

Molecular Dynamics simulation

——从能量最小化到实际模拟

1 基本流程图



动力学过程一般分为三步：能量最小化（minimization）、体系平衡（equilibrium）、实际动力学模拟。

由于我们进行的初始结构来自晶体结构或同源建模，所以在分子内部存在着一定的结构张力，能量最小化就是真正的动力学之前释放这些张力，如果没有这个步骤，在动力学模拟开始之后，整个体系可能会因此变形、散架。

另外，由于动力学模拟的是真实的生物体环境，因此必须使研究对象升温升压到临界值，体系达到平衡，才能做实际的动力学模拟。

2 各流程输入文件

要通过 Amber 软件做动力学模拟，需要明白如何去配置上述过程中的每一步。一般来说就是指定一些键/值对。

1) Minimization

```
&cntrl  
  imin    = 1,  
  ntp     = 1,  
  ntb     = 1,  
  cut     = 9.0,  
  nsnb    = 10,  
  ntr     = 1,  
  maxcyc  = 5000,  
  ntmin   = 1,  
  ncyc    = 2500,  
&end
```

第一行是标题，&cntrl 是起始符，&end 是结束符，也可以用 “/” 表示结束符，中间的键/值对就是参数配置。

上述的参数配置可以归纳如下：

<i>imin=1</i>	Single point energy calculation, do minimization.
<i>ntp=1</i>	这个表明采集计算信息的频率，输出到 mdout 文件中。
<i>ntb=1</i>	采用周期性边界的恒容条件，表明这是一个 NVT 系综，周期性边界在附录中说明
<i>cut=9.0</i>	由于计算能量时需要有一个截断距离，这个参数指定截断距离。
<i>nsnb=10</i>	表示一个非键列表更新的一个频率值
<i>ntr=1</i>	增加限制力，与下面的 Hold the protein fixed 照应或者出现 restraint_wt 标识
<i>maxcyc=5000</i>	能量最小化的算法涉及循环迭代，这里指定迭代次数。
<i>ntmin=1</i>	最小化方法的标志，ntmin=1 表示先用最陡下降法优化，然后用共轭梯度法
<i>ncyc=2500</i>	如上所说，循环迭代的算法不同，此处指定到哪一步第一种迭代算法结束。表示先用 2500 步的最陡下降法优化然后做 2500 步共轭梯度法优化

能量最小化过程连续做两次，第一次与第二次的参数相同

2) Heating (Equilibrium step1)

```
&cntrl
  imin = 0,
  nmropt = 1,
  ntp = 500,
  ntwr = 500,
  ntf = 2,
  ntb = 1,
  cut = 9.0,
  ntr = 1,
  nstlim = 25000,
  nscm = 1000,
  dt = 0.002,
  ntp = 0,
  ntt = 3,
  gamma_ln=2.0
  temp0 = 300.0,
  tempi = 0.0,
  ntc = 2,
  tol = 0.000001,
&end
```

```
&wt
  type = 'TEMP0',
  istep1 = 0,
  istep2 = 20000,
  value1 = 0.0,
  value2 = 300.0,
&end
&wt
  type = 'TEMP0',
  istep1 = 20001,
  istep2 = 25000,
  value1 = 300.0,
  value2 = 300.0,
&end
&wt
  type = 'END',
&end
Hold the Solutes fixed
5.0
RES 1 285
END
END
```

上述的参数配置可以归纳如下：

- | | |
|---------------|--|
| imin=0 | 开始做动力学，而不是能量最小化。 |
| nmropt=1 | 这个参数原本用于指定 nmr 限制，现在被机智地借来用于控制升温过程。 |
| ntwr=500 | 采集计算信息的频率，每隔 100 步打印一次信息写入 restrt 文件中 |
| ntf=2 | 忽略涉及到 H 原子的键的相互作用 |
| nstlim=25000 | 动力学过程的步长，前面说过动力学的原理，它是连续地解牛顿运动方程。动力学的时长会等于步长（nstlim）乘以每一步的时间间隔（如下，dt）。 |
| dt=0.002 | 每一步的时间间隔，单位是 ps。 |
| ntp=0 | 暂时不考虑控制压力 |
| ntt=3 | 用郎之万动力学的方法，且需要设置一个 gamma_ln 值。 |
| gamma_ln=2.0 | 以上两个参数指定如何控制温度。 |
| ntc=2 | 以上两个参数针对氢原子做 shake 限制，主要由于氢原子的振动频率过高，氢原子的运动还存在量子效应。所以涉及到 H 原子的键需要被限制 |
| tempi=0.0 | 初始的模拟温度，单位为 K |
| temp0=300.0 | 升温后的温度，单位为 K |
| tol= 0.000001 | 用 shake 重新设置坐标的一个几何公差 |

升温过程的控制：

这三行配置表示，在 25000 步的升温中，在最初的 20000 步中，温度将从 0K 升到 300K，最后 20001 步到 25000 步，温度会保持在 300K。

3) Equilibrium step2、Production

<pre> &cntrl imin = 0, ntx = 5, irest = 1, ntpr = 500, ntwx = 500, ntf = 2, ntb = 2, cut = 9.0, ntr = 1, nstlim = 25000, dt = 0.002, temp0 = 300.0, tempi = 300.0, ntt = 3, ntp = 1, pres0 = 1.0, taup = 1, ntc = 2, tol = 0.000001, &end Hold the Protein fixed 5.0 RES 1 285 END END </pre>	<pre> &cntrl ntx = 5, irest = 1, ntpr = 500, ntwx = 500, ntf = 2, ntb = 2, cut = 9.0, nsnb = 10, ntr = 1, nstlim = 25000, dt = 0.002, temp0 = 300.0, ntt = 3, ntp = 1, tautp = 0.5, ntc = 2, tol = 0.000001, &end Hold the solutes fixed 5.0 RES 1 285 END END </pre>	<pre> &cntrl ntx = 5, irest = 1, ntpr = 500, ntwx = 500, ntf = 2, ntb = 2, cut = 9.0, nsnb = 10, ntr = 1, nstlim = 25000, dt = 0.002, temp0 = 300.0, ntt = 3, ntp = 1, tautp = 0.5, ntc = 2, tol = 0.000001, &end Hold the solutes fixed 4.0 RES 1 285 END END </pre>
升温动力学模拟	恒温恒逐步释放限制的的动 力学_1	恒温恒逐步释放限制的的动 力学_2
<pre> &cntrl ntx = 5, irest = 1, ntpr = 500, ntwx = 500, ntf = 2, ntb = 2, cut = 9.0, nsnb = 10, ntr = 1, nstlim = 25000, dt = 0.002, temp0 = 300.0, ntt = 3, ntp = 1, tautp = 0.5, ntc = 2, tol = 0.000001, &end Hold the solutes fixed 3.0 RES 1 285 END END </pre>	<pre> &cntrl ntx = 5, irest = 1, ntpr = 500, ntwx = 500, ntf = 2, ntb = 2, cut = 9.0, nsnb = 10, ntr = 1, nstlim = 25000, dt = 0.002, temp0 = 300.0, ntt = 3, ntp = 1, tautp = 0.5, ntc = 2, tol = 0.000001, &end Hold the solutes fixed 2.0 RES 1 285 END END </pre>	<pre> &cntrl ntx = 5, irest = 1, ntpr = 500, ntwx = 500, ntf = 2, ntb = 2, cut = 9.0, nsnb = 10, ntr = 1, nstlim = 25000, dt = 0.002, temp0 = 300.0, ntt = 3, ntp = 1, tautp = 0.5, ntc = 2, tol = 0.000001, &end Hold the solutes fixed 1.0 RES 1 285 END END </pre>
恒温恒逐步释放限制的的动 力学_3	恒温恒逐步释放限制的的动 力学_4	恒温恒逐步释放限制的的动 力学_5

```

&cntrl
ntx      = 5,
irest    = 1,
ntpr     = 500,
ntwx     = 500,
ntf      = 2,
ntb      = 2,
cut      = 9.0,
nsnb     = 10,
ntr      = 0,
nstlim   = 25000,
dt       = 0.002,
temp0    = 300.0,
ntt      = 3,
ntp      = 1,
tautp    = 0.5,
gamma_ln = 2.0,
ntc      = 2,
tol      = 0.000001,
&end
RES 1 285
END
END

```

恒温恒无限制的的动力学_6

```

&cntrl
ntx      = 5,
irest    = 1,
ntpr     = 500,
ntwx     = 500,
iwrap    = 1,
ntf      = 2,
ntb      = 2,
ntr      = 0,
cut      = 9.0,
nsnb     = 10,
nstlim   = 500000,
dt       = 0.002,
temp0    = 300.0,
ntt      = 3,
ntp      = 1,
gamma_ln = 2.0,
ntc      = 2,
taup     = 2.0,
tol      = 0.000001,
&end

```

MD_production

这个注意：这里配置的动力学时间只有 **60ps**，主要是为了演示；一般来说视体系的大小和研究的需要，模拟的时间一般会更长。

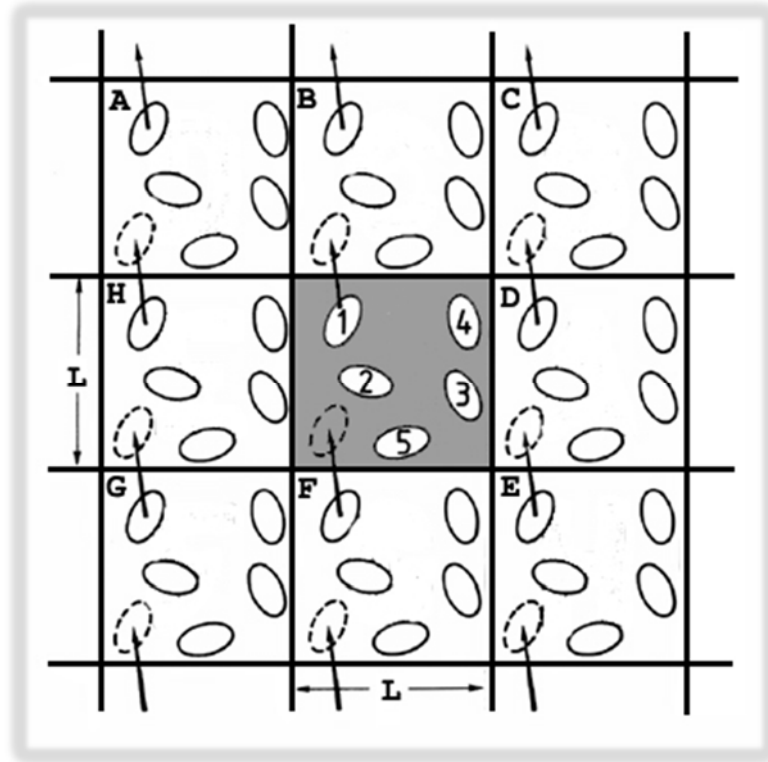
第二步平衡（体系升压）与实际动力学的参数配置是一样的，因为实际动力学便是在恒温恒压下进行的。

上述的参数配置可以归纳如下：

- ntx=5** *Read coordinates and velocities from unformatted **inpcrd** coordinate file*
ntx 等于 5 表示将读取另一种格式的 inpcrd 文件，它往往配合 irest=1（如下）使用。
- irest=1** *Restart previous MD run [This means velocities are expected in the inpcrd file and will be used to provide initial atom velocities]*
irest 等于 1 表明将重启上一次的动力学，也就意味着希望能从输入文件中读取到上一次动力学结束时对应的（除了坐标信息外的）原子速度信息（储存在 rst 文件中）。
- temp0=300.0** *Thermostat temperature. Run at 300K*
动力学过程中需要控制保持的温度，一般在 300K 左右
- ntb=2** *Use periodic boundary conditions with constant pressure*
采用周期性边界的恒压条件。一般动力学即采用此 NPT 系综。
- ntp=1** *Use the Berendsen barostat for constant pressure simulation*
设定控压的算法。
- ig=-1** The seed for the pseudo-random number generator, If ig=-1, the random seed will be based on the current date and time, and hence will be different for every run
- ntr=1** Flag for restraining specified atoms in Cartesian space using a harmonic potential, if ntr > 0.
The restrained atoms are determined by the restraintmask string

附录：

1) 周期性边界条件简介



一个二维盒子排列，分子 1 将从中心盒子（有阴影的盒子）运动到盒子 B 中，这时为了保持中心盒子的粒子守恒，盒子镜像 F 中有一个相应的粒子会进入到中心盒子中来，就好像分子 1 从上缘出去，而从下缘进入。

周期性边界条件使得粒子包裹在无穷多的溶液分子中，这样在使用相对少量的粒子的情况下模拟了真实的生物体系，使动力学模拟的成本大大减少。假设一个有粒子在内部运动的立方体盒子从各个方向被复制，形成了周期性排列的盒子，就像上图的二维盒子展示的一样。

在二维的情况下，每个盒子周围都有 8 个相邻的盒子，而在三维的情况下，每个盒子周围则会有 26 个相邻的盒子。对于镜像盒子，相当于中心盒子朝各个方向平移了，因此镜像盒子中粒子的坐标也就等于中心盒子的坐标加上或减去平移向量： $\mathbf{r} + i\mathbf{x} + j\mathbf{y} + k\mathbf{z}$ ($i, j, k \rightarrow -$

inf, +inf)。如果一个粒子在模拟过程中从某一面离开了盒子，那么相应地就会有一个镜像粒子从另一面进入盒子，始终保持盒子中粒子数目不变。

立方体是最简单的可用于周期性排列的几何形状，视图查看和程序化都很方便；然而，有些情况下另一种盒子形状更适合模拟体系，例如被水分子包围的单个大分子或复合物，往往需要研究的对象也就是被包裹在中心的大分子、复合物，因此周围溶剂（水分子）的计算量越少越好，也就是说尽可能减少水分子的数目。在这种情况下，立方体盒子事实上是最不节省的盒子。

原则上，任何形状的盒子都可以用于周期性模拟，只要通过平移操作，能够填满整个三维空间（也就是中心盒子与镜像盒子间没有空隙）。因此大概有 5 类盒子满足条件：立方体（同样的，长方体也可以）、六方柱、截八面体、菱形十二面体。

另外，在选取合适的盒子形状后，模拟中盒子的大小也是非常重要的，既要盒子足够大，保证模拟的精确，又要尽量减小盒子的大小降低计算量。体系的大小一般要大于体系中相互作用能的计算范围的两倍。