中 国 科 学
1990 年 10 月　　B 辑　　第 10 期

蛋白质结构与功能分析的理论方法*

——I. 定向微扰构象分析

A. M. Hemmings** 常文瑞 梁栋材
(中国科学院生物物理研究所，北京 100080)

摘　要

一种被称为定向微扰构象分析(Derected Perturbation Conformational Analysis)的新的计算技术已被发展并用于蛋白质模型组建和结构功能研究. 所设计的算法可以执行对生物大分子势能表面的有效地局部搜寻，它可被用于从一个初始的或试用结构出发通过低能过渡态结构确定多重能量极小构象. 该算法包含了对由多个非谐势能表面的自由度所描述的系统的稳定过渡态优化的一些改进. 与 N-acetyl, N'-methyldeca-L-alaninamide 的标准分子动力学模拟的结果相比，该算法已被证明在从一个给定的试用结构出发产生交替的平衡结构中是有效的.

关键词：定向微扰，过渡态结构，势能函数

在许多情况下，一个令人关注的问题是预见不同已测结构的单个氨基酸取代的突变体蛋白的结构. 另一方面，特别是在同源蛋白结构的模型组建中，问题的关键是确定使之对结构其余部分相适合的模型的短肽段的构象. 这方面已知的最好例子是免疫球蛋白的模型组建[1]. 如果以下两点假设成立，上述两个基本类似的问题从理论上是容易解决的. 第一，必须能有一种方法可绘制母体分子的构象势能面 (PES)；第二，必须假定在 PES 中作为结果而发生在突变体上的变化是小的，而且它只限于所替代的位置上.

作为第一步逼近，我们所研究的算法是希望能实际应用于指导本实验室的分子设计实验特别是胰岛素结构与功能的研究. 本文所报道的是作为上述总体设计思想的第一步，称为定向微扰构象分析 (DPCA)，用于在一个初始的或试用的结构周围产生一组代表局部能量极小的构象. 从这一总体出发，我们可以利用一些适当的规则，如极小能量规则或质子间距离的 NOE 限制来预测母体的构象.

通过分子力学技术预测点突变对蛋白结构的影响已引起重视. 单从给定的一级结构数据预测蛋白质的整体三维结构也与类似的技术相关. 正如在同源蛋白质中所发现的那样，多个氨基酸取代的模型组建已被广泛地尝试，但最终被验证的很少[2]，而且其模型通常是应用能量

本文 1989 年 12月21日收到，1990 年 4 月26日收到修改稿.
* 国家自然科学基金和中国科学院生物学部重点资助项目.
** 现在地址：Dept. of Crystallography, Birkbeck College, London WC1E 7HX England.

极小化过程来实现的. 对于单个氨基酸取代的情况,问题固然是被简化了,但过程是基本相同的. 该方法的基点是假设由一个氨基酸取代得到的蛋白的稳定构象与母体蛋白的构象非常相近,结构的变化只是局限于被取代位置的周围. 然而,尽管只有局部的结构变化,由于蛋白质结构的复杂性加之标准的能量极小化过程的收敛半径较低,使得只确定局部能量极小值的可能性还是相当多的.

在蛋白质构象分析中,多重极小化问题已引起极大的关注,这也是 DPCA 算法要解决的特殊问题. 它是基于这样的事实: 一个分子的任何构象相互转换过程都要通过该分子的 PES 上的一个静态点(一个过渡态)[3],于是当给定了一个初始能量极小结构,要在这个能量极小结构的周围找出进一步的能量极小值的最有效的办法是确定将这些能量极小分开的过渡态结构. 一旦一个过渡态结构被找到,就可利用标准能量极小化技术来确定另外的能量极小值. 每一个新的能量极小点又可以作为一个新的起点来搜寻下一个过渡态和下一个极小.

对于蛋白质,由于如此巨大的自由度数目并由此而导致获得对 PES 的静态点搜寻的收敛的不可能性,使得该算法直接用于蛋白质是不现实的. 首先我们必须降解问题的维数. 这是通过 DPCA 算法的一个辅助程序 BIGMOL 来实现的,该程序在分析中只使用所研究分子的扭角的子集作为自由度.

一、定向微扰构象分析

由 Crippen 和 Scherage (1971) 提出的和缓坡度法 (The Method of gentlest-ascent) 的基本想法是在 PES 上从一个极小值出发向上走到峰的顶点再向下走到另一个极小值来确定静态点. 这也是本文所描述的技术的基础. 在早期应用中,沿着最低的简正模方向取一短的步长,然后在与搜寻方向正交的子空间中进行能量极小化引导这一过程从一个极小值沿一个低能通道移动. 然而这个过程并不能确定过渡结构,它只表明已经通过这样一个点.

Cerjan 和 Miller[4] 提出了一个改进的 Newton-Raphson 算法,即在过渡态极小化过程中,在通常的步长上强加一个 Lagrange 乘子. 为确保搜寻的稳定性,对该方法的各种改进也已发表 (Simons 等)[5],它们可作为估价这一方法的参照. 已发表的对该方法用于多变量体系的改进方法与现在的工作有许多类似之处,但所引用的例子也许不能用作对该方法的严格验证,如小分子的低频本征的解数目将很低,并与可归因的高频解的绝大多数实际上是分离的,如键的伸展运动. 于是,对于用直角坐标自由度所描述的小分子,在过渡态的极小化中的有效自由度的数目可能很少.

二、DPCA 算 法

给定一个初始的分子构象,DPCA 算法开始用一系列搜寻方式来寻找过渡态. 搜寻的初始方向选定为平行和反平行于系统的N个最低本征值的本征矢量的方向(如$N=3$时给出 6 个搜寻方向). 一旦一个过渡态被找到,紧接着的能量极小化将导出一个新的能量极小构象. 每一个新的构象又可作为下一个过渡态搜寻的起点,于是,分子势能表面的局域搜寻就可以按这一方式进行下去. 该方法的详细过程如下:

(1) 能量极小栈 (EM) 和过渡态栈 (TS) 被初始化.

(2) 初始构象经能量极小后加到 EM 栈中.

(3) 从 EM 中取出一个构象使之成为活动构象同时预置搜寻方向计数器 (SDC).

(4) SDC 被增值.

(5) 沿相应于现存于 SDC 的本征矢量的方向搜寻过渡态.

(6) 如过渡态已被找到,则与在 TS 中的过渡态相比较,如 TS 中还没有则将其加到 TS 中去,反之转执行(10).

(7) 结构将沿最低本征矢量被微扰并使其能量极小; 如果能量极小化过程失败则转执行(10).

(8) 能量极小的结构与存在 EM 中的结构相比较, 1) 如果新的构象与"活动"构象相同,则以二倍于先前的步长重复执行(7); 2) 如新的构象已列入但与"活动"构象不同, 则转执行(10); 3)如新的构象在 EM 中不存在,则将其坐标加到 EM 中去.

(9) 如在 EM 中的构象已达到所需要的最低数目,则退出执行.

(10) 如所需要搜寻的本征矢量的数目已完成,转执行(3)和(4).

对这一基本算法的使用是灵活的,例如它可以通过对 EM 栈的一个简单指令来抑制对那些不感兴趣的能量极小的搜寻,从而使它能更适宜对构象空间搜寻的定向.

三、坐标体系和势能函数

上述方法需要一个构象能量函数势能函数，它可单独作为扭角变化的函数来表达体系的能量. 由 Scherage 及其同事所改进的用于多肽的经验构象能量程序 ECEPP[6] 满足这一要求并已被并入 BIGMOL 程序之中. 程序用由使用者指定的一组独立扭角坐标来表达分子的基本自由度. 这些变量的选择由使用者任定,但为了不至使收敛造成困难,变量总数不要超过 200 个. 相应于所选定的独立扭角坐标的势能函数的一二级偏导数将被解析计算[7,8].

在包含两个分子并以扭角作为独立变量的生物多聚物体系的构象分析中，一个新的问题是要增加 6 个描述两个分子之间关系的外部自由度,它涉及到势能函数的一二级导数的计算. 克服这一问题的建议也已提出,但这一建议的缺点是奇点,不连续地或部分地打断独立变量的树-结构次序,而这种树-结构次序正是 Wako[8] 为计算势能函数的一二级偏导数的精采算法的最基本的部分,为了检验这个问题，我们选择使 6 个独立的外部坐标进入虚键，它可将第一个分子的最后一个原子与第二个分子的第一个原子分开，也可以很容易地插到独立扭角变量的树-结构序中去. 更重要的是,它只需对 Wako 的算法作很小的修改. 关于独立外部变量体系的细节请参照本文的附录.

四、过渡态的优化

在 BIGMOL 中用于过渡态优化的算法是基于 Cerjan 和 Miller[4] 的工作，所设计的改进是为了确保由一个重复出现的严格的非谐势能函数所描述的多自由度体系搜寻的稳定性，这些保证措施是:

(1) 预置一个最大的步长 B_{max} 和一个最大的能量变化量 E_{max},凡超过这些限制而产生的所有步长均被取消.

(2) PES 的局部二次项按如下方法检验: 将能量变化与在理想的谐振表面上按同样步长计算的变化量相比较，如果其比例大于预置的限制 Q_{max}, 则步长乘以 0.7，直到满意为

止.

(3) 对于该算法在连续迭代中有效的本征矢量之间的点乘的模数,预置一个极小值 D_{min},凡点乘小于 D_{min},则步长乘以 0.5 直到满意为止. D_{min} 取 0.8 已被证明是合适的.

对这一基本算法已作过各种试验以判断它在最有效的搜寻中的功用,以下作法已被证明是有效的:

(1) 用一个非键相互作用截断距离 (cutoff) r_{max} 和立方平滑半径 r_{smooth}[9] 来计算势能和导数.

(2) 只对每 NHess 次迭代计算解析的 hessian 矩阵,而对其它的迭代,矩阵是用解析的一级导数矢量和 Powell[10] 算法来修正的.

(3) 对于蛋白质 PES 的合理的谐振区,通常的过渡态优化算法产生的步长在高频本征解上有很小的投影. 由于这个原因,仅对每 NEigen 个迭代计算全部本征解: 对于其余的各步,最低的 Nacvive + 6 个本征解是用二分法获得的[11].

五、能量极小化

程序中包括 3 个优化程序可用于能量极小化: 即共轭梯度 (CG) 优化[12], 带保护的 quasi-Newton-Raphson (QNR) 优化[13]和 Banerjee[14] 等人的旋转函数优化的辅助程序 (BASS). 由于 CG 优化的效率和稳定性,通常它被用于极小化的第一步,接着可用 QNR 和 BASS 精确确定能量极小点. BASS 程序的优点是在任一次迭代中仅需计算最低频的本征值.

六、方 法

为估价 DPCA 方法的效率,我们对一系列简单的同源肽段的构象进行了分析,这些肽段均具有一般蛋白质中常见的短伸展的二级结构. 选择的分子有: 6-L-丙氨酸的 N-乙酰N′-甲酰胺衍生物,6-L-缬氨酸和 10-L-丙氨酸,另外,对只含有单个 L-丙氨酸的衍生物的构象空间进行了更彻底的搜寻.

这些分子的初始模型是用从 ECEPP 力场中得到的标准自身的坐标值来建造的. 所有这些模型均是在构象空间的 α_R 区 (ϕ,ψ = (−60°, −40°)) 的扭角建造的,只有 6-L-丙氨酸的衍生物的一个附加模型是在一般的伸展构象区(−83°, 153°)内建造的. 后者是选用单 L-丙氨酸衍生物的 PES 的分析结果. 6-L-缬氨酸衍生物的 χ_1 扭角全取 180°, 所有的肽基团取 Trans 平面构象.

分子自由度的分配取决于模型,但对每一种情况均包含扭角,对第一个模型,即 6-L-丙氨酸的 N-乙酰, N′-甲酰氨的衍生物共分配 5 个扭角变量(包括 ϕ 和 ψ),在其它模型中,只有每个残基的主链 ϕ 和 ψ 不被限制,于是,对于 6-L-丙氨酸衍生物是 12 个变量,10-L-丙氨酸衍生物是 20 个变量;而 N-acetyl, N′-methyl-hexa-L-Valinamide 模型,除了 ϕ, ψ 自由度外, χ_1 侧链扭角不被限制,共 18 个变量. 每一个单肽模型中的所有肽扭角W均被限制.

这些体系的构象分析是用 BIGMOL 程序实现的.在分子势能函数的计算中使用了力场. 使用 Wako 和 Go[8] 的高效算法计算势能导数,过渡态的优化是在平行和反平行于本征矢量的初始搜寻方向上进行的,所用的特征矢量则取相应于在能量极小初始点上计算的 hessian 矩阵的 3 个最低频本征解. 最大扭角步长为 2.5°, 二次项标准 Q_{max}=1.5 以防止在搜寻中异常

大的分子扰动的出现. 能量极小化依次使用了共轭梯度，quasi-Newton 和 BASS 优化方法以精确确定其能量极小点，一般情况下，最多只需确定 30 个能量极小值.

10-L-丙氨酸衍生物的分子动力学模拟退火（MDSA）是用 AMBER 分子力学程序 (Version 3.0)[15] 实现的，使用以直角坐标表示自由度的全原子力场. 由于分子较小，没有使用非键相互作用的裁断半径. 在计算静电相互作用时介电常数 ε 取 1. 用共轭梯度法完成真空中的能量极小化之后，将分子以 50K 步由 0K 加热到 300K，使用 $\delta = 2$fs 的积分时间步长，总时间为 6ps. 使用同样的时间步长进行 14ps 的模拟，所用的坐标是在能量极小化中每 1ps 中取出的. 在结束时，所有能量极小的数据组相对于剩余的一级导数矢量变化的均方差 (Rms) 小于 0.01kcal · mol^{-1} · Å^{-1}.

全部计算工作是在北京大学信息科学中心的 VAX-11/785 计算机上完成的.

七、结　果

在第一个实例中，我们的目的是监视该方法在具有完好特征的小分子的 PES 分析中相

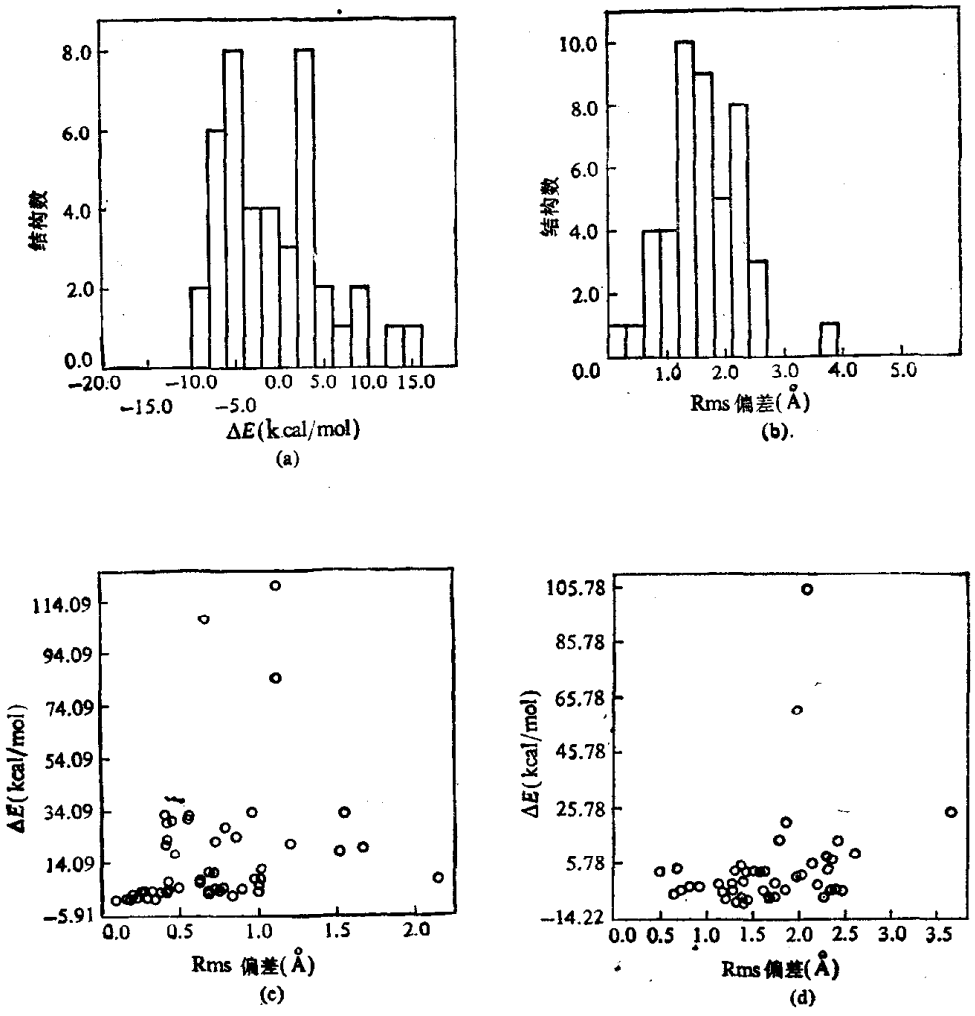


图 1　α_R-型的 N-actyl, N′-methyl-hexa-L-alaninamide 的 DPCA 分析结果

((a) EM 栈取出的构象相对于能量极小初始结构的能量差异分布；(b) EM 栈中取出的构象相对于能量极小初始结构的 Rms 原子偏差的分布：导数计算是基于全部非氢原子；(c)从 TS 栈中取出的转变态结构相对于它们的母体能量极小结构以 Rms 原子偏差为函数的能量差异图；(d) 从 EM 栈中取出的能量极小构象相对于能量极小的初始结构(即 EM 栈中的第一个结构)，以 Rms 原子偏差为函数的能量差异分布图)

对于其它方法的效能特征.从 ϕ,ψ 空间的 α_R 螺旋区的一个分子构象开始,N-乙酰, N'-甲酰胺-L-丙氨酸的 DPCA 算法很快产生了 9 个能量极小构象和将其分开的过渡态。 如果用过渡态搜寻导出能量极小的百分数来表示,其搜寻效率为 80%. 这些构象的细节列于表 1. 能量极小点的几何学与前人用理论计算的结果有非常合理的吻合[16,17]. 在 $C_{7,eq}$ 区中找到的总能量极小点 $(\phi,\psi)=(-83°,135°)$, 位于 α_R 型上方 2.2kcal · mol^{-1} 处.

具有 α 型和伸展型两种显然不同构象的 6-L-丙氨酸衍生物的分子的 PES 图分析结果分别示于图 1 和图 2. 正如所预期的,在伸展型中,更为开放的氢键型式导致更大的柔性和具有相对多的低位能过渡态结构. 这些过渡态结构可以很容易取出来确定其交替的能量极小并能很好地与初始构象相分开. 这一点也可在确定能量极小值之前从能垒分布来推断图 1 和图 2 与起始能量极小结构的最大 Rms 偏离, 对 α_R 型为 3.7 Å, 对 β 型为 2.3 Å. 这些偏差均远大于从标准能量极小化或分子动力学过程的结果. 在 α 型中以 Valyl 代替 Alanyl 的侧链生的产 6-L-缬氨酸衍生物导致 α 螺旋的明显变硬并使得用 DPCA 算法的构象空间搜寻的总数减少. 该体系的过渡态结构的能量和能量极小值的相对分布示于图 3. 该算法可以确定低

表 1 N-acetyl, N'-methyl-L-alaninamide 的势能表面的能量极小(EM)点

EM 标号	Φ	Ψ	$\triangle E$
1	−68	−46	0.0
2	−172	−60	+4.0
3	−96	91	−1.4
4	77	−52	+22.1
5	−83	153	−2.2
6	44	−71	+22.8
7	44	60	+1.1
8	−150	84	−1.0
9	60	174	+5.3

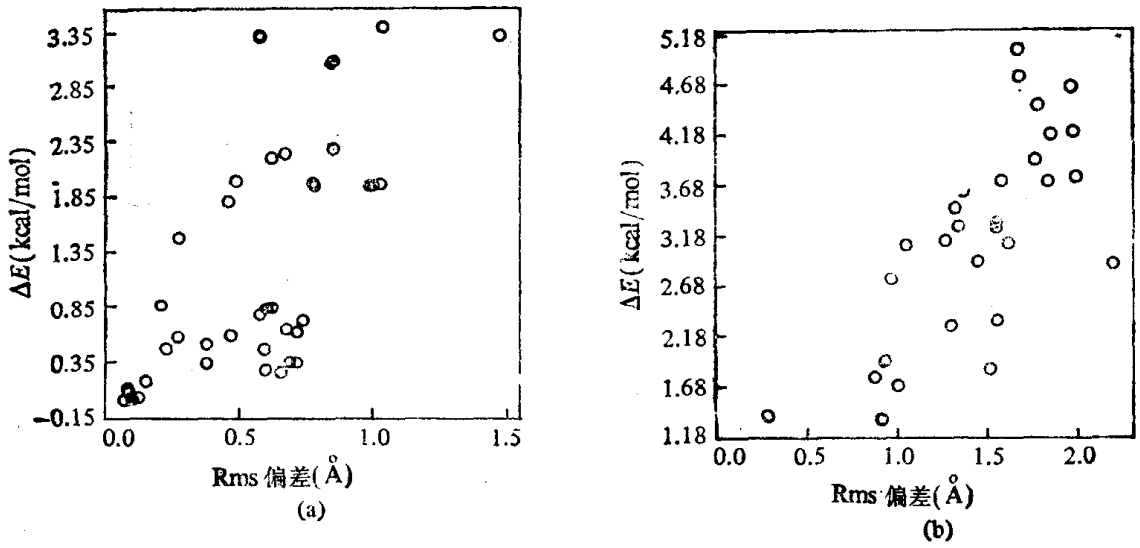


图 2 伸展型的 N-acetyl, N'-methyl-hexa-L-alaninamide 的 DPCA 分析结果
((a) 同图 1 的 (c),(b) 同图 1 的 (d))

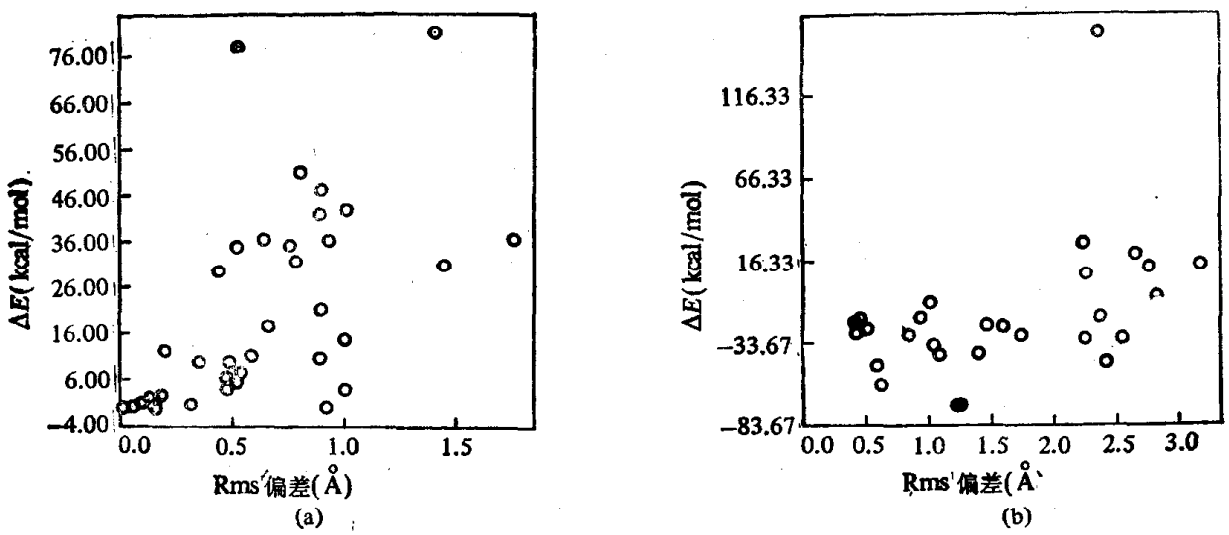


图 3　α_R-型的 N-acetyl, N′-methyl-hexa-L-Valinamide 的 DPCA 分析结果
((a)同图 1 的 (c), (b) 同图 1 的 (d))

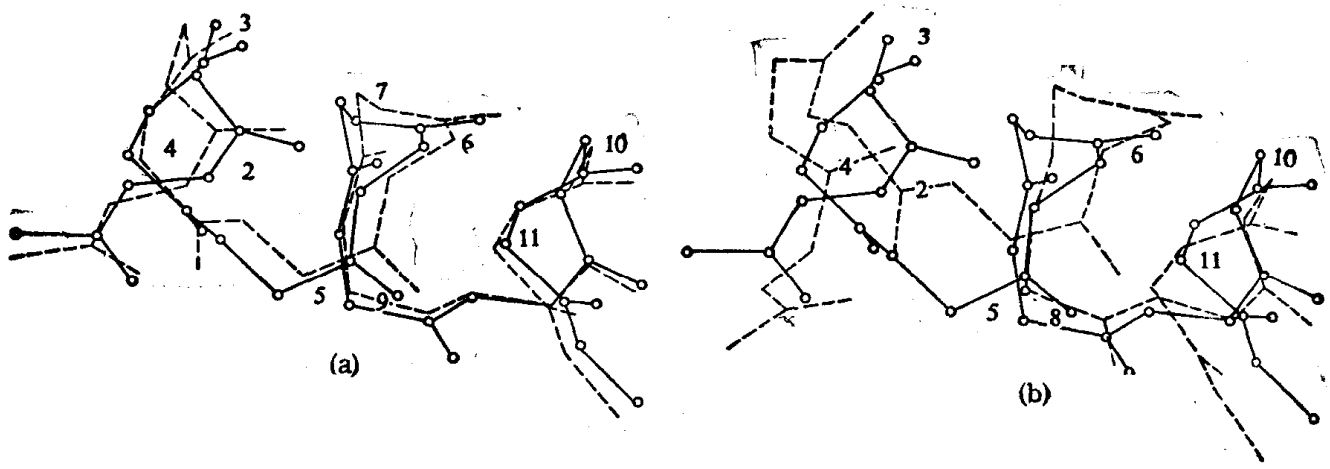


图 4　DPCA 算法产生的构象的比较
((a) deca-L-alanine 衍生物的初始能量极小构象(点线)和从 TS 栈中取出的第 4 个构象(实线)的比较, (b) deca-L-alanine 衍生物的初始能量极小构象(点线)和从 TS 栈中取出的第 5 个构象(实线)的比较)

于能量极小的初始结构 74kcal · mol^{-1} 的交替能量极小构象.

对于 10-L-丙氨酸衍生物,在经过 67 个过渡态极小搜寻得到 32 个过渡态之后,由 DPCA 算法最终找到 30 个明显的能量极小值. 图 4 显示了由 DPCA 算法连续产生的分子 PES 图上的点所代表的分子构象的最小二乘比较的序列结果. 图中示出了从 EM 栈中取出的指标为 1 和 5 的能量极小结构被 TS 栈中的指标为 4 的过渡态结构所分开. 请注意, 过渡态结构与母体结构只有很小的差异 (Rms = 0.4 Å, ΔE = + 23kcal · mol^{-1}), 由它所定位的能量极小结构相对于其初始结构显示出可观察的扰动. 实际上, 这个交替构象的能量比起初始构象低 8kcal · mol^{-1}, 其 Rms 偏差为 1.4 Å. 我们可以看到,螺旋的N端已开始解开,而且随着该模型的过渡态和能量极小的进一步确定,上述过程也将继续.过渡态结构的能量分布与它们的初始能量极小点的关系示于图 5(a), 它揭示了有意义的低能过渡态的数目, 这些低能过渡态

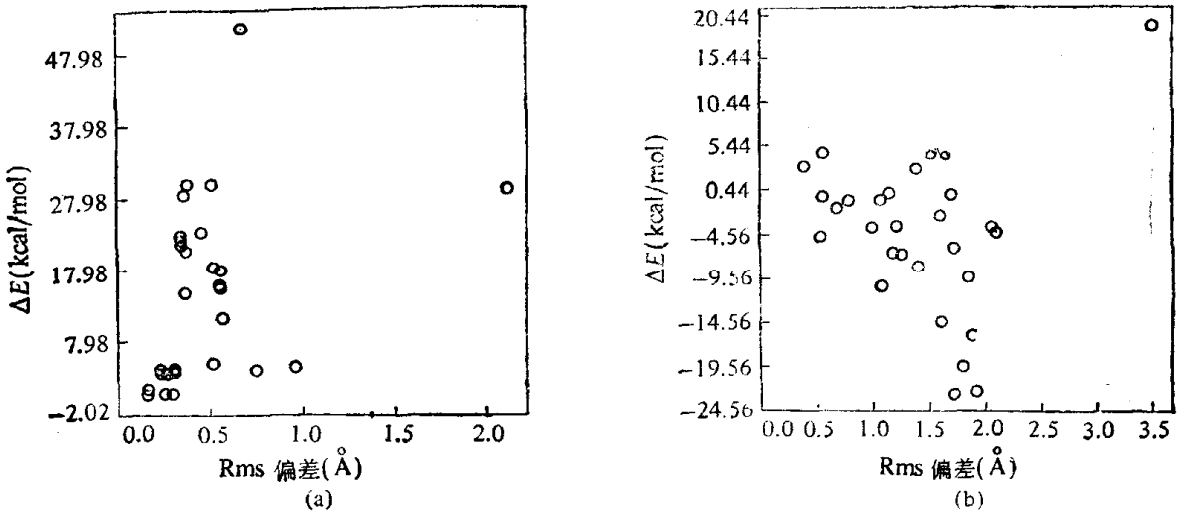


图 5 α_R型的 N-acetyl, N′-methyl-deca-L-alaninamide 的 DPCA 分析结果
((a) 同图 1 的 (c), (b) 同图 1 的 (d))

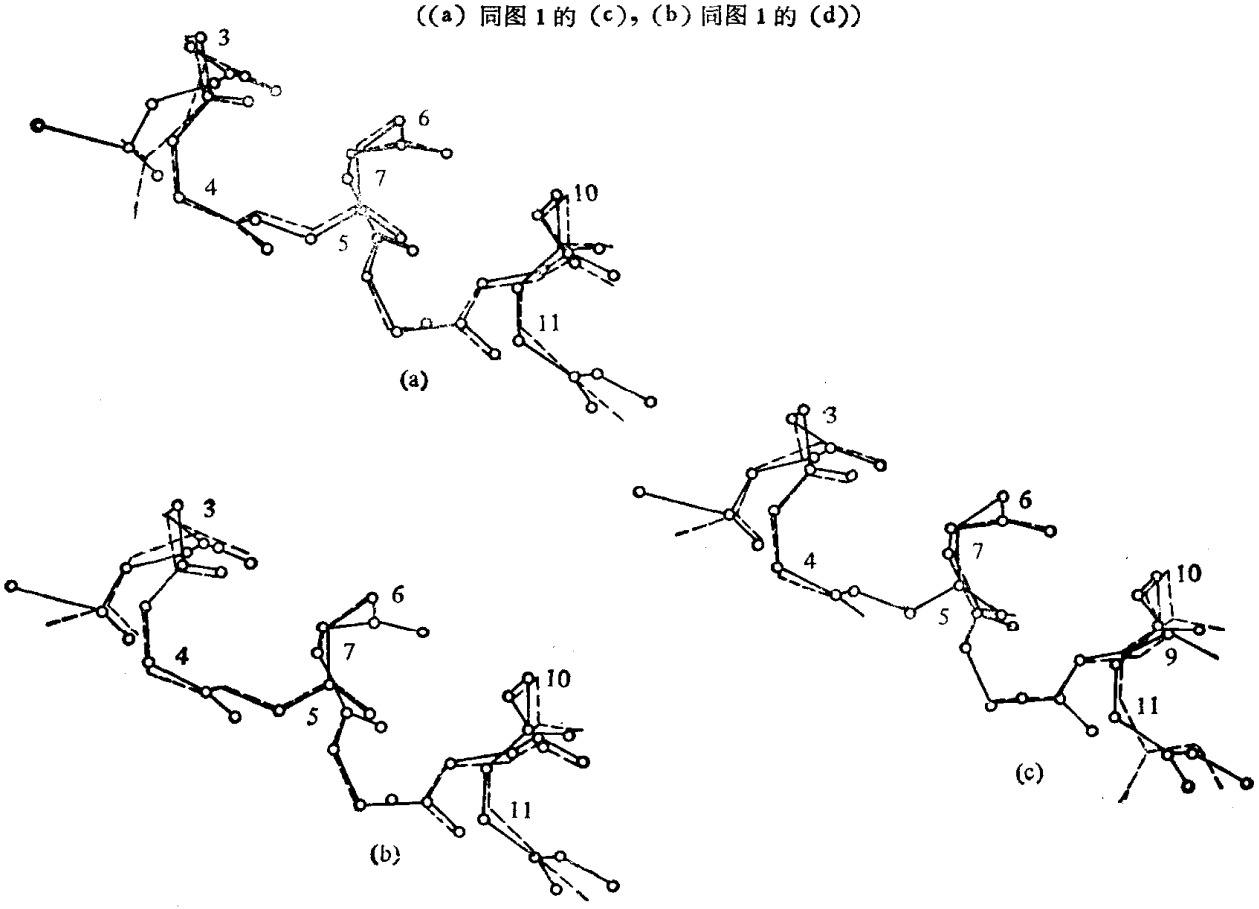


图 6 α_R 型的 N-acetyl, N′-methyl-deca-L-alaninamide 的分子力学模拟退火的结果
((a) t = 0ps (实键)和 t = 6ps (断键)的能量极小构象的比较; (b) t = 0ps (实键)和 t = 9ps (断键)的能量极小构象的比较; (c) t = 0ps (实键)和 t = 19ps (断键)的能量极小构象的比较)

提供了构象相互转变甚至包括视为刚性的螺旋的低能扰动路线. 能量极小构象与从初始螺旋模型导出的结构的能量分布的关系示于图 5(b). 其分布明显地展示出已被划分出来的广泛的分布在构象空间中的平衡构象的区域. 整体能量极小结构位于低于初始能量极小结构 22.5

kcal · mol^{-1}，Rms 为 1.72 Å 处。这个结构相对于初始的 α 螺旋型显示出有较大的扰动。以分子动力学模拟退火实验导出的 N-acetyl，N′-methyl-deca-L-alaninamide 的 15 个能量极小构象与 $t = 0$ps 的结构比较分析，揭示了在经典动力学模拟中可用于该体系的构象空间区域。要提醒的是，该过程对于分子只产生 4 个非常接近的构象，这些典型的构象示于图 6，其差异只出现在 N-acetyl 和 N′-methylamide 基团的取向，而分子的螺旋区并无大的扰动。相对于 $t = 0$ps 的结构，经过 $t = 6.9$ 和 $t = 19$ps 的模拟之后的能量极小化导出的结构，其全部原子的 Rms 偏差分别为 0.6，0.7 和 0.5 Å。当使用 AMBER 的全部原子力场计算时，这 3 个结构分别处在低于初始结构能量 0.7，2.0 和 0.3kcal · mol^{-1} 的状态。

八、讨　论

我们已经介绍了为执行对生物大分子势能表面进行有效搜寻而设计的方法，为此，我们根据需要扩展了 Braun 等人关于生物分子体系的构象能量分析中分子间自由度的说明和规定，我们这里所提出的特殊形式是简单的，而且在使用扭角快速计算势能和导数时不破坏树结构的基础，更重要的是它避免了其它方法产生的不连续性。另外，它提供了一种可能性，即不只仅使用扭角作为大分子基本自由度描述的独立变量，而且也提高了势能和导数的计算效率。当然该方法有待进一步发展，即采用更适宜的势能函数，其加合项应扩展到包括键长、键角和扭角等独立变量。通常使用的以直角坐标表达的分子势能函数，即 CHARMM[9] 和 AMBER[15]，包含有潜在的内部坐标定义项，因而不适于这里的情况。

我们的目的已表明：DPCA 算法可以从一个合理的起始结构（可用模型组建技术产生）出发，在构象空间中，在初始结构的周围产生一组平衡结构，因而它可被用于结构预测。在 10-L-丙氨酸衍生物的分析中，DPCA 算法和模拟退火实验所用的 CPU 时间是很接近的（DPCA4h，DMSA 对于 20ps 的模拟需 6h，能量极小化 4h），但两种方法搜寻的构象空间的量是显然不同的。从 DPCA 算法获得的能量极小的数目以及它们与初始结构的 Rms 偏差远大于 DMSA 的结果。这些数据表明，对于取样构象空间，DPCA 算法比单纯的分子动力学模拟退火更为有效。由 DPCA 算法得到的平衡结构的力能学和构象精确细节将可能受到因约束键长、键角而带来的影响，这是显然的，但其影响的幅度实际上是很小的。

DPCA 算法可能还有潜在的功能，如在定位定向突变之后，相对于母体，酶的特性变化结合常数和速率常数的自由能微扰计算。这可归于计算时间不足产生的收敛问题。我们面临的情况是，DPCA 算法将被用来有效的确定众多的交替构象，而每一个这样的构象都将在整个的微扰模拟中作为一个起点。然而由于相空间取样不须十分严格，每一个微扰模拟比单起点模拟计算所需时间要短，这可作为上述问题的补偿。

显然，这里描述的方法没有考虑溶剂效应、时间平均和其它因素对所获得的分子构象的影响。程序结构设计中考虑到了将离散的水分子加入模型的可能性（某些必要的特定情况）。但一般情况下应避免这样做，因为会导致自由度的急据增加。通过使用更精确的模型，如通过平均力场来包含这些因素的作用，可能是有待研究的更为合理的途径。

在定向微扰构象分析中，蛋白体系比小的寡肽需要更详细地计算某些变化，首先我们可在感兴趣的区域只分配指定的扭角自由度，例如对蛋白的单定位突变体，仅在取代点的一定半径的区域内分配主侧链的扭角来表征其柔性，显然，对于给定的计算条件，自由度越少，所能搜寻

的构象空间的体积越大。下一篇文章我们将给出该方法用于胰岛素母体，突变体和衍生物研究中的一些结果。

A. M. Hemmings 感谢英国皇家学会接受其本人为访问学者的资格，作者对北京胰岛素组全体成员的有益讨论和北京大学信息科学中心在计算机使用方面的慷慨协助，在此表示衷心感谢。

附 录

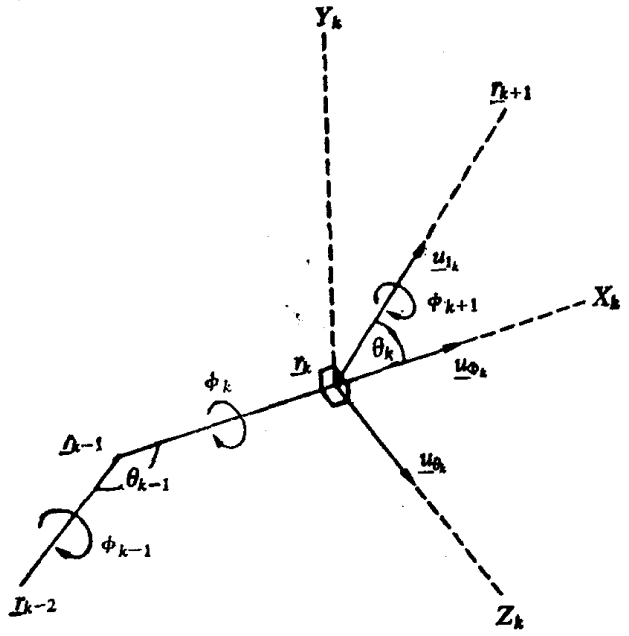
包含两个或多个分子的生物聚合体系的构象能量分析中分子间自由度的说明.

两个或多个分子的相对位置和取向可以有不同的方法来指定. 在两个分子的体系中，我们选择在体系树-结构表达中用 6 个包含分子间虚键的独立的内部坐标来指定分子间的自由度(附录图 1)。我们可按通常的做法将这个 6 个外部参数以如下方式插入二面角的树-结构的次序中去:

$$q = (\phi_1, \phi_2, \cdots, \phi_i, \theta_i, l_i, \phi_{i+1}, \theta_{i+1}, \phi_{i+2}, \cdots, \theta_n), \tag{1}$$

这里 $\phi_j, j = 1, n$ 是键 j 的二面角. 在从可变的内部坐标组 q 产生或修正直角坐标的方法中，当一个变量 q_i 改变时只有两个分子的一部分运动，这一部分是不包含分子的第一个单位的一侧，它被定义为变量 q_i 的运动侧，并假设为 M_i 这组 M_i 的余数，也就是变量 q_i 的固定侧，假设为 $\bar{M}_i$，于是我们可得到

$$M_{\phi_i} = M_{\theta_i} = M_{l_i}$$



附录图 1 多于一个分子的体系的构象分析中使用的外部坐标

(第二个分子的第一个原子 k 的局部坐标系的图解表示. 从第一分子的最后一个原子 $k-1$ 到原子 k 的矢量表示两个分子之间的虚键. 6 个分子间自由度为: $\theta_{k-1}, \phi_{k-1}, l_k, \theta_k, \phi_k$ 和 ϕ_{k+1}. 单位矢量 $u(\phi_k)$ 和 $u(l_k)$ 分别位于矢量 $(r_k - r_{k-1})$ 和 $(r_{k+1} - r_k)$ 的方向上，单位矢量 $u(\phi_k)$ 由下式给出: $U(Q_k) \cdot \sin Q'_i = U(\phi_k) \times U(L_k)$)

相应于(1)式中的独立变量，坐标矢量对给定的扭角 ϕ_i 的导数如下:

$$\frac{\partial r_\alpha}{\partial \phi_i} \begin{cases} (e_\alpha \times r_\alpha - r_i) & if \alpha \in M_i, \\ 0 & if_\alpha \in \bar{M}_i, \end{cases} \tag{2}$$

这里 e_α 是沿着键 i 在其运动侧方向上的单位矢量. 矢量 r_i 是键 i 的任一端的一个原子的位置矢量. 同样，相对于键长 l_i 的导数为:

$$\frac{\partial r_\alpha}{\partial l_i} = \begin{cases} e_\alpha & if \alpha \in M_i, \\ 0 & if \alpha \in \bar{M}_i \end{cases} \tag{3}$$

这里，e_α 是延着键 $i+1$ 在其运动侧方向的单位矢量. 键角的表达式为:

$$\frac{\partial r_\alpha}{\partial \theta_i} = \begin{cases} (e_b \times r_\alpha - r_i) & if \alpha \in M_i, \\ 0 & if_\alpha \in \bar{M}, \end{cases} \tag{4}$$

这里，e_b 是由 $e_b \cdot \sin\theta_i = e_c \times e_a$ 给出的单位矢量.

上述 3 个表达式(2)，(3)和(4)可用在表 2 中定义的命名法通过下面的一般公式简单地加合：

表 2　ϕ_i 和 ψ_i 的一般定义

q	ϕ_i	ψ_i
ϕ_a	e_a	$e_a \times r_a$
θ_a	e_b	$e_b \times r_a$
l_a	0	$-e_c$

$$\frac{\partial r_\alpha}{\partial q_i} = \begin{cases} (\Phi_i \times r_\alpha - \Psi_i) if \alpha \in M_i \\ 0 \quad if \alpha \in M_i \end{cases} \tag{5}$$

势能函数的一二级导数可用下述表达式按文献[8]中建立的方式进行：

$$\frac{\partial E}{\partial q_i} = (\Phi_i, \Psi_i) \cdot \sum_{M_i, \beta \in M_i} C_{\alpha\beta} \begin{pmatrix} r_\alpha \times r_\beta \\ r_\alpha - r_\beta \end{pmatrix} \tag{6}$$

和

$$\frac{\partial^2 E}{\partial q_i \partial q_i} = (\Phi_i, \Psi_i) \cdot R_{ij} \cdot \begin{pmatrix} \Phi_j \\ \Psi_j \end{pmatrix}. \tag{7}$$

参　考　文　献

[1] Chothia, C. et al., *Science*, 233(1986), 755—758.
[2] Novotny, J., Bruccoleri, R. & Karplus, M., *J. Mol. Biol.*, 177(1984), 787—818.
[3] Salem, L., *Electrons in Chemical Reactions: First Principles*, Wiley-Interscience, New York, 1982.
[4] Cerjan, C. J. & Miller, W. H., *J. Chem. Phys.*, 75(1981), 2800—2806.
[5] Simons, J. et al., *ibid.*, 87(1983), 2745—2753.
[6] Momany, F. A. et al., *J. Phys. Chem.*, 79(1975), 2361—2380.
[7] Noguti, T. & Go, N., *J. Phys. Soc. Japan.*, 52(1983), 3685—3690.
[8] Wako, H. & Go, N., *J. Comp. Chem.*, 8(1987), 625—635.
[9] Brooks, B. R. et al., *ibid.*, 4(1983), 187—217.
[10] Fletcher, R., *Practical Methods of Optimisation*, Vol 1, Unconstrained Optimisation, Wiley, New York, 1980.
[11] Wilkinson, J. H. & Reinsch, C., *Handbook for Automatic Computation*, Vol II, Linear Algebra, Springer-Verlag, London, 1971.
[12] Powell, M. D. J., *Math. Prog.*, 12(1977), 241—254.
[13] Fletcher, R., *Computer J.*, 13(1970), 317—322.
[14] Banerjee, A. et al., *J. Phys. Chem.*, 89(1985), 52—57.
[15] Singh, U. C. et al., *AMBER(UCSF) Version* 3.0, Dept. Pharmac. Chem., Univ. California, San Francisco, 1986.
[16] Rossky, P. J., Karplus, M. & Rahman, A., *Biopolymers*, 18(1979), 825—854.
[17] Scarsdale, J. N. et al., *J. Amer. Chem. Soc.*, 105(1983), 3438—3445.