

分子动力学模拟

一. 分子动力学的基本原理

在分子动力学模拟中,体系原子的一系列位移是通过牛顿运动方程积分得到的,结果是一条运动轨迹,它表明了系统内原子的位置与速度如何随时间而发生变化。通过解牛顿第二定律的微分方程,可以获得原子的运动轨迹。方程如下:

$$\frac{d^2 \vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m}$$

这个方程描述了质量为 m_i 的原子 i 在力 F_i 的作用下,位置矢量为 r_i 时的运动方程。其中, F_i 可以由势函数 U 的梯度给出:

$$\vec{F}_i = -\nabla_i U$$

系统的温度则与系统中全部原子的总动能 K 通过下式相联系:

$$K = \frac{1}{2} \sum_{i=1}^N m_i |\vec{v}_i|^2 = \frac{1}{2} (3N - N_c) k_B T$$

N 是原子数, N_c 是限制条件, k_B 是波尔兹曼常数。

二. MD 模拟的积分算法

为了得到原子的运动轨迹,可以采用有限差分法来求解运动方程。有限差分法的基本思想就是将积分分成很多小步,每一小步的时间固定为 δt 。用有限差分运动方程有许多方法,所有的算法都假定位置与动态性质(速度、加速度等)可以用 Taylor 级数展开来近似:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) + \frac{1}{6} \delta t^3 \vec{b}(t) + \frac{1}{24} \delta t^4 \vec{c}(t) + \dots$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \delta t \vec{a}(t) + \frac{1}{2} \delta t^2 \vec{b}(t) + \frac{1}{6} \delta t^3 \vec{c}(t) + \dots$$

$$\vec{a}(t + \delta t) = \vec{a}(t) + \delta t \vec{b}(t) + \frac{1}{2} \delta t^2 \vec{c}(t) + \dots$$

在分子动力学模拟中,常用的有以下的几种算法:

1. Verlet 算法

运用 t 时刻的位置和速度及 $t - \delta t$ 时刻的位置,计算出 $t + \delta t$ 时刻的位置:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) + \dots$$

$$\vec{r}(t - \delta t) = \vec{r}(t) - \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) - \dots$$

两式相加并忽略高阶项,可以得到:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \delta t^2 \vec{a}(t)$$

速度可以通过以下方法得到:

用 $t + \delta t$ 时刻与 $t - \delta t$ 时刻的位置差除以 $2\delta t$:

$$\vec{v}(t) = \left[\vec{r}(t + \delta t) - \vec{r}(t - \delta t) \right] / 2\delta t$$

同理，半时间步 $t + \delta t$ 时刻的速度也可以算：

$$\vec{v}(t + \frac{1}{2}\delta t) = \left[\vec{r}(t + \delta t) - \vec{r}(t) \right] / \delta t$$

Verlet 算法执行简单明了，存储要求适度，但缺点是位置 $r(t + \delta t)$ 要通过小项与非常大的两项 $2r(t)$ 与 $r(t - \delta t)$ 的差相加得到，容易造成精度损失。另外，其方程中没有显示速度项，在没有得到下一步的位置前速度项难以得到。它不是一个自启动算法：新位置必须由 t 时刻与前一时刻 $t - \delta t$ 的位置得到。在 $t=0$ 时刻，只有一组位置，所以必须通过其它方法得到 $t - \delta t$ 的位置。一般用 Taylor 级数：

$$\vec{r}(-\delta t) = \vec{r}(0) - \delta t \vec{v}(0)$$

2. Velocity-Verlet 算法

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t)$$

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

3. Leap-frog 算法

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

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

为了执行 Leap-frog 算法，必须首先由 $t - 0.5 \delta t$ 时刻的速度与 t 时刻的加速度计算出速度 $v(t + \delta t)$ ，然后由方程

$$\vec{v}(t + \delta t) = \vec{v}(t) + \frac{1}{3} \delta t \vec{a}(t + \delta t) + \frac{5}{6} \delta t \vec{a}(t) - \frac{1}{6} \delta t \vec{a}(t - \delta t)$$

计算出位置 $r(t + \delta t)$ 。T 时刻的速度可以由：

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

得到。速度蛙跳过此 t 时刻的位置而得到 $t + 0.5 \delta t$ 时刻的速度值，而位置跳过速度值给出了 $t + \delta t$ 时刻的位置值，为计算 $t + 0.5 \delta t$ 时刻的速的作准备，依此类推。其缺点是位置与速度不同步。这意味着在位置一定时，不可能同时计算动能对总能量的贡献。

三. 分子动力学计算的时间间隔

时间间隔 δt 在积分算法中是一个非常重要的参数。为了充分利用 CPU 时间，尽量选择比较大的时间间隔，但是如果时间间隔太大，就会造成积分过程的不稳定性和不精确性。时间间隔的设置同时依赖于算法和模型的情况。模型本身给时间间隔带来的最大的限制就是最高频率的运动。由于 Verlet 算法要求在每个时间间隔内模型的速度和加速度保持一边，时间间隔就应该低于振动周期的八分之一到十分之一。对大多数的有机模型来讲，最高的振动频率是 C-H 键的伸缩振动，其振动周期的数量级为 10^{-14} s。这样，时间间隔就应该是 0.5-1 fs 左右。如果采用受约束的 SHAKE 或者 RATTLE 算法，可以使用更长的时间间隔。如果研究对象是液态或者固态简单模型，对体系内作用模式不感兴趣，也可以采用一些更长的时间间隔，比如 20 fs。对离子态的材料模型，5 fs 左右是合适的。时间间隔必须跟选择的算法相匹配。比如，ABM4 算法的时间间隔应该是 Verlet 算法的一半左 Runge-Kutta-4

算法则需要比其他算法更短的时间间隔。

四. 生物大分子的相互作用势函数

生物大分子势一个自由度很大，结构复杂的体系，各种物理过程极为丰富。其中的相互作用规律还难以进行统一的处理。但其物理图像是清晰的，电子的运动和骨架有效的分开，成键电子是定域的，价键的概念依然有效。荷电集团之间存在静电相互作用。原子或集团正负电荷中心不重合形成电多极矩并参与静电相互作用。原子或集团的电子云受静电相互的诱导形成电诱导偶极由此产生色散力。静电、范德华相互作用称非键相互作用。此外有共价键德振动及单键德旋转绕共价双键德扭曲等称为成键相互作用。

典型德生物分子动力学的模型势函数可以写成一下形式：

$$V(\bar{r}_1, \bar{r}_2, \dots, \bar{r}_N) = \sum_{\text{bond}} \frac{1}{2} k_b (b - b_0)^2 + \sum_{\text{bond angle}} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{\text{improper dihedral angle}} \frac{1}{2} k_\xi (\xi - \xi_0)^2$$

$$+ \sum_{\text{dihedral angles}} k_\phi [1 + \cos(n\phi - \delta)] + \sum_{\text{nonbonded pairs}(i,j)} \left[\frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} \right]$$

第一项求和是共价键的键能，第二项求和是键角能，第三项求和是非正常二面角能量，第四项求和是正常二面角能量。Kb, k θ , k ϕ , k ξ 代表式中前四项成键相互作用的力常数。式中最后一项式体系中的非键相互作用项，(i,j)表示对非键对求和，C12、C6 式范德华相互作用势中的系数。势函数中参数 qi, qj 分别为第 i 荷第 j 个原子的电荷。但是，对于疏水相互作用，溶剂的极化荷对蛋白质内部静电屏蔽，都需要对势函数加以修正和补充。若要了解溶液中分子的动力学性质则须将模拟的体系放到充分多的水分子做充分动力学模拟。

五. 分子动力学模拟的启动

进行 MD 模拟，必须首先建立系统的初始结构。初始结构可以通过实验数据、理论模型、或两者的结合来获得。除此之外，还可以给每个原子赋予初速度，它可以从一定温度下的 Maxwell-Boltzmann 分布来任意选取：

$$p(v_x) = \left(\frac{m_i}{2\pi k_B T} \right)^{1/2} \exp\left[-\frac{1}{2} \frac{m_i v_x^2}{k_B T} \right]$$

Maxwell-Boltzmann分布给出了质量为mi的原子i在温度T下沿x方向速度为v_{ix}的概率。Maxwell-Boltzmann分布势一中Guassian分布，它可以用随机数发生器得到。大多数随机发生器产生的随机数均匀分布在 0—1 之间，但是可以通过变换得到Guassian分布。均值为<x>和波动值为 δ^2 的Guassian分布的概率为：

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \langle x \rangle)^2}{2\sigma^2} \right]$$

一种方法势首先产生两个在 0—1 之间的随机数 ξ_1 和 ξ_2 。运用下列式子可产生两个数下 x1、x2：

$$x_1 = \sqrt{(-2 \ln \xi_1)} \cos(2\pi\xi_2) \quad x_2 = \sqrt{(-2 \ln \xi_2)} \sin(\pi\xi_1)$$

另一种方法势先产生 12 个随机数 ξ_1 、 ξ_2 、 \dots 、 ξ_{12} ，然后计算：

$$x = \sum_{i=1}^{12} \xi_i - 6$$

这两种方法产生的随机值都服从均值为零，偏差为一个单位的正态分布。初始速度经常被校正以满足总动量为零，为了是总动量为零，分别计算沿三方向的动量总和，然后用每一方向的总定量除以总质量，得到一速度值。用每个原子的速的减去此速度值，即可保持系统的总动量为零。

在建立了系统的初始位形和赋予初始速度后，就具备分子动力学模拟的初步条件了。在每一步中，原子所收到的力通过对势函数的微分可以得到。然后根据牛顿第二定律，计算加速度，再由以上提供的算法即可进行连续的模拟计算了。

六. 分子动力学模拟的系综

采用 MD 模拟，必须在一定的系综下进行，经常用到的系综包括微正则系综、正则系综、等温等压系综和等温等焓系综。

1. 微正则系综(NVE)

是孤立的、保守的系统的系综，在这种系综中，系统沿着相空间中的恒定能量轨道演化。在分子动力学模拟的过程中，系统中的原子数(N)、体积(V)、和能量(E)都保持不变。

2. 正则系综(NVT)

NVT 保持不变，并且总动量为零。恒温下，系统的总能量不是一个守恒量，系统要与外界发生能量交换。保持系统的温度不变，通常运用的方法是让系统与外界的热浴处于热平衡状态。由于温度与系统的动能有直接的关系，通常的做法是把系统的动能固定在一个给定值上。

3. 等温等压系综(NPT)

NPT 保持不变，这种系综是我们常见的系综，许多分子动力学模拟都要在从系综下进行。这时，要保证系统的温度恒定，还要保持它的压力恒定。温度恒定和以前一样，是通过调节系统的速度或加一约束力来实现的。而对压力进行调节，就比较复杂。由于系统的压力 P 与其体积 V 是共轭量，要调节压力值可以通过改变系统的体积来实现。

4. 等压等焓系综(NPH)

NPH 保持不变，焓值通过 $H=E+PV$ 得到。在系综下进行模拟时要保持压力和焓值为固定值。

七. 边界条件

正确处理边界和边界效应对模拟方法时至关重要的，因为它时从模拟相对较少的原子来计算物质的宏观性质的。为了减小有限尺寸的影响，经典的方法是使用周期性边界条件。

周期性边界条件使得可以用相对少的粒子数目来真实的模拟大块的体系，使得粒子仿佛处在一个完整的体系中，假设一个用来模拟的立方单胞，使这个单胞在各个方向上都不断重复，看起来象有周期一样。在二维图像中，每一个单胞都被其他的 8 个单胞所包围；在三维方向上，每一个单胞就会被 26 个单胞所包围。

所有的单胞中的粒子的坐标都可以通过一个整数而得到。当模拟的单胞中的一个粒子由于力的作用而离开这个单胞的时候，就会有另一个和它对应的粒子运动到这个单胞中来，这样，模拟的整个体系的粒子数就会保持不变。

在分子动力学模拟中经常采用的边界条件有：

1. 矩形盒子周期性边界条件
2. 单斜盒子周期性边界条件
3. 去头八面体合租周期性边界条件

八. MD 模拟结果分析方法

MD 模拟的时间能够达到几百皮秒或纳秒，甚至更长，在运行 MD 模拟时，体系的速度和坐标被保存下来。在分析时可以对热力学参数进行计算。热力学参数随时间演化的特性可以用图形显示，一个坐标对应时间，另外一个坐标时感兴趣的物理量，如能量、均方根差、原子位置的涨落，此外还能计算出平均结构与实验数据相对照，这些都有助于在原子水平上直观地理解构象地变化。

1. 均方根差(RMSD)

一种广泛用于检验 MD 模拟正确性的方法时求相对于蛋白质晶体结果地均方根差。目前已知一部分蛋白质地晶体结构与在溶液中的结果存在显著的差异，但人们普遍认为这种差异对于大多数蛋白质时非常小的。蛋白质模拟中常有的质量控制时获得既小又稳定的蛋白质骨架原子的 RMSD 值(通常<0.2nm)。最常见的方法时求出对于参考结构的旋转和平移拟合，原子 i(如 Ca 原子)的集合的 RMSD 可按式计算：

$$RMSD = \left[\frac{1}{N} \sum_i (r_i - r_i^0)^2 \right]$$

其中 r_i 和 r_i^0 分别指拟合后原子i的位置及其参考位置。N为要计算的原子总数。第二种方法无需进行任何拟合，而是利用了距离矩阵：

$$RMSD = \left[\frac{1}{N^2} \sum_i \sum_{j>i} (d_{ij} - d_{ij}^0)^2 \right]$$

其中， d_{ij} 是原子i和j之间的距离，而 d_{ij}^0 是在参考结构中原子i和j之间的距离。这种方法比第一种好，更常用于NMR的研究。

九. 分子力学与分子力场

在做分子动力学模拟时，常用到分子力场参数。

分子力学的经典力学模型:分子中的化学键具有“自然”键长、键角，并由这些键长和键角调节构象，给出核位置的最佳分布，即分子的平衡构型。基于 Born-Oppenheimer近似分子力学计算不显含对电子的处理。并且，量子力学从头算的计算量随基函数数目的四次方递增，半经验方法的计算量随基函数数目的三次方递增，而分子力学的计算量则仅与分子中原子数目的平方成正比。因此，分子力学已成当前研究大分子体系的分子结构、构象平衡与转变的研究。

1). 分子力场的概念

从量子化学中知道，对于任何一个包括原子核和电子的微粒体系，若忽略自旋轨道及其它相关效应，则可用定态(Time-independent)Schrddinger方程来描述它的运动

$$\hat{H}\Psi(R, r) = E\Psi(R, r)$$

其中H称之为哈密顿算符，E是体系的总能量， $\Psi(R, r)$ 是核坐标R和电子坐标r的函数。多电子体系的定态Schrodinger方程不易求解，要借助简化近似 Born-Oppenheimer近似，简称B-O近似。这个近似的依据是原子核比电子重得多这一事实，所以电子运动比核快得多。在很好的近似下，当电子运动时，我们可以把核看成是固定的。于是，电子和核运动就可以分开来处理。根据 Bom-Oppenheimer近似，体系的总波函数可以写成：

$$\Psi(R, r) = \Psi_e(r, R) \Phi_N(R)$$

其中, $\psi_e(r, R)$ 是把核坐标作为参变量的电子波函数, $\Phi_N(R)$ 是原子核的波函数。描述电子运动的Schrodinger方程为:

$$\hat{H}\Psi_e(r, R) = E(R)\Psi_e(r, R)$$

在式中, 定义了一个能量 $E(R)$, 它是体系中固定原子核坐标时给定电子状态下的电子本征能量函数, 称为分子势能函数或分子内势能函数 (intramolecular potential energy function)。在分子力场中, 通常我们把 $E(R)$ 称为位能面, 或称Born-Oppenheimer面。如果一个解析表达式能拟合这个位能面, 则此解析表达式就称为分子力场, 简称力场。即力场是: 描述分子结构和能量之间的一种数据解析表达式, 包括原子核的坐标和一些可调节的参数。所以, 一个力场的确定就是选择解析函数形式及确定参数

2). 力场能的表示形式

分子力学的基本理论就是一个分子力场由分子内相互作用和分子间相互作用两大部分构成, 分别对应于键伸缩、角弯曲、扭转运动、偶合相互作用以及Vdw相互作用和静电相互作用或氢键能。力场的参数和力常数可从一个分子迁移至相似结构的另一分子。一个力场中诸项的能量表达式基本上是相似的。

$$E_{total} = E_{valence} + E_{crossterm} + E_{nonbond}$$

其中 $E_{valence}$, 是各种成键相互作用能。成键相互作用能由键伸缩、角弯曲、二面角扭转和键角面外弯曲能等子项组成。即是:

$$E_{valence} = E_{bond} + E_{angle} + E_{torsion} + E_{oop} + E_{UB}$$

其中 $E_{crossterm}$ 是键、角之间的偶合 (Coupling), 所相应的能量在力场中称为交叉项 (CrossTerms)。这是因为考虑到分子内键长及键角的变化受到相邻键、角的影响。包括键-键偶合、键-角偶合等。其中 $E_{nonbond}$ 是非键相互作用, 由范德华能、静电效应及氢键能部分组成。具体到各种分子结构和分子势能函数的计算, 由于研究的对象和考虑的因素不同而其解析表达式有所不同

3). 常用的分子力场介绍

A. 第一代经典的分子力场。主要是AMBER, CHARMM, 及CVFF。其特点是: 函数形式简单; 应用的范围比较特定 (大部分适合于生物分子); 优化力场参数的方法比较多, 结果也比较好; 能合理地预测分子结构, 构象性质, 凝聚态性质。

B, 第二代分子力场。CFF91, CFF, PCFF, COMPASS及MMFF94。这一类力场的共同特点是函数形式较复杂, 附加项多、适用范围宽, 特别是大而复杂的分子模型的多能量极小和势垒、优化得到的力场参数比较合理; 能比较好地预测分子结构, 振动频率, 构象性质。

C. 针对周期表中的所有元素的力场。所有的力学参数基于元素, 杂化和化合的规则而产生, 并为许多结构形式所证实。这类力场有: ESFF, UFF-VALBOND, Dreiding FF。特点是: 函数形式简单; 适用的范围广; 能比较合理地预测分子结构; 依照某些规则使力场参数化。Universal力场 (UFF) 是针对整个周期表的分子力学和动态模拟的力场, 力场参数是依靠元素、元素的杂化及化合性而计算出来的, 与电荷平衡算法相结合。

D. 用于特定目的分子力场。如用于优化玻璃质的Glass forcefield用于聚合体的PCFF, 沸石结构的吸附作用的CVFF等。这类的特点自然是只局限于某一特定的范围内, 不能在其他场合中使用。

十. AMBER 练习: 真实水溶剂环境中的腺嘌呤-胸腺嘧啶十聚体(A-DNA)的分子动力学模拟

a) 用nucgen命令创建A-DNA双螺旋结构

i. 建立一个a-dna.nucgen.in文件

```
NUC 1
D
A5  A  A  A  A  A  A  A  A  A3

NUC 2
D
T5  T  T  T  T  T  T  T  T  T3

END
$ABDNA
```

b) 运行nucgen命令创建nuc.pdb文件

```
$AMBERHOME/exe/nucgen -O -i a-dna.nucgen.in -o
a-dna.nucgen.out -d $AMBERHOME/dat/leap/parm/nucgen.dat -p
a-dna.nucgen.pdb
```

这样产生 a-dna.nucgen.out, a-dna.nucgen.pdb 个文件。

c) 创建sander输入文件

下一步是为了给sander创建拓扑文件(prmtop)和坐标文件(inpcrd)。在用xleap创建之前,我们需要给a-dna.nucgen.pdb文件中的第10、11个残基之间加TER 标记。修改后的pdb文件为a-dna.pdb

运行xleap:

```
$AMBERHOME/exe/xleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff99
```

在xleap窗口创建一个工作单元adna打开a-dna.pdb文件:

```
adna = loadpdb a-dna.pdb
```

给我们的系统加抗衡粒子(中性体系):

```
addions adna Na+ 0
```

现在产生中性体系(非溶剂化体系)的拓扑和坐标文件:

```
saveamberparm adna a-dna_charges_only.prmtop a-dna_charges_only.inpcrd
```

下一步用一个截面八面体的水盒子使系统溶剂化:

```
solvateoct adna WATBOX216 8.0 #amber7.0 不支持 TIP3PBOX
```

保存溶剂化后的拓扑和坐标文件:

```
saveamberparm adna a-dna_wat.prmtop a-dna_wat.inpcrd
```

d) 对系统进行能量优化:

在对系统做MD之前要对系统做一次能量优化,为了除去加氢和溶剂化时产生的不良触点。

第一步：优化离子和水的位置。优化的前500步用最陡下降法后500步用共轭梯度法，在DNA 分子(第1~20个残基)上采用500kal/mol的限制作用力。

```
a- dna_min1.in
A-DNA 10-mer: initial minimisation solvent + ions
&cntrl
  imin   = 1,
  maxcyc = 1000,
  ncyc   = 500,
  ntb    = 1,
  ntr    = 1,
  cut    = 10
/
Hold the DNA fixed
500.0
RES 1 20
END
END
```

```
sander -O -i a-dna_min1.in -o a-dna_min1.out -p
a-dna_wat.prmtop -c a-dna_wat.inpcrd -r a-dna_min1.rst -ref
a-dna_wat.inpcrd
```

第二步：全局优化。前1000步采用最陡下降法，后1500步采用共轭梯度法，这时不需要约束力。输入文件如下：

```
a- dna_min2.in
A-DNA 10-mer: initial minimisation solvent + ions
&cntrl
  imin   = 1,
  maxcyc = 2500,
  ncyc   = 1000,
  ntb    = 1,
  ntr    = 0,
  cut    = 10
/
```

```
sander -O -i a-dna_min2.in -o a-dna_min2.out -p
a-dna_wat.prmtop -c a-dna_min1.rst -r a-dna_min2.rst
```

e) 初始平衡

这一步开始运行分子动力学。我们首先运行20ps的恒V加热平衡。采用的方法是：在DNA上加10kal/mol的弱限制力，采用恒V周期边界条件做20ps的MD，起始温度是0K，最终温度是300K，采用Langevin温度控制，在H原子上采用SHAKE限制，步长为2fs。

```
a-dna_md1.in
```


A-DNA 10-mer: 20ps MD with res on DNA

```
&cntrl
  imin = 0,
  irest = 0,
  ntx = 1,
  ntb = 1,
  cut = 10,
  ntr = 1,
  ntc = 2,
  ntf = 2,
  tempi = 0.0,
  temp0 = 300.0,
  ntt = 1,
  nstlim = 10000, dt = 0.002,
  ntp = 250, ntwx = 250, ntwr = 10000
/
```

Keep DNA fixed with weak restraints

10.0

RES 1 20

END

END

```
sander -O -i a-dna_md1.in -o a-dna_md1.out -p a-dna_wat.prmtop -c
a-dna_min2.rst -r a-dna_md1.rst -ref a-dna_min2.rst -x a-dna_md1.mdcrd
```

f) 进一步的平衡

这一步将运行更长的1.8ns的MD，对DNA不加以限制，采用Langevin恒温法使温度保持在300K，采用恒压周期边界条件(1 atm)，对H原子采用SHAKE限制，步长为2fs。

a-dna_md_1800ps.in

A-DNA 10-mer: 1800ps of MD

```
&cntrl
  imin = 0, irest = 1, ntx = 7,
  ntb = 2, pres0 = 1.0, ntp = 1,
  taup = 2.0,
  cut = 10, ntr = 0,
  ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0,
  ntt = 1,
  nstlim = 900000, dt = 0.002,
  ntp = 250, ntwx = 250, ntwr = 10000
/
```

```
$AMBERHOME/exe/sander -O -i a-dna_md_1800ps.in -o a-dna_md2.out -p
a-dna_wat.prmtop -c a-dna_md1.rst -r a-dna_md2.rst -x a-dna_md2.mdcrd
```

g) 结果分析:

从输出文件a-dna_md1.out a-dna_md2.out中可以看书系统的很多特性，如:

系统的势能、动能、总能量、温度、压力、体积、密度、均方根差。

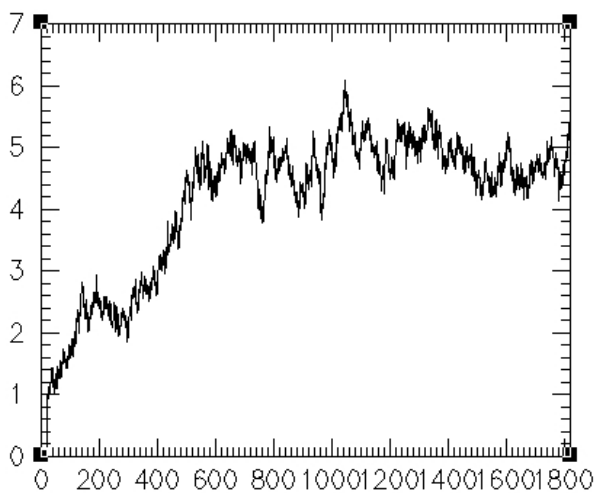
先用process-mdout.perl脚本对a-dna_md2.out进行数据处理，产生一系列summary文件，再用绘图程序xmgrace-5.99.0对上面产生的summary文件绘图。

1. 通过对势能、动能、总能量、温度、压力、体积、密度的分析，在系统得到最初的释放后，势能、动能、总能量、温度、压力、体积、密度保持恒定。这说明在这次模拟中没有出现大的问题。

2. 接下来通过计算主链上原子的RMSd来看看DNA结果发生了什么变化。使用ptraj来对整个分子轨道重新成像并除去分子轨道文件里所有的水。这样做的原因是不是要在VMD显示中隐藏水，而是现在的mdcrd文件要比先前的大得多，如果把所有的东西都加入到VMD中会使VMD运行很慢。在分子动力学模拟中，我们只对DNA主链感兴趣，所以只关注主链原子P,O3',O5',C3',C4',C5'。

```
combine_mdcrds_and_strip.ptraj
trajin a-dna_md1.mdcrd
trajin a-dna_md2.mdcrd
trajout a-dna_0-1820ps_no_wat.mdcrd
rms first out a-dna_0-1820ps.rmsfit @P,O3',O5',C3',C4',C5' time 0.50
center :1-20
image familiar
strip :WAT
ptraj a-dna_wat.prmtop < combine_mdcrds_and_strip.ptraj
```

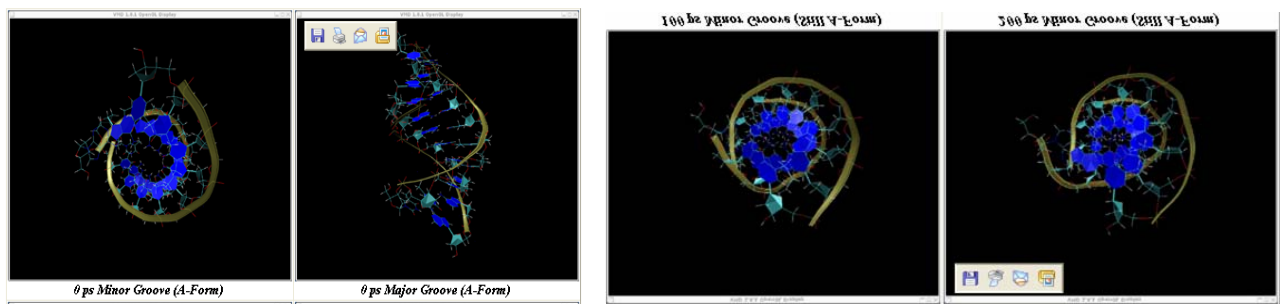
产生2个输出文件：a-dna_0-1820ps.rmsfit, a-dna_0-1820ps_no_wat.mdcrd
用Grace 打开a-dna_0-1820ps.rmsfit文件，可以得到下列图像：



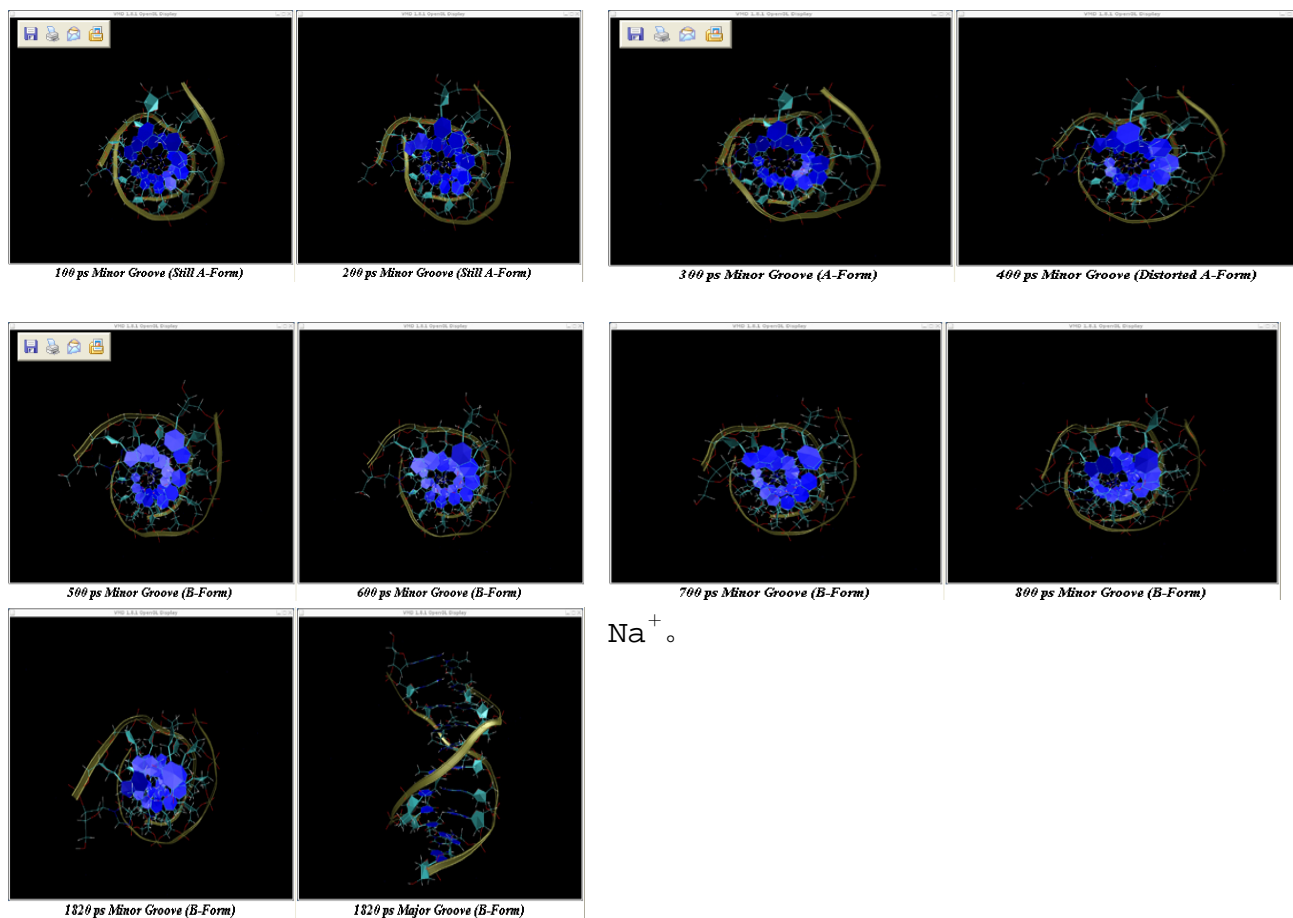
根据图像可以看到：由于没有限制力，从第20ps到500ps RMSd缓慢的增大。从500ps到模拟结束RMSd基本保持在4.5~5.5个埃。这表明：在模拟的前500ps初始构型缓慢的转变成更稳定的构型。这个稳定的构型保持到模拟结束。下面看下A型DNA能否转化成B型DNA。

由于分子轨道中含有水会使VMD运行速度变慢，所以使用去掉水的分子轨道ptraj文件。这样需要使用溶剂化以前的拓扑文件a-dna_charges_only.prmtop。

用VMD装载a-dna_charges_only.prmtop文件，参数选择parm7，然后装载



a-dna_0-1820ps_no_wat.mdcrd文件，参数选择crdbox。这样可以看到DNA链和



可以看到：从300ps到400ps即600—800步构型看是转化。在500ps(1000步)时，构型转化完成。最终构型为B型DNA。