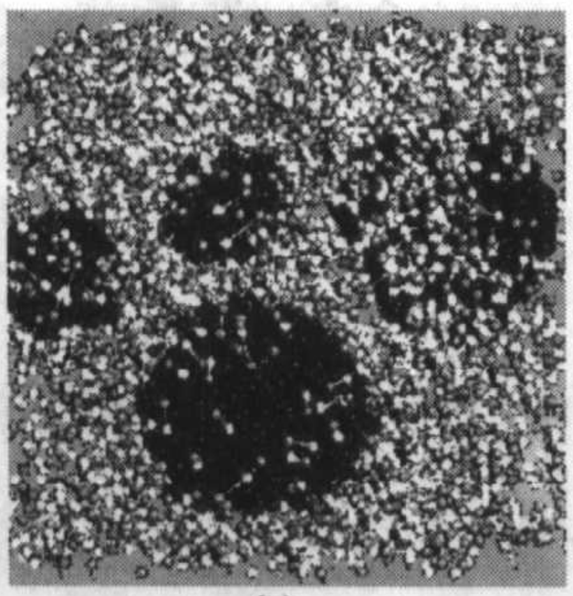
这两种片段的尺寸相近，均为 10Å 左右。将溶剂分子甲苯视为另一种粒子而将金原子视为第四种粒子。如果直接以此种分法执行 DPD 模拟，无法得到金粒子的球状聚集体，系统呈现层状的堆集。Juan 等认为，这模拟的误差是因为金原子的半径与其他粒子的半径相差太大所致。为了在模拟中调整金原子颗粒的大小，Juan 等将含有 4 个金原子的小团簇体(cluster) 视为一个金粒子。此团簇体的结构取自金的晶体结构(No. 225)，其晶格长度为 4.0783Å。这种选取小团簇体的方法，增大计算中金粒子的尺寸，与甲苯溶剂分子的尺寸相近。图 14-15 为利用这样的



(a)



(b)



(c)

**图 14-15　TTOE-Au-甲苯系统的 DPD 模拟结果**

模拟体系的体积（简化单位）为（a）10×10×10；（b）15×15×15；（c）20×20×20

粒子执行 DPD 模拟在三种系统浓度下所得到的结果。

模拟的结果显示，金粒子形成球状的聚集体，其外层为 TTOE 分子所包裹。并且，TTOE 分子的 S 片段与金粒子接合，而其 C 片段则暴露于球状堆集物的表面。表 14-5 为模拟所得的金球粒子的成分与结构。表

**表 14-5　甲苯溶剂中，模拟 TTOE-Au 球状纳米粒子的结构**

| 系统盒长<br>（简化单位） | 物质的量之比<br>（TTOE：Au：甲苯） | 纳米球状聚集物所含<br>金原子的个数 | 系统的结构 |
|---|---|---|---|
| 10×10×10 | 1：4：1.885 | | 层状 |
| | 1：4：2.166 | | 层状 |
| | 1：4：8 | 2060 | 椭圆体聚集 |
| 15×15×15 | 1：4：8 | 550(2nm) | 球状粒子 |
| | | 6170(5nm) | |
| 20×20×20 | 1：4：8 | 1360(3nm) | 球状粒子 |
| | | 1360(3nm) | |
| | | 6920(5nm) | |
| | | 6920(5nm) | |
| 20×20×20 | 1：4：0 | | 层状 |

中数据显示 DPD 方法所模拟的金球粒子半径为 2～5nm，与 TEM 实验所观测的结果一致。

模拟的结果显示溶液的成分与浓度亦为影响金球纳米粒子尺寸的重要因素。此计算实例成功地利用 DPD 方法模拟含有多种差异性成分复杂体系的粒子聚集行为，对包含有机分子与金属原子的复杂体系研究提供新的途径。

## 参 考 文 献


1 Groot R D, Warren P B. J. Chem. Phys., 107, 4423, 1997

2 Kuo M Y, Yang H C, Hua C Y, Chen C L, Mao S Z, Deng F, Wang H H, Du Y R. Chem Phys Chem, 5, 575, 2004

3 Ryikina E, Kuhn H, Rehage H, Muller F, Peggau J. Angew. Chem. Int. Ed, 41, 983, 2002

4 Li D W, Liu X Y, Feng Y P. J. Phys. Chem. B, 108, 11206, 2004

5 Juan S C C, Hua C Y, Chen C L, Sun X, Xi H. Molecular Simulation, 31, 277, 2005

# 第15章 溶解自由能

在许多溶液体系中，溶解自由能的大小与方向决定所形成溶液的浓度与稳定性。但利用分子动力学模拟的方法计算溶液的溶解自由能却往往遇到计算体系过于庞大的问题。系统的庞大主要是因为溶剂分子过多所导致。然而，若是将溶剂分子完全忽略，计算的性质与溶剂分子在真实溶液系统中的性质往往有明显的误差。针对此问题发展出了连续溶剂模型，计算中将溶剂分子的作用视为连续化效应，而将溶剂分子完全省略。利用这种模型，可研究巨大的溶质分子在溶剂中的各种特性，尤以生化分子在水溶液中的模拟最为重要。本章介绍 Born 与 Onsager 的连续溶剂模型及其他两种常用的简化模型。这些模型目前已经与分子动力学结合起来应用。

讨论溶液热力学性质，通常讨论溶解自由能（solvation free energy，$\Delta G_{sol}$），这种自由能为将一个分子由真空移入溶剂自由能的变化。溶解自由能由三个分量组成

$$\Delta G_{sol} = \Delta G_{elec} + \Delta G_{VDW} + \Delta G_{cav} \tag{15-1}$$

式中，$\Delta G_{elec}$为静电作用自由能的贡献；$\Delta G_{VDW}$为溶质分子与溶剂的范德瓦耳斯作用自由能；$\Delta G_{cav}$是在溶剂中形成溶质分子洞穴（cavity）所需的自由能。除了上述的 3 个分量外，亦可加入 $\Delta G_{hyb}$表示溶质分子与溶剂分子间固定的氢键作用。

## 15.1 Born 与 Onsager 模型

1920 年 Born 及 1936 年 Onsager 分别提出的理论为研究溶解效应最重要的贡献。图 15-1 为此二人所提出的模型示意图。Born 的模型为将电荷

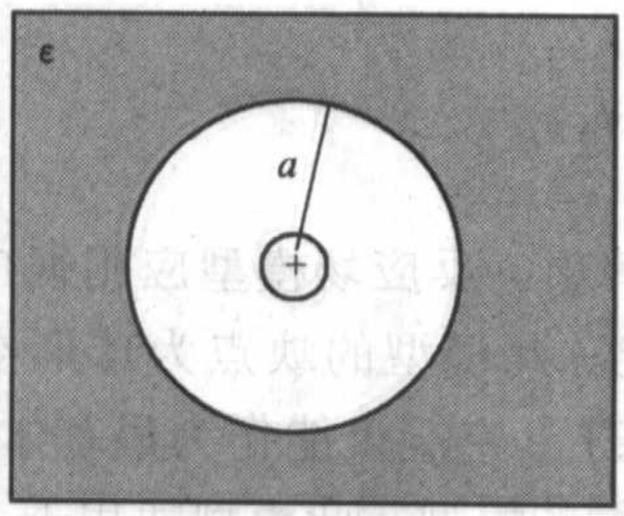



**图 15-1　Born 及 Onsager 的溶液模型**

置入球形的洞穴中，如图 15-1(a)。利用此简单的模型，他导出溶解自由能静电作用的贡献。Onsager 则将 Born 的理论推广至电偶极在球形的空穴中，如图 15-1(b) 所示。

Born 的模型中，离子的 $\Delta G_{elec}$ 等于将离子由真空移至介质所需要做的功。该功等于在两种环境下将离子带电所需做的静电功的差。若介质的介电常数为 $\epsilon$，则在介质中将离子带电所需做的功为 $q^2/2\epsilon a$，$q$ 为离子所带的电荷，$a$ 为空洞的半径。若于真空中，将离子带电所需做的功为 $q^2/2a$。依此，溶解自由能为：

$$\Delta G_{elec} = -\frac{q^2}{2a}\left(1-\frac{1}{\epsilon}\right) \quad (15\text{-}2)$$

式(15-2) 中引用简化单位，故 $4\pi\epsilon_0$ 设为 1。Born 的模型虽然简单但十分有效。模型中必须选择空穴的半径，通常依据晶体的离子半径。但对于碱离子的卤化物则发现若计算与实验吻合，则阳离子半径应较晶体的离子半径大 0.85Å，而阴离子半径则应大 0.1Å。Rashin 与 Honig 检视晶体的电子密度分布，并证实空洞的半径应选择晶体中阴离子的半径与阳离子的共价半径。与实验相合的最佳选择为将阳离子与阴离子的半径增加 7%。

Born 的模型仅适用于带电荷的物种，而 Onsager 的电偶极模型则适用于许多分子。事实上，Onsager 的模型只不过是 Kirkwood 模型的一个特例。Kirkwood 考虑空洞中任意电荷分布的物种，更为通用。空洞中溶质分子的电偶极于其周围介质诱发另一电偶极，以致在空洞中产生诱发电场（称为反应场，reaction field)。反应场与溶质分子电偶极的作用增加了系统的稳定性。Onsager 所推导的反应场大小（magnitude of reaction field）为：

$$\phi_{RF} = \frac{2(\epsilon-1)}{(2\epsilon+1)a^3}\mu \quad (15\text{-}3)$$

式中，$\mu$ 为溶质分子的偶极矩。依此，反应场与溶质间的作用能为 $-\phi_{RF}\mu$。除此之外，可极化的偶极应加上所需做的功，此项大小为 $\phi_{RF}\mu/2$，

类似先前在空洞中将离子带电的做功项。故由静电作用计算的溶解自由能为：

$$\Delta G_{elec} = -\frac{\phi_{RF}\mu}{2} \tag{15-4}$$

若物种带电，则应加上 Born 的电荷项。反应场模型应用很广，量子计算中亦引入此效应以估算溶解自由能。此模型的缺点为球形空洞的假设，因为多数分子的形状并非球形，故球形空洞仅能作为最基本的假设。除了球形空洞外，亦已发展出椭圆形空洞的模型，此模型适用于一些形状类似的分子。球形空洞与椭圆形空洞均需选择其空洞的大小，一般球形空洞的半径可由分子体积 $V_m$ 估算，即：

$$a^3 = 3V_m/4\pi N_A \tag{15-5}$$

式中，$N_A$ 为阿伏伽德罗常数（Avogadro's number）。文献上有一些估算空洞大小的方法，许多已被纳入一般的量子力学商业计算软件中。

## 15.2 连续模型

分子力学所引用的连续模型以通用的 Born 方程式（generalized Born equation，GB）为依据，以计算溶解自由能。设系统含有 $N$ 个粒子，则在介电常数为 $\epsilon$ 的介质中，其静电作用的总自由能为：

$$G_{elec} = \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{\epsilon r_{ij}} - \frac{1}{2}\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{15-6}$$

式中第一项为溶液中粒子间的库仑作用项。可将此项分为：

$$\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{\epsilon r_{ij}} = \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{r_{ij}} - \left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{r_{ij}} \tag{15-7}$$

将此代入式(15-6)，得：

$$G_{elec} = \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{r_{ij}} - \left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{r_{ij}} - \frac{1}{2}\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{15-8}$$

式中的第一项表示真空中的库仑作用项。依此，溶解的自由能可表示为 $G_{elec}$ 与真空的库仑作用的差，即：

$$\Delta G_{elec} = -\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{r_{ij}} - \frac{1}{2}\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{15-9}$$

此式称为通用的 Born 方程式（GB）。通常将上式的两项合写为：

$$\Delta G_{elec} = -\frac{1}{2}\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{f(r_{ij}, a_{ij})} \tag{15-10}$$

式中，$f$ 为粒子间距离 $r_{ij}$ 与 Born 离子半径 $a_i$ 的函数。Still 等建议使用如下的函数：

$$f(r_{ij},a_{ij})=\sqrt{(r_{ij}^2+a_{ij}^2e^{-D})}$$

$$a_{ij}=\sqrt{a_ia_j},D=\frac{r_{ij}^2}{(2a_{ij})^2} \tag{15-11}$$

式中，若 $i=j$，方程式(15-10) 回复最初的 Born 方程式；若两电荷的距离很近（意即形成偶极，$r_{ij}<a_i$ 或 $a_j$），则方程式近似于 Onsager 的结果；若两电荷相离很远（$r_{ij}\gg a_i$ 或 $a_j$）则结果近似于库仑作用与 Born 作用的和。并且可以分析式表示此方程式的微分，因此可将此溶解项包含于利用导数求能量极小值的方法，或是分子动力学的计算。

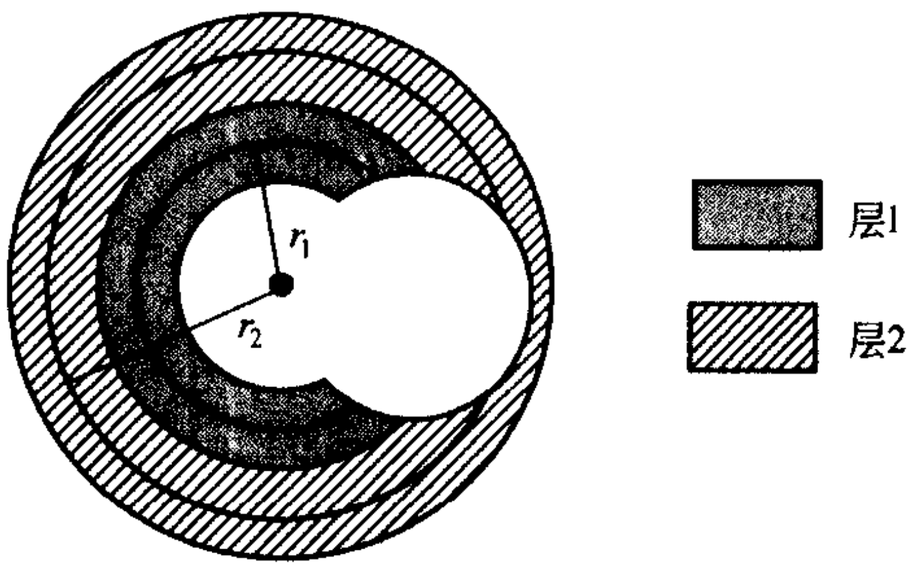



**图 15-2　原子外围球形层示意**

通常计算 $i$ 原子的静电作用能的方法为数值计算法。绕着原子建立一系列的球形层，直到最外层的球能包含整个分子的范德瓦耳斯半径，如图 15-2 所示。此例中，设共有 $M$ 个球层，球层中两个介电质（dielectric）的 Born 静电作用能为：

$$\Delta G_{elec}=-\frac{1}{2}\left(1-\frac{1}{\epsilon}\right)q_i^2\left\{\sum_{k=1}^{M}\frac{A_k}{4\pi r_k^2}\left[\left(\frac{1}{r_k-0.5T_k}\right)-\left(\frac{1}{r_k+0.5T_k}\right)\right]\right.$$

$$\left.+\frac{1}{r_{M+1}-0.5T_{M+1}}\right\} \tag{15-12}$$

式中，$r_k$ 为第 $k$ 层球的半径；$A_k$ 为未与分子范德瓦耳斯表面接触第 $k$ 层球的表面积，每层球的厚度 $T_k$ 随着与原子的距离而增加，即：

$$T_k=(1+F)T_{k-1} \tag{15-13}$$

$F$ 为展开系数。Still 等将 $F$ 及第一层球的厚度 $T_1$ 分别定为 0.5Å 及

1.0Å。式中的最后一项为与分子范德瓦耳斯表面外极性介电质的贡献。可令式(15-13) 与 Born 的原子方程式相等，而得 Born 有效半径 $a$，即：

$$\frac{1}{a_i}=\sum_{k=1}^{M}\frac{A_k}{4\pi r_k^2}\left[\left(\frac{1}{r_k-0.5T_k}\right)-\left(\frac{1}{r_k+0.5T_k}\right)\right]+\frac{1}{r_{M+1}-0.5T_{M+1}} \tag{15-14}$$

实际的分子动力计算中，Born 有效半径改变并不大，因此仅当分子周围非键结的分子数目表（list）改变时才需要重新计算。Still 所推导的式中，需要计算暴露在溶剂中球层表面积 $A_k$，可利用 Wodak 与 Janin 所发展的数值方法，即：

$$A_i=S_i\prod_j(1.0-b_{ij}/S_i) \tag{15-15}$$

式中，$S_i$ 为 $i$ 原子总可及表面积（total accessible surface area)。$i$ 原子的半径为 $r_i$，接触溶剂分子的半径为 $r_s$。将距离 $i$ 原子 $d_{ij}$ 的 $j$ 原子重合（overlap）时所需移除的表面积为 $b_{ij}$，即：

$$S_i=4\pi(r_i+r_s)^2$$

$$b_{ij}=\pi(r_i+r_s)(r_j+r_i-2r_s-d_{ij})\cdot[1.0+(r_i-r_j)/d_{ij}] \tag{15-16}$$

Still 利用如上的方法将溶剂的效应纳入分子力学的计算中，并以此计算分子于溶液中的结构及键结的相对自由能。例如计算 α-氨基酸的 D,L 旋光异构体（enantiomers)，如图 15-3 所示。

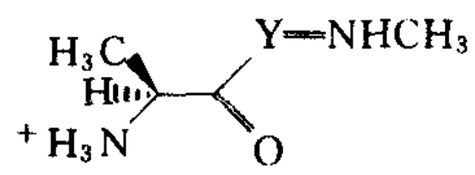



**图 15-3　α-氨基酸的构造式**

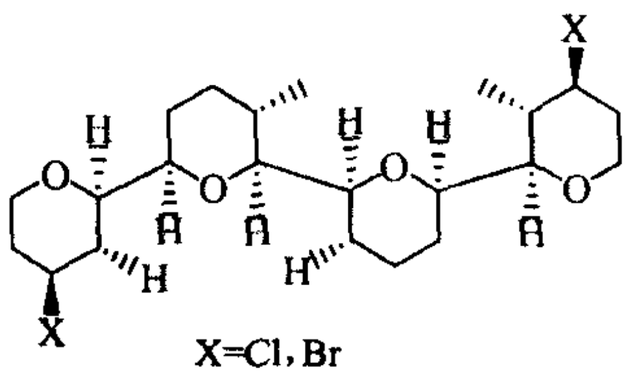



**图 15-4　Podand 多离子体衍生物**

在氯仿溶剂中与 Podand 多离子体（ionphore）键结的相对自由能，得到与实验非常吻合的结果。

Still 以类似的计算方法预测 α-氨基酸的对掌异构物对图 15-4 及图 15-5 所示化合物的光学选择性。

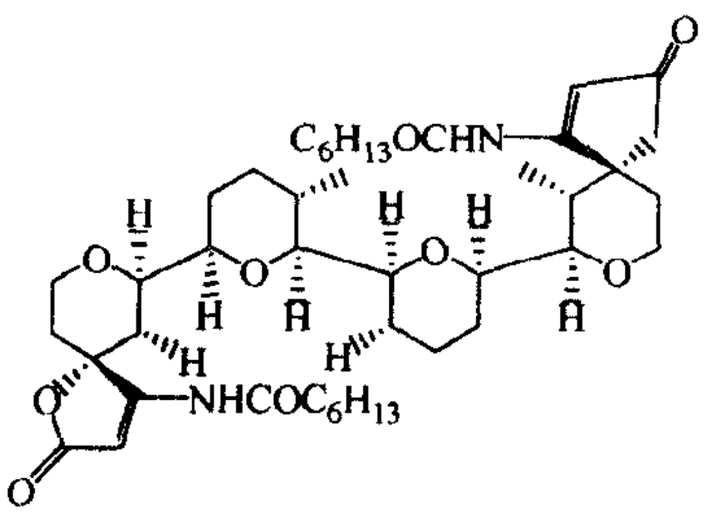



**图 15-5　Podand 多离子体衍生物**

除了正确地预测出旋光异构体的选择性外，计算的键结亲合能（binding affinity）与实验的度量值非常接近，误差仅为 0.3kcal/mol。由此可见

Born 连续溶剂模型的实用性。执行这样系统的分子动力计算，通常至少需要 10ns，如果计算不引用连续溶剂模型，则需耗费非常大量的计算时间。

## 15.3 朗之万偶极矩模型

Warshel 与 Levitt 所发展的朗之万偶极矩模型介于连续溶剂模型与真实溶剂分子计算之间。此模型绕着溶质分子建立一个区域，此区域可为任意形状，如图 15-6 所示。若溶质分子为高分子，则一般以溶剂的可及表面积（accessible surface）为此区域的边界。

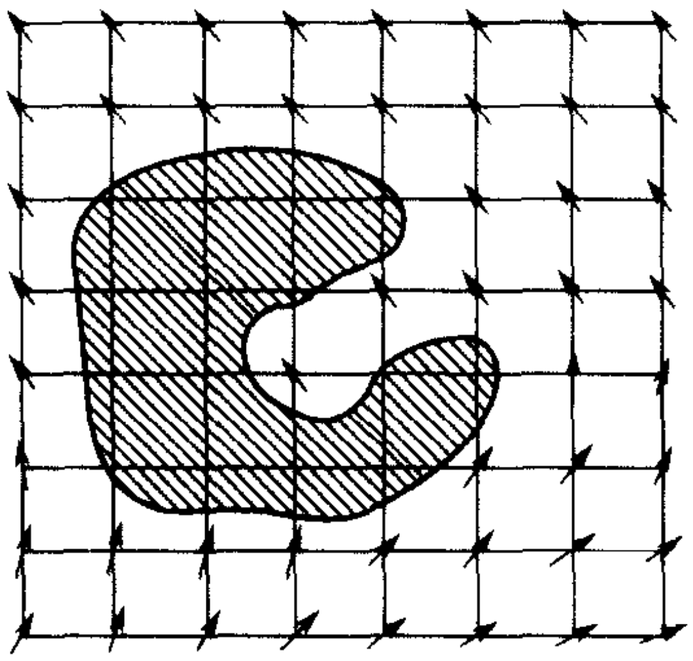


**图 15-6 朗之万偶极矩模型示意**

图 15-6 中，将此区域外的空间分成许多格子点，每一格子点上定义一可以转动的偶极矩。这些偶极矩表示溶剂分子的偶极矩，其相互间的距离依溶剂分子间的距离而定。溶质分子与其他溶剂分子的偶极造成每一偶极位置的电场 $\vec{E}_i$。利用朗之万方程式可决定每一偶极的大小与方向，即：

$$\vec{\mu}_i=\mu_0\frac{\vec{E}_i}{|\vec{E}_i|}\left[\frac{\exp(C\mu_0|\vec{E}_i|/k_BT)+\exp(-C\mu_0|\vec{E}_i|/k_BT)}{\exp(C\mu_0|\vec{E}_i|/k_BT)-\exp(-C\mu_0|\vec{E}_i|/k_BT)}-\frac{1}{C\mu_0|\vec{E}_i|/k_BT}\right] \tag{15-17}$$

式中，$\mu_0$ 是溶剂分子的偶极矩的大小；$C$ 为反映电偶极抵抗重排倾向的参数，其大小或可由其他单独对溶剂的分子动力模拟决定。朗之万偶极的自由能为：

$$\Delta G_{sol}=-\frac{1}{2}\sum_i\vec{\mu}_i\cdot\vec{E}_i^0 \tag{15-18}$$

式中，$\vec{E}_i^0$ 为只由溶质分子的电荷所形成的电场。Warshel 等利用此模式计算许多酶的催化反应，获得成功。

## 15.4 泊松-玻耳兹曼方程式

对于一些生化大分子如蛋白质及 DNA 等的系统，常用以计算静电作用的方法是依据泊松-玻耳兹曼方程式所推导的。这些方法中，将溶质分子视为低介电常数（通常介于 2～4 间）的物种，而将溶剂视为连续的高介电常数介质。泊松（Poisson）方程式描述介电常数为 $\epsilon$，均匀分布的介质中电势 $\phi$ 的变化与电荷密度的关系，即：

$$\nabla^2\phi(\vec{r})=-\frac{\rho(\vec{r})}{\epsilon_0\epsilon} \tag{15-19}$$

将式(15-19) 以简化单位表示，即设 $4\pi\epsilon_0=1$，为：

$$\nabla^2\phi(\vec{r})=-\frac{4\pi\rho(\vec{r})}{\epsilon} \tag{15-20}$$

式中，电荷密度的单位为 $C/m^3$。对于固定介电常数值的介质中的点电荷，Poisson 方程式与库仑静电作用式相同。若介质的介电常数不为常数，而随位置而改变，则库仑作用式不再适用；此时 Poisson 方程式为：

$$\nabla\cdot\epsilon(\vec{r})\nabla\phi(\vec{r})=-4\pi\rho(\vec{r}) \tag{15-21}$$

若系统中具有可移动的离子，则 Poisson 方程式必须加上离子受到电势作用后在溶液中重新分布的修正项。离子的重新分布依据 Boltzmann 方程式为：

$$n(\vec{r})=N\exp[-U(\vec{r})/k_BT] \tag{15-22}$$

式中，$n(\vec{r})$ 为于位置 $\vec{r}$ 离子的数目密度（number density）；$N$ 为系统中离子的数目密度；$U(\vec{r})$ 为自无穷远处将离子移至位置 $\vec{r}$ 的能量变化。将这些效应加入 Poisson 方程式中，则得泊松-玻耳兹曼方程式，即：

$$\nabla\cdot\epsilon(\vec{r})\nabla\phi(\vec{r})-\kappa'\sinh[\phi(\vec{r})]=-4\pi\rho(\vec{r}) \tag{15-23}$$

式中，$\kappa'$ 反比于德拜-休克尔长度（Debye-Hückel length）$\kappa$，即：

$$\kappa^2=\frac{\kappa'^2}{\epsilon}=\frac{8\pi N_Ae^2I}{1000\epsilon k_BT} \tag{15-24}$$

式中，$e$ 为电荷；$I$ 为溶液的离子强度（ionic strength）；$N_A$ 为阿伏伽德罗常数。将式(15-23) 中的双曲线正弦函数（hyperbolic sine function）展开，得：

$$\nabla\cdot\epsilon(\vec{r})\nabla\phi(\vec{r})-\kappa'\phi(\vec{r})\left[1+\frac{\phi(\vec{r})^2}{6}+\frac{\phi(\vec{r})^4}{120}+\cdots\right]=-4\pi\rho(\vec{r}) \tag{15-25}$$

取上展开式中的第一项，得线性泊松-玻耳兹曼方程式，即：

$$\nabla\cdot\epsilon(\vec{r})\nabla\phi(\vec{r})-\kappa'\phi(\vec{r})=-4\pi\rho(\vec{r}) \tag{15-26}$$

式(15-26) 可利用边界元素（boundary element）与有限差分（finite element）的数值方法求解。此处介绍 Warwicker 与 Watson 所发展的有限差分解法，称为泊松-玻耳兹曼有限差分方法（finite difference Possion-Boltzmann，FDPB)。

如图 15-7，为一被边长 $h$ 的立方体所围绕的格子，格子中的电荷为 $q_0$。

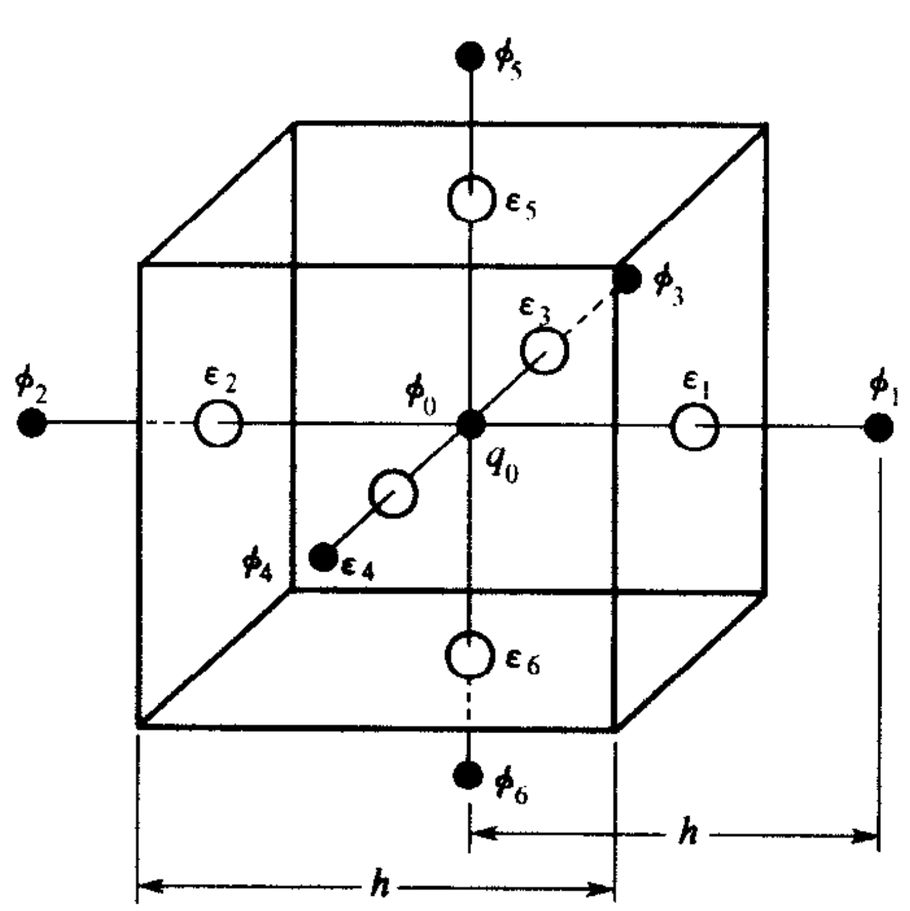



**图 15-7　FDPB 方法的模型示意**

此立方体内相当于一均匀电荷密度（$q_0/h^3$）的系统，$\rho_0 = q_0/h^3$。格子点上的势能为：

$$\phi_0 = \frac{\sum \varepsilon_i \phi_i + 4\pi \dfrac{q_0}{h}}{\sum \varepsilon_i + \kappa'^2_0 f(\phi_0)} \tag{15-27}$$

式中，势能为周围邻近 6 个格子点位能 $\phi_i$ 的和；介电常数 $\varepsilon_i$ 对应于格子点连接的中点。线性的泊松-玻耳兹曼方程式，$f(\phi_0)=1$；若为非线性的泊松-玻耳兹曼方程式，则 $f(\phi_0)=(1+\phi_0^2/6+\phi_0^4/120+\cdots)$。$\kappa'^2$ 可由格子点的离子强度计算。格子点的位能于其邻近格子点的位能相关，故重复计算许多的格子点可得到收敛的格子点位能。

执行泊松-玻耳兹曼计算，必须决定每一格子点的介电常数，因而必须分辨属于溶质或属于溶剂的格子点。通常以分子的表面积或是可及表面积（accessible surface）决定溶质与溶剂的分界（boundary)。所有溶质边界外的格子点均为高介电常数（如水，介电常数为 80）的点，与固定的离子强度。边界内格子点的介电常数，则定为高分子的低介电常数值

(2～4)。通常高分子的介电常数与其自身的极化性（polarisability）与在可变电场下重排内在的偶极的能力相关。若分子为刚性，则不可能重排其内在的偶极（带电基团），因此其介电常数仅与其极化性有关。通常有机液体由其极化性所得的介电常数值约为 2。若分子为非刚性分子，因其位向改变所导致的偶极重排，而使介电常数值改变为 4。晶格点的取法影响泊松-玻耳兹曼计算的准确性。晶格点取的愈密集，计算的结果愈准确，相对的所需的计算时间也愈长。通常的计算将系统分为 $65^3$ 个晶格点。

利用 FDPB 方法可用以获得许多有用的信息。除了获得各点的电势数值外，可利用画图的方法画出分子周围的电势。图 15-8 为 FDPB 方法计算所得胰岛素（trypsin）分子的电势图，图中颜色的深浅表示电势的高低。

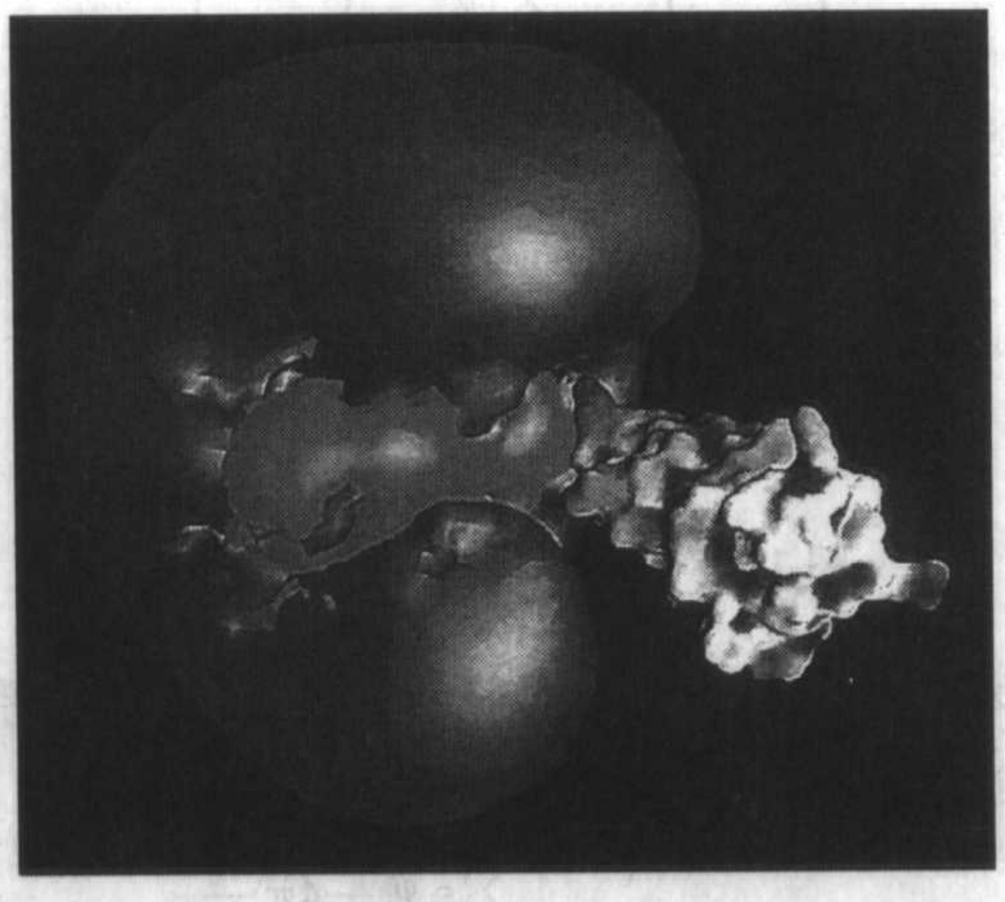


**图 15-8　FDPB 方法计算所得 trypsin 分子的电势图**

此酶分子可以与 trypsin 紧密接合，但是两者均为带正电的物种，因此必须厘清两者的接合方式。若以均匀分布的介电常数 80 计算 trypsin 周围的电势，则必然得到周围电势全为正值的结果。但是，若考虑了边界的介电性，则计算的电势呈现部分负的区间，如图 15-8 所示。这部分负电势的区间极为可接合处。

利用有限差分泊松-玻耳兹曼方法可以计算静电作用对例如溶剂化（solvation）或形成分子间的复合物等过程的影响。静电作用的溶剂化自由能为将粒子自真空移入溶剂中静电作用能的差，若粒子带电量为 $q_i$，溶液中的电势为 $\phi_i$，则静电作用能为 $q_i\phi_i$。计算溶剂化自由能 $\Delta G_{elec}$ 可分别

计算两个系统，其中一个将溶剂的介电常数定为 80（水溶液），另外一个将介电常数定为 1（真空），则：

$$\Delta G_{\text{elec}} = \frac{1}{2} \sum_i q_i (\phi_i^{80} - \phi_i^1) \tag{15-28}$$

式中，自由能为所有溶质分子电荷项的和。

计算两分子结合的自由能，通常假设两分子有相同的分子内介电常数（$\epsilon_m$），可将其过程分为一系列的步骤执行，如图 15-9 所示。

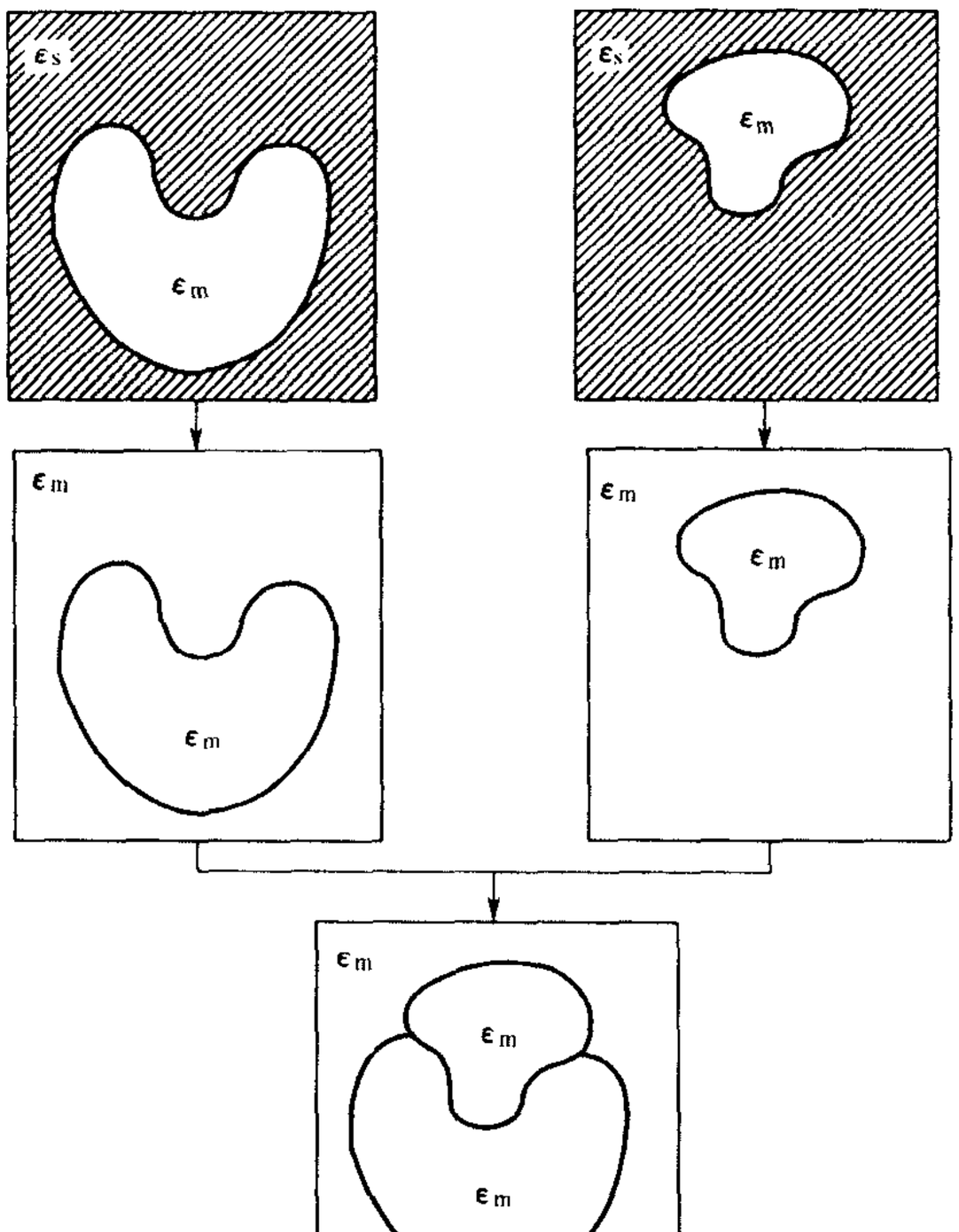



**图 15-9　两分子结合自由能计算示意**

按图中所示，第一步为计算将两个物种分别由溶剂（$\epsilon_s$）移至介质（$\epsilon_m$）中的自由能，第二步为依照库仑定律计算两物种在介质（$\epsilon_m$）中的自由能，第三步为计算复合物由介质（$\epsilon_m$）移至溶剂（$\epsilon_s$）中的自由能。

除了静电作用外，中性或低极性的溶质分子非静电作用的溶剂化自由能一般可综合写为：

$$\Delta G_{cav} + \Delta G_{vdw} = \gamma A + b \tag{15-29}$$

式中，$A$ 为溶剂的总可容面积（accessible area）；$\gamma$ 与 $b$ 为常数。此二常数通常可由一些实验的数据定出。除了以上的形式外，Claverie 等建议空洞项为：

$$\Delta G_{cav} = K_0 + K_1 a_{12} + K_2 a_{12}^2 \tag{15-30}$$

式中，常数 $K$ 与溶剂分子的体积相关；$a_{12}$ 为溶剂分子与溶质分子的平均直径。

欲准确的计算溶解自由能不是一件容易的事，目前仍不断的有人提出新的理论。针对不同的系统，往往需要不同的计算模型。虽然结合量子力学的计算方法可以得到较准确的结果，但是其适用范围只限于非常简单的体系。如何以简单的计算模型得到满意的结果有待进一步的努力。

## 参 考 文 献


1 Born M. Zeitschrift für Physik, 1, 45, 1920

2 Onsager L. J. Amer. Chem. Soc., 58, 1486, 1936

3 Rashin A A, Honig B. J. Phys. Chem., 89, 5588, 1985

4 Still W C, Tempczyrk A, Hawley R C, Hendrickson T. J. Amer. Chem. Soc., 112, 6127, 1990

5 Wodak S J, Janin J. Proc. Nat. Acad. Sci. USA 77, 1736, 1980

6 Warshel A. Computer Modeling of Chemical Reactions in Enzymes and Solutions. New York, John Wiley, 1991

7 Warshel A, Levitt M. J. Mol. Bio., 103, 227, 1976

8 Warwicker J, Watson H C. J. Mol. Bio., 157, 671, 1982

9 Claverie P, Daudey J P, Langlet J, Pullman B, Piazzola D, Huron M J. J. Phys. Chem., 82 405, 1978

# 第16章 生化分子系统的模拟

利用分子模拟方法研究生化分子体系，目前尚在起步阶段，以现今的计算机计算速度与容量模拟研究生化分子体系仍远远不能满足实际的需要，计算的结果与实际体系相比较仍有相当的误差存在。本章介绍几个以分子动力模拟方法研究生化体系的实际计算例子，这些例子是分子模拟计算的综合应用。例子包括：蛋白质分子开折的模拟、细胞膜的融合模拟以及生化分子体系中疏水基团堆积的研究。

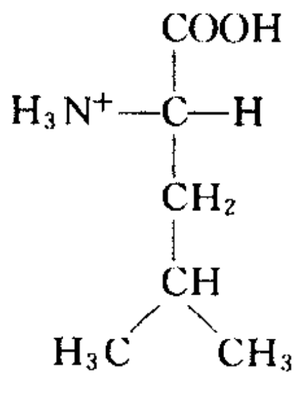
## 16.1 生化分子的结构

生化分子系统的模拟是化学计算中最具挑战性的课题。由于一般的生化分子自身拥有大量的原子数及模拟中必须考虑多量的水分子，并且重要的生物反应动力过程往往为纳秒或微秒以上的缓慢运动，因此模拟这么庞大的系统并非易事。组成生物分子的最基本单位为氨基酸（amino acid），图 16-1 列出 20 种天然氨基酸的结构式与各氨基酸的极性、酸碱性及通用简称。

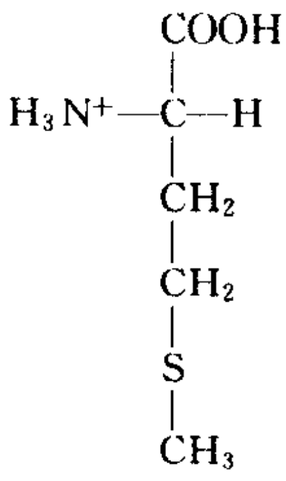
生物分子中氨基酸间脱去一分子水以肽键（peptide bond）互相连接，如图 16-2 所示，脱去水分子的氨基酸单位称为残基（residue）。

所有天然的氨基酸其 $C_\alpha$ 原子的光学性质均相同。由 2 个氨基酸经肽键连接聚成的分子称为二肽（dipeptide），由 3 个氨基酸聚成的分子称为三肽（tripeptide），由 3～10 个氨基酸聚成的分子称为寡肽（oligopeptide），由多个氨基酸聚成的聚合物则称为多肽（polypeptide）。蛋白质（protein）通常由一个或多个的多肽分子所组成，组成的多肽一般包括 40～4000 个氨基酸残基，因此蛋白质的分子量的范围为（4～440）千道尔顿（kD）。

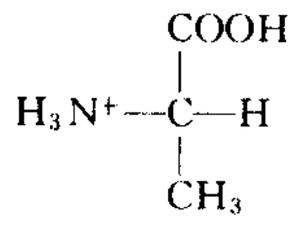
① 非极性-疏水性



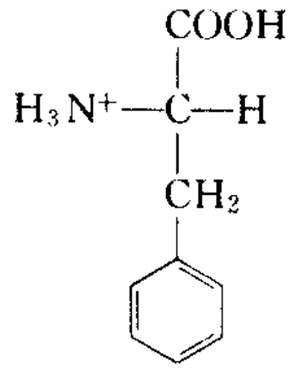
**亮氨酸**
leucine
(Leu, L)



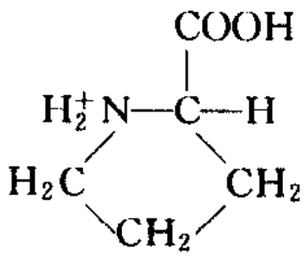
**甲硫氨酸**
methionine
(Met, M)



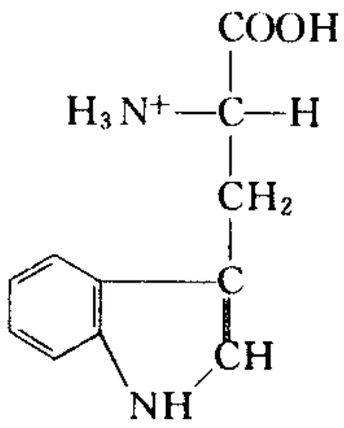
**丙氨酸**
alanine
(Ala, A)



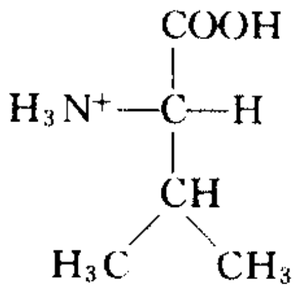
**苯丙氨酸**
phenylalanine
(Phe, F)



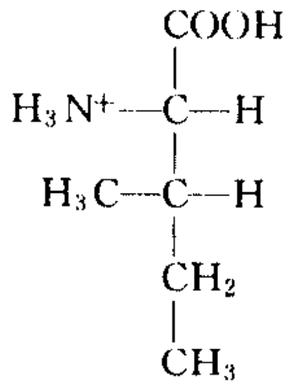
**脯氨酸**
proline
(Pro, P)



**色氨酸**
tryptophan
(Trp, W)

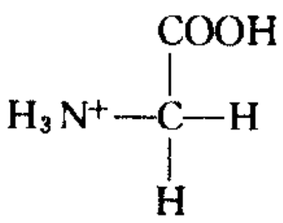


**缬氨酸**
valine
(Val, V)

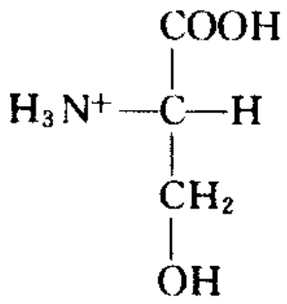


**异亮氨酸**
isoleucine
(Ile, I)

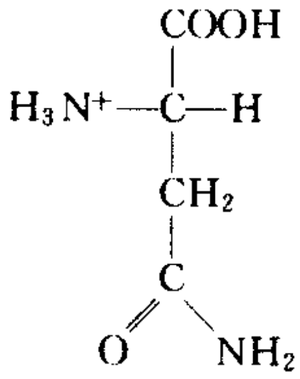
② 极性，无电性，亲水性



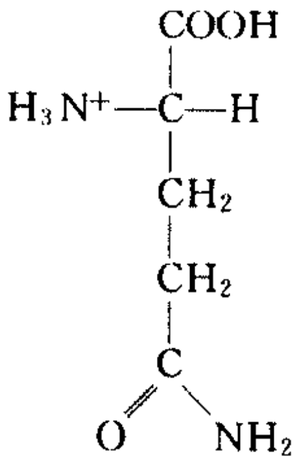
**甘氨酸**
glycine
(Gly, G)



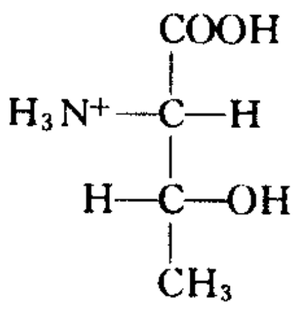
**丝氨酸**
serine
(Ser, S)



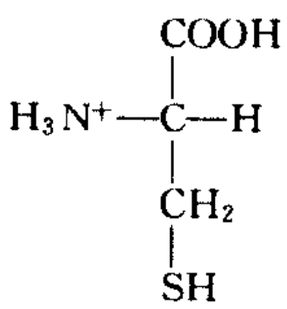
**天冬酰胺**
aspargine
(Asn, N)



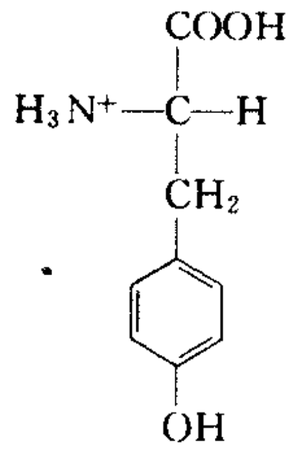
**谷氨酰胺**
glutamine
(Gln, Q)



**苏氨酸**
threonine
(Thr, T)

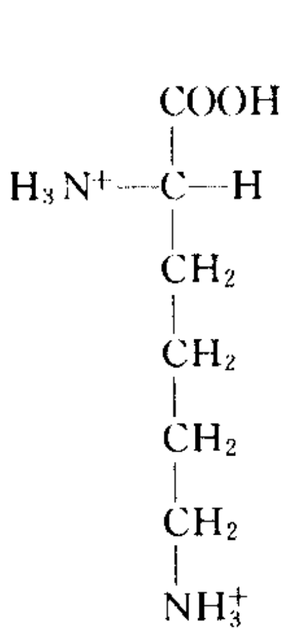


**半胱氨酸**
cysteine
(Cys, C)



**酪氨酸**
tyrosine
(Tyr, Y)

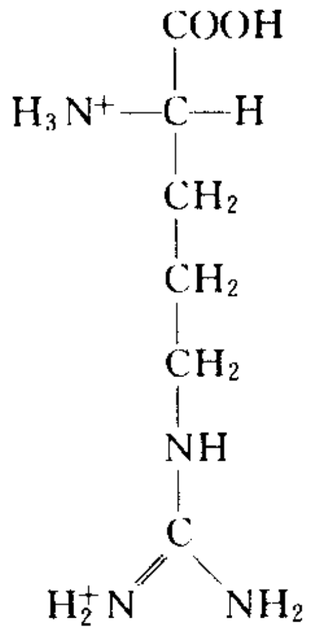
③ 极性-碱性-亲水性



**赖氨酸**

lysine

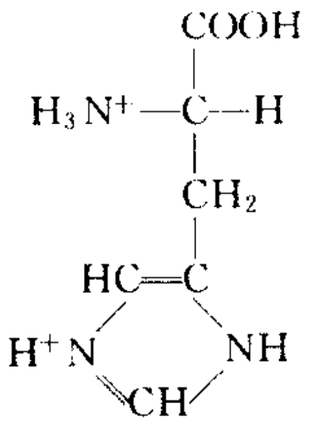
(Lys，K)



**精氨酸**

arginine

(Arg，R)

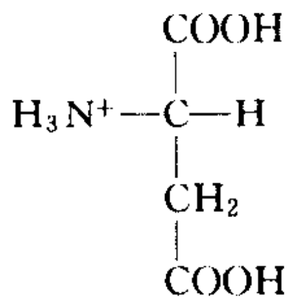


**组氨酸**

histidine

(His，H)

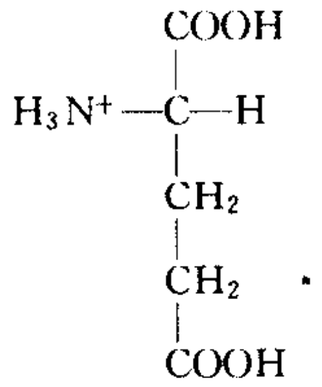
④ 极性-酸性-亲水性



**天冬氨酸**

aspartic acid

(Asp，D)



**谷氨酸**

glutamic acid

(Glu，E)

**图 16-1　天然氨基酸**

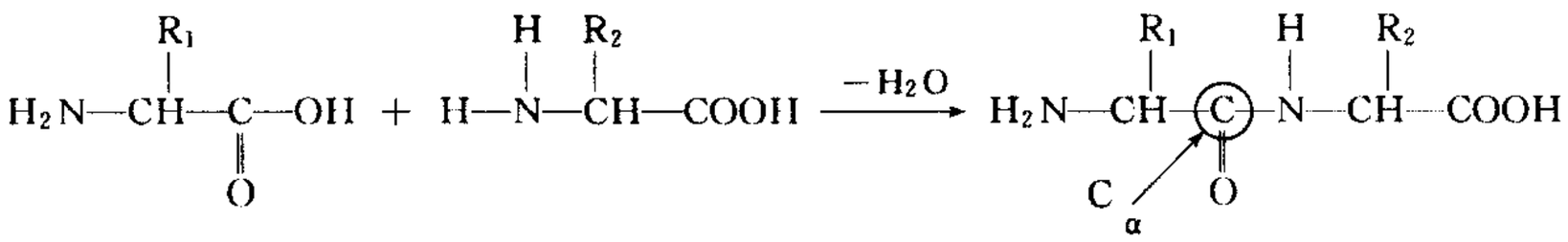


**图 16-2　氨基酸间的肽键**

蛋白质水解后生成氨基酸或其衍生物的蛋白质。蛋白质与多肽分子因其组成的氨基酸分子侧基的形状、大小、电荷分布与形成氢键的能力不同，以致造成不同的生化特性。构成蛋白质的氨基酸排列顺序是一定的，称为氨基酸序列（amino acid sequence)。例如，胰岛素（insulin）由两条分别称为 A 链及 B 链的多肽链所组成，人类胰岛素链的 A 链由 21 个氨基酸依固定的排列顺序所组成；B 链则由 30 个氨基酸依固定的排列顺序所组成。

A 链与 B 链间在固定的位置以双硫键互相衔接，而形成一定的氨基酸顺序结构。这种结构称为蛋白质的一级结构，如图 16-3 所示。

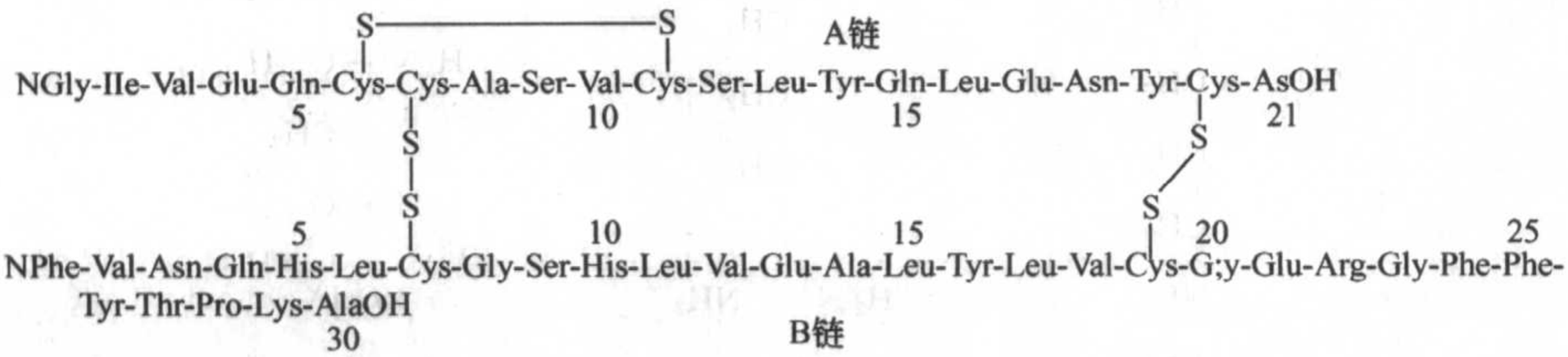



**图 16-3　胰岛素的一级结构**

多肽与蛋白质分子的生化特性与分子的结构有密不可分的关系。此处结构除了组成的氨基酸序列外，还包括分子三度空间的立体排列方式。通常可利用结晶 X 射线衍射或核磁共振方法以决定分子的结构。但以实验方法完全决定多肽或蛋白质分子的立体结构依然非常困难，需要耗费许多的时间。决定结构的速率则远不及发现新的多肽或蛋白质分子的速率。图 16-4 显示生化分子中主链的扭转角。

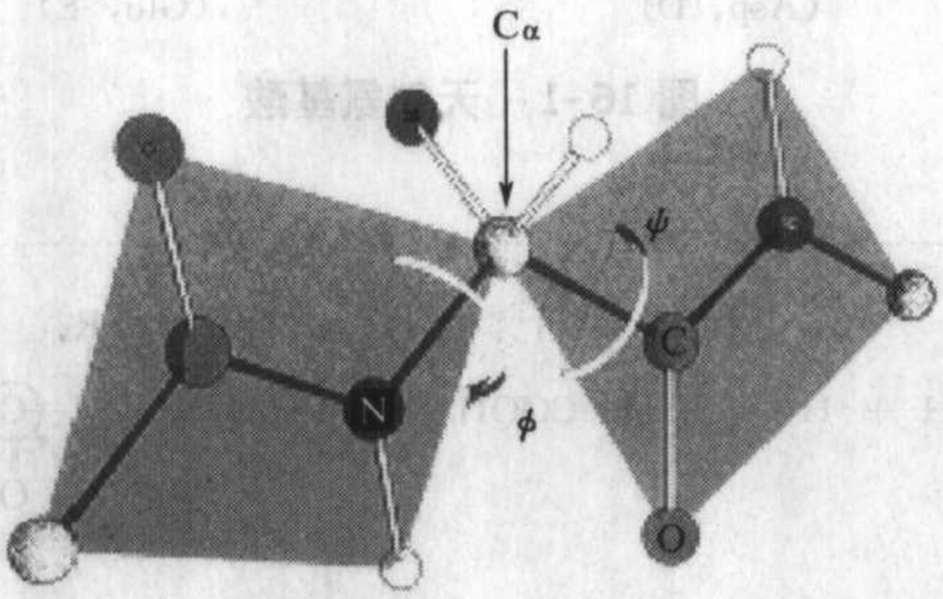



**图 16-4　生化分子中的扭转角**

图中，灰色的平面由 CO—NH 所构成，称为肽平面。通常定联结 $\alpha$ 碳原子的扭转角为 $(\psi,\phi)$。氨基酸间的联结（NH—CO）除脯氨酸（proline）外多为反式形态（扭转角约 180°），脯氨酸则多为顺式结构。生化分子主链的构形基本上由一系列的 $(\psi,\phi)$ 所决定。

实验显示蛋白质分子的结构非常不规则，并无一定的对称性。但是往

往可发现蛋白质分子存在一些常见的结构形状，其中最常见结构形式为 α-螺旋形（α-helical），及 β-股（β-strand）结构形。这两种结构形式称为蛋白质的二级构造。图 16-5 显示 α-螺旋结构的蛋白质。

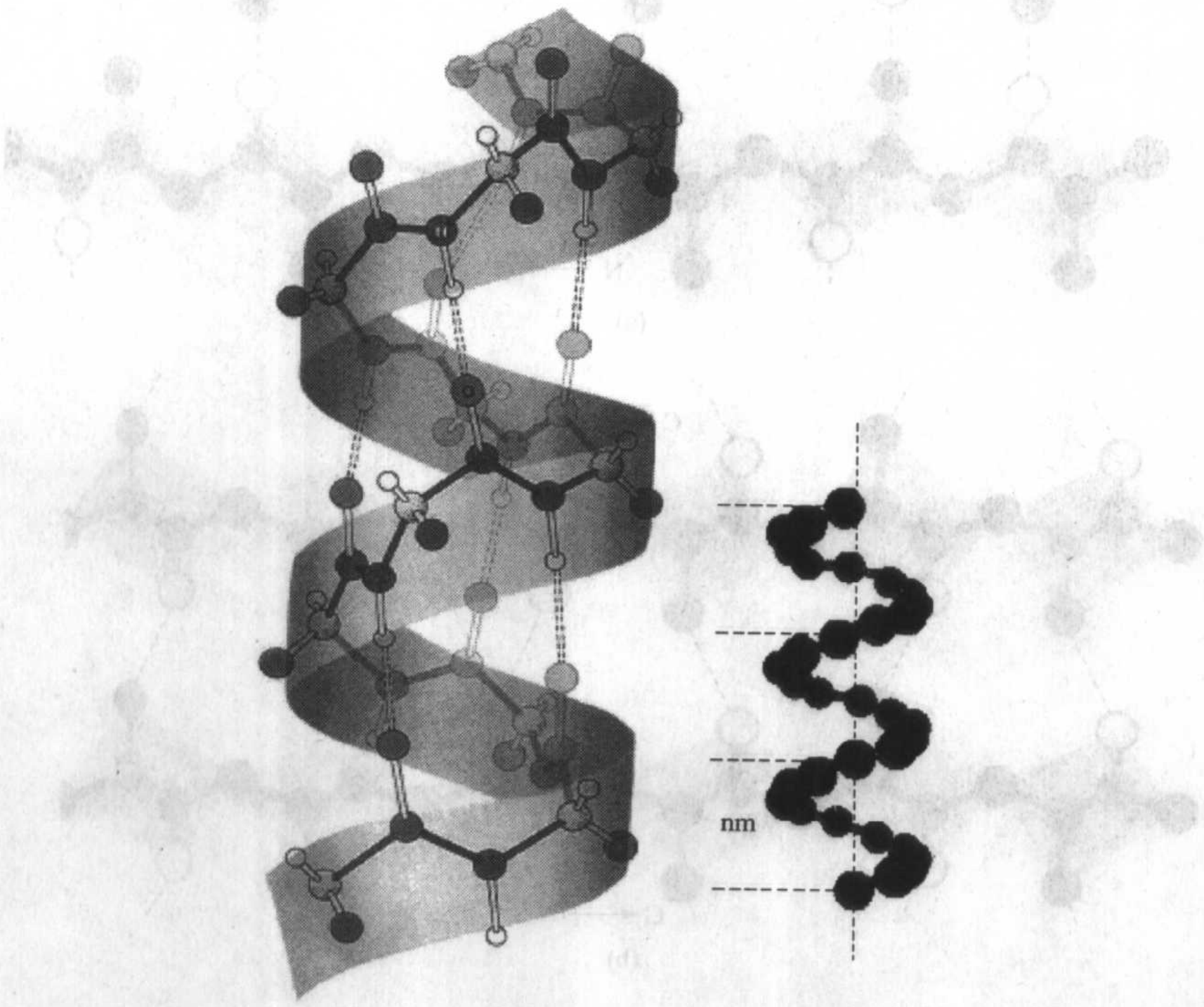



**图 16-5　α-螺旋结构的蛋白质**

通常 β-股构造的蛋白质间的氢键互相作用而形成平行或反平行的层状结构，称为 β-折叠（β-sheet）。图 16-6 显示两种 β-折叠的形式，平行板块的 β-股互相平行，均朝同一方向；而反平行板块的 β-股其 β-股位向相反，NH 基团与 CO 基团相对，形成氢键。

常见的蛋白质分子的立体结构多为纤维状或球状，这样的结构称为蛋白质的三级结构。纤维状蛋白质的多肽链不仅有 α-螺旋构造，还存在折曲扭转的构造，在生成折曲扭转处，为难以形成 α-螺旋构造的氨基酸序列，这些氨基酸为脯氨酸、丝氨酸与异亮氨酸等。球形蛋白质是以不规则的扭曲旋转形成一个球，例如肌红素（myoglobin）。肌红素中有 8 处为 α-螺旋的构造，在扭曲转折处为脯氨酸等难以形成 α-螺旋构造的氨基酸，其三级

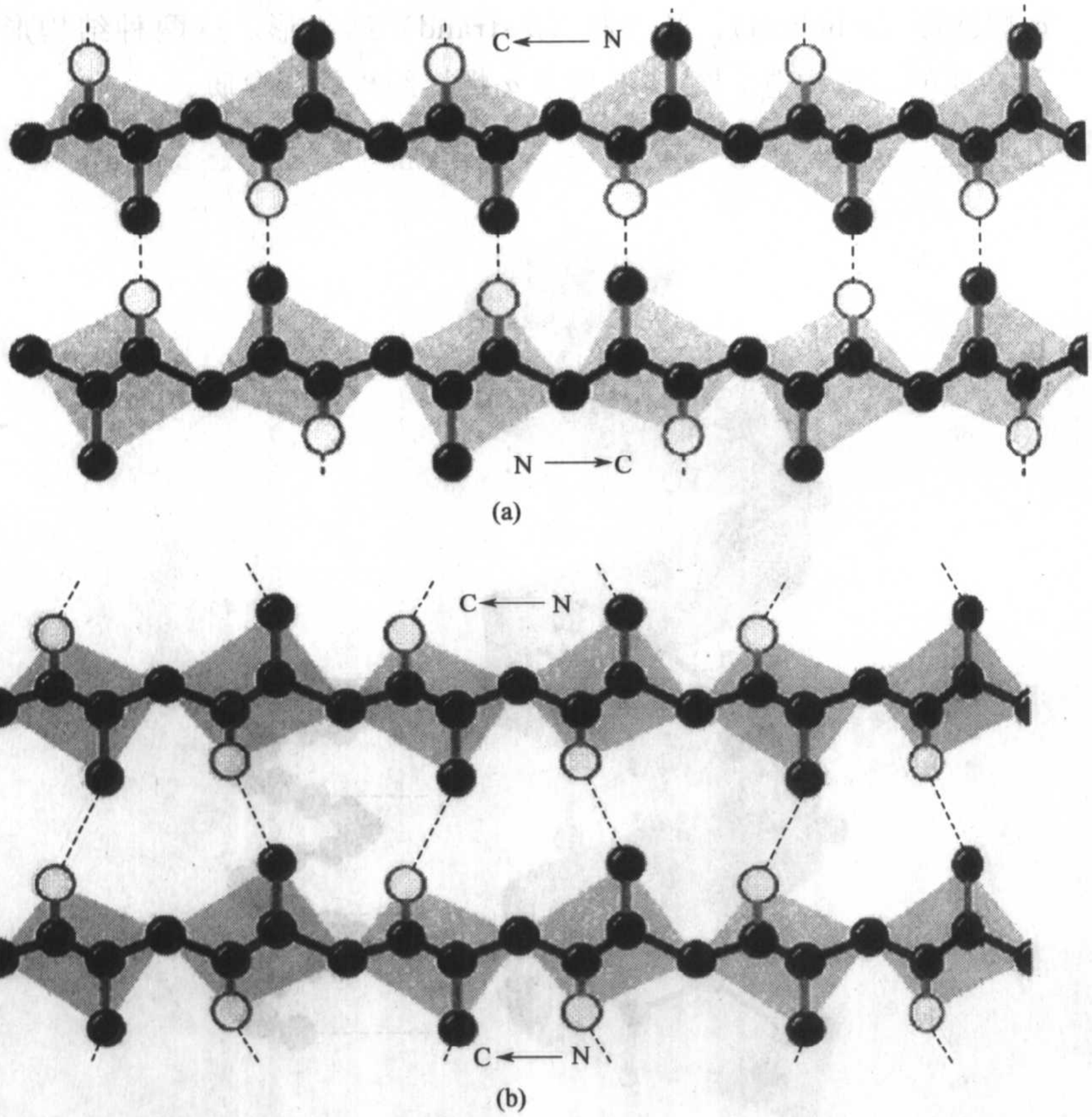



**图 16-6　β-折叠板块的形式**

（a）平行板块；（b）反平行板块

结构如图 16-7 所示。

除了蛋白质的三级结构外，还有四级结构。例如许多多肽是由 2 条或更多的三级结构所组成的。组成四级结构的多肽链称为亚单位（subunit）。以血红素（hemoglobin）为例，此蛋白质由 2 条 $\alpha$ 链与 2 条 $\beta$ 链所组成的四倍体。$\alpha$ 链由 141 个氨基酸所组成，$\beta$ 链则由 146 个氨基酸所组成。血红素的四级结构如图 16-8 所示。

研究生化分子的构造，可以了解形成各种结构的原因，从而理解生化分子的活性。以能量的观点，形成不同构造的生化分子是由于形成后的结构为低能量的稳定结构。但因为通常的生化分子组成复杂，欲建立明确的

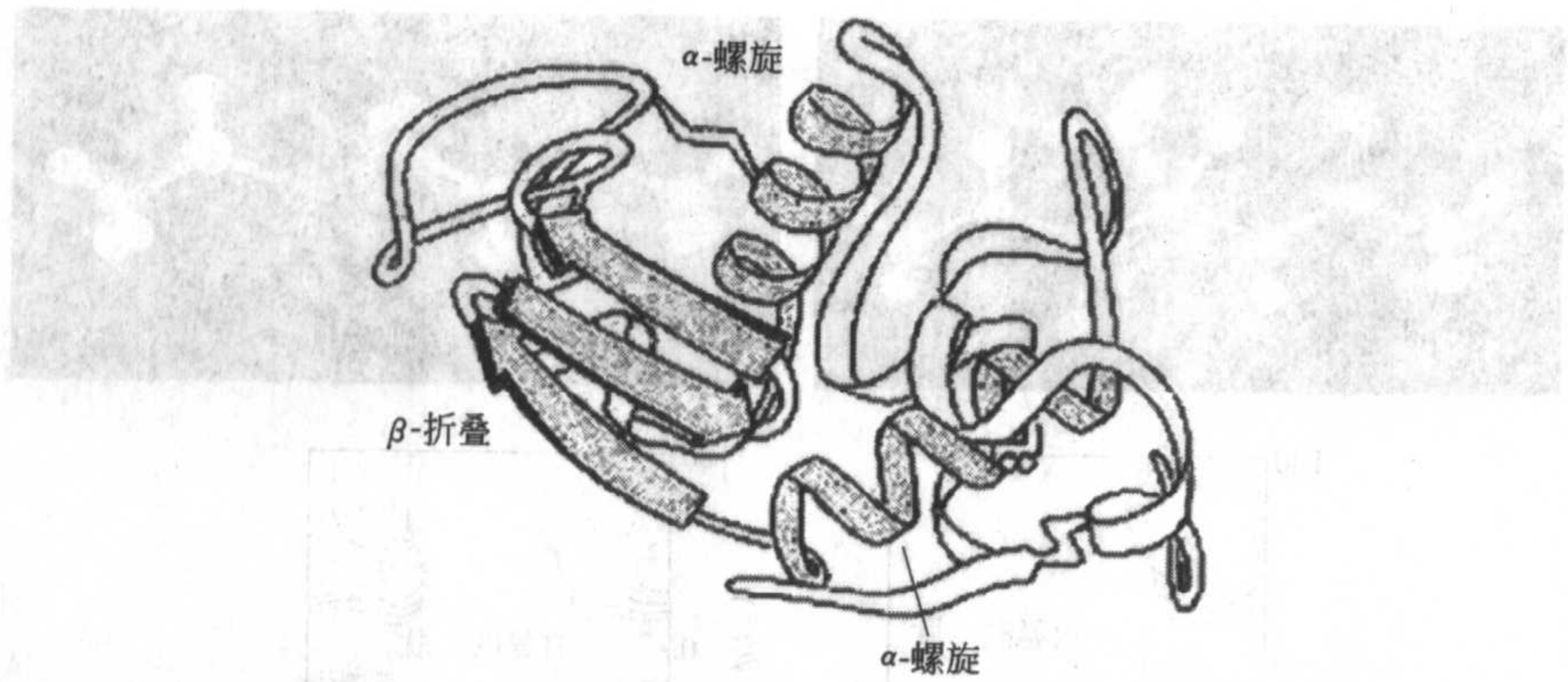



图 16-7　肌红素的三级结构

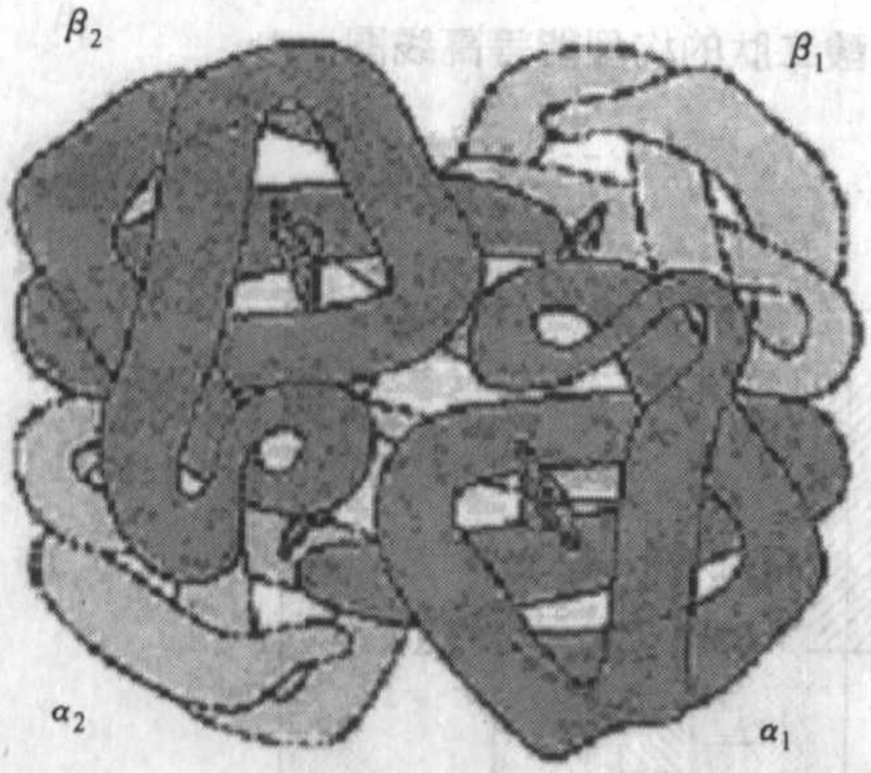



图 16-8　血红素的四级结构

能量与构型关系并非易事。简单的生化分子的构型可由适当的构型能等高线图判断。图 16-9 为两种简单的二肽分子：丙氨酸二肽（alanine dipeptide）与甘氨酸二肽（glycine dipeptide）的构型能等高线图。

图 16-9 根据二面角 $\phi$ 与 $\psi$ 的变化，将最低的构型能量定为零，画出高于较最低能量构型高出 0.5kcal/mol、1.0kcal/mol、1.5kcal/mol、2.0kcal/mol 的等高线。这样的等高线图称为 Ramachandran 图，可由此推断二肽分子可能的构型。图 16-10 为由 X 射线衍射实验所观察许多二肽分子的构型分布图，图中区域颜色的深浅表示分布的多寡。

将此构型分布图与能量等高线图比较，可发现能量愈低处，构型分布愈大；反之，构型出现于高能区的概率很小。这样的能量等高线图有助于 X 射线衍射或核磁共振实验决定生化分子可能的构型。当然，要全面了解生化分子的可能构型，除骨干主链的构型外，也必须考虑氨基酸残基上侧链的可能构型。一些实验显示，侧链的构型通常与骨干主链的构型相关。Summers 的实验发现，某些骨干主链的构型仅对应一定特殊的侧

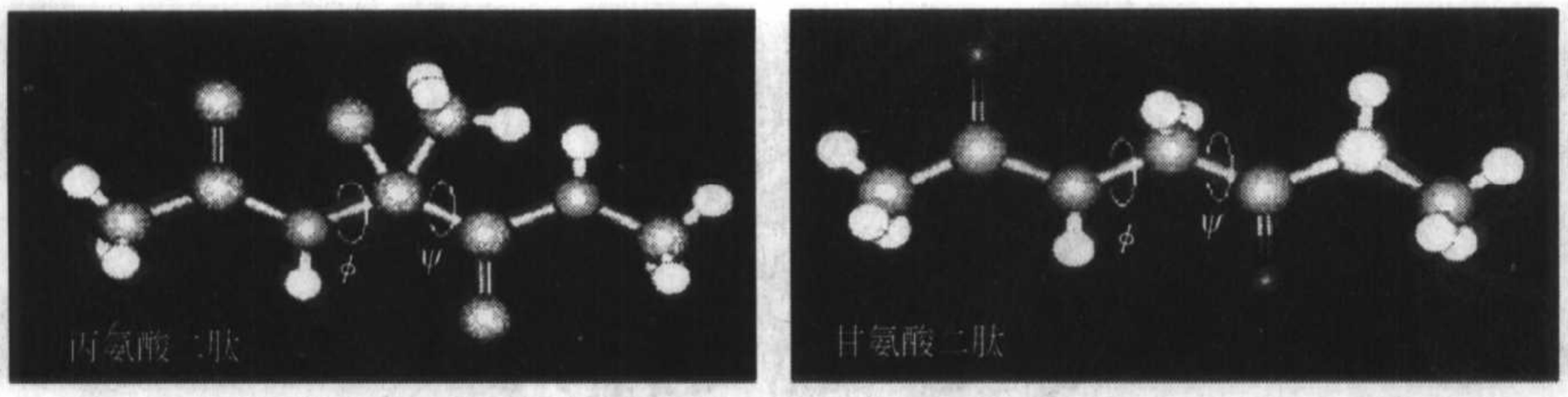



图 16-9　丙氨酸二肽与甘氨酸二肽的构型能等高线图

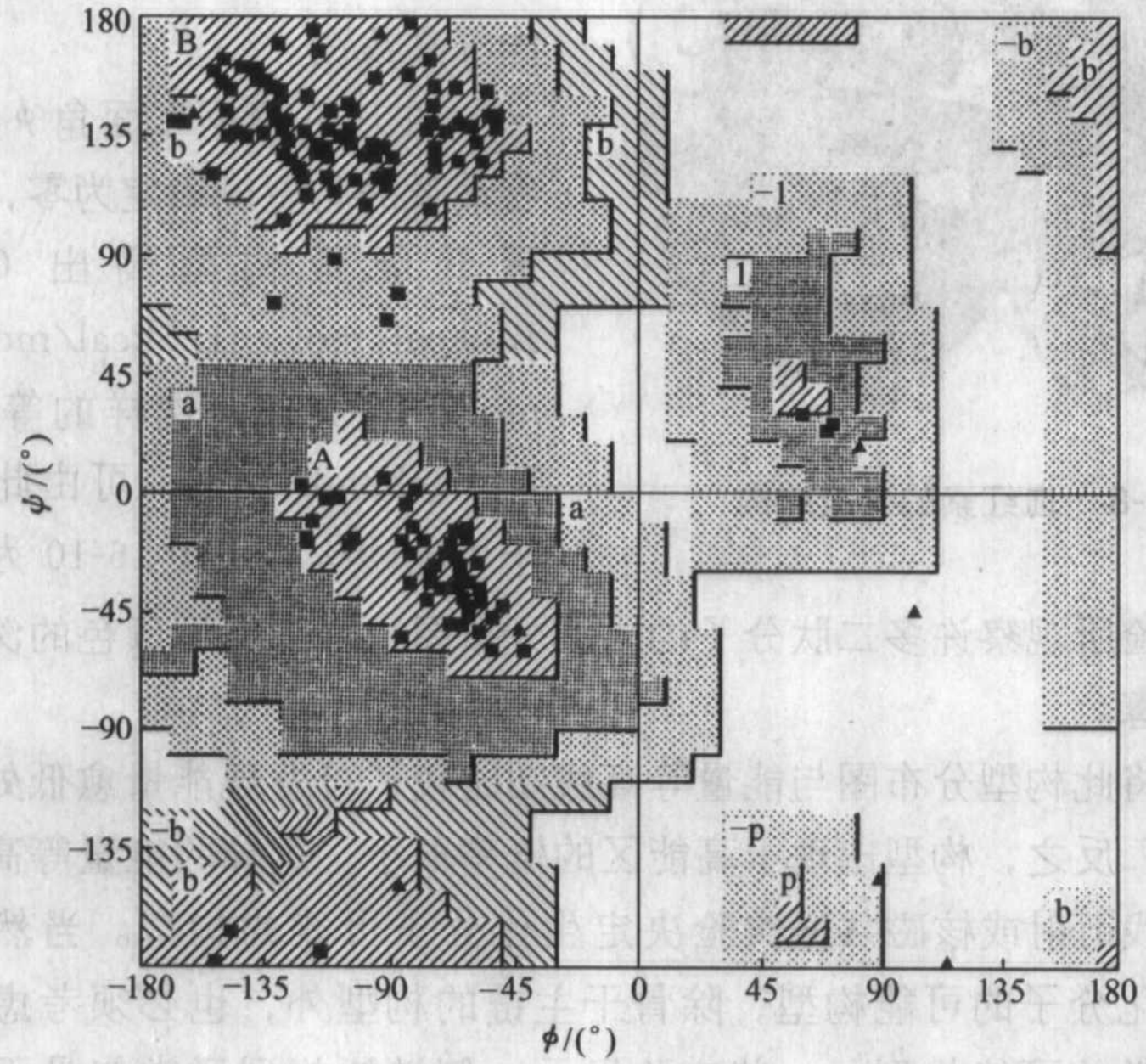



图 16-10　由 X 射线衍射实验所得二肽分子的构型分布图

链构型。

## 16.2 生化分子体系间的亲水性与疏水性

蛋白质分子在水中的折叠（folding）与解折叠（unfolding）为生化中最重要的过程。经由折叠后的蛋白质分子才有一定的生化功能。水溶性的蛋白质分子在水中形成一定的折叠形态，其分子的内部为疏水性（hydrophobic）的残基，如苯丙氨酸、色氨酸、缬氨酸、亮氨酸等，其外部则为亲水性的残基，如赖氨酸与精氨酸。这种残基的排列堆积方式源自疏水效应（hydrophobic effect），是影响蛋白质分子稳定性最重要的因素。所谓的疏水效应是指在极性溶剂中，非极性溶质分子倾向于与非极性分子聚集，而极性溶剂分子则倾向于与极性分子聚集。一般认为疏水效应是因系统的熵值（entropy）改变所造成的，其中，最重要的是溶剂分子的熵值改变，而非溶质分子熵或焓（enthalpy）的变化。通常分子折叠过程中，将蛋白质分子中非极性的氨基酸堆栈在一起的自由能变化对溶剂分子而言为不利的正值。自由能变化的各项中，熵与焓均为不利的值。堆栈过程的焓值为正值，这是由于将非极性的氨基酸由水中移出所需克服极性的水分子与残基侧链间的偶极-诱发偶极间之作用力，因此需作功才能达成。当疏水性的残基堆栈在一起，其间仅有微弱的分散力。堆栈过程熵值的改变为负值，因为未折叠状态较无秩序，而折叠状态较有一定的秩序。因此，以热力学的观点，堆栈过程不利于溶质分子。但堆栈后，原本被蛋白质分子限制住的水分子因疏水基团的靠近而被挤压出。对整个溶剂而言，其乱度增加，熵值的变化为正值，而有利于形成折叠状的蛋白质结构。一般当蛋白质分子堆栈后，分子附近的水分子环绕着非极性的溶质分子形成如同冰山的结构。在此种状态中的水分子，借其相互间的氢键作用，形成有序的排列结构，如图 16-11 所示。

折叠后的蛋白质分子，其非极性部分与水分子形成的界面面积远大于未折叠时的界面面积，如图 16-12 所示。

所以，当蛋白质分子折叠时，其熵值的改变很大。另外，分子折叠时须形成氢键，故焓为负值。以上四种影响自由能变化的因素中，焓值的改变通常不大，最重要的是溶剂水分子与蛋白质分子熵值的改变。这两种熵值的变化，决定蛋白质的结构变化。

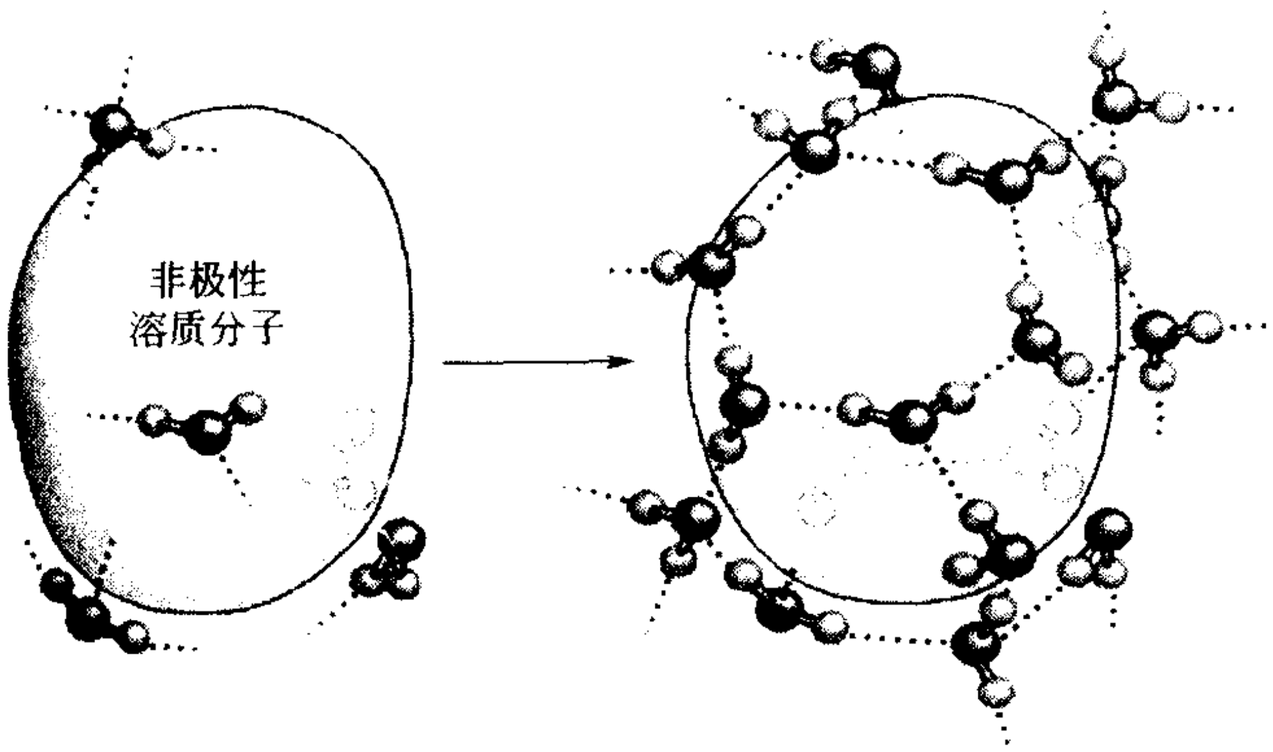



图 16-11　环绕生化分子的水分子排列

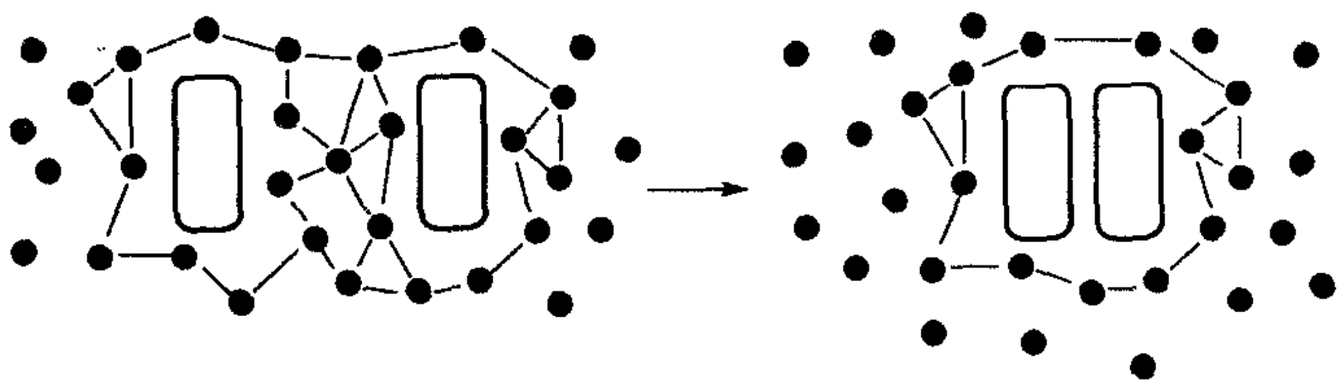


图 16-12　蛋白质分子折叠前后疏水区域的变化

## 16.3　蛋白质分子折叠的模拟

蛋白质分子因折叠而形成特定的结构，这些特定的结构显现出特定的生物机能。因此，欲了解生物体的一些特性及功能，必须先了解蛋白质折叠的形态与过程。但是一个未折叠的蛋白质分子如何折叠成特殊结构的过程却是研究上的一大难题。由能量的观点，一般未折叠与折叠后蛋白质分子间的自由能差仅为 1～15kcal/mol。这样微小的自由能差，却能导致庞大（数千至数万十个原子）的蛋白质分子折叠成固定的形态，是非常令人惊讶的。到底折叠过程中经过哪些过渡状态？环境的因素对折叠的过程有什么样的影响？除了能量因素外是否还有其他的影响因素？实验上如何鉴别折叠的过渡态？种种疑问，迄今仍为生物学研究的重点与努力寻找解答的目标。

分子动力模拟为研究蛋白质分子折叠与开折的动力过程最有效的方法。在模拟的过程中可以清楚地了解折叠（或开折）所历经的过渡态与能量的变化。但是，分子动力模拟方法因受模拟系统原子数目与模拟时间的限制，至今仍仅能模拟小片段蛋白质分子（多肽分子）的折叠过程。此处列举 DeMarco 等近期利用分子动力模拟研究含有 61 个残基的镶边形（engrailed homeodomain，En-HD）小蛋白质分子的折叠与开折过程的工作。图 16-13 为 En-HD 蛋白质分子片段中残基序号 3-56 的晶体结构。

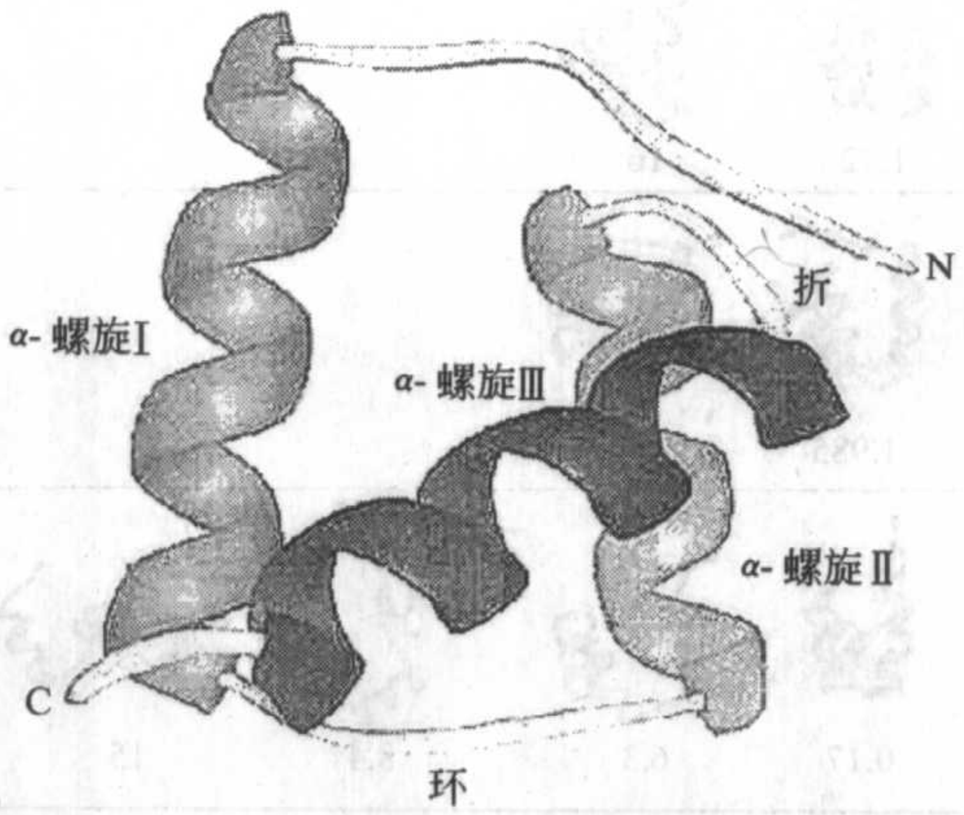



**图 16-13　En-HD 蛋白质分子片段中残基序号 3-56 的晶体结构**

图 16-13 中三个螺旋状片段 HⅠ、HⅡ、HⅢ 的残基序列分别为 SEQLARLKREFNE，ERRRQQLSSE 与 AQIKIWPQNKRAKI，螺旋片段间形成环路（loop）。实验显示，En-HD 分子的折叠与开折过程均十分快速。在 25℃的条件下，其折叠的速率常数为 $k_F = 39000s^{-1}$，开折的速率常数为 $k_U = 1100s^{-1}$。是目前所测量的蛋白质分子中折叠与开折速率最快的。

DeMarco 等利用商业化的生物分子计算软件 ENCAD 执行 En-HD 分子动力模拟计算。计算系统中包含刚性的水分子。水分子间之作用选择 TIP3 力场。图 16-14 为 DeMarco 等模拟所得 En-HD 分子的开折过程的结果。图中的第一行为模拟的温度与系统编号，每个图形下的数字为以纳秒为单位的模拟时间。

图 16-14 中清楚显示出 En-HD 分子在高温（498K）下的开折过程（498_2）。分析计算的结果显示当 En-HD 分子开折时，在最初的 20ns，HⅠ 片段因溶剂分子的作用而离开 HⅡ 与 HⅢ片段。但整根蛋白质分子的

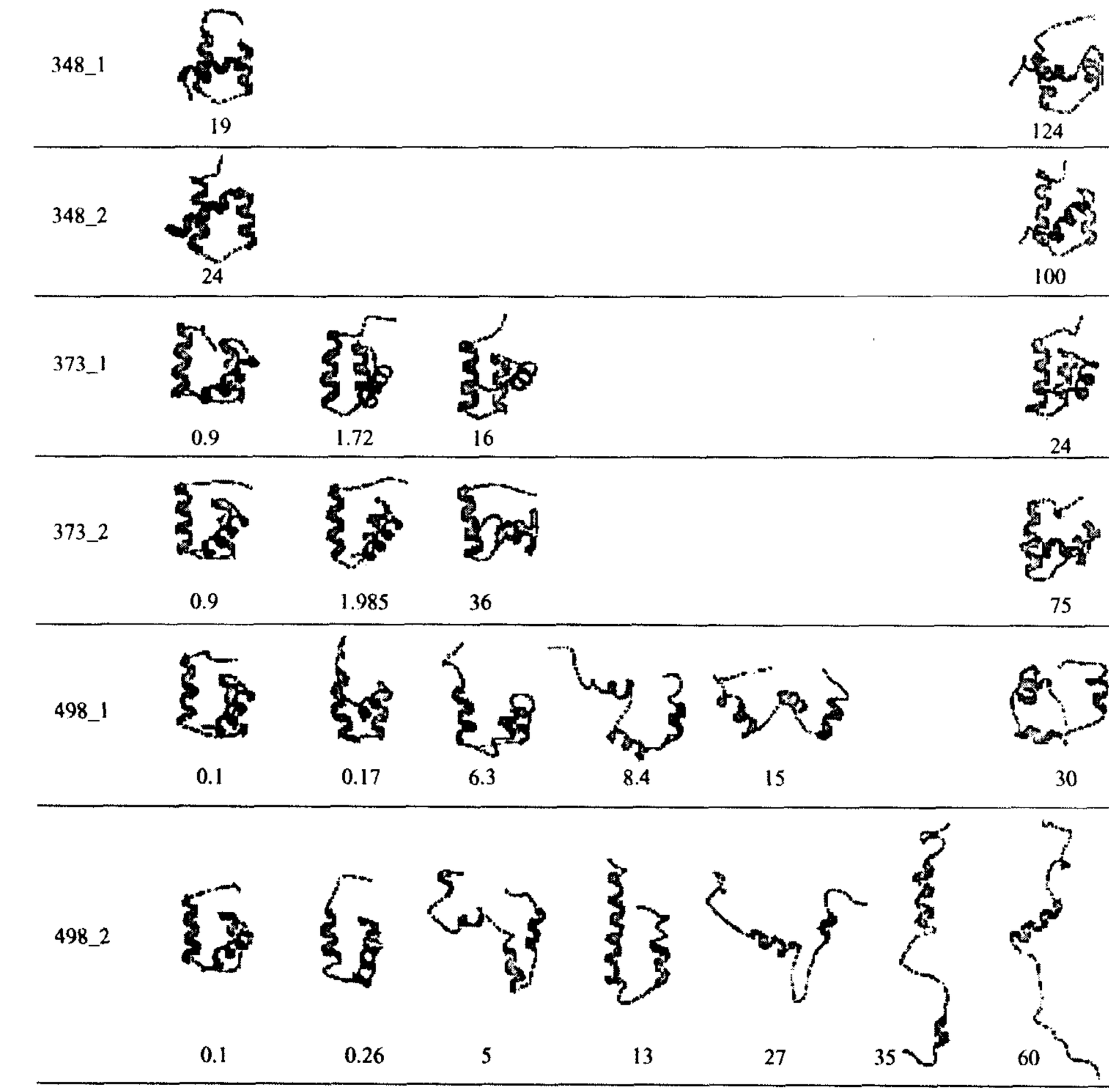



**图 16-14　各温度下 En-HD 分子的开折过程**

螺旋度（helicity）在 11～20ns 间却呈现增加的状态。当时间为 30ns 后，蛋白质分子开始开折至线状。开折过程中最明显的变化为 En-HD 分子片段的螺旋度的改变。HⅡ 片段的螺旋度在开折的过程随时间而减少，表示此片段内的分子间作用力减弱，导致蛋白质链的开折。另外，HⅡ片段的螺旋度却随蛋白质链的开折而大量增加。HⅠ 片段的螺旋度，则先小幅度增加，当链完全开折后，再减少至零。表 16-1 为 498_2 系统中，模拟时间 30ns 后，En-HD 分子上三片段螺旋度的变化量。

表 16-1　En-HD 分子上三片段螺旋度的变化量

| 模拟时间/ns | HⅠ | HⅡ | HⅢ |
|---|---|---|---|
| 30 | 20% | 10% | 30% |
| 35 | 30% | 12% | 60% |
| 40 | 43% | 32% | 60% |
| 45 | 70% | 10% | 8% |
| 50 | 60% | 40% | 12% |
| 55 | 62% | 50% | 20% |
| 60 | 0 | 50% | 8% |

DeMarco 等对 En-HD 蛋白质分子的模拟为典型的开折动态模拟。此种方法以晶体数据库中已折叠的蛋白质分子结构作为模拟的起点，将系统升温研究其开折的过程。但这并不表示此过程为可逆的，即一根线形的伸展形蛋白质分子开始折叠形成最终的堆栈结构并不一定经过相同的折叠途径。

## 16.4　细胞膜的融合模拟

生化分子体系的模拟是非常困难的工作。模拟的结果必须与实验的结果互相参照，以保证模拟的正确性。文献中有大量生化分子体系模拟的论文，此处列举两个较新的分子动力模拟的实例，分别为研究细胞膜的融合与研究生化分子体系的疏水性性质。利用分子动力计算模拟由酯质分子（dimyristoyl phosphatidylcholine，DMPC）所形成的类似细胞膜（cell membrane）的结构与其在水溶液中的稳定性。除了模拟细胞膜外，同时研究融合肽分子（fusion peptide）与细胞膜结合的结构问题。病毒分子与细胞膜的融合，被认为是导致病毒感染的主要原因。模拟结果显示 DMPC 所形成的双层膜在溶液中可维持稳定的结构，V 形的 HX 分子主要以氢键作用与膜形成稳定的结构。膜的结构因 HX 分子的结合产生变形。模拟的结果与核磁共振（nuclei magnetic resonance，NMR）及各类实验的结果吻合。

细胞膜多由双层的脂质分子自组装（self-assembly）而成。所谓自组装意为分子与分子间并无化学键结存在，而是靠范德瓦耳斯作用力，氢键作用力等微弱的作用力形成规则的聚集排列。Huang 等利用分子动力模拟研究由 DMPC 分子（dimristoyl phosphatidylcholine）自组装所形成的类似细胞膜的聚集体的稳定性与此细胞膜和融合肽分子 HA X31 所形成的构型。HA 为一种流感病毒血球凝集素（influenza virus hemagglutnin），X31 为 HA 的一个片段，含有 20 个氨基酸残基。图 16-15 为 DMPC 分子的结构式。

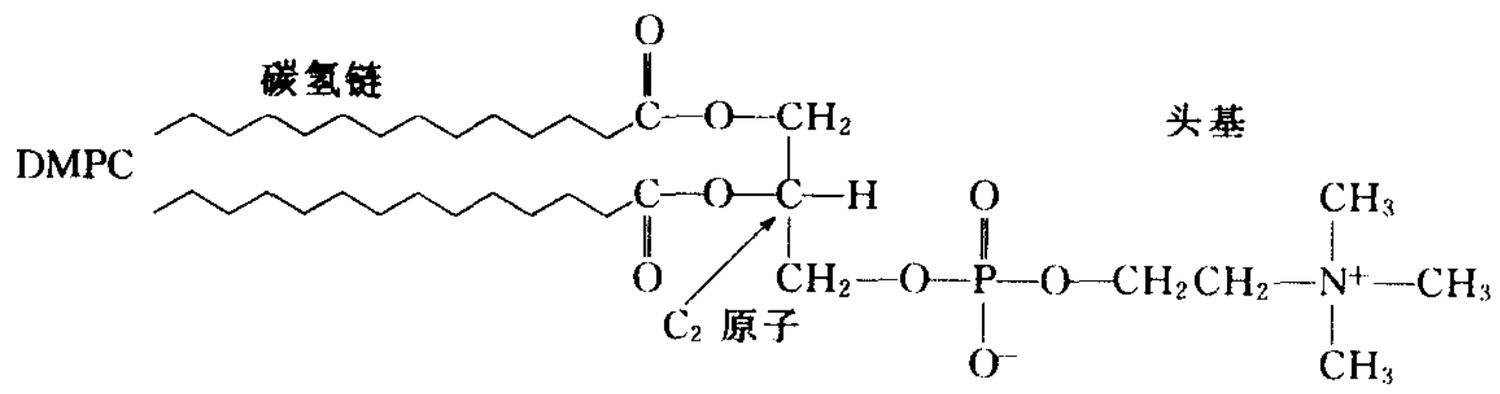



**图 16-15　DMPC 分子的结构式**

此分子包含两根疏水的烷链与磷脂基团，其头端为亲水的铵离子基团。融合肽分子 HA X31 的氨基酸残基序列为（GLFGA-IAGFI-ENGWE-GMIDG）。模拟的双层脂膜中的顶层及底层分别由 60 及 66 根 DMPC 分子自组装而成，膜的截面积约为 360Å$^2$。模拟系统的盒长（$X\times Y\times Z$）为 62.8Å×62.8Å×72Å，除脂质双层膜及融合肽分子外，含有 4800 个水分子，系统的总原子数约为 30000。在此系统中纳入水分子是非常重要的，因为 DMPC 的亲水端与疏水部分与水的作用而导致形成稳定的双层脂膜。若是没有水分子的存在，则基团间互相排斥，不能形成稳定的膜。但纳入如此多的水分子使得计算量增大许多，计算时间增长，而且因为水分子的运动较生物分子快速，所以大部分的模拟为水的运动。图 16-16 为 DMPC-HA X31 模拟系统的示意。

Huang 等先利用分子动力方法模拟 DMPC 双层脂的结构。待所模拟的双层脂呈平衡后，再加入已平衡的水分子及融合肽分子 HA X31。以杆状分子为 HA X31 的起始构型，并根据圆二色性（circular dichroism，CD）实验及电子顺磁共振（electron paramagnetic resonance，EPR）实验的结果将其置于与细胞膜约呈 35° 斜角的位置。计算所采用的为 CHARMM 力场，及定温定压（NPT）系综。将水分子视为刚性分子，其相互间的作用采用 TIP3 模型。计算的积分步长为 2fs，每个系统的模拟时间则为 18ns。利用由 4 台 PC 组装的团簇体进行模拟，实际的模拟时间约为两个月。HA X31 分子的 N 端可能为中性（Ⅰ）或为质子化带正电（Ⅱ）。图 16-17 比较此两种分子经模拟时间 1500ps 后结构的差异。

图 16-17(a) 中深色的分子为肽分子（Ⅱ）的结构，浅色为肽分子（Ⅰ）的结构。图中显示肽分子（Ⅰ）与（Ⅱ）均呈现大幅度的弯曲形状。图 16-17(b) 显示肽分子（Ⅱ）中由疏水基团 Leu2，Phe3，Ala5，Ile6，Ala7，Phe9，Ile10 与 Trp14 等所堆积形成的疏水区。这显示在水中疏水性的基团因其与水的排斥性而导致互相堆积，结果形成了疏水的区域。图 16-18 为系统（Ⅱ）中，双层脂质分子膜的上层分子（1-60）与下层分子

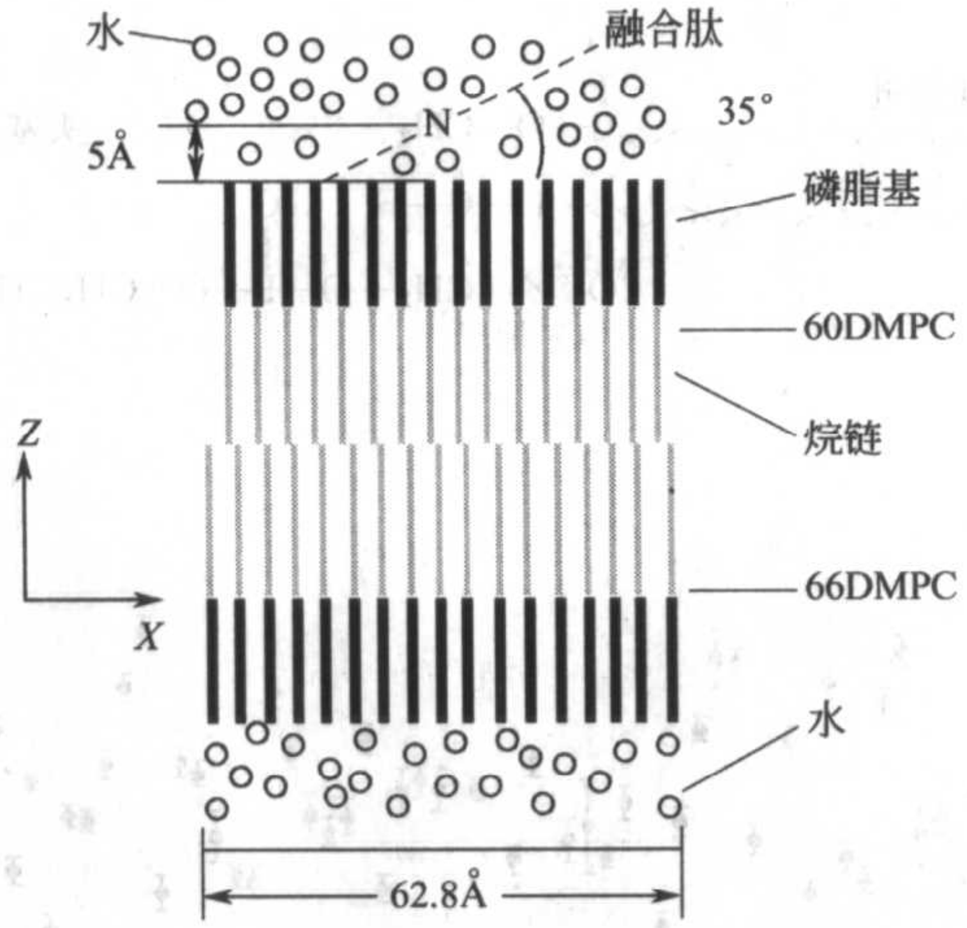



图 16-16　DMPC 双层脂系统与 HA X31 的示意

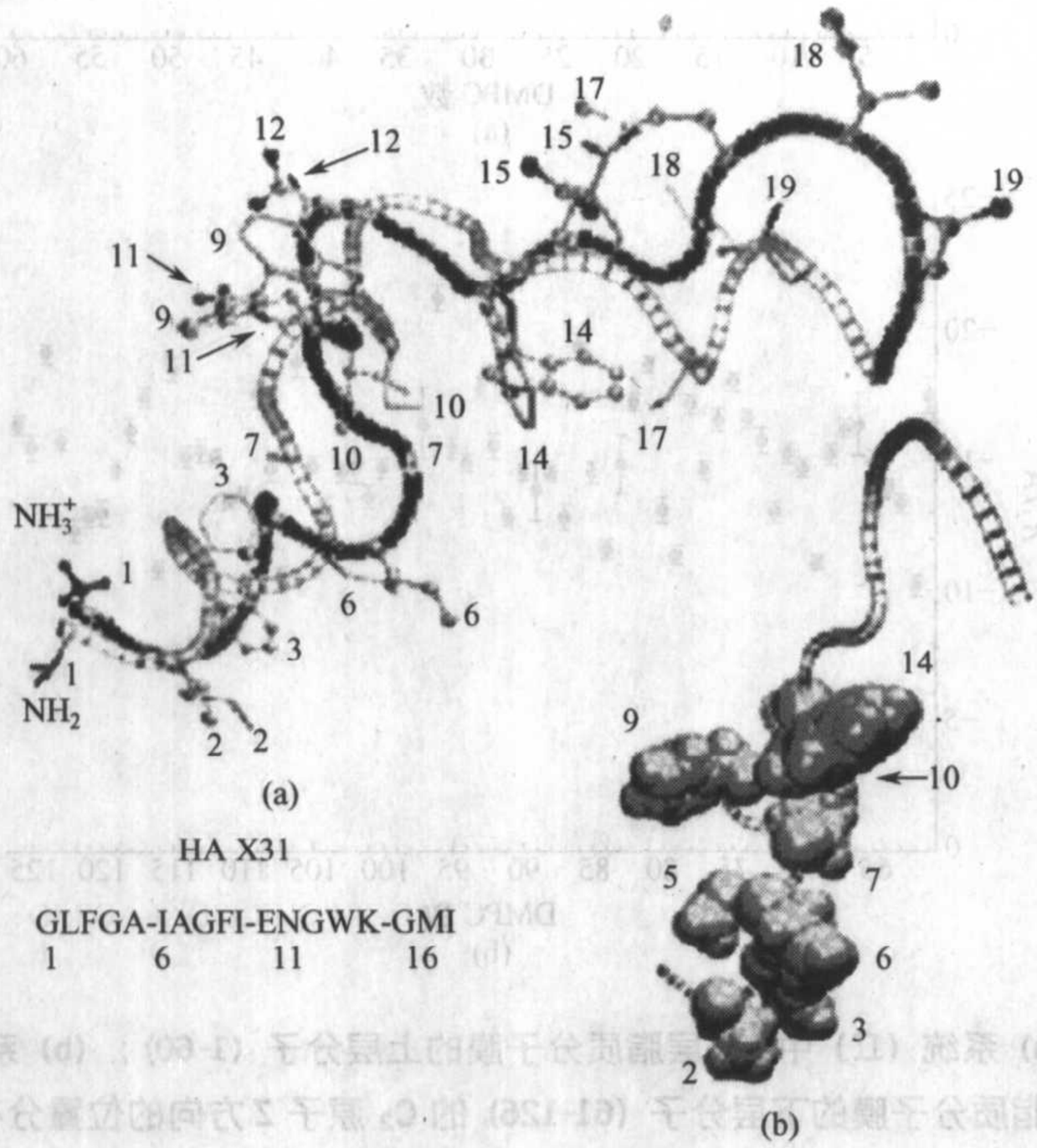
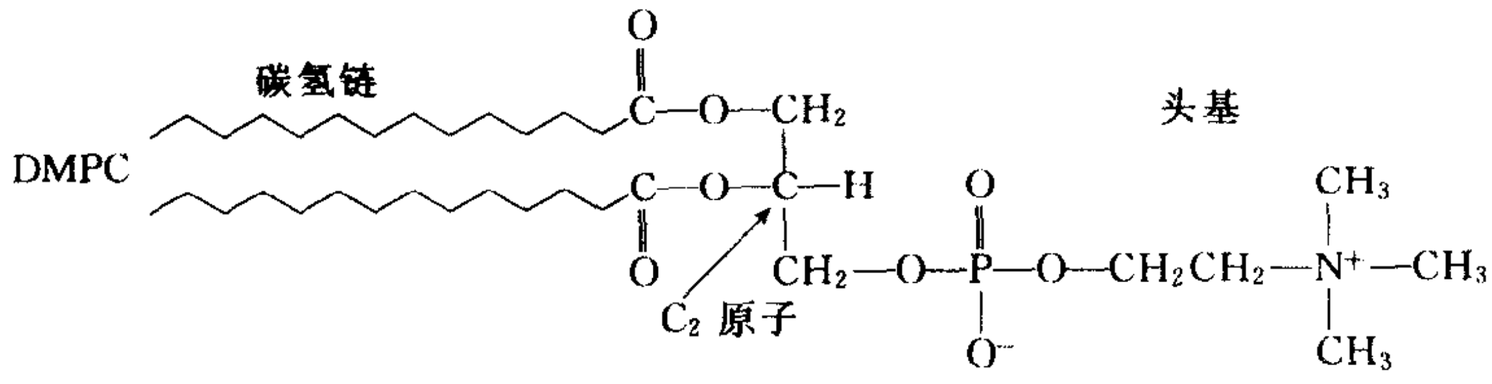



图 16-17　HA X31 分子 I 与 II 的结构变化比较



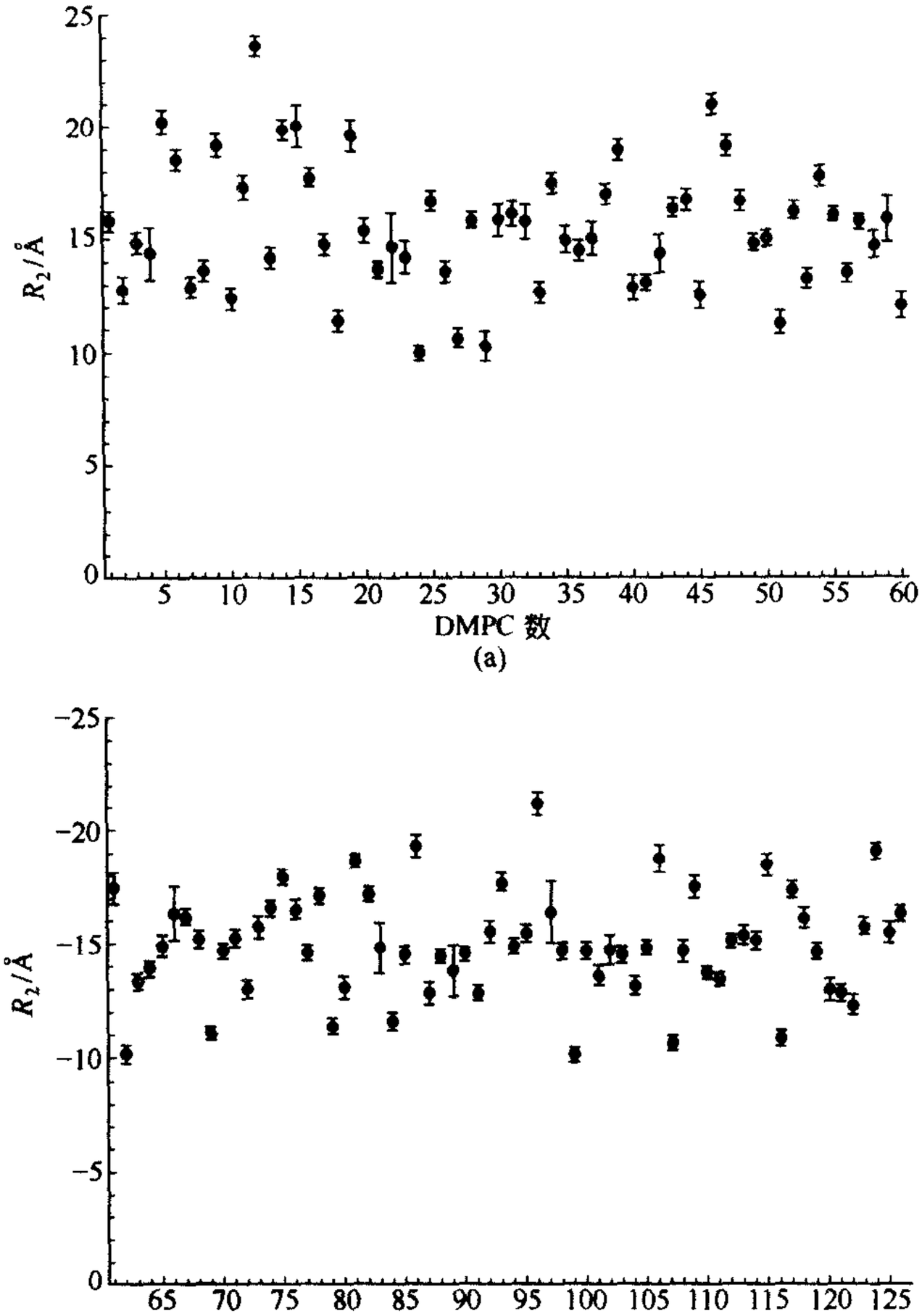




**图 16-18　(a) 系统（Ⅱ）中，双层脂质分子膜的上层分子 (1-60)；(b) 系统（Ⅱ）中，双层脂质分子膜的下层分子 (61-126) 的 $C_2$ 原子 Z 方向的位置分布图**

图中 $R_2$ 为 $C_2$ 原子与中心平面的距离。

(61-126) 的 $C_2$ 原子经模拟后在 $Z$ 方向的位置分布图。此图显示脂质分子 DMPC 所形成的双层膜大致维持稳定的结构，其 $Z$ 方向的标准偏差约为 5～7Å。上层膜分子中，因与水及融合肽分子的作用，导致其偏差较大，形成较为扩展的结构。

图 16-19 为 HA X31 分子（Ⅱ）与双层膜的模拟结果。图中显示肽分子的部分插入膜中（融合）形成弯曲的 V 形结构。Egremov 等曾利用理论计算的方法推测肽分子在细胞膜上的构型为 α-螺旋形，并且与细胞膜平面呈现固定的倾斜角。Lineberg 与 Macosko 等利用低分辨率的 CD 与 EPR 实验均推断融合肽分子在细胞膜上的构型为类似杆状的结构。但是，Han 与 Tamm 利用核磁共振（NMR）的方法却认为融合肽分子 HA X31 在细胞膜 DPC（dodcylphosphocholine）的表面结构应为 V 形。并且，融合肽分子在细胞膜的表面堆积成一团疏水区。

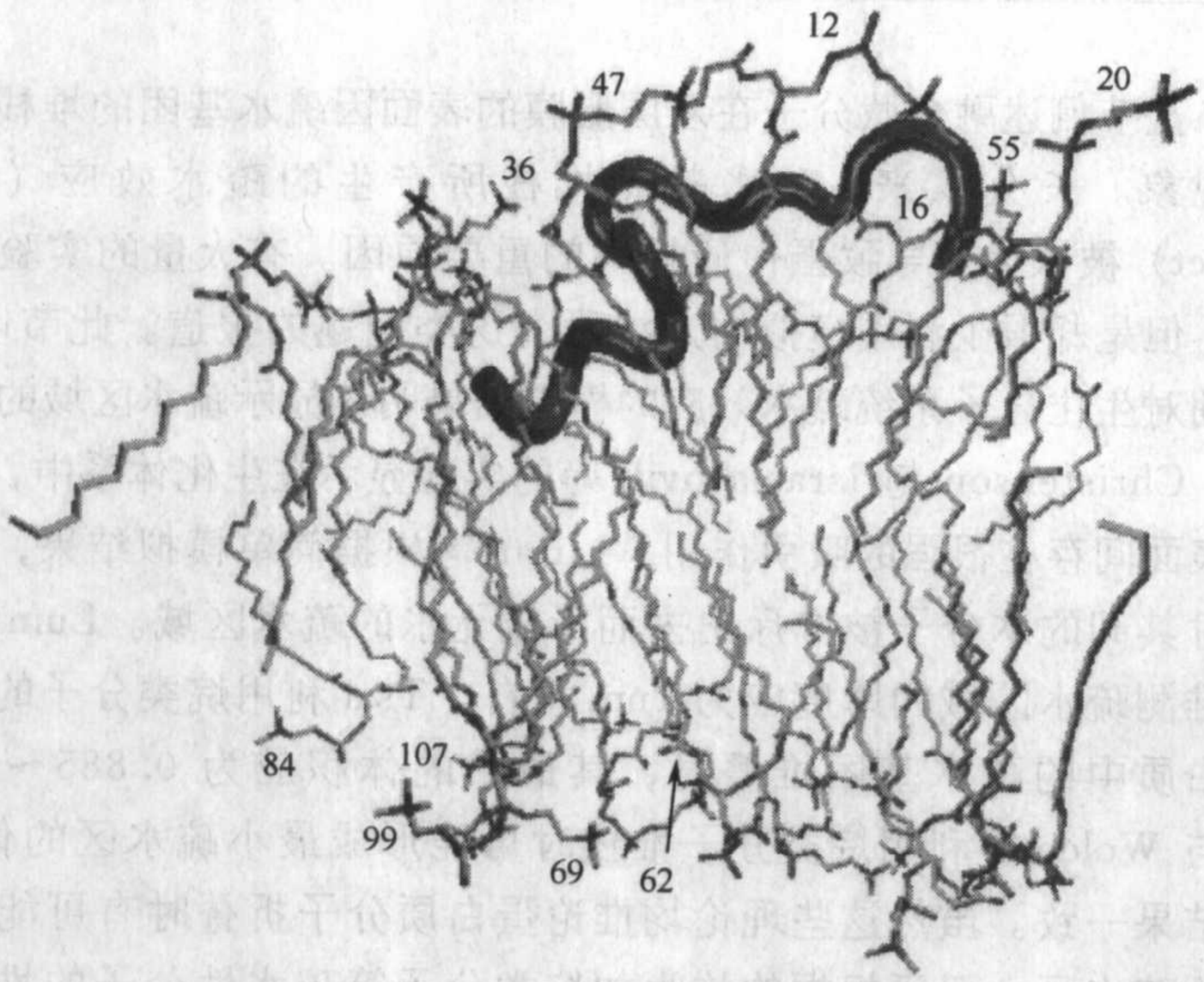



**图 16-19　HA X31 分子（Ⅱ）与双层膜的模拟结构**

Huang 等利用分子动力模拟融合肽分子 HA X31 与 DMPC 双层膜的结果与 Han 的 NMR 实验的所得肽分子的 V 构型一致。并且，模拟的结果也证实了疏水区域的存在。

表 16-2 为比较融合肽分子系统（Ⅰ与Ⅱ）与细胞膜间及水分子的各

种类型作用势能。表中的数据显示，系统Ⅱ的能量低于系统Ⅰ的能量。其主要的原因是因肽分子Ⅱ与 DMPC 分子产生较有利的作用，而导致系统能量的降低。

**表 16-2　HA X-31、DMPC 与水各种类型的能量**

| 能　量 | 系统Ⅰ(700～1400ps) /(kcal/mol) | 系统Ⅱ(700～1500ps) /(kcal/mol) | 百分差异/% |
|---|---|---|---|
| HA X31 | 474.5±15.5 | 404.7±14.2 | −14.7 |
| DMPC | −2556.7±133.0 | −3300.9±151.7 | −29.1 |
| 水 | −33002.4±154.0 | −33506.3±158.5 | −1.5 |
| HA X31-DMPCa | −551.2±12.8 | −771.8±31.5 | −40.0 |
| HA X31-水 | −820.0±28.6 | −825.9±31.2 | −0.7 |
| 水-DMPC | −17823.1±160.7 | −17517.3±209.6 | −1.7 |

# 16.5　疏水基团的堆积

上节例述融合肽分子在双层脂膜的表面因疏水基团的堆栈产生疏水区的现象。长久以来，疏水基团堆栈所产生的疏水效应（hydrophobic effect）被认为是导致蛋白质折叠的重要原因。有大量的实验支持这种推测，但是却很少利用模拟的方法研究这类问题的报道。此节以 Huang 等近期对生化分子系统疏水效应的模拟结果明确显示疏水区域的形成。

Christenson 与 Israelachvili 等的实验显示在生化体系中，两个疏水性的表面间存在很强的吸引作用。Luzar 等根据简单模拟结果，认为界面接近时其间的水分子被排斥出去而形成无水的疏水区域。Lum 等以理论模型推测疏水区域的厚度应为 1nm 左右。Tsai 利用烷类分子的计算推测当蛋白质中的疏水基团堆叠时，其最小的体积应为 0.885～1.075 $nm^3$。这与 Wolde 等利用烷类分子堆栈时可能形成最小疏水区的体积为 $1nm^3$ 的结果一致。虽然这些理论均推论蛋白质分子折叠时有可能形成纳米尺寸的疏水区，但所根据的均为如烷类分子等疏水性分子的堆栈结果，而非真正的模拟蛋白质分子折叠而形成疏水区的过程，以致缺乏实际的证据。针对此，Huang 等利用分子动力学方法，模拟两个 $\alpha_2$D 蛋白质分子在水溶液中堆栈形成疏水区的过程，此模拟提供了形成疏水区域的实际例证。

$\alpha_2$D 蛋白质分子为人工合成专为研究所用的小蛋白质分子，这种蛋白质称为 de Novo 蛋白质（de Novo protein）。$\alpha_2$D 蛋白质分子由两条相同

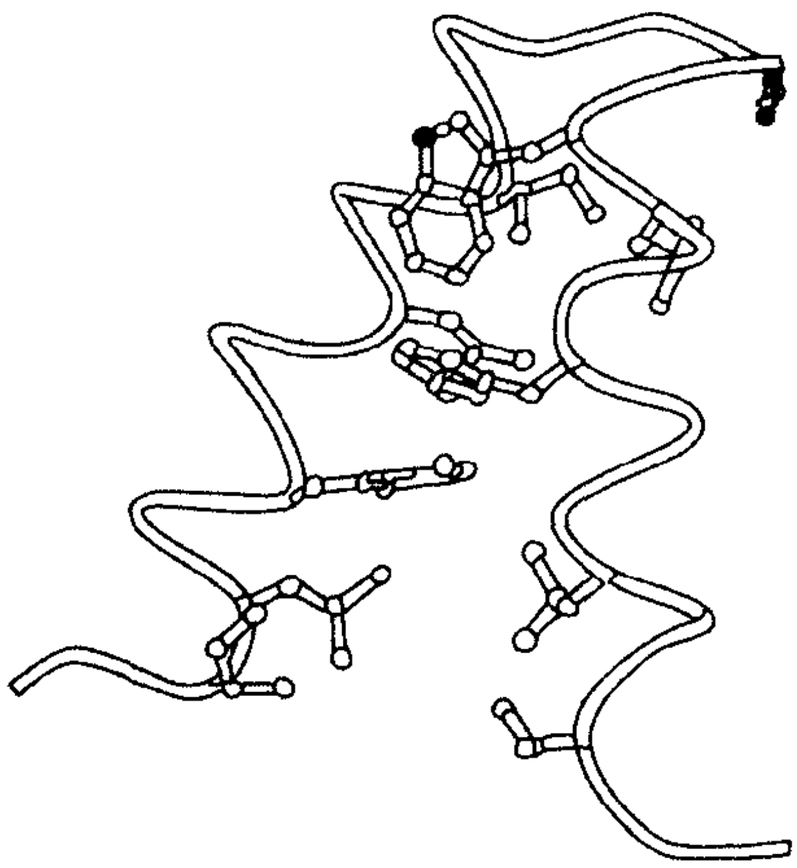


**图 16-20　$\alpha_2$D 分子结构的示意**

序列的 $\alpha$ 螺旋状多肽分子链所组成，每条链含有 35 个残基。多肽分子链的残基序列为 GEVEE-LEKKF-KELWK-GPRRG-EIEFL-HKKFH-ELIKG。整个分子的形状大体呈现 U 字形。在 $\alpha_2$D 分子中，多肽链上的疏水性残基为 Val-3、Leu-6、Phe-10、Leu-13、Trp-14、Ile-22、Leu-25、Phe-29、Leu-32、Ile-33。依据 Kyte 与 Doolie 等所定的疏水亲水尺度，这些疏水性的残基中，又以 Ile、Leu、Val、Phe 等为疏水性最强的基团。图 16-20 为 $\alpha_2$D 分子结构的示意图。

核磁共振 NMR 实验的证据显示，在水溶液中，当两个 $\alpha_2$D 分子堆栈时，会形成两个 U 形分子交错嵌入的结构，而分子链上的疏水基团会堆栈形成疏水区域。图 16-21 显示这种分子的镶嵌，图中所标示的球状区域为疏水区。

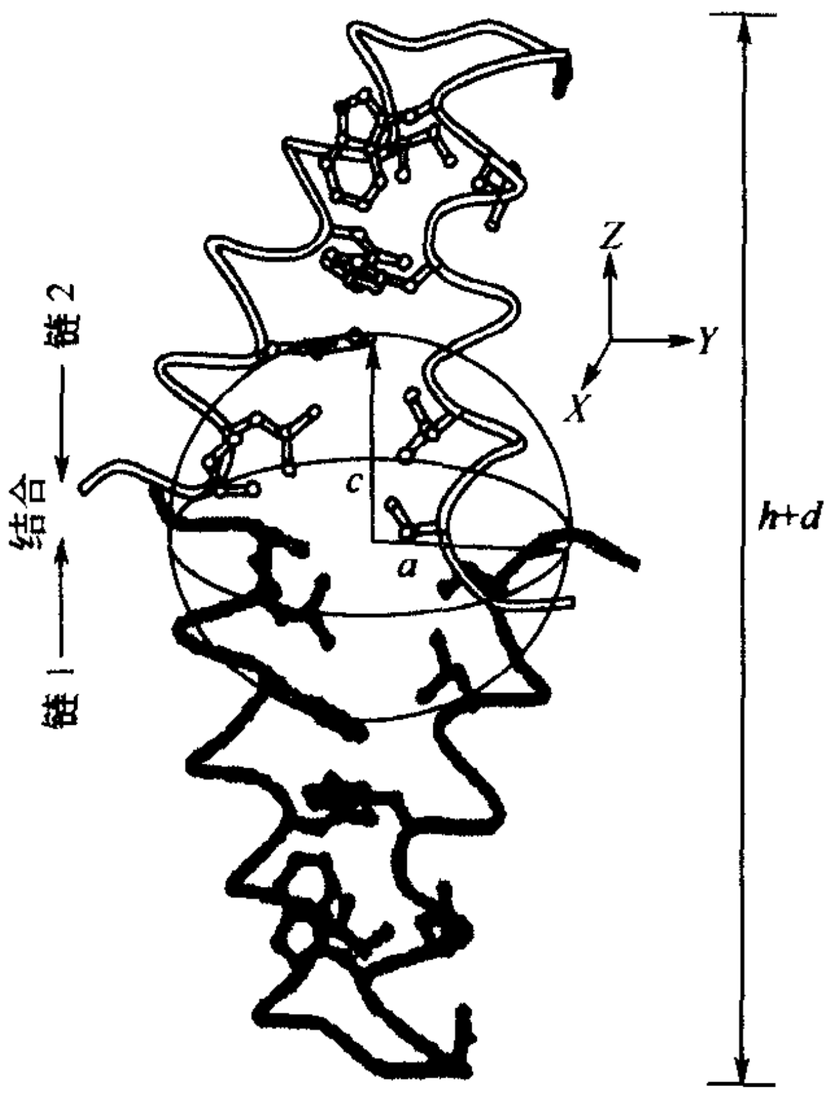



**图 16-21　两个 $\alpha_2$D 分子的镶嵌图**

若以分子动力学方法完全模拟两个 $\alpha_2$D 分子在水溶液中镶嵌与形成疏水区的过程，需要耗费相当长的时间。Huang 等采用的方法为先将两个 $\alpha_2$D 分子置于如图 16-21 中的位置。图中的距离 $d$ 为两个分子上 Phe-29 的距离。再将此二分子聚合物与 7500 个水分子置于模拟的立方形盒中。水分子的起始位置为随机选取，系统中水的密度为 $1g/cm^3$，盒长为 4.0nm×4.0nm×4.0nm。依此方法，建立 8 个类似系统的起始结构，分别对应于 $d$ = 0.6nm，0.8nm，0.9nm，1.0nm，1.1nm，1.2nm，1.6nm 与 2.2nm。当起始结构建立好后，再分别对此 8 个系统进行分子动力模拟。

分子动力模拟选取 GROMOS 力场模拟生化分子与水的作用。将水分子视

为刚性分子，水分子间的作用则选取 SPC 力场。模拟选取的环境为 NPT 系综，温度为 300K，压力为 $P=1\text{atm}$（$1\text{atm}=10^5\text{Pa}$）。每个系统执行 1ns 的动力模拟，所选的积分步长为 2ps。为了确定 U 形 $\alpha_2$D 分子的稳定性，先执行单独一个 $\alpha_2$D 分子在水溶液中的分子动力模拟。经过长达 50ns 的计算后，检视 $\alpha_2$D 分子的结构，发现其仍为 U 字形。因此，可确定此种形状的 $\alpha_2$D 分子为稳定的结构。图 16-22 为各个系统经模拟后，两个 $\alpha_2$D 分子交接处的球状区域中水分子数目的变化。图中的实心圆点为球状区的水分子数，空心圆点为纯水中相等体积的圆球体中的平均水分子数。

图中显示，当 $d$ 小于 1.2nm 时，球状区间装的水分子数目明显减少。当 $d=0.9\text{nm}$ 时，区间中所有的水分子已被排除出去。此模拟的结果明确的显示当两个 $\alpha_2$D 分子堆栈时，其疏水性基团堆栈而形成疏水区。图 16-23

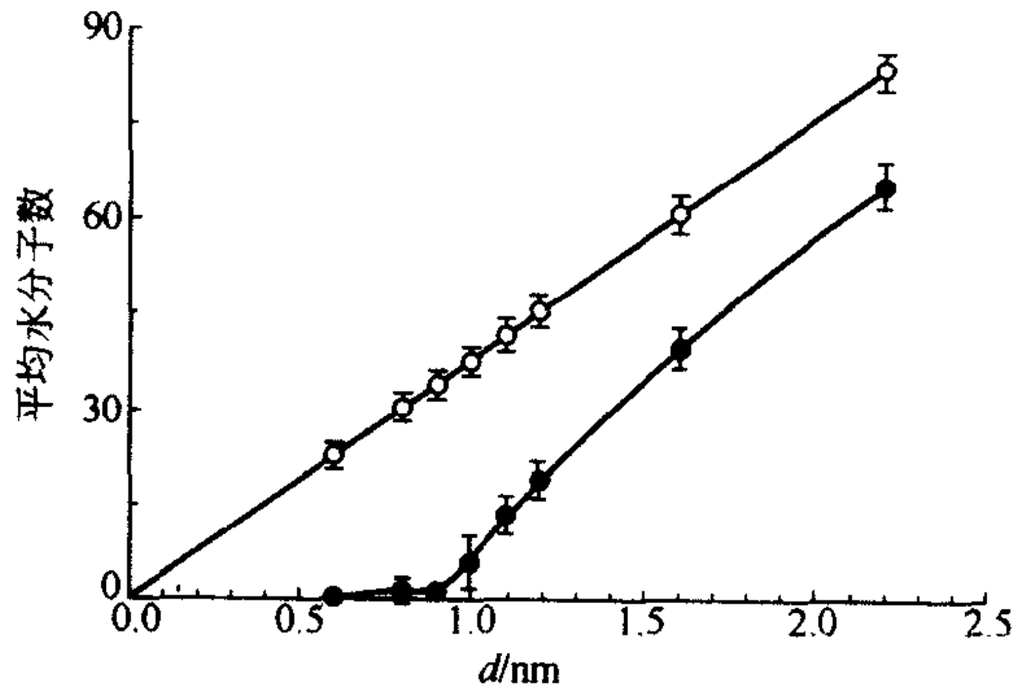



**图 16-22　疏水区中水分子数目的变化**

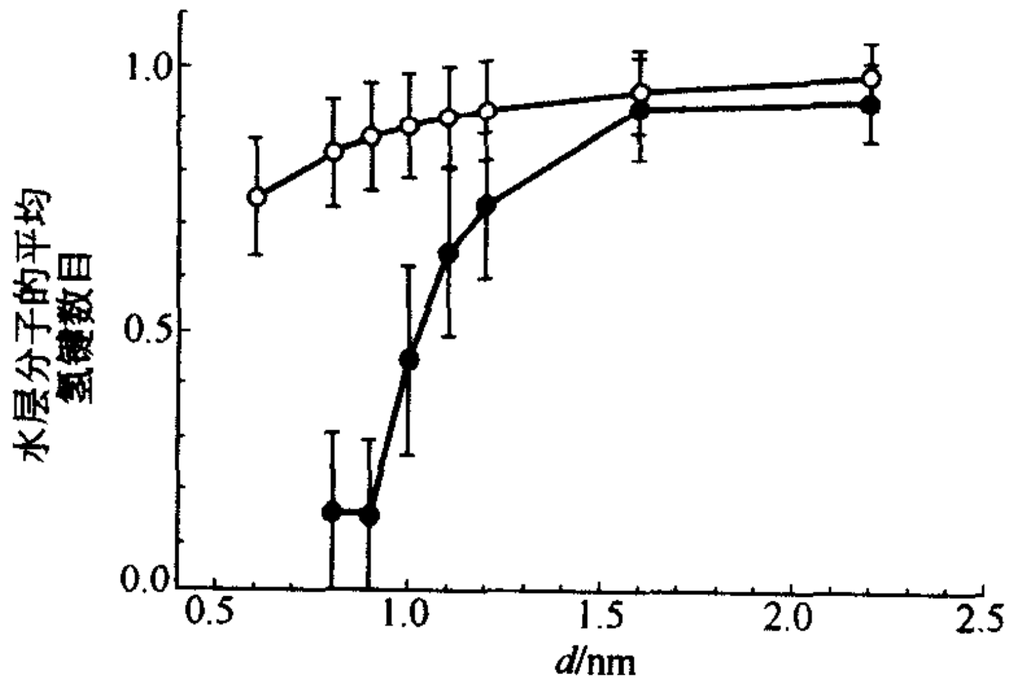



**图 16-23　疏水区中水分子与 $\alpha_2$D 分子间氢键数目的变化**

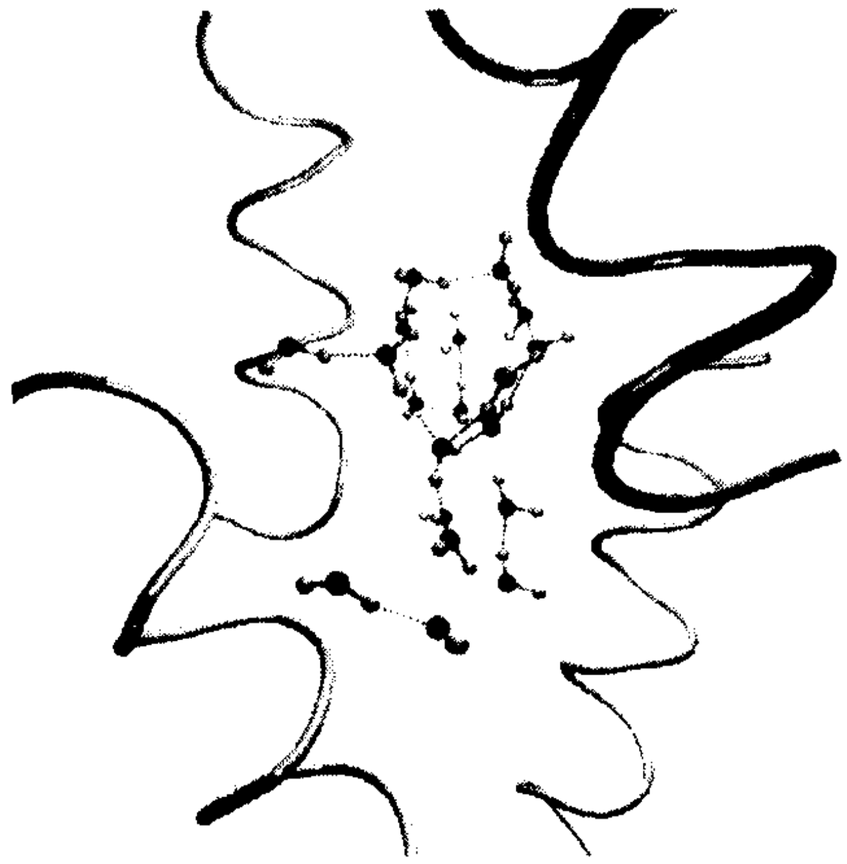


**图 16-24　两个 $\alpha_2$D 分子堆栈区间中所形成的氢键图**

为疏水区中水分子与 $\alpha_2$D 分子间氢键数目的变化。

图中明确地显示当 $d$ 小于 1.2nm 时，氢键的数目急速减少。当 $d=0.9$nm 时，氢键的数目已趋最小值。图 16-24 为两个 $\alpha_2$D 分子堆栈区间中所形成的氢键图，图中只显示出 $\alpha_2$D 分子的骨干。

根据以上的模拟结果，可知两个 $\alpha_2$D 分子堆栈形成的疏水区是由于堆栈时破坏了水分子间及水分子与 $\alpha_2$D 分子间的氢键所致。此外，模拟显示疏水区的体积约为 $1nm^3$，与各种实验与理论推测的结果相吻合。此为首次以分子模拟方法清楚地显示蛋白质分子堆栈形成疏水区动态过程的报告。

## 参　考　文　献


1　DeMarco M L, Alonso D O V, Daggett V. J. Mol. Bio., 341, 1109, 2004

2　Bradshaw J P, Darkes M J M, Harroun T A, Katsaras J, Epand R M. Biochemistry, 39, 6581, 2000

3　Belohorcova K, Qian J, Davis J H. *Biophys. J.* 79, 3201, 2000

4　Chang D-K, Cheng S-F, Trivedi V Deo, Yang S-H. J. Biol. Chem., 275, 19150, 2000

5　Eckert D M, Kim P S. *Annu. Rev. Biochem.* 70, 777, 2001

6　Han X, Bushweller J N, Cafiso D S, Tamm L K. Nature Struct., Biol. 8, 715, 2001

7　Han X, Tamm L K. Proc. Natl. Acad. Sci. USA, 97, 13097, 2000

8　Huang Q, Chen C L, Hermann A. Biophysical J, 87, 14, 2004

9　Christenson H K, Claesson P M. Science, 239, 390, 1988

10　Tsai J, Gerstein M, Levitt M. Protein Sci, 6, 2606, 1997

11　Kyte J, Doolite R. J. Mol. Bio., 157, 105, 1982

12　Huang Q, Ding S, Hua C Y, Yang S C, Chen C L. J. Chem. Phys., 121, 1969, 2004

# 第⑰章 分子模拟与药物设计

药物可以理解为调节人体的生理功能或病理状态的外来物质，简单分为小分子和大分子药物，常见的多为小分子药物。一个药物成药性包括：活性、毒性、药动学特性和化学性质等四个方面。药物设计的目的就是要在分子设计阶段对分子的成药性有一个良好的预测。药物设计的内容可以简单地分成：小分子结构、靶标结构、小分子与靶标之间相互作用等三部分。本章主要是介绍基于小分子（构效关系和药效团模建）和基于靶标（对接和从头设计）的两类计算方法，其原理的实质就是前面各章介绍的分子力学和分子动力学的方法。本章还介绍了这两类方法的实际计算例子，帮助读者更好地理解分子模拟在药物设计中的应用。

## 17.1 小分子的定量构效关系

研究药物结构和活性关系对于药物的发展发挥了重要的作用，现在还是优化先导化合物的重要方法，在药物的设计领域有着不可替代的作用。定量结构-活性关系（quantitative structure-activity relationship，QSAR）是利用数学模式建立结构和活性关系的方法。一般结构和活性的关系式如下：

$$Y=f(x) \tag{17-1}$$

式中，$x$ 是分子的结构性质；$Y$ 是活性；$f$ 是适当的函数。

### 17.1.1 传统的构效关系

最早利用这种关系研究生化活性的为 Hansch。他发展了一个生物活

性和分子电子性质及疏水性质的关系式：

$$\lg(1/C)=k_1\lg P-k_2(\lg P)^2+k_3\sigma+k_4 \tag{17-2}$$

式中，$C$ 是分子的浓度；$P$ 是将分子处于辛醇和水溶液中的分配系数，代表疏水性程度；$\sigma$ 代表 Hammett 取代系数；$k_1$、$k_4$ 等都是常数。在 20 世纪 60～80 年代，Hansch 方程及各种类似或衍生的构效关系式得到广泛应用，比较重要的有 Hansch-藤田模型与 Free-Wilson 模型。

Hansch-藤田模型的通式为：

$$\lg(1/C)=-k_1\pi^2+k_2\pi+k_3\sigma+k_4E_z+k_5 \tag{17-3}$$

式中，$-k_1$，$k_2$，$k_3$，$k_4$ 代表各项系数；$k_5$ 是常数项；$\pi$，$\sigma$，$E_z$ 等分别代表了疏水系数、电子效应和立体参数。

Free-Wilson 模型则假定分子的活性是母体化合物和取代基的活性贡献的和，不论其他位置的取代基的变化如何，每个取代基对生物活性的贡献是固定的并且可以加和，与分子内其他取代基的变化不相关。活性的公式为：

$$\lg(1/C)=\sum a_{ij}X_j+\mu \tag{17-4}$$

式中，$a_{ij}$为正，活性升高，$a_{ij}$为负，则活性降低；$X_j$ 在 $j$ 位置上的取代情况，若此位置有取代物，包括氢，则 $X_j$ 值为 1，没有取代则 $X_j$ 值为 0；$\mu$ 是母体的活性贡献。

### 17.1.2 三维定量构效关系

基于立体结构的构效关系分析可以称为三维定量构效关系，通常包括的是比较分子力场方法（comparative molecular field analysis，CoMFA）；比较分子相似性指数分析法（comparative molecular similarity indices analysis，CoMSIA）；分子形状法（molecular shape analysis，MSA）等。一般最常用的方法是 CoMFA。

CoMFA 方法由 Cramer 于 1988 年提出。此分析方法由一组化合物的构象开始，每个化合物选取一种构象。假定每一个构象都是所对应分子的活性结构。每个构象必须根据一定的方式进行叠合。将叠合好的分子置于空间，产生格点，将探针分子放在每个格点上，计算相互作用能。分析的结果可以用一个 S 的矩阵表示，矩阵的每一列代表分子，每一行代表在每个格点的能量，如图 17-1 所示。

如果有 $N$ 个格点，有 $P$ 个探针分子，则矩阵有 $N\times P$ 行，矩阵的最后一行为分子的活性。活性和分子场之间的相关可以由式(17-5) 决定：

$$\text{activity}=C+\sum_{i=1}^{N}\sum_{j=1}^{P}c_{ij}S_{ij} \tag{17-5}$$

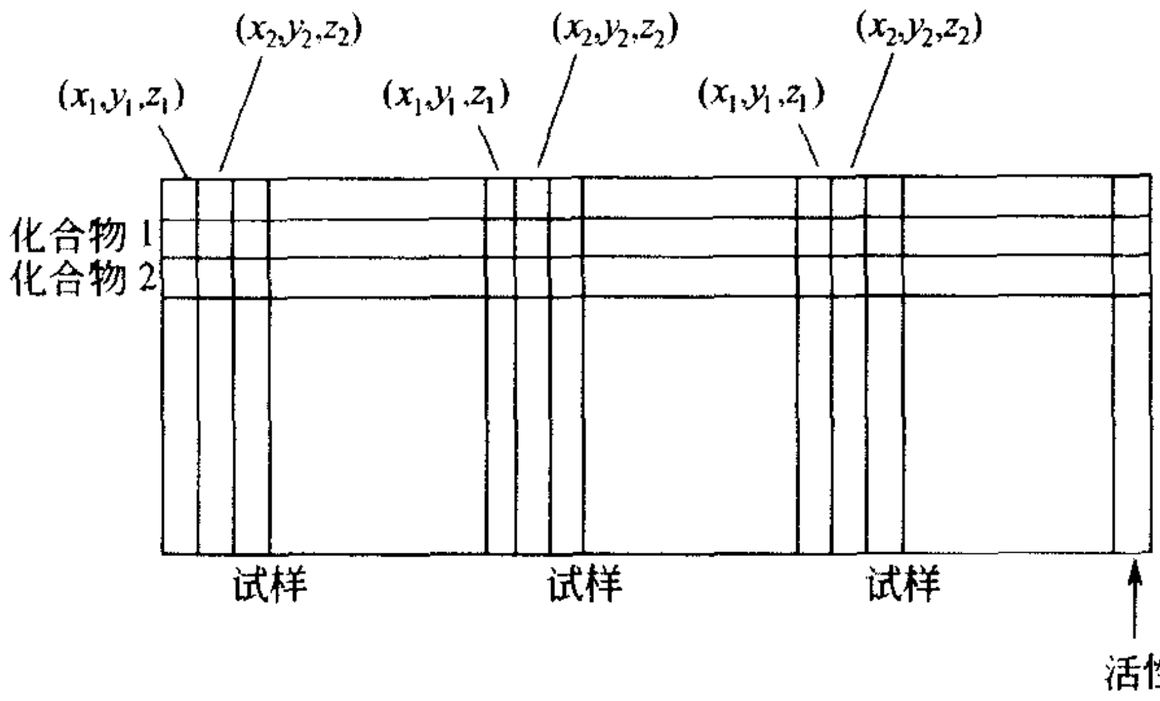



**图 17-1　CoMFA 分析的数据结构**

式中，$c_{ij}$是矩阵中的系数，代表将第 $j$ 个探针分子在第 $i$ 格点。由于每个格点都有一个能量值，静电和立体作用分别由库仑静电作用和 Lennard-Jones 的 6-12 势函数计算得到，将这些能量值作为自变量，求出它们与活性之间的关系即构效关系。

CoMFA 方法由于产生了大量的自变量，通常需要和偏最小二乘法一起使用，才能得到好的结果。CoMFA 近似法在表中的每一行都产生系数，每个格点的系数均表示其对活性的重要性。通常可将相同系数的格子点互相连接成三度空间表面，这些图像有助于了解改变空间体积对结合的影响。目前市面的商业软件 Sybyl 和 Cerius2 都有相应的模块进行 CoMFA 的计算。包括分子构象搜索、构象叠合、分子场分析、相关性计算和预测等过程。Sybyl 公司的叠合软件主要以公共结构为基础进行的。Accelrys 的 Cerius2 则具有多种叠合方式，包括场叠合、柔性叠合和公共结构叠合等。Sybyl 公司的 CoMFA 模块中对分子场分析只包含静电和立体场分析，新近改进模块加了其他的探针。Cerius2 的 QSAR 模块中可以计算四种探针的分子场，还可以根据自己定义的分子计算分子场。Sybyl 公司的 CoMFA 模块采用偏最小二乘法结合多元线性回归分析。Cerius2 的 QSAR 模块引进了多种相关性计算方法，其中遗传算法结合偏最小二乘法，拟合能力非常强。

比较分子相似性指数分析法（CoMSIA）是 Sybyl 公司软件模块，这种方法与比较分子力场方法的最大不同就是在计算能量时采用了与距离相关的高斯函数形式，而不是传统的库仑静电作用和 Lennard-Jones 的 6-12 势函数。一般情况下，CoMSIA 对格点设置的敏感性下降了，计算结构稳定性增加了。CoMSIA 中共定义了五种分子场，包括立体、静电、疏水、

氢键受体、氢键供体等。这五种分子场的计算公式为：

$$A^{q}_{\mathrm{F,K}}(j) = -\sum_{i-1}^{n} \omega_{\mathrm{probe,k}}\omega_{\mathrm{ik}} e^{-ar^{2}_{\mathrm{iq}}} \qquad (17\text{-}6)$$

式中，$i$ 表示分子 $j$ 的原子序号；$\omega_{\mathrm{ik}}$ 表示探针分子的分子场特征；$a$ 为衰减因子；$r_{\mathrm{iq}}$ 表示探针与某个原子 $i$ 之间的距离。

分子形状法（molecular shape analysis，MSA）是 Cerius 软件中的模块，是由 Hopfinger 于 1980 年提出的。目的是找出药物分子空间形状的相似性和差别，以及与活性之间的关系。本方法采用分子形状相关的参数，如：其他分子与参考分子之间的重叠体积、体积差、共同的重叠体积比例、非共同重叠体积比例、分子势场积分差异等。实际计算中不能有效揭示细致信息，结构提示的预测方向性差，应用面受到很大限制。

### 17.1.3 构效关系的数学基础

构效关系其实可以简单地理解为数学上因变量与自变量的函数式。

$$Y = f(x_1, x_2, \cdots, x_n) \qquad (17\text{-}7)$$

式中，$x_1$，…，$x_n$ 等是自变量；$Y$ 是因变量。自变量对应于分子的描述参数（descriptor）、格点的能量等。因变量可以对应于分子的任何性质，如活性、毒性等。为了便于理解，可用将实际计算过程分为以下几个步骤：①因变量的采样；②自变量的采样；③自变量的处理；④相关性分析；⑤相关性模型的确证等内容。下面结合应用中的具体情况加以分析。

**（1）因变量的采样** 活性指标有：$IC_{50}$、$ED_{50}$、抑制率、有效率，指标具体值等。毒性指标如：$LD_{50}$、毒性反应率等。药代性质有：Cmax、AUC、$T_{1/2}$等。理论上各种定量指标都可以作为因变量。实际计算中，根据具体情况，有些指标需要变换。通常的变换是求倒数、求对数或倒数的对数、负对数等。判断变换是否合理主要是看变换能否使变量的分布均匀化。挑选训练集的样本时，因变量方面需要注意的问题有以下一些。

① 变量的分布 样品应尽量均匀分布，或者数据变换以后分布均匀。

② 变量的误差 实验是有误差的，尤其是接近检测限时候，噪声的干扰不能忽视。在取数据作为训练时，要选择准确的数据。不同实验室之间的定量数据存在系统误差，多批不同来源的数据尽量不要一起作为训练集使用。必须一起使用时，要有可靠性的保障。变量的意义：应该选择比较重要的指标作为变量，特别需要注意的是变量的范围要符合专业判断。

**(2) 自变量的采样** 现在的计算软件都提供许多分子特征符可以作为自变量，MOE 软件中采用的特征符非常多，Cerius2 也提供了许多的分子特征符，符号及其意义如表 17-1～表 17-6 所示。

**表 17-1 通用型分子特征描述符**

| 名 称 代 号 | 定 义 |
|---|---|
| diameter | 距离矩阵的最大值，对应于分子中最远两个原子的距离 |
| weight | 分子量 |
| zagreb | Zagreb 指数，所有重原子的直径平方的总和 |
| SlgP | 辛醇/水的分配系数的对数值 |
| VSA | 范德华表面积 |
| VDW_area | VDW 表面积 |
| VDW_vol | VDW 体积 |
| vol | 用格近似计算的范德华体积 |
| lgP(o/w) | 辛醇/水的分配系数的对数值 |
| a_acc | 氢键受体的原子 |
| a_don | 氢键供体的原子数 |
| vsa_acc | 氢键受体的 VDW 表面积总和近似值 |
| vsa_don | 供氢键供体 VDW 表面积总和的近似值 |

**表 17-2 连接指数型分子特征描述符**

| 名称代号 | 定 义 | 名称代号 | 定 义 |
|---|---|---|---|
| chi0v | 原子的 0 级价键连接指数 | Kier1 | 一级 Kappa 形状指数 |
| chi0v_C | 碳原子的 0 级价键连接指数 | Kier2 | 二级 Kappa 形状指数 |
| chi1v | 原子的 1 级价键连接指数 | Kier3 | 三级 Kappa 形状指数 |
| chi1v_C | 碳原子的 1 级价键连接指数 | KierA1 | 一级 a 修正的形状指数 |
| chi0 | 原子的 0 级连接指数 | KierA2 | 二级 a 修正的形状指数 |
| chi0_C | 碳原子的 0 级连接指数 | KierA3 | 三级 a 修正的形状指数 |
| chi1 | 原子的 1 级连接指数 | KierFlex | Kier 分子柔性指数 |
| chi1_C | 碳原子的 1 级连接指数 | | |

**表 17-3 疏水型分子特征描述符**

| 名 称 代 号 | 定 义 |
|---|---|
| a_aro | 芳香原子数 |
| PEOE_VSA_FHYD | PEOE_VSA_HYD 除以总面积 |
| PEOE_VSA_HYD | 总疏水 van der Waals 表面积，以电荷小于 0.2 的原子计算 |
| Q_VSA_FHYD | PEOE_VSA_HYD 除以总面积 |
| Q_VSA_HYD | 总疏水 van der Waals 表面积，以电荷小于 0.2 的原子计算 |
| a_hyd | 疏水原子数 |
| vsa_hyd | 疏水原子 VDW 表面积总和的近似值 |
| ASA_H | 水可及的总疏水原子(以电荷小于 0.2 的原子计算)的表面积 |

**表 17-4　极性型分子特征描述符**

| 名称代号 | 定义 |
|---|---|
| PEOE_VSA_FPOL | Q_VSA_POL 除以总面积 |
| PEOE_VSA_POL | 总极性原子的 VDW 表面积，以电荷大于 0.2 的原子计算 |
| Q_VSA_FPOL | Q_VSA_POL 除以总面积 |
| Q_VSA_POL | 总极性原子的 VDW 表面积，以电荷大于 0.2 的原子计算 |
| apol | 极性原子总数 |
| vsa_pol | 极性原子 VDW 表面积总和的近似值 |
| ASA_P | 水可及的总极性原子(以电荷大于 0.2 的原子计算)的表面积 |
| dipole | 总偶极矩 |
| dipoleX | 偶极矩在 $X$ 方向的分量 |
| dipoleY | 偶极矩在 $Y$ 方向的分量 |
| dipoleZ | 偶极矩在 $Z$ 方向的分量 |

**表 17-5　正电荷型分子特征描述符**

| 名称代号 | 定义 |
|---|---|
| PEOE_PC+ | 原子正的部分电荷总和 |
| PEOE_RPC+ | 原子相对正的部分电荷总和 |
| PEOE_VSA_FPOS | 总的正电荷 VDW 表面积除以总面积 |
| PC+ | 原子正的部分电荷总和 |
| Q_PC+ | 原子正的部分电荷总和 |
| Q_RPC+ | 原子相对正的部分电荷总和 |
| Q_VSA_FPOS | 总的正电荷 VDW 表面积除以总面积 |
| Q_VSA_PPOS | 总的正极性 VDW 表面积，以电荷大于 0.2 的原子计算 |
| RPC+ | 原子相对正的部分电荷总和 |
| ASA+ | 总带正电原子的水可接近的表面积 |
| CASA+ | 正电荷加权的水可接近表面积 |
| FASA+ | ASA+/ASA 比值 |
| FCASA+ | CASA+/ASA 比值 |

**表 17-6　负电荷型分子特征描述符**

| 名称代号 | 定义 |
|---|---|
| PEOE_PC− | 原子负的部分电荷总和 |
| PEOE_RPC− | 原子相对负的部分电荷总和 |
| PEOE_VSA_FNEG | 总的负电荷 VDW 表面积除以总面积 |
| PC− | 原子负的部分电荷总和 |
| Q_PC− | 原子负的部分电荷总和 |
| Q_RPC− | 原子相对负的部分电荷总和 |
| Q_VSA_FNEG | 总的负电荷 VDW 表面积除以总面积 |
| Q_VSA_PNEG | 总的负极性 VDW 表面积，以电荷大于 0.2 的原子计算 |
| RPC− | 原子相对负的部分电荷总 |
| ASA− | 总带负电原子的水可接近的表面积 |
| CASA− | 负电荷加权的水可接近表面积 |
| FASA− | ASA−/ASA 比值 |
| FCASA− | CASA−/ASA 比值 |

上面的分子特征符一般是二维构效关系中常见的参数。在三维构效关系计算中，从探针分子在各格点上的计算能量中，选择变化较大的格点作为自变量。这个过程通常是软件完成的。

研究者还可以自己设定相关参数作为自变量，如：键长、电子分布、原子间距离等。理论上任何与该分子相关的参数都可以作为自变量。然而，药物设计的实际计算中必须注意几个问题。

① 变量的意义　变量的选择应尽量结合专业，并且有明确的物理意义，这样对于预测的结果也有改造的可行性，具有指导意义。

② 样本的数量　为了达到有意义的统计结果，一定要有足够的资料。一般执行回归分析时，每一个参数最少需要五个化合物的数据。

③ 变量的分布　所选用的化合物应符合使参数值均匀分布，且参数间不具相关性的要求。变量的分布范围应符合专业的要求。

除了这些外，还有预测结果考虑到与因变量的实际相关性。

**(3) 自变量的处理**　在多元的相关性分析中，自变量的选择是个很大的困难，当变量间有很强的相关性或变量的数目超过数据的数目时，常用的方法是数学上的降维方法，有主成分分析法（principal components regression，PCR）和偏最小二乘法（partial least squares，PLS)。遗传算法和神经网络算法本身具有较强的处理多自变量的功能，实际计算软件中自变量的处理经常与这些算法结合起来，不需要额外的处理。了解这些方法的基本的原理有助于更好地选择如何处理自变量。

主成分分析方法通过选取对活性影响最大的几个主成分来建立定量构效关系。基本算法是一种空间变换方式，让经过标准化的原始变量线性组合成若干个矢量，要求它们之间相互正交，且第一个变量能反应样本间自变量的最大差异，第二个次之，依此类推。基本原理是，原变量矩阵 $X$ 可以表达为主成分矩阵 $T$，由在本征矢量上的投影所得主成分与矩阵的本征矢量一一对应，即 $XP=T$。

主成分回归可以有效解决共线问题，同时由于去掉了不太重要的主成分，因而可以削弱噪声产生的影响。在主成分分析法中，所处理的仅为 $X$ 矩阵，而对于矩阵 $Y$ 中的信息并未考虑。

偏最小二乘法也是一种主成分分析法，与 PCA 相似，也能排除原始变量的相关性，过滤自变量的噪声，还能过滤因变量的噪声，描述模型所需的特征变量数目比 PCA 少，预报能力更强，更稳定。在回归建模中已经取代了大部分的主成分分析法。其物理思想是将因变量 $Y$ 中不利于建模

的噪声排除掉，将有用的信息部分通过其负载与得分两个矢量和自变量 $X$ 的负载和得分相互交叉回归，以求得兼顾 $X$ 和 $Y$ 都拟合最好的负载 PLS 成分。

选定自变量后，自变量和因变量之间的相关性的计算方法也有很多，常用的有多元线性回归、遗传算法和神经网络算法等等。神经网络算法具有很强的拟合能力，由于它不能提供明确的数学方程，并且过于良好的拟合容易产生假阳性，拟合的结果指导意义不明确，因此，在实际应用中很难提供方向性的预测。

常用以推导 QSAR 方程式的方法为多重线性回归法（multiple linear regression method）。此方法利用最小平方误差拟合（least square fit）找寻最佳的组合系数。以最简单的线性关系式如：

$$y=mx+c \tag{17-8}$$

式中，$y$ 是因变量（dependent variable），即观察的数值；$x$ 是自变量（independent variable）即参数。例如 $y$ 为活性，而 $x$ 为 $\lg P$。回归分析的目的在于找出 $m$ 及 $c$ 值，使由拟合方程式所计算的推测值与度量值间有最小的误差平方和，即：

$$m=\frac{\sum_{i=1}^{n}(x_i-\langle x\rangle)(y_i-\langle y\rangle)}{\sum_{i=1}^{n}(x_i-\langle x\rangle)^2},\quad c=\langle y\rangle-m\langle x\rangle$$

回归方程式经过点（$\langle x\rangle,\langle y\rangle$），$\langle x\rangle$ 和 $\langle y\rangle$ 代表自变量和因变量的算数平均值。拟合的好坏通常由指标数 $R^2$ 决定。$R^2$ 的计算法如下：

$$SS_{\text{mean}}=\sum_{i=1}^{n}(y_i-\langle y\rangle)^2,\quad SS_{\text{error}}=\sum_{i=1}^{n}(y_i-y_{p,i})^2$$

$$R^2=\frac{SS_{\text{mean}}-SS_{\text{error}}}{SS_{\text{mean}}} \tag{17-9}$$

$R^2$ 的范围为 0.0～1.0。若 $R^2=0.0$ 表示拟合的程度最佳，若 $R^2=1.0$，则表示拟合的效果最差。

**（4）相关性分析** 自变量和因变量的相关性，即变数 $x$ 和 $y$ 的相关性可以式(17-10) 估计：

$$r=\frac{\sum_{i=1}^{n}(x_i-\langle x\rangle)(y_i-\langle y\rangle)}{\sqrt{\left[\sum_{i=1}^{n}(x_i-\langle x\rangle)^2\right]\left[\sum_{i=1}^{n}(y_i-\langle y\rangle)^2\right]}} \tag{17-10}$$

式中，$\langle x\rangle$，$\langle y\rangle$ 代表 $x$，$y$ 的算术平均值；$n$ 是化合物的数目。若相关系数 $r=1.0$，表示 $x$ 和 $y$ 完全的正相关，即 $x$ 和 $y$ 坐标在一条直线且斜率是正值。若 $r=-1.0$，表示 $x$ 和 $y$ 是完全的负相关，即 $x$ 和 $y$ 坐标在一条直线且斜率是负值。若 $r=0.0$ 表示 $x$ 和 $y$ 并无任何相关性。线性回归法可扩展到两个自变量以上，就成了多元线性回归。

遗传算法（genetic algorithm）也是一种相关性分析的方法。遗传算法是一种仿生优化算法，通过模仿孟德尔的遗传变异理论，在迭代过程中保存父代中的已有模式，寻找更好的模式；通过模仿达尔文的进化理论，“优胜劣汰，适者生存”，使好的模式具有更多的机会在子代中出现。从数学上理解，遗传算法是一类随机优化搜索算法，基本原理如下：使用二进制串表示个体，若干个个体构成种群，起始种群随机产生，然后以适应度函数（fitness function）对每个个体评分（适应度），根据每个个体的适应度按确定的机制选择一部分个体作为父代，适应度高的个体具有更多机会作为父代参与繁殖过程，未选中的个体自然消亡，父代通过基因重组操作产生新的子代。所有子代按一定的概率进行变异操作，最终形成新一代种群。如此不断迭代，模拟种群的繁殖过程，直到满足停止条件。遗传算法是一种快捷、简便、容错性强的算法，可以智能化地选择合理的变量组合以得到最佳的构效关系模型。

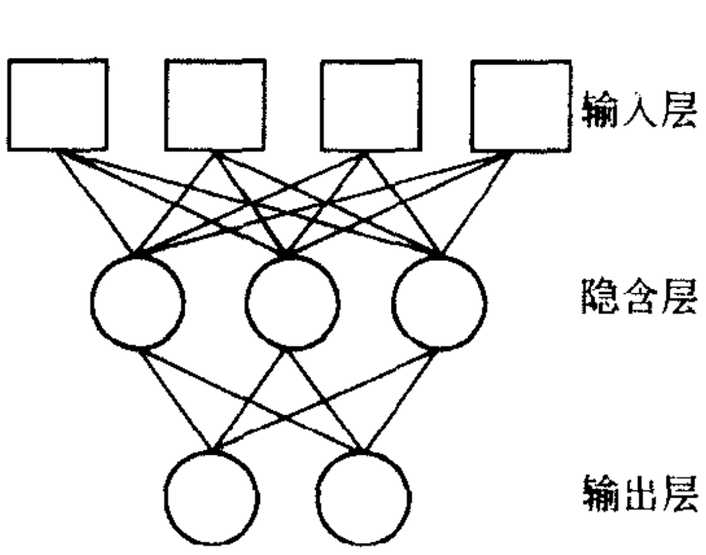



**图 17-2　神经网络的四个输入、三个隐藏和两个输出节点**

神经网络算法（neural network algorithm）是一种产生结构-活性关系的方法。如图 17-2 所示，神经网络在每层有不同的单元，且不同层间两两单元互相连接。

神经网络的每一个单元代表一个0～1的实数。单元的状态（数值大小）由所连接的单元及连接的强度所决定。为了达到理想的作用，必须先行训练神经网络。训练过程中，有许多的输入和输出，每一个输入将会顺着层与层的连接至下一层。传输过程中将调整连接强度直到输入和输出间有最好的一致性。训练完成的网络可用以预测各种性质。

在 QSAR 中输入代表不同参数的值，训练整个网络再现实验所测量的活性。当完成网络训练后，未知化合物的活性可以借由输入适当的参数求得。神经网络算法的拟合能力很强，实际计算中任何数据都能得到较好

的拟合，但并不能保证预测结果的真正可靠性，对模型的验证尤其重要，特别要注意的是训练集和测试集之间的分布的可比性。

**(5) 相关性模型的确证** 对于一个自变量和因变量之间的相关性模型，其合理性必须经过评价后才能应用到预测中去。从统计的角度来看，比较常用的指标是相关系数 $r$［式(17-10)］和残差指标 $R^2$［式(17-9)］。但模型本身的拟合优劣只是一个方面，更关键的是必须进行抽一法交叉验证和测定集的验证。

抽一法交叉验证的过程是把训练集的每个样品删去后，由其他的样本拟合方程，再预测删去样本的结果，将预测结果与实测结果进行分析。测试集的验证是把一部分样本作为训练集，留出一部分样本作为测试集，以训练集的模型来预测测试集的样本，与实测的结果进行比较。通常判断指标是相关系数和残差。这种方法特别要注意测试集样本的分布要比较均匀，并且尽量与真正需要预测的样本比较可比。

构效关系的分析的主要目的是通过已知的样本建立相关性后来预测未知的样本的性质，所以相关模型的应用才是真正的目的。首先要把需要预测的分子的自变量求出来，再代入相关模型或方程求出活性或其他性质。在实际计算中，非常容易被忽略的问题就是相关性的方程或模型的每个自变量是有适用范围的，多维分析时自变量的范围不容易理解，如果需要预测样本的参数不在训练集样本的参数范围，则预测结果很难保证有可靠性。

## 17.2 药效团模型

当受体蛋白质的结构未知的情况下，根据已知系列小分子的活性和结构进行定量构效关系的研究是重要的方法，但是人们在很多情况下还希望对大量的分子进行对比，这时就需要药效团模型了。通常的过程是先把活性较好的分子叠合起来，求出共同的活性构象，抽象成几个点和距离的模型，根据这个模型对数据库进行搜索，来寻找可能的活性化合物。不像对接计算中有蛋白质模板，药效团模型的建立有很多的人为因素，如果模型的建立不正确，数据库的搜索就没有意义了，实际应用需要特别注意药效团模型的合理性，本方法的成功应用通常需要研究者具有很深的专业背景，初级人员不容易用好这个方法，本节简单介绍相关

的原理。

## 17.2.1 药效团模型的建立

药效团模型的建立需注意以下几点。

**(1) 药效团**（pharmacophoric groups）**的特征元素** 理想的药效团是最关键的活性基团和骨架的组合，通常以特征元素的组合来表示。常见的特征元素如：氢键供体、氢键受体、电荷中心、疏水中心和芳环中心等。氢键供体是指杂原子及连接的氢，如羟基、氨基、酰胺、巯基等。氢键受体是指含有孤对电子的杂原子，如 N、O、S、F 等。

**(2) 几何约束** 上述特征元素在三度空间展现的相互几何关系，是药效团的另一个重要的因素。可以距离、角度和二面角来表示，距离限制是最常用的约束形式。

**(3) 药效团识别的基本步骤**

① 选择作用机制相似、活性高的化合物作为药效团模板，刚性的分子有利于确定药效团模型。然后定义每个分子的药效特征元素。

② 对所选的分子进行构象分析，通常的软件会选择几个构象进行叠合，太多的构象对建立药效团带来困难。由于药效与受体之间存在互相诱导，并不一定是最低能量的构象与活性相关。

③ 接下来需要对各种分子叠合进行叠合和药效团的识别。分子叠合是药效团识别过程中最关键的步骤，与三维构效关系中的叠合方法是一致的。Sybyl 公司的叠合软件主要以公共结构为基础进行的。Accelrys 的 Cerius2 具有多种叠合方式，包括场叠合、柔性叠合和公共结构叠合等。对于刚性分子，活性构象接近最低能量构象，比较容易确定。对于柔性分子，最理想的是用每个分子的活性构象进行叠合，实际上无法知道什么样的构象是活性构象，所以多个构象叠合的过程是寻找不同分子的类似构象的过程，最佳的叠合构象可能与活性构象比较接近。

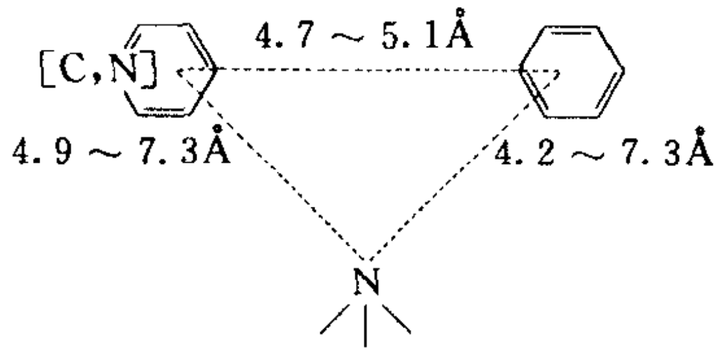



**图 17-3 抗组织胺剂三度空间药效团**

④ 最后是药效团的确证，通过叠合可能得到数种的药效团模型，需要再利用实验数据或其他方法筛选出适当的药效团，专业的经验和判断对确

定药效团是非常重要的，实在不能确定时，采用多个药效团系统。

图 17-3 所示，是一个抗组胺剂的三度空间药效团。包含两个苯环和一个三级氮，这些基团是构成抗组胺剂最重要的基团。

### 17.2.2 药效团识别的软件

目前主要应用的商业药效团识别软件有 Receptor、Apex-3D、DISCO、GASP、CATALYST。

Receptor 是 Cerius2 中的一个模块，主要基于 Marshall 等 1993 年提出的活性类似物方法（active analogue approach，AAA），根据具有结构多样性的活性化合物和非活性化合物的构效关系，寻找对活性起主要作用的基团、构象，构建三维药效团模型。

Apex-3D 是 Biosym 公司推出的系统，Biosym 与 MSI 合并后主要发展 CATALYST 系统，CATALYST 在功能上更强大，可以取代 Apex-3D。Apex-3D 是基于 Valerby Goldender 的逻辑结构分析方法。首先将分子结构参数数值化，对活性和非活性分子进行构象分析及构象叠合，分子叠合的过程是通过比较结构参数实现的，叠合后具有相似性质的子结构在空间上重叠在一起。叠合过程可以产生许多药效团模型，每个模型经过得分来评价。

DISCO 是 Martin 1993 年提出的方法，在 Sybyl 中实现了商业化，主要的建立药效团的步骤是：选择具有活性的分子，尽量体现结构的多样性；进行构象搜索，得到低能构象；采用子结构搜索的方法来定义分子可能的药效元素，采用构象少的结构为参考，与其他分子的每个构象进行叠合，叠合最好的构象作为活性构象，采用族类检查算法来识别叠合构象中药效特征元素的最大公共子结构，得到药效团模型。

GASP 是 Gareth Jones 等提出的方法，目前在 Sybyl 软件中已实现商业化，主要的特点是采用遗传算法来实现药物分子间的柔性叠合，从而找到药效团模型，另外通过范德瓦耳斯能力得分、公共体积得分和相似性得分来评价药效。本方法计算效率很高，通常分子计算相对较准，但定义的药效特征元素较少，不能完全表示的结构可能存在较大的误差。

CATALYST 是一个独立的软件，是基于药效团模型的药物设计系统，是 Accelrys 公司的产品，建立药效团的过程与 DISCO 方法相近，只是药效特征元素更完备，定义的内容包括化学功能、坐标位置、许可误差范围和权重等。药效团的构建以 Occam's Razor 剃刀规则的量化延伸为基础。

## 17.3 分子对接计算

蛋白质结构通常不作为药物设计的内容，而其结构是设计的模板，因此，在设计药物时首先要考虑的是作为模板的蛋白质的结构合理性。基于一种合理的结构设计才有意义。以蛋白质受体为依据进行药物设计的方法主要有分子对接（molecular docking）和从头设计（de Novo）两种。

通常药物-受体结合时，存在静电相互作用、氢键相互作用、范德瓦耳斯相互作用和疏水相互作用，作用的方式需要遵循以下互补匹配规则：几何形状互补匹配、静电相互作用互补匹配（正电荷对应负电荷）、氢键相互作用互补匹配（氢键供体对应氢键受体）、疏水相互作用互补匹配（疏水区对应疏水区）。对接计算就是搜索分子与受体结合位点之间的匹配构象的算法。理想的分子对接方式是药物分子与其受体都应柔性的，目前的对接方法通常仅考虑药物分子的结构的变化，而将受体分子视为完全刚性体。分子对接的目的是得到受体与小分子的最佳结合方式，计算中的两个实际问题是：如何产生快速搜索全构象空间的结合模式，如何正确地评价结合模式。

目前已有多种方法实现了商业化，代表性的分子对接方法见表 17-7。

**表 17-7　代表性的分子对接方法**

| 名　　称 | 构象搜索方法 | 结合评价方法 |
|---|---|---|
| FlexX | 片段生长法 | 半经验自由能 |
| LigandFit | 蒙地卡罗模拟 | 半经验自由能 |
| Glide | 系统搜索 | 半经验自由能 |
| Gold | 遗传算法 | 半经验自由能 |
| Affinity | 蒙地卡罗/MM/MD | 分子力场 |
| Autodock3. 05 | 遗传算法 | 半经验自由能 |
| Dock5 | 片段生长法 | 分子力场 |
| ICM-Dock | 随机全局优化 | 半经验自由能 |
| Fred | 系统搜索 | 半经验自由能 |

下面对文献中比较常用的四种方法简单介绍原理和计算过程。

### 17.3.1 FlexX

FlexX 是德国科学家 Matthias Rarey 等发展的方法，是 Sybyl 软件中的一个重要的商业模块，对接过程如下所述。

① 确定核心结构片段，核心片段是指对分子间相互作用起决定性的基团，其构象应尽量少，然后把配体分子或分为多个片段。

② 放置核心结构，采用形态聚类算法，核心片段和活性位点都看成刚性结构，核心结构放置在活性口袋的过程就相当于配体的三个作用点叠合在活性位点的三个相互作用点上，类似于三角形的匹配。找出所有匹配的三角形，转换坐标，检查配体和受体是否碰撞，去掉不合理的核心结合取向。

③ 当核心结合位置确定后，配体分子的其他部分片段依次生长，在核心结构上，采用树形搜索算法，片段依次连接。

FlexX 的评价函数类似于 Böhm 半经验自由能方程。（Böhm H A. J comput-Aided Mol Des, 1992, 6: 61; Böhm H A. J comput-Aided Mol Des, 1992, 6: 593）

### 17.3.2 LigandFit

这是 Accelery 公司的产品，是 Cerius2 的一个对接模块。首先发现活性位点后，然后进行配体分子的构象搜索，进行对接，计算评分。LigandFit 采用蒙地卡罗模拟来产生配体分子的构象，键长和键角不动，改变二面角，每产生一个新构象后都进行配体小分子与受体蛋白质之间的对接。对接分两步进行，首先结合位点与配体分子的非质量加权的主惯量矩（PMI）按方程进行比较，当 $Fit_{PMI}$ 的值大于阈值或差于上个计算的构象，这个构象不再进行下面的对接计算。

$$ratio_{xy} = P_x / P_y , ratio_{xz} = P_x / P_z , ratio_{yz} = P_y / P_z$$

$$Fit_{PMI} = \sqrt{\Delta ratio_{xy}^2 + \Delta ratio_{xz}^2 + \Delta ratio_{yz}^2} \qquad (17\text{-}11)$$

另外，如果 $Fit_{PMI}$ 好于上步的结果，配体按照 PMI 放进活性口袋，因为 PMI 为标量，配体与受体的结合存在四种可能，计算对每一种位置的对接分数，另外还可以应用刚体优化来计算各种位置的对接分数，如图 17-4。LigandFit 的评价函数有四种：ligscore、PLP、PMF、LUDI。

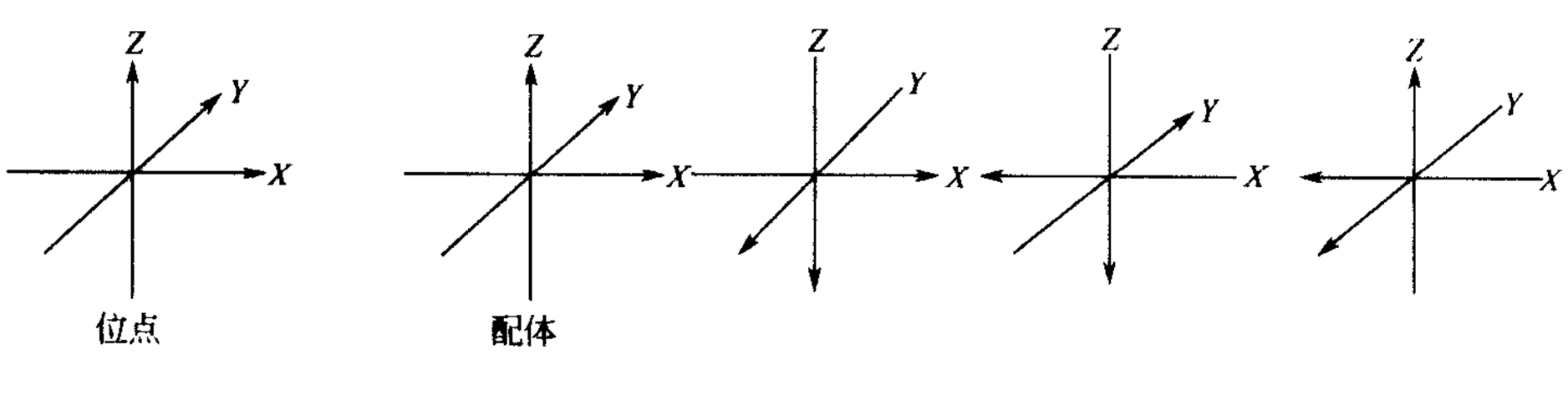



**图 17-4　受体与配体的主惯量矩匹配方式示意**

因为 ligscore 是一种快速的简化的预测蛋白质与配体结合常数的方法，以式(17-12) 表示：

$$pK_1 = A - B \times \mathrm{VDW} + C \times C_{+\mathrm{pol}} - D \times \mathrm{Totpol}^2 \qquad (17\text{-}12)$$

式中，VDW 是 LJ 势精确计算的，也可以采用 grid-based 能力近似法，kcal/mol；$C_{+\mathrm{pol}}$是蛋白质与配体之间的覆盖的极性表面积计数，涉及蛋白质与配体之间的吸引相互作用的面积，$Å^2$；$\mathrm{Totpol}^2$ 是蛋白质与配体之间覆盖的极性表面积平方的计数，涉及蛋白质与配体之间的吸引和排斥相互作用的面积，$Å^4$。

Piecewise Linear Potential（PLP）是一种快速、简单的对接评价函数，与蛋白质-配体结合常数有良好的相关性。较高的负 PLP 分数表明结合较好（较高的结合常数）。

Potential of Mean Force（PMF）是利用 3D 数据库进行的统计分析方法，提供快速和精确的蛋白质和配体的结合自由能。评价函数定义为蛋白质-配体各原子对相互作用自由能的总和。较高的负 PMF 值表明结合较好。

LUDI 评分是基于 Böhm 半经验自由能或其改进计算分数。(Böhm H A. J comput-Aided Mol Des, 1992, 6: 61; Böhm H A. J comput-Aided Mol Des, 1992, 6: 593)

### 17.3.3 Autodock 3.05

Autodock 是 Scripps 的 Olson 科研小组开发的分子对接软件包，对学术用户是免费的，目前版本为 3.05，采用模拟退火和遗传算法来寻找受体和配体的最佳结合位置，用半经验自由能计算方法来评价受体和配体之间的匹配情况。Autodock 只能进行一个分子与一个蛋白质的对接计算，不能进行数据库对接，没有平行化功能。不需要预先知道活性位点。

#### (1) 对接过程

① 准备蛋白质和配体　正确地将蛋白质结构转化成 pdbqs 格式，在 pdb 格式的基础上，加上电荷和溶剂效应。将配体分子转换成 pdbq 格式，在 pdb 基础上加上电荷。并定义搜索的二面角和根片段。

② 计算格点能量　创建 GPF（grid parameter file）和 DPF（docking parameter file）文件，计算格点空间的能量图。格点上保存的是探针原子和受体之间的相互作用能量，通常配体有 C、O、H 三种原子类型，那么

每个格点上需要用这三个探针原子计算范德瓦耳斯作用。静电相互作用采用静电势格点保存。计算去溶剂化能时，格点仅保存了C原子数据。

③ 对接计算　最早的版本中采用了模拟退火来寻找结合方式，3.0以后的版本中加上了拉马克遗传算法（LGA），把遗传算法和局部搜索结合在一起，遗传算法用于全局搜索，局部搜索优化能量。

**（2）评价函数**

$$\Delta G = \Delta G_{\mathrm{VDW}} \sum_{i,j} \left( \frac{A_{\mathrm{ij}}}{r_{\mathrm{ij}}^{12}} - \frac{B_{\mathrm{ij}}}{r_{\mathrm{ij}}^{6}} \right) + \Delta G_{\mathrm{H\text{-}Bond}} \sum_{i,j} E(t) \left( \frac{C_{\mathrm{ij}}}{r_{\mathrm{ij}}^{12}} + \frac{D_{\mathrm{ij}}}{r_{\mathrm{ij}}^{10}} \right) + \Delta G_{\mathrm{ele}} \sum_{i,j} \frac{q_{\mathrm{ij}}}{\varepsilon(r_{\mathrm{ij}}) r_{\mathrm{ij}}} + \Delta G_{\mathrm{tor}} N_{\mathrm{tor}} + \Delta G_{\mathrm{sol}} \sum_{i,j} (S_{\mathrm{i}} V_{\mathrm{j}} + S_{\mathrm{j}} V_{\mathrm{i}}) \mathrm{e}^{(-r_{\mathrm{ij}}^2 / 2\delta^2)} \tag{17-13}$$

式中，$\Delta G_{\mathrm{VDW}}$，$\Delta G_{\mathrm{H\text{-}Bond}}$，$\Delta G_{\mathrm{ele}}$，$\Delta G_{\mathrm{tor}}$和 $\Delta G_{\mathrm{sol}}$都是半经验参数，通过拟合得到，$N_{\mathrm{tor}}$是配体在对接后被约束的可旋转键的数目。

## 17.3.4　DOCK 5.2

DOCK程序是Kuntz小组开发的对接软件，对学术机构是免费的，可以应用于数据库的对接，具有平行化计算能力，应用比较广泛。

**（1）对接操作过程**　具体的对接操作主要为：先将小分子或分子库转换成mol2格式备用，蛋白质结构修改正确后，保存pdb格式，计算MS表面积，进行活性位点特征描写，再建立评分格点，最后进行对接计算。

**（2）评分函数**　DOCK5实行分级的评分策略，一个主要的评分级贯穿所有的评分函数。任何DOCK的评分函数可以被选为初级或二级的评分函数，初级函数用于对接过程中的刚体优化等步骤，通常需要多次调用。二级函数用于最后的优化、评分和排列。DOCK5版本中有基于AMBER力场的分子间能量评分（范德华＋静电）、接触评分、DOCK4中实现的碰撞过滤分以及GB/SA评分（与SDOCK相同）。

## 17.3.5　分子对接计算的注意点

分子对接计算现在是基于受体结构进行药物设计的最主要方法，通常的虚拟筛选主要是指利用分子对接对小分子数据库进行的批量计算（广义的虚拟筛选可以把药物的理论设计都包括进去）。因此，分子对接的实用性较强，应用面很广，已经成为一种常用的工具。在实际应用中需要注意下列问题。

**(1) 小分子的问题** 现在大多的方法都把蛋白质结构固定，配体小分子构象可以改变，理论上对接前小分子不需要考虑构象，可以交给对接过程来处理。通过测试 FlexX，Autodock 等软件，以小分子搜索后的各构象与同一个受体对接，计算结果差异不小，说明起始构象对对接结果有较大影响。所以建议在对接前适当搜索小分子的构象空间。另一个值得注意的问题是分子在人体内可能存在活性形式的变化，如：酯键、某些酰胺键、某些饱和杂原子上的取代基等在体内很快水解，有些结构在体内很快还原或氧化。代谢后的结构与原结构不同，如果在体内主要以代谢物形式存在，则对接时应主要以代谢物的结构进行对接。实际操作中大量结构的手工修改有困难，可以借助一些数据库操作的软件（$C_2$、Sybyl、MOE 等都有相关的软件）进行。另外，还需要根据相应的软件的要求，对分子电荷、氢进行处理。

**(2) 蛋白质的构象问题** 蛋白质受体的最大问题就是以什么标准来判断选择的蛋白质是合理的，所以在确定某个蛋白质可以作为对接的受体前，需要进行一些蛋白质结构的研究，如果报道的晶体或 NMR 数据较多，可以对多个结构进行叠合，考察活性位点的变异情况，如果位点比较一致，位点附件的序列比较保守，则这类结构的可靠性高一些。如果变异较大，需要进行更多的蛋白质结构研究，确定正真正活性相关的构象。选择一些阳性药和阴性药来对接，去掉假阳性较高和假阴性较高的蛋白质结构通常是比较有用的办法。

上面的讨论是基于蛋白质结构刚性的情况来进行的，实际存在于机体中的蛋白质本身与晶体结构存在差异，而且蛋白质之间的结构诱变、配体与蛋白质之间的结构互适应的情况很多，这时蛋白质的结构变化很大，基于晶体结构的对接就无能为力了。所以在实际计算中要多查阅有关蛋白质的生物学研究的结果，印证受体结构的合理性。

**(3) 对接的问题** 最理想的对接应该能够快速搜索到所有的构象空间，并且实现柔性对接，实际计算中每个软件都很难做到，每个软件各有特色，都有比较成功的应用例子，也都有明显的不足，在可能的情况下，尽量采用多个软件进行评价，减少结合模式搜索误差。

对于一个数据库的对接计算是非常耗时间的，每个分子的计算量增加一点，100 万个分子的计算就增加很多，计算能力是限制对接应用的实际困难。人们一方面在改进软件，实现并行化计算，同时计算机的计算能力也在不断提高，虚拟筛选成了高性能计算的重要内容。我们除了增加计算能

力外，对分子进行预筛可以有效降低计算量，通常可以结合专业知识设置一些标准，如分子量范围，lg*P* 范围，氢键数范围等等，通过这些指标虑掉许多分子，从而大大提高筛选效率。“五倍率”是一种最常用的初筛标准。

除了搜索结合模式的正确性、对接的效率，还有评分的正确性问题。尽管各软件采用了各种方法来提高评分的正确性，仔细分析一下软件的计算原理就可以发现，有的软件不考虑蛋白质的氢，或只考虑极性氢，有的软件只考虑形式电荷，有的不考虑离子、关键的辅酶等。每个软件都有自己的简化方式，所以评分的精确性与简化计算之间的差距必须要搞清楚，才能知道每个软件的评分的实际可靠性。通常评分是一个相对定量的指标，不能精确反应实际的相互作用能，在应用中如果要精确定量评价，还应该结合分子动力学等其他方法来进行评价。

## 17.4 数据库搜索

对已有的三维数据库进行搜索，是药物设计中常用的方法，通常可以利用分子对接和药效团对数据库进行大规模的筛选。就像利用某个活性模型对化合物进行筛选一样，上述过程特别是分子对接计算被形象地称为虚拟筛选。基于蛋白质受体或基于药效团的数据库搜索，尽管使用软件不同，都是对已有小分子库进行立体结构的搜索，所以一起讨论。

**(1) 小分子的立体结构数据库** 剑桥结构数据库（Cambridge structure database，CSD)。这个数据库约含有 10 万种化合物的 X 射线衍射的晶体结构。

NCI 数据库是美国国家研究所的数据库，公开发布的是 24 万多个化合物结构，可以免费下载。

ACD 数据库（available chemicals directory）是 MDL 公司的产品，共有约 200 多万个化合物，所有的化合物都可以找到相关的购买信息，是药物虚拟筛选中首选的数据库。

一些大的制药公司都有自己的数据库，一般不对商业公开。另外还有没有立体结构的数据库，需利用计算机计算程序产生低能量可能的分子结构。最通用产生结构的程序为 CONCORD 程序。

**(2) 分子对接法与数据库的搜索** 已知药物的受体后，分子对接法

是非常有用的工具，前面已经比较详细地介绍了对接方法的原理。把三维小分子库中的每个分子与蛋白质受体进行对接，是常用的药物虚拟筛选的手段，从技术上没有更多的特点，只是对接一个分子与多个分子的差别，主要是计算量的差别。像 ACD 这样 200 多万个化合物的分子库，对接一遍需要很长的时间，对接的效率和计算能力就成了一个比较关键的因素。

**(3) 基于药效团的三维数据库的搜索** 得到三度空间药效团结构后，第一种方法是进行数据库搜索，找出具有类似药效团的分子，第二种方法是利用药效团的结构进行全新分子设计。实际工作中，由于药效团结构简化太多，基于此基础的全新设计比较容易产生误差，应用较多的还是三维结构的数据库搜索。

在药物设计的实践中，应用最广的基于药效团的三维数据库的搜索软件是 Unity、CATALYST 和 MOE。一旦药效团模型确定后，就可以利用这些软件的功能对小分子的三维数据库进行搜索，寻找药效特征元素在空间分布类似的结构。从事三维数据库搜寻时，软件本身通常不考虑库中分子的构象可变性，但数据库中分子结构的柔性问题应该加以考虑，最简单的方法就是先搜索分子的构象，产生许多构象储存在数据库中。

## 17.5 从头设计

从数据库中得到的并不保证一定是新的分子，而且可能会偏重某些特殊类的分子。因此，人们就期望利用蛋白质受体或药效团的三度空间结构进行分子的从头设计。基于蛋白质受体结构的从头设计是基于某些与受体结合位点匹配良好的分子片段，把这些片段分子合理地组合起来形成新分子的过程。基于药效团的从头设计是利用药效团的三度空间结构设计新的分子。通常有两个基本的模式：第一个是“由外而内”模式（outside in model），是先利用结合位置判断是由那种官能基才能紧密地结合。再将这些官能基接上结合位置组成分子骨架，再由骨架变成真正的分子。第二个是“由内而外”模式（inside out model），分子由结合处附近开始生长，利用适当的算法借由能量判断控制分子的成长。图 17-5 显示这两种不同模式的差异。

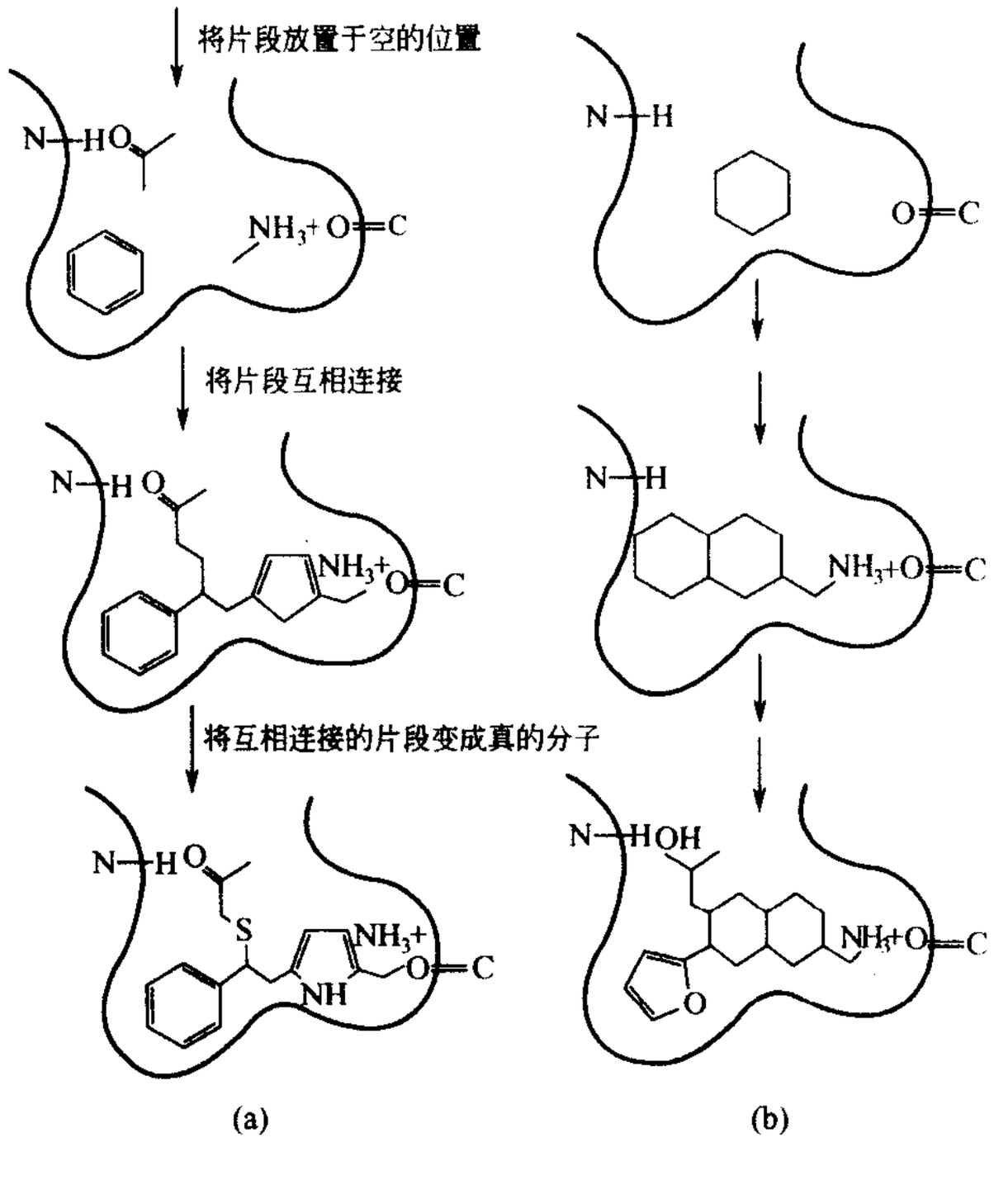



图 17-5　从头设计药物（a）由外而内（b）由内而外

# 17.6　计算实例

## 17.6.1　3D-QSAR 研究实例

本例是一个最普通的 3D-QSAR 的研究过程。根据文献报道的噁唑烷酮化合物结构及活性数据，抽取代表性的化合物 20 个作为训练集，剩余 3 个化合物作为检验集。以训练集分子构建预测活性的模型，经过抽一法交叉验证和检验集的验证，考察模型的可靠性，如果可靠就可以进行预测设计分子的活性。主要步骤如下所述。

第 1 步：分子构建。训练集分子（20 个）和检验集分子（3 个）先以 Chemoffice 软件构建三维结构，保存为 mol 格式，在 Cerius2 系统中，采用 MMFF 力场搜索低能构象，优化构象，保存分子构象。

第 2 步：分子叠合。由于活性构象无法得知，取活性最强分子的低能构象为参照模板，在 Cerius2 软件中，应用 Align 模块叠合，采用的是场叠合方式。叠合完成后，观察叠合图，如果个别分子没有与模板很好叠合，还需手动调整构象，直到所有分子母核结构基本重叠并且性质相近的取代基方向大致相同。叠合完成后的如图 17-6 所示。

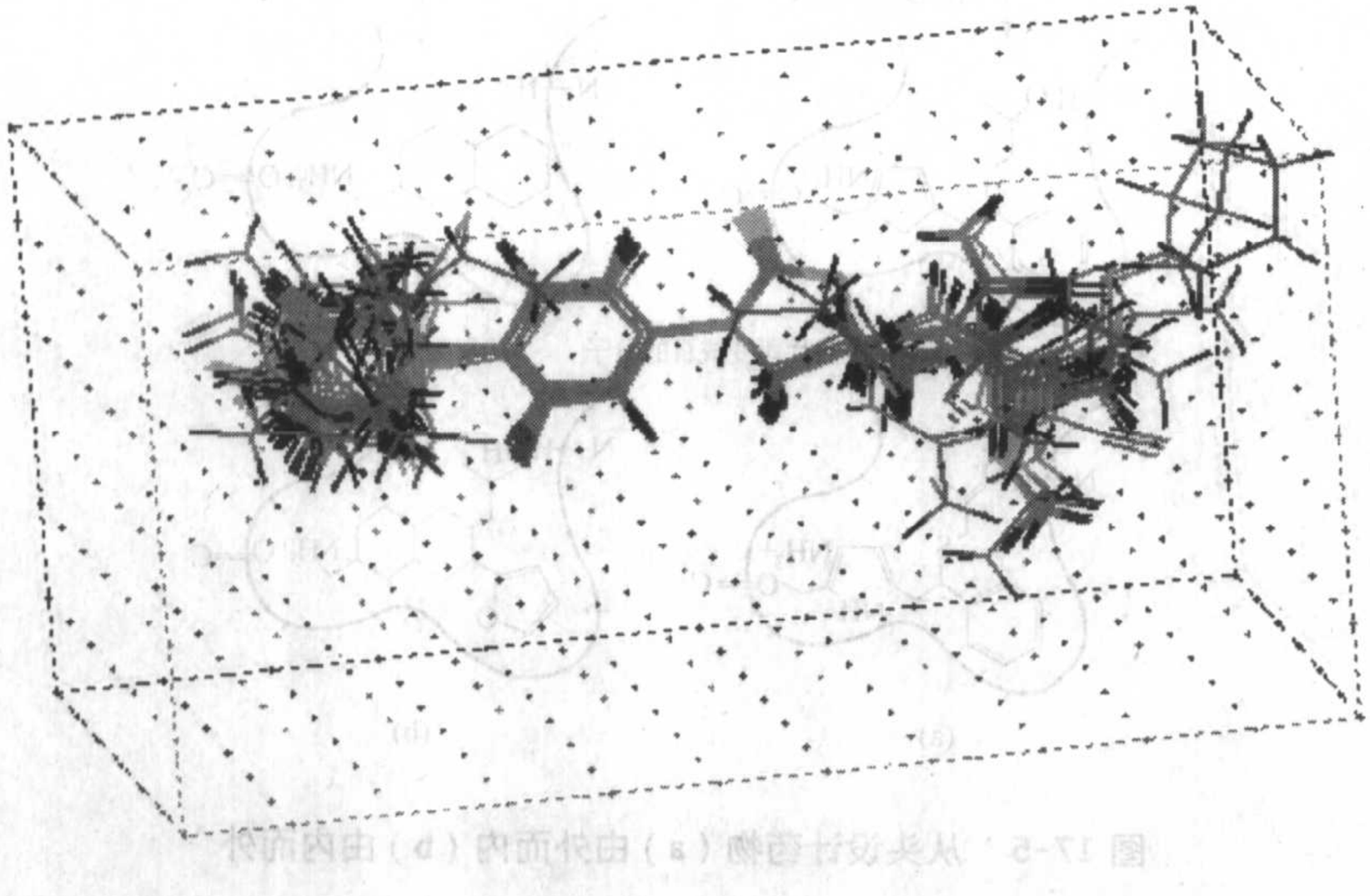


**图 17-6　分子叠合及格点设置**

第 3 步：场分析。Cerius2 的 MFA（molecular field analysis）场分析模块和 CoMFA 法采用相似的原理，参数（格点距离、探针类型）的设置可以采用默认值，选择格点距离 2Å，选择 sp3 杂化的 C＋离子、H＋及 HO－等探针来计算立体场、静电场和氢键场，计算得到叠合后每个分子在各格点与探针的相互作用能。

第 4 步：建立 QSAR 模型。场分析计算完成后，每个格点的数据很多，Cerius2 软件有很多方法拟合相关性方程。采用 G/PLS（遗传算法与偏最小二乘法结合的统计方法）处理数据，可以自动挑选最主要的变量，也可以手动改变变量的定义，活性数值采用 Ln(1/MIC) 形式，是不同探针在格点能量（自变量）的函数。自变量的个数是可调的，需要根据活性的样本来控制，本次拟合的训练集有 20 个样本，所以最合理的自变量数为 4 个。最终得到 QSAR 方程。形式如下：

Activity=−1.88785+0.009576 * “H+/351”−0.037078 * “HO−/346”−0.035981 * “H+/303”−0.041852 * “CH3/322”

拟合的方程的评价结果：$r^2$：0.984；Press：39.961；SD：40.223；LSE：0.033。

第5步：模型的验证。对于模型预测能力的评估一般有两种方法，就是交叉验证和检验集的预测结果。用交叉验证抽一法进行交叉验证，结果为：$q^2$：0.951；Press：0.827；SD：16.944。表明训练集的数据具有良好的相关性，拟合的方程可靠性较高。测定值与预测值具有可比性。由于交叉验证的相关系数较高，检验集样本的预测结果较好，因此，本模型具有较高的可靠性，可以用于类似结构的活性预测，具体方法与检验集样本预测方法完全一致，见表17-8所列。

**表17-8　检验集化合物活性［Ln(1/MIC)］的测定值、预测值及残差**

| 检验集 | 测定值 | 预测值 | 残差 |
|---|---|---|---|
| 1 | −1.833 | −1.236 | −0.579 |
| 2 | 0.248 | 0.282 | −0.034 |
| 3 | −0.445 | 0.165 | −0.610 |

## 17.6.2　对接计算和分子设计实例

对接主要包括的内容，首先是受体的确定和小分子库的建立，然后以适当的方法进行对接计算，寻找小分子与受体的最合理的结合模式，并对结合模式进行评分，根据评分的排序原则筛选出目标结构，具体操作过程如下。

**（1）受体的确定**　本例以PDF晶体结构为基础进行抗菌药物设计和虚拟筛选。文献中报道了48个PDF酶的结构，选择哪些酶作为模板进行对接将会直接影响到设计结果的准确性。因此首先要确定合理的受体进行对接计算。

以代号为1bsk的晶体结构参照，对所有酶进行结构叠合。选择1bsk中的配体6Å以内的氨基酸残基为活性区域（图17-7），比较所有叠合结构在活性区域对应位置残基的分布，表17-9列出了loop区的残基结果。然后计算活性区域对应位置残基α碳的坐标差异（图17-8）。根据活性区域的氨基酸残基及其α碳的坐标差异，将现有的PDF酶分成9类，从每类中选择1～2个结构，共选择10个具有代表性的晶体结构作为对接受体。

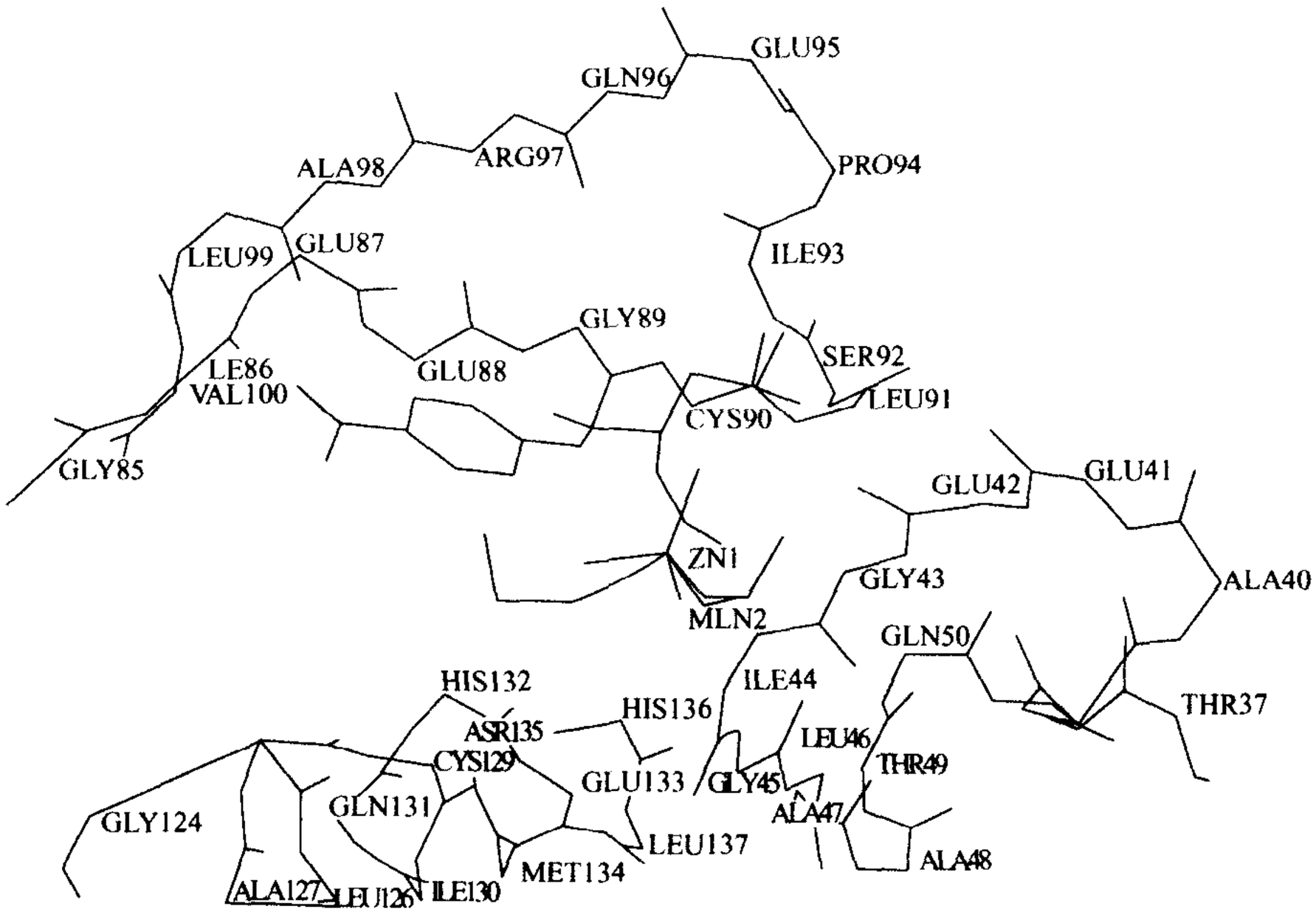



**图 17-7　1bsk 的活性区域氨基酸残基的分布**

**表 17-9　1bsk 的活性区域 loop 区残基组成情况**

| 结构代码 | 分类 | loop 区残基名称 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1bsk | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1BS4 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1BS5 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1bs6 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1BS7 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1BS8 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1BSJ | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1BSZ | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1G2A | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1DEF | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1DFF | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1G27 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1ICJ | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1LRU | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |

续表

| 结构代码 | 分类 | loop 区残基名称 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1XEM | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1XEN | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1XEO | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 2AI8 | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 2DEF | 1 | E41 | E42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1SV2 | 2 | A44 | E45 | G46 | V47 | G48 | L49 | A50 | A51 | P52 | Q53 |
| 1SZZ | 2 | A44 | E45 | G46 | V47 | G48 | L49 | A50 | A51 | P52 | Q53 |
| 1VEV | 2 | A44 | E45 | G46 | V47 | G48 | L49 | A50 | A51 | P52 | Q53 |
| 1VEY | 2 | A44 | E45 | G46 | V47 | G48 | L49 | A50 | A51 | P52 | Q53 |
| 1VEZ | 2 | A44 | E45 | G46 | V47 | G48 | L49 | A50 | A51 | P52 | Q53 |
| 1Y6H | 2 | A44 | E45 | G46 | V47 | G48 | L49 | A50 | A51 | P52 | Q53 |
| 1LM4 | 3 | R56 | S57 | G58 | V59 | G60 | L61 | A62 | A63 | P64 | Q65 |
| 1LMH | 3 | R56 | S57 | G58 | V59 | G60 | L61 | A62 | A63 | P64 | Q65 |
| 1LQW | 3 | R56 | S57 | G58 | V59 | G60 | L61 | A62 | A63 | P64 | Q65 |
| 1Q1Y | 3 | R56 | S57 | G58 | V59 | G60 | L61 | A62 | A63 | P64 | Q65 |
| 2AI9 | 3 | R55 | S56 | G57 | V58 | G59 | L60 | A61 | A62 | P63 | Q64 |
| 1IX1 | 4 | A42 | P43 | G44 | I45 | G46 | L47 | A48 | A49 | T50 | Q51 |
| 1LRY | 4 | A41 | P42 | G43 | I44 | G45 | L46 | A47 | A48 | T49 | Q50 |
| 1N5N | 4 | A42 | P43 | G44 | I45 | G46 | L47 | A48 | A49 | T50 | Q51 |
| 1S17 | 4 | A42 | P43 | G44 | I45 | G46 | L47 | A48 | A49 | T50 | Q51 |
| 1ZXZ | 5 | A45 | P46 | G47 | V48 | G49 | L50 | A51 | A52 | P53 | Q54 |
| 1ZY0 | 5 | A45 | P46 | G47 | V48 | G49 | L50 | A51 | A52 | P53 | Q54 |
| 1ZY1 | 5 | A45 | P46 | G47 | V48 | G49 | L50 | A51 | A52 | P53 | Q54 |
| 1JYM | 6 | S102 | K103 | G104 | I105 | G106 | L107 | S108 | A109 | P110 | Q111 |
| 1RL4 | 6 | S103 | K104 | G105 | I106 | G107 | L108 | S109 | A110 | P111 | Q112 |
| 1RQC | 6 | S103 | K104 | G105 | I106 | G107 | L108 | S109 | A110 | P111 | Q112 |
| 1LM6 | 7 | R68 | G69 | G70 | V71 | G72 | L73 | A74 | A75 | P76 | Q77 |
| 2AI7 | 7 | R67 | G68 | G69 | V70 | G71 | L72 | A73 | A74 | P75 | Q76 |
| 2AIA | 7 | R67 | G68 | G69 | V70 | G71 | L72 | A73 | A74 | P75 | Q76 |
| 2AIE | 7 | R67 | G68 | G69 | V70 | G71 | L72 | A73 | A74 | P75 | Q76 |
| 1WS0 | 8 | A42 | D43 | G44 | V45 | G46 | L47 | A48 | A49 | P50 | Q51 |
| 1WS1 | 8 | A42 | D43 | G44 | V45 | G46 | L47 | A48 | A49 | P50 | Q51 |
| 1LQY | 9 | R56 | P57 | G58 | I59 | G60 | L61 | A62 | A63 | P64 | Q65 |

注：E=GLU；R=ARG；S=SER，A=ALA，G=GLY，D=ASP，P=PRO；K=LYS；L=LEU；Q=GLN；T=THR；V=VAL；I=ILE。

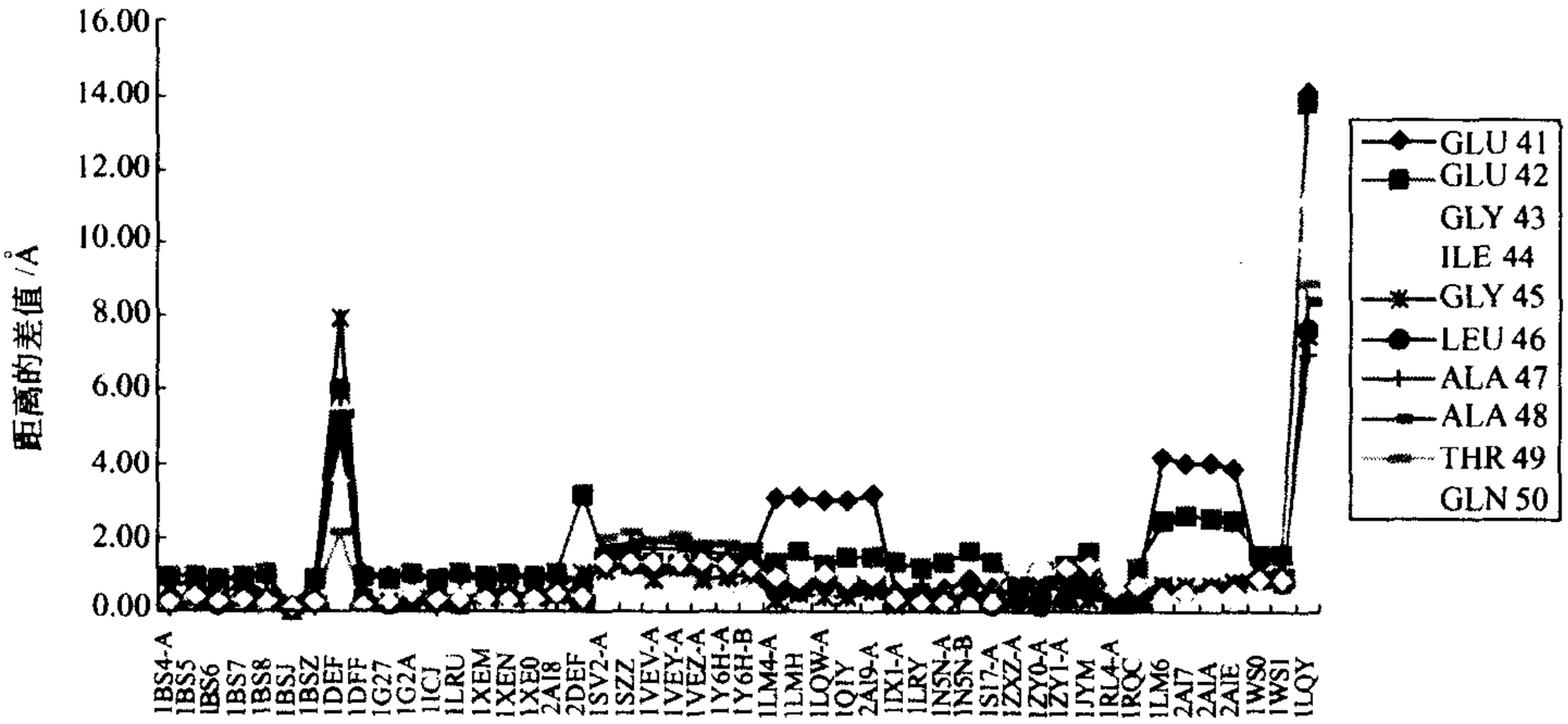



**图 17-8　所有 PDF 与 1bsk 叠合后在活性位点相应残基的 α 碳的坐标差异**

**（2）对接方法的确定**　最理想的对接应该能够快速搜索到所有的构象空间，并且正确地评分。每个软件都有处理方式，计算的结果差异较大，需要适当选择对计算方法，特别是分子很多时，许多分子的评分比较接近，选择分子就是一件非常困难的事，因此，要提前考虑计算方法的适用性。

本例采用 10 个 PDF 晶体结构（1BS8、1BSK、1G27、1LRU、1Y6H、1LM4、1S17、1R14、1LM6 和 1LQY）作为对接受体。从文献中选择一些对应的配体化合物结构（活性大小存在合理的分布）作为阳性对照结构，再选择一些肯定没有相关性的结构作为阴性对照结构，共得到 37 个化合物用于对接计算。分别采用三种对接方法（FlexX，DOCK5 和 MCDOCK）考察了 37 个化合物与 10 个受体的对接结果，每种化合物与 10 个受体对接后的最高得分取平均值。三种方法的结果评分排序的相关性如图 17-9，FlexX 和 MCDOCK 对分子评分的排序具有较好的相关性，DOCK5 则相关性较差。对阳性对照和阴性对照结构的对接结果进行定性分析，活性与 FlexX 和 MCDOCK 的评分的相关性好于 DOCK5。

由于对接评分受配体分子量大小的影响，分子量越大，VDW 作用越强，不同方法对接评分排序与分子量的相关性如图 17-10～图 17-12。图中化合物的顺序是按照对接分数由低往高进行排列后，同时显示评分及分子量，计算了评分线和分子量线的斜率。评分线斜率越大，说明该方法的计

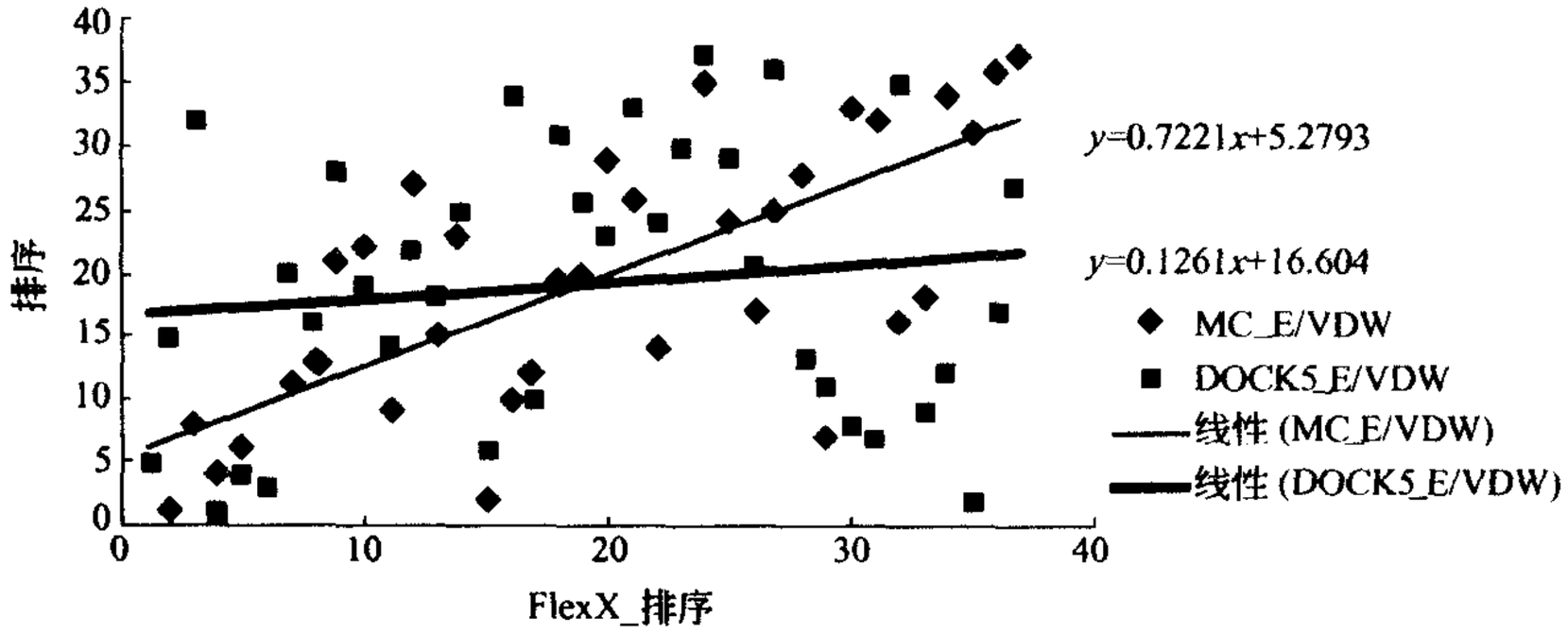



**图 17-9　三种对接方法评分排序的相关性**

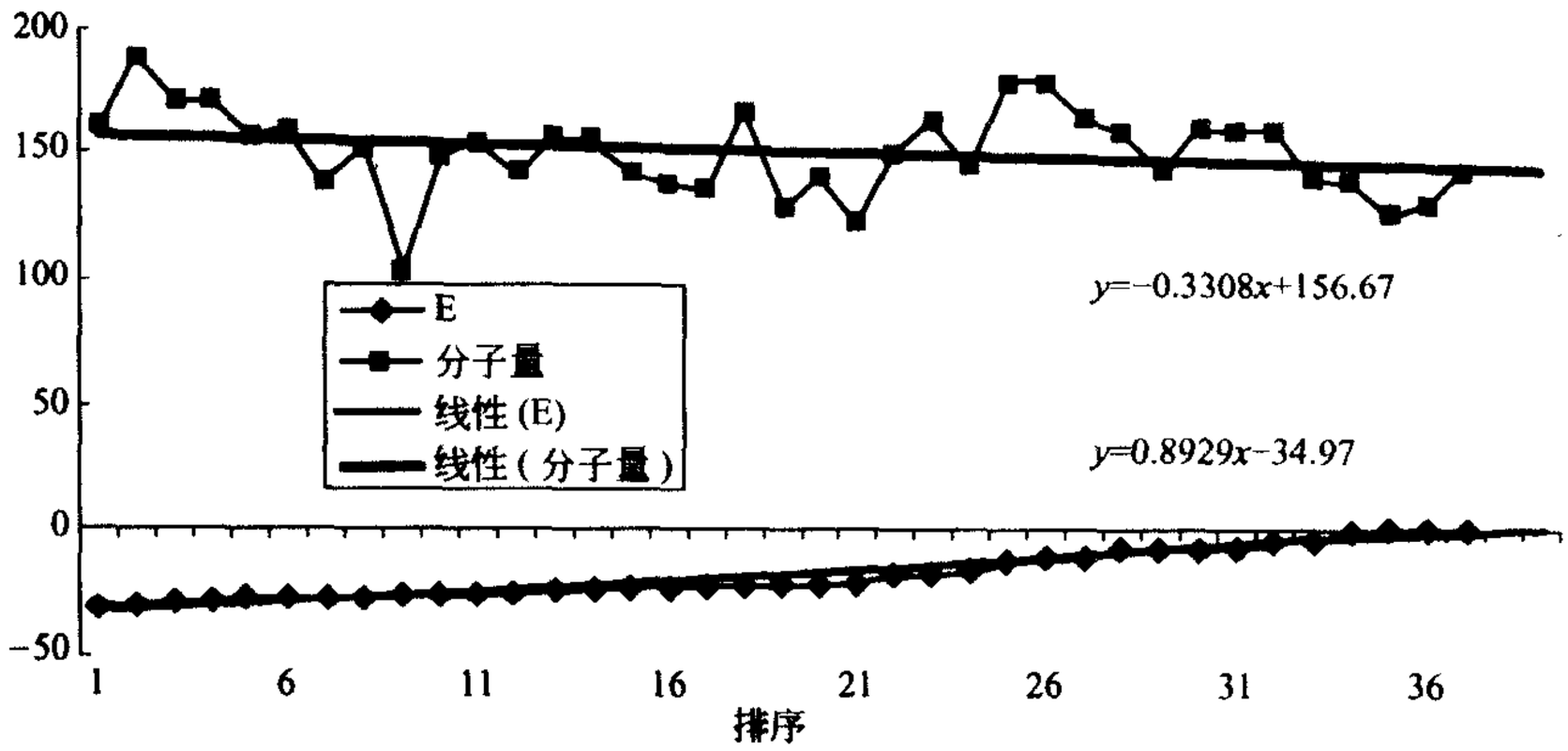



**图 17-10　FlexX 评分直接排序**

算结果表现的差异较大，利于对分子进行评价，相当于仪器分析指标的分辨率较高。分子量曲线拟合的斜率越小，则说明该计算方法受分子大小的影响越小。三种方法中 FlexX 方法受化合物分子量影响最小，而且分辨率最高。MCDOCK 和 DOCK5 的评分容易受分子大小的影响。

有文献报道需要对接分数进行校正，作者考察的校正方法为将对接分数值除以：总原子数开根号值、重原子数开根号值、VDW 体积、VDW 开根号值、VDW 开立方值。计算结果见表 17-10，FlexX 方法经 VDW 体积校正后，受化合物分子量影响最小，而且分辨率最高。MCDOCK 和 DOCK5 的评分，当考虑减少分子量的影响时，可以总原子的根号值校正，当考虑分别率时，可用 VDW 的立方根值进行校正。

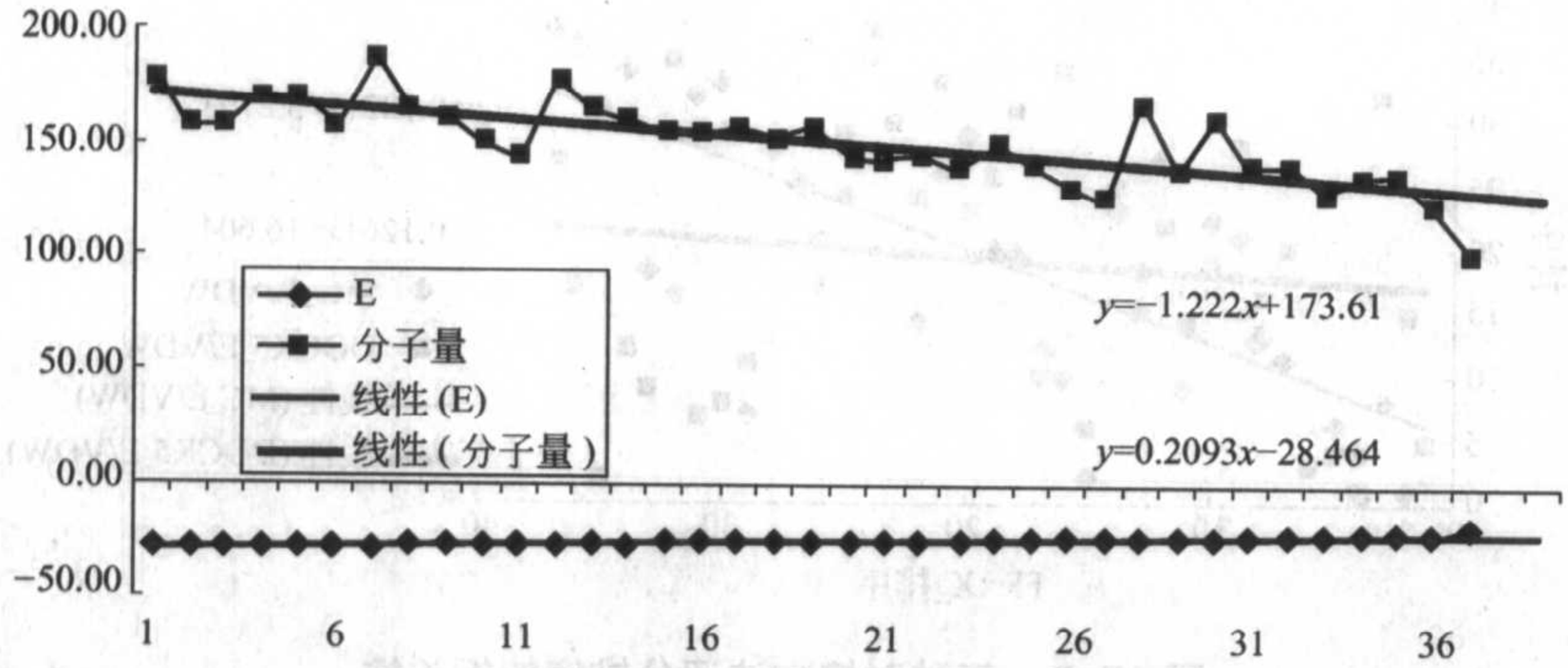



图 17-11　DOCK5 评分直接排序

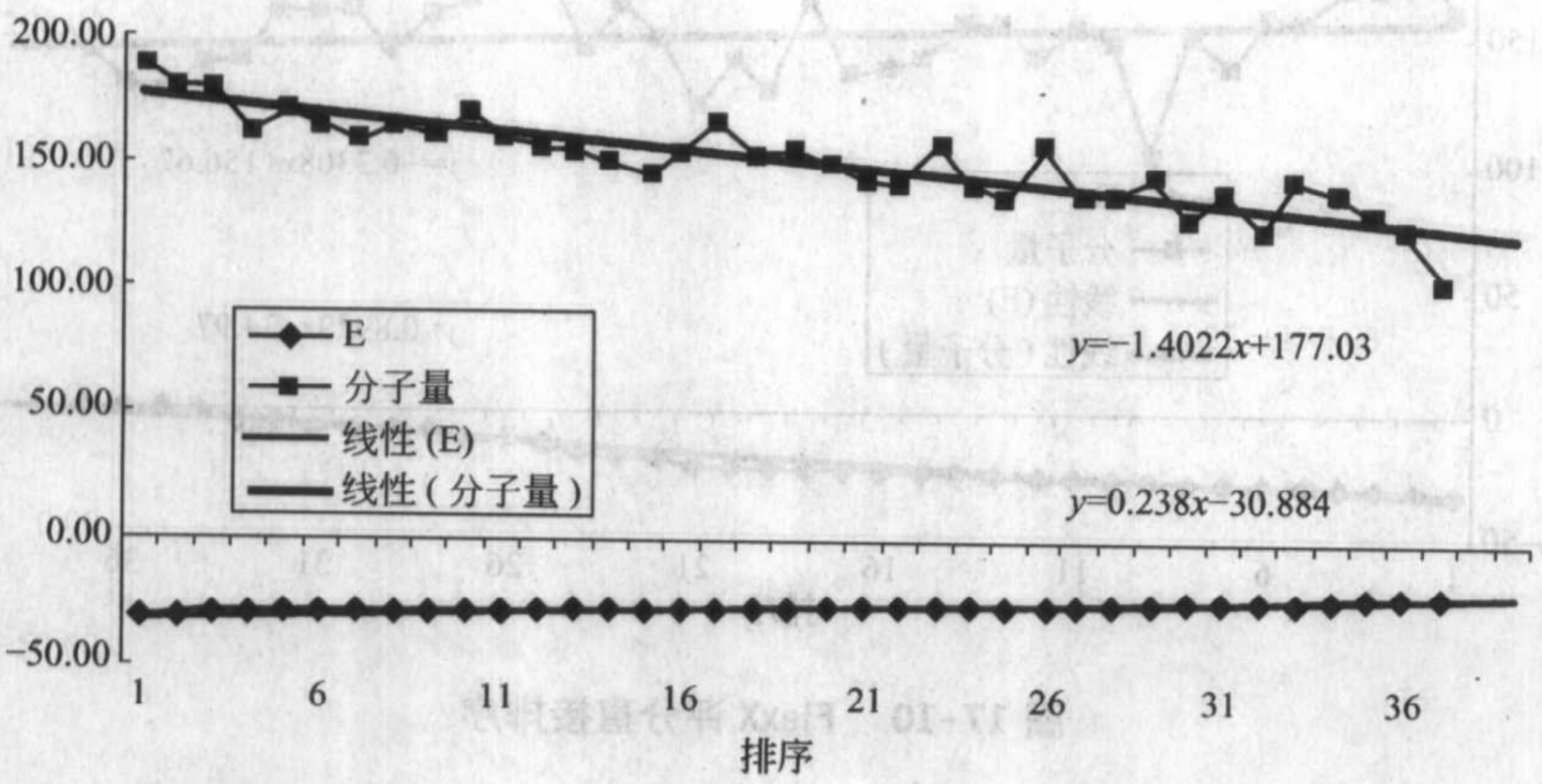



图 17-12　MCDOCK 评分直接排序

表 17-10　不同评分校正方法对排序后分数与分子量线斜率的影响

| 评分校正方法 | 分子量线斜率 | | | 分数线斜率 | | |
|---|---|---|---|---|---|---|
| | FlexX | DOCK5 | MCDOCK | FlexX | DOCK5 | MCDOCK |
| 分数 | −0.33 | −1.22 | −1.40 | 0.89 | 0.21 | 0.24 |
| 分数/总原子根号值 | −0.12 | −0.62 | −0.81 | 0.88 | 0.13 | 0.17 |
| 分数/重原子根号值 | −0.12 | −0.38 | −1.26 | 0.89 | 0.15 | 0.22 |
| 分数/VDW | −0.01 | 0.20 | 0.01 | 0.87 | −0.01 | 0.05 |
| 分数/VDW 根号值 | −0.12 | −0.94 | −1.01 | 0.88 | 0.18 | 0.20 |
| 分数/VDW 立方根值 | −0.15 | −1.03 | −1.24 | 0.89 | 0.20 | 0.22 |

**(3) 配体结构设计与对接计算** 在前面确定 PDF 受体结构和筛选方法的基础上，主要是利用 FlexX 对设计得到配体结构与 PDF 进行对接计算，评价各种形式的结构改造对配体结构与 PDF 结合能力的影响，在筛选的初期，主要考虑为改造留出空间，找出母体，分子量为优先考虑。在筛选后期，主要是从大量分子中以确定具体分子，分辨率为优先考虑。

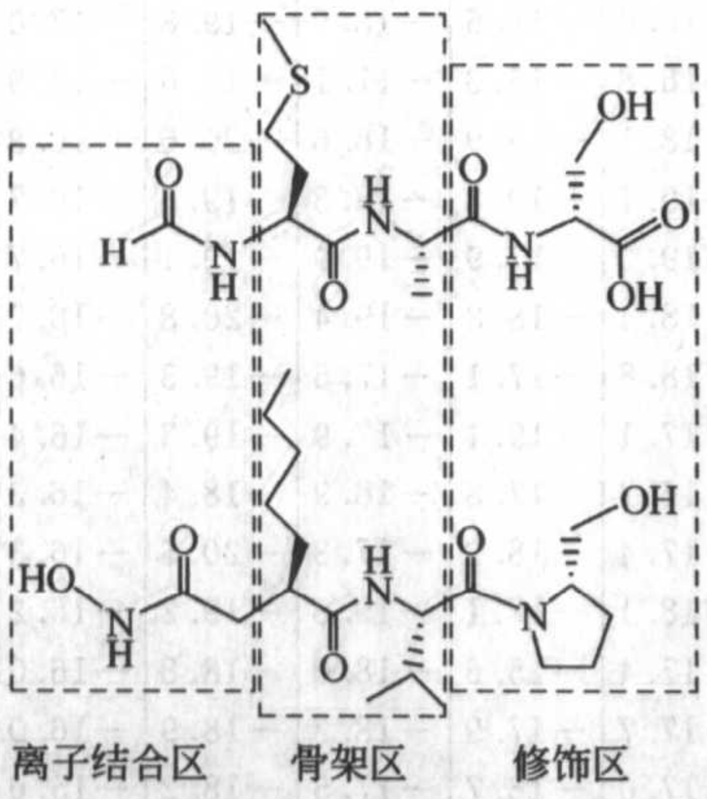



**图 17-13 分子结构分解示意**

根据晶体结构与配体的结合方式，将配体分子分成三部分：离子结合区、骨架区和修饰区如图 17-13。分两条线进行设计，第一条线，首先分别设计离子结合区和骨架区不同结构，进行组合，以 FlexX 筛选出得分负值高的结构，再连接修饰区结构形成完整分子。第二条线，基于已知配体的结构，改变修饰区结构。采用的取代基主要涉及极性、几何和芳香等不同的代表性结构，得到虚拟分子库。所有设计的分子在 MOE 软件中先以 MMF94 力场计算电荷，再进行构象优化，将构象保存为 mol2 格式后直接用于对接计算。采用 Sybyl 7.0 分子模拟软件包的 FlexX/Run muti-ligand 模块，定义受体中离子周围 8.0Å 范围为活性位点，采用经验结合自由能函数作为打分函数，与 10 个受体对接，以平均对接分数计算。

本例是进行母体结构的筛选，评分经 VDW 体积校正后排序，计算取代基和母体的平均分排序，以选择适当的取代基和母体。结果见表 17-11，母体 m38 和 m34，取代基 g9 和 g17 平均得分较佳，由此构建的分子，可以作为下一步改造的基础。

**表 17-11 虚拟分子库与受体的对接分数除以 VDW 体积所得的结果**

| 母体号 | g1 | g6 | g7 | g8 | g9 | g10 | g11 | g12 | g16 | g17 | mean | order |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| m38 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | −20.8 | 1 |
| m34 | −20.1 | −19.1 | −20.1 | −19.1 | −21.4 | −19.8 | −19.5 | −18.3 | −20.1 | −20.7 | −18.0 | 2 |
| m50 | −18.9 | −20.6 | −18.8 | −20.0 | −22.1 | −20.1 | −18.9 | −18.2 | −20.9 | −21.3 | −17.7 | 3 |
| m40 | −19.4 | −19.2 | −19.2 | −18.2 | −20.4 | −19.1 | −18.4 | −18.7 | −18.3 | −21.6 | −17.7 | 4 |
| m53 | −18.4 | −18.6 | −18.7 | −18.3 | −21.0 | −20.5 | −18.4 | −19.0 | −20.6 | −21.3 | −17.4 | 5 |

续表

| 母体号 | g1 | g6 | g7 | g8 | g9 | g10 | g11 | g12 | g16 | g17 | mean | order |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| m28 | −20.4 | −19.8 | −18.3 | −19.5 | −20.5 | −18.7 | −18.3 | −18.2 | −19.7 | −19.7 | −17.4 | 6 |
| m37 | −19.0 | −17.7 | −19.2 | −18.5 | −19.9 | −19.3 | −19.9 | −18.1 | −20.6 | −20.3 | −17.4 | 7 |
| m51 | −16.8 | −19.8 | −17.8 | −18.3 | −21.4 | −21.3 | −18.6 | −18.6 | −18.9 | −19.8 | −17.0 | 8 |
| m4 | −18.5 | −17.1 | −18.5 | −17.9 | −19.5 | −17.7 | −18.8 | −15.3 | −17.1 | −17.6 | −16.9 | 9 |
| m31 | −17.9 | −17.0 | −18.1 | −18.5 | −21.3 | −17.9 | −18.1 | −18.9 | −18.6 | −20.6 | −16.8 | 10 |
| m32 | −17.5 | −17.1 | −18.1 | −17.6 | −21.2 | −17.2 | −19.1 | −19.3 | −18.5 | −19.6 | −16.7 | 11 |
| m16 | −17.3 | −17.8 | −19.9 | −18.2 | −19.5 | −18.2 | −19.0 | −17.9 | −19.9 | −19.1 | −16.7 | 12 |
| m33 | −18.2 | −18.4 | −18.4 | −17.6 | −19.6 | −17.4 | −18.4 | −18.8 | −19.4 | −20.8 | −16.7 | 13 |
| m10 | −17.9 | −18.0 | −18.8 | −17.8 | −19.4 | −18.2 | −18.8 | −17.1 | −17.5 | −19.3 | −16.6 | 14 |
| m52 | −16.5 | −17.3 | −17.0 | −18.0 | −21.2 | −19.9 | −17.1 | −19.1 | −17.9 | −19.3 | −16.4 | 15 |
| m42 | −17.5 | −18.7 | −18.3 | −17.5 | −19.1 | −17.5 | −16.7 | −17.8 | −18.9 | −18.4 | −16.3 | 16 |
| m39 | −17.6 | −17.8 | −17.8 | −16.4 | −20.3 | −18.2 | −17.4 | −18.2 | −17.3 | −20.5 | −16.3 | 17 |
| m36 | −16.4 | −16.6 | −16.9 | −17.4 | −18.9 | −16.8 | −18.1 | −17.1 | −19.5 | −19.2 | −16.2 | 18 |
| m22 | −17.8 | −17.2 | −17.0 | −17.7 | −18.6 | −18.7 | −17.4 | −15.6 | −18.0 | −18.8 | −16.0 | 19 |
| m55 | −17.3 | −15.7 | −17.1 | −18.1 | −18.9 | −18.0 | −17.7 | −17.9 | −18.3 | −18.9 | −16.0 | 20 |
| mean | −16.8 | −16.8 | −16.9 | −16.7 | −18.6 | −17.1 | −17.0 | −16.7 | −17.5 | −18.2 | −15.6 | |
| order | 7 | 8 | 6 | 9 | 1 | 4 | 5 | 10 | 3 | 2 | | |

## 参 考 文 献


1 陈凯先，罗小民，蒋华良. 中国科学院院刊，(4)，265，2000

2 蒋华良，胡增建，陈建忠等. 化学进展，10 (4)，427，1998

3 Taylor R D，Jewsbury P J，Essex J W. Journal of Computer-aided Molecular Design，16，151，2002

4 Dammkoehler R A，Karasek S F，Shands E F B，Marshall G R. Journal of Computer-aided Molecular Design，3，3，1984

5 Sheridan R P，Nilakantan R，Dixon J S，Venkataraghavan R. Journal of Medicinal Chemistry，29，899，1986

6 陈凯先，蒋华良，嵇汝运. 计算机辅助药物设计——原理、方法及应用. 上海：上海科学技术出版社，2000

7 Kramer B，Rarey M，Lengauer T. Proteins：Structure，Function，and Genetic，37，1，1999

8 Pan Y，Huang N，Cho S，MacKerell Jr A D. J Chem Inf Comput Sci，43，267，2003

9 Peng H，Huang N，Qi J，Xie P，Xu C，Wang J，Yang C. Bioorg Med Chem Lett，13，3693，2003

10 Muegge I，Enyedy I J. Curr Med Chem，11，693，2004

11 徐筱杰，侯廷军，乔学斌，章威. 计算机辅助药物分子设计. 北京：化学工业出版社，2004

# 附录A 经典运动力学

## A-1 牛顿运动力学

依照牛顿运动定律，粒子所受的力等于此粒子的质量乘以加速度，即：

$$\vec{F}=m\vec{a} \tag{A-1}$$

或以动量 $\vec{p}$ 表示为：

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t}=\dot{\vec{p}}=\vec{F} \tag{A-2}$$

以上为熟知的牛顿运动方程式 (Newtonian equation of motion)。若质量与时间无关，则可将式(A-2) 写成：

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t}=m\frac{\mathrm{d}\dot{\vec{r}}}{\mathrm{d}t}=m\ddot{\vec{r}}=m\vec{a} \tag{A-3}$$

式中，$\vec{r}=(x,y,z)$ 为粒子的位置。若 $\vec{F}$ 为位置的函数 $\vec{F}(x,y,z)$ 则式(A-3) 为一个二次微分方程式。因此，若知粒子的起始位置 $x(0)$，$y(0)$，$z(0)$，则可由此方程式的解得到粒子的位置与时间的关系 $x(t)$，$y(t)$，$z(t)$。

以如图 A-1 的简谐振子的运动为例。

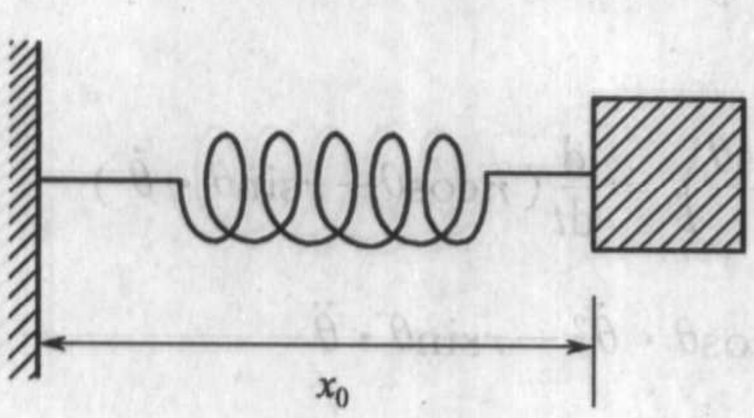



**图 A-1　简谐振子示意**

设 $x_0$ 为弹簧未拉伸的长度，则依照虎克定律 (Hook's law)，弹簧所系质点所受的力为：

$$F=-k(x-x_0) \tag{A-4}$$

依式(A-2)，得运动方程式：

$$m\ddot{x}=-k(x-x_0) \tag{A-5}$$

定义 $\xi=x-x_0$，则式(A-5) 可写成

$$\frac{d^2\xi}{dt^2}+\frac{k}{m}\xi=0 \tag{A-6}$$

此微分方程式的解为：

$$\xi(t)=A\sin\omega t+B\cos\omega t \tag{A-7}$$

式中，$\omega=\sqrt{\frac{k}{m}}$为振荡子的振动频率。可将式(A-7) 化为

$$\xi(t)=\sqrt{A^2+B^2}\left(\sin\omega t\cdot\frac{A}{\sqrt{A^2+B^2}}+\cos\omega t\cdot\frac{B}{\sqrt{A^2+B^2}}\right)$$

$$=C(\sin\omega t\cdot\cos\phi+\cos\omega t\cdot\sin\phi)=C\sin(\omega t+\phi) \tag{A-8}$$

$\phi$ 通常称为“相角”(phase angle)。

再以一例说明牛顿运动方程式的解法。设一个二维平面运动的粒子，粒子与原点为库仑吸引力。粒子所受的力为：

$$\vec{F}=-K\frac{\vec{r}}{r^3} \tag{A-9}$$

式中，$K$ 为库仑常数；$r$ 为粒子与原点的距离，$r=(x^2+y^2)^{1/2}$。

将 $r$ 代入式(A-9)，得牛顿运动方程式，

$$m\ddot{x}=F_x=-\frac{Kx}{(x^2+y^2)^{1/2}}$$

$$m\ddot{y}=F_y=-\frac{Ky}{(x^2+y^2)^{1/2}} \tag{A-10}$$

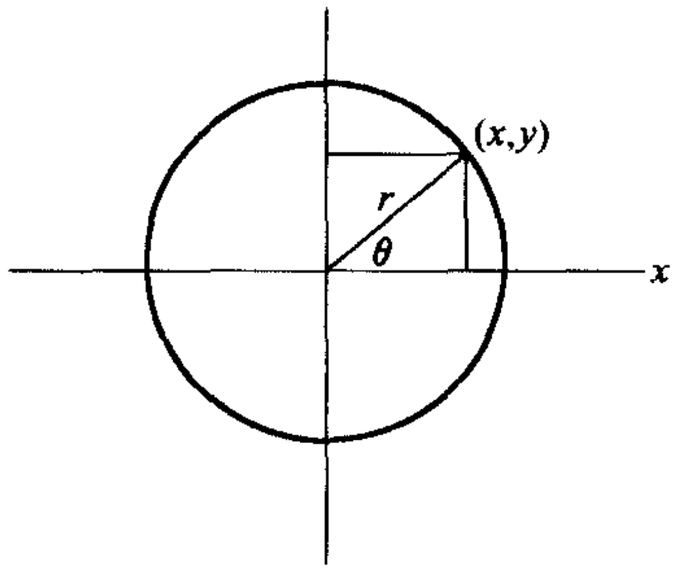



图 A-2 二维平面

由上，$x$ 的微分方程式含有变量 $y$，$y$ 的微分方程式含有变量 $x$，此二方程式互相耦合 (coupling)，不易求解。因此，必须设法将变量 $x$，$y$ 转换，使成为去耦合 (decoupling) 的方程式才能得解。将笛卡儿坐标换为极坐标，如图 A-2，即：

$$x=r\cos\theta,\quad y=r\sin\theta \tag{A-11}$$

代入式(A-10)，得：

$$\ddot{x}=\frac{d^2x}{dt^2}=\frac{d}{dt}\left(\frac{dx}{dt}\right)=\frac{d}{dt}\left(\frac{dr\cos\theta}{dt}\right)=\frac{d}{dt}(\dot{r}\cos\theta-r\sin\theta\cdot\dot{\theta})$$

$$=\ddot{r}\cos\theta-2\dot{r}\sin\theta\cdot\dot{\theta}-r\cos\theta\cdot\dot{\theta}^2-r\sin\theta\cdot\ddot{\theta}$$

$$\ddot{y}=\ddot{r}\sin\theta+2\dot{r}\cos\theta\cdot\dot{\theta}-\dot{r}\sin\theta\cdot\dot{\theta}^2+r\cos\theta\cdot\ddot{\theta}$$

故得极坐标 $r$，$\theta$ 的运动方程式，即：

$$\left[m(\ddot{r}-\dot{\theta}^2 r)+\frac{K}{r^2}\right]\cos\theta-m(r\ddot{\theta}+2\dot{\theta}\dot{r})\sin\theta=0$$

$$\left[m(\ddot{r}-\dot{\theta}^2 r)+\frac{K}{r^2}\right]\sin\theta+m(r\ddot{\theta}+2\dot{\theta}\dot{r})\cos\theta=0 \tag{A-12}$$

将式(A-12) 中第一式乘以 $\cos\theta$，第二式乘以 $\sin\theta$ 相加后得：

$$m(\ddot{r}-\dot{\theta}^2 r)+\frac{K}{r^2}=0 \tag{A-13}$$

将此式与式(A-12) 合并，得

$$m(r\ddot{\theta}+2\dot{\theta}\dot{r})=0 \tag{A-14}$$

此式可化为对时间的微分式，即：

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\theta})=0 \tag{A-15}$$

此式的意义为：

$$mr^2\dot{\theta}=\text{常数} \tag{A-16}$$

而 $mr^2\dot{\theta}=L$ 为质点的角动量，故式(A-16) 物理意义为此运动粒子的角动量为恒定常数。此系统中，粒子受力方向朝向原点，称为向心力 (central force)。如此类的系统，其角动量必为常数。将 $mr^2\dot{\theta}=L$ 代入式(A-13)，消去 $\dot{\theta}$ 得仅与 $r$ 有关的径向方程式 (radial equation)，即：

$$m\ddot{r}-\frac{L^2}{mr^3}+\frac{K}{r^3}=0 \tag{A-17}$$

如上的过程即为将耦合的微分方程式去耦合。此微分方程式即使无简单的分析解亦可用数值方法解，得 $r(t)$。将之代入式(A-12) 可得 $\theta(t)$。此例引用极坐标的转换解运动方程式，虽然较为麻烦，但较解笛卡儿坐标系统的运动方程式却简单得多。这显示牛顿运动方程式，仅适用于笛卡儿坐标系：

$$m\ddot{q}=F_q \tag{A-18}$$

式中，$q$ 为笛卡儿坐标。

## A-2 拉格朗日运动力学

前面指出，牛顿力学仅适用于卡式坐标系。拉格朗日运动力学 (Lagrange dynamics) 却可通用于任何形式的坐标系。于笛卡儿坐标系，令 $K$

为粒子的动能，即：

$$K(\dot{x},\dot{y},\dot{z})=\frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2) \tag{A-19}$$

$U(x,y,z)$ 为势能。粒子于位置 $(x,y,z)$ 所受的力为：

$$\vec{F}=-\nabla U=-\left(\vec{i}\,\frac{\partial U}{\partial x}+\vec{j}\,\frac{\partial U}{\partial y}+\vec{k}\,\frac{\partial U}{\partial z}\right) \tag{A-20}$$

此粒子的牛顿运动方程式为：

$$m\ddot{x}=-\frac{\partial U}{\partial x},\ m\ddot{y}=-\frac{\partial U}{\partial y},\ m\ddot{z}=-\frac{\partial U}{\partial z} \tag{A-21}$$

定义新的函数 $L$ 为：

$$L(x,y,z,\dot{x},\dot{y},\dot{z})\equiv K(\dot{x},\dot{y},\dot{z})-U(x,y,z) \tag{A-22}$$

函数 $L$ 称为拉格朗日量 (Lagrangian)。根据此定义，得：

$$\frac{\partial L}{\partial \dot{x}}=\frac{\partial K}{\partial \dot{x}}=m\dot{x}$$

$$\frac{\partial L}{\partial x}=-\frac{\partial U}{\partial x}$$

$$\cdots \tag{A-23}$$

引用以上关系，将牛顿方程式改写成：

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}}\right)=\frac{\partial L}{\partial x} \tag{A-24}$$

对 $y$，$z$ 亦可得到类似的关系。此三个方程式称为卡式坐标系的拉格朗日运动方程式 (Lagrange equation)。拉格朗日运动方程式可适用于任何坐标系，其一般式为：

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right)=\frac{\partial L}{\partial q_j},\qquad (j=1,2,3,\cdots) \tag{A-25}$$

式中，$q_j$ 为系统的坐标，可为笛卡儿坐标、极坐标或任何形式的坐标。

以前述二维平面上库仑中心作用力运动的粒子为例，以极坐标表示粒子的动能及势能为

$$K=\frac{m}{2}(\dot{x}^2+\dot{y}^2)=\frac{m}{2}(\dot{r}^2+r^2\dot{\theta}^2)$$

$$U=-\frac{K}{r} \tag{A-26}$$

拉格朗日量为：

$$L(r,\theta,\dot{r},\dot{\theta})=\frac{m}{2}(\dot{r}^2+r^2\dot{\theta}^2)+\frac{K}{r} \tag{A-27}$$

此系统的两个极坐标的拉格朗日方程式为：

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{r}}\right)=\frac{\partial L}{\partial r}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}}\right)=\frac{\partial L}{\partial \theta} \tag{A-28}$$

将式(A-27) 代入，得

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\dot{r})=mr\dot{\theta}^2-\frac{K}{r^2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\theta})=0 \tag{A-29}$$

此结果与前节所得的式(A-16) 和式(A-17) 相同，但此处由拉格朗日方程式直接得出，不需经过复杂的变换。

再以一例说明拉格朗日力学的应用。图 A-3 为三质点的振动系统，设质点的质量均为 $m$，弹簧的弹力常数均为 $k$。

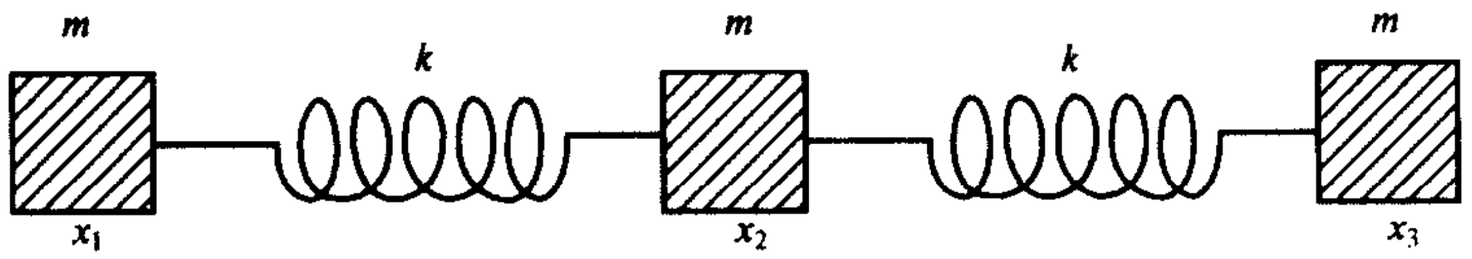



**图 A-3　三质点的振动系统**

此系统的势能为：

$$U=\frac{1}{2}k(x_{12}-x_{12}^0)^2+\frac{1}{2}k(x_{23}-x_{23}^0)^2 \tag{A-30}$$

式中，$x_{12}=x_2-x_1$，$x_{23}=x_3-x_2$；$x_{12}^0$与$x_{23}^0$分别为质点 1，2 与质点 2，3 间的平衡距离。系统的动能为：

$$K=\frac{1}{2}m\dot{x}_1^2+\frac{1}{2}m\dot{x}_2^2+\frac{1}{2}m\dot{x}_3^2 \tag{A-31}$$

若定义新的坐标，$R_1=\Delta r_{12}=x_{12}-x_{12}^0$，$R_2=\Delta r_{23}=x_{23}-x_{23}^0$，则系统的势能可重新表示为：

$$U=\frac{1}{2}kR_1^2+\frac{1}{2}kR_2^2 \tag{A-32}$$

新坐标与笛卡儿坐标间的关系为：

$$\dot{R}_1=\Delta\dot{r}_{12}=\dot{x}_2-\dot{x}_1$$

$$\dot{R}_2=\Delta\dot{r}_{23}=\dot{x}_3-\dot{x}_2 \tag{A-33}$$

此振动系统的质心守恒，其关系式为：

$$m\dot{x}_1+m\dot{x}_2+m\dot{x}_3=0 \tag{A-34}$$

将式(A-32)～式(A-34) 合并写为：

$$\begin{bmatrix}\dot{R}_1\\ \dot{R}_2\\ 0\end{bmatrix}=\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 1\\ m & m & m\end{bmatrix}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} \tag{A-35}$$

由反矩阵得：

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}-2/3 & -1/3 & 1/M\\ 1/3 & -1/3 & 1/M\\ 1/3 & 2/3 & 1/M\end{bmatrix}\begin{bmatrix}\dot{R}_1\\ \dot{R}_2\\ 0\end{bmatrix} \tag{A-36}$$

式中，$M$ 为系统的总质量，$M=3m$。将此式代入式(A-31)，得动能的表示式为：

$$K=\frac{m}{3}(\dot{R}_1^2+\dot{R}_2^2+\dot{R}_1\dot{R}_2) \tag{A-37}$$

依此，得定义于新坐标 $R_1$，$R_2$ 的拉格朗日方程组：

$$\frac{m}{3}(2\ddot{R}_1+\ddot{R}_2)+kR_1=0$$

$$\frac{m}{3}(\ddot{R}_1+2\ddot{R}_2)+kR_2=0 \tag{A-38}$$

此方程组之解为：

$$R_1=A_1\cos 2\pi\nu t$$

$$R_2=A_2\cos 2\pi\nu t \tag{A-39}$$

式中，$A_1$，$A_2$ 为振幅；$\nu$ 为振动的频率。此例中，$R_1$ 与 $R_2$ 称为系统振动的内坐标 (internal coordinate)。由此种坐标的拉格朗日运动方程式可直接解出振动的频率。

## A-3 哈密顿运动力学

与拉格朗日运动力学一样，哈密顿运动力学 (Hamiltonian dynamics) 可适用于任何坐标系。设运动系统含有 $N$ 个粒子，描述每个粒子的位置需要 3 个坐标。定义动量为：

$$p_j=\frac{\partial L}{\partial \dot{q}_j} \qquad (j=1,2,\cdots,3N) \tag{A-40}$$

并称动量 $p_j$ 与坐标 $q_j$ 彼此共轭 (conjugate)。考虑最简单的单粒子系统，定义哈密顿量 (Hamiltonian) 为：

$$H(p_1,p_2,p_3,q_1,q_2,q_3)=\sum_{j=1}^{3} p_j\,\dot{q}_j - L(\dot{q}_1,\dot{q}_2,\dot{q}_3,q_1,q_2,q_3) \tag{A-41}$$

拉格朗日量与哈密顿量的不同在于前者为$\dot{q}_j$ 与 $q_j$的函数，后者为 $p_j$ 与 $q_j$ 的函数。以笛卡儿坐标系为例，即：

$$p_x=\frac{\partial L}{\partial \dot{x}}=\frac{\partial K}{\partial \dot{x}}=m\dot{x}\ ,\ p_y=m\dot{y}\ ,\ p_z=m\dot{z} \tag{A-42}$$

将此代入式(A-41)，得哈密顿量：

$$\begin{aligned} H &=(m\dot{x}\cdot\dot{x}+m\dot{y}\cdot\dot{y}+m\dot{z}\cdot\dot{z})-(K-U) \\ &=\frac{1}{2m}(p_x^2+p_y^2+p_z^2)+U(x,y,z) \end{aligned} \tag{A-43}$$

式(A-43) 表示系统的哈密顿量即为系统的总能量

$$H=K(p_j)+U(q_j)=\text{总能量} \tag{A-44}$$

哈密顿运动方程式为：

$$\frac{\partial H}{\partial p_j}=\dot{q}_j$$

$$\frac{\partial H}{\partial q_j}=-\dot{p}_j \quad (j=1,2,3,\cdots,3N) \tag{A-45}$$

哈密顿力学的重心在于式(A-45)，可由此式得到对应于任何形式坐标的动量。哈密顿力学证实对任何的坐标系统，总能量的守恒。执行实际的计算则并不需要解哈密顿运动方程式。以上三质点的振动系统为例，系统的拉格朗日量为：

$$L=K-U=\frac{m}{3}(\dot{R}_1^2+\dot{R}_2^2+\dot{R}_1\,\dot{R}_2)-\left(\frac{1}{2}kR_1^2+\frac{1}{2}kR_2^2\right) \tag{A-46}$$

将式(A-45) 和式(A-46) 代入式(A-45) 可得与内坐标共轭的动量分别为：

$$\begin{aligned} p_1&=\frac{\partial L}{\partial \dot{R}_1}=\frac{2m}{3}\dot{R}_1+\frac{m}{3}\dot{R}_2 \\ p_2&=\frac{\partial L}{\partial \dot{R}_2}=\frac{2m}{3}\dot{R}_2+\frac{m}{3}\dot{R}_1 \end{aligned} \tag{A-47}$$

因此，利用内坐标对时间的微分可以计算对应的动量。例如在分子动力模拟中，可利用此关系检视特殊形式的运动。

# 附录 B 统计力学

统计力学（statistical mechanics）为一门实用的应用科学，可用以研究各种不同的多分子体系的特性。这些多分子体系包括：气体、液体、溶液、电解液、聚合物、金属等。统计力学由分子的微观（microscopic）性质出发，可计算及预测复杂系统的宏观（macroscopic）行为。利用统计力学可研究如系统的相变化、光谱、DNA 的转换及各种热力学特性。统计力学为分子模拟的基础，本章简单叙述一些统计力学的基本观念、原理与应用。

## B-1 基本统计原理

设 $n$ 个不同的样品（sample），度量得各样品的特性 $X_i$，$i=1$，…，$n$，则此类样品特性 $X$ 的平均值（average）或期望值（expectation value）为：

$$\overline{X} = \langle X \rangle = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{B-1}$$

若度量 $n$ 个样品得到 $j$ 个度量值，$j \leqslant n$，特殊的度量值 $X_j$ 的个数为 $n_j$，则：

$$n_1 + n_2 + \cdots n_j = \sum_{i=1}^{j} n_i = n$$

$$\overline{X} = \frac{n_1 X_1 + n_2 X_2 + \cdots + n_j X_j}{n_1 + n_2 + \cdots + n_j} = \sum_{i=1}^{j} \frac{n_i X_i}{n} = \sum_{i=1}^{j} P_i X_i \tag{B-2}$$

式中，$P_i = \frac{n_i}{n}$ 为 $X_i$ 出现的概率（probability）。

度量平均值的方差（variance）为：

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \overline{(X - \overline{X})^2} \tag{B-3}$$

方差为样品的度量值与平均值的差异性；方差愈小，表示各度量值与平均值间的差异愈小，统计的结果愈佳。方差可表示为更方便于计算的形式，即：

$$\sigma^2 = \frac{1}{n}\left[\sum_{i=1}^{n}(x_i)^2 - \frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right] \tag{B-4}$$

式中，$\sigma$ 称为标准误差（standard deviation），为方差的正值平方根，即：

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \langle x\rangle)^2} \tag{B-5}$$

计算机模拟中常需要计算一些特定值群体的分布（distribution），与理论推导的分布函数比较。其中最重要的理论分布函数为正态分布（normal distribution），正态分布或称为高斯分布（Gaussian distribution）。一般的正态分布函数其概率密度函数为：

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp[-(x-\langle x\rangle)^2/2\sigma^2] \tag{B-6}$$

式(B-6）右方指数项前的常数称为归一化因子（normalization factor），此因子使得 $f(x)$ 由 $-\infty \sim +\infty$ 的积分值为 1，此表示在整个空间出现的概率为 1。或将式(B-6）表示为：

$$f(x) = \sqrt{\frac{\alpha}{\pi}}\exp[-\alpha(x-\langle x\rangle)^2] \tag{B-7}$$

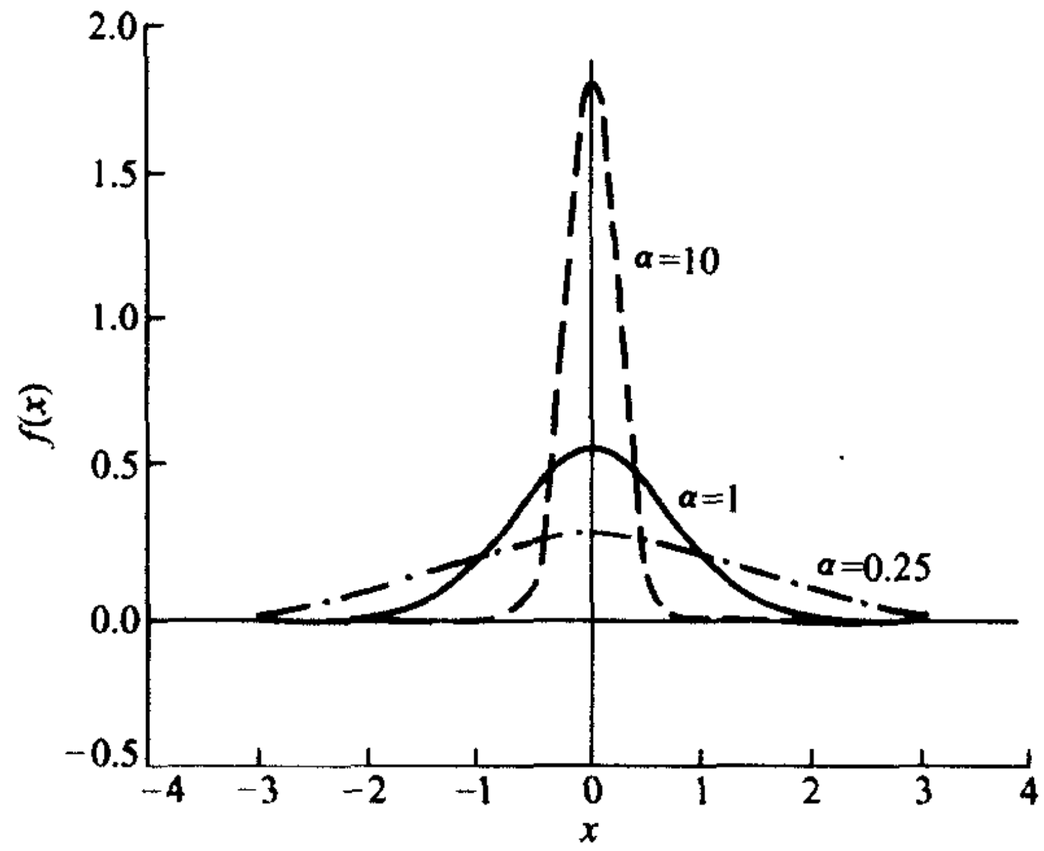



**图 B-1　三种不同 α 值的正态分布曲线**

$\alpha=1/2\sigma^2$。图 B-1 显示平均值为零的三种方差的正态分布。若方差大于 1（$\alpha$ 小），则曲线较平缓，若方差小于 1（$\alpha$ 大）则分布曲线较为尖锐。

因为这些高斯函数均为归一化的函数，故图中分布曲线下的面积均为 1。

## B-2 系综

一些具有同样条件系统（system）的集合称为系综（ensemble）。例如正则系综（canonical ensemble），为一群具有相同分子数目 $N$、相同体积 $V$ 与相同温度 $T$ 的系统的集合，以符号（$N,V,T$）表示。除正则系综外，统计力学所讨论的系综尚有微正则系综（micro-canonical ensemble）（$N,V,E$），$E$ 为系统的能量；巨正则系综（grand canonical ensemble），（$V,T,\mu$），$\mu$ 为系统的化学势能（chemical potential）；等温定压系综（isothermal-isobaric ensemble）（$N,T,p$），$p$ 为系统的压力等各种不同的系综。系综为统计力学最重要的观念，系统一切的统计特性与热力学特性都是以系综为起点推导得到的。实际应用上，不同的体系需选取适当的系综，例如正则系综适用于研究定温定容且无分子数目变化的问题。系综如同统计学的样品（sample），样品数愈多，统计的结果愈佳。组成系综的系统，系综愈大，所含的分子数目愈多，误差愈小，统计结果愈佳。

## B-3 配分函数

各种不同的系综有其对应的配分函数（partition function）。依统计力学，系统的特性可利用配分函数算出。正则系综的配分函数为：

$$Q(N,V,T)=\sum_i e^{-E_i/k_BT} \tag{B-8}$$

式中，$k_B$ 为玻耳兹曼常数（Boltzmann constant），$k_B=1.3806\times10^{-16}\text{erg/(mol}\cdot\text{K)}$；$E_i$ 为（$N,V,T$）系统的能量。系统能量为 $E_i$ 的概率为：

$$P_i=\frac{e^{-E_i/k_BT}}{\sum_j e^{-E_j/k_BT}}=\frac{e^{-E_i/k_BT}}{Q} \tag{B-9}$$

设 $E_i < E_k$，对应于 $E_i$ 的系统数目为 $n_i$，对应于 $E_k$ 的系统数目为 $n_k$，则：

$$\frac{\bar{n}_k}{\bar{n}_i}=\frac{P_k}{P_i}=\frac{e^{-E_k/k_BT}}{e^{-E_i/k_BT}}=e^{-(E_k-E_i)/k_BT}=e^{-\Delta E/k_BT} \tag{B-10}$$

式中，$\Delta E=E_k-E_i$，为两种状态能量的差值；$\bar{n}_k=\langle n_k\rangle$ 为平均值的符号。

由式(B-10)，正则系综的任何物理量 $B$ 的平均值为：

$$\bar{B}=\sum_i P_i B_i=\sum_i \frac{B_i e^{-E_i/k_BT}}{Q}=\frac{1}{Q}\sum_i B_i e^{-E_i/k_BT} \tag{B-11}$$

利用此关系及热力学公式，可推导出正则系综 $Q(N,V,T)$ 的各种热力特性，如亥姆霍兹自由能（Helmholtz free energy）$A$、熵（entropy）$S$、压力 $p$、化学势能等。由统计力学计算的为系综的力学平均值，但若系综够大，数目够多，则此力学的平均值与热力学度量的值一致。

$$\bar{A}=-k_BT\ln Q$$

$$\bar{E}=k_BT^2\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V}$$

$$\bar{S}=k_B\ln Q+k_BT\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V}$$

$$\bar{p}=k_BT\left(\frac{\partial \ln Q}{\partial V}\right)_{N,T}$$

$$\bar{\mu}=-k_BT\left(\frac{\partial \ln Q}{\partial N}\right)_{V,T} \tag{B-12}$$

配分函数为统计力学的中心，由以上公式可知，若知道系统的配分函数，则可经由计算得到系统的各种特性，表 B-1 列出几个系综的热力学性值与配分函数的关系。

**表 B-1　各种热力学性质与配分函数的关系**

| 微正则系综 $\Omega(N,V,E)$ | 巨正则系综 $\Xi(V,T,\mu)$ | 等温恒压系综 $\Delta(N,T,p)$ |
|---|---|---|
| $\bar{S}=k_B\ln\Omega$ | $\bar{p}V=k_BT\ln\Xi$ | $\bar{G}=-k_BT\ln\Delta$ |
| $\frac{1}{k_B\bar{T}}=\left(\frac{\partial\ln\Omega}{\partial E}\right)_{N,V}$ | $\bar{S}=k_B\ln\Xi+k_BT\left(\frac{\partial\ln\Xi}{\partial T}\right)_{V,\mu}$ | $\bar{S}=k_B\ln\Delta+k_BT\left(\frac{\partial\ln\Delta}{\partial T}\right)_{N,p}$ |
| $\frac{\bar{p}}{k_BT}=\left(\frac{\partial\ln\Omega}{\partial V}\right)_{N,E}$ | $\bar{N}=k_BT\left(\frac{\partial\ln\Xi}{\partial\mu}\right)_{V,T}$ | $\bar{V}=-k_BT\left(\frac{\partial\ln\Delta}{\partial p}\right)_{N,T}$ |
| $\frac{\bar{\mu}}{k_BT}=-\left(\frac{\partial\ln\Omega}{\partial N}\right)_{V,E}$ | | $\bar{\mu}=-k_BT\left(\frac{\partial\ln\Delta}{\partial N}\right)_{T,p}$ |

# B-4 统计涨落

度量的样品值与平均值差异的幅度称为统计涨落（fluctuation）。依系综的观念，统计涨落为系统各别的值与整个系综平均值差异的幅度。以正则系综为例，由式(B-4)，能量的方差为：

$$\sigma_{\mathrm{E}}^2=\overline{(E-\bar{E})^2}=\bar{E}^2-\bar{E}^2 \tag{B-13}$$

由式(B-2)，式(B-13) 可改写为

$$\sigma_{\mathrm{E}}^2=\sum_j P_{\mathrm{j}}E_{\mathrm{j}}^2-\bar{E}^2 \tag{B-14}$$

由式(B-12) 平均值的表示式，可得：

$$\begin{aligned}\sum_j P_{\mathrm{j}}E_{\mathrm{j}}^2&=\frac{1}{Q}\sum_j E_{\mathrm{j}}^2\mathrm{e}^{-E_{\mathrm{j}}/k_{\mathrm{B}}T}=-\frac{k_{\mathrm{B}}}{Q}\frac{\partial}{\partial(1/T)}\sum_j E_{\mathrm{j}}\mathrm{e}^{-E_{\mathrm{j}}/k_{\mathrm{B}}T}\\&=-\frac{k_{\mathrm{B}}}{Q}\frac{\partial}{\partial(1/T)}(\bar{E}Q)=-k_{\mathrm{B}}\frac{\partial\bar{E}}{\partial(1/T)}-k_{\mathrm{B}}\bar{E}\frac{\partial\ln Q}{\partial(1/T)}\\&=k_{\mathrm{B}}T^2\frac{\partial\bar{E}}{\partial T}+\bar{E}^2\end{aligned} \tag{B-15}$$

将此结果代入式(B-14)，得能量的统计涨落与能量的平均值的关系，即：

$$\sigma_{\mathrm{E}}^2=k_{\mathrm{B}}T^2\left(\frac{\partial\bar{E}}{\partial T}\right)_{N,V} \tag{B-16}$$

若将 $\bar{E}$ 视为热力学的能量，则 $\left(\frac{\partial E}{\partial T}\right)_{N,V}=C_{\mathrm{V}}$，为分子定容比热容（molar heat capacity of constant volume）。由此得方差为：

$$\sigma_{\mathrm{E}}^2=k_{\mathrm{B}}T^2C_{\mathrm{V}} \tag{B-17}$$

能量的相对误差为标准误差除以平均能量，即：

$$\frac{\sigma_{\mathrm{E}}}{\bar{E}}=\frac{(k_{\mathrm{B}}T^2C_{\mathrm{V}})^{1/2}}{\bar{E}} \tag{B-18}$$

以单原子理想气体为例，常温下，其能量与定容比热容的值分别为：

$$\bar{E}=\frac{3}{2}Nk_{\mathrm{B}}T,\ C_{\mathrm{V}}=\frac{3}{2}Nk_{\mathrm{B}} \tag{B-19}$$

故 $\frac{\sigma_{\mathrm{E}}}{\bar{E}}$ 的量级为 $O(N^{-1/2})$。由此，系统的分子数目愈多，相对误差愈

小。利用统计涨落可得到其他的一些热力学性质，如巨正则系综（$V,T,\mu$）中分子数目的统计涨落为：

$$\sigma_{\mathrm{N}}^2=\frac{\bar{N}^2 k_{\mathrm{B}} T\kappa}{V} \tag{B-20}$$

式中，$\kappa$ 为等温压缩系数（isothermal compressibility），即：

$$\kappa=-\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{N,T} \tag{B-21}$$

等温恒压系综（$N,T,p$）中，其焓的统计涨落为：

$$\sigma_{\mathrm{H}}^2=k_{\mathrm{B}} T^2 C_{\mathrm{p}} \tag{B-22}$$

式中，$H$ 为系统的焓（enthalpy）；$C_{\mathrm{p}}$ 为分子定压比热容（molecular heat capacity at constant pressure）。

## B-5 统计力学与模拟计算的关系

系综各种特性的平均值称为系综平均值（ensemble average）。若特性 $x$ 为连续的值，（如经典力学的动能，速度……），则：

$$\langle x\rangle=\int_{-\infty}^{\infty} x w(x)\mathrm{d}x \tag{B-23}$$

式中，积分的范围为所有可能的 $x$ 值，$w(x)\mathrm{d}x$ 称为 $x$ 的概率配分函数（probability distribution function）。$w(x)\mathrm{d}x$ 的意义为于范围 $x\to x+\mathrm{d}x$ 间 $x$ 的概率。例如麦克斯韦-玻耳兹曼配分函数（Maxwell-Boltzmann distribution function），为理想气体速度的配分函数，即：

$$w(v_{\mathrm{x}},v_{\mathrm{y}},v_{\mathrm{z}})\mathrm{d}v_{\mathrm{x}}\mathrm{d}v_{\mathrm{y}}\mathrm{d}v_{\mathrm{z}}=\left(\frac{m}{2\pi k_{\mathrm{B}} T}\right)^{3/2}\exp\left\{-\frac{m}{2k_{\mathrm{B}} T}(v_{\mathrm{x}}^2+v_{\mathrm{y}}^2+v_{\mathrm{z}}^2)\right\}\mathrm{d}v_{\mathrm{x}}\mathrm{d}v_{\mathrm{y}}\mathrm{d}v_{\mathrm{z}} \tag{B-24}$$

由此配分函数，可得各方向速度平方的平均值：

$$\langle v_{\mathrm{x}}^2\rangle=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v_{\mathrm{x}}^2 w(v_{\mathrm{x}},v_{\mathrm{y}},v_{\mathrm{z}})\mathrm{d}v_{\mathrm{x}}\mathrm{d}v_{\mathrm{y}}\mathrm{d}v_{\mathrm{z}}=\frac{k_{\mathrm{B}} T}{m}$$

$$\langle v_{\mathrm{x}}^2\rangle=\langle v_{\mathrm{y}}^2\rangle=\langle v_{\mathrm{z}}^2\rangle=\frac{k_{\mathrm{B}} T}{m} \tag{B-25}$$

一个系统的特性 $x$ 亦可能与时间相关，$x(t)$。由时间所得的平均

值为：

$$\bar{x} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)\mathrm{d}t \tag{B-26}$$

式中，积分表示所有时间的范围。依据各态历经假设（ergodic hypothesis），系综的平均值与长时间的平均值应相等，即：

$$\bar{x} = \langle x \rangle \tag{B-27}$$

依此假说，平均值可由许多系统的特性计算，或是计算单独一个系统长时间的平均值。模拟计算系统的特性所依据的即是统计学的原理，蒙地卡罗模拟（Monte Carlo simulation，MC）为计算系综的平均值，而分子动力模拟（molecular dynamics simulation）与布朗动力模拟（Brownian dynamics simulation）为计算单独一个系统长时间的平均值。

## B-6 统计力学与量子力学的关系

量子力学可准确的计算出单独分子的能量。利用这些分子的能量可以计算气态分子系统的配分函数。以双原子分子系统为例，根据量子力学所得双原子分子 $AB$ 的移动，振动与转动的能量，可得分子的配分函数为：

$$q_{\mathrm{trans}} = \left(\frac{2\pi M k_{\mathrm{B}} T}{h^2}\right)^{3/2} V$$

$$q_{\mathrm{vib}} = \frac{\mathrm{e}^{-\frac{1}{2}\frac{h\nu}{k_{\mathrm{B}}T}}}{1-\mathrm{e}^{-\frac{h\nu}{k_{\mathrm{B}}T}}}$$

$$q_{\mathrm{rot}} = \frac{8\pi^2 I k_{\mathrm{B}} T}{h^2} \tag{B-28}$$

利用分子的配分函数与 $\left(\frac{\partial E}{\partial T}\right)_{N,V} = C_{\mathrm{V}}$ 的关系，可计算理想双原子分子气体的各种热力学性质。例如，分子振动的定容比热容为：

$$C_{\mathrm{V,vib}} = nR\left(\frac{h\nu}{k_{\mathrm{B}}T}\right)^2 \frac{\mathrm{e}^{\frac{h\nu}{k_{\mathrm{B}}T}}}{(\mathrm{e}^{\frac{h\nu}{k_{\mathrm{B}}T}}-1)^2} \tag{B-29}$$

式中，$R$ 为理想气体常数。图 B-2 为计算的摩尔振动比热容，图中显示高温时，$\bar{C}_{\mathrm{V,vib}} \sim R$。

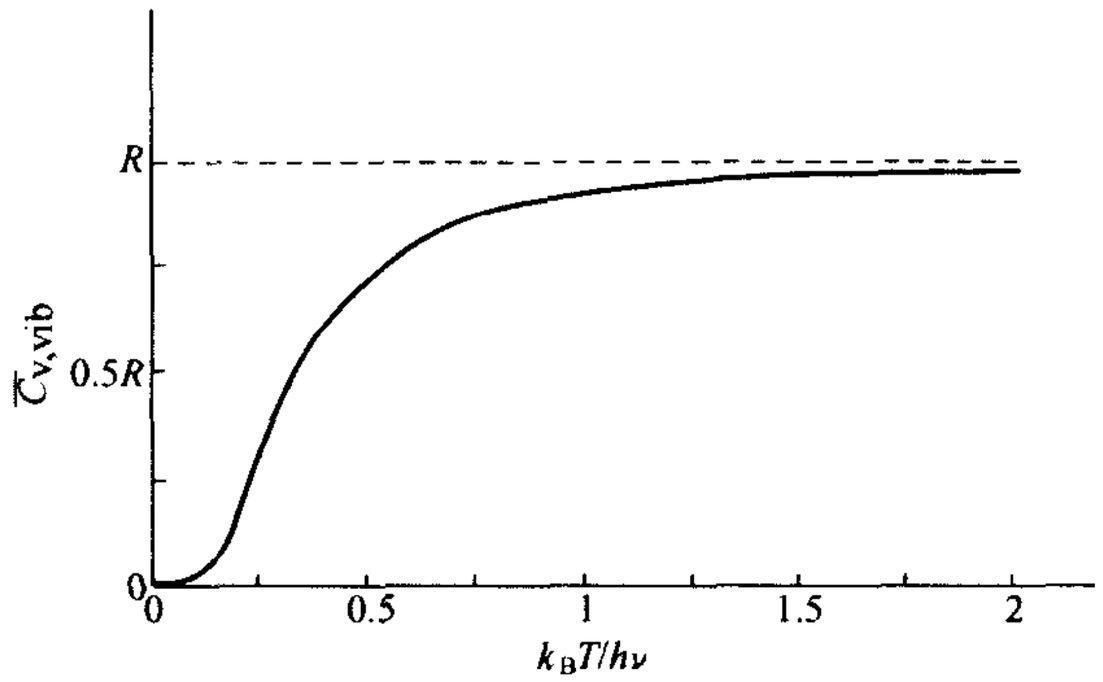



图 B-2　双原子分子气体的摩尔振动比热容与温度的关系图

# B-7　经典力学与统计力学的关系

统计力学主要的依据为配分函数，由量子力学可准确的计算出分子的能阶，因此可利用 $x$ 式得到气体的各项热力学性质。但对于液体，固体等系统，分子间的作用力不容忽略，系统的配分函数不能由其成分分子的配分函数组合而成。依量子力学的方法，计算这样系统的能阶必须解整个系统的薛定谔方程式（Schrödinger equation），而系统通常含有大量的分子数目，以目前量子力学的发展尚无法办到。所幸，系统的配分函数除了由量子力学得到外，还可利用经典力学的方法。在高温与低压的状态下（分子间作用力小），这两种方法所得的配分函数结果一致。依经典力学，系统的哈密顿量（Hamiltonian）为：

$$H=\sum_{i=1}^{N}\frac{1}{2m_i}(P_{x,i}^2+P_{y,i}^2+P_{z,i}^2)+V(x_1,y_1,z_1,\cdots,x_N,y_N,z_N) \tag{B-30}$$

系统的配分函数为：

$$Q=\frac{1}{h^{sN}}\int\cdots\int e^{-\frac{H}{k_BT}}dP_{x,1}\,dP_{y,1}\,dP_{z,1}\cdots dP_{x,N}dP_{y,N}dP_{z,N}dx_1\,dy_1\,dz_1\cdots dx_N\,dy_N\,dz_N \tag{B-31}$$

式中，$s$ 为每个分子所具有的自由度（degree of freedom）；$h$ 为普朗克常数（planck constant）。利用此配分函数，计算任何物理量的平均

值为：

$$\langle A\rangle = \frac{\iint A e^{-\frac{H}{k_B T}} \mathrm{d}p\mathrm{d}q}{\iint e^{-\frac{H}{k_B T}} \mathrm{d}p\mathrm{d}q} \tag{B-32}$$

式中，$\mathrm{d}p\mathrm{d}q$ 为所有动量与坐标的积分。依此，由系综各个系统的分子于不同位置及不同动量的能量（$H$）可计算 $A$ 的平均量。这样的平均量称为系综平均量（ensemble average)。实际应用上，可以利用不同的方法，产生系综的各种系统（如蒙地卡罗方法），计算每一系统 $A$ 的量，再应用式(B-32) 求出其平均量。

根据各态历经假设，系综的平均量应等于单一系统长时间的平均量。由此，可由分子动力模拟的方法，选定一系统，计算各时间 $A$ 的量，而得到统计的平均值，即：

$$\overline{A} = \lim_{T\to\infty} \frac{\sum A(t)}{T} \tag{B-33}$$

在分子模拟计算中，统计力学应用很广。几乎所有的性质计算都必须引用统计力学的观念。本章介绍的为平衡系统的统计力学，所计算的统计量与时间无关。此外，尚有非平衡系的统计力学，所计算的统计量与时间有关联性。非平衡系的统计力学在计算动力特性时非常重要。有兴趣的读者可参考 McQuire 的著作“Statistical Mechanics”等书。

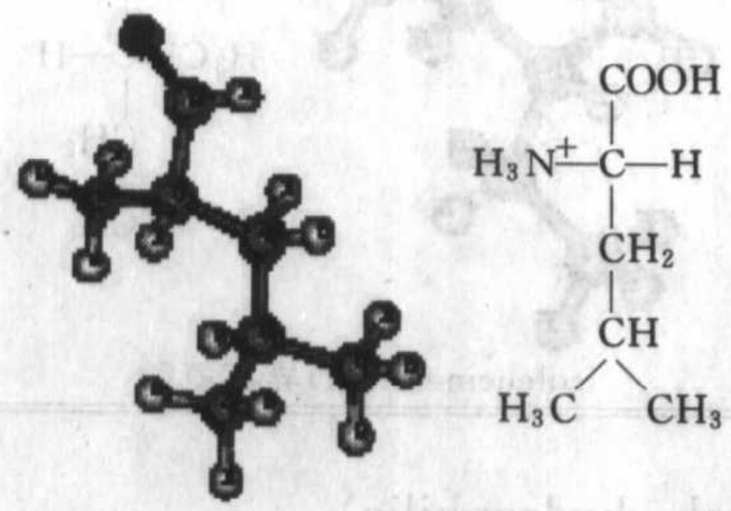
## 参 考 文 献


1 McQuire D A. “Statistical Thermodynamics”, Harper Row, New York, USA, 1973

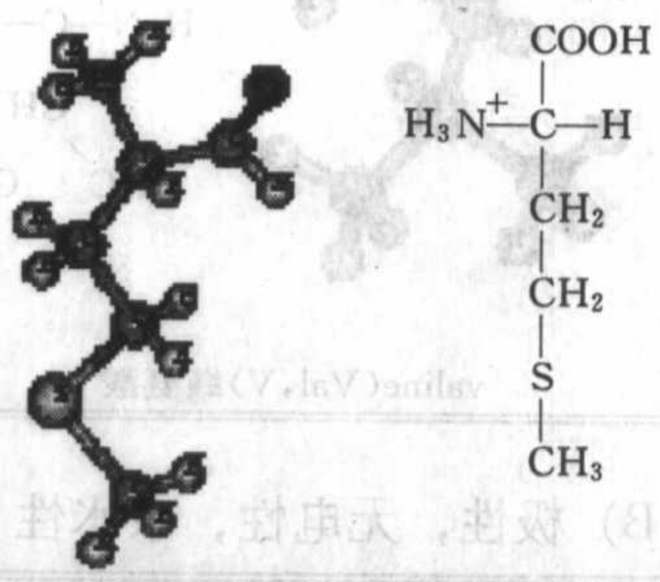
2 McQuire D A. “Statistical Mechanics”, Harper Row, New York, USA, 1973

# 附录C
# 天然存在的氨基酸的结构

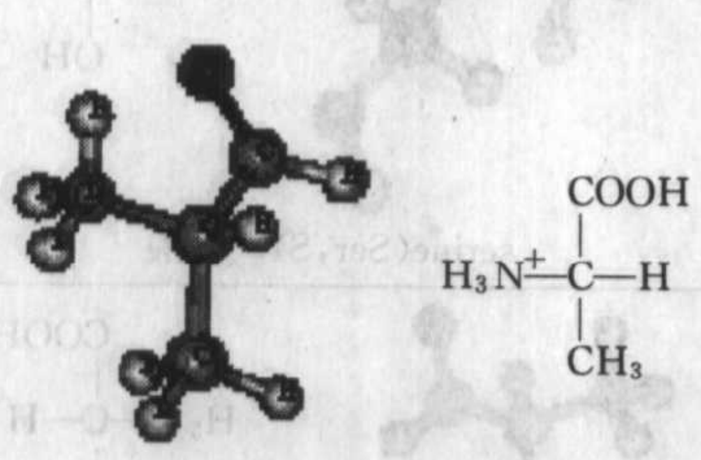
## (A) 非极性-疏水性 (non-polar, hydrophobic)



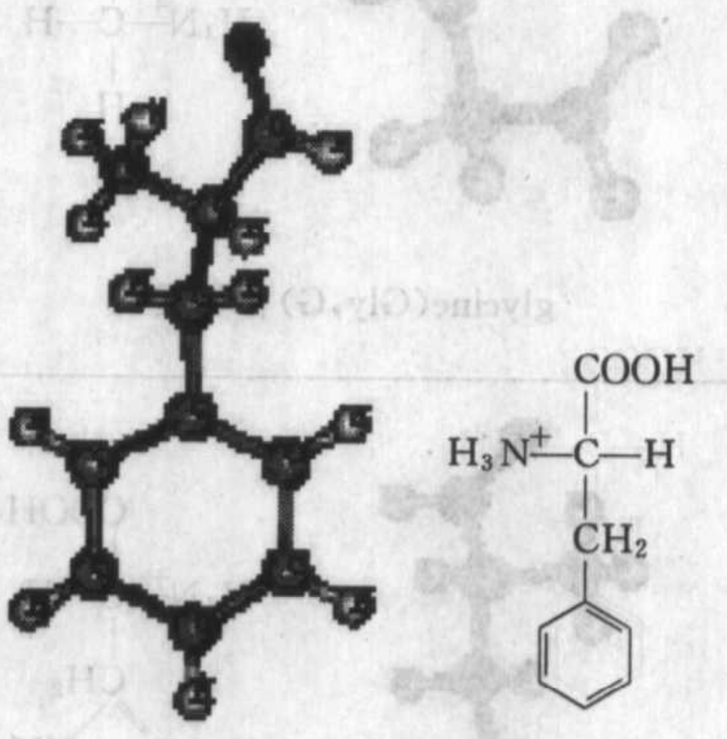

leucine (Leu, L) 亮氨酸



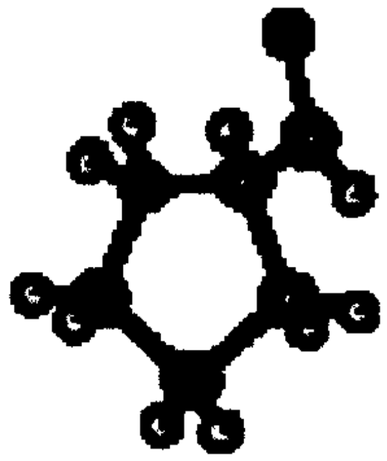

methionine(Met, M) 甲硫氨酸



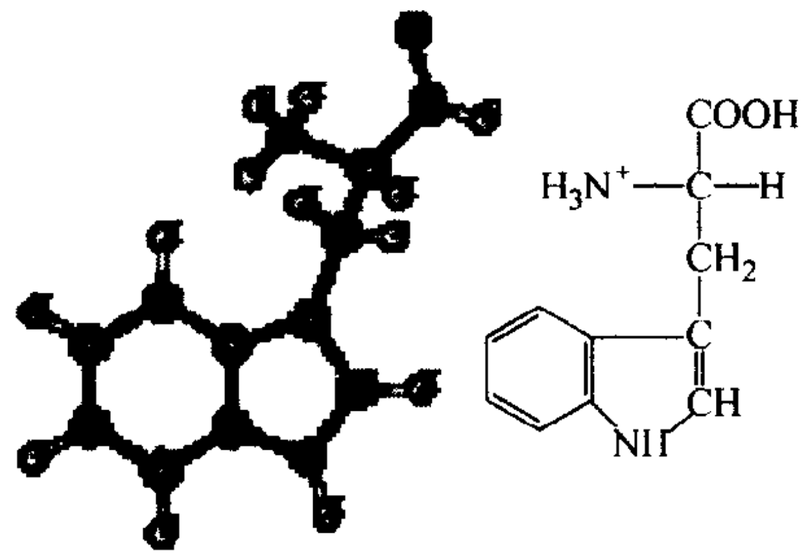

alanine(Ala, A)丙氨酸



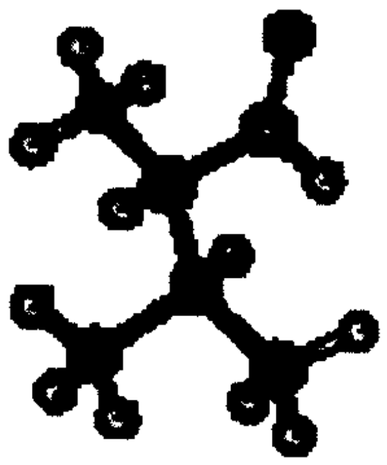

phenylalanine(Phe, F)苯基丙氨酸



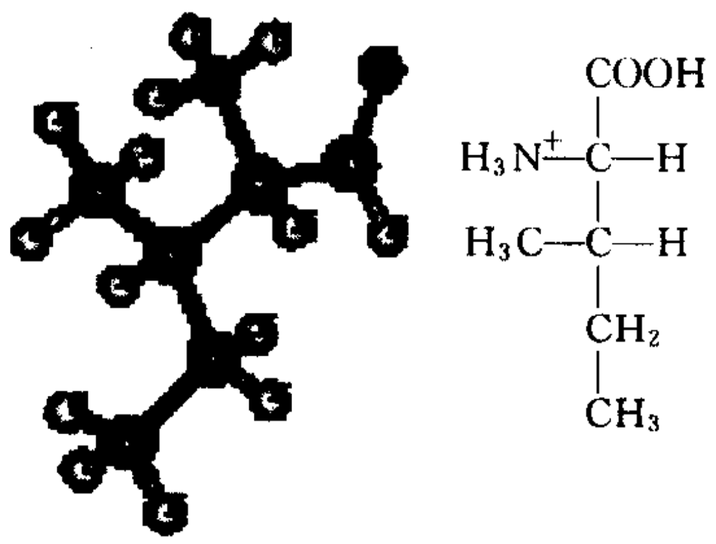
proline(Pro,P)脯氨酸



tryptophan(Trp, W)色氨酸

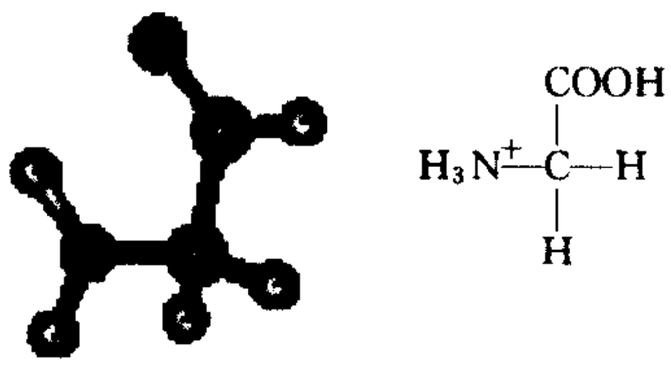


valine(Val,V)缬氨酸

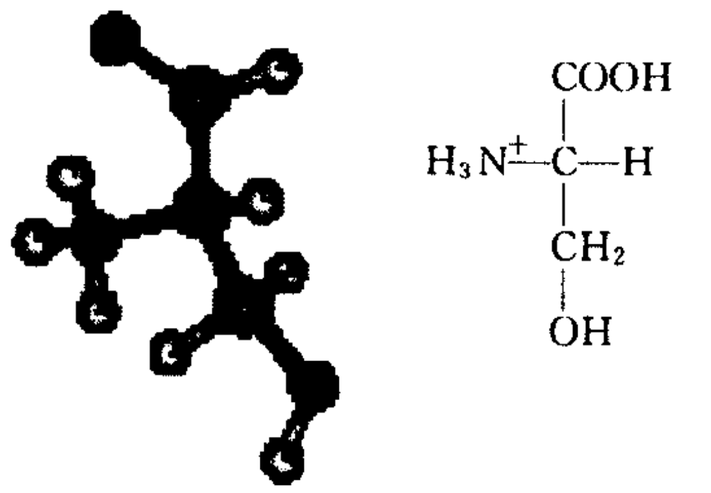


isoleucine(Ile,I)异亮氨酸

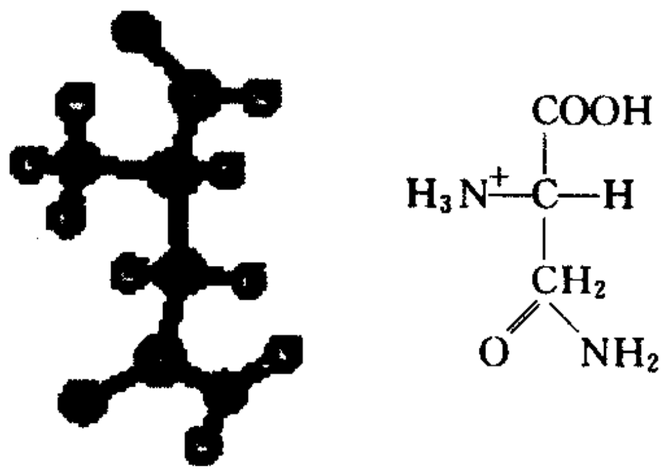
(B) 极性，无电性，亲水性 (polar, neutral, hydrophilic)



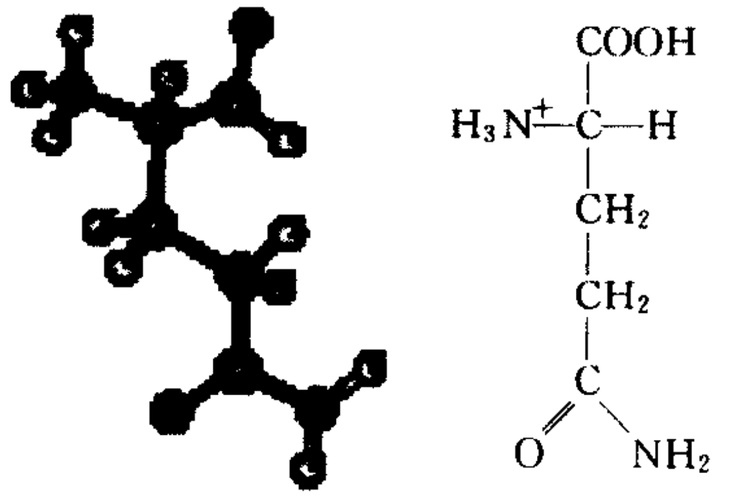
glycine(Gly,G)甘氨酸



serine(Ser,S)丝氨酸

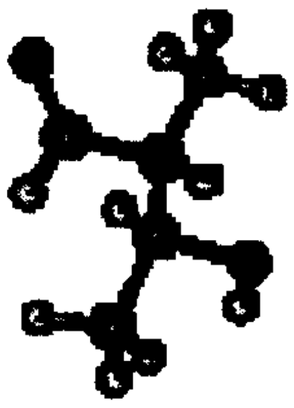


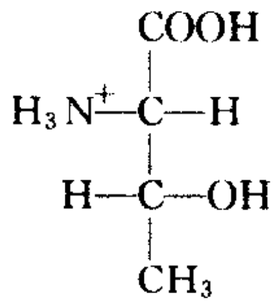
asparagine(Asn,N)天冬酰胺



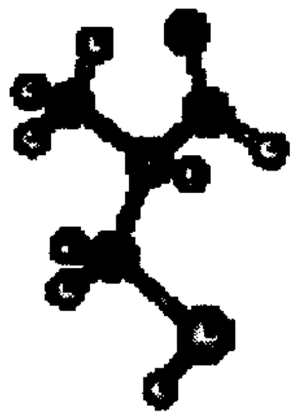
glutamine(Gln,Q)谷氨酰胺

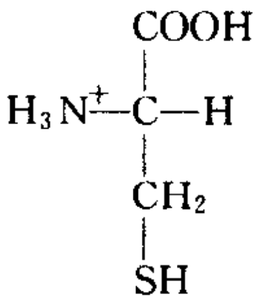
2



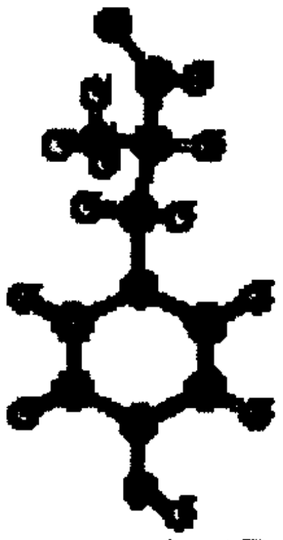


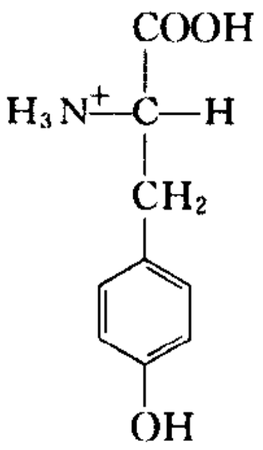
threonine(Thr,T)苏氨酸





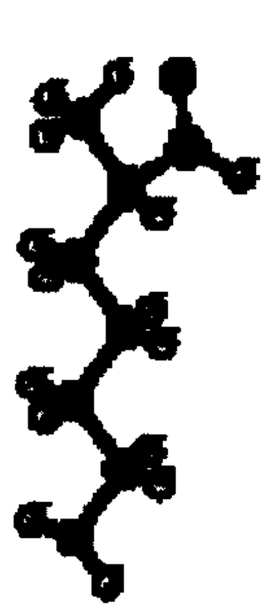
cysteine(Cys,C)半胱氨酸

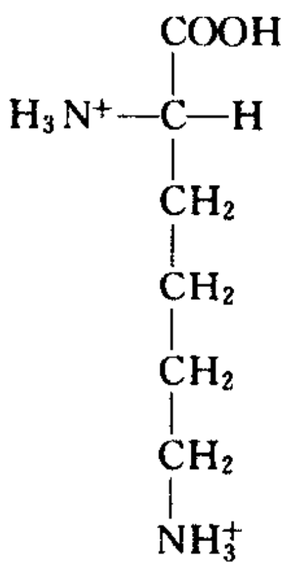




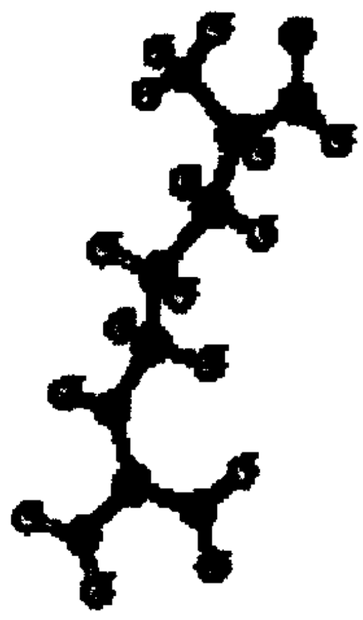
tyrosine(Tyr,Y)酪氨酸

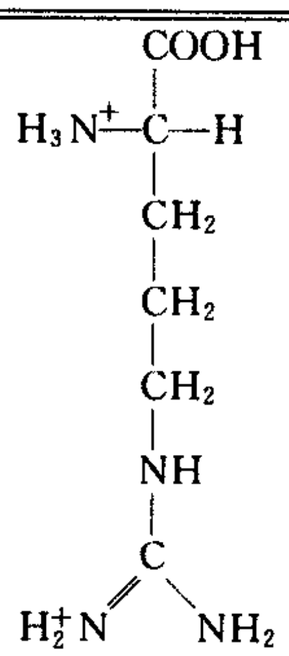
## (C) 极性-碱性-亲水性 (polar, basic, hydrophilic)



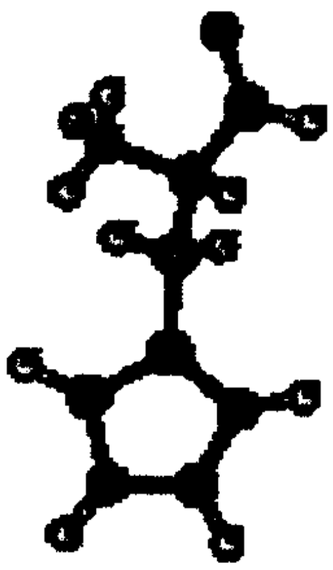


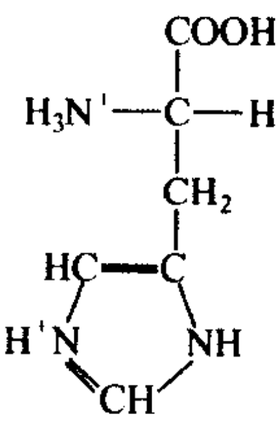
lysine(Lys,K)赖氨酸





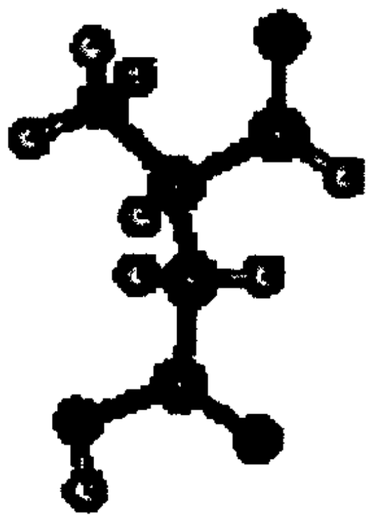
arginine(Arg,R)精氨酸

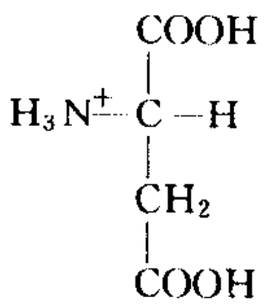




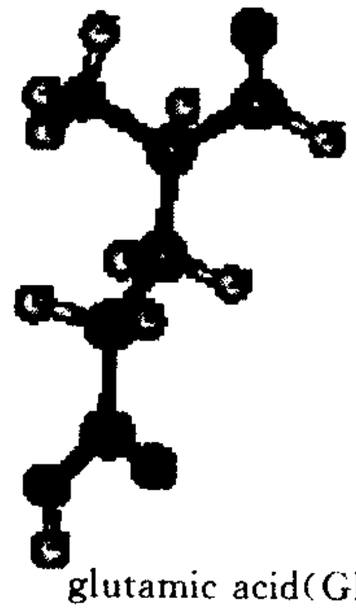
histidine(His,H)组氨酸

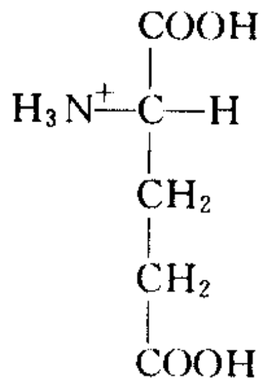
（D）极性-酸性-亲水性（polar，acidic，hydrophilic）






aspartic acid(Asp,D)天冬氨酸






glutamic acid(Glu,E)谷氨酸

# 索引（按汉语拼音排序）


## A

阿伏伽德罗常数（Avogadro's number） 258
阿仑尼乌斯（Arrhenius） 230
氨基酸序列（amino acid sequence） 269
鞍点（saddle point） 51

## B

八极（octapole） 153
半经验分子轨道计算法（semi-empirical molecular orbital method） 2
棒模式（bead-rod model） 197
保守力（conservative force） 240
本征值（eigen value） 51
比较分子力场方法（comparative molecular field analysis，CoMFA） 289
比较分子相似性指数分析法（comparative molecular similarity indices analysis，CoMSIA） 289
边界元素（boundary element） 263
扁盘形（orblate） 102
变分原理（variation principle） 2
遍历链（ergodic chain） 171
标准生成焓（standard enthalpy of formation） 59
标准误差（standard deviation） 325
波函数（wave function） 1
玻恩-奥本海默近似（Born-Oppenheimer approximation） 3
玻耳兹曼常数（Boltzmann constant） 326
玻璃转换温度（glass transition temperature） 147

泊松-玻耳兹曼有限差分方法（finite difference Possion-Boltzmann，FDPB） 263
布朗动力（Brownian dynamics） 182
布朗动力计算（Brownian dynamics，BD） 223
布朗动力模拟（Brownian dynamics simulation，BD） 4，330
布朗粒子（Brownian particle） 221
布朗运动（Brownian motion） 182
部分电荷量（partial charge） 18
部分原子电荷（partial atomic charge） 152

## C

残基（basic residues） 219
测不准原理（uncertainty principle） 1
层状型液晶（sematic liquid crystal） 203
插入（insert） 215
缠绕（chain crossing） 199
超额化学势（excess chemical potential） 179
超额自由能（excess free energy） 137
成对加成（pair-wise additive） 107
重叠构象（eclipse） 50
从头设计（de Novo） 300
从头算计算法（ab initio method） 2

## D

单体（monomer） 189
蛋白质折叠（protein folding） 217
得布罗意波长（de Broglie wavelength） 178
等温定压系综（isothermal-isobaric ensemble） 326
等温吸附线（isotherm） 212
等温压缩系数（isothermal compressibility） 329
等效构象（equivalent conformation） 139
第一媒合层（primary solvation shell） 180
电离能（ionization potential） 1
电子密度（electron density） 1
电子亲和力（electron affinity） 1
电子顺磁共振（electron paramagnetic resonance，EPR） 280
迭代计算（iteration） 200
迭代计算法（iterative method） 98

定标（scaling） 209
定量结构-活性关系（quantitative structure-activity relationship，QSAR） 288
定容比热容（heat capacity at constant volume） 108
定温定压系综（isothermal-isobaric ensemble） 176
度规张量（metric tensor） 84
度量化坐标（scaled coordinates） 83，177
对称伸缩（symmetric stretch） 40
对称数（symmetry number） 141
对角线化（diagonalization） 51
对角线块状牛顿-拉森法（Block-diagonal Newton-Raphson method） 33
对位交叉构象（staggered） 50
多离子体（ionphore） 260
多肽（polypeptide） 267
多糖类（polysaccharides） 16
多维势能面（multi-dimensional potential surface） 58
多重线性回归法（multiple linear regression method） 295

## E

二次区域（quadratic region） 54
二极（偶极，dipole） 153
二聚体（dimmer） 134
1,2-二氯乙烷（1,2-dichloroethane，DCE） 106
二面角（dihedral angle） 7
二面角扭曲项（torsion angle term） 7
二肽（dipeptide） 267

## F

反复试验（trial and error） 187
反空间（reciprocal space） 158
反六角形（inverted-hexagonal） 250
反式构象（anti） 50
反向量（reciprocal vector） 159
反应场大小（magnitude of reaction field） 257
反应能量曲线（energy profile） 52
反应途径（reaction pathway） 52
反应坐标（reaction coordinate） 52
范德瓦耳斯作用力（van der Waals force） 7

方差（variance） 324
非定域化（delocalization） 1
非对称伸缩（asymmetric stretch） 40
非极性长距离（apolar long-range） 205
非键结（non-bonding） 7
非均向性（anisotropic） 19
非均向性形状参数（anisotropic shape parameter） 20
非线性光学（non-linear optics） 42
沸石（zeolite） 18，212
分布（distribution） 325
分界（boundary） 263
分子定容比热容（molar heat capacity of constant volume） 328
分子定压比热容（molecular heat capacity at constant pressure） 329
分子动力模拟（molecular dynamics simulation，MD） 4，330
分子洞穴（cavity） 256
分子对接（molecular docking） 300
分子固定轴（molecular-fixed axis） 92
分子轨道（molecular orbital） 2
分子间自由能（inter free energy） 62
分子力学（molecular mechanics，MM） 24
分子力学方法（molecular mechanics，MM） 3
分子内自由能（intra free energy） 62
分子形状法（molecular shape analysis，MSA） 291
分子异构体（isomer） 58
封闭系统（closed system） 209

## G

概率（probability） 324
概率配分函数（probability distribution function） 329
刚性分子（rigid molecule） 90
高斯分布（Gaussian distribution） 78，325
高斯随机运动过程（Gaussian random process） 223
格胞多极法（cell multipole method，CMM） 161
各向同性（isotropic） 224
功率谱（power spectrum） 118
共轭（conjugate） 323
共轭梯度法（conjugate gradient） 29

构群加成法（group increment method） 59
构群加成能（group increment energy，GE） 59
构象（conformation） 1，58，122
构象能（steric energy，SE） 58
构象异构体（conformer） 49，58
构形自由能（conformational free energy） 129
构型（configuration） 169
构型能量（configuration energy） 179
构型选择（configurational-bias） 212
固化点（solidify point） 143
寡肽（oligopeptide） 267
归一化（normalized） 114
归一化因子（normalization factor） 325
硅酸沸石（zeolite silicalite） 212
过渡态（transition state） 130

## H

哈密顿量（Hamiltonian） 121，331
哈密顿运动力学（Hamiltonian dynamics） 322
亥姆霍兹自由能（Helmholtz free energy） 128，186，327
焓（enthalpy） 58，120
耗散力（dissipative force） 240
耗散粒子动力计算（dissipative particle dynamics，DPD） 240
核磁共振（nuclei magnetic resonance，NMR） 279
宏观（macroscopic） 324
虎克定律（Hook's law） 39，317
互补错误函数（complementary error function） 158
化学势能（chemical potential） 177，210，326
缓慢成长法（slow growth method） 127
回转半径（gyration radius） 100，190

## J

迹（trace） 191
积分步程（integration time step） 75
吉布斯系综（Gibbs ensemble） 210
吉布斯自由能（Gibbs free energy） 62，128
极化性（polarisability） 264

极限值（limiting value） 172
极性长距离（polar long-range） 205
几何优化构象（geometry optimized configuration） 24
计算机集群（PC cluster） 4
坚硬度（hardness） 18
剪切应力（shear stress） 109
简并度（degeneracy） 37
简单点电荷力场（simple point charge，SPC） 133
简化单位（reduced unit） 76
简化质量（reduced mass） 75
简正模式（normal mode） 37
简正频率（normal frequency） 38
简正坐标（normal coordinate） 37
键角弯曲项（angle bending term） 7
键结亲合能（binding affinity） 260
键伸缩项（bond stretching term） 7
交互相关函数（cross correlation function） 114
接受（accept） 173
结构（structure） 1
结构积分（configuration integral，$Z$） 128，169
截断半径（cut-off radius） 73
截断项（truncation） 163
解折叠（unfolding） 275
介电常数（dielectric constant） 124，152，164
介电质（dielectric） 259
介观（meso-scale） 240
近似牛顿-拉森法（Quasi Newton-Raphson method） 33
晶格模式（lattice model） 188
径向方程式（radial equation） 319
径向分布函数（radial distribution function，RDF） 110
净原子电荷（net atomic charge） 152
静压（hydrostatic pressure） 109
镜像电荷（image charge） 160
局部结构（local structure） 192
巨正则系综（grand canonical ensemble） 177，326
聚集体（aggregates） 242
聚醚醚酮（polyether ether ketone，PEEK） 104

聚碳酸酯类（polycarbonates） 16
卷积定理（convolution theory） 115
绝对自由能（absolute free energy） 129
绝热投像（adiabatic mapping） 56
均方根位移（root mean square displacement，RMSD） 182
均方根误差（root-mean-square deviation，RMS） 36
均方位移（mean square displacement，MSD） 112
均一分布（uniform distribution） 187

## K

开关函数（switching function） 15，74
可调步程寻找法（arbitrary step approach） 31
可用表面积（accessible surface area） 226，261
可转移的分子间势能（transferable intermolecular potential，TIP） 133
客体分子（guest molecules） 212
空间固定轴（space-fixed axis） 92
空间位阻能量（steric energy） 49
孔洞（micropores） 212
库仑作用项（columbic interaction term） 8
快速多极法（fast multipole method） 161
快速傅里叶变换（fast Fourier transform，FFT） 115
扩散区间（diffusive regime） 224
扩散系数（diffusion constant） 113，182

## L

拉格朗日乘子（Lagrangian multiplier） 18
拉格朗日未定乘子（Lagrange undetermined multiplier） 91
拉格朗日运动方程式（Lagrange equation） 320
拉格朗日运动力学（Lagrange dynamics） 319
拉伸应力（tensile stress） 109
朗格缪尔薄膜（Langmuir film） 144
朗格缪尔-布罗杰特膜（Langmuir-Blodgett，LB） 145
朗之万（Langevin） 222
勒让德多项式（Legendre polynomial） 119，235
离平面振动项（out-of-plane bending term） 8
离子强度（ionic strength） 262
力常数矩阵（force constant matrix） 29

力场（force field） 3
力场导向蒙地卡罗（force bias MC） 181
立方晶格（cubic lattice） 188
立体效应（steric effect） 64
联合原子模型（united atom model） 99
链含过程（chain-enrichment process） 194
链状双性体（amphiphiles） 144
量子力学（quantum mechanics） 1
“劣质”溶剂（bad solvent） 190
邻位交叉式构象（gauche） 50
临界密度（critical density） 212
临界温度（critical temperature） 212
零点能量（zero point energy） 37
零点振动能（zero-point energy） 58
流动效应（hydrodynamic effect） 227
六角形（hexagonal） 249
螺旋度（helicity） 278
螺旋状（gyroid） 249

## M

马克夫链（Markov chain） 170
麦克斯韦-玻耳兹曼分布（Maxwell-Boltzmann distribution） 80
麦克斯韦-玻耳兹曼配分函数（Maxwell-Boltzmann distribution function） 329
麦克斯韦分配（Maxwell distribution） 223
蒙地卡罗（Monte Carlo，MC） 167
蒙地卡罗计算法（Monte Carlo method） 3
蒙地卡罗模拟（Monte Carlo simulation，MC） 330
密度泛函数理论（density functional theory，DFT） 3
明智蒙地卡罗（smart Monte Carlo） 183
摩擦力（frictional force） 221
末端距（end-to-end distance） 100，190

## N

内能（internal energy） 58
内坐标（internal coordinate） 24，322
能量最小化（energy minimization） 24
能量最优化（energy optimization） 26

黏度（viscosity） 120
牛顿-拉森法（Newton-Raphson method） 33
牛顿运动方程式（Newtonian equation of motion） 317
扭曲结构（twist） 50
扭转角驱动法（torsion angle driving） 56

## O

欧拉角（Euler angle） 92
偶极（dipole） 11
偶极矩（dipole moment） 1
耦合（coupling） 318
耦合系数（coupling constant） 123

## P

排除（reject） 173
排除体积（excluded volume） 192
排列方向（director） 205
配分函数（partition function） 121，178，326
配位数（coordination number） 243
片段（segment） 242
偏最小二乘法（partial least squares，PLS） 294
平衡（equilibrium） 172
平均密度（bulk density） 111
平均值（average） 324
普朗克常数（planck constant） 331

## Q

期望值（expectation value） 324
亲水性（hydrophilic） 144
倾斜度（steepness） 28
球与弹簧模式（bead-spring model） 199
球状结构（globular structure） 197
区域密度（local density） 111
去耦合（decoupling） 318
权重因子（weighting factor） 241
全对称伸缩（totally symmetric stretch） 119

## R

热导率（thermal conductivity coefficient） 120
热力摩擦系数（thermodynamic friction coefficient） 82
热力微扰法（thermodynamic perturbation method） 122
热平衡（thermal equilibrium） 79
溶剂化（solvation） 264
溶剂化自由能（solvation free energy） 14
溶剂介电模型（solvent dielectric model） 164
溶解自由能（solvation free energy，$\Delta G_{sol}$） 256
柔软度（flexibility） 191

## S

三聚氰氨树脂（melamine resins） 16
散乱状态（random-coil） 198
散射（scattering） 119
熵（entropy） 62，209，327
蛇形滑动模式（slithering snake model） 194
神经网络算法（neural network algorithm） 296
升华热（heat of sublimation） 63
生成焓（enthalpy of formation） 58
生命期（life time） 230
势能面（potential energy surface） 24
适应度函数（fitness function） 296
疏水效应（hydrophobic effect） 275
疏水性（hydrophobic） 144，275
数目密度（number density） 262
数值密度（number density） 146
双变量（bivariant） 225
双螺旋核酸（double-strand nucleic acid） 228
双曲线正弦函数（hyperbolic sine function） 262
双向取样（double-wide sampling） 125
双性（amphiphilic） 242
水合层（hydration shell） 117
水合能（hydration energy） 81
水合数目（hydration number） 112
斯托克斯定律（Stokes' law） 221

四极（quadrupole） 153
四面体晶格（或称为钻石晶格，diamond lattice） 188
速度校正法（rescaling velocity method） 84
酸性基点（acid sites） 212
随机动力计算（stochastic dynamics，SD） 223
随机力（random force） 241
随机力微分方程式（stochastic differential equation） 222
随机乱踢（random kick） 200
随机数（random number） 167
随机搜寻（random search） 200
随机选择矩阵（stochastic matrix） 172
随机作用力效应（stochastic effect） 182
随意乱动模式（random walk） 190

## T

肽键（peptide bond） 267
泰勒级数（Taylor series） 28
梯度（gradient） 28
体系介电常数（bulk dielectric constant） 165
跳蛙方法（leap frog method） 69
通用的 Born 方程式（generalized Born equation，GB） 258
统计比重（statistical weight） 193
统计力学（statistical mechanics） 324
统计涨落（fluctuation） 108，177，205，328
图谱密度（spectra density） 118
退火淬火（annealing-quenching） 202，218

## W

弯曲度（curvature） 28
弯曲转折（turn） 220
微胞（micelles） 145，242
微观（micro-scale） 240，324
微观可逆（microscopic reversibility） 173
微扰（perturbation） 123
微扰计算法（perturbation method） 123
微正则系综（micro-canonical ensemble） 79，326
位力（virial） 177

位力系数（virial coefficient） 198
位向相关函数（orientation auto-correlation function） 232
稳定度（stability） 52
无定形（amorphous） 80

## X

系统（system） 326
系综（ensemble） 326
系综平均量（ensemble average） 332
系综平均值（ensemble average） 329
限制计算法（constraint dynamics） 96
线性回归法（linear regression） 60
线性寻找法（line search） 30
线状函数（line-shape function） 118
相（phase） 206
相对穿透系数（relative permittivity） 146
相关函数（correlation function） 114
相关时间（correlation time） 232
相角（phase angle） 318
相空间（phase space） 122
相空间轨迹（phase space trajectories） 174
相图（phase diagram） 208
相转换点（phase-transition point） 143
向后取样法（backward sampling） 125
向列型液晶（nematic liquid crystal） 203
向前取样法（forward sampling） 125
向心力（central force） 319
消除函数（destruction function） 178
校正预测法（predictor-corrector method） 71
协同运动（cooperative motion） 106
新生函数（generation function） 179
信封状结构（envelope） 50
虚频（imaginary frequency） 53
旋光异构体（enantiomers） 260
薛定谔方程式（Schrödinger equation） 2，331

## Y

压力张量（pressure tensor） 84

亚单位（subunit） 272
样品（sample） 324，326
摇摆振动（wegging） 53
液晶（liquid crystal） 20，145，203
一般性分子动力计算（conventional molecular dynamics） 79
一致性价力场（consistent valence force field） 15
一致性力场（consistent force field，CFF） 16
移除（remove） 215
移动（translation） 40
遗传算法（genetic algorithm） 296
阴电性（electronegativity） 18
应力-应变曲线（stress-strain curve） 110
应力张量（stress tensor） 109
荧光去极化（fluorescence depolarization） 228
“优良”溶剂（good solvent） 190
优先取样蒙地卡罗（preferential sampling MC） 180
油脂分子（lipid） 249
有限差分（finite difference） 224
有限差分热力积分法（finite difference thermodynamic integration，FDTI） 126
有效介电常数（effective dielectric constant） 11
有序（ordered） 203
原子重合（overlap） 260
运动轨迹（trajectory） 68，104

## Z

张量（tensor） 154
折叠（folding） 275
真空穿透系数 $\varepsilon_0$（permittivity in vacuum） 164
真实空间（real space） 158
振动模式（vibration mode） 37
正态分布（normal distribution） 325
正则系综（canonical ensemble） 326
脂质分子（lipid） 145
质量中心（center of mass，CM） 90
质权坐标位移（mass-weighted displacement） 38
秩序参数（order parameter） 145，205
中和（neutralising） 157

中心力力场（central force，CF） 133
中央多极展开法（central multiple expansion） 153
重要取样法（importance sampling） 170
周期性边界条件（periodic boundary condition） 72
周期性镜像（periodic mirror image） 72
主成分分析法（principal components regression，PCR） 294
主轴（principal axes） 155
主转动惯量（principal moment of inertia） 191
驻留长度（persistence length） 191
转动惯量（moment of inertia） 91
转动相关函数（rotational correlation function） 119
转换矩阵（transition matrix） 171
转置向量（transpose） 29
自身相关函数（auto correlation function） 114
自由度（degree of freedom） 331
自由能（free energy） 121
自组装（self-assembly） 279
自组装团簇体（self-assembled cluster） 42
总可及表面积（total accessible surface area） 260
阻碍移动（hindered translation） 118
阻尼力方法（damped force method） 84
最低极小值（global energy minimum） 26
最近镜像方法（nearest mirror image） 72
最速下降法（steepest descents） 29
最小平方误差拟合（least square fit） 295
最优化结构（optimized structure） 26
作用场（reaction field） 159

## 其他

AMBER（assisted model building with energy minimization）力场 14
Hessian 矩阵（Hessian matrix） 29
Lennard-Jones（LJ）势能 9
Rosenbluth 统计比重（Rosenbluth statistical weight） 214